Medical Colleges and Front-line Health Workers

Fulvia Budillon* and Sabareesh Ramachandran †

September 15, 2024

Abstract

Proximity to providers is essential for non-tradable services like health and education. However, the distribution of providers per capita is skewed towards cities with better amenities. Training the locals to become providers could be more effective than moving qualified providers to remote areas. In this paper we study the impact of opening a medical college in a district on the access to care, take-up of care, and in turn on health outcomes in the same district. A two-way fixed effects regression is used to study the 5x growth in medical colleges in India between 1980 and 2020. One additional batch of students graduating from a college in the district is associated with a 4.3pp increase in health facilities and health workers. This increase is almost entirely in the private sector in urban areas. This leads to an increase in healthcare visits by pregnant women. However, we do not see any significant improvements in morbidity or mortality outcomes. Reforming public hiring to post some of the additional health workers in well-equipped public health facilities may be a promising strategy to make improved access translating into better health outcomes.

*UCSD. Email: fbudillo@ucsd.edu

†University of Chicago. Email: sabareesh@uchicago.edu

1 Introduction

Proximity to service providers is essential to access non-tradable services like education and healthcare. However, high-skilled teachers and doctors would prefer to work in areas with better amenities leading to a skewed distribution of providers per capita within a country. For instance, the number of doctors per 1000 people varies from 0.5 to 2 across different states in India[1]. This is one potential reason for poor health and education outcomes in rural areas (Banerjee et al., 2004). Governments could offer incentives to work in remote areas or mandate rural service, but these are very expensive and only partly effective[2]

In this paper we study an alternative policy option: to set up training centres in the remote areas and equip the locals to serve there themselves. We study the 5x growth in number of medical colleges across India between 1980 and 2020, many of which were the first to open in that district. We look at the impact of opening a medical college on the number of health facilities and health workers in that district, and in turn on the take-up of care, and health outcomes. In a fully integrated labor and education market with free mobility the location of the college would have no effect on the distribution of doctors. But proximity to the colleges and lower commuting costs may induce more locals to enroll in these colleges and in turn serve in their home districts. Moreover, during their medical education, all students may build connections in the district and find a job locally more easily[3].

We combine multiple datasets to construct a district x year level panel to estimate the impacts using a two-way fixed effect (TWFE) specification. We use a public dataset from National Medical Commission (NMC) to determine the location and year of opening of every medical college in India. Four rounds of the Economic Census data are used to count the

[1]Data from Indian Medical Register (of India, 2018). See Figure 1. This is at the state level, there are further disparities across districts within a state, and across villages within a district.

[2]In Odisha, India, the government offers incentives equal to the salary to attract doctors to remote districts (Hindu, 2015). Even students who have signed a contract to serve in rural areas in China have a low willingness to do so (Hu et al., 2023). Even if providers are mandated to serve in rural areas, we see a higher level of absenteeism among them due to their long commute from home (Chaudhury et al., 2006).

[3]Trust among patients is essential for doctors to practise independently. They would not be able to move to a new city and start practising.

number of health facilities and the number of employees in these facilities (including doctors, nurses, and other personnel). Four rounds of the Demographic Health Survey (DHS) are used to estimate the take-up of care by pregnant women and health outcomes for infants and adults. We define the treatment variable as the number of batches of students who have graduated from a medical college in the district. We also control for the time-invariant average treatment effects due to the college opening and for potential spillover effects on nearby districts.

We find that one additional batch of students graduating from a medical college in a district is associated with 31.5 additional health facilities in the same district. This is a large effect of 4.3% relative to a mean of 722 facilities in districts where colleges hadn't yet opened in 1990. This effect is in addition to a time-invariant effect of 200 facilities when the college opens. The number of employees at health facilities increases by 145 with each graduating batch[4]. The increase is driven by those providing western allopathic medicine with no effects on number providing alternative systems of medicine. This is consistent with the fact that the colleges only train students in western medicine.We find no impact on the number of public facilities or public providers, as all the increase is observed in the private sector. There has been no change in the public hiring process through which this could adjust. Nearly all the additional providers locate themselves in the urban areas where they might be able to command a higher price.

The significant increase in the number of providers and facilities in these districts should lead to an increase in the capacity of local healthcare, which might increase the take-up of care. We see an average 0.016 additional antenatal care (ANC) visits to pregnant women in the district (on a base of 2.48 visits in 1990) each year with a batch graduating. We see a significant increase in the share of ANC visits done by doctors and a reduction in the share by nurses. The share of deliveries assisted by doctors also increases by 0.2 pp each year and the share assisted by nurses decreases. While the ANC visits increase among rural and urban women, the share of deliveries assisted by doctors increases only among urban woman.

