

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Chatterjee, Somdeep

Article

From better schools to better nourishment: Evidence from a school-building program in India

IZA Journal of Labor Policy

Provided in Cooperation with: IZA – Institute of Labor Economics

Suggested Citation: Chatterjee, Somdeep (2020) : From better schools to better nourishment: Evidence from a school-building program in India, IZA Journal of Labor Policy, ISSN 2193-9004, Sciendo, Warsaw, Vol. 10, Iss. 1, pp. 1-13, https://doi.org/10.2478/izajolp-2020-0002

This Version is available at: https://hdl.handle.net/10419/298745

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.



https://creativecommons.org/licenses/by/4.0/

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU

Somdeep Chatterjee*

From better schools to better nourishment: evidence from a school-building program in India¹

Abstract

This is a short paper analyzing the potential effects of a targeted school-building program on health indicators. The Kasturba Gandhi Balika Vidyalaya (KGBV) program in India intended to build residential schools for girls from historically disadvantaged sections of the society, providing a unique multifaceted policy setting with tenets of gender equality, affirmative action, and infrastructure reform in education. Exploiting the potentially exogenous cross-sectional variations generated by the institutional features of implementation of this intervention, I run triple-difference regressions to find that the program led to increases in body mass index (BMI) among the underweight. There seems to be a positive correlation between KGBV exposure and probability of being in the "healthy" band of BMI indicators.

Current version:	February 13, 2020
Keywords:	health; education for girls; BMI; affirmative action
JEL codes:	125, 128, 112, 115
Corresponding author:	Somdeep Chatterjee
	somdeep@iiml.ac.in

I thank seminar participants at Department of Economics, University of Calcutta, an anonymous reviewer, and the editor, Nuria Rodriguez-Planas, for useful comments and suggestions.

Business Environment (Economics) Group, Indian Institute of Management Lucknow, IIM Road, Prabandh Nagar, Lucknow 226013, Uttar Pradesh, India

[©] The Author(s). 2020 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Cite as: Somdeep Chatterjee. *IZA Journal of Labor Policy* (2020) 10:2.

1 Introduction

Empirically identifying the causal effects of education on outcomes of interest is difficult because of obvious endogeneity issues. While most papers use mandated regulations, e.g., compulsory schooling reforms, some like Duflo (2001) and Chin (2005) use infrastructure reforms to generate identifying variations. Interestingly, very little seems to be known about the direct effects of such infrastructural reforms on health indicators even though such effects seem very likely. A potential channel through which schooling reforms may lead to better health could be the switch away from child labor, much of which is usually physically demanding, and a reform that keeps children in schools is likely to prevent the incidence of such labor. The other obvious channel is access to better sanitation and hygiene, which is likely to be brought about by reforms providing better schooling infrastructure.

Among the little evidence that exists, Breirova and Duflo (2004) find causal effects of the INPRES program of Indonesia on mortality and long-term fertility decisions although the context of their paper is to study the impacts of parental education on child mortality. In this paper, I study a school-building program from India, namely, the Kasturba Gandhi Balika Vidyalaya (KGBV) initiative, to answer the question: does better schooling infrastructure lead to better health of the affected individuals? In a sense, the idea is to study the direct impacts of an infrastructural reform on the short-term health status of those potentially affected.

The intention of the KGBV program was to build residential schools for girls in Grades 6–8 from historically disadvantaged sections of society in *educationally backward blocks* identified based on predefined literacy thresholds all over India. The primary idea was to increase the levels of educational achievement among girls in the country and, since the program was targeted toward the scheduled castes (SCs) and scheduled tribes (STs), which are the marginalized sections of the Indian population, the program can essentially be viewed as an affirmative action in the field of elementary education. Chatterjee (2017) finds that KGBV led to an increase in enrollment and reading test scores of kids potentially exposed to the program, in what constitutes, to date, the only direct causal estimates of the program, to the best of my knowledge. Reduced-form effects of this program on health indicators have not been estimated. From what is previously known, this reform did not explicitly include any stated features to improve the health of the girls going to KGBV schools. The stated objectives were mainly to increase literacy and enrollment. Consequently, the estimates in this paper can also be considered potential spillover effects of the program.

