# Retain, promote or support: How to reduce inequality in elementary education

Madhuri Agarwal

December 8, 2019

**Abstract**

This study provides evidence on the policy of retention on educational outcomes in post-primary education. Retention implies asking students to repeat the grade if they fail in a grade-level exam. Those in favor of retention believe that a threat of repeating a grade is an effective tool to motivate parents, teachers, and students and thus improve education quality. Previous studies from developed countries have found retention to be an expensive and largely ineffective method to help students progress through school. The objective of this study is to find how retention affects student outcomes, especially in a developing country context. In 2010, India launched a large scale education reform, the Right to Education Act (2010), with a no-detention policy clause that abolished the practice of retention for all elementary grades across the country. I use difference in difference method to estimate the effect of the policy on student drop-out rates and learning outcomes. I exploit variation in exposure to the policy due to differences in pre-existing trends in repetition rates across districts. It was found that the no-detention policy helped in reducing drop out rates in elementary education especially for children from economically poor background. Although, the policy negatively affected learning outcomes. Further research is needed to understand the mechanism behind the fall in the quality of education after the implementation of the policy.

**Keywords:** Retention, elementary education, inequality in education
**JEL Codes:** I2, I24, I28

# 1 Introduction

Over the last few decades, access to primary education has improved significantly in many parts of the world. The net enrolment ratios in primary education increased by at least 20 percentage points from 1999 to 2012 in 17 countries (UNESCO, 2015). However, many developing countries are still far from reaching the SDG 4 of *inclusive and equitable quality education for all*. There exist huge inequalities in educational opportunity, for instance, in countries such as Ghana, India, Mozambique, Nigeria, and Zambia, the poorest quintile accounts for 30% to 40% of the out-of-school population (UNESCO, 2009). Also, estimates from regional student assessments across the world reveal that even though children are enrolled in schools they have failed to acquire even the most basic literacy and numeracy skills.[1]. Therefore, it is important to identify policies to reduce high drop-out rates and improve learning in order to help children progress successfully through school irrespective of their socio-economic background.

Understanding the existence of a large population of less privileged children who are lagging behind their peers, this study looks at the impact of retention as a policy tool to enable students to progress through school. Proponents of retention believe that asking weak students to repeat will help them in master the skills which they were unable to learn during the failed year. Another reason advocated for retention is to improve monitoring and accountability in schools. Accordingly, a threat of repeating a grade due to failure in an exam is believed to motivate students and parents and send a strong signal to teachers to pay attention to low-achieving students.

What does the evidence say? Studies from developed countries have found retention to be an expensive and largely ineffective method to help weak students (Jimerson et al., 2002). Other studies that found a positive effect of retention report effect size that are either small in magnitude or fade-out over time (Nunes et al., 2018; Roderick and Nagaoka,

[1]Around 250 million children across the world are unable to acquire basic reading and mathematics skills and more than half of them have stayed in school for at least four years (Altinok et al., 2014; UNESCO, 2013)

2005). Bernard and Michaelowa (2006) provide evidence in a developing country context using a panel data of students in primary grades in Senegal. Grouping students according to initial achievements in national tests and controlling for family background they found an insignificant effect of repetition on the performance of weaker students and negative effects for those in the higher performance group. A recent study from India by Ahsan et al. (2018) which is very closely related to this study also focuses on the affect of the no-detention policy in India on learning outcomes of students. They find that removal of retention helped in improving learning outcomes of students. The authors suggest that the possible mechanism for improved learning levels could be due to increased motivation of low ability students caused by the removal of fear of failure in exams.

The above evidence broadly indicates that retention negatively affects a student's decision to continue education, however the endogenity of repeating a grade with motivation and student ability might lead to overestimation of the negative effect. For instance, students who lack motivation or are "less able" might not want to continue education beyond a certain level. These students are also more likely to fail in assessments and thus repeat the grade. If one does not account for this association between student's initial ability and retention one might exaggerate the effect of the policy on early school leaving rates. This paper provides additional evidence on the effect of retention in elementary schools on education outcomes like the decision to drop out of school and learning levels in a developing country context. The study also focuses on understanding the differential impact of retention depending on the socio-economic status of the children. In order to overcome the endogenity of retention with student motivation/ability, I make use of a large scale education reform in India: the no-detention policy of the Right to Education Act (RTE) 2010. I employ difference in difference method to estimate the effect of the no-detention policy on student drop-out rates and learning outcomes. Though the policy was implemented in all states across India at the same time, I exploit variation in exposure to the policy due to differences in pre-existing trends in repetition rates across districts.