However, the increased access of care does not apply to all health services, and it does not lead to significant health improvements. We see a significantly lower probability of in-

[4]Each batch consists of 100 students on average

stitutional delivery and vaccination, perhaps due to doctors being able to assist deliveries at home. There is a small but not significant increase in infant and adult mortality rates. There is a marginal but not significant improvement in morbidity measures like incidence of cough, fever, and diarrhea. This suggests that better access to providers alone may not improve infant health outcomes, which may require complementary improvements in other inputs and hygiene (Weaver et al., 2024).

We show that these results are not sensitive to the assumption that the treatment effects are proportional to the number of batches graduated. Allowing for a non-parametric evolution of treatment effects in time we find qualitatively similar results. We also show that the results are robust to multiple alternative definitions of the spillover terms. While the interpretation is not necessarily causal, the spillovers are negligible beyond a distance of 50 km. We also find similar results when we ignore the spillovers and use the estimator proposed by Callaway and Sant'Anna (2021). We also show parallel trends in the pre-treatment periods between districts where a college opens earlier and where it opens later. While the opening of the college itself seems to have an effect on the number of providers, the impact of students graduating is much larger.

This paper contributes to the literature on distribution of high-skilled workers like doctors. Better amenities in urban areas, and agglomeration benefits influence high-skilled workers preference to be in cities (Arntz et al., 2023; Diamond, 2016; Jeworrek and Brachert, 2022). This is a problem in the case of non-tradable services like healthcare and has been long discussed (Newhouse et al., 1982; Rosenthal et al., 2005). Dingel et al. (n.d.) suggests there may be increasing returns to scale in provision of advanced care and hence such a distribution may be efficient. However, it is desirable to make primary care accessible to all. Inducing doctors to serve in rural areas through incentives and mandates are only partly effective. The required wage differential is very large making this expensive (Holte et al., 2015; Scott et al., 2013; Swami and Scott, 2021). In this paper we explore a second policy option: we show that opening colleges in remote areas could lead to an increase in health workers in those areas. This takes advantage of potential heterogeneous location preferences in the population, as students studying in remote colleges may be more willing to serve there.

This paper also relates to a literature on the local impacts of higher educational institutions. Across countries, when a college opens nearby, we find improvements in college enrollment among students and in turn labor market returns (Frenette, 2006, 2009; Lapid, 2018). Jagnani and Khanna (2020) also see more enrollment in primary and secondary schools and a crowd in of services like electricity and roads. Large university expansions in China and Vietnam have also been found to have lead to an increase in firm productivity, capital investments in firms, and structural transformation towards high-skill industries in the neighbourhood. This leads to greater returns for the educated (Vu and Vu-Thanh, 2024; Che and Zhang, 2018). In this paper we show that setting up medical colleges has large externalities on access to and take up of healthcare locally. These improvements likely benefit all people in those districts and need to be factored in to the returns to starting such colleges.

This paper also contributes to an active policy debate on what governments can do to increase concentration of providers in remote areas. The government in South Korea is increasing the number of medical seats in existing colleges hoping to achieve this while doctors are protesting this move claiming it would only lead to more doctors in the urban areas (Pacific, 2024). In India, the National Medical Commission has proposed new medical colleges only be opened in states with few colleges [5]. Muralidharan (2024) suggests that hiring public workers locally may make them more regular to their workplace and be accountable to the population they serve. In this paper we show that this is possible: opening higher educational institutions does lead to more providers in that district. Reforming the public hiring process like creating district-level cadres will allow the state to capitalize on their presence in the district.

2 Context: Medical Education in India

After high school, students who want to pursue a medical career can enroll in a Bachelor in Medicine and Bachelor in Surgery (commonly referred to by the abbreviation MBBS). The usual duration of an MBBS degree is five and half years and includes a year residential

[5] The move was intended to balance the distribution of medical colleges across the country. The proposal is to cap the number of seats to 1 per 10,000 population.

internship. After completing this degree the student is eligible to practice allopathic medicine. There are degrees in other forms of medicine like ayurveda, unani, siddha, and homeopathy. These courses are usually offered in other colleges that specialize in these schools of medicine.