Since the program was essentially implemented in certain regions based on whether female literacy rates in that region were less than the national average, a potentially attractive source of exogenous variation to identify the causal effects would be to compare these regions to others before and after program implementation. However, as Chatterjee (2017) argues, this methodology would lead to confounding estimates as there were other contemporary programs introduced based on this criteria in the country. Therefore, following Chatterjee (2017), I use a triple-difference estimation strategy exploiting plausibly exogenous cohort-level variation in exposure to the KGBV program to identify causal effects on health status. I use body mass index (BMI) as a proxy outcome variable for the health status.

2 Background

2.1 The KGBV program: building residential schools for girls

The KGBV program was introduced by the Indian government in the year 2004–2005 for improvement in educational status of the historically marginalized sections of the Indian population, viz., SCs and STs. While 75% of all KGBV seats were reserved for minority girls, the remaining 25% was kept open also for families below the poverty line, irrespective of minority status. Implementation of the program was nationwide and carried out in all regions classified as educationally backward blocks (EBBs). A block is an administrative division smaller than a district but bigger than a village. A block is considered to be an EBB if the female rural literacy in the block is below the national average and if the gender gap in literacy is above the national average based on the 2001 census.

Census figures suggest that roughly 25% of the Indian population consists of the SCs and STs. The state of Punjab has the highest percentage of SCs (approximately 29%), whereas Mizoram has the highest percentage of STs (95%). India also has a very unfavorable sex ratio for women, with only about 74 out of a total of 593 districts – as per the census of 2001 – having at least as many women compared to men. While it is quite possible that marginalized sections of the Indian society have higher prevalence of malnutrition, the implementation of KGBV however, did not make health status a salient feature for consideration of program penetration. In general, since the program was implemented in the EBBs, it is not unlikely that the health status of these regions would have been poor especially if we assume a positive association between literacy and health.

An evaluation report by the Planning Commission of India (Niti Aayog 2015) points out that 3,609 KGBVs have been sanctioned throughout the country. Around 69% of the teachers have had some sort of prior training and the majority of them have either a postgraduate degree or a professional qualification (such as Bachelor of Education [B.Ed.]). The report also suggests that about 80% of the schools are equipped with computer facilities and have access to fully functional libraries. KGBV is a voluntary program; therefore, the reform should not be confused with other standard compulsory schooling reforms prevalent elsewhere in the world. The KGBV program, implemented in 2004–2005, targeted girls in middle school in India, which corresponds to the age group of 11–14 years. Therefore, during the period of the survey used in the study (conducted in 2011–2012; described in the following section), girls aged 11 years are the youngest to be ever-affected by KGBV and those aged 22 years during the survey must be the oldest that have been exposed ever to KGBV.

A reason why this program serves as an interesting case study for estimating the effects of education on health is from the perspective of the policymaker in a developing country. KGBV was a mix of an infrastructure reform, a gender-equality reform, and affirmative action, and therefore, the effects of such a three-pronged policy in a large economy such as India maybe relevant in terms of replicability elsewhere in the developing world. Since these schools are essentially residential in nature, it is not unlikely that greater enrollment would naturally lead to better health through nutritional channels. For instance, some news reports suggest that in the state of Telangana, which has 475 KGBVs catering to 80,000 underprivileged kids, nonvegetarian items have been included in the weekly menus with the idea of increasing the intake of protein, leading to better nourishment, which otherwise would not have been affordable for these families (https://www.thehindu.com/news/cities/ Hyderabad/mutton-on-menu-for-girls-of-kgbv-schools/article22320963.ece). This provides a potential channel through which KGBV exposure may lead to better BMI for the malnourished kids.

2.2 From education to better health: establishing the link

While the link between education and health has been widely studied, dating back to Grossman (1972), empirically identifying the causal effects of education on health has relied on finding relevant instruments for education, and the commonly used approach is through exploiting schooling reforms (see Arendt 2005, 2008; Brunello et al. 2013, 2016; Parinduri 2017, and so on). The central idea of such a strategy is that a schooling reform is unlikely to affect health through channels other than education.

