Online Appendix For

Access to Colleges, Human Capital, and Empowerment of Women

Sheetal Sekhri

Md Amzad Hossain

Pooja Khosla

A.1 Effect Size: Increase in College Enrollment vs Increase in Capacity

In this section, we compute the effect size of college enrollment in levels to assess if this is plausible given the expansion in college capacity.[44]

1. ***Capacity Count Based on Official UGC Statistics****: Based on our estimates, and using an average approximate number of seats of 1,417 per college, the approximate increase in enrollment capacity per district would be approximately 9,659 over a period of eight years (2009-2016)*

   The AISHE data does not provide cross-walks so we cannot estimate the exact change in capacity or number of seats in districts over years. Hence, we approximate capacity and acknowledge that it is not precise. Our capacity count is based on a report published by the University Grants Commission (UGC) of India, the nodal body regulating higher education. In their 2008 publication titled "Higher Education in India: Issues Related to Expansion, Inclusiveness, Quality and Finance", UGC summarized average enrollment per college by states based on a sample of 1400 colleges across India.[45] Based on these state-wise average survey based enrollment per college (assuming in the survey colleges, available seats must be atleast as large as enrollment), we computed the average capacity of Indian colleges by taking the weighted average of the average enrollment capacity per college of each state, where the weight is based on the population aged between 20-24 in these states. This count is 1,417.

   Based on our estimates in Table II, an additional 0.192 colleges are added each year. So in year 1, 0.192 additional colleges are operating. This will generate an additional approximate enrollment capacity of 0.192×1,417 in the first year (where we use the fact that average enrollment per college in India is 1,417 students). In year 2, an

additional 0.192 colleges are operational, which implies a total of 0.192×2 (=0.384) more colleges relative to the baseline year. This in turn implies an additional capacity of 0.384×1,417 in the second year relative to the baseline. Continuing in this manner, in year 8, there are 0.192×8×1,417 additional seats. Summing the capacity increases from years 1 through 8 yields an estimated 9,659 more seats created over the 8 year period. The calculations for the estimated capacity increases are as follows:

Capacity increase = (0.192×1×1,417) [Year 1] + (0.192×2×1,417) [Year 2] + (0.192×3×1,417) [Year 3] +(0.192×4×1,417) [Year 4]+ (0.192×5×1,417) [Year 5] + (0.192×6×1,417) [Year 6] + (0.192×7×1,417) [Year 7] +(0.192×8×1,417) [Year 8] = 9,659

Note that intake in colleges happens only in year 1 (the first year of degree programs, though very limited transfer students are accepted in years 2 or 3 of the Bachelor's program). So seats or capacity imply year 1 capacity, not the aggregate for all years that the student is in college.

2. ***Enrollment:*** *Based on our estimates, and using an average baseline male, and female population of 86,241 and 80,932 aged 20-24 per treated district, respectively, the increase in college enrollment per district is approximately 13,327 students over a period of eight years (2009-2016).*

   As mentioned in our paper, we restrict our sample to women who stayed in the same place all their lives (we do this to ensure that we know the women's region of birth and do not include migrants in our sample). Such women comprise 17 percent of 18+ aged women (as per the survey data we used). Thus, we impose this restriction on the age-relevant female population from the Census of India of 2011. According to the Census of India, 2011, there were around 80,932 females aged 20-24 per treated districts. That implies a total of (80,932*0.17) =13,758 women per district who have likely stayed in

the same place. Combined with the 2.1 annual pp increase in enrollment, roughly an additional 289 (=13,758*0.021) females should have enrolled in college each year which would give us a total of 2,311 (=289*8) females enrolled in college during the eight years after the intervention among the non-movers.

We next estimate the effect for those females who moved (in-migrants). we find an annual increase of 1.1 pp in college enrollment (Appendix Table A.16). This implies an additional 739=(0.011*80932*0.83) females should have enrolled in college each year in this sample. For 8 years, this would add up to 5,912 (=739*8) females enrolled in college. Then, we sum the enrollment from both of these samples (i.e non-movers and movers) and resulting enrollment of 8,223=(2,311+5,912) females during the eight years after the intervention is plausible given the approximated capacity.

The average relevant college-going male population per treated district is 86,241. For men who always stayed at the same place (65% of the sample), Table V, column (4) shows an effect on the likelihood of college enrollment of about 0.6 pp per year.This implies an additional 2,690=(0.006*86241*0.65*8) males enrolled in college during the eight years after the intervention. For men who always stayed at the same place (35% of the sample), we find an increase of 0.1 pp per year in the likelihood of college enrollment.This implies an additional 2,414=(0.01*86241*0.35*8) males enrolled in college during the eight years after the intervention. Adding enrollment from both of these samples, we find an additional 5,104 =(2,690+2,414) males would have enrolled in college during the eight years after the intervention.

Adding enrollment for both males and females, we find an additional 13,327 =(8,223+5,104) males and females would have enrolled in college during the eight years after the intervention. We summarize this discussion in Appendix Table A.5.

3. We found that the increase in college enrollment per district is approximately 13,327

students over a period of eight years (2009-2016). For the same duration, the approximate increase in capacity per district is 9,659. The enrollment figure we obtain exceeds the approximate capacity figure. We do not have data to examine the precise increase in capacity. On scrutinizing various reports and news articles, we think there are two reasons why this approximation of capacity might be understating the available seats. Below we summarize these reasons:

(a) Some of the colleges have evening as well as morning sessions as per the AISHE reports. For example, in Tamil Nadu, during our sample period, state colleges had 2 sessions offering the same courses: morning college sessions offered classes in the morning shift between 8.45 am to 1.15 pm and the evening shift started at 1.30 pm and ended at 6.05 pm (Hindustan Times, 2021). There are 218 state colleges in Tamil Nadu in our sample. We do not have data on what is the fraction of colleges offering morning and evening sessions overall for India. So, our capacity or available seats is plausibly underestimated due to this.

(b) Colleges that were established prior to the policy sometimes have vacant seats (Indian Express, 2014) especially in technical education which is a field chosen predominantly by boys. These vacant seats vary yearly and across regions. For example, 42 percent of the seats in technical colleges in Maharashtra were vacant in 2018 (Hindustan Times, 2019). In our sample, 17% of colleges are technical or engineering. As Table A.4, reveals, the policy leads to an increase in the number of students completing high school (37 percent increase in attainment is accruing from K-12). Some of these marginal students, especially the boys who are propelled by the policy to complete high school, might be enrolling in these other institutions. We do not have data to scrutinize this possibility but a share of increase in men's enrollment could be on this account.

Another important point to note based on this discussion is that since our av-

erage capacity computation utilizes average enrollment in colleges in a survey conducted by the UGC (see discussion above ***Capacity Count based on Official UGC Statistics***), our capacity or number of seats available is plausibly an underestimate.