2.1 Medical colleges

There are about 700 medical colleges in India as of 2023. The initial medical colleges were started by the state or central government. Subsequently non-profit trusts and charitable societies were allowed to set up colleges. The principle that medical education cannot be a profit-making venture prevented private bodies from starting colleges until this was relaxed in 2017 (Mahal and Mohanan, 2006).

The process of setting up a private medical college is lengthy and cumbersome. The college needs to be attached to an existing public or private hospital where students would receive practical training. The hospital needs to have at least 300 beds and the capacity to increase number of beds based on the number of students enrolled. The private entities need to own a land at least 20 acres big where the college can be set up (10 acres for big cities). They need to have required infrastructure for the college, the attached teaching hospital and residential quarters for students. They also need to provide bank guarantees to establish their ability to run a college. They first need to get an approval from the university they wish to be affiliated with and the state government where the college is located. They can then submit the request to the central government. The request is forwarded to the Medical Council of India (now renamed National Medical Commission) who conduct the technical scrutiny of the application and request clarifications from the applicant. Once all requirements are met the private entity finally receives the approval to start the college. The entire process could take multiple years.

2.2 Medical admissions

Process of admissions to medical colleges were determined by individual state governments. Starting in the 1980s, they conducted a common entrance exam for the colleges in the state. Students scoring above the cutoff for a college were granted admission to that college. 80%

of the seats were reserved for the domiciles of that state. While domicile rules vary across states, usually the student should have completed a few years of education in the state or one of their parents should be a resident of the state for them to be considered a domicile. In states with multiple medical colleges, the students get to choose the college they study and proximity to their homes is one of the important factors in this choice. Since 2016 a centralized exam conducted across the country is used to determine admissions to medical colleges. However, the domicile restrictions were retained as before. 15% of the seats in private colleges are also reserved under a management quota. Admissions to these seats are at the discretion of the college administration.

Fees to pursue a medical degree in public college are relatively low (about 500 USD per year) while fees in private colleges are very high (about 30,000 USD per year). In addition to the fees, students' costs include living expenses near their college or commuting costs to and from the college. Hence students might choose to study in colleges closer to their homes to reduce these costs.

2.3 Medical jobs in India

After completing their MBBS degrees students may pursue a post graduation degree or other specialized degrees. The number of post graduation seats in the country is less than half of the number of undergraduate seats. Admissions to the post-graduate seats are based on another entrance test conducted nationally.

Public recruitment of doctors usually happens at the state level. State governments conduct exams to award jobs. Doctors that are awarded the job are added to the state public service cadre, and they may be assigned to any location in the state. Public doctors are also often transferred between different locations.

With limited public recruitment, many doctors work in private hospitals and clinics, or set up their own private practice. Jobs in the private sector are usually obtained through connections the student can leverage besides their performance in their degree. Trust among their potential patients is essential for a doctor to set up a private practice. This makes it difficult for a doctor to go to a new city and set up a clinic. Hence they are more likely to return to their home district to start a clinic or work in a hospital.

3 Data

Colleges: We use the list of medical colleges approved by the National Medical Council. This public database contains the address of the college, the year it received the approval, number of seats, and the administration type (public, private, or trust). We assume that the year of approval is the year the first batch of students were admitted to that college. We also assume that no medical colleges have closed down in this period. We assign the colleges to the district they are located in in the year 2011.

The number of medical colleges increases from 112 in 1980 to 555 in 2020. Of the 112 colleges in 1980, 99 were government-run colleges and 13 were privately managed. Of the 555 colleges in 2020, 275 were public colleges and 280 were privately managed. Hence, of the 443 colleges that opened between 1980 and 2020, 40% are government run and 60% are privately managed.

Figure 2 presents the distribution of these colleges across the country in 1980 and 2020. While in 1980 there are colleges across the country, most of them are located in the state capitals and in the bigger cities with few colleges in smaller cities and remote districts. The number of districts with a medical college increased from 89 in 1980 to 294 in 2020.

Health facilities and workers: We use four rounds of the Economic Census to compile outcome measures. The Economic Census is the complete count of all entrepreneurial units located in the country involved in any economic activities. We use the three digit NIC codes of each establishment to count the number providing health services. This includes all organizations providing health and medical services like hospitals, dispensaries, sanatoria, nursing homes, and maternal and child welfare clinics. We further use the four digit codes to classify these institutions by the type of medicine practiced: allopathic, ayurvedic, unani, homeopathy, and others. We group the last four categories as non-allopathic for convenience.