The other very important government intervention widely studied in the field of education is improving schooling infrastructure. For instance, the INPRES school construction program in Indonesia (Duflo 2001) and "Operation Blackboard" in India (Chin 2005) have been found to have significant effects on various measures of education. However, very little seems to be known about the direct effects of such infrastructural reforms on health indicators.

Why is it important to know the direct effect of this schooling infrastructure policy on health? This is because countries like India usually run against strict budgets in terms of development expenditure on health and education. For instance, while India spends about 4%–6% of its gross domestic product (GDP) on education, it is only able to spend about 2% of its GDP on health, compared to the much higher shares in developed nations, such as 18% for the USA (see https://thewire.in/health/indias-defence-budget-is-nearly-five-times-the-health-budget and https://www.crfb.org/papers/american-health-care-health-spending-and-federal-budget). Consequently, if spillovers exist from a reform in one sector to another, designing and implementation of policy becomes a lot more efficient. The purpose of this paper is to study if such spillovers actually exist, using KGBV as an example case. Chatterjee (2017) has already evaluated the impact of the KGBV program on educational outcomes. The motivation of this paper is that in the presence of potential spillovers to other outcomes such as health, the overall assessment of the impact of the policy maybe underestimated if one does not take into account the unintended consequences as well.

In this paper, we use measures of BMI as the outcome variable and as a proxy indicator of health status. This choice of variable is motivated by two factors. First, KGBV schools are residential in nature and, as a result, meals and dietary supplements provided at these schools are likely to be a lot different from the standard nutritional intake at home. Considering the high incidence of malnutrition, better intake in schools is most likely to manifest in improved health through nutritional status. BMI is the commonly accepted metric for measures of health along this dimension. Unfortunately, we do not observe caloric intake in the data set, and the results can therefore not be validated through a more accurate channel. Second, as these KGBV schools require kids to be in residence, the likelihood of parents sending their kids off to child labor is minimized and this may lead to better BMI measures. The majority of existing evidence on the impact of residential schools on health and related outcomes is somewhat aberrant. In general, it has been associated with the mental trauma of being away from one's family (Schaverien 2015) or exposure to cohorts alienated from the society or from marginalized backgrounds, leading to poorer labor market consequences and poorer lifestyle choices (Kaspar 2014). The overall impact of sociocultural dimensions and how it matters in terms of impacts on BMI have been sparsely studied in the context of a residential school, with the notable exception of Cardoso and Caninas (2010). This is because most of the work that child laborers do involve physical labor in potentially unhealthy and hazardous environments for this age group, which are likely to have an impact on their health in terms of expending relatively more calories than that consumed. As a result, their BMI would be at a low level in the counterfactual condition.

The paper contributes to the literature in three major ways. First, to the best of my knowledge, this is the only paper looking at the direct effects of a school infrastructure policy on BMI. Second, considering the unique context of the policy, this paper presents new estimates of how a mix of affirmative action, gender equality, and infrastructure building in education may affect health indicators. Third, most of the works on schooling reforms used as instruments for education [(with the exception of the studies by Parinduri (2017) and Breirova and Duflo (2004)] have studied the context of developed countries. However, the basic link between health and education assumes greater policy significance in the context of developing countries because of tighter budgets and the potential for spillovers and, therefore, potential efficiency gains. This paper contributes to the literature by pointing out this *positive externality* of an education policy on the health sector.

3 Econometric methodology

3.1 Data

Data come from the nationally representative India Human Development Survey - II (IHDS II) - 2011–12. I restrict my sample to individuals in the age group of 6–30 years. I use an earlier round of the survey conducted in 2004–05 for falsification. I use the data on the measured and observed heights and weights of individuals from the survey to construct BMI measures. Each individual's height and weight were measured twice. I calculate the simple average of these metrics and construct the BMI using the standard definition: the ratio of weight (in kilograms) to height squared (in meters). For all practical purposes, a BMI in the range of 18.5–25 is considered healthy. Any value below this range implies that the individual is underweight, while any value >25 is considered overweight or obese. I use BMI as one of the outcome variables in my analysis to measure the effects of KGBV on health behavior. For the underweight cohort, the mean BMI is roughly around 15. The data also contain several demographic variables that I use as controls. Summary statistics for our sample are reported in Table 1.