Recently, the Indian parliament passed a bill that allows states to abolish the no-detention policy in schools. Accordingly, contrary to the earlier policy under RTE (2010), states in India can now ask a child to repeat a grade in case of failure in the grade-level exam. This move of removing the no-detention clause of RTE Act, 2010 has been criticised by educationalists as they fear it might lead to an increase in drop out rates and negatively affect children from disadvantaged backgrounds. It is therefore important to understand how the policy might affect education outcomes. Evidence from this study could help states in analyzing the affects of removing the no-detention policy.

The paper is divided into seven sections. In the following section I provide a brief descriptive analysis of retention in elementary schools in India before the RTE Act was implemented in 2010. In section three the no-detention policy of the RTE Act 2010 is discussed with details about the recent amendment. Section four and five provide details of the data and empirical strategy respectively. In section six I discuss the results followed by conclusion in the last section.

# 2 Retention in elementary schools before RTE (2010)

This section provides a brief descriptive analysis on retention in elementary schools in India before RTE Act was implemented in 2010. This will help us undestand the socio-economic implications of retention. I look at the following questions:

- What factors at the individual and household level affect the probability of repeating a grade? Do people from disadvantaged background are more likely to repeat grade?
- How does retention in early grades affect the likelihood of continuing education?

In this section I use data from a nation-wide representative panel survey called the Indian Human Development Survey (IHDS 2005,2011) to look at the repetition status of a child before the no-detention policy was introduced. IHDS collects information on various socio-economic indicators both for rural and urban households. It also collects data on educational

indicators of children in the age group 8-11[2]. These include basic reading and math scores which we take as a proxy for the child's initial ability, the number of grades completed, and a question on the child's repetition status[3]. It was found that in 2005, the average repetition rate for children in the age group 8-11 was around 10 percent. [4]

In table 1 I present results from linear probability regression models with repetition status (Repeat=Yes) and decision to drop out (Dropout=Yes) as the binary dependent variables in columns 1 and 2 respectively. The regression in column 1 shows various factors at the individual and household level which might be associated with repeating a grade. It was found that the probability that one was retained decreases (0.01 units) as the learning score in basic language and arithmetic test increases by one unit. Also, a child presently enrolled in a government school is more likely to have repeated a grade. This does not necessarily imply that government schools have higher repetition rates compared to private schools. It could also be the case that the child shifted from a private to a government school after she was asked to repeat a grade. Since I do not have information on the child's enrolment history and the grade in which she was asked to repeat I cannot be sure whether the child was retained in the same school in which she is currently enrolled or not. Though, the results indicate that there might exist more repeaters in a government school or in other words a larger proportion of disadvantaged students compared to private schools. Lastly, I find children from backward castes are slightly more likely to repeat a grade. This indicates that apart from student ability, the social status of the household might also influence the child's progress through school. Other variables like parents' education level, household assets, etc. were not found to be statistically significant.

I next look at the relationship between retention and the probability of dropping out

[2]This age group ideally corresponds to grades 3-6.

[3]Children were asked if they have ever repeated a grade in school and how many times. Unfortunately, we do not know the age or grade when the child was asked to repeat a grade.

[4]Though NSSO also collects information on repetition in their education surveys, we use IHDS data for two reasons *first* it is a panel data which enables us to follow the same child over two time periods *second* IHDS collects information on child's learning level. Understanding the strong association between child's ability and detention it is important to account for the initial differences in learning.

of school. From IHDS round two data (2011) I can identify the current education status of children (now aged 14-17) who were surveyed in round one (age 8-11). Using the initial learning scores and other socio-economic indicators form round one data I want to analyse how a child who was retained in primary grades in 2005 progressed through school. The panel data enables us to follow the child and check his educational status in 2011 (six years later). The dependent variable, *Dropout*, takes value 1 if a child who was enrolled in school in round one has dropped out by the end of round two. Column 2 in table 1 presents the results from a linear regression. It was found that children who repeated a grade when they were in primary school (Repeat=Yes) are 5 percent more likely to dropout. Not surprisingly, I also find initial learning levels are important predictors of future success in school. It is important to note that private tuition and other household-level socio-economic indicators are highly correlated with the decision to continue education. This indicates that educational opportunities are highly dependent on one's family background and more needs to be done to make education accessible for all.

