Does a Legal Ban on Pre-Natal Sex Determination Improve Female Educational Attainment? Evidence from India

Aditi Sanjay* (University of Warwick)
Subhashish Dey** (University of Warwick)

(Version: September 2019)

Abstract

The study attempts to empirically estimate the causal impact of a legal ban on pre-natal sex determination on female educational attainment in the Indian context. A differences-in-differences setup is employed to exploit the exogenous policy variation created by the implementation of the Pre-Conception and Pre-Natal Diagnostic Techniques (PCPNDT) Act, 1996. The paper establishes the economic channels linking the direct effect of the Act on the odds of occurrence of a female birth to its indirect effect on female educational attainment. The results of the analysis report that the Act has significantly increased the probability of a female birth by 2.56% and years of education completed by 2.15 years. Hence, this study contributes to the scarce literature on the long-term consequential impact of this legislation and to the relevance of demographic policies in shaping human capital formation.

Key words: PCPNDT, pre-natal sex determination, educational attainment, differences-in-differences

*: aditi.sanjay@warwick.ac.uk (Corresponding Author)

**: subhasish.dey@warwick.ac.uk

1. **Introduction**

1.1 Background

Deep-rooted son preferences across several Asian countries has given rise to persistent discrimination against the girl child. While several studies have concluded that female infants are biologically advantaged in the form of higher probability of survival, the sex ratio[1] still remains favourable to males. (Arnold et al 1998, Echávarri and Ezcurra 2006). The desire to continue the family name, the dowry system and the tradition of sons being the source of financial support to parents after retirement have been responsible for widespread sex-selective abortions and female infanticides. (Westley and Choe,2007). This becomes a cause for concern as having an excess of males has serious and far-reaching consequences. Researchers have shed light on the possible exclusion of women from the labour force as well as the growing prevalence of "male towns" where single men exert dominance across most fields and disciplines (Boserup,1970). Some of the other repercussions include a fall in birth rates, imbalance in the marriage market and brutal crimes committed against women. (Caldwell,2001) It can thus be established that the sex ratio[1] has long-term implications for the future of an economy.

Statistics compiled by the World Health Organisation on Asian countries reveal that over the two decades spanning between 1990 and 2010, Bhutan and North Korea registered an increase in the ratio of males to females by 8% and 1% respectively. On the other hand, Sri Lanka, Nepal, Maldives and Bangladesh emerged to have shown the most improvement in this aspect. By 2011, India (940 females per 1000 males) and China (926 females per 1000 males) were the only SEAR (South-East Asia Region) nations left facing the dilemma of a largely skewed sex ratio despite high economic growth. (United Nations World Population Prospects, 2011).

Since the 1901 Census of India, all states exhibited a falling sex ratio with the exception of Kerala. The legalization of abortion under the Medical Termination of Pregnancy Act (1972) resulted in the widespread practice of sex-selective abortions. This marked the onset of the continuing decline in female births exhibiting a steep fall from 0.941 in 1961, to 0.930 in 1971. The 1980s witnessed rampant use of ultrasound technologies as access become more widespread and affordable. This brought about unusually low sex ratios, with the lowest recorded figure of 0.927 in 1991. Historically, Southern and North Eastern states characterized by higher levels of education and a matrilineal society have the most gender balance. Kerala's sex ratio has been consistently high at 1.084. North Indian states with innate son preferences figure at the lower end of the spectrum among which Haryana and Punjab have been assigned the "gender critical" status. (Census of India,2011). This declining sex ratio can have a long run impact on the educational attainment of the girl child, with possible consequences on the labour market. Hence, a low sex ratio has an economic consequence via the effect of educational attainment on labour force participation. As mentioned before, son preference could result in the girl child becoming more susceptible to pre-natal and post-natal exploitation. However, a ban on pre-natal sex determination could potentially shut down avenues of pre-natal discrimination.

To reverse these worrying trends, governments have pro-actively launched demographic policies that imposed legal bans on pre-natal sex determination. One such legislation is the Pre-Conception and Pre-Natal Diagnostic Techniques (PCPNDT) Act that was introduced in India in 1994 and came into effect on 1st January, 1996. It extends to all states of India except the state of Jammu and Kashmir (henceforth J&K). This is due to the special status that was granted to J&K under Article 370 and 35A of the Indian Constitution which outlines the provisions under which the state possesses autonomy in internal administration and in framing separate legislations from the rest of the country.

[1] Sex ratio is defined as the number of females per 1000 males. This is the definition of sex ratio adopted by the Census of India and will be used throughout this paper unless otherwise specified.

The Act banned the practice of pre-natal sex determination except for cases related to genetic abnormalities and sex-linked disorders. All medical centres are strictly prohibited from utilising diagnostic techniques ranging from ultrasound to In-Vitro Fertilisation (IVF) and revealing the gender of an unborn child. As the primary aim of the law was to reduce pre-natal discrimination towards female foetuses, this motivates us to discover whether this intervention has been able to produce meaningful outcomes for females in the long run.

Social scientists have attempted to study the social and economic implications of gender imbalance on fertility, mortality and demographic composition. The success of the PCPNDT Act has not been unequivocally unidimensional. Some studies estimating the impact of the Act on the child sex ratio, infant mortality and other health indicators including malnutrition, weight and height have found it to be significantly positive (Nandi and Deolalikar 2013). Conversely, others have yielded results proving that the ban was ineffective due to the dwindling sex ratio in the 2001 Census. (Arnold et al, 2004; Jha et al 2006; Visaria, 2008).

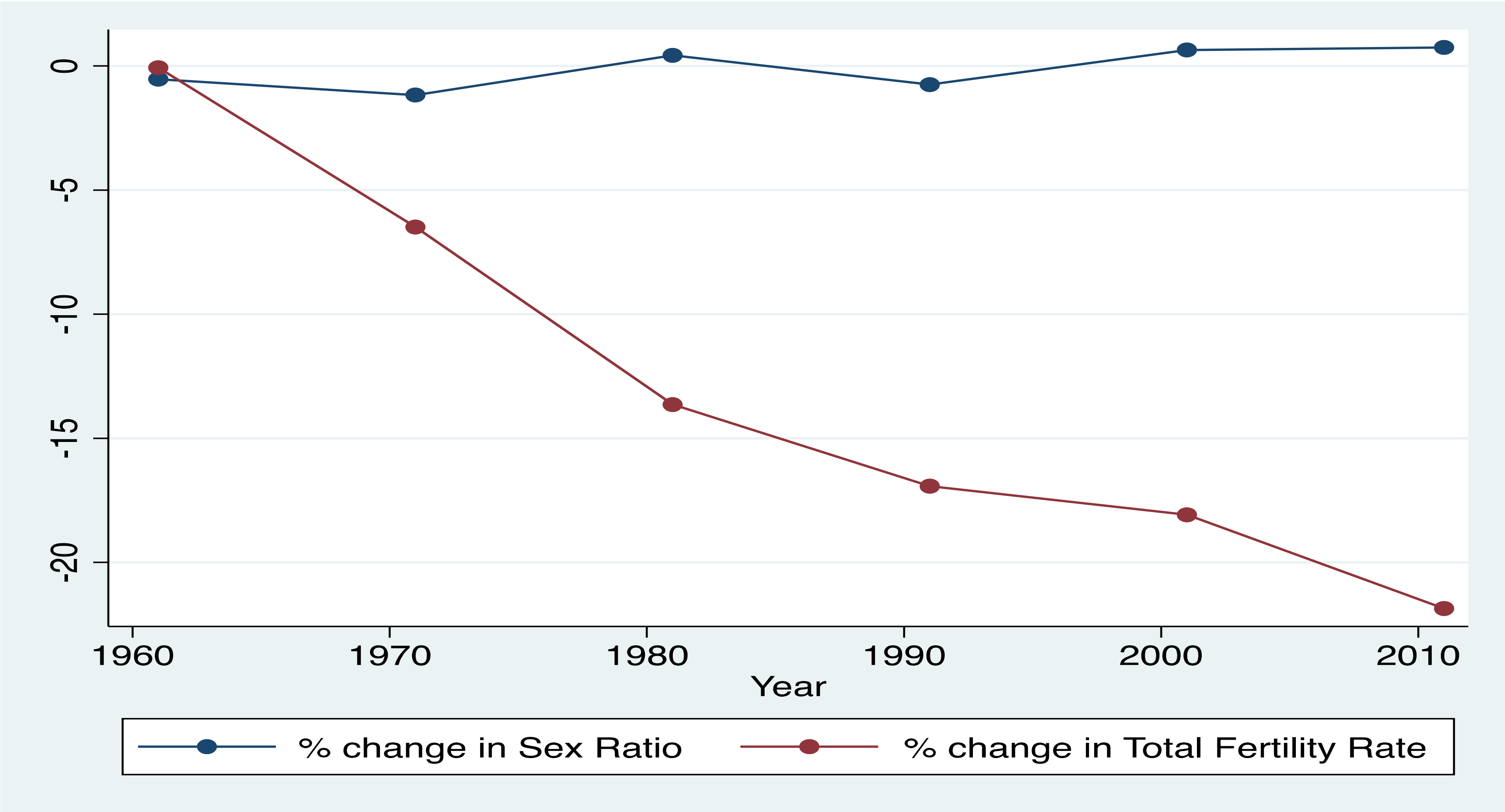
The state of work around the effect of the PCPNDT Act has been so far on the determination of its short term to medium term effects. As the Act has been in operation from 1996, it is a good time now to see its long-term effect on human capital formation for females in general. This prompts us to ask the question of whether a link can be established between the short-term and long-term consequences of the Act. From an economic point of view, an interesting outcome to analyse would be that of educational attainment which in turn has a direct effect on labour market outcomes. Accepting the fact that the law has a significant positive impact on the sex ratio (Nandi and Deolalikar, 2013), this study determines whether this act has any bearing on girls' educational outcomes in the implementing states in comparison to the control state. In doing so, this study examines through various channels, the relevance of such demographic policies in shaping human capital formation.

Although micro level data is utilized, this analysis sheds light on the impact of the law at the macro level. Most impact studies of the PCPNDT Act have looked at the intended effects on females relative to males (Nandi and Deolalikar, 2013). However, this study concentrates on determining its causal effect on girls' education for treated states (all Indian states excluding the state of J&K) in comparison to the control state (J&K). In order to make the causal effect, if any, more convincing, the treated sample is narrowed down to include bordering districts in J&K and Himachal Pradesh.