3.2 BMI: are KGBV cohorts healthier?

Figure 1 presents a snapshot of the density of BMI across the sample. The two vertical lines at BMI = 18.5 and BMI = 25 form a band, indicating the healthy zone contained within. It

is evident from the figure that the healthy band seems to have a higher density for cohorts with potentially more exposure to KGBV. This is further confirmed by the *t*-tests reported in Table 2, which show that KGBV cohorts have a higher probability of having a healthy BMI.

	Mean	Standard deviation
	(1)	(2)
BMI	18.79	13.22
BMI (if underweight)	15.43	1.94
Age 11–22 years (=1, if yes)	0.51	0.49
SC/ST (=1, if yes)	0.74	0.44
Female (=1, if yes)	0.49	0.49
Age (in years)	17.68	7.08
Education (among males; highest number of years)	7.97	4.90
Education (among females; highest number of years)	5.66	5.21

Table 1	Summary	statistics
---------	---------	------------





Notes: The figure plots the density of BMI values. The area between the two vertical lines is the potential healthy zone, i.e., BMI between 18.5 and 25. The panel on the right plots the densities for the affected cohort, i.e., girls of lower castes in the affected age cohort, and the left-hand side panel includes everyone else.

	KGBV	Non-KGBV	Δ
	(1)	(2)	(1)-(2)
Healthy BMI	0.3088	0.2311	0.0777
Standard deviation	0.0034	0.0015	$H_0: \Delta = 0$
			<i>p</i> -value < 0.001

Table 2 Probability of being in the healthy BMI zone

3.3 Estimation

As in Chatterjee (2017), I use cohort-level variation in exposure to KGBV in a triple-difference estimation framework over three different cross sections, viz., gender, age cohort, and caste. Although using cross-sectional regional variation based on reach of the program may seem appealing, for reasons described by Chatterjee (2017), I refrain from doing so. Such a methodology has been used by Debnath (2012), but it is unlikely to capture the KGBV effects uniquely as another program simultaneously affected these regions. Debnath (2012), as a result, estimates the joint effects of the two programs, but the method used by Chatterjee (2017), which I follow here, can uniquely identify the reduced-form KGBV effects. Since only girls from backward castes and in Grades 6-8 were eligible, an interaction of the three indicator variables for these cross sections can potentially provide exogenous variation in exposure to the program. I additionally use region (village) fixed effects and cluster all standard errors at the village level, improving on the Chatterjee (2017) design. In the study by Chatterjee (2017), robust standard errors were left unclustered and regional fixed effects were not used, essentially making the empirical framework of this paper much stronger. The village indicator is essentially the place of residence and, in these societies, due to informal insurance considerations, migration rates are very low (Munshi and Rosenzweig 2016) and, therefore, less of an issue here in the context of selection into KGBV localities.

I propose to run the following specification as my main model, largely replicating the methodology of Chatterjee (2017):

$$Y_i = \alpha_r + \beta_1 \cdot KGBV + \beta_2 \cdot (girl^*affected) + \beta_3 \cdot (girl^*disadvantaged)$$

+ $\beta_4 \cdot (affected^*disadvantaged) + \beta_5 \cdot girl + \beta_6 disadvantaged + \beta_7 \cdot affected + \gamma \cdot X + Ú$ (1)

where α_r represents the regional fixed effects, *girl* is a dummy variable for females, *affected* is the dummy variable for the age cohort 11–22 years, and *disadvantaged* is a dummy for marginalized castes. The interaction of the three cross-sectional dummy variables captured by KGBV generates potentially exogenous variation in access under the assumption that the difference in the difference-in-differences of mean values of outcome *Y* along the three cross sections is statistically indistinguishable from zero in the absence of the intervention. The controls for age, education of male and female household heads, and size of the household are represented by *X*. The outcome is represented by *Y*, which – for most of our regressions – is going to be BMI.