# 3 Right to Education and no-detention policy, 2010

Section 16 of the Right of Children to Free and Compulsory Education (RTE) Act, 2010 stipulates that 'No child admitted in a school shall be held back in any class or expelled from school till the completion of elementary education'. The elementary stage of schooling covers grades 1 to 8. This was considered an important feature of the act to remove the fear of failure and reduce drop out rates in elementary education. Also, instead of traditional exams the RTE aimed at using Comprehensive and Continuous Evaluation (CCE) for evaluating children to improve their learning. One of the most common misconception about the no-detention policy is that it is perceived to be equivalent to having no exams or any other forms of assessments to evaluate children. Many state representatives have expressed concerns about the failure to implement CCE leading to subsequent discontinuation

of any evaluation. The report of the CABE[5] sub-committee on CEE and no-detention acknowledged the reluctance of various stakeholders in continuing with the no detention policy (Bhukkal, 2015). The report states *"At this stage, it would be prudent to re-iterate the need for assessment of learning outcomes and make it consequential by linking it to promotion or otherwise to the next class beyond Grade V."*

Some of the recommendations suggested by the sub committee are (PIB, 2017).

- There should be an examination at Class 5. It should be left to the States and UTs to decide whether this exam will be at the school, block, district or state Level.
- If a child fails then allow the child an opportunity to improve. There should be additional instruction provided to children and the child should be given an opportunity to sit for another exam. If the child is unable to pass the exam in the second chance, then detain the child.

This move of removing the no-detention clause of RTE has been criticised by educationalist as they fear it might lead to increase in drop out rates and negatively effect children from disadvantaged backgrounds. It is therefore important to analyse the affects of retention in order to understand how the proposed policy might influence student outcomes.

# 4 Data

For the main analysis, I combine Annual Status of Education Report (ASER, 2007-14) and District Information on Systems in Education (DISE) data. ASER is a nation-wide household level annual survey conducted by the NGO Pratham since 2005. It collects information on educational achievement of children aged 5-16 from every rural district in the country. Since the survey is conducted in the household it includes both enrolled and out of school children. Apart from basic reading and arithmetic tests, the survey also provides information on

[5]The Central Advisory Board on Education (CABE) is the highest advisory board at the central and the state level in the field of education.

individual's education status (dropout, grade enrolled, school type etc.) and other household socio-economic indicators . DISE is a school based annual survey with details on various school level indicators of elementary education (enrolment rate, repetition rate etc.). For my analysis I merge aggregate district level DISE with ASER data to create a repeated cross section of individuals in the age group 12-14 over the years 2007-14.

Table 2 gives a brief description of the dataset. The target group is children in the age group 12-14. The outcome variables include *Drop out*: whether or not the child is enrolled in school, and *Learning level*: a combined score of basic language and arithmetic levels. ASER administers an oral language test with five levels: read nothing (0), read letter (1), read word (2), read para (3), and read story (4). The child is marked on the highest reading level. For arithmetic the levels are: cannot identify single digits (0), identify single digit numbers (1), identify double digit numbers (2), can subtract (3), and can divide (4). I also create a pre-post dummy variable called*Post* which takes value zero for year before RTE (2007-2009) and one from the year RTE was implemented (2010-2014). Lastly, the variable *Rep. intensity cat.* is generated by arranging district level average repetition rates in elementary grades in the year 2009 (*Rep. rate 2009*) from lowest to highest and grouping them into four quartiles. This variable is generated to compare the variation in treatment effects by intensity category.

# 5 Empirical strategy

In order to estimate the effect of the policy on educational outcomes using the difference in difference methods we not only need exogenous time variation (before and after 2010) but also cross-section variation across groups (for instance regions where the policy was implemented and regions where it was not implemented). Since the no-detention policy was rolled out across the whole country in the year 2010 I exploit cross-section differences across districts caused by the variation in the average repetition rates before the implementation

of the policy. Those districts with higher average repetition rates in the year 2009 would be affected more by the no-detention policy compared to those with lower repetition rates. Thus, I exploit the variation in treatment intensity as measured by average district-level repetition rates in the year 2009 to compare changes in outcomes before and after the policy. It is true that repetition rates prior to the implementation of the policy are not randomly assigned. Differences in repetition rates across districts might be associated with the quality of education in the region or other socio-economic factors which can bias our estimates. The empirical strategy is, therefore, to look at whether there is a break in any pre-existing differences in the trend of the outcomes when the policy was implemented in the year 2010. I estimate the following equations:

$$y_{idst} = \sum_{t=2007}^{t=2014} \alpha_t Year + \beta Reprate_d + \sum_{t=2007}^{t=2014} \delta_t Year_t * Reprate_d + \gamma_1 X_i + \gamma_2 X_d + \phi_s + \epsilon_{idst} \quad (1)$$

where $y_{idst}$ is the outcome variable for individual $i$ in district $d$, state $s$ and, year $t$. I include year $Year$ and state fixed effects $\phi_s$. $X_i$, $X_d$, refer to individual/family, and district level co-variates respectively. $\epsilon_{idst}$ is the error term. In equation 1 The term $Year_t * Reprate_d$ is an interaction of years with average repetition rate of the district in the year 2009. This variable ($Reprate_d$) gives us the treatment intensity. The coefficients of this interaction term denoted by $\delta_t$ gives us the difference in trends in outcome variables due to treatment intensity and is our main variable of interest.

