Tenure Reform, Selection and Student Achievement in Developing Countries: Evidence from Peru

Wilfried Adohinzin

August 2023

1 Motivation and research questions

A large literature including papers such as Rivkin, Hanushek, and Kain (2005), Rockoff (2004), Araujo, Carneiro, Cruz-Aguayo, and Schady (2016) and Chetty, Friedman, and Rockoff (2014) documents significant returns to having a good teacher both in childhood and in adulthood. While this is a universally acknowledged fact, many countries, especially in the developing world still struggle to attract and retain high quality teachers (Chelwa, Pellicer, & Maboshe, 2019). In Sub-Saharan Africa, Mulkeen, Chapman, and DeJaeghere (2005) report attrition rates between 5 and 30 percent of teachers.

In recent years, many policy interventions have been implemented with the goal to improve teacher quality and reduce turnover in developing countries. These interventions include teacher training (Aber et al., 2017; Lucas & Mbiti, 2014); teacher performance pay (Bobba, Ederer, Leon-Ciliotta, Neilson, & Nieddu, 2021; Chelwa et al., 2019; Glewwe, Ilias, & Kremer, 2010; Pugatch & Schroeder, 2018); teacher monitoring (Duflo, Hanna, & Ryan, 2012). However, little is known about whether the offer of job protection to teachers could attract and retain high quality teachers, improve the quality of teachers and translate into improved learning for students (Glewwe, Shen, Sun, & Wisniewski, 2020)[1].

In this paper, I combine rich administrative dataset on teacher hiring applications merged with classroom level information on students and a regression discontinuity design to explore the impact of a recent teacher reform in Peru that granted labor protection to teachers who passed a national

---

[1] I use the terms tenured and permanent interchangeably throughout the paper.

contest and a local evaluation. Specifically, my paper first explores the impact of this reform on the selection of teachers into and out of the teaching labor force; and second, I analyze whether teacher effort responds to the receipt of tenure.

A reform that grants tenure to teachers on the basis of their merit can affect student learning through two major pathways. On the one hand, the offer of tenure can have effects on the selection of teachers into and out of the teaching workforce. (Ng, 2021; Strunk, Barrett, & Lincove, 2017). If high quality teachers perceive tenure as a positive job amenity, they are more likely to remain within the teaching force when granted job security. On the other hand, if denied, as they face good outside options, they are more likely to exit the public teaching force. Similarly, low value teachers (who might not have remained in the teaching force in the absence of tenure) might be more likely to remain in the teaching force after the offer of tenure (Hanushek, 2016). The second major pathway through which job protection can have effects on students is through teacher effort. The offer of tenure can increase the attractiveness of the teaching profession and reduce the stress associated with the threat of being fired. This could therefore have a positive effect on student outcomes. However, job protection may also be associated to a reduction in accountability which would translate into negative outcomes for students.

2   Background and empirical strategy

In Peru, like most developing countries, teachers are commonly hired under two types of contracts: temporary (docentes contratados) and permanent contracts (docentes nombrados). Permanent teachers are hired for an indefinite period while temporary teachers are hired for a fixed period (typically a year) and their renewal is contingent upon achieving satisfactory results (Bobba et al., 2021). Permanent can also advance in the career ladder faster through periodical evaluations while permanent are stuck on an entry level wage (Bertoni, Elacqua, Méndez, & Santos, 2020).[2] Permanent and temporary teachers also differ on several observable dimensions. On average, permanent teachers are more likely to be, younger and to have a university degree instead of an institute

---

[2] Entry level teachers in 2015 earned an average RIM of US $510 (Bobba et al., 2021) Though all newly recruited teachers receive the same basic wage irrespective of their contract status, permanent teachers can climb the ladder when they fulfill a certain number of requirements. However, bonuses and other incentives are in place for any type of teachers who are willing to work for example in remote areas or in multigrade schools for example (Bertoni, Elacqua, Hincapié, Méndez, & Paredese, 2019).

degree. Table 1 summarizes these differences between permanent and temporary teachers.