I present a brief summary of the identification strategy in Table 3. The treatment group, as identified by the *affected* dummy variable described above, consists of girls who have ever been exposed to the policy. Since the KGBV program was intended for girls in middle school and the middle school in India roughly corresponds to the 11–14 age group, we consider only those girls as affected by the KGBV who are currently in that age group or would have been in

Age in data set	6–10 years	11	12	13	14	15	16	17	18	19	20	21	22 and above
Age in policy year	0–3 years	4	5	6	7	8	9	10	11	12	13	14	15 and above
Exposure to policy	Not yet exposed	Cur exp	rent	:ly d		Pre	viou	Isly	ехро	osed	l		Never exposed

Table 3 Summary of identification strategy

that age group after the introduction of the policy. Since our sample includes all students in the age group of 6–30 years, we consider the 6- to 10-year-old kids as part of the control cohort as they are yet to be in middle school. Moreover, the girls in the age group of \geq 22 years would have potentially completed middle school by the time the policy was implemented. Considering that Indian schools mostly followed a no grade-detention policy up to middle school, this is a fairly innocuous assumption. Girls aged 11–14 years are currently likely to be in middle school and the ones \leq 22 years and \geq 14 years would have completed middle school post-KGBV intervention. As a result, these girls are considered the treated cohort.

To make sure that this cohort convergence is meaningful for estimating the causal effect of KGBV on BMI, I additionally run two other specifications as follows. This makes identification of the control group much more intuitive. First, I restrict the sample to only include girls (as KGBV would only have affected them) and then run a standard difference-in-difference across the other two dimensions:

$$BMI_{i} = \alpha_{i} + \theta_{i} \cdot (affected^{*}disadvantaged) + \theta_{i} \cdot disadvantaged + \theta_{i} \cdot affected + \sigma \cdot X + \upsilon$$
(2)

Here, θ_1 is the effect of the intervention on the affected cohort's BMI among girls from the disadvantaged sections of the society compared to girls from other sections. Then, I restrict the sample to only the disadvantaged groups (who are also the only ones potentially affected by KGBV) and run the following specification:

$$BMI_{i} = \alpha_{r} + \phi_{1} \cdot (girl^{*}affected) + \phi_{2} \cdot girl + \phi_{3} \cdot affected + \omega \cdot X + \mu$$
(3)

Here, ϕ_1 is the differential effect of the intervention on the affected cohort's BMI between the girls and boys of the disadvantaged sections of the society.

4 Results

In this section, I present the results from the estimation and falsification exercises. I also report findings from robustness checks.

4.1 Impact of KGBV on BMI

I use the BMI of individuals in the age group of 6–30 years as the main dependent variable to check for any effects of the program on this health indicator at the extensive and intensive margins. The choice of this variable is almost obvious. Since the channel through which we expect KGBV to affect the health status is either improved access to health and sanitation and enhanced nutrition through better diet in the residential schools or through a reduction in child labor, it is most likely that any health effects would show up on how well nourished the individual is. As a result, the BMI seems to be the best approximation of any such measure.

I do not find any extensive margin effects, as reported in Column 1 of Table 4. KGBV did not lead to any change in the probability of being malnourished (BMI < 18.5). However, the estimation results of Equation 1 in Column 2 indicate significant intensive margin effects. KGBV seems to have led to an improvement in the health status of the malnourished individuals. I find that with KGBV exposure, the BMI index is higher for the malnourished category

	Extensive margin	Body mass index (if BMI < 18.5		
	BMI < 18.5	KGBV effect	Falsification	
	(1)	(2)	(3)	
KGBV	-0.0043 (0.0125)	0.1901** (0.0922)	0.0742 (0.1902)	
R^2	0.33	0.41	0.45	
Observations	90,988	33,406	16,756	
Mean of dependent variable	0.37	15.43	15.10	

Table 4	Effects of KGBV (on health status: tri	ple-difference	methodology
100000	Elicets of Koby (on neath status, th	pic annerence	methodology

