FROM CASH TO CLASSROOMS: DIRECT AND INDIRECT EFFECTS OF A CONDITIONAL CASH TRANSFER PROGRAM ON GIRLS' EDUCATION IN WEST BENGAL, INDIA

Thesis submitted in partial fulfillment of the requirements for the Degree of Master of Science in Economics and Management of Government and International Organizations at Bocconi University

October 2024

By

Kanika

Supervisor: Professor Thomas Le Barbanchon

Acknowledgements

I would like to express my deepest gratitude to my dissertation supervisor, Professor Thomas Le Barbanchon, for his invaluable guidance throughout this journey. His mentorship not only helped shape the direction of my research but also reignited my passion for econometrics during my time at Bocconi. I am deeply appreciative of his insight, patience, and encouragement. I am also grateful to the ASER Centre for sharing the data used in the empirical analysis.

I'm deeply thankful to my parents for their unwavering love and support throughout this journey. They have been my role models in every way, and I wouldn't be where I am today without their constant belief in me. To my brother Rahul, thank you for always inspiring me to persevere and for being a source of motivation when I needed it most. To my friends at Bocconi, thank you for being wonderful companions through everything—from stressing together over exams to our impromptu aperitivos after class. You made this MSc journey feel lighter, more fun, and filled with so much warmth.

Lastly, a special thank you to my puppy Max, who kept me company during the long nights of work and always reminded me to take breaks. His presence brought so much comfort and joy and made even the hardest moments feel a little easier.

Abstract

This paper evaluates the causal impact of the Kanyashree Prakalpa (KP), a conditional cash transfer (CCT) program introduced in West Bengal, India, in 2013. The ongoing program aims to increase school enrollment and delay marriage for adolescent girls from socio-economically disadvantaged backgrounds. The program provides an annual scholarship to 13- to 18-year-old girls, conditional on being enrolled in school and remaining unmarried. Using a difference-in-difference (DD) and difference-in-difference-in-differences (DDD) framework, I analyze the direct effect of KP on eligible girls and the indirect effect on their younger and ineligible siblings. The analysis focuses on school enrollment, learning outcomes, and household educational investments for these two groups. I find a positive, though modest, effect on school enrollment and numeracy skills of eligible girls, especially those with more disadvantaged backgrounds and live closer to formal schools. I also find a positive indirect effect on enrollment of younger and ineligible male siblings. However, I find no significant indirect effects on learning outcomes and household educational investments of younger ineligible siblings. Overall, I find that KP program plays a limited role in affecting the education of girls and addressing the issue of gender gap in education in West Bengal. This paper contributes to the literature by being the first to explore indirect effects of KP on ineligible siblings and by demonstrating how the program's impact varies based on maternal education and proximity to schools.

Table of Contents

I. Introduction

In a developing country like India, adolescent girls often face a multitude of problems like school dropout, child marriage, and poor mental and physical health (Bergstrom & Ozler, 2022). Although there has been a notable decline in the incidence of child marriage in India —from 54% to 26% between 1992 and 2015, it continues to pose a significant challenge for the educational attainment and well-being of young girls (Nanda et al., 2022). Given this context, formal education is one of the most widely recognized tools for tackling such societal issues. Staying enrolled in school beyond the primary level is not only associated with a lower incidence of child marriages but also improved health outcomes and greater economic opportunities for girls (Llyod & Young, 2009). For instance, Psacharopoulos and Patrinos (2004) find that each year of education beyond the primary level leads to an increase of about 10% in an individual's earnings, which may be critical for addressing intergenerational inequality and poverty.

However, despite achieving near universal primary school enrollment, there continues to be a critical gender gap in secondary schooling in India (Murlidharan & Prakash, 2017). Both demand- and supply-side constraints have contributed to this gender gap. One possible demand-side constraint is the differential treatment of sons and daughters by their parents. Many studies have found pro-male bias in both the decision to send the child to school and the household expenditure on education (Azam & Kingdon, 2011; Lancaster et al., 2008; Bhatkal, 2012). This bias may be due to perceived lower economic returns to girls' education than boys, lingering cultural and gender norms, and early marriage (Kingdon, 2002). A supply-side constraint could be the lack of availability of a same-gender school in rural areas that parents may prefer for their adolescent girls rather than a co-educational school due to safety concerns or strong cultural beliefs (Azam & Kingdon, 2011).

To address the challenge of gender bias in education and school accessibility, policymakers have focused on both demand- and supply-side interventions to bridge the gender gap in secondary education. Supply-side interventions include constructing new schools to reduce the time and distance cost of attending school. Demand side interventions include conditional cash transfer programs (CCTs) where either the girl herself or her parents receive cash conditional on the girl staying enrolled in school. In this paper, I analyze the impact of the 'Kanyashree Prakalpa' (henceforth, KP) conditional cash transfer program on enrollment, learning levels, and private household investment in education. This program was introduced in West Bengal (a state in the Eastern part of India) in October 2013. The program covers 13–18-year-old girls from economically disadvantaged backgrounds and includes conditionalities such as staying enrolled and regularly attending school and remaining unmarried throughout the eligibility period.

Many studies have found a positive effect of CCTs on enrollment, but their impact on learning outcomes and private household investment in education remains ambiguous. Moreover, the broad literature on CCTs primarily focuses on the direct effect of the policy on the beneficiary siblings and not so much on the non-beneficiary siblings in the same household. I contribute to this literature by evaluating the direct effect of KP on eligible girls and the indirect effects on both male and female ineligible younger siblings. Understanding both the direct and indirect effects of such a policy and their plausible mechanisms is important due to several reasons. First, the presence of these effects would have consequences for the future design of such policies and programs. Second, given the significant economic investments made by the West Bengal government into the program, it is imperative to understand if it actually helps in increasing girls' enrollment and learning levels or if it has negative consequences on non-beneficiary siblings.

Using individual-level schooling and test scores data from the Annual Status of Education Reports (ASER), I conduct both difference-in-difference (DD) and difference-in-difference-in-differences (DDD) analysis to isolate the causal impact of the KP program by comparing outcomes across time between West Bengal and neighboring control states. While using DD estimation strategy, I find no effect of the program on eligible girls. However, I do find a modest positive spillover effect on school enrollment of younger boy siblings but no significant spillover effects on younger girl siblings. While using DDD estimation strategy, I find a very modest increase in enrollment and math learning levels for the eligible girls suggesting that KP plays a limited role in affecting the education of girls and addressing the issue of gender gap in education in West Bengal.

The rest of the paper is organized as follows. Section II outlines the context and the program in detail. Section III describes the data and the empirical strategy. Section IV discusses the main results and heterogeneity. Section V presents several robustness checks and Section VI concludes. Tables and figures with the prefix "A" are in the appendix attached at the end.

II. Context and Program Description

According to Census 2011, West Bengal is the fourth most populous state in India. The state faces numerous social and economic challenges, with child marriage being one of the most prominent issues. Over 40 percent of ever-married women in the state have been married before the legal age of 18 (Sen & Thamarapani, 2023). Child marriage is a serious human rights violation and may lead to several adverse outcomes for the girl. It may lead to teenage pregnancy, high incidence of anemia, and child stunting (Dutta & Sen, 2020).

Given these severe consequences, it is crucial to address the issue of child marriage. With formal education being one of the most important tools to tackle issues like child marriage, West Bengal needed a policy that would encourage and incentivize the girls as well as their families to keep the girls enrolled in school while remaining unmarried (Dutta & Sen, 2020). Therefore, the Government of West Bengal launched the Kanyashree Prakalpa (KP) scheme in October 2013.

KP is a Conditional Cash Transfer program that covers 13–18-year-old girls from economically disadvantaged backgrounds with conditionalities such as staying enrolled and regularly attending school and remaining unmarried until 18. The program has two main objectives. The first objective is to promote continuation of education by providing an annual scholarship to 13- to 18-year-old girls who are enrolled in school and are unmarried. This financial support is intended to alleviate some of the demand-side constraints like poverty that often lead parents to withdraw their daughters from school as they reach secondary school. While primary schooling in India is free until Grade 8[1] under the Right to Education (RTE) Act, secondary schooling is often only partly subsidized. This may create financial barriers for families and may push the girls to drop out before entering secondary school. The second objective is delaying marriage, which is achieved by a substantial one-time grant at age 18, conditional upon remaining unmarried.

Furthermore, KP has two main financial components. The first component is the annual scholarship of INR 750 (now INR 1000 or $12). To receive the scholarship, the girl should be between the ages of 13-18 years old, remain unmarried, and be enrolled in a formal school, madrasa[2], or an equivalent open school course[3]. The girl should also regularly attend her classes, and this is confirmed by a certificate from the head of her educational institution at the time of the application. Furthermore, the annual family income of the girl should be less than INR 1,20,000 (~ $1430). However, this requirement is waived for disabled or orphaned girls.
The second component is the one-time grant of INR 25,000 (~ $300) that the girl receives when she is between 18-19 years old and continues to meet the enrollment and marital status requirements described above. The program aims to achieve financial inclusion of girls by transferring cash into bank accounts where the eligible girl herself is the account holder. KP has covered over 7,525,819 unique beneficiaries in West Bengal ever since its inception in 2013 and remains one of the most successful CCTs in India (Dutta & Sen, 2020).

KP has gained both national and global recognition for its impact. It was awarded the United Nations Public Service Award in 2017 for its success in reducing child marriage and early dropouts among adolescent girls (Dutta & Sen, 2020). Various studies have either qualitatively or quantitatively evaluated the KP program and have found positive outcomes like improved retention of girls in schools and delayed marriages. Das and

[1] Grade 8 is usually associated with age 12-13 years in India.
[2] Madrasa is a religious Islamic school.
[3] Open school provides a flexible learning environment where students can choose their study schedules and subjects as per their convenience. This is an alternative for children who are unable to join a traditional or formal school.

Sarkhel (2023) use household data from the Annual Status of Education Report (ASER) survey (2008 – 2018) to analyze the effect of KP and find a 7-percentage point increase in government school enrollment for girls and a modest improvement in their lower-order learning outcomes. However, the authors observe a decline in higher-order learning skills (especially math), which they attribute to inadequate school resources and teacher absenteeism. Similarly, Dey and Ghosal (2021) analyze the effect of KP using data from multiple rounds of the National Family Health Survey (NFHS) and find a significant 6% increase in secondary or higher educational attainment among 13–18-year-old girls in West Bengal. Additionally, Sen and Thamarapani (2023) use NFHS-4 (2015-16) data to assess the likelihood of girls being enrolled in or completing secondary and higher secondary school. They report that KP-eligible girls are 12% more likely to be enrolled in secondary school and 7% more likely to be enrolled in higher secondary school compared to non-eligible girls. However, to the best of my knowledge, no studies have analyzed the indirect effects of KP on non-beneficiary siblings and examined how the program's impact differs across different demographic groups, such as those with different levels of maternal education and proximity to formal schools.

This paper aims to contribute to the existing literature on CCTs by conducting a detailed analysis of both the direct and indirect effect of KP on eligible girls and their ineligible siblings. By exploring the indirect effects on ineligible siblings and performing a heterogeneity test based on maternal education and proximity to schools, I aim to provide a broader and more comprehensive understanding of KP's impact and its effectiveness in addressing the socio-economic challenges it set out to curb in West Bengal.

III. Data and Empirical Strategy

A. Data source and scope

My primary data source is the household data from the Annual Status of Education Report (ASER) survey between 2008 and 2022 (ASER Centre, 2009–2022). The survey was conducted annually until 2014, but it was later switched to an alternate-year cycle. Therefore, the data is available for the years 2008 to 2014, 2016, 2018, and 2022. Notably, data for 2020 is missing due to disruptions caused by COVID-19. The ASER survey is a national-level household survey that collects data on 3–to 16-year-old children only in rural areas. It provides comprehensive information on children's socio-economic background, enrollment in both 'traditional[4]' and 'non-traditional[5]' educational institutions, and reading and math learning levels.

[4] Traditional educational institutions include private and government/ government aided schools

[5] Non – Traditional institutions include madrasa, vocational/technical institutions, open school courses.

Sampling Method: The ASER survey uses a two-stage sampling design which ensures that the survey is representative at the district level[6]. In the first stage, to ensure that sampling units (villages) with bigger populations have a higher chance of being selected, the villages are selected using the probability proportional to size (PPS) method. This method is advantageous when sampling units vary greatly in size because it ensures that individuals in larger sites have the same probability of being included in the sample as those from the smaller sites. In the second stage, 20 households are randomly selected within each village, which ensures that all children in these households are covered independent of whether they are enrolled in formal schools or non-traditional schools like madrasas. Lastly, the survey includes 30 villages per district and provides reliable estimates of schooling status and basic learning levels for children aged 3 to 16 years old at the district, state, and national levels.

Assessment Method: In addition to collecting information on school enrollment status, the survey collects data on children's foundational literacy and numeracy skills. Since basic literacy and numeracy skills have been found to be positively associated with labor market outcomes later in life (Meehan et al., 2023; McIntosh & Vignoles, 2001), it is important to see whether KP affects learning outcomes among eligible girls as well. ASER uses an internationally recognized testing tool designed to capture the highest level that each child can comfortably achieve instead of testing individual grade-level competencies, which may be more subjective across regions. These tests are conducted individually with the surveyor at home, which ensures a more comfortable environment for the child with minimal cognitive load. Moreover, the test is conducted in one of the 19 local Indian languages and is adaptive to the child's ability to ensure that the child does not necessarily have to attempt all the test levels. Reading tests include tasks like letter recognition and reading simple texts, and arithmetic tests focus on number recognition and basic operations like subtraction and division. The content of tests is in line with the state-mandated curricula to ensure it is relevant for the child. Despite language differences, this standardized approach allows for reliable and valid comparisons of foundational skills across time and regions.

Data selection Motivation: I chose the ASER dataset for this study because of its extensive coverage and standardized testing approach, which ensures comparability of data across different regions and years in India. To the best of my knowledge, ASER is the only survey in India that provides comprehensive data on educational outcomes like enrollment, learning levels, and private investment in education combined with socio-economic outcomes for children at both the household and the village level. Therefore, it is particularly well-suited for evaluating the impact of the KP on girls' education in West Bengal.

[6] A district in India is an administrative division similar to a county in the United States.

Data Limitations: While the intervention began in October 2013, data for the years 2015, 2017, 2019, 2020, and 2021 is missing. Given that COVID-19 may have had a negative impact on school enrollment and learning outcomes between 2020 and 2022 (Guariso & Nyqvist, 2023), this gap in the data poses a significant challenge in analyzing the dynamic effects of KP on these outcomes. Moreover, the KP program covers 13–18-year-old girls in West Bengal, but the ASER survey only collects information on children aged 3 to 16. Nevertheless, the biggest limitation of the ASER dataset for this study is that it does not include indicators of enrollment in the KP program or household income levels, both of which are crucial for determining the child's eligibility into the program. These limitations may hinder my ability to assess the causality of the policy's impact accurately.

However, it is important to note that despite these limitations, alternative data sources provide some reassurance. Using the Indian Human Development Survey (IHDS) from 2011-12, I find that about 90.3% of rural households in West Bengal have an annual income below the KP eligibility threshold of INR 120,000 (~$1600)[7]. This high proportion implies that most rural households with the presence of an eligible girl (13–16-year-old) would qualify for the program. Moreover, I find similar income distributions in control states, which further support the comparability of treated and control groups for this study. Therefore, the lack of explicit data on enrollment in KP or the income threshold is partly mitigated by this additional analysis using an alternative dataset.

Important variables: The first outcome of interest is school enrollment, which is a binary variable that takes the value of 1 if a child is currently enrolled in any type of educational institution or 0 if the child was never enrolled/has dropped out. The assessment of reading skills has ordinal ranking with an increased level of difficulty— recognition of letters, reading of words, reading a short paragraph (a grade 1 level text), and reading a short story (a grade 2 level text). Similarly, the math level of the child is based on ordinal ranking— recognition of single-digit numbers, recognition of double-digit numbers, subtraction of two-digit numbers with a borrowing, and division of a three-digit number by one digit. Therefore, for the purpose of this study, I have coded Math and Reading levels on a scale of 0 to 4 (inclusive) to indicate the five progressive levels measured through the ASER assessment tools. The last outcome variable that I consider is child tuition, which is a binary variable that takes the value 1 if the child is currently enrolled in private tuition classes outside of regular school or 0, if not enrolled.

[7] The dataset and replication code are available upon request.

B. Identification Strategy and Estimating Equations

I exploit two sources of variation to isolate the causal impact of the KP program: (i) the eligibility criterion as the program restricts access to 13- to 16-year-old girls who are enrolled in an educational institution and are unmarried and (ii) similar neighboring states that did not have such a conditional cash transfer program for girls.