I also present results from equation 2 where I interact the continuous intensity variable ($Reprate_d$) with a dummy variable $Post$ grouping all the years into before ($< 2010$,0) and after the policy was implemented ($>= 2010$,1).

$$y_{idst} = \alpha Post_t + \beta Reprate_d + \delta Post_t * Reprate_d + \gamma_1 X_i + \gamma_2 X_d + \phi_s + \epsilon_{idst} \quad (2)$$

# 6 Results and discussion

In figure 1 I plot the average repetition rate for the whole country. It shows that the repetition rates had been falling even before RTE was implemented in 2010 [6]. For our analysis, it is important to see if the decline varies by treatment intensity before and after the policy. To check this I regress average repetition rate for the years 2007-2014 on the interaction of year with repetition intensity category variable (Rep. intensity cat.). I also include other district level co-variates. In figure 2 I plot the average marginal effect on repetition rate by repetition intensity categories for each year. It shows the trends of the difference in repetition rate before and after the policy depending on the intensity category. The base category is the 1st quartile and each line represents different intensity category (2nd 3rd and 4th quartile respectively). One can notice that the significant differences between repetition rate categories before the policy disappears by the year 2011 (one year after the policy) and eventually converges to zero. Also, the fall is steepest for the 4th quartile category (high intensity). Thus, no-detention policy leads to a fall in repetition rates (treatment) and the decline was larger for districts with high existing repetition rates (treatment intensity).

In figure 3 and 4 I plot the results from 1 for our outcome variables: probability to drop out and learning scores respectively. The graphs plot the trends in difference by treatment intensity $\delta_t$ from year 2008-2014. In figure 3 before 2010 there was no significant difference in dropout rate by treatment intensity but since 2010 until 2014 there is a significant difference in droprate. The negative $\delta_t$ implies that the fall in droprate was much higher for high intensity districts compared to low intensity district. Thus, the no-detention policy in 2010 had a significant negative effect on drop-out rate. To understand the magnitude of the effect one can look at the results from equation 2 in Table 3. In column one the dependent variable drop-out is an indicator variable for whether the child is enrolled in school or not. The linear probability model gives us the probability whether the child is in school or not. The variable $Post_t * Reprate_d$ is the main variable of interest. It shows that the probability that one is

[6] Except for a small spike in repetition rates in the year 2011.

not enrolled in school dropped by 0.2 percent after the policy for those regions that were affected the most.

Coming next to the effect on learning outcomes once again I plot the $\delta_t$ from equation 1 in figure 4. We can now see that until 2010 there was no significant difference in learning scores by treatment intensity. Although, from 2011 until 2014 one can notice a significant difference in learning outcomes. Once again the negative $\delta_t$ implies that the fall in learning levels was much higher for high intensity districts compared to low intensity districts. Thus, the no-detention policy in 2010 had a significant negative effect on learning outcomes. To understand the magnitude of the effect one can look at the results from equation 2 in Table 3. In column 2 the dependent variable learn is a score ranging from 0 to 8. The variable $Post_t * Reprate_d$ is the main variable of interest. It shows the learning scores dropped by -0.027 points after the policy for those regions that were affected the most.

In Table 3 I also present the results from learning outcomes depending on the school the child is enrolled in (columns 3 and 4). It can be seen that the negative effect on learning outcomes is significant for government schools but not for private schools.

## 6.1 Heterogeneous effects

In Tables 4 and 5 in further explore whether the affect of no-detention policy varies by individual's economic status and gender. I use variable $HH.Poor$ as a proxy for the economic status of the household. It is generated from the information on the type of the house collected during the ASER survey. It is a categorical variable which states whether the household lives in a katcha, semi katcha or pukka house. I use this indicator as a continuous variable with a high score representing that the household is poorer (0=pukka,1=semi kutcha and 2=kutcha). The coefficient of the interaction variable $Year_t * Reprate_d * HH.poor_i$ can be seen in table Table 4. I shows that the fall in droprate was higher for poor households compared to rich households (column 1). Also, the fall in learning level was higher for poor and unlike the overall effect the negative learning outcomes are also seen in case of private

schools.

Table 5 reports the heterogenous effects by gender. We dont see any statistically significant difference for any of the outcome variables.

# 7 Conclusion

The improvement in access to primary education across many developing countries has met with serious challenges to provide inclusive, quality education for all. Recent years have witnessed a shift in the focus of education policy towards improving the quality of education. For instance, the Indian parliament recently passed a bill that allows states to abolish the no-detention policy in schools in a bid to improve declining quality of education. It is not yet clear though how this policy might affect educational opportunity and learning outcomes. This study thereby provides evidence on the effect of retention on educational outcomes in post-primary education.