Notes: Column 1 suggests no effects of the program at the extensive margins of health; Column 2 presents the results on the intensive margin effects, whereas Column 3 reports the falsification results at the intensive margin. Therefore, the sample in columns 1-2 is restricted to only individuals in the age group of 6–30 years in the IHDS-II data set (2011–12), which is a period after the KGBV was implemented. In Column 3, the sample is from the pre-policy period, i.e., IHDS-I. All the columns report results from different regressions. Column 1 is a regression on the full sample. Columns 2 and 3 are the results from restricted subsamples. For columns 2 and 3, the sample is restricted to only the low-BMI individuals, with BMI <18.5. It only makes sense for an increase in BMI for this subcategory of the population. The extensive margin of this measure is essentially the outcome variable in Column 1. The coefficient KGBV is the causal effect of the KGBV program on outcomes, as described in the section on the triple-difference estimation strategy. All regressions include the regional (village) fixed effects and control for the relevant baseline dummy variables and double interactions. Additional controls are age, family size, and education of household head, both female and male. Robust standard errors clustered at the regional (village) level are reported in parentheses. **p < 0.05.

by 0.19 points, which is roughly 1.25% compared to the mean. One concern could be that religion is an omitted variable and the behavior of individuals may be different by religious identity. However, because the caste categorization is largely prevalent only among Hindus, it is unlikely to be a cause of major concern, considering that the majority of the sample consists of Hindus either way. Regressions including religion as a control do not change the magnitude of the effect. Standard errors are marginally higher keeping the effect sizes significant at the 95% level of confidence.

The causal estimate holds under the assumption that in the counterfactual condition, i.e., in the absence of the KGBV, this estimated difference in BMI would be statistically indistinguishable from zero. Since this is an assumption about the counterfactual condition, there is no way to test this statistically. However, as per standard norms in practice, one might run placebo regressions to provide some support to this assumption. Following Chatterjee (2017), I use the IHDS-I data (published in 2004–05) for falsification and confirm that there are no effects in the pre-policy year for the significant intensive margin variable. Essentially, I run exactly the same regression using the same outcomes and controls from a survey conducted before the implementation of KGBV would have taken place. As a result, the estimated coefficient should not be significant if there are no preexisting differences along the dimensions used in the triple-difference analysis. In Column 3 of Table 4, I find that not only is the point estimate insignificant for this falsification exercise, it is also much smaller in magnitude, which is reassuring in terms of support for the assumptions required to sustain the identification strategy. I also perform a robustness check (results not reported here but available upon request) by running

	Body mass index (if BMI < 18.5)			
	Females	Disadvantaged castes		
	(1)	(2)		
affected*disadvantaged	0.1984*** (0.0716)			
girl*affected		0.3916*** (0.0427)		
R^2	0.47	0.44		
Observations	17,177	25,799		
Mean of dependent variable	15.56	15.41		

Table 5 Cohort convergences: difference-in-differences

Notes: Sample includes 6- to 30-year–old individuals in IHDS-II with BMI values <18.5. In Column 1, the subsample is restricted to only females. The estimated coefficient gives the effect of the intervention on the affected cohort's (disadvantaged groups) BMI compared to other groups. In Column 2, the subsample is restricted to only the disadvantaged groups. So, the estimated coefficient gives similar cohort effects for girls relative to boys, only in this subcategory. All regressions include regional (village) fixed effects and controls for the relevant baseline dummy variables. Additional controls are age, family size, and education of household head, both female and male. Robust standard errors clustered at the regional (village) level are reported in parentheses. ***p < 0.01.

the same regressions on a sample from states that did not have a single EBB based on the 2001 census and hence were potentially not having any exposure to the KGBV program. The *p*-value of the significance test for the coefficient on BMI for our main specification is 0.93, indicating that there is absolutely no impact. This may be considered as a placebo experiment to support the main analysis.

In Table 5, I present the results from regressions described in Equations 2 and 3 above in columns 1 and 2, respectively. In Column 1, I restrict the sample to only girls and run a difference-in-difference regression along the other two dimensions to find very similar differentially affected cohort effects on the BMI for disadvantaged kids relative to the kids of the general castes. The point estimate is very similar in magnitude to the one estimated using the triple-difference method. In Column 2, I restrict the sample to disadvantaged kids and find a similar positive cohort effect for the BMI of girls relative to boys, although the magnitude is somewhat larger. The fact that the estimates across these specifications are not very different provides support for the identification strategy and suggests that the identification strategy that relies on this cohort convergence in a cross-sectional setting does make sense.