While the immediate result of the Act was the change in the number of female births, it could have triggered a paradigm shift in household behavior. Firstly, it could be argued that the skewed sex ratios would result in a scarcity of brides, early marriages and higher chances of females withdrawing from education. (Tuljapurkar et al. 1995). However, the introduction of the Act could restore gender balance in the marriage market and improve female educational attainment. Secondly, the Act could potentially alter family size. An unintended consequence in the form of the "son-stopping rule"[2] might yet again distort the natural sex ratio of 1.05 boys per girl. This means that these "unwanted" girls would be more likely to be born into larger families, receive a much smaller share of total resources and would experience little improvement in educational attainment in comparison to their male siblings. (De Souza and Marteleto,2013). However, a ban on pre-natal sex determination will prevent parents from performing sex-selective abortions and would therefore increase the likelihood of a female birth. (Nandi,2015). Also, financial constraints would limit the number of off-spring born to a couple, enabling the girl child to receive a greater share of total resources leading to an improvement in her education. It is these arguments that will formulate the primary hypotheses of this study.

[2] The sex of a new born child is dependent on the sex composition of older siblings. Examining the effect of family composition on parity progression, Arnold et al (1998) show that in northern states with high son preference , women with two sons were least likely to have another child as the demand for sons has been met. (Punjab, Haryana , Uttar Pradesh and Rajasthan). The results of their analysis also reveal that at third and higher parities, women with no sons was the group with the highest probability of wanting to have more children in the hopes of conceiving a son. This provides strong evidence of son preference influencing parents' decision to adopt the son-stopping rule.

Figure 1 : Trends in Sex Ratio and Total Fertility Rate in India



Notes: Data has been sourced from the Census of India. The percentage change in sex ratio and percentage change in total fertility rates have been plotted against the respective census years spanning between 1961 and 2011.

1.2 Research Question and Objective

This study uses education in years for females in the age group 15-19 years (girls born before the Act were born between 1973 and 1978 and those born after the Act were born between 1996 and 2001) as the proxy for educational attainment. The main objective is to determine whether the Act has had a causal effect on improving female educational attainment for the rest of India. This dissertation will discuss and test the validity of the arguments mentioned above using a differences-in-differences framework. It will utilise survey data from National Family Health Survey (NFHS) waves one, two and four to determine the effects for a longer time period. Another unique feature of this study is extending the earlier focus on Maharashtra and neighbouring states to the whole of India, by examining the policy from the perspective of J&K and the remaining states. To the best of our knowledge, this will be the first analysis of the PCPNDT law's impact on educational outcomes. Furthermore, it will attempt to contribute to existing literature by discussing in breadth the channels through which any long -term impacts should occur over time. Finally, it will provide insight into the status of girls at the macro level using micro-level data.

Section 2 will comprise of a discussion of related topics covered in existing literature. This will be followed by a description of the data source and variables utilised in the study in Section 3. Section 4 outlines the methodology adopted while Section 5 will cover the results, mechanisms, robustness checks, validity checks and limitations. Section 6 will summarise the analysis along with the findings, and conclude with comments on scope for future extensions.

2. Literature Review

Economists have long been interested in examining the role of parental investments in children in order to determine their long-run returns. There exists a wealth of literature on the effects of demographic policies on education, employment and health, with the gender gap often being revisited. This section will explore the main topic of research and related sub-topics thematically from various angles and contexts.

2.1 Impact on the Child Sex Ratio over time

Sex selective abortions and female infanticides have long been responsible for the phenomenon of "Missing Women" (Sen,1990). One of the immediate results of such attitudes and practices is the highly skewed sex ratios in these societies. Nandi and Deolalikar (2013) proved that the PCPNDT Act indeed did have a significant positive impact on child sex ratios. Using data from the 1991 and 2001 Censuses, the empirical strategy exploited the regional variation in the timing of the law's implementation. The regions of study included districts in the state of Maharashtra (pre-treated sample) and districts in the neighbouring states (newly treated sample). Village and town fixed effects regressions also yielded similar results which concluded that the absence of the law would have reduced the number of female births by 106,000.

2.2 Gender Differentials in Health Outcomes

Additional channels of post-natal gender discrimination were further examined in Nandi (2015) to estimate the secondary effects of the PCPNDT Act. The main hypothesis tested was that parents' unchanged son preference would cause a decline in chances of survival of females. A differences-in-differences model with state and year fixed effects estimated that while the law had been successful in increasing the probability of a female birth, it had no significant impact on reducing female infant mortality. Hu and Schlosser (2015) explore this topic from the perspective of family size and intra-family transfer of resources. The argument made by the authors is that girls are more likely to be born into households with weaker son preferences and would experience increased post-natal investments and improved health outcomes. The results reveal decreasing family size and increased duration of breast-feeding for girls, thereby providing evidence to support the aforementioned mechanism.

2.3 Birth Order Effects

Harkonnen (2013) utilizes a German data set and sibling fixed effects models to analyze the magnitude of birth order effects on educational attainment. His findings are consistent with most empirical literature in this context; first-born or older siblings are more intelligent and therefore, proceed to qualify for higher levels of education. He also shows that birth order effects are weaker for females, reflecting traditional gender inequalities, even within the same family. In conjunction with parents' education, their economic resources also play a vital role in amplifying or dampening the magnitude of these effects. (Behrman et al, 1986; Conley 2004, Gratz 2018).

2.4 Fertility Decline

The One-Child Policy in China inevitably resulted in a substantial drop in fertility levels. Ebenstein (2010) causally inferred that although population reduction was the main goal, a by-product of this reform was a highly skewed sex ratio. This was shown by using regional variation in the fines imposed for extra children born during the tenure of the policy. Jayachandran (2017) hypothesizes that the desire for a smaller family size pressurizes parents to practice sex-selective abortions. She proceeds to demonstrate that a fall in fertility rate in India has been associated with 30-35% of the increase in the desired male to female sex ratio. An analysis of the Danish population revealed that sex preferences did not determine fertility rates in one-child families. However, two and three child families with all children belonging to the same sex, were the key drivers behind the increase in fertility rate. (Engholm, et al,1999). Hence, a common consensus in the literature points to the establishment of a causal relationship between fertility rates and sex preference.

2.5 Quantity-Quality Trade-Off Theory

A Q-Q trade-off is encountered when it becomes more expensive to invest in child quality with an increase in quantity of children subject to a household's budget constraint. (Becker and Lewis,1973). This seminal paper paved the way for numerous extensions of the theory and new measures of parental investments. Dang and Rogers (2015) instrument for family size using the distance to the nearest family planning centre in Vietnam and show that a decline in the fertility rate was associated with a sharp rise in educational attainment. In Brazil, using a twin birth instrumental variable strategy, Marteleto and De Souza (2012) found no uniform effects of family size on education over time. Angrist and Schlosser (2010) provide evidence of the absence of a causal link between the two in Israel attributing this to a plausible adjustment of parents' spending and in their extensive and intensive margins of labour supply in response to an increase in family size. However, as a common identification strategy cannot be applied to all countries and contexts, the evidence on the quantity-quality trade-off still remains inconclusive.

2.6 Gender Gaps in Education

A majority of the OECD nations has witnessed women overtaking men in educational attainment. The gender gap has been reversed in the United States with the female high school and college graduation rates surpassing that of male students by 5 percentage points and 7 percentage points respectively. (Autor and Wasserman, 2013). Conversely, in the developing world, governments have implemented nation-wide policies as well as invested in necessary infrastructure with the goal of uplifting and educating the disadvantaged girl child. The construction of latrines in government schools as part of the School Sanitation and Hygiene Education programme (SSHE) significantly increased school enrolment of girls in rural India. (Adukia, 2017). Girls in China have inadvertently benefited from higher levels of education as a result of the One-Child Policy which sought to control the country's explosive population growth (Lee,2011)[3].

3. Data

3.1 Description of the Data Source

This dissertation utilizes the household member records and birth records datasets from three rounds of the National Family Health Survey (NFHS): NFHS-I (1992-1993), NFHS-II (1998-1999) and NFHS-IV (2015-2016). They are nation-wide surveys compiled by the DHS (Demographic and Health Surveys) Program containing a wealth of data on individuals and households across a range of health, education and development indicators. All the information on the variables were collected through separate questionnaires for men, women and households. Different households (70% from rural areas) were interviewed across the total of four time waves with no unique identifier to match them across the rounds. Hence, the data has been organised and analysed in a repeated cross section structure. The observations from the state of Maharashtra have been dropped as this legislation was implemented in this state in 1988. The large sample size and rich controls ensure high statistical power and increased precision of results. Another advantage of this dataset is the consistent availability of information on key variables for over two decades which enables the study of the long-term impact of the PCPNDT Act. Despite all measures taken by the DHS to enhance data quality, survey data is still subject to measurement error, reporting and recollection bias which could compromise on the reliability of results.

[3] Lee (2011) revealed that girls in one-child households received significantly more years of schooling than their female counterparts in multiple-children households. Additionally, the results reported no difference between years of schooling received by both genders, translating into the policy's valuable contribution towards improving the status of women.

3.2 Description of Variables

Dataset 1: Birth Records

The Birth Records dataset comprises of three rounds of the survey: NFHS-I, NFHS-II and NFHS-IV. The second wave was chosen to estimate both the immediate effect of the law on the probability of a female birth since its implementation in 1996, and the fourth wave was included to determine the long- term impact of the same. It is a record of 2,02,000 births containing individual level information on child's and parent's characteristics such as birth order of the child sex of the child, fertility preference, mother's education, abortion and exposure to family planning. The main outcome variable, sex of the newborn child is binary, taking up 0 for males and 1 for females. The treatment variable in the first equation is a dummy variable which takes up a value of 0 for Jammu and Kashmir and 1 for the rest of India. The dummy for time indicator is 0 for years prior to 1996 and 1 for 1996 and after. Controls at the individual level include mother's education in years, birth order of the child, age of mother at first birth, ideal number of boys and girls, and total children ever born which are continuous. Dummies were created for marital status, whether mother had an abortion or not, indicator for whether the previous child was wanted or not and exposure to family planning. Categorical variables include fertility preference and use of contraception. Household level controls include caste ,tribe, type of place or residence (rural or urban) , age of the respondent, marital status, household wealth, number of household members ,sex of the household head, and number of children. State level controls chosen are fertility rate, total population and literacy rate from the Census years of 1991 and 2011. A household wealth index was constructed for both datasets from the following parameters: presence of a toilet, electricity supply, ownership of a television, radio, refrigerator, car, bicycle and scooter. The Principal Component Analysis method was used to compute wealthscores and additionally, wealth quintiles were generated and assigned to rank each household from richest to poorest. Each asset was weighted by the sample weight to derive precise measures of the estimates. The descriptive statistics for this dataset are presented in Table 1.