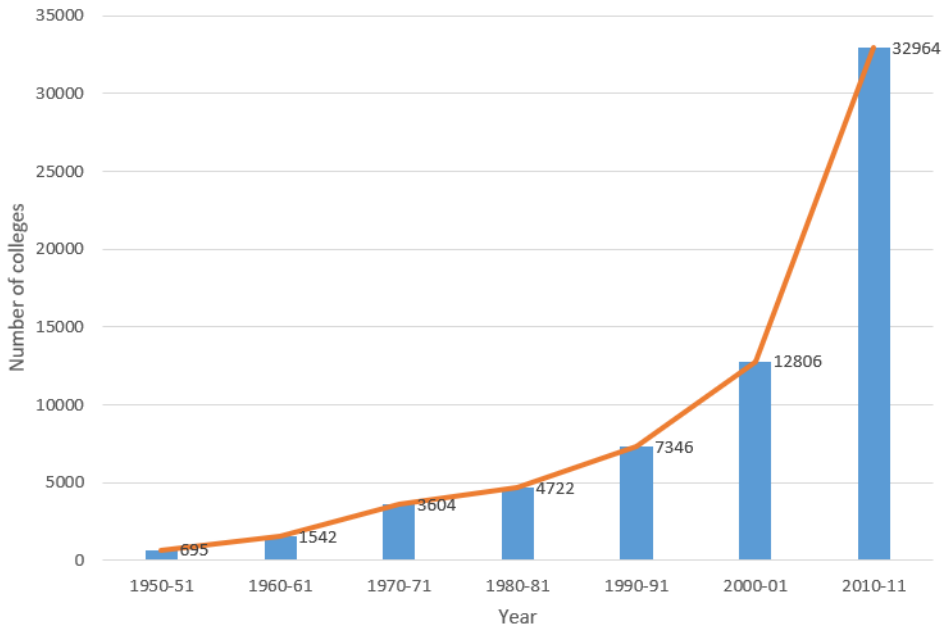
Our data do not allow us to specifically ascertain how much of the enrollment that exceeds the approximate capacity of the new colleges, as shown above, is on account of these two reasons.

A.2 Effect Size: Years of Schooling

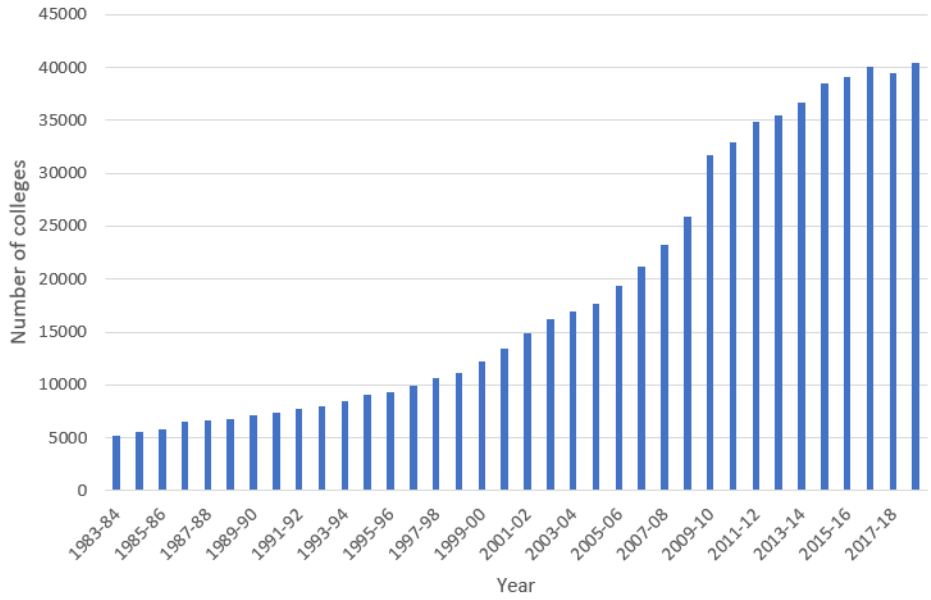
Table 3 shows that the number of years of schooling increased by 0.10 each year for the treated districts post-policy period. This translates to an increase of 0.82 years of schooling over the eight years of post-policy period. In this subsection, we try do disentangle how much of this increase is due to increase in college years and how much is due to an increase in the years of education during middle/high school years (K-12). In Appendix Table A.4, we show the effect of grant policy on number of years of schooling (Column 1), number of years of schooling for individuals who have not enrolled into college (Column 2), and number of years in college (Column 3). The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). Column 3 shows the effect of college construction on the number of college years completed. The coefficient is 0.0664 which indicates an increase of 0.53 years of college education (8*0.0664) over the eight-year period. This is comparable to the US (Doyle and Skinner, 2016).

Column 2 indicates that the increase in the total years of schooling includes an increase in the years of education during middle/high school years (K-12). We find a 0.30 (=0.038*8) years increase in schooling happens for school students in the post-policy period between 2009 and 2016. The sum of the effect: 0.53 + 0.30 is what we get from an aggregate estimate in column 1. There is an incentive effect of college construction. In the literature, other studies have documented that college construction affects the incentives to go to school even if individuals do not enrol in college or complete a college degree (Jagnani and Khanna, 2020).

Appendix Figures and Tables



(a) Number of colleges in different decades



(b) Number of colleges in different years

Figure A.1: Total Number of Colleges over Time

Notes: Data for all years except 1999-2000 and 2000-01 are from different rounds of University Grant Commission (UGC) annual reports. For those two years, the numbers were extrapolated using the growth rate of colleges between 1998-99 and 2001-02.