In 2012, Peru adopted the *Ley de Reforma Magisterial (LRM)*, a teacher reform that changed the mode of hiring of permanent teachers in Peruvian public schools.[3] As part of the reform, entering the profession as a permanent teacher is based on passing a national (centralized) exam *Prueba Única Nacional (PUN)* and a local (decentralized) evaluation.[4] The national written exam is composed of three subjects: logical reasoning, reading comprehension and a curricular and pedagogical knowledge. Applicants are required to score at least a 60 percent grade in each subject to be eligible for the decentralized stage. In the second stage, eligible applicants submit their ranked preferences among the available vacancies.[5] The Ministry of Education uses these preferences and the PUN scores to assign candidates to their different schools. Selected candidates are invited to the local evaluation which consists of an assessment of the applicant's resume, an interview, and an in-classroom demonstration. A 60 percent grade is also required for the local evaluation to pass the local stage. Finally, scores from both the national phase and the local evaluation are weighted and summed up and positions are awarded to the highest scoring applicants. Applicants who fail at any step of the process can enter a third phase of the process where they compete for unfilled vacancies and temporary jobs. Figure 1 gives an overview of the hiring process.

In this paper, I explore the impact of receiving a tenured position on teacher outcomes. Since receiving tenure is endogenous, I exploit the scores obtained by the applicant in the national exam in a regression discontinuity design. Essentially, the empirical strategy can be summarized by a comparison of applicants who scored just above the thresholds and end up with permanent positions to applicants who scored just below and end up with temporary positions. While the likelihood of getting a tenured position should be 0 for teachers who fail the national test, this probability is likely not 1 for applicants who succeed. Generally, applicants with scores below the threshold are certain not to be awarded a permanent position. However, applicants with scores just above the threshold who pass the local stage get a permanent contract. This warrants a fuzzy regression discontinuity design, akin to estimating local average treatment effect (LATE) with an IV (Lee & Lemieux, 2010).

---

[3] Prior to 2012, the hiring of teachers was done by local jurisdictions

[4] While the reform was implemented in 2012, the first evaluation was done is 2015 with 77,598 applicants for nearly 19,000 job positions in primary schools

[5] During the 2015 hiring process, a maximum of five preferences were allowed. However, this was changed to an unlimited number starting from 2017

I estimate the following empirical specification:

$$Y_{jgst} = \theta_{gs} + \mu_t + \beta D_{jgst} + f(z_j) + \delta X_{jgst} + \varepsilon_{jgst} \tag{1}$$

Where D is a dummy variable equal to 1 if teacher j in grade g, school s in year t has a permanent contract; $f(.)$ is a smooth function of the distance between the score of applicant j and the cutoff faced; Y represents some outcome for teacher j. z_j is the running variable. X is a vector of teacher level covariates that include age, gender, disability status, teacher maternal language fixed effects, whether the teacher has a university degree and experience dummies.

There are generally at least two caveats to address regarding this identification strategy. The first is that the RD design in this analysis is based on three running variables. Since applicants are required to pass *each* of the three subsections of the PUN test, this is a case of multiple assignment regression discontinuity (MRD). Intuitively, my identification strategy will consider three different sets of compliers. Each of these sets will include individuals who passed two of the subsections and I will be comparing applicants who barely failed the third test to applicants who barely passed it. For example, one set of compliers will include only applicants who passed both the logical reasoning and the reading comprehension assignments. Within that sample, I will be comparing applicants just above to applicants just below the 60% threshold for curricular and pedagogical knowledge. The intuition in this setting is that conditioning on passing two sections of the test, applicants just below and applicants just above the threshold for the third section are similar on both observable and unobservable determinants of teacher effectiveness.

The second caveat is typical of any regression discontinuity design that uses scores to assign treatment to units. If there is evidence of manipulation either by applicants, or administrators, my estimates are likely to be biased. To deal with these issues, I implement a more recent iteration of the standard McCrary (2008) test. More specifically, I implement manipulation tests proposed by Cattaneo, Jansson, and Ma (2017). I present results of these in Figure 3,. In each of these figures, I plot histograms of scores on subsample of individuals who passed the other two tests. For example, the top left figure plots a histogram of logical reasoning scores for a sample of applicants who passed both the reading comprehension and the curricular knowledge subjects. On top of each figure, I add the p-value from the Cattaneo et al. (2017) describing whether the discontinuity at the

cutoff is statistically significant. Examining these histograms, we see p-values of the manipulation test are all above 10 percent except for the logical reasoning sample.[6]