***Difference-in-Differences**:* Given 13–16-year-old girls in West Bengal as the treated group, I can consider 13–16-year-old girls in the neighboring states (Odisha, Jharkhand, and Chhattisgarh) to be the control group. These neighboring states are similar to West Bengal in terms of proximity and socio-economic and cultural factors[8]. I adopt a linear probability model for the DD and DDD estimation, following Muralidharan & Prakash (2017) and Anukriti (2018).

In the DD specification, the difference between the 13–16-year-old girls of West Bengal and the control states after the implementation of the program (after 2013) is compared to the same difference before the program was implemented (before 2013). The corresponding difference-in-differences (DD) estimating equation is as follows:

$$(1)\ y_{ist} = \beta_0 + \beta_1.\ post_t + \beta_2.\ \mathrm{WB}_s + \beta_3.\ (post_t \times \mathrm{WB}_s) + \vartheta X_{ist} + \gamma_t + (\partial_s \times t) + \epsilon_{ist}$$

Here, y_{ist} is the outcome variable of interest for the i^{th} child from state s measured at time t. The first outcome variable for enrollment is a dummy that takes the value 0 if the i^{th} child has either dropped out or was never enrolled in school and the value 1 if the child is enrolled in any type of traditional or non-traditional educational institution. For learning outcomes (both math and reading levels), y_{ist} takes integer values from 0 to 4, where 0 means no learning skills and 4 implies the highest level of learning. For child tuition, y_{ist} is a binary variable that takes the value 0 if the i^{th} child is not enrolled in any private tuition classes and 1 if the child is enrolled in private tuition classes. 'WB_s' is a dummy that takes the value 1 if the i^{th} observation comes from West Bengal. The '$post_t$' dummy compares the outcomes for the post-program years (after 2013) when it takes the value 1 to the same before 2013 when it takes the value 0. The coefficient of interest is β_3, which measures the intent-to-treat (ITT) by modeling the potential exposure of the eligible population (13–16-year-old girls in WB) to the KP program. X_{ist} includes child-, household-, and village-level characteristics. The child level controls include the age of the child and whether their mother went to school. The household level controls include the

[8] Please see Figure A1 in the appendix for a state-wise map of India.

number of household members, the presence of a pucca[9] house, the presence of electricity, the possession of a phone, and the presence of a toilet. The village-level controls include the presence of government secondary and middle schools, a private school, a private health clinic, a bank, and a pucca road in the village. γ_t are the year-fixed effects. Moreover, since Bilinski and Hatfield (2020) suggest that the default DD estimation equation should allow for linear trend differences, I also include state-specific linear time trends ($\partial_s \times t$) in my specification. As a robustness check, I also test if my results remain robust to the addition of higher-order state-specific trends. This test is discussed in more detail in Section V. Lastly, the regression is weighted to be representative at the state level.

Now, even though we get an estimate of the program using the DD method, certain broader trends can still affect the DD estimate. Olden and Moen (2022) posit that while the Difference-in-Differences (DD) estimator is well-understood and extensively studied in the literature, it often sacrifices degrees of freedom and may offer less detailed information compared to the Difference-in-Difference-in-Differences (DDD) estimator in some cases. Therefore, a triple difference approach can be adopted here to ensure the reliability of the results and eliminate any other broad trends that might be biasing the results.

Triple difference: Following Muralidharan and Prakash (2017), I use a DDD regression specification that compares Group A (13–16-year-old girls) and Group B (13–16-year-old boys) in the treatment state before and after KP implementation as the first double difference, with Group A (13–16-year-old girls) and Group B (13–16-year-old boys) in the control states before and after KP implementation as the second double difference. The reasoning behind decomposing my DDD estimator as the difference between these two double differences mentioned above is that I assume that relatively time-invariant biases like pro-male bias in education are similar in treated and control states and, therefore, get differenced out in the DDD estimator. Consequently, I estimate the following DDD model:

$$(2)\ y_{ist} = \beta_0 + \beta_1 . \mathrm{WB}_s + \beta_2 . post_t + \beta_3 . \mathrm{Female}_i + \beta_4 (post_t \times \mathrm{Female}_i) + \beta_5 (post_t \times \mathrm{WB}_s) + \beta_6 (\mathrm{WB}_s \times \mathrm{Female}_i) + \beta_7 (\mathrm{WB}_s \times post_t \times \mathrm{Female}_i) + \vartheta X_{ist} + \gamma_t + (\partial_s \times t) + (\theta_i \times t) + \epsilon_{ist}$$

Here, ($\theta_i \times t$) is the gender-specific linear time trends. I add it to the specification in addition to the state-specific linear time trends to account for any underlying differences in enrollment and learning trends between males and females over time, which may arise due to policy interventions like Beti Bachao Beti Padhao[10] or

[9] A pucca house is a durable structure with walls and a roof made from materials like burnt bricks, cement, and reinforced concrete. In contrast, a kutcha house is made from less durable materials such as mud, bamboo, or thatch.

[10] Beti Bachao Beti Padhao (BBBP) is a nationwide initiative launched by the Government of India in 2015 aimed at promoting the survival, protection, and education of the girl child. The program seeks to improve the welfare of girls through various campaigns and interventions, which

other unobserved factors affecting gender-specific outcomes. Including this interaction ensures that the estimated treatment effect of KP is not confounded by these broader gender-specific trends that could vary independently of the intervention. The rest of the variables used in equation (2) have already been mentioned while discussing the DD specification. In this model, the coefficient of interest is β_7, which represents the causal ITT effect of the KP program on the eligible girls. Now, before I discuss the main results from these specifications, it is important to discuss the main identifying assumptions required to get a causal ITT estimate.

C. Identifying Assumptions

The two key identifying assumptions in both DD and DDD settings are Parallel trends and No Anticipation. Together, these assumptions allow the identification of the ITT in this study.

***Parallel Trends*:** The validity of the DD estimator relies on the parallel-trend assumption, which states that the outcome variable should move parallelly between the treated and control groups in the absence of treatment (Angrist & Pischke, 2009). In the context of a DDD estimator, Older and Moen (2022) argue that even though the triple difference is the difference between two difference-in-differences, it does not need two parallel assumptions. Instead, it requires the relative outcomes of group A (13–16-year-old girls) and group B (13–16-year-old boys) in the treatment state (West Bengal) to trend in the same way as the relative outcome of group A (13–16-year-old girls) and group B (13–16-year-old boys) in the control states (Odisha, Chhattisgarh and Jharkhand).

I first check the parallel trends graphically by comparing the group means across time. The trends are presented in Figures A2 through A5 in the appendix. The trends appear broadly parallel, but at times converge or diverge. Consequently, I also conduct regression analysis to check the validity of the parallel trend assumption for both DD and DDD specifications. I test for parallel trends in boys' and girls' enrollment as well as learning outcomes in the six years prior to the program (2008–2013) using the ASER data and regression specifications discussed above. I find that the parallel trends assumption holds for all the outcomes using both specifications as shown in Tables A1 and A2 in the appendix.

To further reinforce these findings, I conduct additional tests. I run the regression specifications with individual year dummies and conduct a joint F test for each of the interaction coefficients. This test also yields

could positively influence school enrollment rates for girls. These concurrent gender-focused efforts might confound the results of my analysis by independently affecting the enrollment trends in girls with respect to boys that the Kanyashree Prakalpa program aims to impact.

insignificant results for all outcomes using both DD and DDD specifications. I report these in the appendix (Tables A3 and A4). Therefore, I conclude that trends are parallel in the pre-treatment period between West Bengal and control states (Odisha, Chhattisgarh, and Jharkhand) for enrollment, learning outcomes and private tutoring of 13-16 years old girls and boys for DD specification (1) and DDD specification (2). It is still possible that testing for parallel trends may not be sufficient due to low power or the presence of unobserved confounders (Roth et al., 2023). Nevertheless, the graphical mean comparison of the outcome variables, along with the regression analysis helps strengthen the case.

Additionally, in Table A5 in the appendix, I do a balancing test on all the covariates used in the DD and DDD estimations. I conduct this exercise to test for the presence of significant socioeconomic differences between West Bengal and the control states in the pre-program years. Notably, some variables show imbalances across the treated and control groups, like the proportion of children whose mothers have had some schooling, the proportion of households with a toilet, and the proportion of villages with a pucca road, a private clinic, and a private school. The remaining variables exhibit non-significant differences. Consequently, to ensure that my estimates of the effects of KP are not confounded by these preexisting socioeconomic differences between the treated and control samples, I include these imbalanced variables in the regression analysis to control for any remaining selection bias.

No anticipation: The no-anticipation assumption is often hidden in DD and DDD settings as opposed to the parallel trend assumption (Roth et al., 2023). It states that the treatment (KP in this case) has no causal effect prior to its implementation. This is important for the identification of the ITT because if this assumption is not fulfilled, then the change in outcome for the treated group between the pre-and post-period could potentially reflect not only the causal effect of the policy in the post period but also its anticipatory effect in the pre-period (Abbring & van den Berg, 2003; Malani & Reif, 2015). To test for this assumption, I conduct placebo tests where I change the year of intervention from 2013 to either 2010 or 2011 and calculate the ITT while restricting my dataset until 2013. The DD and DDD estimations are run based on these placebo intervention years and the treatment effects are insignificant for all outcomes using both specifications. These results are discussed in more detail in Section V. Moreover, since KP was rolled out universally in October 2013, making it available to every eligible girl from the start, there was no staggered implementation that could have led to anticipatory behavior. Therefore, this universal and simultaneous roll-out of the KP program, combined with the results from the placebo intervention year tests, provides strong evidence against the presence of anticipatory effects, thereby bolstering the credibility of my identification strategy.

Standard Errors and Inference*:* For all the regressions in my analysis, I cluster standard errors at the level of treatment assignment, which is the state level. Abadie et al. (2022) suggest that standard errors should be clustered at the level of treatment assignment as the DD and DDD estimands are dependent on potential outcomes, and the sampling of potential outcomes is determined only by the assignment mechanism. Moreover, clustering at the state level instead of at the state-year level may help to avoid potential issues of serial correlation (Angrist & Pischke, 2009; Bertrand et al., 2004). However, given the small number of clusters in this study (four), standard cluster-robust errors, which rely on asymptotic properties, may lead to incorrect inferences (Roth et al., 2023; Olden & Moen, 2022).

To address the concern of incorrect inferences due to the small number of clusters, a commonly recommended solution is the use of the wild cluster bootstrap method (Cameron et al., 2008). However, MacKinnon and Webb (2018) argue that this method can be problematic in settings with a very small number of treated clusters, such as in DD and DDD designs. Instead, they recommend using a subcluster wild bootstrap procedure, where the bootstrap data generating process (DGP) clusters at a finer level, such as the district or household level. The authors argue that this procedure performs well when there is a small number of treated clusters.

Additionally, in my study, where West Bengal has a significantly different number of observations compared to some of the control states, it is also important to account for the variation in cluster sizes. Therefore, in addition to clustering at the state level in the main specification, I also sub-cluster at the district level, which performs well when cluster sizes vary (MacKinnon & Webb, 2018). Lastly, I conduct inference based on wild bootstrapped p-values that are sub-clustered at district levels. I report standard errors clustered at the state level for completeness but do not use them for inference. Instead, I use the p-values generated from the bootstrap procedure (using the *boottest* command) are used for statistical inference, following the recommendations by Roodman et al. (2018).

IV. Results

A. Direct Effects

(i) Difference-in-differences (DD) Estimation

Enrollment: I first estimate the direct effect of KP on the enrollment rate of eligible girls. Given the policy structure and related literature on CCTs, I hypothesize that the eligible girl should unambiguously experience higher enrollment. Math and reading levels are non-targeted outcomes and represent whether the program also improved learning levels. Table 1 shows results from DD specification (1). Each column represents a layer of additional variables. The inclusion of socio-economic controls and linear trends slightly increases the magnitudes of the coefficients. Moreover, given the significant imbalance in socio-economic factors between the treated and control states, it is important to progressively add controls to account for the selection bias[11]. Therefore, I focus on interpreting the results from column (3). Additionally, I perform a robustness check using state-specific higher order time trends to further validate the results. This robustness check is discussed in detail in Section V.

Table 1 – Difference-in-differences (DD) Estimate of the Impact of KP program on Enrollment of Eligible Girls

	Dependent Variable: Enrollment		
	(1)	(2)	(3)
WB x Post	0.0267	0.0434	0.0567
	(0.022)	(0.000)	(0.004)
	[0.538]	[0.273]	[0.256]
WB	0.0256	0.0061	-0.0162
Post	0.0533	0.0933	-0.0041
Observations	100,063	100,063	75,998
Control group mean	~	~	0.853
R-squared	0.013	0.0252	0.103
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables. I check whether the change in coefficient magnitude is due to the inclusion of control variables or the reduction in sample size by running the regression on the same sub-sample as in column (3), but without the control variables. The coefficient is lower in magnitude, suggesting that the control variables are capturing important variation that would otherwise be underestimated.

[11] I check whether the change in coefficient magnitude is due to the inclusion of control variables or the reduction in sample size (due to missing observations) by running the regression on the same sub-sample as in column (3), but without the control variables. I find the coefficient to be lower in magnitude, suggesting that the control variables are capturing important variation that would otherwise be underestimated.

In Table 1, the main parameter of interest is the interaction term (WB × POST), which shows an estimate of 5.7 percentage points. This indicates that 13–16-year-old girls exposed to KP in West Bengal are 5.7 percentage points more likely to be enrolled in school than girls in control states. Relative to the control group mean for enrollment, which is 85.30%, this effect translates to an approximate 6.6% increase in enrollment for eligible girls exposed to KP. However, the point estimate is insignificant at the conventional levels.

Moreover, conducting the regression analysis while including and excluding 2022 survey data, may be particularly insightful in the context of the COVID-19 pandemic. COVID-19 has been widely reported to have negatively impacted enrollment and learning outcomes (Guariso & Nyqvist, 2023). By conducting analyses both with and without the 2022 data, I aim to explore how the pandemic may have influenced the long-term impact of the KP program. The observed increase in the program's effect (from 5.7 percentage points to 11.7 percentage points) when 2022 data is excluded (as shown in Table A18) may indicate that COVID-19 had an adverse effect on enrollment. However, this interpretation remains tentative, as the missing data for several key years (2019 to 2021) limits my ability to fully test this hypothesis and comprehensively assess KP's effects during the pandemic.

Reading level: The point estimate in column (3) of Table 2 suggests that the eligible girls are 6.6 percentage points less likely to score one level higher in reading than without the program. Relative to the standard deviation (S.D.) of the average reading level, the effect size is -0.06 SD, which is quite small. However, these results are insignificant at the conventional levels. It is important to note that the results include children who are out of school and are not impacted by the selection of children into school.

When excluding 2022 data, the point estimate in column (2) of Table A18 indicates a modest negative effect of 1.62 percentage points on reading outcomes of the eligible girls. However, this result is insignificant at the conventional levels. Therefore, I conclude that the program did not have a meaningful impact on the reading outcomes of the eligible girls when compared to untreated girls in the control states.

Math Level: The point estimate in column (3) of Table 3 suggests that the eligible girls are 9.5 percentage points less likely to score one level higher in math than without the program. Relative to the standard deviation (S.D.) of the average math level, the effect size is -0.09 SD, which is quite small. However, these results are insignificant. I conclude that the program did not have a meaningful impact on the math outcomes of the eligible girls when compared to untreated girls in control states.

When excluding 2022 data, the point estimate in column (3) of Table A18 suggests that the eligible girls are 10.8 percentage points more likely to score one level higher in math than without the program. However, the result is statistically insignificant, and I conclude that excluding 2022 data does not provide evidence of a significant impact of the KP program on math outcomes for eligible girls.