The census data also includes the number of employees in each facility. We do not see any characteristics of these employees other than their gender. The employees may be doctors, nurses, pharmacists, accountants, sanitary staff, or any others who work in these facilities. We argue that the total number of workers in the facility is a reasonable proxy for the number of health workers in the district. We aggregate all census data up to the district level.

Healthcare take-up and outcomes: The DHS is a nationally representative household survey that provides data on population, health, and nutrition in developing countries[6]. We use four rounds of the DHS to study outcomes related to healthcare take-up and health status. These are the first, second, fourth, and fifth rounds administered in 1992-93, 1998-99, 2015-16, and 2019-21 respectively. We do not include data from the third round since it does not contain district identifiers.

The respondents are also asked if they suffered from fever, cold, cough, or diarrhoea in the last 12 months[7]. We calculate the mortality rate in the entire population as the share of the sample households reported dead in each year. Since the survey asks about deaths in the preceding 5 years, we are able to calculate the adult mortality rates for each of the 5 years preceding the survey. The women respondents are also asked about all their pregnancies in the last 5 years. They are asked about the number of ante-natal care home visits, the type of providers during these visits, the vaccines given to the infant, and whether the infant survived. We compute the infant mortality rate each year following the component death probabilities method used to construct demographic indicators from the DHS survey data (Elkasabi, 2019).

4 Econometric Strategy

We are interested in studying the impact of opening medical colleges in a district on the number of health workers in the same district. We capitalize on the staggered opening of medical colleges across the country to estimate the treatment effects using a two-way fixed effect specification. The unit of analysis is district x year. In Figure 3 we see the staggered opening of the first college in a district across four decades.

[6] The DHS is called NFHS in India and is administered by an independent organization called Indian Institute of Population Sciences.

[7] The respondents are also asked for the number of times they visited a health facility for care in the last 12 months. However, this question is framed differently in different rounds. Hence we do not include this here.

In the TWFE specification we account for the fact that the number of health workers is a stock variable that accumulates over time post treatment. We also acknowledge that the effects are not limited to the same district and likely have a spillover effect on neighboring districts within the same state.

We estimate the parameters in the following specification:

$$y_{dt} = \theta_t + \eta_d + \beta batches_{dt} + collegeopen_{dt} + \gamma spillovers_{dt} + \varepsilon_{dt} \tag{1}$$

where y_{dt} is the outcome of interest, θ_t are year fixed effects, and η_d are district fixed effects. $batches_{dt}$ is the treatment variable that is equal to the number of batches graduated from a college in that district. We assume that the treatment effects aggregate over time linearly here. In alternative specifications we also estimate the heterogeneous impact over time non-parametrically using indicators for 0-5 years, 5-10 years, 10-15 years and >15 years since treatment. $collegeopen_{dt}$ is an indicator that is 1 if there is a college in the district. This variable captures the time-invariant average effect of opening a college. While this term is also a treatment effect, we do not emphasise this in the results since it is a more mechanical effect of opening a large institution in the district and less interesting economically.

$spillovers_{dt}$ is a variable that captures the total spillovers of a college opening on all other districts. This is defined as the number of batches graduated in colleges in other districts weighted by the inverse of the distance from that college. It is defined as:

$$spillovers_{dt} = \sum_{\text{colleges } c} \frac{\mathbb{1}[state(c) = state(d)\ \&\ district(c) \neq d] * \mathbb{1}[opening(c) < t] * (t - opening(c))}{(distance(district(c), d))^2}$$

Here, we make three assumptions about the spillovers: (i) the spillovers are restricted to the districts in the same state of the college and decay with the square of the distance from the college; (ii) the spillover effects aggregate linearly over time (or are proportional to the number of batches of students that have graduated); (iii) the total spillovers on a district is the sum of spillovers from all colleges in the state. In alternative specifications we vary these assumptions in the spillover definition.

We exclude districts which are always treated (have a college in the first time period) and never treated (don't have a college even in 2020). Besides the above functional form assumptions, we also assume parallel trends between counterfactual outcomes in the districts with a college and the observed outcome in the districts without a college as yet conditional on the spillover and fixed effects.

We test this assumption by comparing trends for districts where colleges opened earlier with districts where colleges opened later. We focus on the subset of colleges that opened between 2000 and 2020. We compare districts where a college first opened between 2000 and 2010 with districts where a college first opened between 2010 and 2020. We compare the outcomes of these districts in the previous periods using the following specification:

$$y_{dt} = \theta_t + \eta c_d^{2000-10} + \beta_t c_d^{2000-10} + \nu_{dt} \tag{2}$$

where $c_d^{2000-10}$ is an indicator that is 1 if a college first opened in that district in 2000-10 and 0 otherwise.