Figure A.2: College Enrollment in Different Years

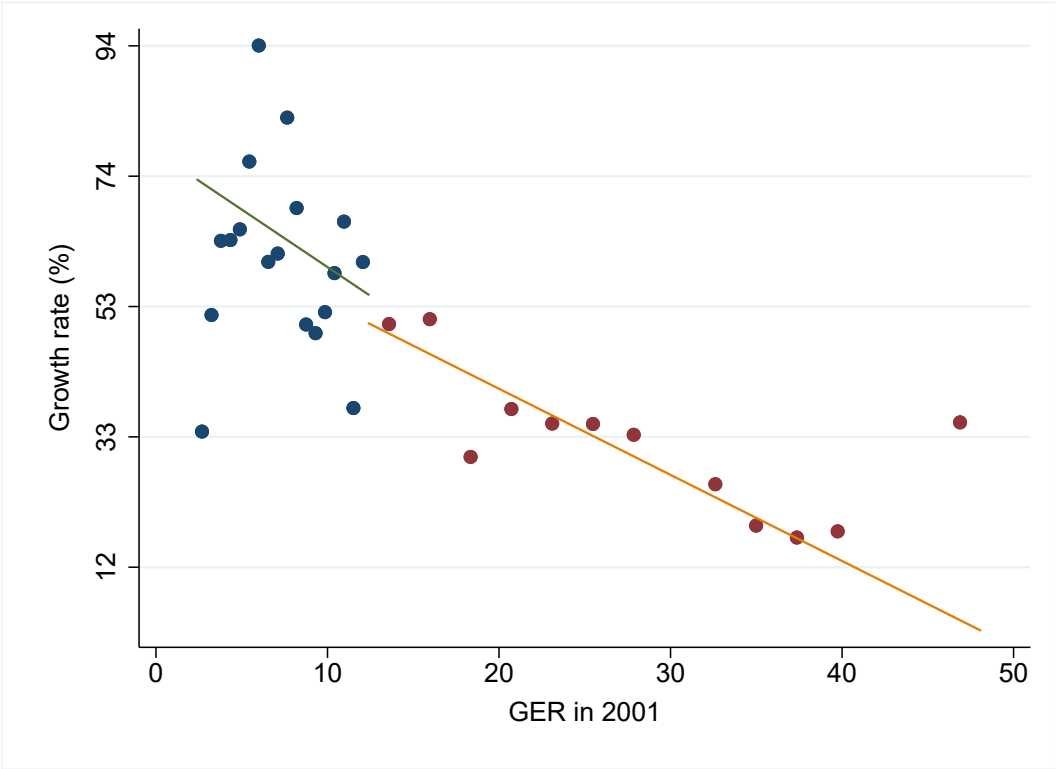


Figure A.3: 2001 GER and Growth in Colleges

Notes: The figure shows a linear fit from a regression of growth of colleges on gross enrollment ratio (GER) at the district level in 2001 separately estimated on the left and right of the GER eligibility threshold of 12.4. The dependent variables is the growth rate of total colleges between 2008 and 2018. Source: Authors' calculation.

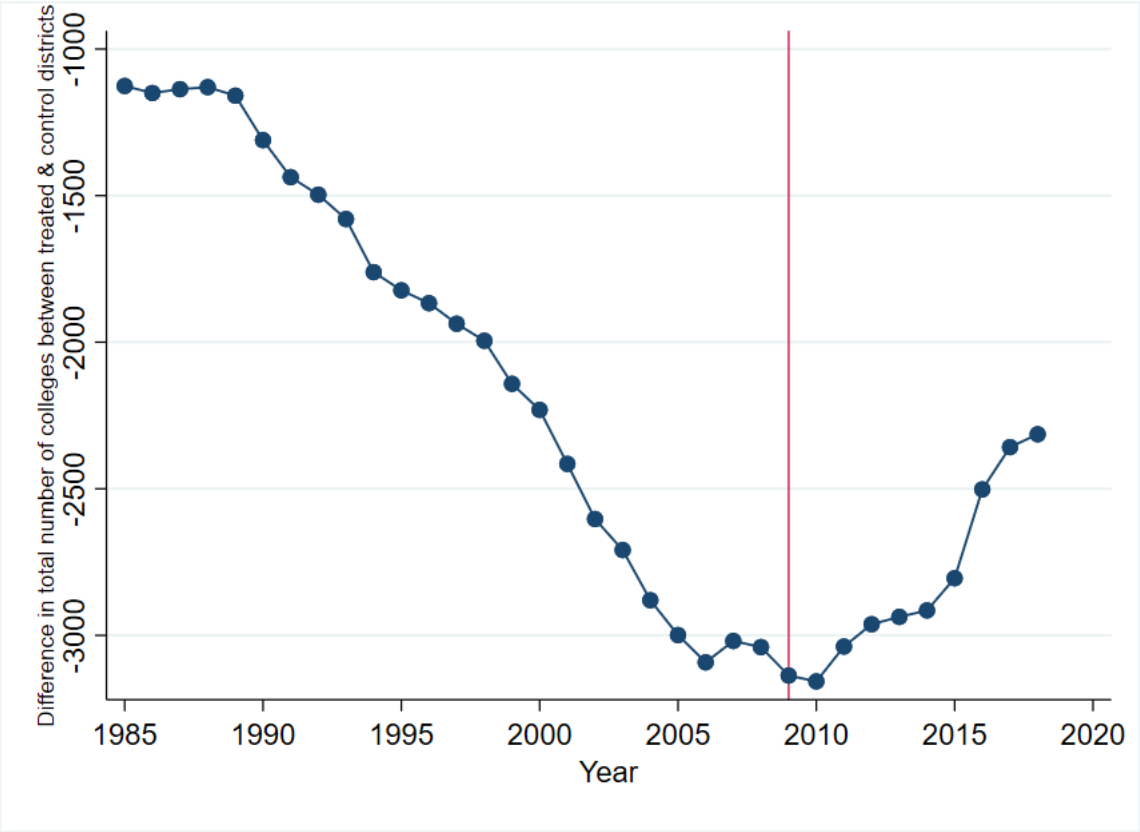


Figure A.4: Difference in the Total Stock of Colleges between the Treated and Control Districts

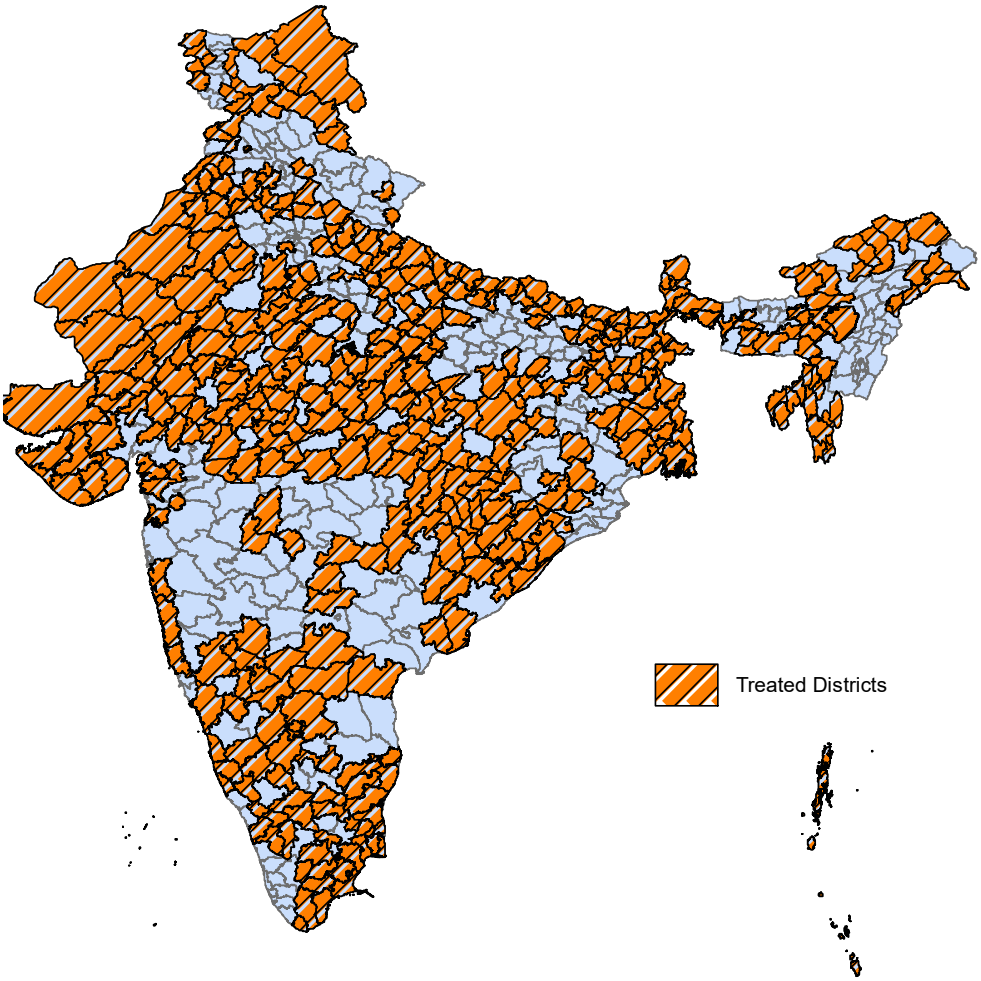
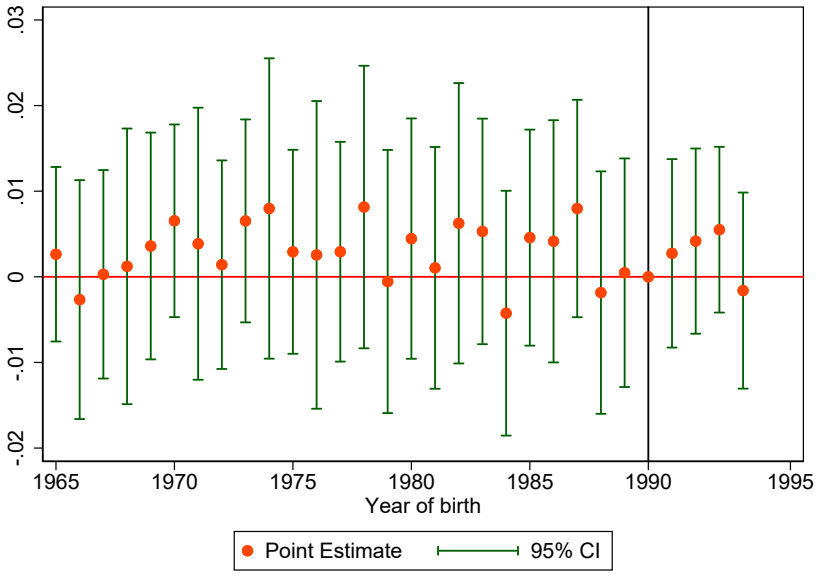