4.2 Choice of age cohorts

In the above analysis, identification is critically reliant on the choice of age cohorts. In other words, the control group for this *quasi-experimental* design includes the boys who do not belong to the disadvantaged SC/ST castes and who are aged 6–10 years and 23–30 years. A concern could be that the choice of this age cohort-based control group is not meaningful. To alleviate such concerns, I run cohort-specific regressions of BMI for individuals with BMI <18.5 in a difference-in-difference framework using just dummies on whether the individual is disadvantaged, if the gender is female, and their interaction, apart from all the other usual controls with the age 6 cohort as the omitted category.



Figure 2 Triple differences along gender and disadvantaged dummies by age.

Figure 2 plots the estimated coefficients for the estimated regressions by age. Each point on the graph represents the estimated triple difference for that particular age cohort relative to the omitted age 6 years. So, instead of *KGBV*, which is essentially *girls*disadvantaged*affected* relative to the unaffected, points on the graph in Figure 2 represent *girls*disadvantaged*affected* relative to age 6. The vertical lines represent the 95% confidence intervals. If the above identification strategy is meaningful, then one would expect these coefficients to be significant for the affected cohorts only. This is largely what we see in the figure. The coefficients become significant at age 11, which is the first age cohort in the treated group, and the coefficients are insignificant for all younger cohorts as they were unaffected by the KGBV. Roughly around age 22, the coefficient comes down to almost zero, which is again the eldest cohort likely to be affected. All coefficients from age 23 and above seem to be largely insignificant, providing support to the identification strategy.

4.3 Further robustness checks

A remaining concern with the above analysis could be that the short-run and long-run effects are mixed because the age spans of the people in the sample implies that the estimation is done for people in school age as well as older-than-school age. Furthermore, household composition variables are usually good controls for school-aged children but may not be so for adults. This is because household composition is potentially endogenous to education. I thank an anonymous referee for pointing this out and motivating the robustness check exercise. As a result, in this section, I report results from the above analysis by breaking down the sample to younger cohorts, to potentially include closer-to-school-age people in the control group.

Table 6 presents these results. Columns (1)–(3) report the estimated effect sizes for BMI. For comparison, Column 2 from Table 2 has been reproduced as Column 1 in Table 4. It is found that the estimates mostly hold up even after restricting the sample to younger (and closer-to-school age) cohorts. There still seems to be a positive effect on BMI among the underweight children across the board. This exercise potentially addresses some of the concerns mentioned above.

4.4 Are KGBVs allotted based on health status?

Another concern regarding the potential validity of the above empirical exercise would be with regard to the penetration of the KGBV program. Is it possible that introduction and

	Body mass index				
	Age < 30	Age < 25	Age < 22		
	(1)	(2)	(3)		
KGBV	0.1901** (0.0922)	0.1832* (0.0951)	0.2131** (0.0985)		
R ²	0.41	0.41	0.39		
Observations	33,406	30,723	29,389		

Table 6	Robustness	check:	restrictive	subsamples

Notes: All columns report results from different regressions. The coefficient *KGBV* is the causal effect of the KGBV program on outcomes, as described in the section on estimation strategy. All regressions include regional (village) fixed effects and controls for relevant baseline dummy variables and double interactions. Additional controls are age, family size, and education of household head, both female and male. Robust standard errors clustered at the regional (village) level are reported in parentheses. **p < 0.05; *p < 0.1.

	KGBV grants received	No KGBV grants received	Δ
	(1)	(2)	(1)-(2)
Mean BMI (women aged 15–49 years)	20.283	20.987	-0.704
Standard deviation	0.334	0.204	$H_0: \Delta = 0$ p-value = 0.12

Table 7 Comparison of states by mean BMI based on KGBV penetration

implementation of KGBV are driven by initial differences in health status? If this is likely, then a potential selection bias may confound the above estimates. The policy targeted the historically marginalized sections of the Indian population. Therefore, even though the caste identity of an individual, i.e., whether one is from an SC/ST household or not, is random, the fact that KGBVs could have been prioritized in areas with a low base in terms of health indicators is problematic because the above strategy would then overestimate the true effect of the program.