Dataset 2 : Household Member Records

This dataset consists of the first and the fourth waves of the survey, NFHS-I and NFHS-IV, covering years from 1992 to 2016. It comprises of individual level data on household members who were interviewed on factors related to their education levels, health and household wealth. Individuals selected for this study are females in the 15-19 age cohort in order to allow for sufficient time for them to reach a substantial age for which meaningful measures of their educational attainment could be estimated. Examining the dependent variable, education in years, we can see that it is continuous in nature. Education in years completed by an individual ranges from 0 to 18 years. The treatment and time dummies are the same as which has been adopted in the first dataset. The set of controls at the individual and household levels cover caste, tribe, type of place of residence (rural or urban), age of the respondent, marital status, household wealth, number of household members, sex of the household head, and number of children under the age of five. However, information on parents' educational attainment is not available in this dataset. Therefore, Household Head's Education is used as a proxy for parent's education, because in most Indian households, the head is the primary breadwinner and therefore, his or her socio-economic background determines the future allocation resources for the other members of the family. Dummies were created to indicate whether an individual resides in a rural or an urban area. State level controls were collected from the Census of India tables for the years 1991 and 2011. They include total population, sex ratio and literacy rate and control for time-varying characteristics. A household wealth index was also constructed utilising the same method in dataset 1. Table 2 reports the summary statistics of dataset 2 which will be used for the education regressions.

Table 1: Summary Statistics of Dataset 1 (Birth Records)

	Summary Statistics	
Variable	Mean	Standard Deviation
Dependent Variable		
Sex of Child	0.479	0.499
Independent Variables		
PCPNDT Act	0.966	0.179
Post- PCPNDT Act	0.813	0.389
State	18.336	11.160
Caste or Tribe	1.495	0.821
Employed or Not	0.305	0.460
Mother's Education in Years	4.330	4.865
Age at First Birth	20.138	3.703
Ever Had an Abortion	0.051	0.220
Total Ever Born	3.571	1.933
Exposure to Family Planning	0.437	0.500
Ideal Number of Boys	2.399	9.408
Ideal Number of Girls	2.065	9.431
Last Child Wanted	1.265	0.650
Fertility Preference	3.036	1.125
Birth Order of Child	2.589	1.702
Age of Household Head	44.408	13.264
Sex of Household Head	1.110	0.313
Household Wealth Quintiles	2.888	1.432
Literacy Rate	66.729	11.830
Total Population	6.99e + 07	6.11e +07
Fertility Rate	2.623	0.723

Notes: The summary statistics in this table have been generated from the first, second and fourth waves of the NFHS survey. The means and standard deviations have been reported for the dependent variable and all covariates.

Table 2: Summary Statistics of Dataset 2 (Household Records)

Summary Statistics		
Variable	Mean	Standard Deviation
Dependent Variable		
Education in Years	8.109	3.943
Independent Variables		
PCPNDT Act	0.965	0.183
Post-PNDT Act	0.845	0.361
State	19.066	11.3
Age	16.947	1.393
Caste or Tribe	1.528	0.902
Rural / Urban Indicator	0.264	0.440
Marital Status	0.183	0.447
Number of Household Members	6.455	2.827
Number of Children under the age of five	0.421	0.826
Household Head Education In Years	5.409	6.534
Age of Household Head	48.628	11.703
Sex of Household Head	1.129	0.336
Household Wealth Quintiles	2.947	1.402
Total Population	7.28e+07	6.34e+07
Sex Ratio	0.937	0.038
Literacy Rate	68.592	10.511

Notes: The summary statistics in this table have been generated from the first and fourth waves of the NFHS survey. The means and standard deviations have been reported for the dependent variable and all covariates.

4. Methods and Identification Strategy

4.1 Evaluating the Impact of the PCPNDT Act on Female Education: Differences-in-differences strategy

The Differences-in-differences setup has been utilised in numerous studies to causally infer answers related to policy questions. The identification strategy of this paper exploits the exogenous variation through the implementation of the PCPNDT Act across India. A quasi-experiment modelled around the differences-in-differences framework compares educational attainment between J&K and the other states of India. Conditional on controlling for observables, this allows for any gap in the outcomes between the treatment and control groups to be attributed to the treatment.

Table 3 : Potential Outcomes Framework for Differences-in-Differences[4]

	Pre-PCPNDT Act (T=0)	**Post PCPNDT Act (T=1)**
Rest of India (S_i=1)	E [Y_{io} (0) | S_i=1]	E [Y_{i1} (1) | S_i=1]
Jammu and Kashmir (S_i=0)	E [Y_{io} (0) | S_i=0]	E [Y_{i1} (0) | S_i=0]

Notes: Y_i stands for the outcome variable. S_i and T are the group and time indicators respectively. Z is the treatment indicator given in parentheses which is 1 only for the treated group in the post-treatment period.

[4] The table below gives the values for treatment indicator Z across the group and time dimensions respectively.

	Pre-PCPNDT Act (T=0)	**Post PCPNDT Act (T=1)**
Rest of India (S_i=1)	Z_{i0}=0 (untreated)	Z_{i1}= 1(treated)
Jammu and Kashmir (S_i=0)	Z_{i0}=0 (untreated)	Z_{i0}=0 (untreated)

The setup of the study is organised in two waves of randomly sampled survey data, covering two time periods (1992-1993 and 2015-2016), with the cross-sectional units comprising of the states. The Average Treatment on the Treated (ATET) in the post treatment period can thus be computed as E [Y_{i1} (1) | S_i=1] - E[Y_{i1} (0) | S_i=0].

This can be estimated by the differences-in-differences equation:

Equation 1

$$\boldsymbol{Y_{ist} = \beta_o + \beta_1 PCPNDTAct_s + \beta_2 Post_t + \beta_3 (PCPNDTAct_s \times Post_t) + \beta_4 X_{ist} + \beta_5 S_{st} + e_{ist}}$$

Where ß$_0$ denotes the average value of the outcome variable of the control group at baseline. PCPNDT Act is a dummy with ß1 capturing for any differences between the rest of India and J&K before the policy intervention. Post (ß2) is a time dummy indicator and accounts for factors that could potentially change the years of education variable over time regardless of the enactment of the law. X comprises of the time variant and time invariant individual level controls and S comprises of the state level controls mentioned in Section 3. The differences-in-differences estimator is defined by the interaction term coefficient ß3 derived from differences in outcomes across two dimensions: group and time. Any correlation with error term eist is adjusted for by clustering standard errors robust at the level of treatment assignment which is at that of the states.

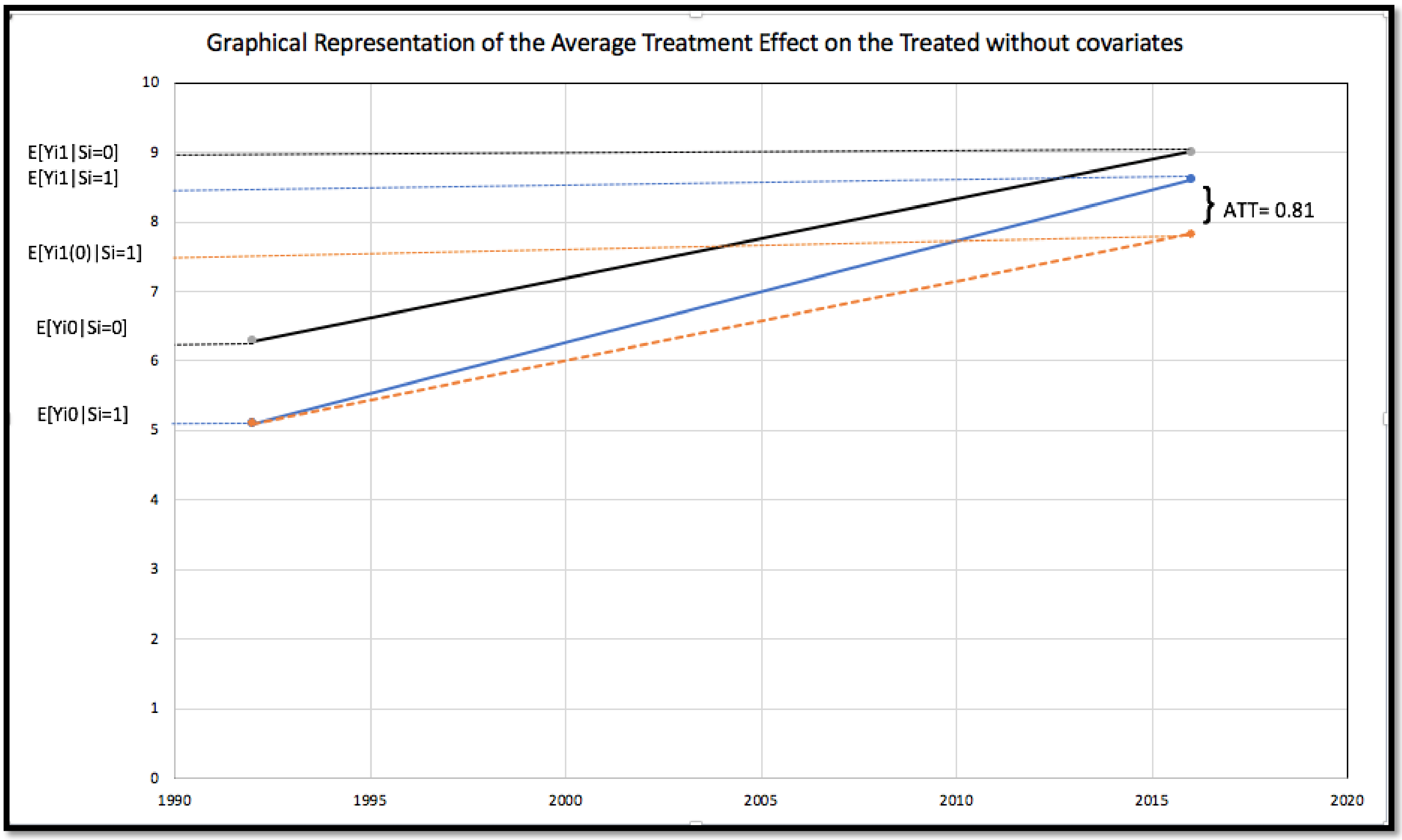
Table 4 presents the difference in the means of the outcome variable between the treated and control states before and after 1996. The crucial requirement of this study design cited in Hastings (2004) is comparability between the two groups. This has been satisfied as individuals from the two groups have similar levels of education before the Act. This difference between Jammu and Kashmir and the other states has narrowed down arguably after the enactment of law. At the preliminary stage, this favours our hypothesis that the treatment indeed did have a positive effect on the educational outcomes of females for the treated states in India. The following sections will determine whether this translates into a causal effect. This motivates the discussion on the validity of the standard differences-in-differences model through an evaluation of the crucial assumptions that are required to be satisfied in order for the results of this study to hold.

Table 4: Differences in Means of Years of Education

	Rest of India	Jammu and Kashmir	Difference	Difference-In-Difference
Pre-PNDT Act	5.101	6.284	1.183	
				0.81
Post-PNDT Act	8.637	9.009	0.372	

Notes: The sample is from the NFHS dataset and consists of 150,160 ever-married females in the 15-19 age cohort. The first two columns comprise of the means of the treated and control group at the baseline (1992) and after the implementation of the PCPNDT Act (2016). The Average Treatment Effect on the Treated (ATET) is calculated as the difference in the difference in means of both groups.