In order to causally estimate the effect of retaining students on their educational outcomes I use a large scale education reform in India, the Right to Education Act (2010) no-detention policy, that abolished the practice of retention for all elementary grades across the country. The difference in difference method allows me to estimate the effect of the policy on student drop-out rates and learning outcomes. I exploit variation in exposure to the no-detention policy due to differences in pre-existing trends in repetition rates across districts.

As intended no-detention policy helped in reducing drop out rates in elementary education especially for children from an economically poor background. I find that the policy helped in reducing the drop-out rates in post-primary education by 0.2 percent. Thus, the no-detention policy helped in reducing inequality in educational opportunities.

Next, looking at the effect of the policy on learning outcomes it was found one year following the policy there was a significant decline in the learning levels of children in the post-primary age group by -0.027 points. Also, the decline was relatively larger for economically

disadvantaged children compared to the rich. The probable mechanism for the negative effect of the could be (a) less motivated student and teachers because of removal of high stake exams (a) congestion or (b) inability to implement the new form of Continuous Comprehensive Evaluations (CCE) to assess student learning instead of regular exams. Further research is needed to understand the true mechanism behind the fall in the quality of education after the implementation of the policy.

# References


Ahsan, Md Nazmul, Rakesh Banerjee, and Siddharth Hari (2018). "Social Promotion and Learning Outcomes: Evidence from India". In: *USC-INET Research Paper* 18-14.

Altinok, Nadir, Claude Diebolt, and Jean-Luc Demeulemeester (2014). "A new international database on education quality: 1965–2010". In: *Applied Economics* 46.11, pp. 1212–1247.

Bernard, Jean-Marc and Katharina Michaelowa (2006). "Managing the Impact of PASEC Projects in Francophone Sub-Saharan Africa". In: *Cross-National Studies of the Quality of Education: Planning Their Design and Managing Their Impact*. Ed. by Kenneth N Ross and Ilona Jurgens Genevois. Paris: IIEP, UNESCO. Chap. 11, pp. 229–240.

Bhukkal, Geeta (2015). *Report of CABE Sub-Committee on Assessment and Implementation of CCE*. Ministry of Human Resource and Development.

Jimerson, Shane R, Gabrielle E Anderson, and Angela D Whipple (2002). "Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school". In: *Psychology in the Schools* 39.4, pp. 441–457.

Nunes, Luis Catela, Ana Balcão Reis, and Carmo Seabra (2018). "Is retention beneficial to low-achieving students? Evidence from Portugal". In: *Applied Economics*, pp. 1–12.

PIB (2017). *Press Information Bureau,Government of India, Ministry of Human Resource Development:Change in No-Detention Policy*. `http://pib.nic.in/newsite/PrintRelease.aspx?relid=159006`. Accessed: 2019-03-17.

Roderick, Melissa and Jenny Nagaoka (2005). "Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless?" In: *Educational evaluation and policy analysis* 27.4, pp. 309–340.

UNESCO (2009). *Global Monitoring Report 2009: Overcoming inequality: why governance matters*. UNESCO, Paris.

— (2013). *Education for All Global Monitoring Report 2013/4 Teaching and Learning: Achieving Quality for All*. UNESCO, Paris.

— (2015). *Global Monitoring Report 2015: Education for all 2000–2015: Achievements and Challenges*. UNESCO, Paris.

# A Tables

**Table 1** Linear probabilty model for children in age group 8-11 in IHDS 2005

| | (1) Y=Repeat | (2) Y=Drop out |
|---|---|---|
| Learning level | -0.010** | -0.035*** |
| | (0.004) | (0.004) |
| Female | -0.008 | 0.015 |
| | (0.009) | (0.014) |
| Govt school=Yes | 0.032*** | 0.012 |
| | (0.012) | (0.014) |
| Tution=Yes | 0.007 | -0.044*** |
| | (0.014) | (0.017) |
| Age when enroled | -0.001 | 0.014*** |
| | (0.004) | (0.004) |
| HH assets | -0.002 | -0.010*** |
| | (0.001) | (0.001) |
| Mother edu | -0.002 | -0.007*** |
| | (0.001) | (0.002) |
| HH size | 0.001 | -0.003 |
| | (0.001) | (0.002) |
| Muslim=Yes | -0.002 | 0.187*** |
| | (0.017) | (0.028) |
| Backward caste=Yes | 0.027* | 0.001 |
| | (0.014) | (0.017) |
| Repeat=yes | | 0.053** |
| | | (0.022) |
| Grade Fixed Effect | Yes | Yes |
| District Fixed Effect | Yes | Yes |
| $R^2$ | 0.017 | 0.124 |
| No. of Obs. | 4035 | 4035 |
| No. of clusters | 355 | 355 |