To measure teacher productivity, I build value added measures for all teachers in the main sample. I follow Bau and Das (2020) and estimate the following lagged score model separately for math and language:

$$Y_{ijs,t} = \beta_0 + \beta_1 Z_{ijs,t} + \beta_2 X_{ijs,t} + \mu_{j,t} + \varepsilon_{ics,t} \tag{2}$$

In this specification, $Y_{ics,t}$ is the math or reading score of student i taught by teacher j; $Z_{ics,t}$ is a vector that includes prior grades in math and language score; $X_{ijs,t}$ indicates a vector of student/background characteristics that include age, gender, whether the student has repeated a class before, a socioeconomic index, dummies for the level of education of the mother and dummies for maternal language; $\mu_{js,t}$ is the effect of being assigned to teacher j in in school s in year t; these include implicitly year effects; and $\varepsilon_{ics,t}$ represents the error term. I follow Biasi (2021) by defining a time-varying value added that allows me to assess changes in teachers productivity, but I also use pre-tenure value added to analyze selection effects of tenure. For example, for teachers who took the test in 2015, I use their value added in earlier years as a measure of pre-tenure value added. By distinguishing between pre-tenure value added and post tenure value added, I can separate pure composition effects from changes in teacher productivity. Before estimating our value added models, I normalize the test scores from each exam so they have mean zero and standard deviation one.

3   Data

In this paper, I merge several administrative different datasets obtained from the Peruvian Ministry of Education over the period 2012 to 2019. The first data source is the teacher job application dataset for the years 2015, 2017 and 2018. The data include teacher-level information for both the centralized exam stage and the local evaluation stage as well as background demographic details on teachers. The second dataset includes complete records for all teachers employed in the public sector over the 2012-2019 period. These records include the school where the teacher works, the

---

[6] The logical reasoning portion of the test is the easiest to crack. In fact, 71 percent of all applicants pass it while less than one out of three applicants pass reading comprehension or curricular knowledge.

type of contract; the number of hours taught, the workday, the salary level. The third data source is a student-level dataset and includes information at the year-school-classroom-subject level. The fourth data source includes standardized mathematics and reading results at national tests for select grades during select years. The dataset also comes with rich family background information collected through a questionnaire addressed to parents and students in some cases.[7] The fifth data source is the School Census (Census Educativo) from 2012 to 2019. This dataset collects school level information on various school characteristics and amenities.[8]

4 Preliminary results and next steps

4.1 First stage effects: Does passing the national exam affect the likelihood of receiving a permanent position?

In order to assess whether the teacher evaluation scores are able to predict assignment to the treatment near the cutoff, I plot the probability of receiving tenure as a function of my running variables and for each set of compliers. More specifically, I estimate Equation 1, where the outcome variable is a binary variable for whether applicant j has a permanent position in the year after they took the test. Figure 4 present the results of these estimations. Across all three figures, it is clearly noticeable that the probability of receiving tenure jumps discontinuously at the cutoff. Estimates are all statistically significant at the 1 percent level and applicants who barely cross the threshold are 40 to 49 percentage points more likely to get a tenured job. In the next steps, I will focus on analyzing reduced form and second stage impacts on teacher outcomes. These outcomes include exit rates and teacher productivity. Particularly, focusing on differential effects on exit rates by pre-tenure quality will allow me to analyze the selection effects of the tenure reform.

---

[7] While the evaluation is organized only for schools with five or more students, the dataset covers around 98 percent of the student population.

[8] This includes the number of students enrolled, computers, libraries, sport facilities, access to internet and presence of banks

References

Aber, J. L., Torrente, C., Starkey, L., Johnston, B., Seidman, E., Halpin, P., . . . Wolf, S. (2017). Impacts after one year of "healing classroom" on children's reading and math skills in drc: Results from a cluster randomized trial. Journal of Research on Educational Effectiveness, 10(3), 507–529.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. The Quarterly Journal of Economics, 131(3), 1415–1453.

Bau, N., & Das, J. (2020). Teacher value added in a low-income country. American Economic Journal: Economic Policy, 12(1), 62–96.