We claim that under these assumptions β is the treatment effect of opening a medical college in that district. It is the difference in potential outcomes of that district if that particular college was not opened versus if that college was opened holding all other college openings the same.

The interpretation of the spillovers term is however difficult. The districts receiving a higher value of the spillovers may not have parallel trends with the districts receiving a lower value of the spillovers. For example, the districts receiving a higher spillover may be those near bigger cities where many colleges open in this period. The districts closer to the cities may not be comparable with the districts further away. Hence we do not analyse the coefficients of the spillover term in this paper. We include them only to remove the bias in the treatment effect due to the spillovers.

Robustness: Callaway and Sant'Anna estimator

To ensure the results are robust to the specification, we also use the estimation procedure proposed by Callaway and Sant'Anna (2021). We define the dependent variable as an indicator of whether there is a medical college in the district in that year. We estimate the heterogeneous treatment effects over time. Consider the data generating process given by:

$$y_{dt} = \theta_t + \eta_d + \sum_{\ell} \beta_\ell \mathbf{1}\left\{t - g_d = \ell\right\} + v_{dt}$$

where y_{dt} is the outcome of interest in district d in year t, θ_t are year fixed-effects, η_d are district fixed effects, g_d is the year in which district d had the first medical college opening. The β_ℓ parameters are the average treatment effect ℓ years from the college opening.

We again use the not-yet-treated districts as our comparison group. For the Economic Census outcomes, we define a time period as 8 years long. Hence a college may have opened in any of those 8 years. We estimate the average treatment effects in the next period in all such districts.

5 Results

We first look at the impacts on the outcomes from the Economic Census. We estimate the coefficient β in the specification in equation 1. In Table 1 we see the results on the number of health facilities in the district. One additional batch of students graduating from a college in the district is associated with 31.5 additional health facilities in the district. This is a large effect relative to the mean of 722.1 facilities in 1990 in districts with no medical colleges yet. We see that most of these additional facilities are private facilities, practicing allopathic medicine in urban areas. We see an increase of 26.9 urban health facilities, 31.4 private health facilities, and 33.1 allopathic health facilities with each batch of graduating students. There is a direct effect of opening a college as well-we see an average 199.6 additional facilities in the district before any batches graduate.

In Table 2 we see the treatment effects on the number of people working in health facilities in the district. This is the sum of number of employees working in all health facilities in the district. One additional batch of students graduating from a medical college in the district is associated with 144.7 additional employees at health facilities. 74.7 of these workers are male and 70 are female. The additional workers are mostly in the private sector (138.4) and in urban areas (141.4).

In Table 3 we look at the outcomes from DHS surveys in India. In the first column we see a small but significant increase of 0.016 in the number of ANC visits made to pregnant women. We also see a slight decline in the share of institutional childbirth in the district. This may be due to doctors travelling to the homes to assist in deliveries. Perhaps as a consequence, we see a lower probability of the mother receiving a tetanus injection before birth.

In Table 4 we see a small but not significant increase in the share of ANC visits attended by doctors and a decrease in the share attended by nurses and others. This suggests a substitution between nurses and doctors happening here. Similarly, in Table 5 we see a significant 0.2pp increase in the share of deliveries assisted by doctors and a decrease in the share assisted by nurses.

In Appendix Table 8 we see that the increase in ANC visits is mostly rural areas. These visits can be scheduled in advance and hence the provider or the woman can travel. However, the increase in share of deliveries assisted by doctors is primarily in urban areas.

In the first four columns of Table 6 we see a small but insignificant decline in the incidence of diseases like cough and fever. In the last two columns we see the effects on infant and adult mortality. We rule out even a small decline in infant mortality rate. While not significant, the point estimates are positive.

Parallel pre-trends: In Table 7 we look at the pre-trends between the districts where a college first opened between 2000 and 2010 and districts where a college first opened between 2010 and 2020. These coefficients are estimated according to the specification in equation 2. We see that the coefficients in the pre-period of 1998 are small and insignificant across all variables. The coefficients in 2005 are also small although some districts may have had a college opened by then. In figure 4 we see the same coefficients for the DHS outcome variables as well. Here again we see that the coefficients until 2000 are small and mostly insignificant.