To alleviate these concerns, I conduct a very simple analysis on data collected from a different source for a time period concurrent with the implementation of the policy. I use statelevel data on the disbursement and sanction of KGBV funds from the government as of December 2005. The data have been compiled from the answers to a question asked in the lower house of the Indian Parliament on March 7, 2006. I compare the states with no funds received to the ones that received some KGBV funds. For comparing these states, I look at the state-level average BMI levels of women in the age group of 15–49 years (majority of this age cohort is part of our analysis), excluding pregnant women and women having given birth in the preceding 2 months, for the year 2005–06. These data come from the Ministry of Health and Family Welfare, Government of India. Both these variables are collected from the compilations available at the online repository, *Indiastat.com*.

The results from this analysis are reported in Table 7. While the mean BMI levels in states that have received KGBV funds do appear to be numerically smaller than that of states not receiving any KGBV funds, the difference is only marginal and statistically indistinguishable from zero. Therefore, it is very unlikely that the government was prioritizing KGBV in states

based on BMI, which is our primary outcome variable here. This exercise provides reasonable support to the validity of our empirical framework.

5 Conclusions

Building residential schools for disadvantaged girls in India appears to have led to significant improvements in BMI among the potentially malnourished people in areas potentially exposed to the program. The probability of having a healthy BMI seems to be higher for individuals potentially affected by the policy. Since the program studied in this paper was a targeted education reform, much of these effects can be interpreted as the ancillary reduced-form effects of the program on health. One of the channels through which these effects may operate could be that better education leads to better awareness about hygiene and sanitation, and this leads to better observed health effects. Other channels may include a decline in child labor and access to better nutrition in the residential school setup.

References

- Arendt, J.N. (2005): Does Education Cause Better Health? A Panel Data Analysis Using School Reforms for Identification. Economics of Education Review 24(2), 149-160.
- Arendt, J.N. (2008): In Sickness and in Health? Till Education do us Part: Education Effects on Hospitalization. Economics of Education Review 27(2), 161-172.
- Breirova, Lucia; Esther Duflo (2004): The Impact of Education and Fertility and Child Mortality: Do Fathers Really Matter Less than Mothers? NBER Working Paper.
- Brunello, Giorgio; D. Fabbri; M. Fort (2013): The Causal Effect of Education on Body Mass: Evidence from Europe. Journal of Labor Economics 31(1), 195-223.
- Brunello, G.; M. Fort; N. Schneeweis; R. Winter-Ebmer (2016): The Causal Effect of Education on Health: What is the Role of Health Behaviors? Health Economics 25(3), 314-336.
- Cardoso, H.F.; M. Caninas (2010): Secular Trends in Social Class Differences of Height, Weight and BMI of Boys from Two Schools in Lisbon, Portugal (1910?2000). Economics & Human Biology 8(1), 111-120.
- Chatterjee, Somdeep (2017): Getting Girls to Schools! Assessing the Impacts of a Targeted Program on Enrollment and Academic Performance. The B. E. Journal of Economic Analysis and Policy 17(1).
- Chin, Aimee (2005): Can Redistributing Teachers Across Schools Raise Educational Attainment? Evidence from Operation Blackboard in India. Journal of Development Economics 78(2), 384-405.
- **Debnath, Sisir** (2012): School Subsidy for Girls and Gender Gap in Enrollment. Working Paper (Unpublished Manuscript).
- Duflo, Esther (2001): Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. American Economic Review 91(4), 795-813.
- Grossman, M. (1972): On the Concept of health Capital and the Demand for Health. Journal of Political Economy 80(2), 223-255.
- Kaspar V. (2014): The Lifetime Effect of Residential School Attendance on Indigenous Health Status. American Journal of Public Health 104(11), 2184-2190.
- Munshi, Kaivan; Mark Rosenzweig (2016): Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. American Economic Review 106(1), 46-98.
- Niti Aayog (2015): Evaluation Study on Kasturba Gandhi Balika Vidyalaya (KGBV), PEO Report No 228, Program Evaluation Organization.
- Parinduri, R.A. (2017): Does Education Improve Health? Evidence from Indonesia. The Journal of Development Studies, 53(9), 1358-1375.
- Schaverien, J. (2015): Boarding School Syndrome: The Psychological Trauma of the 'Privileged' Child. London, Routledge.