*Notes*: We use grade fixed effect and district-fixed effects. Standard errors are clustered at district level. *p<0.1;**p<0.05;***p<0.01

**Table 2** Descriptive statistics I- Merged ASER and DISE data

| | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Individual and household level variable* | | | | | |
| Drop out | 1,160,952 | 0.04 | 0.20 | 0 | 1 |
| Grade enrolled | 1,037,069 | 7.20 | 1.40 | 4 | 10 |
| Learning level[a] | 1,061,366 | 6.60 | 1.93 | 0 | 8 |
| School=govt. | 1,080,177 | 0.73 | 0.45 | 0 | 1 |
| Gender=female | 1,150,447 | 0.47 | 0.50 | 0 | 1 |
| Age | 1,160,952 | 12.91 | 0.83 | 12 | 14 |
| HH. size | 1,145,214 | 6.52 | 3.09 | 1 | 99 |
| Mother school | 1,121,492 | 0.47 | 0.50 | 0 | 1 |
| HH. type | 991,298 | 1.00 | 0.83 | 0 | 2 |
| *District level variables* | | | | | |
| Year | 1,160,952 | 2010.33 | 2.29 | 2007 | 2014 |
| Post[b] | 1,160,952 | 0.59 | 0.49 | 0 | 1 |
| Rep. rate 2009 | 1,157,488 | 3.98 | 3.52 | 0.00 | 22.77 |
| Rep. intensity cat.[c] | 1,157,488 | 2.50 | 1.12 | 1 | 4 |
| Log Total schools | 1,160,952 | 7.50 | 0.76 | 3.33 | 9.17 |
| Log Population | 1,150,882 | 14.12 | 0.92 | 8.98 | 16.22 |
| Log Enrolment prev. yr | 1,160,952 | 10.17 | 0.89 | 4.77 | 12.20 |
| Log Avg. Rep. rate | 1,116,736 | 0.42 | 1.56 | −6.21 | 4.27 |
| Female Lit. rate | 1,150,882 | 54.68 | 15.05 | 18.60 | 98.28 |
| Years | 8 | | | | |
| States | 31 | | | | |
| Districts | 567 | | | | |
| N | 1,160,952 | | | | |

*Notes*: This table gives the summary statistics of ASER (age 12-14) and DISE (grades 1-8) merged data from 2007-14.

[a] Learning level is generated by combining basic reading (0-4) and arithmetic levels (0-4).

[b] Post is a dummy variable which takes value 0 for year before RTE (2007-2009) and 1 from the year RTE was implemented (2010-2014)

[c] Rep. intensity cat. is generated by arranging district level average repetition rates in elementary education in the year 2009 from lowest to highest and grouping them into four quartiles.

**Table 3** Difference in difference results for Upper-Primary Age (12-14)

| | (1) Droprate | (2) Learn | (3) Learn Govt | (4) Learn Pvt |
|---|---|---|---|---|
| Post | -0.006*** | -0.258*** | -0.300*** | -0.212*** |
| | (0.001) | (0.030) | (0.033) | (0.026) |
| Reprate | 0.002*** | -0.009 | -0.001 | -0.004 |
| | (0.001) | (0.008) | (0.007) | (0.008) |
| Post*Reprate | -0.002*** | -0.027*** | -0.031*** | -0.009 |
| | (0.000) | (0.007) | (0.007) | (0.006) |
| Learn level | -0.023*** | | | |
| | (0.001) | | | |
| Gender | 0.006*** | -0.165*** | -0.101*** | -0.051*** |
| | (0.001) | (0.012) | (0.010) | (0.011) |
| Age | 0.022*** | 0.268*** | 0.331*** | 0.236*** |
| | (0.001) | (0.004) | (0.005) | (0.006) |
| HH size | -0.000 | -0.001 | -0.002 | 0.001 |
| | (0.000) | (0.002) | (0.002) | (0.002) |
| Mother edu | -0.018*** | 0.708*** | 0.536*** | 0.426*** |
| | (0.001) | (0.015) | (0.012) | (0.014) |
| Log tot sch. | 0.002 | 0.050 | 0.134** | -0.084 |
| | (0.003) | (0.065) | (0.062) | (0.068) |
| Log tot pop. | 0.008*** | -0.034 | 0.015 | 0.055 |
| | (0.003) | (0.064) | (0.058) | (0.063) |
| Log enrol. | -0.009*** | -0.065 | -0.193*** | -0.030 |
| | (0.003) | (0.062) | (0.057) | (0.055) |
| Female Lit | -0.000*** | 0.008*** | 0.001 | 0.002 |
| | (0.000) | (0.001) | (0.001) | (0.001) |
| Log reprate | 0.001*** | 0.031*** | 0.036*** | 0.025*** |
| | (0.000) | (0.008) | (0.008) | (0.008) |
| State Fixed Effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.080 | 0.096 | 0.101 | 0.081 |
| No. of Obs. | 954655 | 954655 | 658627 | 244110 |
| No. of clusters | 565 | 565 | 565 | 565 |