6 Discussion

In this paper we find that opening a medical college leads to a significant increase in the number of health facilities and health workers in the area. This in turn leads to an increase in the take up of healthcare like ANC visits. However we reject even small changes in health outcomes like infant mortality rate and incidence of fever and diarrhoea.

The census data does not allow us to distinguish between health clinics run by quacks and those managed by qualified doctors. However, we do see more ANC visits conducted by doctors. We also see an increase in facilities providing western medicine and not alternate systems of medicine. Both of these suggest the additional health workers in the district are not all quacks. An increase in the number of qualified doctors could lead to a substitution of care sought from quacks with care sought from doctors. Further, Das et al. (2022) show that the quality of care provided by quacks is better in places where quality of care provided by qualified providers is better, perhaps driven by competition. Hence even if the increase in number of doctors is only a fraction of the observed treatment effects, there can be a large increase in access to quality care.

In this version of the paper, we are unable to identify the mechanisms at play here. Students who live near these colleges may choose to attend them and then practice near their home districts. Students from across the state may come to the college and choose to stay in the same district for work. We will be able to separate these mechanisms using additional data available with the government.

Reducing disparities in the distribution of health providers is of considerable policy interest in both developing and developed countries. Government of South Korea announced in 2024 that they would be increasing the number of medical seats to have sufficient doctors serving in the remote areas of the country. In India, a new regulation restricts opening of new medical colleges to states with a doctor per 1000 people ratio below 1.

However, disparities in the distribution exist at multiple levels. In India, in figure 1 we see the disparities across states. Within each state though, the districts containing big cities may have more doctors. Further within each district, the doctors may live in more urban areas. While the opening of colleges in remote districts can reduce the former two imbalances, the last one persists. We see that the increase in providers and facilities is concentrated in the urban areas of the district. While people in the rural areas can travel about 15-20 kms to reach the urban areas, the severely ill may not be able to do so.

Improving health outcomes is very difficult. The marginal patients seeking care may not be the severely ill. Further, Weaver et al. (2024) show that there are complementarities between factors like healthcare, nutrition, and sanitation in how they affect outcomes. Hence improving any one factor alone may not move the needle significantly.

More specifically, there may be complementarities between provider knowledge and the equipment and infrastructure at the facilities. While we see the largest increase in *private* providers and facilities, the *public* facilities are better funded and may have better equipment. Assigning the additional providers to public facilities in rural areas may hence be more effective.

One reason this may not be happening currently is due to the structure of public hiring in India. State governments hire doctors at the state level and may assign them to any facility in any district in the state. This mechanism does not take advantage of the doctors' preferences to be in some districts than others. As a result most doctors vie for postings in the state capital. The state could establish district-level cadres of doctors, recruiting and appointing them within the same district. This approach might offer significant benefits for the medical professionals, as it provides the certainty of employment within a single district. Consequently, doctors would not need to relocate their families and could confidently invest in their local communities. In this paper we show that colleges expand the pool of candidates available in the district. The increased availability of qualified doctors directly supports the feasibility and success of a strategy of local recruitment and retention of health professionals.

References

Arntz, Melanie, Eduard Brüll, and Cäcilia Lipowski, "Do preferences for urban amenities differ by skill?," *Journal of Economic Geography*, May 2023, *23* (3), 541–576.

Banerjee, Abhijit, Angus Deaton, and Esther Duflo, "Wealth, Health, and Health Services in Rural Rajasthan," *American Economic Review*, April 2004, *94* (2), 326–330.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers, "Missing in Action: Teacher and Health Worker Absence in Developing Countries," *Journal of Economic Perspectives*, February 2006, *20* (1), 91–116.

Che, Yi and Lei Zhang, "Human Capital, Technology Adoption and Firm Performance: Impacts of China's Higher Education Expansion in the Late 1990s," *The Economic Journal*, 2018, *128* (614), 2282–2320. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecoj.12524.

Das, Jishnu, Benjamin Daniels, Monisha Ashok, Eun-Young Shim, and Karthik Muralidharan, "Two Indias: The structure of primary health care markets in rural Indian villages with implications for policy," *Social Science & Medicine (1982)*, May 2022, *301*, 112799.

Diamond, Rebecca, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *The American Economic Review*, 2016, *106* (3), 479–524. Publisher: American Economic Association.