Bertoni, E., Elacqua, G., Hincapié, D., Méndez, C., & Paredese, D. (2019). Teachers' preferences for proximity and the implications for staffing schools: Evidence from peru. Education Finance and Policy, 1–32.

Bertoni, E., Elacqua, G., Méndez, C., & Santos, H. (2020). Teacher hiring instruments and teacher value added: Evidence from peru (Tech. Rep.). IDB Working Paper Series.

Biasi, B. (2021). The labor market for teachers under different pay schemes. American Economic Journal: Economic Policy, 13(3), 63–102.

Bobba, M., Ederer, T., Leon-Ciliotta, G., Neilson, C., & Nieddu, M. G. (2021). Teacher compensation and structural inequality: Evidence from centralized teacher school choice in perú (Tech. Rep.). National Bureau of Economic Research.

Cattaneo, M. D., Jansson, M., & Ma, X. (2017). Simple local regression distribution estimators. Unpublished. http://www-personal. umich. edu/˜ cattaneo/papers/Cattaneo-Jansson-Ma_2017_LocPolDensity. pdf.

Chelwa, G., Pellicer, M., & Maboshe, M. (2019). Teacher pay and educational outcomes: Evidence from the rural hardship allowance in zambia. South African Journal of Economics, 87(3), 255–282.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. American economic review, 104(9), 2633–79.

Duflo, E., Hanna, R., & Ryan, S. P. (2012). Incentives work: Getting teachers to come to school.

American Economic Review, 102(4), 1241–78.

Glewwe, P., Ilias, N., & Kremer, M. (2010). Teacher incentives. American Economic Journal: Applied Economics, 2(3), 205–27.

Glewwe, P., Shen, R., Sun, B., & Wisniewski, S. (2020). Teachers in developing countries. The Economics of Education, 371–389.

Hanushek, E. (2016). Ending tenure to more easily fire bad teachers helps everyone. The New York Times.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of economic literature, 48(2), 281–355.

Lucas, A. M., & Mbiti, I. M. (2014). Effects of school quality on student achievement: Discontinuity evidence from kenya. American Economic Journal: Applied Economics, 6(3), 234–63.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. Journal of econometrics, 142(2), 698–714.

Mulkeen, A., Chapman, D. W., & DeJaeghere, J. G. (2005). Recruiting, retaining, and retraining secondary school teachers and principals in sub-saharan africa. secondary education in africa (seia) thematic study# 4. gec working paper series 2005/# 3. ERIC.

Ng, K. (2021). The effects of teacher tenure on productivity and selection.

Pugatch, T., & Schroeder, E. (2018). Teacher pay and student performance: evidence from the gambian hardship allowance. Journal of Development Effectiveness, 10(2), 249–276.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. Econometrica, 73(2), 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American economic review, 94(2), 247–252.

Strunk, K. O., Barrett, N., & Lincove, J. A. (2017). When tenure ends: The short-run effects of the elimination of louisiana's teacher employment protections on teacher exit and retirement. Education Research Alliance Technical Report.

Table 1: Descriptive statistics – Permanent and Temporary teachers

	Temporary			Permanent		
	Mean	Sd	N	Mean	Sd	N
Teacher is male	.53	.49	19,314	.53	.49	963
Teacher's Age	38.82	7.06	19,314	35.77	6.88	963
Teacher has a university degree	.48	.49	19,314	.68	.47	963
Teacher has an institute degree	.54	.49	19,314	.34	.47	963
Teacher has both degrees	.02	.15	19314	.02	.14	963
Teacher is disable	.37	.48	19,314	.21	.41	963
No public teaching experience	.10	.30	19,314	.12	.32	963
No private teaching experience	.46	.49	19314	.36	.47	963
Pub. Experience less than 5 years	.45	.49	19,314	.51	.50	963
Priv. Experience less than 5 years	.40	.49	19,314	.47	.49	963
Pub. Experience between 6 and 10 years	.31	.46	19,314	.26	.44	963
Priv. Experience between 6 and 10 years	.10	.30	19,314	.13	.33	963
Pub. Experience more than 10 years	.08	.28	19,314	.07	.26	963
Priv. Experience more than 10 years	.03	.16	19,314	.02	.13	963
Hours worked	26.91	5.26	19,314	29.69	2.85	963

Notes: The sample is composed of teachers who were teaching in second grades of secondary public schools who either had a temporary or a permanent contract in 2016–2019. This is the final sample of teachers who applied for permanent positions during the previous year and are matched to second grade classrooms of secondary school. Data come from the Ministry of Education of Peru.