*Notes*: We use cluster robust standard errors at district level. **p<0.05;***p<0.01

**Table 4** Heterogenous effect-HHpoor for Upper-Primary Age (12-14)

| | (1) Dropout | (2) Learn | (3) Learn Govt | (4) Learn Pvt |
|---|---|---|---|---|
| Post | -0.001 | -0.214*** | -0.247*** | -0.163*** |
| | (0.001) | (0.032) | (0.037) | (0.030) |
| Reprate | 0.001*** | -0.027*** | -0.022*** | -0.013 |
| | (0.000) | (0.008) | (0.008) | (0.008) |
| Post*Reprate | -0.001*** | -0.000 | 0.001 | -0.003 |
| | (0.000) | (0.006) | (0.007) | (0.006) |
| HH poor | 0.007*** | -0.283*** | -0.207*** | -0.177*** |
| | (0.001) | (0.021) | (0.021) | (0.020) |
| Post*Hhpoor | -0.004*** | -0.039 | -0.044 | -0.019 |
| | (0.001) | (0.022) | (0.025) | (0.023) |
| HHpoor*Reprate | 0.000 | 0.015*** | 0.016*** | 0.013*** |
| | (0.000) | (0.004) | (0.004) | (0.005) |
| Post*Reprate*HHpoor | -0.001** | -0.019*** | -0.022*** | -0.016*** |
| | (0.000) | (0.005) | (0.005) | (0.005) |
| Learn level | -0.022*** | | | |
| | (0.001) | | | |
| Gender | 0.006*** | -0.169*** | -0.109*** | -0.056*** |
| | (0.001) | (0.012) | (0.011) | (0.012) |
| Age | 0.021*** | 0.272*** | 0.334*** | 0.243*** |
| | (0.001) | (0.004) | (0.006) | (0.006) |
| HH size | -0.000 | -0.005*** | -0.007*** | -0.001 |
| | (0.000) | (0.002) | (0.002) | (0.002) |
| Mother edu | -0.015*** | 0.651*** | 0.512*** | 0.410*** |
| | (0.001) | (0.014) | (0.012) | (0.014) |
| State Fixed Effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.080 | 0.114 | 0.113 | 0.093 |
| No. of Obs. | 807785 | 807785 | 554979 | 210860 |
| No. of clusters | 565 | 565 | 565 | 565 |

*Notes*: We use cluster robust standard errors at district level. $^{**}p<0.05;^{***}p<0.01$

**Table 5** Heterogenous effect-Gender for Upper-Primary Age (12-14)

| | (1) Dropout | (2) Learn | (3) Learn Govt | (4) Learn Pvt |
|---|---|---|---|---|
| Post | -0.006*** | -0.234*** | -0.272*** | -0.202*** |
| | (0.001) | (0.029) | (0.033) | (0.026) |
| Reprate | 0.002*** | -0.013 | -0.004 | -0.009 |
| | (0.000) | (0.008) | (0.007) | (0.008) |
| Post*Reprate | -0.002*** | -0.029*** | -0.033*** | -0.012** |
| | (0.000) | (0.006) | (0.007) | (0.006) |
| Female | 0.007*** | -0.186*** | -0.104*** | -0.093*** |
| | (0.001) | (0.018) | (0.016) | (0.021) |
| Post*Female | -0.000 | -0.049*** | -0.055*** | -0.023 |
| | (0.001) | (0.015) | (0.017) | (0.021) |
| Female*Reprate | 0.000 | 0.010*** | 0.006** | 0.013*** |
| | (0.000) | (0.003) | (0.003) | (0.004) |
| Post*Reprate*Female | -0.000 | 0.004 | 0.003 | 0.008 |
| | (0.000) | (0.003) | (0.003) | (0.004) |
| Learn level | -0.023*** | | | |
| | (0.001) | | | |
| Age | 0.022*** | 0.268*** | 0.331*** | 0.237*** |
| | (0.001) | (0.004) | (0.005) | (0.006) |
| HH size | -0.000 | -0.001 | -0.002 | 0.001 |
| | (0.000) | (0.002) | (0.002) | (0.002) |
| Mother edu | -0.018*** | 0.708*** | 0.536*** | 0.426*** |
| | (0.001) | (0.015) | (0.012) | (0.014) |
| State Fixed Effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.080 | 0.096 | 0.102 | 0.082 |
| No. of Obs. | 954655 | 954655 | 658627 | 244110 |
| No. of clusters | 565 | 565 | 565 | 565 |