Figure 2 : Average Treatment Effect on the Treated without controlling for covariates



Notes: The above figure displays the magnitude of the Average Treatment Effect on the Treated. The dotted line plots the unobserved outcome in the treatment period i.e. what would have happened to education in the treated states in the absence of the PCPNDT Law. Therefore, in the absence of the treatment, if the treated states exhibit similar trends to J&K after 1996, it can be established that the parallel trends assumption has been satisfied.

4.1.1 Validity Tests for Differences-in-Differences

1. **Parallel Trends**

One of the key assumptions behind the validity of the differences-in-differences strategy is that of the parallel trends which demands a parallel trend in the outcome variable for treated and control groups before the treatment assignment. The most commonly adopted method to test this is generating pre-treatment trends to show that the both the groups have been behaving in a similar manner and that any resulting changes in outcomes can be attributed to the treatment. It has been widely adopted in studies examining the effects of policy interventions on labour market outcomes. (Wolfers 2003). However, the non-availability of data on more than two periods before the treatment assignment renders this method infeasible in this study. Additionally, Propensity Score Matching (PSM) in conjunction with differences-in-differences cannot be employed in this case as it matches treatment and control observations on the conditional probability of being exposed to the treatment only at the individual level, when our treatment assignment is in fact, at the state level. It would therefore fail to eliminate heterogeneity in unobservable characteristics between states which is the level of treatment assignment in this analysis. The solution to proving this assumption stems from the principle that conditional on controlling for time variant and time invariant covariates that would capture differential changes between J&K and the rest of India, it could be expected that time variant factors other than the outcome variable would behave similarly in the absence of the law which has been done in this study.

2. **Falsification Test**

Given the data structure of the analysis, the falsification test can be conducted using two methods, namely assigning placebo treatment and control groups and creating a placebo outcome. The alternative treatment and control groups were redefined by random assignment (coin tossing). A differences-in-differences estimation was once again run on these "false" groups. Theoretically, the treatment should have an insignificant effect as the true control and treatment groups have been shuffled to create noise. If a significant non-zero effect is observed from the placebo law, it can be concluded that our findings are biased which dampens the validity of the differences-in-differences strategy. However, it could also be attributed to the picking up of a false effect. A suitable placebo outcome is one that is highly unlikely to be affected by the law, but may still have a channel through which effects can be produced. This implies that the effect "does not exist" when it "should not exist". Adoption of the differences-in-differences strategy can be validated, if plotting these placebo outcomes against time gives parallel trends in the treatment and control groups because there is no effect of the treatment. In this study, the following placebo outcomes were selected for running the second test; urban area indicator, household's source of drinking water and age of household head.

3. **Is Selection Based On Past Outcomes?**

There could be a selection problem arising from the other states being specifically targeted for the treatment due to persistently high levels of sex-selective abortions. However, the choice for Jammu and Kashmir to be exempted from the policy can be attributed solely to political and institutional factors and not to existing trends in education and pre-natal discrimination, eliminating channels through which endogeneity might occur in this regard.

4. **Exogenous Variation**

Government policies are often sources of exogenous variation and form an ideal setting for a natural experiment by random assignment of the treatment (Meyer, 1995). However, the validity of the study relies on the intervention affecting specific groups and not others. Additionally, biased estimates can be derived if unobserved political factors generate variation in the outcomes, subsequently confounding the true effect. Controlling for variables that have been responsible for the policy intervention can eliminate this bias to an extent (Cook & Tauchen, 1982). In this evaluation of the impact of the PCPNDT Act, there is a clear-cut demarcation between the treated and control groups which can be easily distinguished from each other. Additionally, we have included a rich set of covariates that would have, theoretically, been one of the factors that had brought about the law, such as total number of children born, access to abortion facilities and birth rates. Hence, this substantiates the choice of our experimental design.

4.1.2 Robustness Check

Neighbouring Districts of Jammu and Kashmir

The results obtained could be challenged on the grounds that they do not capture for the true effect of the law as the vast observed and unobserved heterogeneity across states may not be sufficiently controlled for. The combination of several geographical, social and cultural factors would render the treated and control groups incomparable, causing us to speculate whether any false state effects have been picked up. This will be tested by moving closer to J&K and narrowing down the sample to include only regions that border the control state.[5] The neighbouring districts comprise

[5] Muralidharan and Prakash (2013) analyse the effect of an educational programme in the Indian states of Bihar and Jharkhand on girls' secondary school enrolment. They administer a robustness check wherein the sample is trimmed down to only include the bordering districts to eliminate the heterogeneity concern.

of Lahul and Spiti and Chamba in the state of Himachal Pradesh and the districts of Doda, Kargil, Kathua and Kishtwar in J&K. Regression Equation 1 will be run for this smaller sample to ensure robustness of the findings.

4.2 Evaluating the Impact of the PCPNDT Act on the Probability of a Female Birth: A Probit Estimate

The main aim of the study is to investigate the impact of the PCPNDT Act on girls' educational attainment. However, it has been postulated that this impact on education could occur only through the direct effect that the Act has on the probability of a female birth. Hence, in order to defend this economic channel, it becomes necessary to determine the direct effect of the Act on the odds of a female birth and then proceed to demonstrate the mechanism by which the long-term impact on female education operates. As the dependent variable is binary, the probit model serves as a good fit for this estimation. It is a non-linear regression framework that utilises the standard cumulative distribution function (c.d.f) to predict conditional probabilities; Pr (Y=1|X). The treatment effect is therefore not constant across the sample of individuals from different treated states owing to the outcome variable being bounded between 0 and 1. This is an inherent feature of non-linear models. (Athey and Imbens, 2006)

This gives the differences-in-differences Probit model equation:

Equation 2 :

$$\boldsymbol{E}\,[\boldsymbol{Y} \mid \boldsymbol{T},\boldsymbol{S},\boldsymbol{X}] = \boldsymbol{\Phi}(\boldsymbol{\alpha T} + \boldsymbol{\beta S} + \boldsymbol{\gamma TS} + \boldsymbol{X\theta})$$

"The treatment effect on the treated at the time of treatment" (Puhani, 2011) given X is derived as:

$$\boldsymbol{\tau}\,[\boldsymbol{Y} \mid \boldsymbol{T}=\mathbf{1},\boldsymbol{S}=\mathbf{1},\boldsymbol{X}] = \boldsymbol{E}[\boldsymbol{Y}^{\mathbf{1}} \mid \boldsymbol{T}=\mathbf{1},\boldsymbol{S}=\mathbf{1},\boldsymbol{X}] - \boldsymbol{E}[\boldsymbol{Y}^{\mathbf{0}} \mid \boldsymbol{T}=\mathbf{1},\boldsymbol{S}=\mathbf{1},\boldsymbol{X}] = \boldsymbol{\Phi}\,(\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}+\boldsymbol{X\theta}) - \boldsymbol{\Phi}\,(\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{X\theta})$$

Unlike in the linear differences-in-differences model, the cross-difference term is not informative to derive the causal effect of the treatment. This was pointed out by Puhani (2011) who postulated that it is the "cross difference of the conditional expectation of the observed outcome Y minus the cross difference of the conditional expectation of the counterfactual outcome Y^o ". The treatment effect in this case cannot be inferred by directly interpreting the coefficient estimates. As Φ (.) is a monotonic function, only the sign of the treatment effect can be interpreted by that of the interaction term and not it's magnitude. It is calculated as the marginal increase in the coefficient of interest, the cross-difference term γ. In other words, the marginal change in a variable is not constant. It is dependent on not only the treatment coefficient, but also on all the controls in the equation. A significant effect is defined by an increase or decrease in the z-score of the probability of the outcome variable which is different from zero at the 5% level.

This study will determine the probability of the birth of a girl child , the "desired outcome" (Y=1) conditional on covariates (X) , with α and β being defined as the time and group effects which are not constant across the group and time dimensions respectively. We will now proceed to analyse the results in order to conclude whether a causal relationship can be drawn between the introduction of the PCPNDT Act and female education. The robustness of the findings will be verified and addressed in the following sections.

5. Results and Discussion

5.1 Results from Exploratory Analysis

This analysis conducts a simple comparison of means of the outcome variable and covariates in the treated and control samples through the t-test. We test whether the effect of the treatment on the dependent variable and covariates is statistically significant. In other words, the t-statistic or mean difference of the outcome variable must be greater than or equal to the critical value of 1.96 at 95% confidence. The treatment should not have a significant effect on the covariates. It is found that some of the covariates are statistically significantly different. However, we can attribute it to the absence of other controls and the fact that it is only a univariate analysis. Hence this motivates a multivariate regression analysis which is highlighted in the confirmatory analysis section. Appendix I reports the results for the same for the sample comprising of the J&K and the treated neighbouring districts in the state of Himachal Pradesh.

Table 5 : Comparison of Means: J&K and the Rest of India

Variable	Mean in Control Sample	Mean in Treated Sample	Difference in Means	T- Statistic
Education in Years	8.559	8.093	0.466	8.404
Post-PNDT Act	0.834	0.845	-0.011	-2.219
Age	17.006	16.945	0.061	3.125
Caste or Tribe	2.350	1.499	0.851	68.004
Rural / Urban Indicator	0.169	0.267	-0.098	-15.823
Marital Status	0.071	0.187	-0.116	-18.530
Number of Household Members	6.545	6.452	0.092	2.334
Number of Children in household	0.349	0.424	-0.074	-6.3874
Household Head Education In Years	5.507	5.405	0.102	1.110
Age of Household Head	49.323	48.602	0.720	4.378
Sex of Household Head	1.091	1.131	-0.039	-8.438
Wealth Quintiles	3.238	2.936	0.301	15.287
Total Population	1.17e+07	7.50e+07	-6.32e+07	-72.140
Sex Ratio	0.889	0.939	-0.049	-95.828
Literacy Rate	61.145	68.861	-7.716	-52.658

Notes: A simple comparison of means of the treated and control states via the t-test gives the values for the t-statistic for the dependent variable and all the covariates utilised in Equation 1.

5.2 Results from Confirmatory Analysis

5.2.1 Impact on the Birth of a Female Child

As the outcome is binary, the results will be expressed as percentages. From Table 7 , it can be comprehended that for individuals from the rest of India, given that all information on independent variables remain constant, the average predicted probability of a female birth is 50.48% and for Jammu and Kashmir, this is 49.45%. However, it is observed that the time effect, alone, would cause a decline in probability for the treated states had the law not been introduced. The predictive margins of the interaction term have stimulated a rise in the chances of female birth for the treated states to 51.64%, thereby emphasizing the requirement of the interaction with the law.

The Average Marginal Effect can be calculated as the difference of the predictive margins at the mean. The interpretation of these estimates are similar to that of the standard Ordinary Least Squares (OLS) Regressions. The implementation of the law (or in other words a unit increase in the independent variable) has increased the probability of a female birth by 2.56%. The statistical significance of the probit model is denoted by the z-score, which is the coefficient divided by the standard error. The z-score of the interaction term is high (greater than the critical value of 1.96), implying that the effect was different from zero at the 5% level. The complete set of the average marginal effects generated from the probit regression is reported in Appendix II.