Table 2 – Difference-in-differences (DD) Estimate of the Impact of KP program on Reading Level of Eligible Girls

Dependent Variable: Reading Level			
	(1)	(2)	(3)
WB x Post	-0.0546	-0.1300	-0.0659
	(0.027)	(0.000)	(0.022)
	[0.0496]	[0.271]	[0.690]
WB	-0.0587	-0.2082	-0.3020
Post	-0.0170	0.0329	-0.3385
Observations	84,809	84809	63,628
Control group mean	~	~	3.53
R-squared	0.002	0.0117	0.0859
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

Table 3 – Difference-in-differences (DD) Estimate of the Impact of KP program on Math Level of Eligible Girls

Dependent Variable: Math Level			
	(1)	(2)	(3)
WB x Post	-0.1334	-0.1923	-0.0946
	(0.082)	(0.000)	(0.025)
	[0.473]	[0.184]	[0.643]
WB	-0.0920	-0.1087	-0.2368
Post	-0.3160	-0.3625	-0.8223
Observations	84,630	84,630	63,501
Control group mean	~	~	3.20
R-squared	0.032	0.062	0.145
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

Private investment in education: In addition to examining the direct impacts of KP on enrollment and learning outcomes, it is also important to explore whether the program has led to increased private household investment in the education of eligible girls. One such indicator of private educational investment is the likelihood of girls receiving private tutoring. In India, private tutoring is a common practice outside regular school hours and is often seen as essential for achieving better academic results. However, access to private tutoring is not evenly distributed in India. Children from urban areas, private schools, and better economic backgrounds are more likely to be enrolled in private tutoring (Azam, 2016). Additionally, there is a pro-male bias in access to tutoring. Azam finds that boys are not only more likely to attend private tutoring but are also more likely to receive higher financial investments in tutoring when compared to girls. This bias seems to be driven by the perception that boys have higher returns on educational investments than girls, as boys are traditionally expected to be the primary earners in many households (Azam, 2016).

Given this context, analyzing the impact of KP on private tutoring can provide insights into whether the program has succeeded in shifting household investment patterns towards girls and potentially reduced the existing gender gap in educational support. The following analysis focuses on whether the KP program has increased the probability of eligible girls receiving private tutoring and whether this effect persists when accounting for the disruptions caused by the COVID-19 pandemic.

The point estimate in column (3) of Table 4 suggests that when including data from 2022, the coefficient for the impact of the KP program on the likelihood of eligible girls receiving private tutoring is - 0.0049, or - 0.49 percentage points. However, this result is not statistically significant at the conventional levels. When I exclude the 2022 data, as shown in Table A18, the coefficient increases and becomes positive (0.034 or 3.4 percentage points). Nevertheless, it remains statistically insignificant.

Although the coefficient remains insignificant in both cases, the shift from a small negative value to a positive one when excluding 2022 is noteworthy. Given the control group's mean for private tutoring is 0.30, or 30%, this 3.4 percentage point increase translates to an approximate 11% relative increase in the likelihood of receiving private tutoring. This suggests that when 2022 is excluded, the KP program may have had a more pronounced positive impact on private tutoring. Nevertheless, this interpretation is tentative as the estimate is not statistically significant.

When including 2022, the negative coefficient likely reflects the broader economic disruptions caused by the COVID-19 pandemic. During the pandemic, many poor households faced significant economic uncertainty, which may have impacted their ability to invest in children's education (Andrew & Salisbury, 2023). This uncertainty caused by the pandemic may have dampened the observable effect of the KP program on private tutoring. However, it is important to emphasize that these interpretations are tentative and testing these mechanisms is beyond the scope of this paper due to missing data for key years (2019-2021).

Table 4 – Difference-in-differences (DD) Estimate of the Impact of KP program on Private Tutoring of Eligible Girls

Dependent Variable: Private Tutoring			
	(1)	(2)	(3)
WB x Post	-0.0354	-0.0397	-0.0049
	(0.023)	(0.000)	(0.010)
	[0.428]	[0.129]	[0.948]
WB	0.4774	0.7576	0.7103
Post	0.0191	0.0243	-0.1511
Observations	78,590	78,590	66929
Control group mean	~	~	0.30
R-squared	0.203	0.291	0.365
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

(ii) Triple-difference (DDD) Estimation

Enrollment: I estimate the direct effect of KP on the eligible girls using the DDD approach. Table 5 shows results from DDD specification (3). As discussed in the DD results, the interpretation focuses on results from Column (3). Additionally, I perform a robustness check using state- and gender-specific higher-order tie trends to further validate the results. These robustness analyses are discussed in detail in Section V.

The point estimate in column (3) of Table 5 indicates a modest enrollment gain for the eligible girls of about 0.53 percentage points relative to if the KP had not occurred. Given the control group's mean enrollment rate of 85.29%, this translates to an approximate 0.62% increase in enrollment due to the program. In contrast to the DD results, this estimate is statistically significant at the 90% level, albeit much smaller.

Table 5 – Triple-difference (DDD) Estimate of the Impact of KP program on Enrollment of Eligible Girls

Dependent Variable: Enrollment			
	(1)	(2)	(3)
WB x Post x Female	0.0021	0.0011	0.0053*
	-0.001	-0.001	-0.001
	[0.452]	[0.664]	[0.051]
WB	-0.0402	-0.0615	-0.0794
Post	0.0451	0.0633	-0.0125
Female	-0.0019	0.0066	0.0042
Observations	203,711	203,711	153,019
R-squared	0.012	0.022	0.103
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state and gender- specific. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

Again, it is important to consider the change in the estimate when excluding 2022, which may reflect the negative impact of COVID-19. When including 2022 data, the coefficient is 0.53 percentage points. Excluding 2022 increases the coefficient to 1.77 percentage points. The observed increase in the program's effect when 2022 data is excluded, as shown in Table A19, may indicate that COVID-19 had an adverse effect on enrollment. However, this interpretation remains tentative, as the missing data for several key years (2019 to 2021) limits my ability to fully test this hypothesis and comprehensively assess KP's effects during the pandemic.

Reading Level: Column (3) of Table 6 suggests that the eligible girls are about 4 percentage points more likely to score one level higher in reading than without the program. Relative to the standard deviation (S.D.) of the average reading level, the effect size is 0.04 SD, which is quite small. However, these results are insignificant. This result is similar to the DD results for reading levels. Also, it is important to remember that these results include children who are out of school and are not impacted by the selection of children into school. Moreover, the significant drop in observations for learning outcomes compared to enrollment raises concerns about potential endogeneity due to missing data. I address this concern through robustness checks to ensure the reliability of the results in Section V.

When excluding 2022 data, Column (2) of Table A19 suggests that eligible girls are about 5.9 percentage points more likely to score one level higher in reading than without the program. However, these results are statistically insignificant, and I conclude that the program did not have a meaningful impact on the reading outcomes of eligible girls.

Math Level: Column (3) of Table 7 suggests that the eligible girls are about 6.41 percentage points more likely to score one level higher in math than without the program, and this estimate is significant at the 90% level. Relative to the standard deviation (S.D.) of the average math level, the effect size is 0.06 SD, which is quite modest.

When excluding 2022 data, Column (3) of Table A19 suggests that eligible girls were 8.6 percentage points more likely to score one level higher in math than without the program. This corresponds to an effect size of about 0.08 SD, which is both statistically significant at the 95% confidence level and higher than the estimate obtained when including data up to 2022. This result may indicate that learning outcomes declined due to the disruptions caused by COVID-19. However, this interpretation remains tentative, as the missing data for several key years (2019 to 2021) limits my ability to fully assess KP's impact during the pandemic.

Table 6 – Triple-difference (DDD) Estimate of the Impact of KP program on Reading Level of Eligible Girls

Dependent Variable: Reading Level			
	(1)	(2)	(3)
WB x Post x Female	0.0352	0.0362	0.0399
	(0.011)	(0.008)	(0.006)
	[0.528]	[0.415]	[0.161]
WB	-0.1234	-0.2553	-0.3523
Post	-0.0843	-0.0639	-0.4329
Female	-0.0216	0.0021	0.0046
Observations	168,420	168,420	124,188
R-squared	0.005	0.014	0.0792
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state and gender- specific. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

Table 7 – Triple-difference (DDD) Estimate of the Impact of KP program on Math Level of Eligible Girls

Dependent Variable: Math Level			
	(1)	(2)	(3)
WB x Post x Female	0.0435	0.0446	0.0641*
	(0.020)	(0.011)	(0.008)
	[0.667]	[0.495]	[0.079]
WB	-0.0873	-0.0658	-0.1783
Post	-0.3249	-0.4384	-0.9061
Female	-0.1123	-0.1144	-0.1194
Observations	168,106	168,106	123,961
R-squared	0.037	0.065	0.140
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state and gender- specific. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

Private investment in education: Table 8 shows a positive impact of KP on the likelihood of eligible girls receiving private tutoring. When including data from 2022, the program's impact is estimated at 5.18 percentage points, though this result is not statistically significant. When excluding the 2022 data, as seen in Table A19, the coefficient slightly decreases to 4.74 percentage points but becomes statistically significant at the 95% level.

The higher estimate for private tutoring when including 2022 survey data as opposed to excluding it contrasts with what I found with the DD specification. One possible explanation for the estimate being higher when including 2022 data is that some households, despite the increased financial constraints and heightened economic uncertainty, prioritized the children's private tutoring, perhaps due to concerns over disrupted schooling during the pandemic. Therefore, a lower but significant estimate when excluding 2022 data may mean that this analysis is capturing the consistent impact of the KP program under more stable economic conditions. Nevertheless, these interpretations are tentative and warrant further investigation. Exploring the interaction between economic shocks like COVID-19 and conditional cash transfer programs like KP lies beyond the scope of this paper but offers valuable insights for future research.

Table 8 – Triple-difference (DDD) Estimate of the Impact of KP program on Private Tutoring of Eligible Girls

Dependent Variable: Private Tutoring			
	(1)	(2)	(3)
WB x Post x Female	0.0551*	0.0531	0.0518
	(0.007)	(0.006)	(0.006)
	[0.094]	[0.116]	[0.105]
WB	0.4404	0.7370	0.6788
Post	0.0206	0.0042	-0.1816
Female	-0.0409	-0.0326	-0.0375
Observations	157,443	157,443	133,396
R-squared	0.177	0.273	0.345
Linear Trends	No	Yes	Yes
Socio-economic controls	No	No	Yes

Notes: This table reports results from DD specification (1) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state and gender- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics. The reduction in sample size in column (3) is due to missing observations for several control variables.

B. Indirect Effects

Policy interventions such as conditional cash transfers (CCTs) are designed to directly impact the human capital development of the targeted individuals like children and adolescents who meet specific eligibility criteria. However, the effects of these programs often extend beyond the directly treated individuals and influence the broader household dynamics and resource allocation (Ferreira et al., 2009). While the primary focus of much of the literature on CCTs has been on their direct effects on the beneficiaries, there is a growing recognition of the importance of understanding the indirect or spillover effects on ineligible siblings within the same household.

Based on a simple model of schooling decisions discussed by Ferreira et al. (2009). I now discuss the possible household dynamics when one of the siblings directly benefits from a cash transfer and ineligible siblings do not. The authors predict that a child-specific cash transfer conditional on enrollment will unambiguously increase enrollment among eligible children. However, the cash transfer has an ambiguous effect on school enrollment among ineligible siblings. Their model highlights three different effects of a child-specific CCT: a substitution effect, an income effect, and a displacement effect. The cash transfer subsidizes the cost of schooling for the eligible child in a financially constrained household. This, in turn, leads to higher enrollment of the eligible child via substitution effect as they are substituting away from work. In households that can only afford to enroll one child in school, either simply due to poverty or needing the child for household or agricultural work, the CCT causes a displacement effect. This is where the eligible child replaces the ineligible child in school to meet the program's conditionality. Lastly, if CCT lifts the financial barrier on the parents by raising them above the income threshold so they can afford to send more than one child to school, they can send both eligible and ineligible siblings to school. All three effects predict an unambiguous increase in the likelihood of the eligible sibling enrolling in school. However, the impact on ineligible siblings remains ambiguous. The displacement effect predicts a negative effect of the CCT on the likelihood of school enrollment of ineligible siblings, while the income effect predicts a positive effect. The substitution effect does not affect the likelihood of school enrollment of ineligible siblings.

Consequently, a large body of literature on CCTs has found a positive effect on school enrollment of the eligible sibling (Baird et al., 2010; Ferreira et al., 2009). However, the evidence on the impact of CCTs on ineligible siblings remains ambiguous, as predicted by the model discussed above. For instance, Lincove and Parker (2015) find a positive spillover effect of a CCT on school enrollment of ineligible girl siblings but not for boy siblings, while Camilo and Zuluaga (2022) and Barrera-Osorio et al. (2011) find a negative spillover effect on school enrollment for both girl and boy ineligible siblings. Some studies, like Ferreira et al. (2009), find no impact of the CCT on school enrollment of ineligible siblings. Another important aspect is whether these dynamics translate to learning outcomes and private household investment in education (private tuition) of the ineligible siblings. Not many studies investigate these dynamics, which may be due to a lack of standardized data on cognitive skills or test scores. Nevertheless, the existing studies do not find any significant effects of CCTs on learning outcomes and test scores of ineligible siblings (Fiszbein et al., 2009; Gaentzsch, 2020).

In the context of the KP program, studying these spillover effects on school enrollment, learning outcomes, and household investment decisions is particularly important as it can inform future policy designs aimed at improving the overall well-being of all children in the treated households. Moreover, the following analysis also contributes to the relatively under-explored literature on the spillover effects of CCTs in India.

In my spillover analysis, I focus on households that include both an eligible girl (aged 13-16) and at least one ineligible younger sibling (aged 9-12)[12]. I then compare the educational outcomes of these ineligible younger siblings, indirectly affected by the KP program, to similarly aged children in comparable households from control states. By doing this, I intend to assess the broader impact of the KP program on ineligible children within the same household in West Bengal. Specifically, I aim to estimate the effect of having a KP eligible older sister on enrollment, learning outcomes (reading and math levels), and private educational investment (private tuition) of these younger siblings. To analyze the spillover effects of the KP program on younger siblings in households with eligible girls, I employ the following DD regression model:

$$(3)\ y_{ist} = \beta_0 + \beta_{1.}\,\mathrm{WB}_s + \beta_2.\,post_t + \beta_3\,(post_t \times \mathrm{WB}_s) + \vartheta X_{ist} + \gamma_t + (\partial_s \times t) + \epsilon_{ist}$$

Here, the sample is restricted to households with at least one eligible girl (13-16 years old) and at least one younger sibling (9-12 years old). Similar to previous specifications, y_{ist} is the outcome variable of interest (enrollment, learning outcomes, and private tuition). 'WB$_s$' is a dummy that takes the value 1 if the ith observation comes from West Bengal. The '$post_t$' dummy compares the outcomes for the post-program years (after 2013) when it takes the value of 1 to the same before 2013 when it takes the value of 0. X_{ist} includes child-, household-, and village-level characteristics. γ_t is the year fixed effects and $(\partial_s \times t)$ is the state-specific linear time trends. The coefficient of interest in this regression specification is β_3, which represents the interaction term (WB x post). This coefficient captures the Intent-to-Treat (ITT) effect of the KP program on the younger, ineligible siblings in households that contain an eligible girl in West Bengal. Essentially, β_3 assesses whether the presence of an eligible older sister leads to any changes in the educational outcomes of

[12] The parallel trends assumption for all outcomes are satisfied for both boy and girl ineligible siblings and shown in appendix in Tables A6 and A7.

her younger siblings compared to similar households in other states. This analysis sheds light on the broader impacts of the KP program on household dynamics and educational investments beyond its effects on eligible girls in West Bengal.

Enrollment: Column (1) of Table 9 suggests that the ineligible younger girl siblings experience a negative spillover effect on their enrollment. The coefficient suggests that the ineligible younger girl siblings in West Bengal are 0.85 percentage points less likely to be enrolled in school compared to those in control states. However, the point estimate is insignificant at conventional levels. In contrast, column (2) of Table 9 suggests that that the ineligible younger boy siblings experience a positive and significant spillover effect on their enrollment. Specifically, the ineligible younger boy siblings in West Bengal are 4.23 percentage points more likely to be enrolled in school compared to those in control states. Relative to the control group mean for enrollment for this analysis, which is 96.4%, this effect translates to an approximate 4.4% increase in enrollment for ineligible younger boy siblings indirectly exposed to KP.

This positive spillover effect on boys could be driven by an income effect. If a CCT raises the household income above a certain threshold, parents may have enough financial resources to send both the eligible older girl and the ineligible younger boy to school. The increased household income due to the CCT might lead to a reallocation of resources, benefiting boys as well, especially in households with a pro-male bias.

On the other hand, the negative impact on enrollment of younger girl siblings suggests a possible displacement effect. Parents may be reallocating resources toward the education of the eligible older girls at the expense of their younger sisters. With limited resources, parents may prioritize the schooling of the older girl to ensure meeting the program conditionalities which may reduce the likelihood of the younger girl attending school. Additionally, the decrease in enrollment for younger girl siblings could be due to the increased likelihood of them taking on household or agricultural labor responsibilities as their older sisters are now attending school more regularly due to the program's conditionalities. This is consistent with findings from studies on other CCTs, such as the "Más Familias en Acción" program in Colombia. Camilo and Zuluaga (2022) find that non-beneficiary girl siblings for this CCT in Colombia were at a higher risk of being involved in child labor than boy siblings as the household's labor needs persisted despite the older siblings' school attendance requirements. However, my estimate for enrollment of younger girl siblings is insignificant. Therefore, this interpretation remains tentative.