Figure A.5: Treated Districts

Note: Treated districts are colored orange while the control districts are colored grey.

(a) Migration for education



(b) Migration for work

Figure A.6: Event Study of the Effect of Grant Policy on Out-Migration

Notes: The figure plots the annual coefficients on treatment from a Difference-in-Difference regression analogous to the form described in Equation 1. The dependent variable in Panel (a) and Panel (b) is the likelihood of migrating into another district for education and work, respectively.

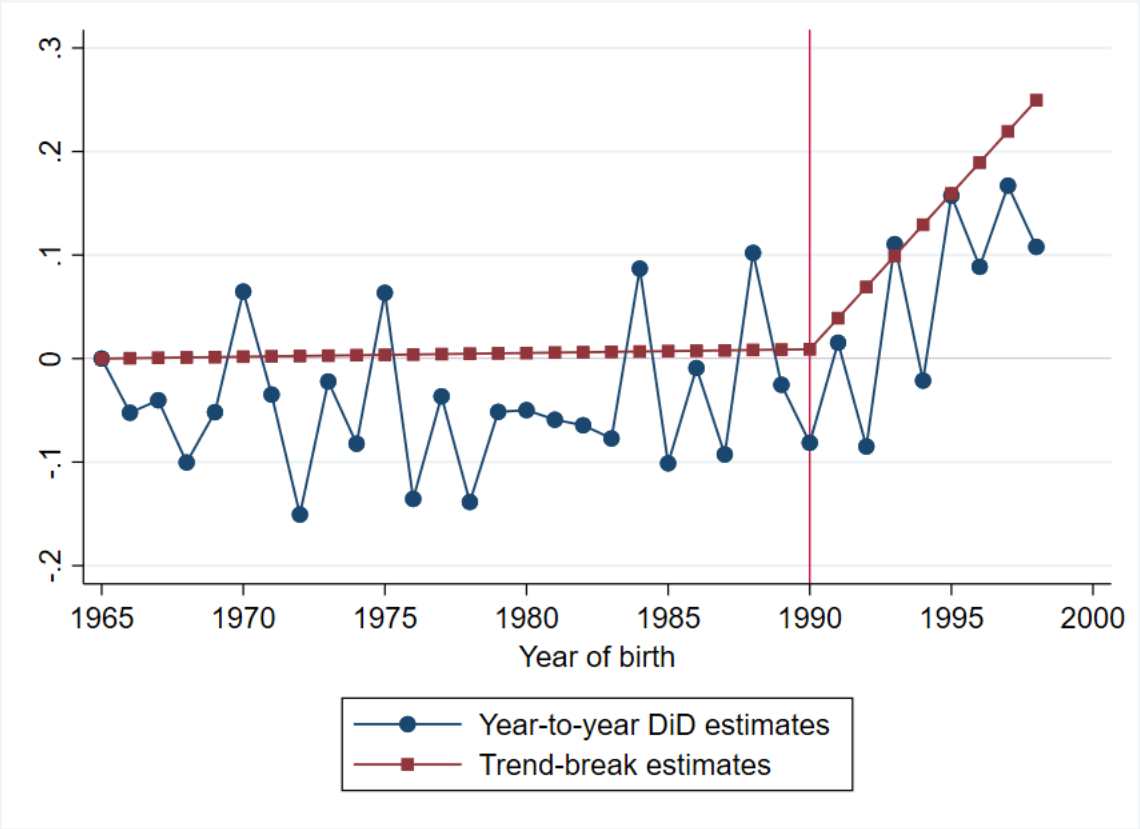


Figure A.7: Female Labor Market Participation by Treatment Status

Notes: The series "Year-to-year DID estimates" graphs the annual coefficients on treatment from a Difference-in-Difference regression analogous to the form described in Equation 1. The series "Trend break estimates" graphs the annual coefficient implied by the trend break model in Column 1 of Table 9. In both cases, the dependent variable is the likelihood that a married woman participates in the labor market.

Table A.1: Covariate Balance between Treatment and Control Districts

VARIABLES	Control mean	Difference between Treatment and Control	Standard error of difference	Observations
Panel A: Household Data				
Respondents' profile				
Years of schooling	7.47	-1.386***	0.235	40,466
Highest level of education is primary	0.12	0.039***	0.010	40,466
Highest level of education is secondary	0.41	-0.024	0.019	40,466
Enrolls into college	0.20	-0.078***	0.011	40,466
Finishes college	0.18	-0.071***	0.011	40,466
Age at marriage	19.27	-0.377*	0.225	27,147
Marriage before 18 years of age	0.37	0.037*	0.020	27,147
Member of workforce	0.40	-0.007	0.024	6,978
Participate in agricultural work	0.17	0.013	0.020	6,978
Participate in non-agricultural work	0.23	-0.021	0.017	6,978
Respondents' husbands' profile				
Years of schooling	7.33	-0.884***	0.299	5,017
Highest level of education is primary	0.14	0.026	0.017	5,017
Highest level of education is secondary	0.51	-0.037*	0.022	5,017
Enrolls into college	0.13	-0.036**	0.015	5,017
Member of workforce	0.96	-0.020**	0.009	4,951
Participate in agricultural work	0.39	0.053*	0.029	4,951
Participate in non-agricultural work	0.57	-0.073**	0.030	4,951
Panel B: District-level Information from 2001 Census				
Male-female ratio	0.93	0.005	0.005	578
Percentage of SC/ST population	0.15	0.006	0.007	578
Percentage literate	0.59	-.09***	0.009	578
Gross enrollment ratio in college	18.08	-9.93***	0.412	578
Labor force participation rate	0.39	.036***	0.005	578

Notes: Data for Panel A is from the NFHS survey. The analysis is only for the pre-policy years, i.e., cohorts born in or before 1990. Only a subset of the surveyed women were asked about their employment and (if married) their husbands' education and employment. Data for Panel B comes from the 2001 Census of India.