Table 6: Average Marginal Effects on the Probability of a Female Birth

	Average Marginal Effect (dF/dx)	z-score	Robust Standard Error
PCPNDT Act	0.0102	1.55	.006578
Post	-.01929	-3.38	.0057152
PCPNDT Act x Post	.02559	4.88	.0052452

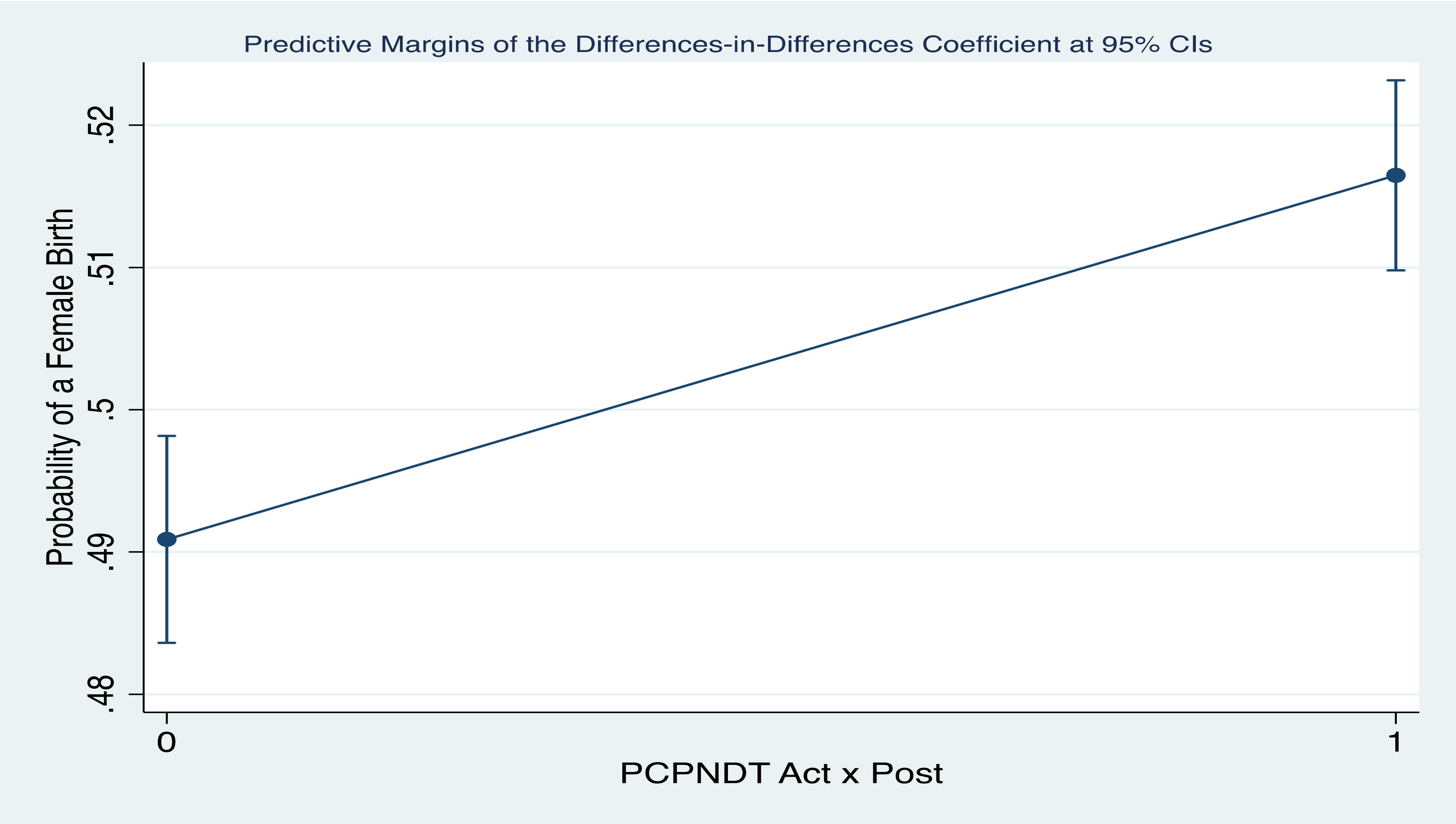
Notes: The Average Marginal Effects have been computed to provide estimates of the probability of the occurrence of a female birth. Robust standard errors are clustered at the state level. Number of observations is 2,01,876.

Table 7: Predictive Margins of the Probability of a Female Birth

	Predictive Margins	Robust Standard Error
Rest of India	0.504815	.0025443
Jammu and Kashmir	0.494595	.0055262
Post-PCPNDT Act	0.4958	.0037742
Pre-PCPNDT Act	0.5150	.0035574
Treated Group	0.5164	.0034057
Untreated Group	0.4908	.0037086

Notes: The table reports predictive margins derived from the probit regressions. Number of observations is 2,01,876

Figure 3: Predictive Margins of the Probability of a Female Birth



Note: The figure presents the impact of the differences-in-differences interaction term on the predicted probability of the birth of a female child from the values of the predicted margins at the mean of the probit model.

5.2.2 Impact on Education

The first model in Column (1) comprises of the basic specification, Column (2) controls for individual characteristics and Column (3) includes both individual and state level covariates as well as state dummies. A cursory glance at Column (1) reveals that the differences-in-differences coefficient is equivalent to the Average Treatment Effect on the Treated (ATET) of 0.81 years represented in Graph 1 in Section 4. This is the crude impact of the law on education without controlling for any other observables that could potentially affect the outcome and it is significant at the 5% level.

Examining the positive constant term in Column (3), it is shown that, individuals in J&K received on average, 3.141 years of education before the enactment of the law and this coefficient is insignificant. The coefficient of the treatment group indicator (PCPNDT Act) gives the difference in the outcome variable between the treated and control states in the pre-treatment period. This difference is negative, implying that in the pre-treatment period, the rest of India was lagging behind J&K in years of education received by 1.127 years. However, this coefficient is again insignificant. Therefore, the difference in education between the two groups exante is inconclusive. The differences-in-differences estimator, denoted by the interaction term coefficient (2.154) , captures the causal effect of the Act. It is positive and significant at the 0.1% level. This explains that the rest of India experienced an increase in educational attainment by 2.154 years between the pre-PCPNDT Act period and post-PCPNDT Act period, relative to the change in years of education that occurred in J&K (-1.706).

This incites the summary of the effects. It is revealed that the sex of household head being male contributes to a lower likelihood of females receiving more education, with a significant decrease equivalent to 4.08 months. Females from the wealthiest households (wealth quintile 5) benefit from higher resource allocation and therefore, receive 2.947 more years of education than those hailing from the poorest households (wealth quintile 1). Individuals belonging to a Scheduled Tribe have continue to remain a disadvantaged community with a decline in education by 5.32 months, significant at the 5% level. Married females experience a significant marginal decline in educational attainment of 2.072 years in comparison to their unmarried counterparts, which could be attributed to instances of girls terminating their education after marriage.

Additionally, Table 8 provides the interpretation of the differences-in-differences coefficients. It was already shown that before the enactment of the law, J&K had higher levels of education than the rest of India. However, after the introduction of the law, the difference in education between the treatment and control groups is positive (1.027), suggesting that females in the rest of India have received more education than those in J&K. Looking at the group dimension, the difference in the outcome variable for the treatment group between the pre-law and post-law periods is positive, denoted by a 0.448 increase in years of education. Conversely, individuals in the control group experienced a significant decrease in education received between the two time periods shown by the coefficient -1.706. Therefore, it can be inferred that while female educational attainment in J&K was ahead of the rest of the country before 1996, the other Indian states experienced a significant increase in education while J&K witnessed a decline in the same over time. Hence, the findings reaffirm the main hypothesis of the study ; the increase in female education in the treatment group can be attributed to the Act. The difference-in-difference estimate indeed shows a positive effect of the Act for the rest of India when compared to J&K. However, one has to interpret these results with caution because we did not explicitly prove the parallel trend assumption in the pre-treatment period and J&K has experienced a deterioration in education over time, bringing us to the conclusion that J&K is not a highly credible control group.

Table 8 : Interpretation of the Coefficients of the Linear Differences-in-Differences Equation

	Rest of India (Treatment)	Jammu and Kashmir (Control)	Difference
Pre PCPNDT Act	$\beta_o + \beta_1 = 2.014$	$\beta_o = 3.141$	$\beta_1 = -1.127$
Post PCPNDT Act	$\beta_o + \beta_1 + \beta_2 + \beta_3 = 2.462$	$\beta_o + \beta_2 = 1.435$	$\beta_1 + \beta_3 = 1.027$
Difference	$\beta_2 + \beta_3 = 0.448$	$\beta_2 = -1.706$	$\beta_3 = 2.154$

Notes: The table presents the interpretation of the coefficients for the standard linear differences-in-differences model. Coefficients reported are the same as in Equation 1.

Table 9: Impact of the PCPNDT Act on Years of Education for Females

	(1)	(2)	(3)
Constant	6.284***	-3.540***	3.141
	(3.52e+12)	(-7.66)	(0.46)
PCPNDT Act	-1.183**	-0.252	-1.127
	(-2.84)	(-1.00)	(-1.76)
Post	2.725***	1.725***	-1.706*
	(1.52e+12)	(19.60)	(-2.40)
PCPNDT Act x Post	0.810*	0.401	2.154***
	(2.49)	(1.69)	(8.26)
Household Head Education		0.101***	0.101***
		(6.51)	(6.43)
Household Members		-0.0543**	-0.0268*
		(-3.67)	(-2.07)
Number of Children under the age of 5		-0.375***	-0.367***
		(-14.21)	(-12.87)
Sex of Household Head being Male		-0.385***	-0.340***
		(-5.17)	(-4.82)
Wealth Quintiles			
Second Lowest		1.452***	1.336***
		(12.25)	(12.20)
Middle		2.231***	2.026***
		(12.62)	(12.94)
Second Highest		2.697***	2.520***
		(15.67)	(16.37)
Highest		3.088***	2.947***
		(23.21)	(18.76)
Rural/Urban Indicator		-0.0874	-0.0941
		(-0.66)	(-0.68)
Scheduled Caste		-0.171	-0.158
		(-0.96)	(-1.12)
Scheduled Tribe		-0.546**	-0.444*
		(-3.27)	(-2.66)
Marital Status Indicator		-2.176***	-2.072***
		(-27.37)	(-34.33)
Individual Controls	No	Yes	Yes
State Controls	No	No	Yes
Number of Individuals	149911	148832	148832
Adjusted R-squared	0.1038	0.3287	0.3462

Notes: This table reports results obtained from regressions on education. T-statistic is in parentheses, where * p<0.05 , ** p<0.01 and *** p<0.001. Standard errors are clustered robust at the state level.