Lastly, the positive spillover observed for boys but not girls could be due to the pre-existing pro-male bias in education in rural India. This bias may lead parents to allocate additional resources from the CCT to the education of boys over girls, even if both are ineligible for KP program benefits.

Table 9 – Difference-in-differences (DD) Estimate of the Impact of KP program on Enrollment of Ineligible Younger Siblings

Dependent Variable: Enrollment		
	(1)	(2)
	Girls	Boys
WB x Post	-0.0085	0.0423**
	(0.001)	(0.001)
	[0.351]	[0.019]
WB	0.0034	-0.0382
Post	0.0181	-0.0303
Observations	28,547	31,358
Control group mean	0.959	0.964
R-squared	0.032	0.029
Linear Trends	Yes	Yes
Socio-economic controls	Yes	Yes

Notes: This table reports spillover analysis results from DD specification (3) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Learning Outcomes: For learning outcomes, the coefficients for reading and math levels in Tables 10 and 11 are negative, but insignificant for both girls and boys. For reading levels, the coefficient for younger girls is -0.2396, while for boys it is -0.0167. Relative to the standard deviation (S.D.) of the average reading level, the effect size is -0.19 SD for girls and -0.01 SD for boys. Similarly, for math, the coefficients are -0.1926 for girls and -0.2329 for boys. Relative to the standard deviation (S.D.) of the average math level, the effect size is -0.17 SD for girls and -0.20 SD for boys. These results suggest that the presence of an eligible older sister may be associated with worse learning outcomes for younger siblings. Nevertheless, these interpretations remain tentative as the coefficients are insignificant.

Table 10 – Difference-in-differences (DD) Estimate of the Impact of KP program on Reading Level of Ineligible Younger Siblings

Dependent Variable: Reading Level		
	(1)	(2)
	Girls	Boys
WB x Post	-0.2396	-0.0167
	(0.017)	(0.026)
	[0.187]	[0.964]
WB	-0.2580	-0.3388
Post	-0.5853	-0.7515
Observations	24,749	26,633
Control group mean	2.99	3.03
R-squared	0.156	0.150
Linear Trends	Yes	Yes
Socio-economic controls	Yes	Yes

Notes: This table reports spillover analysis results from DD specification (3) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Table 11– Difference-in-differences (DD) Estimate of the Impact of KP program on Math Level of Ineligible Younger Siblings

Dependent Variable: Math Level		
	(1)	(2)
	Girls	Boys
WB x Post	-0.1926	-0.2329
	(0.014)	(0.011)
	[0.177]	[0.130]
WB	-0.0181	-0.0433
Post	-0.7093	-0.7414
Observations	24,668	26,546
Control group mean	2.66	3.03
R-squared	0.171	0.170
Linear Trends	Yes	Yes
Socio-economic controls	Yes	Yes

Notes: This table reports spillover analysis results from DD specification (3) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Table 12– Difference-in-differences (DD) Estimate of the Impact of KP program on Private Tutoring of Ineligible Younger Siblings

Dependent Variable: Private Tutoring		
	(1)	(2)
	Girls	Boys
WB x Post	-0.0297	-0.0109
	(0.011)	(0.010)
	[0.503]	[0.834]
WB	0.6469	0.6527
Post	-0.1415	-0.1681
Observations	25,418	28,044
Control group mean	0.243	0.289
R-squared	0.328	0.313
Linear Trends	Yes	Yes
Socio-economic controls	Yes	Yes

Notes: This table reports spillover analysis results from DD specification (3) when including 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

***Private investment in education*:** For private tuition investment, the coefficients are negative and insignificant for both girls (-0.0297) and boys (-0.0109) as seen in Table 12. This suggests that KP may be associated with reducing private educational investments for younger siblings, particularly girls. However, this interpretation is tentative as the estimates are insignificant at the conventional levels.

C. Discussion

The difference in results between the DD and DDD specifications provides important insights into the effectiveness of the KP program. In the DD analysis, where I compare the educational outcomes of 13–16-year-old eligible girls in West Bengal with those in the control states before and after the program, the results are insignificant. One potential explanation could be the choice of control states (Odisha, Chhattisgarh, and Jharkhand), which may not serve as true controls. It is possible that these states have implemented their own education or gender-targeted programs, which may not have been properly documented or publicized on the government websites. Nevertheless, I do a meticulous search for the list of education and gender-targeted policies in these states and find only one policy in Chhattisgarh, which may impact some of the outcomes I look at in my study. The Saraswati Bicycle Supply Scheme was introduced in 2004-05 in Chhattisgarh and aimed to provide bicycles to socio-economically disadvantaged girls from grades 9 to 12. As demonstrated by Muralidharan and Prakash (2017), a conditional-kind transfer (CKT) scheme such as Saraswati Bicycle Supply Scheme that targets adolescent girls may significantly impact their educational outcomes. Nevertheless, as a robustness check, I use the Synthetic Control Method (SCM) to create a synthetic version of West Bengal that closely matches the treated state based on some pre-intervention characteristics by using other Indian states as the donor pool. The results from the SCM analysis match my main findings from the DD regression specification (1), which provides further confidence in the validity of my choice of the original control states.

Meanwhile, the DDD results show a different picture. In the DDD analysis, I compare Group A (13–16-year-old girls) and Group B (13–16-year-old boys) in the treatment state (West Bengal) to Group A (13–16-year-old girls) and Group B (13–16-year-old boys) in the control states before and after the KP implementation. Using this estimation method, I find positive and statistically significant, albeit very small, effects on enrollment and math level of eligible girls. Decomposing the DDD estimand into two double differences as shown in Tables A8 and A9 may offer a clearer insight into KP's gender specific impact. The first double difference compares the educational outcomes for 13–16-year-old girls and boys in West Bengal before and after the KP implementation, while the second double difference compares the same outcomes for 13–16-year-old girls and boys in the control states before and after the KP implementation. By subtracting the second double difference from the first, the DDD approach allows us to interpret the observed results as evidence that the KP program has modestly reduced the gender gap in education in West Bengal compared to the control states. This result is very important given the significant pro-male bias in education that is still rampant, especially in rural India (Muralidharan & Prakash, 2017; Azam & Kingdon, 2011). The positive and significant results from the DDD specification might be interpreted as evidence that the KP program is not only promoting school enrollment among adolescent girls but may also be addressing broader social and cultural norms that have traditionally favored boys over girls. Additionally, the importance of modest improvements in basic numeracy skills of the eligible girls cannot be understated as better foundational literacy and numeracy skills are often associated with higher likelihood of finishing school, future employability and higher wages for young people (Meehan et al., 2023; McIntosh & Vignoles, 2001). While examining the long-term effects of KP on labor market outcomes is beyond the scope of this study, it remains an important avenue for future research.

However, it is important to mention a possible threat to the DDD estimation strategy. There is a possibility of intra-household spillovers to 13–16-year-old boys in West Bengal who have a KP-eligible sister in the same household. This may be concerning because it may be the case that boys may drop out of school to replace their sisters to do household chores, which would result in the DDD estimates being upward biased. However, the chances of this dynamic are quite low given the strong patriarchal culture in West Bengal and the control states. Moreover, several studies find a strong pro-male bias in secondary education (Muralidharan & Prakash 2017; Azam & Kingdon, 2011). A more likely possibility is that boys are more likely to remain in school by observing their sisters staying enrolled in school or due to the cash transfer alleviating some financial barriers for the family that may lead to increased enrollment of not just the beneficiary girl child but also the non-beneficiary male sibling (income effect). I also observe a similar dynamic during the spillover analysis of KP where I find a positive effect on enrollment of ineligible younger boys. Therefore, to the extent that there are positive spillovers from 13–16-year-old eligible girls to 13–16-year-old ineligible boys, the estimated effects using DDD are likely to be the lower bound of the true effect.

The spillover analysis for ineligible younger siblings provides a further understanding of the broader impact of KP beyond its direct effect on the eligible girl. The findings suggest that ineligible younger boy siblings may experience positive spillover effects on enrollment due to the KP program. This effect is potentially driven by an income effect as discussed in the theoretical framework discussed by Ferreira et al. (2009) and a pro-male bias in education within the household, as proposed by Azam & Kingdon (2011). In contrast, younger girls may face negative spillover effects on their enrollment, potentially due to a displacement effect. However, this interpretation is tentative as the point estimate is insignificant. Moreover, given the insignificance of my estimates for learning outcomes and private investment in education, I conclude that I do not observe any significant spillover effects due to KP on these outcomes for ineligible younger siblings. Nevertheless, I do find evidence of a modest positive spillover effect on school enrollment for younger boy siblings who are indirectly exposed to KP through their eligible older sisters.

Another important dimension to consider is the potential impact of the COVID-19 pandemic on my overall findings. Many studies have found that the pandemic caused widespread disruptions to children's education due to school closures and heightened economic uncertainty among households (Jena, 2020; van Cappelle et al., 2021; Guariso & Nyqvist, 2023). Consequently, my estimates are higher when I include data only until 2018 than when I include 2022. This may suggest that COVID-19 may have dampened KP's impact on the educational outcomes of eligible girls. However, the absence of data for 2019–2021 limits my ability to fully assess how the pandemic influenced the long-term effects of KP. Therefore, these interpretations remain tentative. Nevertheless, further studies could delve deeper into how economic disruptions caused by health emergencies like COVID-19 interact with policies like KP, as these shocks may temporarily reduce the effectiveness of such programs. Understanding these dynamics will be important to ensure that cash transfer policies like KP continue to support children's educational outcomes even in the face of unforeseen challenges.

D. Heterogeneity

The impact of the KP program may vary across different socio-economic and infrastructural contexts. To better understand these variations, I conduct a heterogeneity analysis focusing on two key dimensions: maternal education and the availability of schools in the child's village.

The first dimension focuses on the educational background of the child's mother. Specifically, I consider whether the mother has received any education. Literature shows that maternal education is strongly linked to higher likelihood of school enrollment for children, especially girls (Sathar & Lloyd, 1994; Dutta, 2014). This suggests that the effect of the KP on enrollment might be stronger for girls whose mothers have some education compared to the girls whose mothers have no education. This difference in enrollment could potentially influence learning outcomes as well. Additionally, maternal education is associated with better learning outcomes for children (Harding et al., 2015). Therefore, it is important to examine the treatment effect heterogeneity across this dimension, as girls with different levels of maternal education may experience the program’s effects differently.

The second dimension concerns the child's proximity to a local school. Distance to school is often described as one of the most substantial supply-side constraints on female adolescent education in developing countries (Sipahimalani, 1999; Lavy, 1996; Glick & Sahn, 2007). This suggests that the effect of KP on eligible girls’ enrollment might be stronger for girls who live closer to schools. Consequently, this difference in enrollment could potentially influence learning outcomes as seen in Peteros et al., 2022. Therefore, it is important to examine the treatment effect heterogeneity based on eligible girls' proximity to schools as the girls living closer to schools may experience the program's effects differently than the girls who live far away from them.

To explore these dimensions, I divide the sample based on the mother's level of education and the presence of a school in the village. I then generate the treatment effects using the DD and DDD specifications as outlined earlier. The results of this analysis are presented in Tables 13 and 14.

(i) Maternal Education

The first dimension of heterogeneity analysis considers the level of household disadvantage of the eligible girl, specifically the mother's educational level. Panel A1 represents girls whose mothers have no education and Panel A2 represents girls whose mothers have some education.

Enrollment**:** For the DD specification in Table 13, the coefficient for girls whose mothers have no education (Panel A1) is 0.142, while the coefficient for girls whose mothers have some education (Panel A2) is 0.007. However, none of these estimates are statistically significant at the conventional level. Nevertheless, the magnitude and direction of the coefficients suggest that KP may be more effective in increasing enrollment rates for more disadvantaged girls whose mothers have no education than girls whose mothers do have some education.

In comparison, for the DDD specification in Table 14, there is a significant impact on enrollment for girls from more disadvantaged backgrounds. i.e., girls whose mothers have no education. In Panel A1, the coefficient is 3.39 percentage points, indicating a substantial increase in enrollment for these girls. In contrast, Panel A2 shows a much smaller effect of 0.12 percentage points, which is not statistically significant. This suggests that KP may be particularly effective in increasing school enrollment among more disadvantaged girls. The control group baseline enrollment rate for girls with educated mothers is 94.36%, while for girls with uneducated mothers, it is only 82.41%. This stark difference highlights the program's ability to make a meaningful impact, particularly among girls who face greater disadvantages due to their mothers' lack of education.

Reading Level: For the DD specification (1) in Table 13, the coefficient for girls whose mothers have no education (Panel A1) is -0.1379 or an effect size of -0.12 SD, and for girls whose mothers have some education (Panel A2), the coefficient is -0.1910 or an effect size of -0.25 SD. Neither of these results is statistically significant. The negative sign is consistent with the overall trend observed in the DD specification, where KP was linked to a decrease in reading levels among eligible girls, although these results were not statistically significant. While maternal education is generally associated with better learning outcomes for children, in this context, it appears that a mother's education alone may not be sufficient to enhance the daughter's learning levels. Nevertheless, since the coefficients are insignificant, these interpretations remain tentative.

In contrast, for the DDD specification (2) in Table 14, the results indicate a larger effect in Panel A1 than Panel A2. Eligible girls whose mothers have no education are 8.63 percentage points more likely to score one level higher in reading with the KP program compared to those without the program which translates to an effect size of 0.08 SD. In comparison, eligible girls whose mothers have some education, are only 1.47 percentage points more likely to score one level higher which translates to an effect size of 0.02 SD. However, neither result is statistically significant. While these findings suggest that the program may have a stronger impact on reading skills among the relatively more disadvantaged girls, the lack of statistical significance means these effects are not robust enough to draw definitive conclusions. Overall, I avoid inferring a general trend in KP's impact on reading levels across different levels of maternal education.

Math Level: For the DD specification (1) in Table 13, the coefficient for girls whose mothers have no education (Panel A1) is -0.1453 or an effect size of -0.12 SD, and for girls whose mothers have some education (Panel A2), the coefficient is -0.1783 or an effect size of -0.19 SD. Again, neither of these results is statistically significant. The negative sign is consistent with the overall trend observed in the DD specification, where KP was linked to a decrease in math levels among eligible girls, although these results were also not statistically significant. Similar to the results for reading levels, maternal education did not lead to better learning outcomes for daughters, as the coefficients remained negative. Nevertheless, since the coefficients are insignificant, these interpretations remain tentative.

In contrast, for the DDD specification (2) in Table 14, the impact on math levels is more conclusive. In Panel A1, the positive and significant point estimate of 0.1108 suggests that the KP program led to an 11.08 percentage point increase in the probability of more disadvantaged girls scoring one level higher in math or an effect size of approximately 0.1 SD. Panel A2, however, shows a smaller point estimate of 0.0281 or an effect size of 0.03 SD, which is not significant at any conventional level. This indicates that the program has a particularly strong effect on improving math skills among girls who are more disadvantaged in terms of their mother's education level.

Private investment in education: For the DD specification (1) in Table 13, the coefficient for girls whose mothers have no education (Panel A1) is -0.0872, while for girls whose mothers have some education (Panel A2), the coefficient is -0.0347. Neither of these estimates is statistically significant. The negative sign is consistent with the overall trend observed in the DD specification, where KP was linked to a decrease in the likelihood of receiving private tuition among eligible girls, although these results were also not statistically significant.

In contrast, for the DDD specification (2) in Table 14, there is a significant positive impact on the probability of receiving private tutoring for more disadvantaged girls whose mothers have no education. The point estimate in Panel A1 for girls whose mothers have no education is 0.0619, which is statistically significant at the 90% level. Meanwhile, in Panel A2, for girls whose mothers have some education, the point estimate is 0.0322 and is insignificant at the conventional levels.

The DDD results suggest that the KP program may be particularly beneficial for girls from more disadvantaged backgrounds whose mothers have no education. The significant increase in enrollment, math outcomes, and private tutoring for these girls, as observed in the DDD specification, may be attributed to their having fewer educational opportunities and resources to begin with, as seen in much lower baseline enrollment rates for these girls (82.41%). The financial support and incentives provided by the program likely play a crucial role in enabling these girls to continue their education, which they might otherwise have been forced to leave behind due to financial constraints. Moreover, the larger effects observed in math outcomes suggest that consistent and organized schooling as a condition for eligibility in the KP program may have led to significant improvements in math skills among these disadvantaged girls. These results suggest that the KP program is more successful in reducing the gender gap in education in more disadvantaged groups.