Table A.2: College Expansion as a Function of 2001 GER

	Number of colleges constructed				
VARIABLES	(1)	(2)	(3)	(4)	(5)
GER in 2001 x (T-1985) trend	0.00476***	0.00347**	0.00338**	0.00357**	0.00324**
	(0.00155)	(0.00135)	(0.00149)	(0.00151)	(0.00145)
GER in 2001 x (T-2009) trend	-0.0167***	-0.0136***	-0.0142***	-0.0148***	-0.0135***
	(0.00446)	(0.00429)	(0.00440)	(0.00446)	(0.00428)
GER in 2001 x Post	-0.00788	-0.00398	0.00583	0.00488	0.00551
	(0.0140)	(0.0118)	(0.0139)	(0.0140)	(0.0140)
Control	None	Night lights	Poverty gap	Squared poverty gap	Head count ratio
Observations	20,955	20,955	20,394	20,394	20,394
R-squared	0.445	0.450	0.457	0.455	0.460
F-stat 1	13.48	9.924	11.07	11.54	10.50
Sig. level	0.000261	0.00671	0.00415	0.000726	0.00545
F-stat 2	7.055	5.044	5.535	5.773	5.256
Sig. level	0.000933	0.00171	0.000931	0.00328	0.00126

Notes: Data is from the 2018 AISHE. The unit of analysis is a district-year. GER in 2001 is the gross enrollment ratio based on the 2001 Census of India. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients γ_1 and γ_2 in Equation 2 is 0. F-stat 2 reports the F statistic from a test of the joint significance of γ_1 and γ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.3: RDD Estimates of the Effect of the Grants Policy on Cumulative Colleges per 100,000 Population

	Cumulative number of colleges per 100,000 population							
VARIABLES	2008	2009	2010	2011	2012	2013	2014	2015
Panel A: Polynomial of order 1								
Treatment	-0.206	-0.237	-0.259	-0.277	-0.289	-0.302	-0.278	-0.319
	(0.342)	(0.355)	(0.362)	(0.371)	(0.389)	(0.399)	(0.404)	(0.430)
Observations	213	222	231	231	231	231	234	243
R-squared	0.007	0.007	0.007	0.006	0.006	0.007	0.009	0.013
Bandwidth	2.758	2.911	2.996	3.044	3	3.027	3.130	3.189
Panel B: Polynomial of order 2								
Treatment	-0.026	-0.030	-0.058	-0.107	-0.089	-0.084	-0.117	-0.158
	(0.269)	(0.290)	(0.302)	(0.308)	(0.322)	(0.326)	(0.330)	(0.351)
Observations	334	342	334	331	334	342	350	354
R-squared	0.020	0.022	0.018	0.013	0.015	0.019	0.020	0.023
Bandwidth	4.532	4.645	4.532	4.462	4.537	4.658	4.755	4.825
Panel C: Polynomial of order 3								
Treatment	-0.073	-0.108	-0.130	-0.107	-0.150	-0.140	-0.130	-0.187
	(0.337)	(0.368)	(0.380)	(0.394)	(0.414)	(0.425)	(0.431)	(0.465)
Observations	288	289	286	283	284	282	283	283
R-squared	0.017	0.014	0.011	0.011	0.009	0.010	0.011	0.011
Control Mean	2.679	2.813	2.901	2.959	3.017	3.067	3.149	3.203
Bandwidth	3.819	3.823	3.811	3.722	3.759	3.706	3.747	3.736
Control mean	2.679	2.813	2.901	2.959	3.017	3.067	3.149	3.203

Notes: The college data is from the 2018 AISHE. The unit of analysis is a district-year. In each regression we use optimal bandwidth as proposed by Calonico et al. (2017). Treatment is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.4: Effect of the Grants Policy on Number of Years of Schooling

VARIABLES	(1) Number of years of schooling	(2) Number of years of schooling for individuals who have not attended college	(3) Number of years in college
Grant policy x Overall trend	-0.0115	-0.00217	-0.0128***
	(0.0126)	(0.00908)	(0.00377)
Grant policy x Post Policy trend	0.103***	0.0383*	0.0664***
	(0.0274)	(0.0225)	(0.0121)
Grant policy x Post	0.251*	0.134	-0.0305
	(0.135)	(0.150)	(0.0528)
Observations	110,467	88,690	110,467
R-squared	0.443	0.415	0.298
F-stat 1	18.69	3.394	31.10
Sig. level	1.79e-05	0.184	2.45e-07
F-stat 2	9.388	1.697	15.59
Sig. level	9.59e-05	0.0659	3.63e-08

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions control for household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.5: Effect Size of Grants Policy on College Construction and College Enrollment

	Effect Size	8-year effect	Levels over 8 years
Increase in College Capacity	0.192	1.54 colleges	9,659 seats
Increase in enrollment for Women (Full Unrestricted sample)	0.0142	An additional 1,195 women enroll each year	9,556 students
Enrollment (weighted average for non-movers and in-migrants)			
(a) Increase in Female College Enrollment (Non-movers:Women who always lived in the same place)	0.021	An additional 289 women enroll every year	2,311 students
(b) Increase in Female College Enrollment (in-migrants-Women who did not always live in the same place)	0.011	An additional 739 women enroll every year	5,912 students
(c) Increase in Male College Enrollment (Non-movers:Men who always lived in the same place)	0.006	An additional 336 men enroll every year	2,690 students
(d) Increase in Male College Enrollment (in-migrants-men who did not always live in the same place)	0.012	An additional 302 men enroll every year	2,414 students
Total [(a) +(b) +(c) + (d)]			13,327 students

Notes: Estimates are based on authors' calculation.