5 Tables

6 Figures
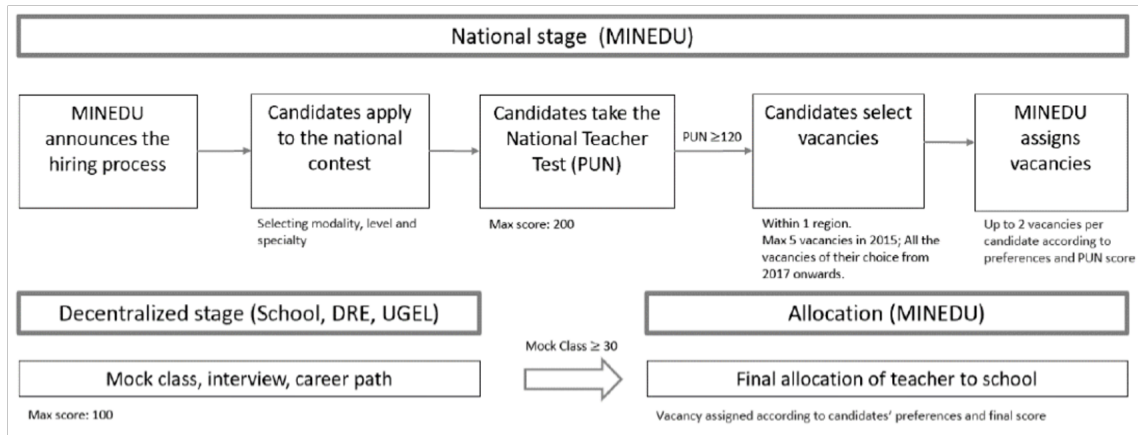
Source: Bertoni et al. (2019)

Figure 1: Overview of the hiring process under the Ley de Reforma Magisterial (LRM)

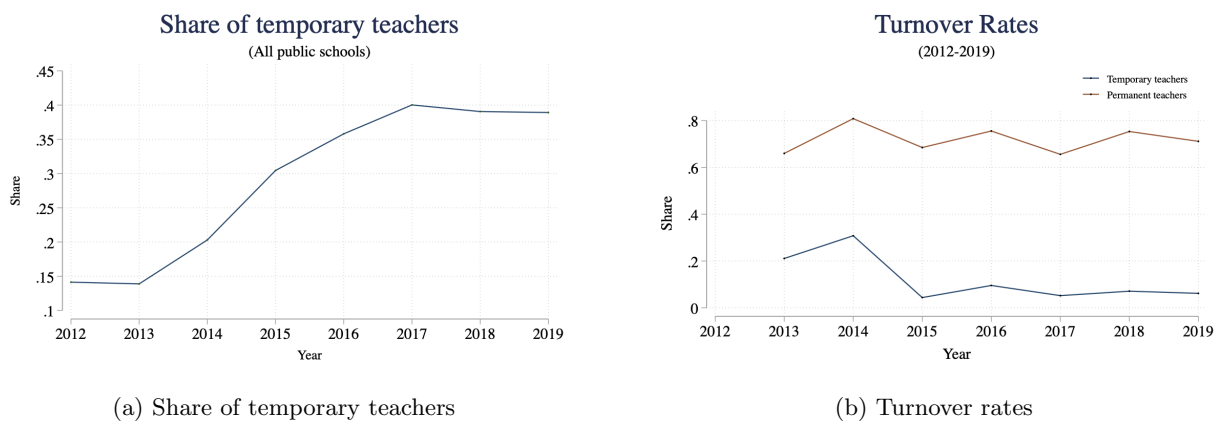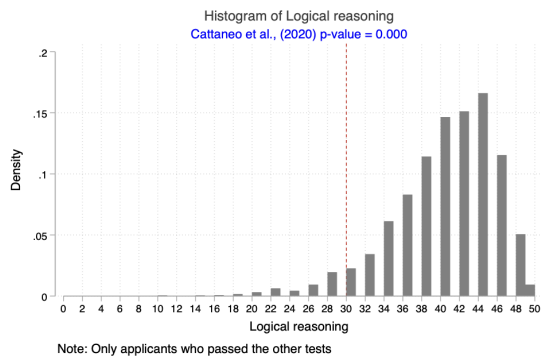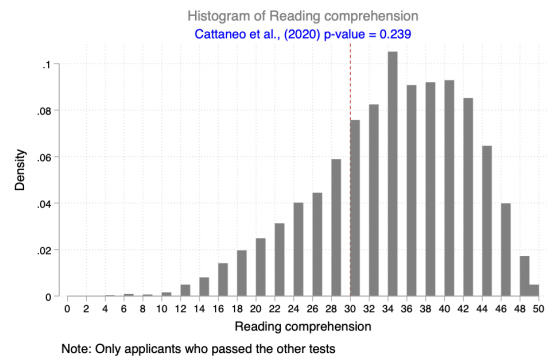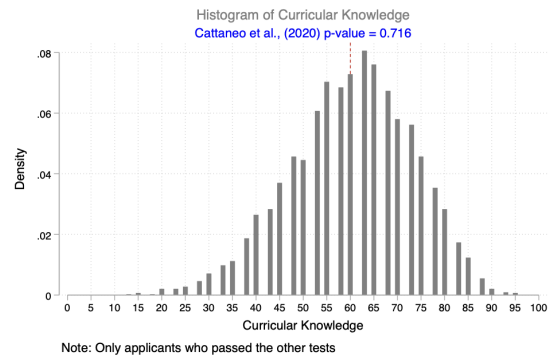(a) Share of temporary teachers