Conversely, the smaller and often insignificant effects observed for girls whose mothers have some education indicate that while the program is still beneficial, its impact may be less dramatic in households where the baseline educational attainment and resources are already relatively higher. These findings suggest that targeted interventions like KP are crucial for reducing socio-economic disparities and helping the most disadvantaged groups.

Table 13– Difference-in-differences (DD) Estimates of the Heterogenous Impact of KP program on Eligible Girls Based on Maternal Education and Proximity to Schools

	(1)	(2)	(3)	(4)
Dependent Variables	Enrollment	Reading Level	Math Level	Private Tuition
Panel A1: Mother School No				
WB x Post	0.1420	-0.1379	-0.1453	-0.0872
	[0.250]	[0.516]	[0.548]	[0.281]
Observations	39,477	32,378	32,299	33,176
Panel A2: Mother School Yes				
WB x Post	0.0073	-0.1910	-0.1783	-0.0347
	[0.798]	[0.246]	[0.239]	[0.400]
Observations	36,521	31,250	31,202	33,753
Panel B1: Schools Present No				
WB x Post	0.0694	-0.0379	-0.0613	0.0127
	[0.335]	[0.818]	[0.784]	[0.803]
Observations	28,282	23,536	23,487	24,766
Panel B2: Schools Present Yes				
WB x Post	0.0459	-0.0939	-0.0848	-0.0203
	[0.149]	[0.750]	[0.712]	[0.844]
Observations	51,237	43,080	42,991	44,999

Notes: This table reports results from DD Specification (1) for eligible girls. Each cell reports the coefficient on the WB - Post interaction term. Each column represents a specific outcome variable, and the panels show the dimensions of heterogeneity. Panel A varies whether the child's mother went to school and Panel B varies whether a child’s village has a middle, secondary or private school. p-values from the wild bootstrap procedure sub-clustered at the district level are reported in brackets.

(ii) Proximity to Schools

The second dimension of heterogeneity analysis examines the impact of KP based on a child's proximity to schools. Panel B1 represents villages without government middle or secondary school or private school. Panel B2 represents villages where at least one of these schools exists.

Enrollment: For the DD specification (1) in Table 13, the coefficient for villages without any school (Panel B1) is 0.0694, while for villages with at least one school (Panel B2), the coefficient is 0.0458. However, neither of these estimates is statistically significant. A slightly lower coefficient for when at least one school is present in the village is counterintuitive and may mean that the presence of a school in a village may not be enough to drive the increase in enrollment among eligible girls. Nevertheless, this interpretation remains tentative because the estimates are insignificant.

In contrast, for the DDD specification (2) in Table 14, the results show that the presence of a school increases the impact on enrollment, with a coefficient of 0.70 percentage points in Panel B1, compared to 0.20 percentage points in Panel B2. However, neither of these effects is statistically significant, suggesting that while the presence of a school might marginally improve enrollment, the observed effect is not strong enough to be conclusive.

Reading Level: For the DD specification (1) in Table 13, the coefficient for villages without any school (Panel B1) is -0.0379 or an effect size of -0.03 SD, and for villages with at least one school (Panel B2), the coefficient is -0.0910 or an effect size of -0.09 SD. Neither of these results is statistically significant. The negative signs are consistent with the overall trend observed in the DD specification, where KP was associated with a decrease in reading levels among eligible girls. This trend persists regardless of the presence of a school, indicating that simply having a school nearby is not sufficient to improve reading outcomes. Nevertheless, these interpretations remain tentative, as the coefficients are not statistically significant.

In contrast, for the DDD specification (2) in Table 14, the presence of a school has a significant and positive impact on the reading levels of the eligible girls. Panel B2, where a school is present, shows a substantial positive impact of 8.44 percentage points or an effect size of 0.09 SD, statistically significant at the 95% level. In contrast, Panel B1, where no school is present, shows a very small but negative impact of -0.39 percentage points or an effect size of 0.002 SD, though this result is not significant. The significant positive effect in Panel B could be because the proximity to schools facilitates better access to educational resources, which in turn enhances reading outcomes.

Math Level: For the DD specification (1) in Table 13, the coefficient for villages without any school (Panel B1) is -0.0613 or an effect size of -0.05 SD and for villages with at least one school (Panel B2), the coefficient is -0.0848 or an effect size of -0.08 SD. Again, neither of these results is statistically significant. The negative sign is consistent with the overall trend observed in the DD specification, where KP was linked to a decrease in math levels among eligible girls, although these results were also not statistically significant. Similar to the results for reading levels, the presence of a school does not appear to lead to better math outcomes. Nevertheless, these interpretations remain tentative, as the coefficients are not statistically significant.

In contrast, for the DDD specification (2) in Table 14, Panel B2 shows a significant positive effect of 8.81 percentage points or an effect size of 0.08 SD in villages with schools, which is significant at the 95% level. However, Panel B1 shows a smaller, insignificant positive effect of 1.24 percentage points or an effect size of 0.01 SD. This result suggests that the presence of a school modestly enhances the program's impact on math outcomes. This is likely due to the presence of formal schools increasing the likelihood of attending school and supporting better learning environments and outcomes.

Private investment in education: For the DD specification (1) in Table 13, the coefficient for villages without any school (Panel B1) 0.0127 and for villages with at least one school (Panel B2), the coefficient is -0.0203. Neither of these estimates is statistically significant. The negative sign is consistent with the overall trend observed in the DD specification, where KP was linked to a decrease in the likelihood of receiving private tuition among eligible girls, although these results were also not statistically significant.

In contrast, for the DDD specification (2) in Table 14, there is a significant positive impact on the probability of receiving private tutoring for girls in villages with at least one school. The point estimate in Panel B2 is 0.055, which is statistically significant at the 95% level. Meanwhile, in Panel B1 where the village has no school, the point estimate is 0.054, which is not significant at any conventional level.

In sum, the modest positive effects on learning outcomes and private tutoring, as observed in the DDD specification (2), suggest that KP has a more significant impact on reading, math outcomes, and private educational investments in villages where schools are present. One possible explanation is that proximity to a formal school facilitates easier enrollment and regular attendance, which may improve learning outcomes and increase the demand for private tuition (Peteros et al., 2022). Nevertheless, the weaker and sometimes negative effects observed in villages without schools suggests that the lack of educational infrastructure poses significant barriers that the KP program alone may not be able to overcome.

Table 14– Triple-difference (DDD) Estimates of the Heterogenous Impact of KP program on Eligible Girls Based on Maternal Education and Proximity to Schools

	(1)	(2)	(3)	(4)
Dependent Variables	Enrollment	Reading Level	Math Level	Private Tuition
Panel A1: Mother School No				
WB x Post x Female	0.0339**	0.0863	0.1108**	0.0619*
	[0.0069]	[0.428]	[0.011]	[0.064]
Observations	80,428	63,941	63799	66808
Panel A2: Mother School Yes				
WB x Post x Female	0.0012	0.0147	0.0281	0.0322
	[0.741]	[0.731]	[0.587]	[0.486]
Observations	72,591	60,247	17,961	20,299
Panel B1: Schools Present No				
WB x Post x Female	0.0020	-0.0039	0.0124	0.0544
	[0.675]	[0.927]	[0.899]	[0.117]
Observations	57,149	46,109	46,011	49,396
Panel B2: Schools Present Yes				
WB x Post x Female	0.0070	0.0844**	0.0881**	0.0545**
	[0.222]	[0.036]	[0.033]	[0.034]
Observations	103,058	84,109	83,968	89,771

Notes: This table reports results from DD Specification (1) for eligible girls. Each cell reports the coefficient on the WB - Post interaction term. Each column represents a specific outcome variable, and the panels show the dimensions of heterogeneity. Panel A varies whether the child's mother went to school and Panel B varies whether a child's village has a middle, secondary or private school. p-values from the wild bootstrap procedure sub-clustered at the district level are reported in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

In areas without schools, the program's financial support alone may not be able to fully address the challenges that adolescent girls face in accessing education. The absence of schools in a village creates significant barriers, including safety concerns for girls traveling longer distances. These concerns can be particularly daunting in rural settings in a developing country like India. Muralidharan and Prakash (2017) argue that proximity to schools is crucial to improve the enrollment of girls in secondary education. Having a school nearby reduces the time and safety concerns associated with attending school and helps adolescent girls remain in school (Muralidharan & Prakash, 2017). The lack of nearby schools may also exacerbate parental reluctance to invest in girls' education, which may be driven by societal norms and financial considerations. As Dutta and Sen (2020) explain, parents often discount the future benefits of educating daughters much more than their sons. They tend to view the immediate costs (both financial and social) of educating daughters as too high, especially when the returns seem uncertain (Azam, 2016). In such cases, the perceived risks of sending girls to distant schools may outweigh the potential educational gains, which may limit the effectiveness of programs like KP.

Overall, the results from the heterogeneity analysis based on proximity to school shows that KP is more successful in reducing the gender gap in education in areas with better access to educational infrastructure.

V. Robustness Checks

Changing the Intervention Years*:* The first robustness check involves reassigning the intervention to alternate placebo years. Using the ASER data, I restrict my sample to 2008 to 2013. Next, I reassign the intervention to 2010 and 2011 as the placebo policy years. The control states remain the same. Now, if my results are indeed capturing the causal impact of KP on the educational outcomes of the eligible girls, then I should not find any significant effects in the placebo regressions. Tables A10 and A11 in the appendix show the results of this robustness check using both DD and DDD specifications. Indeed, I do not find any significant effects on any of the outcomes, giving credibility to my estimation strategy and findings.

Placebo Outcome Variables: This test focuses on using alternative outcome variables that should not be affected by the KP program. Specifically, I examine the highest level of education attained by the mother and father of the child (measured by the highest class attended) as placebo outcomes. These outcomes are unlikely to be influenced by the KP program since the educational attainment of parents is typically determined long before the program's implementation and is not directly targeted by the intervention.

The results for DD and DDD specifications in Tables A12 and A13 in the appendix show no significant effects on either the mother's or father's educational attainment. As anticipated, the KP program did not have any observable impact on these outcomes, confirming that the program's effects are specific to the intended target group and do not extend to unrelated variables. This finding further supports the credibility of the study's results by demonstrating that the KP program's impact is not merely a statistical artifact affecting other aspects of household dynamics unrelated to the program.

Quadratic and Cubic Trends: I conduct additional analyses by including higher-order trends in the DD and DDD specifications. For the DD specification, I replaced state-specific linear trends with quadratic and cubic trends to capture potential non-linear state-level changes over time that might influence the outcomes. For the DDD specification, I similarly introduced quadratic and cubic trends and applied them to state-specific and gender-specific trends to account for more complex dynamics that could affect the interaction between treatment and control groups.

The results of these robustness checks are presented in Tables A14 and A15 in the appendix. For the DD specification, none of the key interaction terms (WB x Post) were significant, consistent with the findings from my main results in Tables 1 to 4. This suggests that the inclusion of higher-order state-specific trends does not alter the main conclusions of the analysis and reinforces the robustness of the findings to changes in the trend specification.

For the DDD specification, the results remained consistent with the main findings, with most of the coefficients comparable in magnitude and significance to those observed in the original analysis in Tables 5 to 8. This stability across specifications indicates that the positive effects observed in the main results are not driven by the choice of trend specification, further validating the reliability of the estimates.

Overall, these robustness checks suggest that the results are not sensitive to the inclusion of higher-order trends. The results remain consistent across different model specifications, which reinforces my confidence in the initial findings.

Missing Observations: The number of observations drops significantly when shifting from enrollment to learning outcomes as the dependent variable. This is because a substantial portion of children in the ASER dataset is missing data on learning outcomes. It is important to note that the learning outcomes data includes children who are out of school, so there's no concern about selection bias related to school attendance. However, the missing data might still bias my results if it's not missing at random. For instance, if children with missing data on learning outcomes systematically come from the treated state after the program was implemented, then it would bias my results. To address this, I created a missing dummy variable for each learning outcome (reading and math), where the variable is 1 if the data is missing for a child and 0 otherwise. I then used these dummy variables as outcome variables in both the DD (1) and DDD (2) specifications. The results in Tables A16 and A17 in the appendix show that the coefficients for the key interaction terms (WB x Post) and (WB x Post x Female) were insignificant, which means that the missing data is unlikely to be systematically missing and that it does not significantly affect my findings.

Synthetic Control Method: In my DD and DDD specifications, I chose the control states based on their geographical, cultural, and socio-economic proximity to West Bengal. However, pre-existing differences between the treated and control states could still bias the estimates. To address this possibility, I employ the Synthetic Control Method (SCM), following the approach used in the literature (Abadie & Gardeazabal, 2003; Abadie et al., 2010)

SCM allows me to compute a synthetic control state, which serves as a counterfactual to the treated state - West Bengal. By conducting a graphical analysis of the treatment effect, I observe the differences between the synthetic control state and the actual treated state across the years. SCM may offer a superior strategy compared to the classic DD approach because it uses a data-driven method to create a linear combination of states from a suitable donor pool of control states, resulting in a synthetic state that closely resembles West Bengal (Abadie, 2021).

For this analysis, I restrict the donor pool to states that did not implement any education-related CCT programs[13]. The synthetic state is constructed based on predefined criteria, including socio-economic household and village-level variables[14] and some lagged outcomes. To build the Synthetic West Bengal, I use ASER data from 2008 to 2022, collapsing it into a state-year panel by taking the average of these variables for each state in each year.

Figures A6 – A8 in the appendix plot the trends of mean enrollment and learning outcomes of the eligible girls for West Bengal and synthetic West Bengal. The results of the synthetic control methodology reaffirm my main findings from the DD specification (1). The synthetic state matches West Bengal before 2013 and then diverges afterward. Panel A shows an increase in the enrollment rate in the post-policy period. In contrast,

[13] These states are: Arunachal Pradesh, Assam, Goa, Gujarat, Jharkhand, Maharashtra, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Pondicherry, Rajasthan, Chhattisgarh, Sikkim, Tripura, Uttar Pradesh, Uttarakhand, Kerala.

[14] The variables include number of household members, presence of pucca house; presence of electricity; possession of a phone, and presence of a toilet, presence of government secondary and middle schools, private school, private health clinic, bank, and pucca road in the village

Panel B and C show negative effects on both reading and math scores in the post-policy period. This matches the main finding in Tables 1 to 3 for the DD specification. Consequently, as the results in the main analysis were insignificant when using DD specification (1), I do not conclude any treatment effect from this robustness exercise.

VI. Conclusion

Young girls in developing countries like India often face significant barriers to formal education due to the high incidence of child marriage, resulting in early school dropouts (Sen & Thamarapani, 2023). This perpetuates an intergenerational cycle of poverty and disadvantage, which can be broken by ensuring proper education of girls well beyond the primary level (Llyod & Young, 2009). On one hand, India has heavily invested in improving the public education infrastructure to combat supply-side constraints to female education. On the other hand, several states like West Bengal have made significant investments in conditional cash transfer programs like the Kanyashree Prakalpa (KP) program to tackle the demand-side constraints. While previous studies on similar CCTs have found an increase in school enrollment among the targeted beneficiaries, the indirect effects on non-beneficiary siblings have remained ambiguous. This paper investigates the direct and indirect causal impact of KP on the education of eligible girls and their ineligible siblings.

Using the DD estimation strategy, I find a positive yet insignificant impact of KP on enrollment of eligible girls. In contrast, while using the DDD estimation strategy, I find a positive and significant effect of KP on the enrollment of eligible girls. This result is in line with existing research as well as the theoretical prediction as discussed in the model of schooling decisions in Ferreira et al., (2009). Moreover, the effect is more pronounced among girls whose mothers have no education and who live closer to schools. While previous studies have not found any impact on children's learning outcomes as they are not the direct focus of the CCTs, I find a positive and significant effect of KP on eligible girls' math outcomes. These results suggest that KP has been effective in bridging the gender gap in both enrollment and learning outcomes in West Bengal. I also conduct a multitude of placebo tests to further validate the robustness of my findings.

In terms of indirect effects of KP, I find mixed results for ineligible younger siblings. While I find a modest positive spillover effect on the school enrollment of younger boy siblings, there is no significant effect on younger girl siblings. The presence of negative estimates for younger girls' enrollment may hint towards the presence of a displacement effect, which theorizes that an eligible child replaces the ineligible child in school to meet the program's conditionalities. However, these estimates for girls are statistically insignificant. Overall, I conclude that while there is some evidence of positive spillover effects on enrollment of younger boys, there are no substantial spillover effects on younger siblings' learning outcomes or private educational investments.