Dingel, Jonathan I, Joshua D Gottlieb, Maya Lozinski, and Pauline Mourot, "Market Size and Trade in Medical Services," *Working Paper.*

Elkasabi, Mahmoud, "Calculating fertility and childhood mortality rates from survey data using the DHS. rates R package," *PloS one*, 2019, *14* (5), e0216403.

Frenette, Marc, "Too Far to Go On? Distance to School and University Participation," *Education Economics*, March 2006. Publisher: Routledge.

– , "Do universities benefit local youth? Evidence from the creation of new universities," *Economics of Education Review*, June 2009, *28* (3), 318–328.

Hindu, The, "Incentives for doctors serving in remote areas of Odisha," *The Hindu*, April 2015.

Holte, Jon Helgheim, Trine Kjaer, Birgit Abelsen, and Jan Abel Olsen, "The impact of pecuniary and non-pecuniary incentives for attracting young doctors to rural general practice," *Social Science & Medicine*, March 2015, *128*, 1–9.

Hu, Dan, Baisong Zhang, Mingyu Huang, Min Liu, Xiulong Xia, Yanli Zuo, and Xiaoyun Liu, "Evaluation of a medical education policy with compulsory rural service in China," *Frontiers in Public Health*, February 2023, *11*, 1042898.

Jagnani, Maulik and Gaurav Khanna, "The effects of elite public colleges on primary and secondary schooling markets in India," *Journal of Development Economics*, September 2020, *146*, 102512.

Jeworrek, Sabrina and Matthias Brachert, "Where to go? High-skilled individuals' regional preferences," *IWH Discussion Paper*, 2022.

Lapid, Patrick A, "Expanding College Access: The Impact of New Universities on Local Enrolments," *Job Market Paper, University of California, Berkeley*, 2018.

Mahal, Ajay and Manoj Mohanan, "Growth of private medical education in India," *Medical Education*, October 2006, *40* (10), 1009–1011.

Muralidharan, Karthik, *Accelerating India's Development* 2024.

Newhouse, Joseph P., Albert P. Williams, Bruce W. Bennett, and William B. Schwartz, "Does the Geographical Distribution of Physicians Reflect Market Failure?," *The Bell Journal of Economics*, 1982, *13* (2), 493–505. Publisher: [RAND Corporation, Wiley].

of India, Govt, "Medical Infrastructure in India," July 2018.

Pacific, The Lancet Regional Health – Western, "Junior doctor strikes in South Korea: more doctors are needed?," *The Lancet Regional Health: Western Pacific*, March 2024, *44*, 101056.

Rosenthal, Meredith B, Alan Zaslavsky, and Joseph P Newhouse, "The Geographic Distribution of Physicians Revisited," *Health Services Research*, December 2005, *40* (6 Pt 1), 1931–1952.

Scott, Anthony, Julia Witt, John Humphreys, Catherine Joyce, Guyonne Kalb, Sung-Hee Jeon, and Matthew McGrail, "Getting doctors into the bush: General Practitioners' preferences for rural location," *Social Science & Medicine*, November 2013, *96*, 33–44.

Swami, Megha and Anthony Scott, "Impact of rural workforce incentives on access to GP services in underserved areas: Evidence from a natural experiment," *Social Science & Medicine*, July 2021, *281*, 114045.

Vu, Khoa and Tu-Anh Vu-Thanh, "Higher Education Expansion, Labor Market, and Firm Productivity in Vietnam," *Working Paper*, 2024.

Weaver, Jeffrey, Sandip Sukhtankar, Paul Niehaus, and Karthik Muralidharan, "Cash Transfers for Child Development: Experimental Evidence from India," January 2024.

7 Tables

	Dep. Var. : Number of Health Facilities						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Rural	Urban	Public	Private	Allopathic	Non-Allop.
Graduate batches	31.5***	4.6*	26.9***	0.3	31.4***	33.1***	-1.6
	(4.2)	(2.7)	(2.6)	(1.3)	(3.8)	(3.8)	(2.8)
College opening	199.6***	69.5***	130.1***	22.8**	174.9***	127.0***	72.5***
	(30.4)	(20.0)	(19.1)	(9.3)	(27.2)	(27.3)	(20.3)
Spillover	37.1***	4.1	33.0***	-5.0*	42.1***	29.3***	7.8
	(9.1)	(6.0)	(5.7)	(2.8)	(8.2)	(8.2)	(6.1)
Observations	2,054	2,054	2,054	2,054	2,054	2,054	2,054
Mean Dep. Var.	722.1	431.6	290.5	126.3	595.2	551.9	170.3