Table A.6: Effect of the Grants Policy on Men's Educational Attainment (Full Sample)

VARIABLES	(1) No of years of schooling	(2) Highest level of education is primary	(3) Highest level of education is Secondary	(4) Enrolls into college	(5) Finishes college	(6) Number of years in college
Panel A: With socio-economic controls						
Treatment x Overall trend	0.0219***	1.86e-05	0.00277***	4.10e-05	0.000189	0.000627
	(0.00524)	(0.000379)	(0.000574)	(0.000429)	(0.000394)	(0.00160)
Treatment x Post Policy trend	0.0677***	-0.00114	-0.00654**	0.00850***	0.00553**	0.0216***
	(0.0219)	(0.00156)	(0.00264)	(0.00238)	(0.00217)	(0.00780)
Treatment x Post	-0.195	-0.00974	0.0229	-0.0286**	-0.00367	-0.0274
	(0.124)	(0.00914)	(0.0158)	(0.0134)	(0.0123)	(0.0438)
Observations	104,036	104,036	104,036	104,036	104,036	104,036
F-stat 1	17.88	0.546	2.264	14.59	8.105	9.703
Sig. level	2.70e-05	0.460	7.23e-06	0.000147	0.00456	0.00192
F-stat 2	17.93	0.274	12.06	7.921	5.141	5.910
Sig. level	2.65e-08	0.760	0.133	0.000400	0.00610	0.00286
Panel B: Without socio-economic controls						
Treatment x Overall trend	0.0207***	1.11e-05	0.00276***	-7.10e-06	0.000142	0.000427
	(0.00592)	(0.000386)	(0.000581)	(0.000461)	(0.000423)	(0.00171)
Treatment x Post Policy trend	0.0699***	-0.00106	-0.00685***	0.00876***	0.00573***	0.0224***
	(0.0228)	(0.00158)	(0.00265)	(0.00240)	(0.00220)	(0.00790)
Treatment x Post	-0.163	-0.0107	0.0243	-0.0275**	-0.00245	-0.0219
	(0.136)	(0.00916)	(0.0156)	(0.0138)	(0.0127)	(0.0454)
Observations	104,471	104,471	104,471	104,471	104,471	104,471
F-stat 1	17.37	0.472	2.676	15.25	8.380	10.07
Sig. level	3.51e-05	0.790	9.57e-06	0.000104	0.00392	0.00158
F-stat 2	15.50	0.236	11.77	8.145	5.109	5.941
Sig. level	2.67e-07	0.492	0.102	0.000322	0.00629	0.00278

Notes: Data is from the NFHS man's survey. The unit of analysis is at the individual level. **The analysis is for full sample of men (both who always lived in the same place and who did not).** Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.7: Effect of the Grants Policy on Educational Attainments with Different Controls

VARIABLES	(1)	(2)	(3)	(4)	(5)
Panel A: Number of years of schooling					
Grant policy x Overall trend	-0.0134	-0.0104	-0.00930	-0.00899	-0.00907
	(0.0147)	(0.0150)	(0.0149)	(0.0150)	(0.0150)
Grant policy x Post Policy trend	0.0988***	0.0727**	0.103***	0.103***	0.103***
	(0.0342)	(0.0343)	(0.0353)	(0.0353)	(0.0353)
Grant policy x Post	0.321*	0.354**	0.308*	0.301*	0.299*
	(0.166)	(0.174)	(0.171)	(0.171)	(0.171)
Observations	110,467	109,837	107,024	107,024	107,024
R-squared	0.288	0.289	0.290	0.290	0.290
F-stat 1	10.61	5.540	12.09	12.21	12.21
Sig. level	0.00118	0.0189	0.000542	0.000509	0.000509
F-stat 2	10.61	5.540	12.09	12.21	12.21
Sig. level	0.00118	0.0189	0.000542	0.000509	0.000509
Panel B: College enrollment					
Grant policy x Overall trend	-0.00360***	-0.00281***	-0.00363***	-0.00358***	-0.00358***
	(0.00105)	(0.00106)	(0.00109)	(0.00111)	(0.00111)
Grant policy x Post Policy trend	0.0207***	0.0164***	0.0217***	0.0218***	0.0218***
	(0.00356)	(0.00359)	(0.00367)	(0.00368)	(0.00369)
Grant policy x Post	-0.0421***	-0.0245	-0.0449***	-0.0456***	-0.0457***
	(0.0160)	(0.0165)	(0.0166)	(0.0166)	(0.0166)
Observations	110,467	109,837	107,024	107,024	107,024
R-squared	0.194	0.196	0.198	0.198	0.197
F-stat 1	34.19	20.58	35.75	36.34	36.37
Sig. level	7.99e-09	6.82e-06	3.78e-09	2.85e-09	2.80e-09
F-stat 2	34.19	20.58	35.75	36.34	36.37
Sig. level	7.99e-09	6.82e-06	3.78e-09	2.85e-09	2.80e-09
Control	**None**	**Night light**	**Poverty gap**	**Squared poverty gap**	**Head count ratio**

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.8: Effect of the Grants Policy: Sensitivity to Different Sample Periods

	Number of colleges constructed			
VARIABLES	(1)	(2)	(3)	(4)
Grant Policy x Overall trend	-0.0396***	-0.0497***	-0.0450*	-0.0725**
	(0.0140)	(0.0185)	(0.0242)	(0.0356)
Grant Policy x Post-Policy trend	0.182***	0.192***	0.187***	0.215***
	(0.0479)	(0.0513)	(0.0558)	(0.0659)
Grant Policy x Post	-0.0246	0.0600	0.0299	0.163
	(0.191)	(0.201)	(0.215)	(0.245)
Observations	24,130	20,955	17,780	14,605
R-squared	0.420	0.444	0.480	0.510
F-stat 1	13.42	13.37	13.29	13.18
Sig. level	0.000269	0.000277	0.000289	0.000305
F-stat 2	7.208	7.049	6.649	6.593
Sig. level	0.000803	0.000938	0.00139	0.000305
Sample Period	1981-2017	1985-2017	1990-2017	1995-2017

Notes: Data is from 2018 AISHE. The unit of analysis is a district-year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients γ_1 and γ_2 in Equation 2 is 0. F-stat 2 reports the F statistic from a test of the joint significance of γ_1 and γ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.9: Changes in Political Affiliation by Treatment Status

	Affiliation with ruling alliance	
VARIABLES	(1)	(2)
Grant Policy x Overall trend	0.00787	0.0189
	(0.0154)	(0.0148)
Grant Policy x Post Policy trend	-0.00611	-0.0176
	(0.0185)	(0.0161)
Grant Policy x Post	-0.0159	-0.0156
	(0.0228)	(0.0237)
Control	None	State-specific trend
Observations	6,686	6,686
R-squared	0.370	0.560
F-stat 1	0.0680	0.0634
Sig. level	0.794	0.801
F-stat 2	0.222	0.883
Sig. level	0.801	0.414

Notes: Data is from the Election Commission of India. The sample period is 2005-2017. The unit of analysis is a district-year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. All regressions include district and year fixed effects. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.10: Changes in Political Affiliation and Migration by Treatment Status

	Affiliation with ruling alliance		Migrated in last 5 years	
VARIABLES	(1)	(2)	(3)	(4)
Grant Policy x Post	-0.00743	0.00564	0.00710	0.00709
	(0.0308)	(0.0281)	(0.00680)	(0.00481)
Post	-0.0445*	-0.0517**	-0.0257***	-0.0257***
	(0.0252)	(0.0222)	(0.00606)	(0.00429)
Constant	0.387***	0.386***	0.0318***	0.0318***
	(0.00581)	(0.0133)	(0.00156)	(0.00110)
Control	District-specific trend	State-specific trend	HH FEs	District FEs
Observations	6,686	6,686	77,092	77,092
R-squared	0.725	0.558	0.532	0.045