Nevertheless, there are limitations to these findings. While I conducted multiple tests to check for the assumption of parallel trends and controlled for potential violations by including covariates and linear trends, it is still possible that these tests were underpowered to detect any significant deviations. Moreover, potential spillover effects on 13–16-year-old boys in West Bengal pose a threat to the validity of the DDD estimation strategy. However, it is more likely that there are positive spillovers (income effect) than negative spillovers (displacement effect), meaning it is more likely that these ineligible boy siblings are more inclined to enroll or stay enrolled in school rather than drop out because their eligible sister replaces them in school to fulfill the program conditionality. Therefore, given the presence of strong pro-male bias in education, especially in rural India, the positive spillover or the income effect likely prevails, which means my DDD results may represent a lower bound of the true treatment effect (Muralidharan & Prakash, 2017; Azam & Kingdon, 2011).

Additionally, there are a few data limitations. The data has a few missing years between 2008 and 2022 and lacks specific data on program uptake and household income, which are essential to determine program eligibility and assess the long-term impact of KP, especially during COVID-19. However, I conduct supplementary tests with alternative datasets and restrict my sample based on the eligibility criteria of KP to partially mitigate these data limitations.

This paper makes three key contributions to the literature on CCTs. First, there are no previous studies that examine the indirect effects of KP on ineligible siblings and assess treatment effect heterogeneity based on dimensions like maternal education and proximity to schools. I find evidence of positive spillover effects on enrollment of younger boy siblings but no substantial spillover effects on younger siblings' learning outcomes or private educational investments. Moreover, the results from heterogeneity analysis suggest that the program's impact is indeed stronger for eligible girls with uneducated mothers and those who live near formal schools. Second, I find a positive, albeit very small, effect of KP on eligible girls' enrollment, which aligns with other CCT studies, and the theoretical framework discussed by Ferreira et al., (2009). Moreover, unlike previous studies that found no effect of CCTs on learning outcomes, I find a positive effect of KP on math outcomes of eligible girls using the DDD specification. This finding suggests that KP has been successful in bridging the gender gap in both enrollment and learning outcomes in West Bengal. Third, I find that KP reduces pro-male bias in household educational investments. Using the DDD specification and restricting my sample until 2018, I find a positive impact of KP on the likelihood of eligible girls receiving private tuition.

Nevertheless, despite the positive and statistically significant findings when using the DDD estimation strategy, the point estimates are quite small. One possible explanation for the low point estimates, especially for learning outcomes, could be related to the enforcement of the eligibility criterion. While the program requires that girls not only be enrolled but also attend school regularly, this may not be strictly enforced. Given the challenges with administrative oversight, it is possible that some girls are enrolled without meeting the regular attendance requirement. This might explain the limited impact on learning outcomes. Additionally, Dutta and Sen (2020) find that initially KP had a more substantial impact on girls' education and the incidence of child marriage, but it seems to have dampened over time. One possible explanation for this deterioration in the positive impact of the KP could be the low cash transfer amounts, which may not be sufficient to induce proper behavioral changes in the eligible girls and their parents. Another reason could be the lack of proper infrastructure and quality teachers in schools. Future research can focus on how these institutional contexts affect KP's impact on adolescent girls' education and well-being. Moreover, my analysis of the effect of COVID-19 on KP was incomplete due to missing data for key years (2019 to 2021). Therefore, future research can further investigate how shocks like COVID-19 interact with policies like KP. Understanding these dynamics is crucial to ensure that expensive CCT policies like KP continue to support children's education and well-being even in the face of unforeseen challenges.

References

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425. https://doi.org/10.1257/jel.20191450

Abadie, A., & Gardeazabal, J. (2003). The economic costs of Conflict: A Case Study of the basque country. *American Economic Review*, *93*(1), 113–132. https://doi.org/10.1257/000282803321455188

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2022). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, *138*(1), 1–35. https://doi.org/10.1093/qje/qjac038

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Abbring, J. H., & van den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, *71*(5), 1491–1517. https://doi.org/10.1111/1468-0262.00456

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Anukriti, S. (2018). Financial incentives and the fertility-sex ratio trade-off. *American Economic Journal: Applied Economics*, *10*(2), 27–57. https://doi.org/10.1257/app.20150234

Azam, M. (2016). Private tutoring: Evidence from India. *Review of Development Economics*, *20*(4), 739–761. https://doi.org/10.1111/rode.12196

ASER Centre. (2009–2022). *Annual Status of Education Report (Rural)* [Data file]. New Delhi: ASER Centre. Available on request from ASER Centre.

Azam, M., & Kingdon, G. G. (2011). Are girls The fairer sex in India? revisiting intra-household allocation of Education expenditure. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1849461

Baird, S., Chirwa, E., McIntosh, C., & Özler, B. (2010). The short-term impacts of a schooling conditional cash transfer program on the sexual behavior of young women. *Health Economics*, *19*(S1), 55–68. https://doi.org/10.1002/hec.1569

Barrera-Osorio, F., Bertrand, M., Linden, L. L., & Perez-Calle, F. (2011). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia. *American Economic Journal: Applied Economics*, *3*(2), 167–195. https://doi.org/10.1257/app.3.2.167

Bergstrom, K., & Özler, B. (2022). Improving the well-being of adolescent girls in developing countries. *The World Bank Research Observer*, *38*(2), 179–212. https://doi.org/10.1596/40250

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. https://doi.org/10.1162/003355304772839588

Bilinski, A., & Hatfield, L. A. (2020). Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions. *arXiv*. https://arxiv.org/abs/1805.03273

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, *90*(3), 414–427. https://doi.org/10.1162/rest.90.3.414

Camilo, K., & Zuluaga, B. (2022). The effects of conditional cash transfers on schooling and child labor of Nonbeneficiary siblings. *International Journal of Educational Development*, *89*, 102539. https://doi.org/10.1016/j.ijedudev.2021.102539

Das, U., & Sarkhel, P. (2019). Conditional cash transfer for secondary education: Impact assessment of the kanyashree program in West Bengal. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3455215

Desai, S., Vanneman, R., & National Council of Applied Economic Research, New Delhi. (2005). *India Human Development Survey (IHDS), 2005* [Data set]. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR22626.v12

Dey, S., & Ghosal, T. (2021). Can conditional cash transfer defer child marriage? Impact of Kanyashree Prakalpa in West Bengal, India (Warwick Economics Research Papers Series No. 1333) [Unpublished working paper]. University of Warwick, Department of Economics.

Dutta, A., & Sen, A. (2020). *Kanyashree Prakalpa in West Bengal, India: Justification and evaluation* (Final Report No. S-35321-INC-1). International Growth Centre.

Dutta, S. (2014). Parental Education and Family Aspect of School Enrollment in RuralL India. *Journal of Rural Development*, *33*(1), 111–126.

Ferreira, F. H., Filmer, D., & Schady, N. (2009). Own and sibling effects of conditional cash transfer programs: Theory and evidence from Cambodia. *Policy Research Working Papers*. https://doi.org/10.1596/1813-9450-5001

Fiszbein, A., Schady, N., Ferreira, F. H. G., Grosh, M., Keleher, N., Olinto, P., & Skoufias, E. (2009). *Conditional cash transfers: Reducing present and future poverty. World Bank policy research report*. World Bank. https://openknowledge.worldbank.org/entities/publication/db93c3fe-1810-5834-a9da-c1386caa0323

Gaentzsch, A. (2020). Do conditional cash transfers (ccts) raise educational attainment? an impact evaluation of Juntos in Peru. *Development Policy Review*, *38*(6), 747–765. https://doi.org/10.1111/dpr.12468

Glick, P., & Sahn, D. E. (2007). Cognitive skills among children in Senegal: Disentangling the roles of schooling and family background. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.832145

Guariso, A., & Nyqvist, M. (2023). The impact of the COVID-19 pandemic on children's learning and wellbeing: Evidence from India. *Journal of Development Economics*, *164*, 103133. https://doi.org/10.1016/j.jdeveco.2023.103133

Harding, J. F., Morris, P. A., & Hughes, D. (2015). The relationship between maternal education and children's academic outcomes: A theoretical framework. *Journal of Marriage and Family*, *77*(1), 60–76. https://doi.org/10.1111/jomf.12156

Jena, P. K. (2020). Impact of pandemic covid-19 on education in India. *International Journal of Current Research*, *12*(7), 12582–12586. https://doi.org/10.31235/osf.io/2kasu

Kingdon, G. G. (2002). The gender gap in educational attainment in India: How much can be explained? *Journal of Development Studies*, *39*(2), 25–53. https://doi.org/10.1080/00220380412331322741

Lancaster, G., Maitra, P., & Ray, R. (2008). Household expenditure patterns and gender bias: Evidence from selected Indian States. *Oxford Development Studies*, *36*(2), 133–157. https://doi.org/10.1080/13600810802037803

Lavy, V. (1996). School supply constraints and children's educational outcomes in rural Ghana. *Journal of Development Economics*, *51*(2), 291–314. https://doi.org/10.1016/s0304-3878(96)00416-6

Lincove, J. A., & Parker, A. (2015). The influence of conditional cash transfers on eligible children and their siblings. *Education Economics*, *24*(4), 352–373. https://doi.org/10.1080/09645292.2015.1019431

Lloyd, C., & Young, J. (2009). New lessons: The power of educating adolescent girls—a girls count report on adolescent girls. *Population Council*. https://doi.org/10.31899/pgy17.1011

MacKinnon, J. G., & Webb, M. D. (2018). The Wild Bootstrap for few (treated) clusters. *The Econometrics Journal*, *21*(2), 114–135. https://doi.org/10.1111/ectj.12107

Malani, A., & Reif, J. (2015). Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. *Journal of Public Economics*, *124*, 1–17. https://doi.org/10.1016/j.jpubeco.2015.01.001

McIntosh, S., & Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers*, *53*(3), 453–481. https://doi.org/10.1093/oep/53.3.453

Meehan, L., Pacheco, G., & Schober, T. (2023). Basic reading and mathematics skills and the labour market outcomes of young people: Evidence from Pisa and linked administrative data*. *Economic Record*, *99*(327), 473–491. https://doi.org/10.1111/1475-4932.12755

Muralidharan, K., & Prakash, N. (2017). Cycling to school: Increasing secondary school enrollment for girls in India. *American Economic Journal: Applied Economics*, *9*(3), 321–350. https://doi.org/10.1257/app.20160004

Nanda, P., Das, P., & Datta, N. (2022). Education, sexuality, and marriageability: Overlapping tropes in the lives of adolescent girls in Haryana, India. *Journal of Adolescent Health*, *70*(3). https://doi.org/10.1016/j.jadohealth.2021.10.018

Olden, A., & Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*, *25*(3), 531–553. https://doi.org/10.1093/ectj/utac010

Peteros, E. D., Ypil, S. C., Vera, J. V., Alcantara, G. A., Fulgencio, M. D., Plando, D. B., & Jr., L. B. (2022). Effects of school proximity on students' performance in Mathematics. *Open Journal of Social Sciences*, *10*(01), 365–376. https://doi.org/10.4236/jss.2022.101028

Psacharopoulos, G., & Patrinos *, H. A. (2004). Returns to investment in education: A further update. *Education Economics*, *12*(2), 111–134. https://doi.org/10.1080/0964529042000239140

Roodman, D., MacKinnon, J. G., Nielsen, M. O., & Webb, M. (2018). Fast and wild: Bootstrap inference in Stata using boottest. *Queen's Economics Department Working Paper No. 1406*. Queen's University. https://www.stata.com/meeting/canada18/slides/canada18_Webb.pdf

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, *235*(2), 2218–2244. https://doi.org/10.1016/j.jeconom.2023.03.008

Sathar, Z. A., & Lloyd, C. B. (1994). Who gets primary schooling in Pakistan: Inequalities among and within families. *The Pakistan Development Review*, *33*(2), 103–134. https://doi.org/10.30541/v33i2pp.103-134

Sen, G., & Thamarapani, D. (2023). Keeping girls in schools longer: The Kanyashree Approach in India. *Feminist Economics*, *29*(4), 36–64. https://doi.org/10.1080/13545701.2023.2263542

Sipahimalani, V. (1999). Education in the rural Indian household: The impact of household and school characteristics on gender differences. *National Council of Applied Economic Research* (Working Paper No. 68).

van Cappelle, F., Chopra, V., Ackers, J., & Gochyyev, P. (2021). An analysis of the reach and effectiveness of distance learning in India during school closures due to covid-19. *International Journal of Educational Development*, *85*, 102439. https://doi.org/10.1016/j.ijedudev.2021.102439

Appendix

Figure A1: State-wise Map of India

Note: This official map of India and its states is obtained from www.mapsofindia.com. West Bengal being the treated in this analysis, it seems justified to use Jharkhand, Odisha and Chhattisgarh as control states. Both Odisha and Jharkhand share borders with West Bengal. While Chhattisgarh doesn't share a border with West Bengal, all the three states are quite similar to West Bengal in terms of cultural factors. Therefore, this helps in justifying the adoption of use Jharkhand, Odisha and Chhattisgarh as the control states for this analysis.