(b) Turnover rates

Figure 2: Share of temporary teacher and turnover rates over 2012-2019
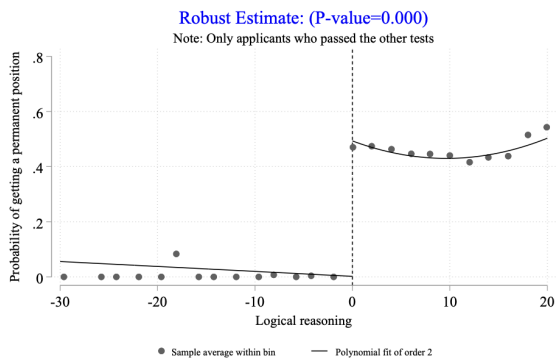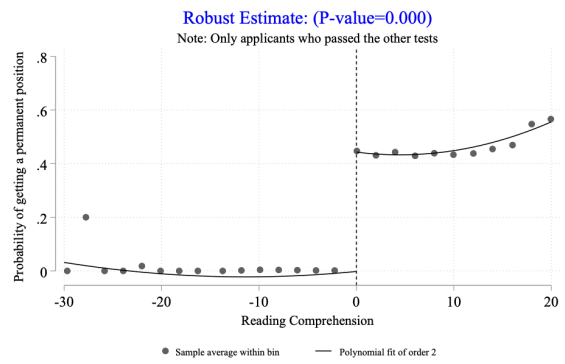
(a) Logical Reasoning

(b) Reading Comprehension
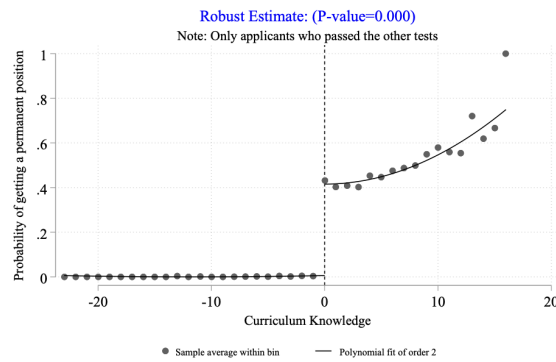
(c) Curricular Knowledge

Figure 3: Conditional Histograms of scores Cattaneo et al. (2017)

(a) Logical Reasoning

(b) Reading Comprehension

(c) Curricular Knowledge

Figure 4: Effect of crossing the threshold for each subject