Notes: Political affiliation data is from the Election Commission of India. The sample period is 2005-2017. The unit of analysis is a district by year. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for years 2009 and after and 0 otherwise. Columns 1 and 2 present estimated coefficients from the difference-in-difference estimates of the intervention (college construction grants) on the fraction of total constituencies in a district that are affiliated with the ruling alliance. Columns 1 and 2 control for district fixed effects. Migration data is from two rounds of IHDS survey. Columns 3 and 4 present estimated coefficients from the difference-in-difference estimates of the intervention (college construction grants) on the probability of in-migration in last five years. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.11: Effect of the Grants Policy on Out-Migration

VARIABLES	(1) Migration for employment	(2) Migration for education
Grant policy x Overall trend	-3.32e-05	-0.000205**
	(0.000126)	(9.94e-05)
Grant policy x Post Policy trend	-0.000721	0.000863
	(0.00139)	(0.00319)
Grant policy x Post	0.00223	-0.00329
	(0.00430)	(0.00940)
Observations	54,504	54,504
R-squared	0.015	0.032
F-stat 1	0.303	0.0425
Sig. level	0.582	0.120
F-stat 2	0.198	2.136
Sig. level	0.821	0.837

Notes: Data is from 2012 IHDS tracking sheet data. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1994 (18 years or older in 2012–the year of the survey). The post trend goes from cohort birth years 1991 to 1994. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.12: Spillover Analysis: Effect of the Grants Policy on Educational Attainment of Girls

VARIABLES	(1) No of years of schooling	(2) Highest level of education is primary	(3) Highest level of education is Secondary	(4) Enrolls into college	(5) Finishes college
Treatment x Overall trend	0.0370**	0.000933	0.00440**	-0.000543	-0.000310
	(0.0186)	(0.000924)	(0.00193)	(0.00180)	(0.00184)
Treatment x Post Policy trend	0.0882**	-0.00196	-0.0176***	0.0190***	0.0121*
	(0.0419)	(0.00190)	(0.00606)	(0.00579)	(0.00643)
Treatment x Post	0.388	0.00565	0.128***	-0.0651**	-0.0207
	(0.260)	(0.0103)	(0.0336)	(0.0273)	(0.0252)
Neighboring controls x Overall trend	0.0621***	0.00147	-0.000214	0.00369*	0.00330*
	(0.0200)	(0.000923)	(0.00202)	(0.00188)	(0.00191)
Neighboring controls x Post Policy trend	-0.0232	-0.000724	0.000748	-0.00259	-0.00595
	(0.0438)	(0.00187)	(0.00634)	(0.00612)	(0.00667)
Neighboring controls x Post	0.164	-0.000507	0.0684**	-0.0234	-0.0205
	(0.271)	(0.00998)	(0.0342)	(0.0285)	(0.0263)
Observations	110,467	110,467	110,467	110,467	110,467
R-squared	0.439	0.069	0.203	0.285	0.257
F-stat 1	11.26	0.388	6.439	14.54	5.126
Sig. level	0.000840	0.645	0.921	0.000339	0.192
F-stat 2	7.693	0.697	4.464	8.090	3.433
Sig. level	0.000500	0.498	0.0114	0.000151	0.0329
F-stat 3	0.997	0.212	0.00980	0.0452	0.239
Sig. level	0.318	0.252	0.0119	0.832	0.0239
F-stat 4	5.304	1.382	0.00796	2.482	1.656
Sig. level	0.00519	0.533	0.992	0.0844	0.625

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. Treatment is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Neighboring control is a dummy variable that takes a value of 1 if the control district is a neighbor to any treatment district. The omitted category is non-neighboring control districts. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in the augmented version of Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. F-stat 3 and F-stat 4 are the corresponding test statistics (analogous to F-stat 1 and F-stat 2, respectively) for the coefficients on the interactions of the neighboring controls dummy with the overall and the post policy trends. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.13: Effect of the Grants Policy on Husband's Education and Employment

VARIABLES	(1) No of years of schooling	(2) Enrolled into college	(3) Member of workforce	(4) Participate in agricultural work	(5) Participate in non-agricultural work
Panel A: All women					
Grant policy x Overall trend	-0.0255	-0.00360**	0.000872	0.00126	-0.000393
	(0.0250)	(0.00153)	(0.000975)	(0.00199)	(0.00193)
Grant policy x Post Policy trend	-0.0605	0.00695	0.0181*	-4.05e-05	0.0182
	(0.149)	(0.0105)	(0.00990)	(0.0142)	(0.0144)
Grant policy x Post	1.404*	0.0757	-0.0709*	0.00441	-0.0753
	(0.770)	(0.0508)	(0.0424)	(0.0667)	(0.0715)
Observations	6,074	6,074	6,075	6,075	6,075
R-squared	0.250	0.151	0.188	0.301	0.289
F-stat 1	0.326	0.105	3.645	0.00769	1.600
Sig. level	0.568	0.746	0.0567	0.930	0.448
F-stat 2	0.615	2.784	2.069	0.209	0.804
Sig. level	0.541	0.0626	0.127	0.812	0.206
Panel B: Women with at least a secondary education					
Grant policy x Overall trend	-0.0230	-0.00505*	-0.000426	0.00104	-0.00146
	(0.0316)	(0.00286)	(0.00147)	(0.00304)	(0.00298)
Grant policy x Post Policy trend	0.213	0.0264*	0.00881	-0.00912	0.0179
	(0.174)	(0.0159)	(0.0128)	(0.0170)	(0.0181)
Grant policy x Post	0.0546	0.0391	-0.0310	-0.0141	-0.0169
	(0.736)	(0.0747)	(0.0563)	(0.0819)	(0.0898)
Observations	3,066	3,066	3,089	3,089	3,089
R-squared	0.375	0.310	0.296	0.334	0.343
F-stat 1	1.234	1.851	0.429	0.230	0.849
Sig. level	0.267	0.0843	0.513	0.836	0.357
F-stat 2	0.881	2.486	0.264	0.180	0.548
Sig. level	0.415	0.174	0.768	0.632	0.578

Notes: Data is from the NFHS woman's survey. Information on the husband's education and employment was collected for a subset of all surveyed women. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. No socio-economic controls are included. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.14: Effect of the Grants Policy on Male Labor Force Participation

VARIABLES	(1) All	(2) Urban	(3) Rural
Panel A: With socio-economic controls			
Grant Policy x Overall trend	0.000489	0.000652	0.000261
	(0.000385)	(0.000827)	(0.000405)
Grant Policy x Post Policy trend	-7.06e-05	-0.00282	0.00163
	(0.00229)	(0.00426)	(0.00263)
Grant Policy x Post	0.0352**	0.0348	0.0303*
	(0.0152)	(0.0324)	(0.0176)
Observations	72,934	18,125	54,809
R-squared	0.371	0.405	0.372
F-stat 1	0.0359	0.270	0.553
Sig. level	0.850	0.643	0.575
F-stat 2	0.878	0.442	0.554
Sig. level	0.416	0.603	0.457
Panel B: Without socio-economic controls			
Grant policy x Overall trend	0.000624	0.000902	0.000288
	(0.000381)	(0.000818)	(0.000405)
Grant policy x Post Policy trend	8.33e-05	-0.00283	0.00166
	(0.00231)	(0.00434)	(0.00264)
Grant policy x Post	0.0318**	0.0289	0.0300*
	(0.0152)	(0.0328)	(0.0176)
Observations	73,255	18,207	55,048
R-squared	0.359	0.392	0.364
F-stat 1	0.101	0.206	0.579
Sig. level	0.750	0.650	0.447
F-stat 2	1.500	0.700	0.628
Sig. level	0.224	0.497	0.534