Figure A2: Trends in Enrollment

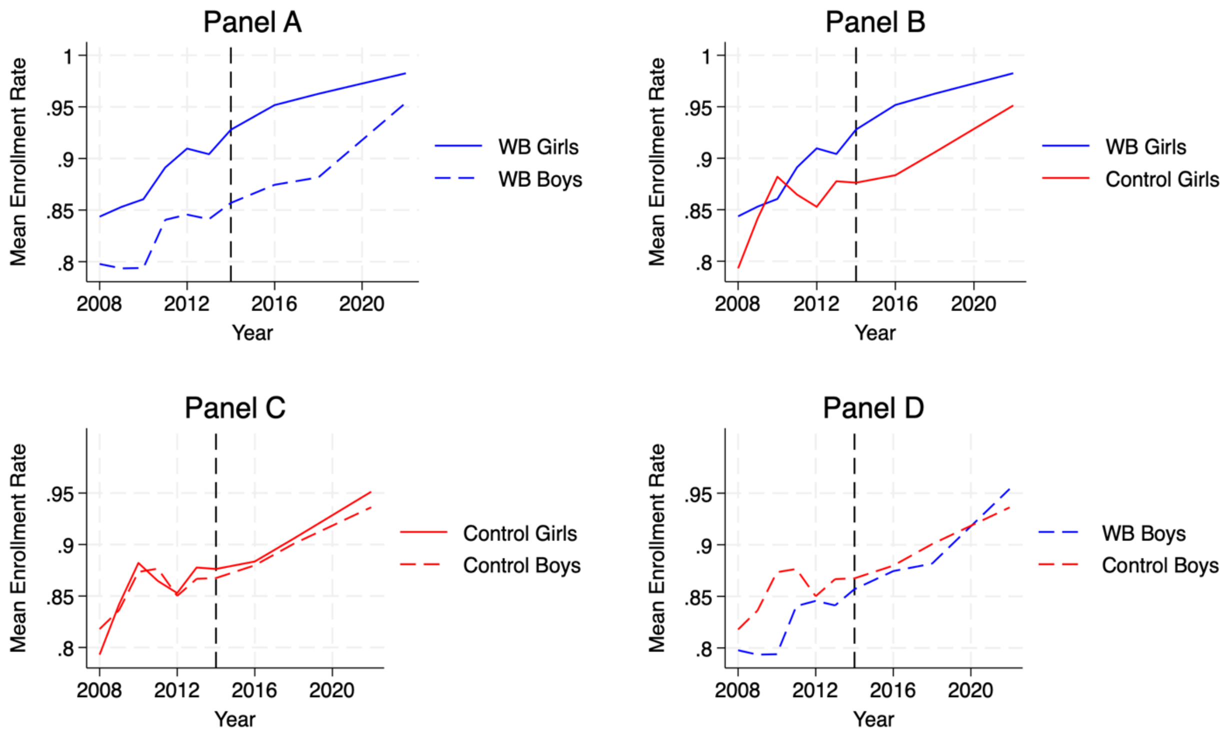

Figure A3: Trends in Reading level

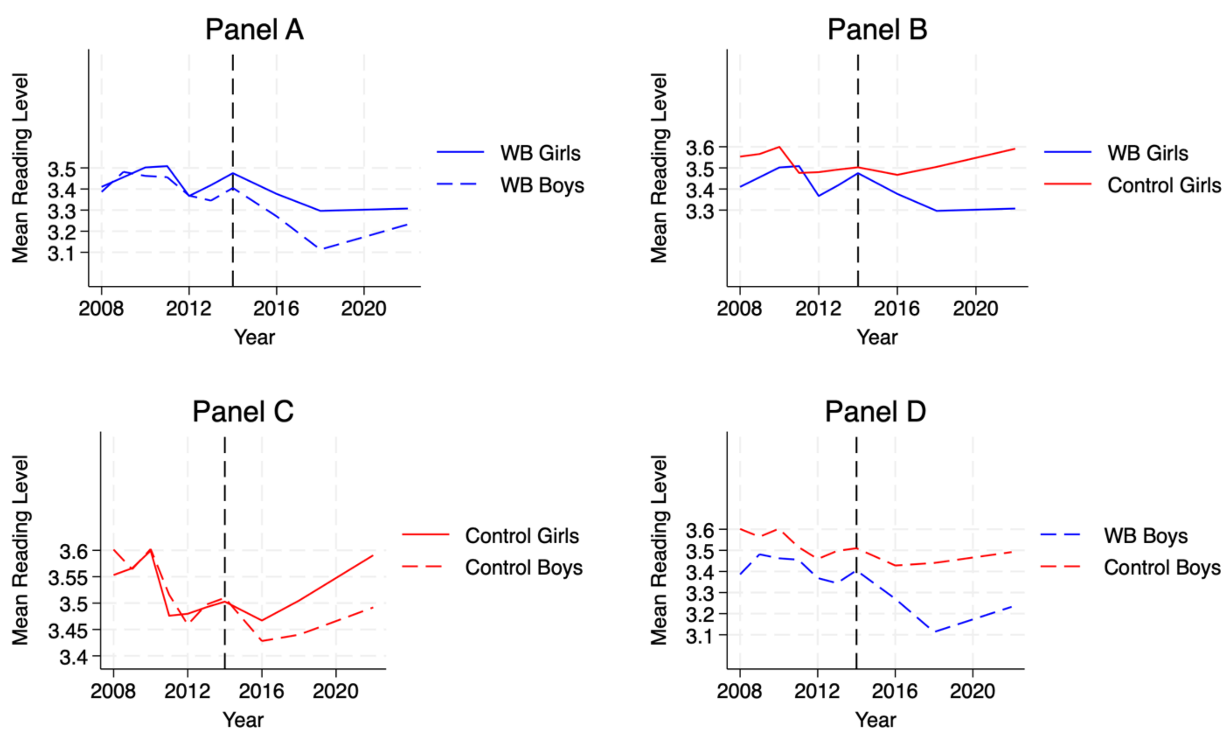

Figure A4: Trends in Math Level

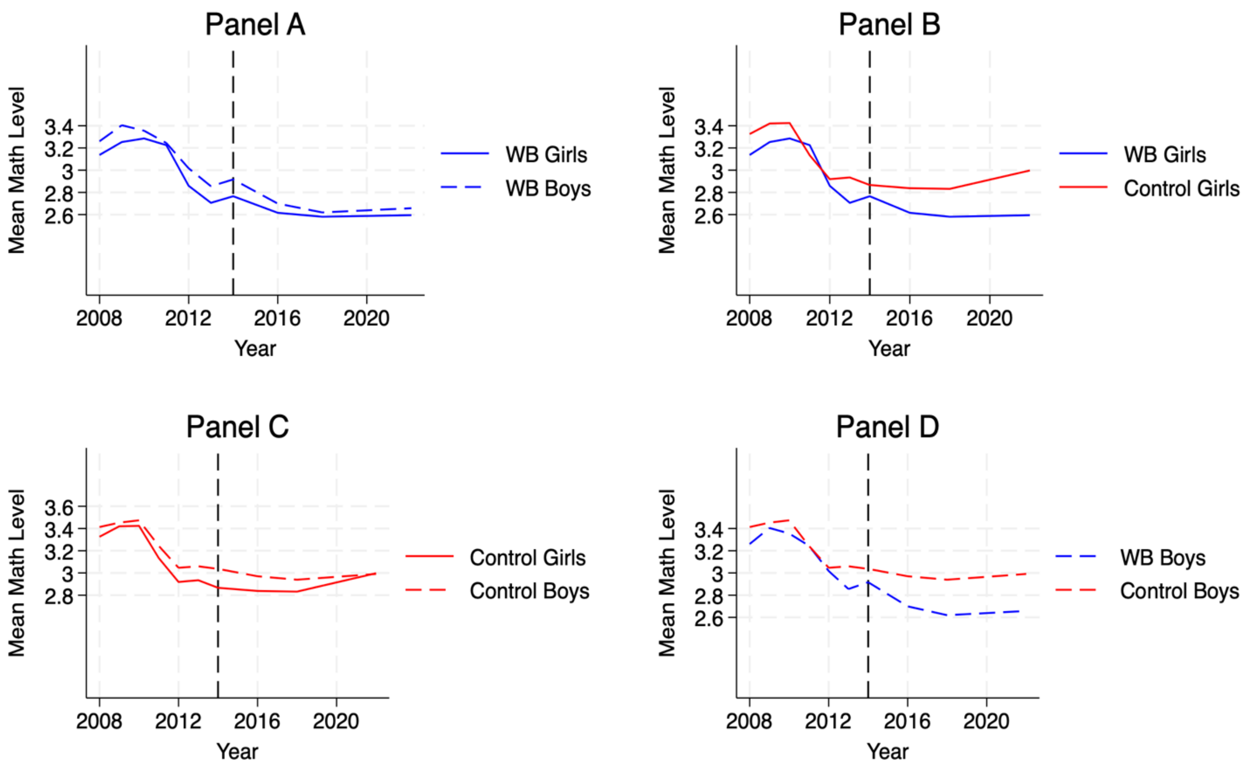

Figure A5: Trends in Private Tuition

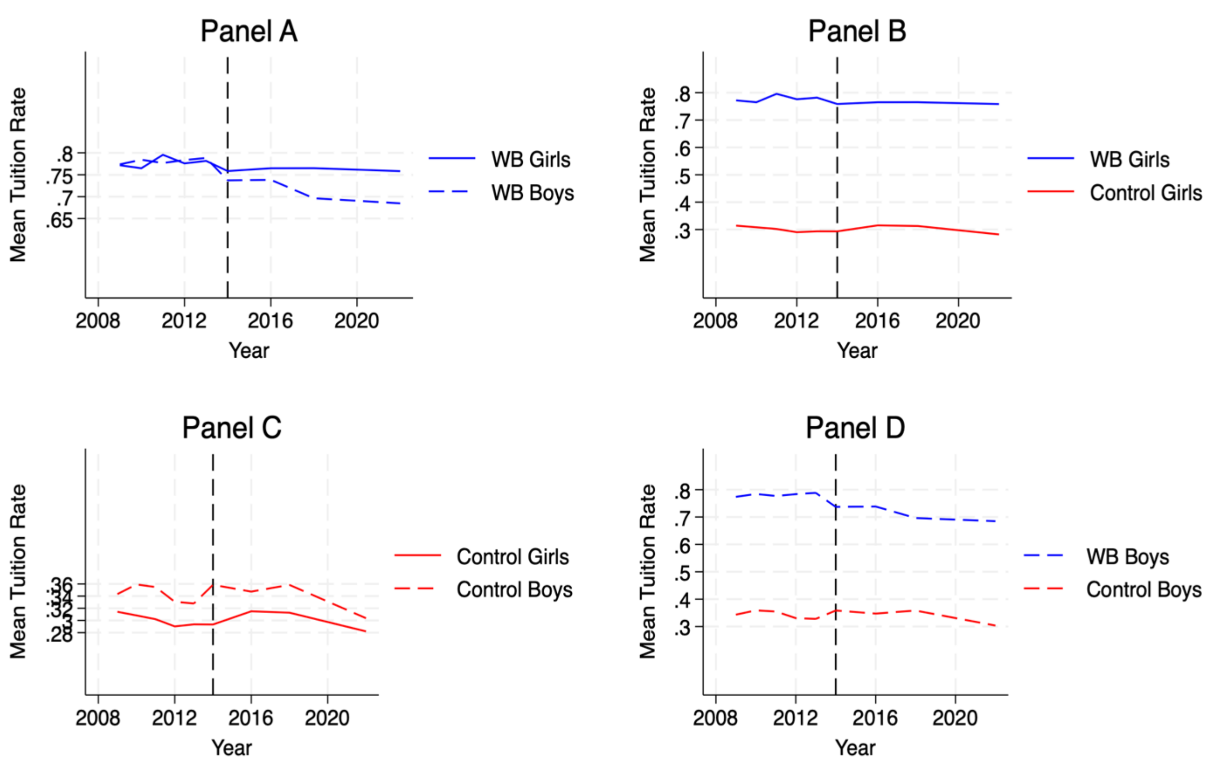

Table A1: Parallel trends for Eligible Girls using DD specification 1

Variables	(1) Enrollment	(2) Reading Level	(3) Math Level	(4) Private Tuition
WB x Year	0.0008	0.0096	0.0094	0.0045
	(0.002)	(0.009)	(0.039)	(0.007)
	[0.837]	[0.656]	[0.907]	[0.758]
Observations	37,849	32,580	32,473	31,464
R-squared	0.101	0.080	0.132	0.395

Notes: This table reports results from DD specification (1) for years 2008 to 2013. Each cell corresponds to a different regression and displays the coefficient on the state-year interaction term with the five years being coded as Year 1 to 5. Robust standard errors clustered at the state level are in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets. The regression is weighted to be representative at the state level and contains year fixed effects. Covariates include year (coded as 1 to 5), West Bengal dummy, child's age, and mother's education.

Table A2: Parallel trends for Eligible Girls using DDD specification 2

Variables	(1) Enrollment	(2) Reading Level	(3) Math Level	(4) Private Tuition
WB x Year x Female	-0.0013	0.0029	0.0090	-0.0011
	(0.001)	(0.005)	(0.005)	(0.002)
	[0.524]	[0.832]	[0.578]	[0.670]
Observations	116,266	100311	100078	79464
R-squared	0.079	0.0450	0.0853	0.239

Notes: This table reports results from DDD specification (2) for years 2008 to 2013. Each cell corresponds to a different regression and displays the coefficient on the state-year-gender interaction term with the five years being coded as Year 1 to 5. Robust standard errors clustered at the state level are in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets. The regression is weighted to be representative at the state level and contains year fixed effects. Covariates include year (coded as 1 to 5), West Bengal dummy, female dummy, child's age, and mother's education.

Table A3: Year-Wise Parallel trends for Eligible Girls using DD specification

Variables	(1) Enrollment	(2) Reading Level	(3) Math Level	(4) Private Tuition
WB x 2008	0.0037	0.1087	-0.0985	0.0637
WB x 2009	-0.0456	0.1020	-0.1304	0.0846
WB x 2010	-0.0557	0.1733	-0.0148	0.0933
WB x 2011	-0.0165	0.2356	0.1481	0.1377
WB x 2012	-0.0157	0.0934	0.0234	0.1094
WB x 2013	-0.0232	0.1618	-0.1527	0.0991
F-test p-value	[0.512]	[0.952]	[0.809]	[0.840]
Observations	96,003	81,355	81,195	76,024
R-squared	0.084	0.055	0.116	0.245

Notes: This table reports results from DD specification (1) where 'Year' has been replaced with individual year dummies. Each cell corresponds to a different regression and displays the coefficient on the state-year interaction. The p-values from the F-test of joint significance using the wild bootstrap sub-clustered at the district level are in brackets. The regression is weighted to be representative at the state level. Covariates include individual year dummies, West Bengal dummy, child's age, and mother's education.

Table A4: Year-Wise Parallel trends for Eligible Girls using DDD specification

Variables	(1) Enrollment	(2) Reading Level	(3) Math Level	(4) Private Tuition
WB x 2008 x Female	0.0753	0.0864	-0.0148	0.0390
WB x 2009 x Female	0.0615	0.0065	-0.1141	0.0315
WB x 2010 x Female	0.0591	0.0906	0.0193	0.0395
WB x 2011 x Female	0.0676	0.1066	0.1008	0.0871
WB x 2012 x Female	0.0624	-0.0137	-0.0334	0.0257
WB x 2013 x Female	0.0637	0.1241	-0.0072	0.0338
F-test p-value	[0.712]	[0.726]	[0.944]	[0.598]
Observations	195,616	161,701	161,416	152,373
R-squared	0.086	0.054	0.111	0.220

Notes: This table reports results from DDD specification (2) where 'Year' has been replaced with individual year dummies. Each cell corresponds to a different regression and displays the coefficient on the state-year-gender interaction. The p-values from the F-test of joint significance using the wild bootstrap sub-clustered at the district level are in brackets. The regression is weighted to be representative at the state level. Covariates include individual year dummies, West Bengal dummy, gender dummy, child's age, and mother's education.

Table A5: Covariate Balance Test

Variables	West Bengal Mean	SD	N	Control States Mean	SD	N	Difference
Enrollment	0.85	0.36	20,777	0.85	0.35	101,194	0
Reading Level	3.43	1.06	17,140	3.54	0.98	88,109	-0.11
Math level	3.14	1.07	17,050	3.25	1.05	87,938	-0.11
Private Tuition	0.78	0.41	13,991	0.32	0.47	68,440	0.46
Child-level characteristics							
Female	0.50	0.01	20777	0.47	0.03	101194	0.03
Age	14.43	0.05	20777	14.39	0.04	101194	0.03
Mother went to school	0.50	0.03	20090	0.42	0.02	96176	0.08
Household-level characteristics							
Availability of electricity	0.62	0.14	20666	0.65	0.09	100427	-0.03
Availability of Cemented House	0.19	0.02	20556	0.17	0.01	99890	0.02
Availability of toilet	0.45	0.01	16747	0.16	0.02	82886	0.28
Possession of Phone	0.57	0.11	16701	0.47	0.09	82432	0.09
Household Size	5.80	0.19	20666	6.36	0.12	100098	-0.56
Village-level characteristics							
Availability of Private School	0.29	0.03	16463	0.25	0.02	79148	0.04
Availability of Secondary School	0.25	0.06	16533	0.21	0.07	91740	0.04
Availability of Middle School	0.31	0.12	19966	0.65	0.04	95282	-0.35
Availability of Private Clinic	0.27	0.09	14378	0.26	0.08	82633	0.01
Availability of Bank	0.23	0.02	20373	0.15	0.02	98127	0.07
Availability of Pucca road	0.51	0.08	20472	0.70	0.03	98528	-0.19

Notes: This table reports the summary statistics of 13–16-year-olds in West Bengal and control states in the pre-period (2008-2013). 'N' denotes the number of observations and 'SD' represents the corresponding standard deviation of that group.