5.3 Results from Robustness Check

Table 10 presents results for the narrowed down sample (including J&K and only the neighbouring districts in Himachal Pradesh as the treated group) in Column (1) and Column (2) shows the results for the full sample (J&K and the rest of India). It can be seen from the treatment indicator in Column (1), that the difference in education in years between the treated and control groups in the pre-treatment period is negative. This suggests that the treated neighbouring districts had significantly lower levels of education than J&K before the law**.** The differences-in-differences estimate is positive and significant. This explains that education of females in treated districts in Himachal Pradesh has increased by 3.263 years between the pre-treatment period and treatment periods relative to the change in education in J&K (increase of 1.598 years of education). This significant increase in education in the treated neighbouring districts can therefore be attributed to the Act. Hence, we find that the treatment effect of the Act is positive and significant in both the narrowed-down and full samples making our results robust. The full set of results along with the interpretation of coefficients are reported in Appendix III.

Table 10: Impact of the PCPNDT Act on Neighbouring Districts vs Full Sample

	Neighbouring Districts and J&K	**Rest of India and J&K**
Constant	-1.563	3.141
	(-1.84)	(0.46)
PCPNDT Act	-2.515*	-1.127
	(-17.63)	(-1.76)
Post	1.598*	-1.706*
	(21.19)	(-2.40)
PNDT Act x Post	3.263*	2.154***
	(51.82)	(8.26)
Number of Individuals	1208	148832
Adjusted R-squared	0.2861	0.3462

Notes: The table presents a comparison of the effect of the Act on female education between the neighbouring district sample and the original sample. The t-statistic is given in parentheses where * p<0.05 , ** p<0.01 and *** p<0.001. Standard errors are clustered robust at the state level.

5.4 Results from Placebo Tests

5.4.1 First Placebo Test

Columns (1) and (2) in Table show that the placebo treatment has a positive insignificant effect on the outcome variable, education in years. The addition of both individual and state controls in Column (3) changes the sign of the placebo treatment effect but still remains insignificant. Hence, this reaffirms the validity of the differences-in-differences strategy employed in our study.

5.4.2 **Second Placebo Test**

Table reports the results for the regressions run using placebo outcomes. Columns (1) , (2) and (3) present the treatment effect on source of drinking water, age of household head and place of residence being urban respectively. We find that the treatment has had an insignificant effect on all the three placebo outcomes. This again emphasizes the validity of our identification strategy.

Table 11: Effect of the Placebo Treatment on Female Education

	(1)	(2)	(3)
Constant	5.096***	-3.835***	4.197
	(8.16)	(-6.46)	(0.49)
Placebo Treatment	0.00916	0.0483	-0.618
	(0.01)	(0.12)	(-0.95)
Post	3.560***	2.082***	0.974
	(7.20)	(5.81)	(1.38)
Placebo Treatment x Post	-0.239	0.109	0.0360
	(-0.39)	(0.28)	(0.13)
Household Head Education		0.102***	0.102***
		(6.54)	(6.35)
Number of Household Members		-0.0454**	-0.0278*
		(-3.60)	(-2.11)
Number of Children under the age of five		-0.383***	-0.370***
		(-13.80)	(-12.55)
Sex of Household Head being Male		-0.400***	-0.354***
		(-5.32)	(-4.91)
Wealth Quintiles			
Second Lowest		1.446***	1.342***
		(11.98)	(12.01)
Middle		2.168***	2.004***
		(12.96)	(13.17)
Second Highest		2.634***	2.504***
		(16.69)	(16.26)
Highest		3.063***	2.956***
		(22.64)	(18.19)
Rural/Urban Indicator		-0.0950	-0.0969
		(-0.68)	(-0.67)
Scheduled Caste		-0.128	-0.0615
		(-0.67)	(-0.34)
Scheduled Tribe		-0.439*	-0.365
		(-2.29)	(-1.92)
Marital Status		-2.192***	-2.092***
		(-28.71)	(-31.91)
Individual Controls	No	Yes	Yes
State Controls	No	No	Yes
Number of Individuals	144756	143689	143689
Adjusted R-squared	0.1008	0.3224	0.3365

Notes: The table provides estimates of the effect of the placebo treatment on female education. The t- statistics are in parentheses, where * p<0.05 ** p<0.01*** p<0.001. Standard Errors are clustered robust at the state level.

Table 12: Effect of the Treatment on Placebo Outcomes

	Source of Drinking Water (1)	Age of Household Head (2)	Residence in Urban Area (3)
Constant	-14.18	19.45	-1.108
	(-0.34)	(1.90)	(-1.49)
PCPNDT Act	2.159	-1.522	-0.0991
	(0.68)	(-1.68)	(-1.60)
Post	-3.201	3.615*	-0.249**
	(-0.71)	(2.32)	(-3.35)
PCPNDT Act x Post	1.293	-0.0393	0.0491
	(0.60)	(-0.03)	(1.33)
Adjusted R-squared	0.0796	0.1194	0.2108
Number of Individuals	149056	149077	149077

Notes: The table provides the estimates of the effect of the PCPNDT Act on a set of placebo outcomes. The t- statistics are in parentheses, where * $p<0.05$ ** $p<0.01$*** $p<0.001$. Standard Errors are clustered robust at the state level.

5.5 Mechanisms

The impact of the PCPNDT Act on female education would be expected to occur only via its direct effect on the sex ratio, specifically the odds of occurrence of a female birth. However, in the presence of the ban on sex-selective abortions, parents will no longer be able to control the gender of children being born to them. In other words, the Act could trigger a shift towards the natural sex ratio by increasing the probability of a female birth. We can therefore expect the positive impact of the PCPNDT Act on female education to occur through either one or both of the following channels.

A highly skewed sex ratio could potentially create widespread scarcity of brides in the marriage market. This could result in higher instances of early marriages and child marriages and a lower likelihood of women completing their education. It would thus worsen the socio-economic status of women. (Jiang et al. 2007). The PCPNDT Act through its effect on increasing likelihood of a female birth, could resolve the issue of bride scarcity by restoring gender balance. This describes the first channel wherein the fall in instances of early marriages that can be attributed to the Act, enables women to complete their education without any hindrances.

Another important consequence of a falling sex ratio is the reduction in the number of women of reproductive age. (Miller 1981, Caldwell 2001). This could have been one of the socio-economic factors that stimulated the downward movement in fertility rate in India. This trend has been illustrated in Figure 1 in Section 1. The decline in fertility continued well into the 1990s before the PCPNDT Act was introduced. This implies that in each new birth cohort, fewer number of children are being born to a set of parents. The core aim of the Act was to correct gender imbalance. Resource constraints would also contribute to the reduction in the total number of children born to them as well as limit the excess number of girls born due to the desire for a son. (Das Gupta 2005). Girls born after the implementation of the Act would now be competing for resources with a fewer number of siblings which could increase their educational attainment in years. To summarise, the increased probability of a female birth reinstated by the law against the prevailing background of fertility decline provides the second channel through which higher educational investments, and consequently improvement in female educational attainment operate. Hence, the main findings corroborate the two hypotheses of the study.

5.6 Limitations

A major limitation of the paper lies in the fact that the parallel trend assumption is not testable in practice. Hence, the results of any differences-in-differences analysis should be interpreted with caution. Secondly, as this study is based on observational data, there could always be unobserved heterogeneity and omitted variables that could bias the findings. It could also be argued that J&K could be a bad control group because it is a noisy state. This could lead to the potential risk of a downward or upward bias due to the effect of this noise showing up in our results. Finally, a noteworthy point that could deem our findings questionable is the time duration of the study. As the analysis spans for more than two decades, there exists a possibility of time effects confounding estimates. Although all available time-varying factors have been controlled for, the risk of bias cannot be entirely wiped out.

6. Conclusion

An evaluation of the PCPNDT Act was conducted to estimate the long-run impact of a nation-wide policy intervention. Using individual level observational data from the NFHS and a differences-in-differences identification strategy, the paper empirically examines whether the law has had a causal effect on the education of females, and determines whether this effect, if any, can be attributed to the Act. The results quantify the direct and indirect effects of the law; the probability of a female birth and educational attainment respectively. The probability of a female birth has significantly increased, echoing the findings of Nandi and Deolalikar (2013). It was also estimated that there has been a positive and significant impact on education received by females in the treated states. While the Act has indeed had a positive impact on female education for the treated states, these results need to be interpreted with caution as it was discovered that J&K which had higher levels of education than the rest of India prior to the Act, experienced a fall in education over time after the law, rendering it to be a less credible control group.

This paper builds on to existing literature and analyses this legislation from the angle of human capital formation. It is the first attempt to examine the long-term consequences of this demographic policy in the form of educational attainment. Another unique aspect of the study is the in-depth exploration of the possible economic channels linking the likelihood of a female birth to female education. Hence, this research and its findings have tried to contribute to the long-run effects of a policy response in reference to the global south.

One of the possible extensions that can be made to this study is including males in the sample and introducing a triple difference approach, facilitating the estimation of the gender gap in education. In the context of India, where son preference is prevalent, it would also be prudent to calculate separately, this gender gap in the cases of one-child families and families with two or more children. This will enable the identification of the causal impact of the Act on females relative to males by birth order and sibling sex composition which would be an interesting topic for future research. Hence, this study provides an impetus to address and examine the implications of the PCPNDT law and other similar policies around the world in the context of new indicators and cohorts.

7. **Appendices**

Appendix I

Table 13 : Comparison of Means- J&K And Neighbouring Districts

Variable	Mean in Control Sample	Mean in Treated Sample	Difference in Means	T- Statistic
Education in Years	8.897	8.629	0.268	0.947
Post-PNDT Act	0.830	0.778	0.051	2.031
Age	17.004	16.844	0.159	1.723
Caste or Tribe	1.725	1.260	0.465	6.953
Rural / Urban Indicator	0.114	0.462	0.068	3.493
Marital Status	0.082	0.108	-0.025	-1.251
Number of Household Members	6.437	6.039	0.398	2.471
Number of Children in household	0.349	0.424	-0.074	1.275
Household Head Education In Years	5.849	4.046	1.802	3.748
Age of Household Head	50.057	51.864	-1.806	-2.248
Sex of Household Head	1.120	1.221	-0.101	-4.357
Wealth Quintiles	2.863	3.049	-0.185	-2.0429
Total Population	1.17e+07	6490109	5235192	49.146
Sex Ratio	0.889	0.972	- 0.083	-9.8e+02
Literacy Rate	61.013	78.62	-17.607	-21.220

Notes: A simple comparison of means of the treated and control states via the t-test gives the values for the t-statistic for the dependent variable and all the covariates utilised in Equation 1.