Notes: Data is from the NFHS man's survey. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.15: Effect of the Grants Policy on Fertility and Child Preference

VARIABLES	(1) Total children ever born	(2) Number of living children	(3) Ideal number of children	(4) Son prefernece	(5) Knows about modern contraceptive method	(6) Uses modern contraceptive method
Panel A: With socio-economic controls						
Grant Policy x Overall trend	0.00331	0.00420	-0.00254	-0.000183	-7.53e-05	0.000376
	(0.00517)	(0.00482)	(0.00250)	(0.000731)	(0.000410)	(0.00112)
Grant Policy x Post Policy trend	-0.00514	-0.00133	0.00228	0.00282**	0.000778	0.00490
	(0.0153)	(0.0141)	(0.00403)	(0.00142)	(0.000849)	(0.00563)
Grant Policy x Post	-0.144**	-0.141**	-0.0334	-0.0177**	-0.00418	0.00480
	(0.0701)	(0.0674)	(0.0271)	(0.00864)	(0.00465)	(0.0264)
Observations	34,482	34,482	108,855	108,673	110,467	34,457
R-squared	0.415	0.415	0.325	0.069	0.101	0.231
F-stat 1	0.0170	0.0489	0.00637	4.841	0.853	0.864
Sig. level	0.896	0.662	0.596	0.0281	0.356	0.636
F-stat 2	0.208	0.413	0.518	2.441	0.451	0.453
Sig. level	0.812	0.825	0.936	0.0879	0.637	0.353
Panel B: Without socio-economic controls						
Grant policy x Overall trend	0.00130	0.00235	-0.00249	-0.000181	-8.67e-05	0.000358
	(0.00560)	(0.00526)	(0.00254)	(0.000736)	(0.000413)	(0.00111)
Grant policy x Post Policy trend	-0.00478	-0.00173	0.00264	0.00287**	0.000763	0.00493
	(0.0156)	(0.0142)	(0.00420)	(0.00145)	(0.000848)	(0.00562)
Grant policy x Post	-0.174**	-0.169**	-0.0373	-0.0183**	-0.00374	0.00340
	(0.0712)	(0.0688)	(0.0277)	(0.00875)	(0.00468)	(0.0266)
Observations	34,482	34,482	108,855	108,673	110,467	34,457
R-squared	0.327	0.322	0.312	0.064	0.095	0.229
F-stat 1	0.0622	0.00235	0.00202	4.873	0.788	0.864
Sig. level	0.803	0.902	0.618	0.0276	0.375	0.353
F-stat 2	0.0522	0.103	0.481	2.451	0.426	0.448
Sig. level	0.949	0.961	0.964	0.0871	0.653	0.639

Notes: Data is from the NHFS woman's survey. Columns (1), (2), and (6) are based on the ever-married women sample, whereas Columns (3), (4) and (5) are based on the whole sample. The unit of analysis is at the individual level. Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.

Table A.16: Effect of the Grants Policy on Educational Attainment of Girls

VARIABLES	(1) No of years of schooling	(2) Highest level of education is primary	(3) Highest level of education is Secondary	(4) Enrolls into college	(5) Finishes college	(6) Number of years in college
Panel A: With socio-economic controls						
Treatment x Overall trend	-0.0157***	0.00204***	0.00137**	-0.00200***	-0.00187***	-0.00703***
	(0.00543)	(0.000364)	(0.000536)	(0.000353)	(0.000324)	(0.00130)
Treatment x Post Policy trend	0.130***	-0.00490***	-0.00239	0.0111***	0.0128***	0.0448***
	(0.0192)	(0.000967)	(0.00168)	(0.00157)	(0.00179)	(0.00647)
Treatment x Post	-0.00573	0.00198	0.0379***	-0.0261***	-0.0170***	-0.0676***
	(0.0707)	(0.00447)	(0.00864)	(0.00714)	(0.00655)	(0.0244)
Observations	540,306	540,306	540,306	540,306	540,306	540,306
F-stat 1	43.36	10.05	0.445	44.59	49.35	47.36
Sig. level	9.49e-11	1.54e-09	0.0368	0	0	0
F-stat 2	22.82	20.95	3.320	25.94	25.49	24.04
Sig. level	2.68e-10	0.00159	0.505	5.28e-11	0	8.61e-11
Panel B: Without socio-economic controls						
Treatment x Overall trend	-0.0176***	0.00204***	0.00114**	-0.00198***	-0.00186***	-0.00694***
	(0.00583)	(0.000366)	(0.000547)	(0.000361)	(0.000331)	(0.00133)
Treatment x Post Policy trend	0.143***	-0.00504***	-0.00186	0.0116***	0.0132***	0.0467***
	(0.0201)	(0.000977)	(0.00170)	(0.00158)	(0.00181)	(0.00657)
Treatment x Post	0.00918	0.00227	0.0398***	-0.0264***	-0.0172**	-0.0687***
	(0.0791)	(0.00451)	(0.00876)	(0.00743)	(0.00683)	(0.0254)
Observations	540,306	540,306	540,306	540,306	540,306	540,306
F-stat 1	50.17	10.99	0.220	49.44	52.71	50.89
Sig. level	0	0.000970	0.112	0	0	0
F-stat 2	25.64	21.02	2.193	27.12	26.65	25.48
Sig. level	0	1.44e-09	0.639	0	0	0

Notes: Data is from the NFHS woman's survey. The unit of analysis is at the individual level. **The sample consists of women who did not always stay in the same place.** Grant Policy is a dummy variable taking a value of 1 if a district is eligible for the college construction grants, i.e., the GER is lower than the national average in 2001. Post is an indicator variable that takes the value 1 for individuals born in or after 1991 (18 years or younger in 2009, when the intervention took place) and 0 otherwise. The overall trend goes from cohort birth year 1965 to cohort birth year 1998 (18 years or older in 2016–the year of the survey). The post trend goes from cohort birth years 1991 to 1998. The variable *number of years in college* is a categorical variable which takes the value of zero if an individual does not attend college, and values between one and five for individuals attending college corresponding to the year they are in college (1, 2, and 3 for the first, second, and third year of college, respectively; 4 for year one in Masters and 5 for year two in Masters or beyond). All regressions include district and year fixed effects. Socio-economic controls include household assets, number of family members, rural/urban dummy, and caste. F-stat 1 reports the F statistic from a test that the sum of coefficients λ_1 and λ_2 in Equation 3 is 0. F-stat 2 reports the F statistic from a test of the joint significance of λ_1 and λ_2. Standard errors are clustered at the district level. *** indicates significance at 1, ** at 5, and * at 10 percent level.