Table A6: Year-Wise Parallel trends for Ineligible Younger Girl Siblings (Spillover Analysis)

	(1)	(2)	(3)	(4)
Variables	Enrollment	Reading Level	Math Level	Private Tuition
WB x Year	0.0115	-0.0489	-0.0778	0.0334
	(0.002)	(0.008)	(0.013)	(0.009)
	[0.421	[0.698]	[0.719]	[0.645]
Observations	18,079	16,117	16,047	15,652
R-squared	0.033	0.158	0.170	0.361

Notes: This table reports results from spillover analysis to ineligible girl siblings using DD specification (3) for years 2008 to 2013. Each cell corresponds to a different regression and displays the coefficient on the state-year interaction term with the five years being coded as Year 1 to 5. Robust standard errors clustered at the state level are in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Table A7: Year-Wise Parallel trends for Ineligible Younger Boy Siblings (Spillover Analysis)

	(1)	(2)	(3)	(4)
Variables	Enrollment	Reading Level	Math Level	Private Tuition
WB x Year	-0.0355	0.0029	-0.0108	-0.0006
	(0.001)	(0.008)	(0.009)	(0.008)
	[0.415]	[0.991]	[0.939]	[0.996]
Observations	20,365	18,049	17,988	17,741
R-squared	0.028	0.148	0.172	0.336

Notes: This table reports results from spillover analysis to ineligible boy siblings using DD specification (2) for years 2008 to 2013. Each cell corresponds to a different regression and displays the coefficient on the state-year interaction term with the five years being coded as Year 1 to 5. Robust standard errors clustered at the state level are in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Table A8: Difference-in-differences (DD) Estimate of the Impact of KP program on 13–16-year-old girls compared to 13–16-year-old boys in West Bengal

	(1)	(2)	(3)	(4)
Variables	Enrollment	Reading Level	Math Level	Private Tuition
WB x Post	0.0093	0.0946**	0.0565*	0.0556**
	(0.380)	(0.010)	(0.073)	(0.001)
	[0.409]	[0.022]	[0.086]	[0.001]
Observations	24,102	18,990	18,917	21,405
R-squared	0.123	0.127	0.189	0.120

Notes: This table reports results from DD specification (1) where Group A (13–16-year-old girls) and Group B (13–16-year-old boys) in the treatment state (West Bengal) are compared before and after KP implementation in 2013. Each cell reports the coefficients on key outcome variables. Robust standard errors clustered at the district level are in parentheses and p-values from the wild bootstrap procedure are in brackets. The regressions contain year fixed effects. Linear trends include district-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Table A9: Difference-in-differences (DD) Estimate of the Impact of KP program on 13–16-year-old girls compared to 13–16-year-old boys in Control states

	(1)	(2)	(3)	(4)
Variables	Enrollment	Reading Level	Math Level	Private Tuition
WB x Post	0.006*	-0.0144**	-0.0145	0.0043
	(0.084)	(0.013)	(0.013)	(0.471)
	[0.081]	[0.002]	[0.301]	[0.497]
Observations	128,917	105,044	105,044	111,991
R-squared	0.118	0.139	0.138	0.152

Notes: This table reports results from DD specification (1) where Group A (13–16-year-old girls) and Group B (13–16-year-old boys) in the control states are compared before and after 2013. Each cell reports the coefficients on key outcome variables. Robust standard errors clustered at the district level are in parentheses and p-values from the wild bootstrap procedure are in brackets. The regressions contain year fixed effects. Linear trends include district-specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Robustness Checks

Table A10- Changing the Intervention Year for Difference-in-differences (DD) Estimates of the Impact of KP program on Eligible Girls

	(1)	(2)	(3)	(4)
Dependent Variables	Enrollment	Reading Level	Math Level	Private Tuition
Panel A: 2010 as Placebo Year				
WB x Post	0.0354	0.0546	0.2338	0.0020
	(0.003)	(0.004)	(0.008)	(0.003)
	[0.251]	[0.206]	[0.115]	[0.925]
Observations	37,849	32,580	32,473	31464
R-squared	0.101	0.080	0.132	0.395
Panel B: 2011 as Placebo Year				
WB x Post	0.0322	0.0360	0.1351	-0.0398
	(0.002)	(0.015)	(0.010)	(0.008)
	[0.111]	[0.759]	[0.358]	[0.621]
Observations	37,849	32,580	32,473	31,464
R-squared	0.101	0.080	0.132	0.395

Notes: This table reports placebo test results from DD Specification (1) where the program intervention year has been changed to from 2013 to 2010 (Panel A) and 2011 (Panel B). Each cell reports the coefficient on the West Bengal-Post interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A11- Changing the Intervention Year for Triple difference (DDD) Estimates of the Impact of KP program on Eligible Girls

	(1)	(2)	(3)	(4)
Dependent Variables	Enrollment	Reading Level	Math Level	Private Tuition
Panel A: 2010 as Placebo Year				
WB x Post x Female	0.0005	0.1481	0.2122	0.0204
	(0.005)	(0.049)	(0.036)	(0.020)
	[0.968]	[0.283]	[0.114]	[0.636]
Observations	78,989	66,761	66583	65,048
R-squared	0.098	0.072	0.123	0.378
Panel A: 2010 as Placebo Year				
WB x Post x Female	0.0055	0.0844	0.1367	0.0183
	(0.006)	(0.048)	(0.039)	(0.014)
	[0.769]	[0.413]	[0.178]	[0.546]
Observations	78,989	66,761	66,583	65048
R-squared	0.098	0.072	0.123	0.378

Notes: This table reports placebo test results from DDD Specification (2) where the program intervention year has been changed to from 2013 to 2010 (Panel A) and 2011 (Panel B). Each cell reports the coefficient on the West Bengal-Post-Gender interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A12- Difference-in-differences (DD) Placebo Outcome Test

	(1)	(2)
Dependent Variables	Mother's Highest Edu Level	Father's Highest Edu Level
WB x Post	0.2924	0.0699
	(0.023)	(0.030)
	[0.330]	[0.781]
Observations	35,915	48661
R-squared	0.102	0.124

Notes: This table reports placebo test results from DD Specification (1) where the placebo outcomes are highest education level attained by the child's mother and father. Each cell reports the coefficient on the West Bengal-Post interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A13- Triple difference (DDD) Placebo Outcome Test

	(1)	(2)
Dependent Variables	Mother's Highest Edu Level	Father's Highest Edu Level
WB x Post x Female	0.1024	-0.1857
	(0.026)	(0.026)
	[0.253]	[0.159]
Observations	72,358	91,488
R-squared	0.126	0.134

Notes: This table reports placebo test results from DDD Specification (2) where the placebo outcomes are highest education level attained by the child's mother and father. Each cell reports the coefficient on the West Bengal-Post- Gender interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A14- Difference-in-differences (DD) Estimates of the Impact of KP program on Eligible Girls with Higher Order Trends

	(1)	(2)	(3)	(4)
Dependent Variables	Enrollment	Reading Level	Math Level	Private Tuition
Panel A: Quadratic trends				
WB x Post	0.0918	-0.0166	0.1303	-0.0056
	(0.005)	(0.022)	(0.030)	(0.011)
	[0.172]	[0.940]	[0.687]	[0.938]
Observations	75,998	63,628	63,501	66,929
Panel B: Cubic Trends				
WB x Post	0.0918	-0.0166	0.1303	-0.0056
	(0.005)	(0.022)	(0.030)	(0.011)
	[0.172]	[0.940]	[0.686]	[0.938]
Observations	75,998	63,628	63,501	66,929

Notes: This table reports placebo test results from DD Specification (1), where higher-order (quadratic and cubic) trends replace the linear trends to check the robustness of the results against changes in the underlying specification. Each cell reports the coefficient on the West Bengal-Post interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A15- Triple difference (DDD) Estimates of the Impact of KP program on Eligible Girls with Higher Order Trends

	(1)	(2)	(3)	(4)
Dependent Variables	Enrollment	Reading Level	Math Level	Private Tuition
Panel A: Quadratic trends				
WB x Post x Female	0.0053*	0.0399	0.0641*	0.0518
	(0.001)	(0.006)	(0.008)	(0.006)
	[0.051]	[0.161]	[0.079]	[0.105]
Observations	153,019	124,188	123,961	133,396
Panel B: Cubic Trends				
WB x Post x Female	0.0053*	0.0399	0.0641*	0.0518
	(0.001)	(0.006)	(0.008)	(0.006)
	[0.051]	[0.161]	[0.079]	[0.105]
Observations	153,019	124,188	123,961	133,396

Notes: This table reports placebo test results from DDD Specification (2), where higher-order (quadratic and cubic) trends replace the linear trends to check the robustness of the results against changes in the underlying specification. Each cell reports the coefficient on the West Bengal-Post- Gender interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A16- Difference-in-Differences (DD) Estimates of the Impact of the KP Program on Missing Learning Outcomes for Eligible Girls

	(1)	(2)
Dependent Variables	Missing Reading Level	Missing Math Level
WB x Post	0.0084	0.0046
	(0.006)	(0.006)
	[0.912]	[0.951]
Observations	76,224	76,224
R-squared	0.033	0.033

Notes: This table reports placebo test results from DD Specification (1), where the outcome variables are missing dummy variables for reading and math learning outcomes. Each cell reports the coefficient on the West Bengal-Post interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Table A17- Triple difference (DDD) Estimates of the Impact of the KP Program on Missing Learning Outcomes for Eligible Girls

	(1)	(2)
Dependent Variables	Missing Reading Level	Missing Math Level
WB x Post x Female	-0.0042	0.0009
	(0.006)	(0.006)
	[0.842]	[0.967]
Observations	153,436	153,436
R-squared	0.0535	0.0534

Notes: This table reports placebo test results from DDD Specification (2), where the outcome variables are missing dummy variables for reading and math learning outcomes. Each cell reports the coefficient on the West Bengal-Post- Gender interaction term. Each column represents a specific outcome variable. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1.

Figure A6 – Synthetic Control Method Trends for Enrollment (West Bengal vs Synthetic West Bengal)

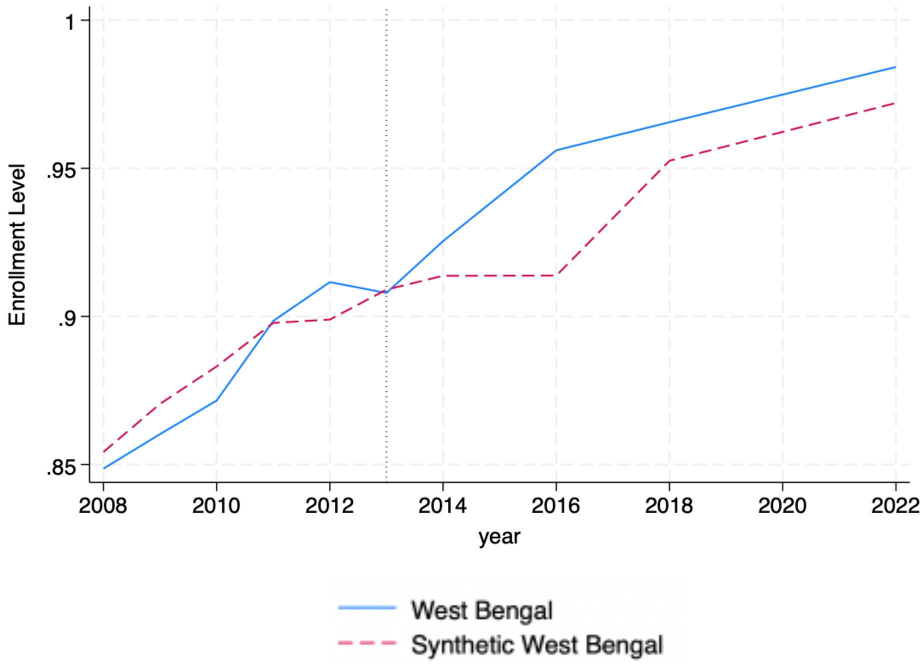

Figure A7– Synthetic Control Method Trends for Reading Level (West Bengal vs Synthetic West Bengal)

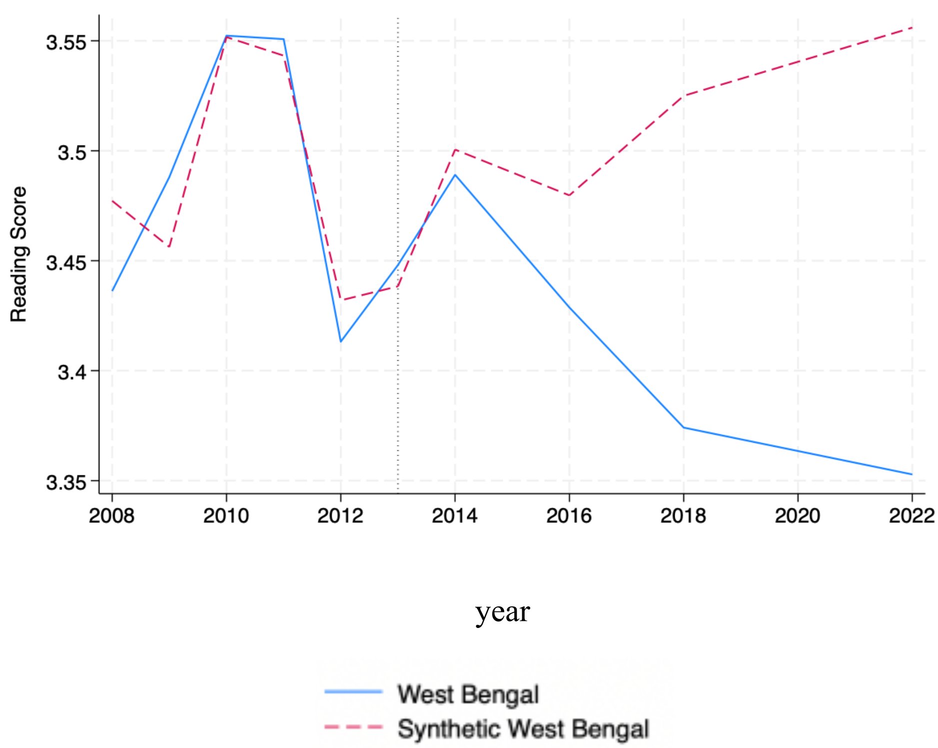

Figure A8 – Synthetic Control Method Trends for Math Score (West Bengal vs Synthetic West Bengal)

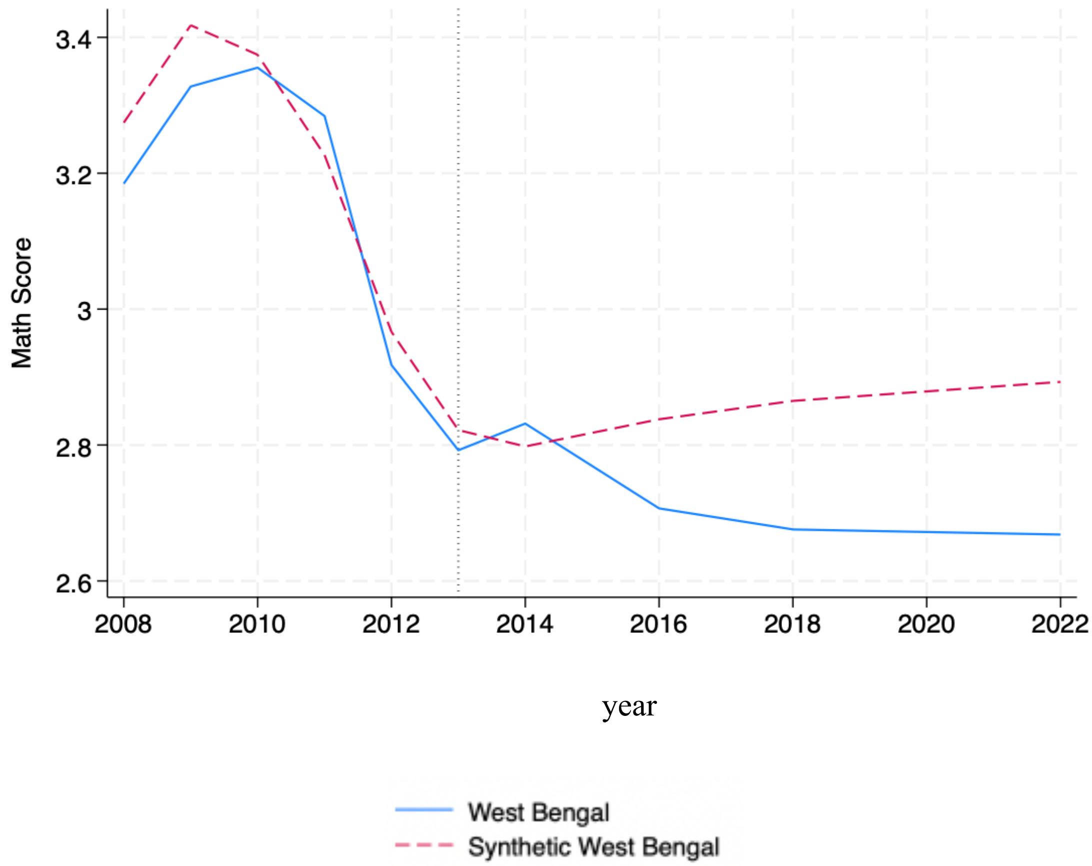

Table A18- Difference-in-Differences (DD) Estimates of the Impact of KP program on Eligible Girls Without 2022

Dependent Variables	(1) Enrollment	(2) Reading Level	(3) Math Level	(4) Private Tuition
WB x Post	0.1172	-0.0162	0.1081	0.0341
	(0.005)	(0.025)	(0.028)	(0.008)
	[0.193]	[0.940]	[0.617]	[0.458]
Observations	63,588	52911	52,787	54,795
R-squared	0.102	0.085	0.149	0.378

Notes: This table reports results from DD specification (1) when excluding 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.

Table A19- Triple difference (DDD) Estimates of the Impact of KP program on Eligible Girls Without 2022

Dependent Variables	(1) Enrollment	(2) Reading Level	(3) Math Level	(4) Private Tuition
WB x Post x Female	0.0177**	0.0590	0.0859**	0.0474**
	(0.000)	(0.014)	(0.010)	(0.004)
	[0.0019]	[0.341]	[0.075]	[0.036]
Observations	129,616	104,760	104,541	110,588
R-squared	0.102	0.079	0.144	0.359

Notes: This table reports results from DDD specification (2) when excluding 2022 and each cell reports the coefficient on key variables. Robust standard errors clustered at the state level are reported in parentheses and p-values from the wild bootstrap procedure sub-clustered at the district level are in brackets, where ***p<0.01, **p<0.05 and *p<0.1. The regressions are weighted to be representative at the state level and contain year fixed effects. Linear trends include state- specific time trends. Socio-economic controls consist of child age, mother's schooling status, household size, whether the house is cemented or not, whether the household has electricity, a TV and a toilet, whether the village has electricity, a cemented road, a bank, schools, health clinics.