Appendix II

Table 14: Average Marginal Effects of Probit Regression Estimating Probability of a Female Birth

Variable	Average Marginal Effect	z-score	Robust Standard Error
PCPNDT Act	0.0102013	1.55	0.006578
Post	-0.019295	-3.38	0.005715
PCPNDT Act x Post	0.0255993	4.88	0.005245
Rural/Urban Indicator	-0.0083721	-2.61	0.003206
Sex of Household Head Male	-0.0706052	-9.78	0.007175
Sex of Household Head Female	-0.0664223	-6.81	0.00968
Wealth Quintile- 1	0.0636012	0.56	0.11406
Wealth Quintile- 2	0.0677491	0.59	0.113582
Wealth Quintile- 3	0.0739051	0.65	0.113349
Wealth Quintile-4	0.0754723	0.66	0.112681
Wealth Quintile-5	0.0770203	0.67	0.114682
Scheduled Caste	-0.0008368	-0.07	0.011698
Scheduled Tribe	-0.0294318	-2.5	0.011747
Other Caste	0.0039014	0.37	0.010459
Mother Currently Employed	-0.0075697	-2.31	0.003283
Ever Had an Abortion	0.0077505	2.05	0.003779
Total Ever Born	0.0435964	10.01	0.004353
Family planning- No	0.018093	0.35	0.051972
Family Planning - Yes	0.0185656	0.36	0.051576
Mother's Education	0.0004686	1	0.00047
Desire to have more children	0.0933417	10.59	0.008745
Birth Order1	0.293326	5.17	0.053364
Birth Order 2	0.2994867	5.45	0.050213
Birth Order 3	0.2726049	5.36	0.045552
Birth Order 4	0.242104	5	0.043355
Birth Order 5	0.213789	4.88	0.039772
Last Child Wanted Then	0.084083	2.58	0.032342
Last Child Wanted Later	0.1079655	3.3	0.03208
Last Child Not Wanted	0.0787942	2.34	0.033363
Ideal Number of Boys	-0.0614668	-15.5	0.003966
Ideal Number of Girls	0.0609391	15.55	0.003919
Age at First Birth	0.0006158	1.33	0.000462
State Literacy Rate	-0.0004926	-2.19	0.000225
State Population	2.81E-11	0.48	5.83E-11
Fertility Rate	-0.0200988	-3.91	0.005139

Note: The complete set of results of the probit regression on probability of a female birth are reported. It presents the average marginal effects, z-scores and robust standard errors. Standard errors are clustered at the state level.

Appendix III

Table 15 : Interpretation of the Coefficients of the Linear Differences-in-Differences Equation for the narrowed down sample of J&K and Neighbouring districts in Himachal Pradesh

	Neighbouring Districts in Himachal Pradesh (Treatment)	Jammu and Kashmir (Control)	Difference
Pre PCPNDT Act	$\beta_o + \beta_1 = -4.078$	$\beta_o = -1.563$	$\beta_1 = -2.515$
Post PCPNDT Act	$\beta_o + \beta_1 + \beta_2 + \beta_3 = 0.783$	$\beta_o + \beta_2 = 0.035$	$\beta_1 + \beta_3 = 0.748$
Difference	$\beta_2 + \beta_3 = 4.861$	$\beta_2 = 1.598$	$\beta_3 = 3.263$

Notes: The table presents the interpretation of the coefficients for the standard linear differences-in-differences model. Coefficients reported are the same as in Equation 1.

The constant term $\boldsymbol{\beta_0}$ is negative and insignificant. Females in J&K had relatively higher levels of education than the neighbouring treated districts before the introduction of the Act, which is denoted by the negative coefficient of $\boldsymbol{\beta_1}$. Both the treated districts in Himachal Pradesh , and untreated districts in J&K escalated significantly in education received by 4.861 years and 1.598 years respectively between the pre-treatment and treated periods. Finally, the coefficient of the differences -in-differences estimate $\boldsymbol{\beta_3}$ implies that the treated districts experienced an increase in education of 3.263 years relative to the increase in education in the control districts. This estimate is significantly different from zero at the 5% level. It can be observed that J&K, from having higher levels of education than neighbouring treated districts at the baseline, experienced an increase in the level of education between the pre-treatment and treatment periods. It was earlier observed in the full sample that , from having higher levels of education than the rest of India before the treatment, J&K exhibited a decline in education between the pre-treated and treatment periods. This would mean that J&K is a more credible control group for the narrowed down treated sample than for the full treated sample comprising of the other Indian states.

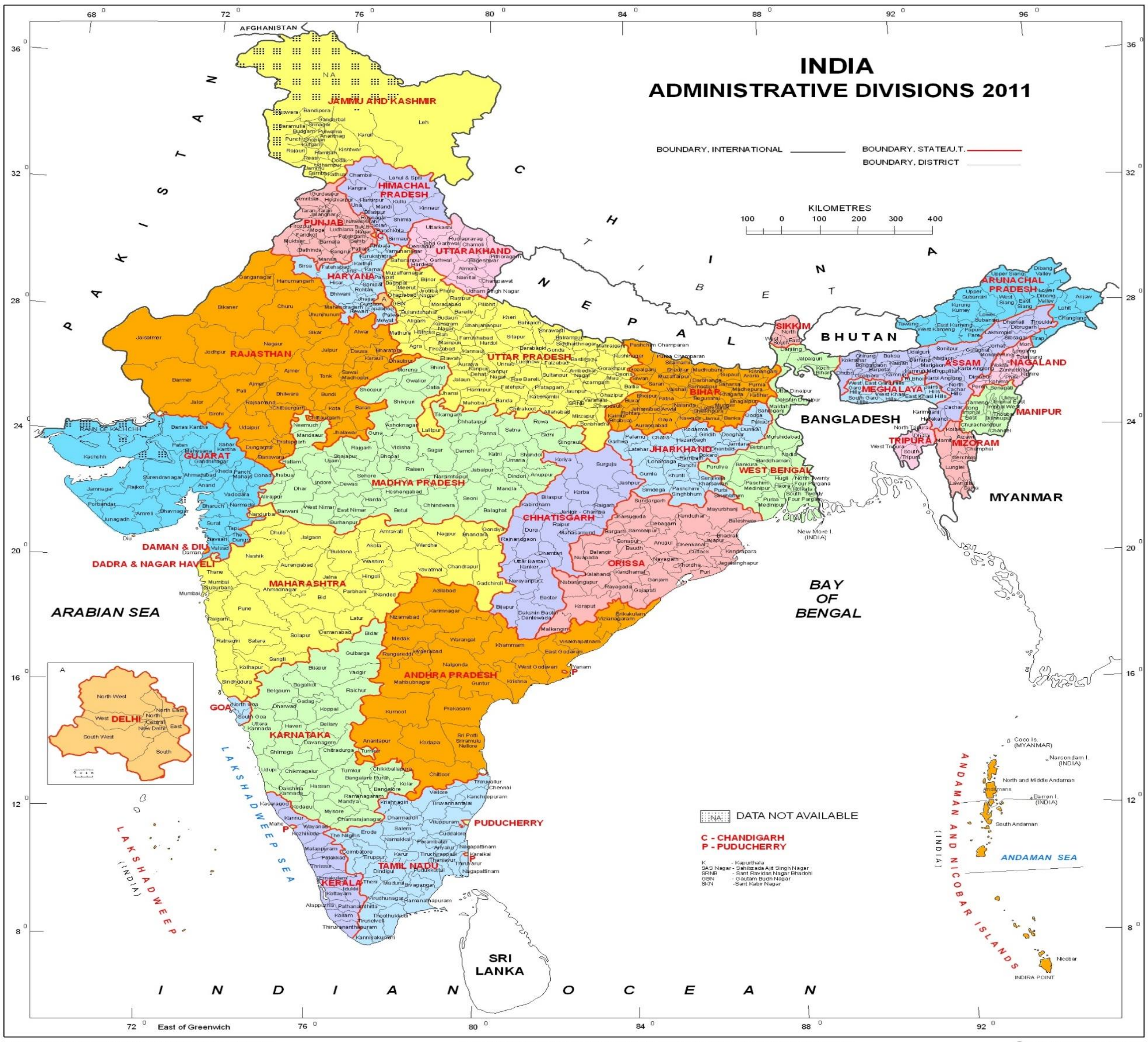
Table 16: Impact of the PCPNDT Act on Neighbouring Districts vs Full Sample

	Neighbouring Districts and J&K	Rest of India and J&K
Constant	-1.563	3.141
	(-1.84)	(0.46)
PCPNDT Act	-2.515*	-1.127
	(-17.63)	(-1.76)
Post	1.598*	-1.706*
	(21.19)	(-2.40)
PNDT Act x Post	3.263*	2.154***
	(51.82)	(8.26)
Household Head Education	0.0249	0.101***
	(1.78)	(6.43)
Household Members	0.00124	-0.0268*
	(0.02)	(-2.07)
Number of Children	-0.497	-0.367***
	(-1.64)	(-12.87)
Sex of Household Head being Male	-0.247	-0.340***
	(-1.24)	(-4.82)
Wealth Quantiles		
Second Lowest	2.299*	1.336***
	(24.74)	(12.20)
Middle	1.842	2.026***
	(3.13)	(12.94)
Second Highest	2.714	2.520***
	(9.33)	(16.37)
Highest	2.719*	2.947***
	(15.65)	(18.76)
Rural/Urban Indicator	0.917*	-0.0941
	(31.17)	(-0.68)
Scheduled Caste	-0.0578	-0.158
	(-0.24)	(-1.12)
Scheduled Tribe	-0.161	-0.444*
	(-0.77)	(-2.66)
Marital Status Indicator	-2.618*	-2.072***
	(-3.31)	(-34.33)
Number of Individuals	1208	148832
Adjusted R-squared	0.2861	0.3462

Notes: The table reports the full set of results of the robustness check. It presents a comparison of the effect of the Act on female education between the neighbouring district sample and the original sample. The t-statistic is given in parentheses where * p<0.05 , ** p<0.01 and *** p<0.001. Standard errors are clustered robust at the state level.

Appendix IV

Figure 4 : Administrative Map of India



Based upon Survey of India map with the permission of the Surveyor General of India.

The territorial waters of India extend into the sea to a distance of twelve nautical miles measured from the appropriate base line.

The interstate boundaries between Arunachal Pradesh, Assam and Meghalaya shown on this map are as interpreted from the North-Eastern Areas (Reorganisation) Act, 1971, but have yet to be verified.

Notes: The Map of India was sourced from the Census of India, 2011. The treatment states cover the entire country with the exception of the northern most state of Jammu and Kashmir.