*Notes*: We use cluster robust standard errors at district level. **p<0.05;***p<0.01

# B Figures

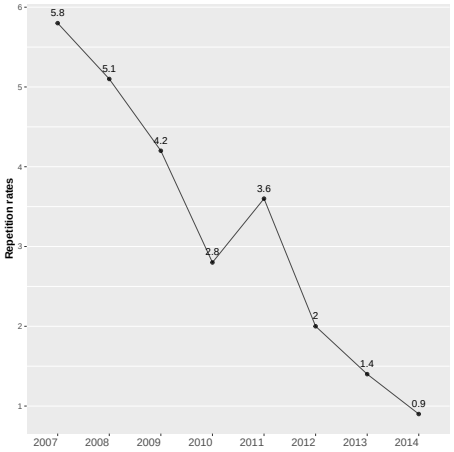


**Figure 1** Trends in average repetition rate from 2007-2014.

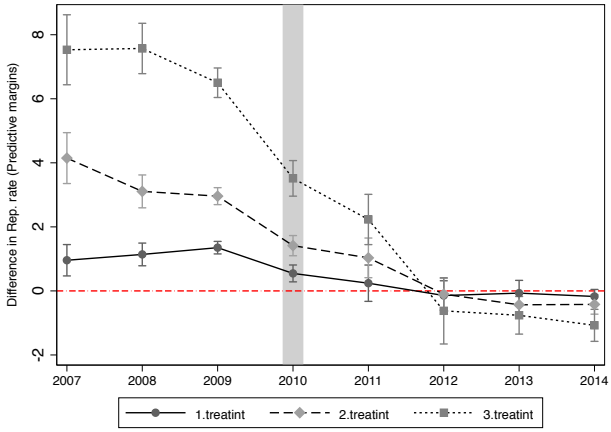


**Figure 2** Trends in Difference in log repetition rate by treatment intensity: The figure compares the average marginal effect on log repetition rate for each year (2008-2014) between low and high repetition intensity districts. The base category is the 1st quartile and each line represents different intensity category (2nd 3rd and 4th quartile respectively). The shaded region is the year when no-detention policys was implemented.

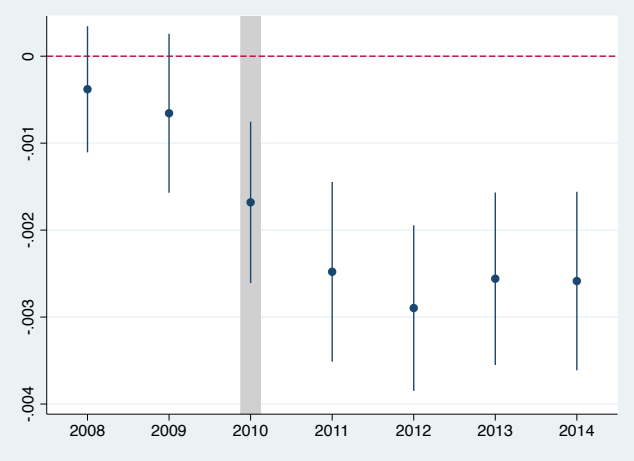


**Figure 3** Coef plot ($\delta_t$) for Dropout rate. The shaded region is the year when RTE was implemented.

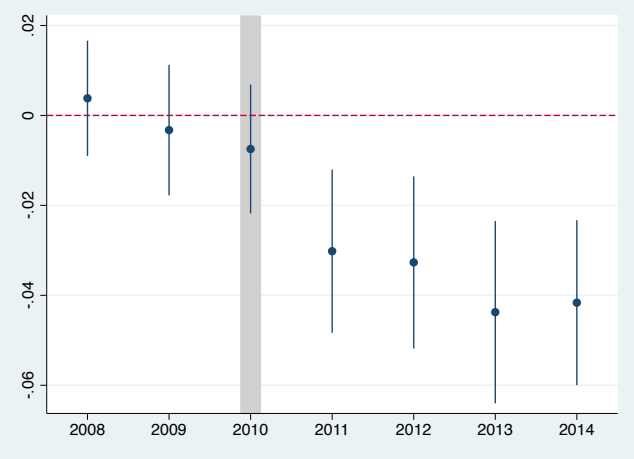


**Figure 4** Coef plot ($\delta_t$) for Learning level. The shaded region is the year when RTE was implemented.