Figure 5: Administrative Map of Jammu and Kashmir

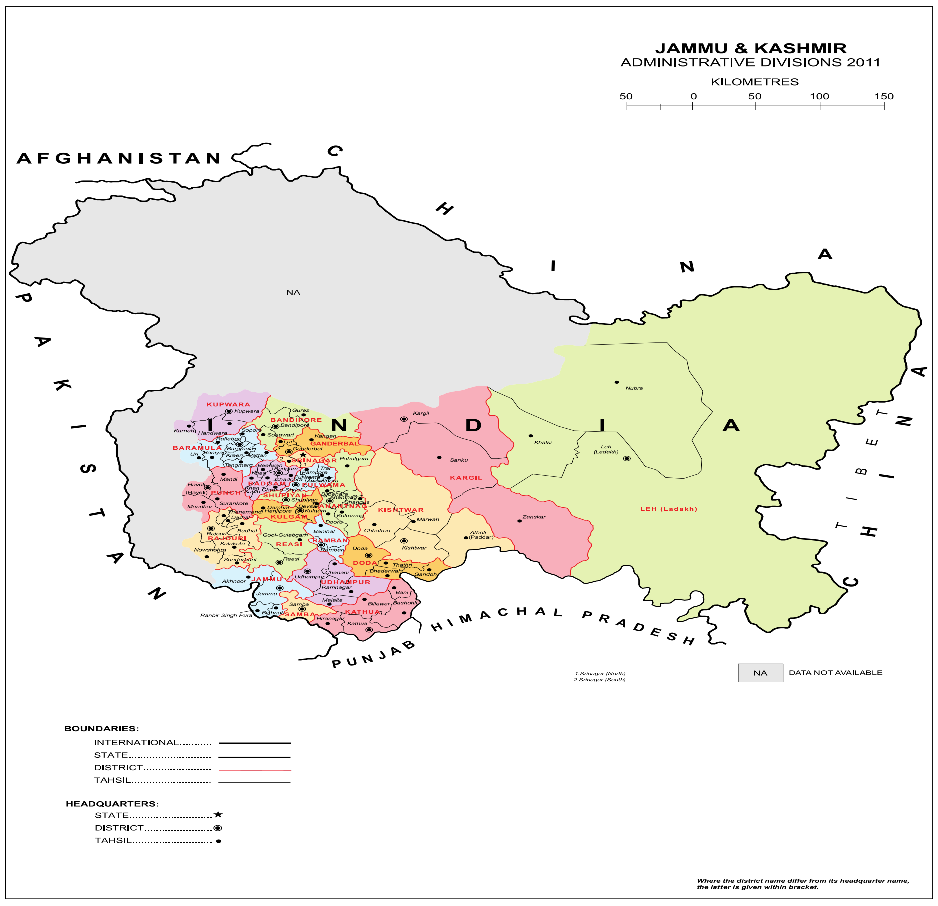


Figure 6: Administrative Map of Himachal Pradesh

Note: The maps of Himachal Pradesh and J&K have been sourced from Census of India (2011). This depicts the geographical area of the control state and its neighbouring treatment state including the bordering districts in our analysis. Additional regressions are run on education using this narrowed down sample as a robustness check.

8. References

1. Angrist, J, Lavy, V & Schlosser, A 2010, "Multiple Experiments for the Causal Link between the Quantity and Quality of Children" *Journal of Labor Economics*, vol. 28, no. 4, pp. 773-823.
2. Arindam Nandi ,2015 "The Unintended Effects of a Ban on Sex-Selective Abortion on Infant Mortality: Evidence from India" *Oxford Development Studies, 2015 Vol. 43, No. 4, 466–482, http://dx.doi.org/10.1080/13600818.2014.973390*
3. Arindam Nandi , Anil B. Deolalikar 2013. "Does A Legal Ban On Sex-Selective Abortions Improve Child Sex Ratios? Evidence From A Policy Change In India "*Journal of Development Economics 103 (2013) 216–228*
4. Arnold, F., M. K. Choe, and T. K. Roy. 1998. "Son preference, the family-building process and child mortality in India", *Population Studies 52(3): 301–315*
5. Arnold F, Sunita Kishor, T.K. Roy, 2004. "Sex-Selective Abortions in India" , *Population and Development Review* , Volume 28 Issue 4.
6. Athey, S. & Imbens, G. W., 2006. "Identification and inference in nonlinear difference-in differences models", *Econometrica,* March, 74(2), pp. 431-497.
7. Autor, D, Wasserman, M, 2013. "Wayward Sons: The Emerging Gender Gap in Labour Markets and Education" [Washington, District of Columbia]: Third Way, 2013.
8. Avraham Ebenstein ,2010. "The "Missing Girls" of China and the Unintended Consequences of the One Child Policy" , *The Journal Of Human Resources* Volume 45, pp. 87-115.
9. Becker, G. S. & Lewis, H. G., 1973. "On The Interaction Between The Quantity And Quality Of Children." *Political Economics*, 84(4), pp. 143-162.
10. Caldwell JC. 2001. "What Do We Know About Asian Population History? Comparisons Of Asian And European Research. In Asian Population History", ed. T Liu, J Lee, DS Reher, O Saito, W Feng, pp. 3–23. Oxford:Oxford Univ. Press.
11. Das GuptaM. 2005. "Explaining Asia's Missing Women": *Population Dev. Review* 31(3):529–35
12. Ester Boserup 1970. *"Woman's Role in Economic Development",* New York: St. Martin's Press,
13. Hai-Anh H. Dang and F. Halsey Rogers ,2015. "The Decision to Invest in Child Quality over Quantity: Household Size and Household Investment in Education in Vietnam" *The World Bank Economic Review, Vol. 30, NO. 1, pp. 104 – 142 doi:10.1093/wber/lhv048 Advance Access Publication August 25, 2015*
14. Jacobsen R., H Moller, G. Engholm, 1999. "Fertility Rates In Denmark In Relation To The Sexes Of Preceding Children In The Family". *Human Reproduction,*14 (4),1127-1130.
15. Jayachandran, Seema. 2017. "Fertility Decline and Missing Women." *American Economic Journal: Applied Economics*, 9 (1): 118-39.
16. Jere R. Behrman and Paul Taubman,1986. "Birth Order, Schooling, and Earnings", *Journal of Labor Economics,*,Vol. 4, No. 3, Part 2: The Family and the Distribution of Economic Rewards (Jul., 1986), pp. S121-S14
17. Jha, P., R. Kumar, P. Vasa, N. Dhingra, D. Thiruchelvam, And R. Moineddin 2006: "Low male-to-female sex ratio of children born in India: national survey of 1.1 million households," *The Lancet,* 367(9506), 211–218
18. Jiang Q, Attan´e I, Li S, Feldman M. 2007. "Son preference and the marriage squeeze in China". See Attan´e & Guilmoto 2007b, pp. 347–63
19. Juho Härkönen,2013. Birth Order Effects on Educational Attainment and Educational Transitions in West Germany Article Navigation, *European Sociological Review,* Volume 30, Issue 2, April 2014, Pages 166–179, https://doi.org/10.1093/esr/jct027
20. Junhong, C., 2001 "Prenatal sex determination and sex-selective abortion in rural central China" *Population and Development Review* 27 (2), 259–281.
21. Justin Wolfers,2003. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results", NBER Working Paper No. 10014 Issued in October 2003
22. Karthik Muralidharan & Nishith Prakash, 2013. "Cycling to School: Increasing Secondary School Enrollment for Girls in India," Working papers 2013-24, University of Connecticut, Department of Economics.
23. Hastings, Justine, S. 2004. "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California", *American Economic Review*, 94 (1): 317-328.
24. Lee, M.-H., 2011. "The One-Child Policy and Gender Equality in Education in China: Evidence from Household Data", *J Fam Econ Issues*, October, Volume 33, pp. 41-52.
25. Leela Visaria, 2008 "Improving the Child Sex Ratio: Role of Policy and Advocacy", *Economic and Political Weekly,*Vol. 43, No. 12/13 (Mar. 22 - Apr. 4, 2008), pp. 34-37

26. Letícia J. Marteleto, Laetícia R. de Souza, 2012. “The changing impact of family size on adolescents' schooling: assessing the exogenous variation in fertility using twins in Brazil”, *Demography,* 2012 Nov;49(4):1453-77. doi: 10.1007/s13524-012-0118-8.
27. Letícia J. Marteleto, Laetícia R. de Souza, September 2013, “The Implications of Family Size for Adolescents’ Education and Work in Brazil: Gender and Birth Order Differences “*Social Forces* 92(1) 275–302, doi: 10.1093/sf/sot069
28. Lisa R. Inchani and Dejian Lai,2008. “Association of Educational Level and Child Sex Ratio in Rural and Urban India” , *Social Indicators Research,* Vol. 86, No. 1 (Mar., 2008), pp. 69-81
29. Luojia Hu and Analia Schlosser,2015. “Prenatal Sex Selection And Girls’ Well-Being: Evidence From India”, *The Economic Journal,* 125 (September), 1227–1261. Doi: 10.1111/ecoj.12259 © 2015 Royal Economic Society.
30. Meyer, B. D., 1995. “Natural and quasi-experiments in economics”, *Journal of Business and Economic Statistics,* Volume 13, pp. 151-161.
31. Michael Gratz,2018. “Competition in the Family: Inequality between Siblings and the Intergenerational Transmission of Educational Advantage.” *Sociological Science* 5: 246- 269.
32. Miller, B. (1981), “The Endangered Sex” Ithaka: Cornell University Press.
33. Nancy Qian 2009, “Quantity-Quality And The One Child Policy: The Only-Child Disadvantage In School Enrolment In Rural China” , *Working Paper 14973 http://www.nber.org/papers/w14973*
34. Philip J. Cook and George Tauchen, 1982. “The Effect of Liquor Taxes on Heavy Drinking” *The Bell Journal of Economics,* Vol. 13, No. 2 (Autumn, 1982), pp. 379-390
35. Priti Kalsi 2015, “Abortion Legalization, Sex Selection, and Female University Enrolment in Taiwan*” Economic Development and Cultural Change , Vol. 64, No. 1 (October 2015), pp. 163- 185*
36. Puhani, P. A., 2011. “The Treatment Effect, the Cross Difference, and the Interaction Term in Non-Linear 'Difference-in-Differences' Models”, *Economic Letters,* December, Volume 115, pp.85-87.
37. R.A. Echávarri, & Ezcurra, R.2006. “Education And Gender Bias In The Sex Ratio At Birth: Evidence from India”, *Demography* (2010) 47: 249.
38. Rune Jacobsen, Henrik Møller and Gerda Engholm, 1999. “Fertility Rates In Denmark In Relation To The Sexes Of Preceding Children In The Family “, *Human Reproduction vol.14 no.4 pp.1127–1130*
39. S Anukriti Sonia Bhalotra Hiu Tam, 2016. “On the Quantity and Quality of Girls: New Evidence on Abortion, Fertility, and Parental Investments*”, IZA Discussion Paper No. 10271*
40. Sen, A. (1990), “More than 100 million women are missing”, *New York Review of Books*, 37(20).
41. Tim Dyson 2012. “Causes and Consequences of Skewed Sex Ratios “*Annual Review of Sociology, 2012. 38:443–61, 2012*
42. Tuljapurkar S, Nan L, Feldman M. 1995. “High sex ratios in China’s future”, *Science 2*67(5199):87476
43. Westley SB, Choe MK. 2007. “How does son preference affect populations in Asia?”, *East-West Cent. Asia Pac.* Issues 84:1–12
44. Yamamoto, Teppei 2016. “Introduction to Causal Inferences” Keio University, Spring 2016.