Table 1: Impact of colleges on health facilities

This table presents the results on the number of health facilities and by type of facilities. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Number of Health Workers						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Female	Male	Rural	Urban	Public	Private
Graduate batches	144.7***	70.0***	74.7***	3.2	141.4***	6.3	138.4***
	(24.2)	(8.7)	(20.2)	(19.5)	(13.7)	(5.4)	(23.2)
College opening	1288.5***	762.5***	526.0***	326.3**	962.2***	134.2***	1154.2***
	(175.9)	(63.0)	(146.3)	(141.4)	(99.3)	(38.9)	(168.6)
Spillover	167.4***	66.4***	101.0**	32.6	134.8***	-9.4	176.9***
	(52.7)	(18.9)	(43.8)	(42.4)	(29.8)	(11.7)	(50.5)
Observations	2,054	2,054	2,054	2,054	2,054	2,054	2,054
Mean Dep. Var.	2159.0	661.6	1497.3	1025.5	1133.4	756.6	1402.4

Table 2: Impact of colleges on workers in health facilities

This table presents the results on the number of workers employed in the health facilities. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Health Services Uptake				
	(1)	(2)	(3)	(4)	(5)
	ANC visits	First ANC	Birth in Facility	Tetanus Injection Before Birth	Tetanus Injection Before Pregnancy
Graduate batches	0.016***	0.003	-0.004***	-0.014***	0.004
	(0.005)	(0.003)	(0.001)	(0.002)	(0.009)
College opening	0.144***	0.010	-0.023***	-0.113***	-0.052
	(0.049)	(0.028)	(0.005)	(0.015)	(0.059)
Spillover	0.083***	0.051***	-0.018***	-0.081***	0.033
	(0.015)	(0.008)	(0.002)	(0.004)	(0.034)
Observations	8,763	8,599	9,078	8,764	4,497
Mean Dep. Var.	3.648	3.753	0.573	1.686	1.149

Table 3: Impact of colleges on take-up of healthcare

This table presents the results on different measures of healthcare take-up. The first column is the number of ANC visits during their pregnancy. The second column is the month of the first ANC visit. The third column is an indicator of whether the child was delivered in a facility. The fourth and fifth columns are indicators for whether the mother received a tetanus injection. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Share of Antenatal Care Visits					
	(1) Doctor	(2) Nurse/Anm	(3) TBA	(4) CHW	(5) Aganwadi	(6) Asha
Graduate batches	0.001	-0.004***	0.000	-0.000	-0.001	-0.003***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
College opening	0.009	-0.019**	-0.003	-0.002*	-0.011	-0.002
	(0.009)	(0.009)	(0.006)	(0.001)	(0.009)	(0.008)
Spillover	-0.005*	-0.008***	-0.005**	0.005***	-0.014***	-0.015***
	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.002)
Observations	8,764	8,764	8,764	4,760	6,880	6,880
Mean Dep. Var.	0.502	0.359	0.017	0.006	0.117	0.089

Table 4: Impact of colleges on share of ANC visits by provider type

This table presents the results on share of ANC visits assisted by different providers. Multiple providers may be present during the same visit. The providers in each column are doctors, nurses/ANM, traditional birth attendant, community health worker, anganwadi worker, and ASHA worker respectively. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Share of Deliveries Assisted					
	(1) Doctor	(2) Anm	(3) TBA	(4) Friends	(5) Other	(6) None
Graduate batches	0.002***	-0.006***	0.002***	0.006***	0.000	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
College opening	0.012*	-0.077***	-0.021***	0.010	-0.005	-0.004
	(0.006)	(0.007)	(0.007)	(0.007)	(0.004)	(0.003)
Spillover	0.006***	-0.026***	0.001	0.013***	-0.001	0.002**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Observations	8,764	8,764	8,764	8,764	6,880	8,764
Mean Dep. Var.	0.407	0.453	0.242	0.366	0.040	0.016

Table 5: Impact of colleges on share of deliveries assisted by provider type

This table presents the results on share of deliveries assisted by different providers. Multiple providers may be present during the delivery. The providers in each column are doctors, nurses/ANM, traditional birth attendant, community health worker, anganwadi worker, and ASHA worker respectively. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Morbidity and Mortality					
	(1)	(2)	(3)	(4)	(5)	(6)
	Diarrhea	Fever	Cough	Diabete	Infant Mortality	Adult Mortality
Graduate batches	-0.035	-0.043	-0.088	0.007	0.016	0.003
	(0.058)	(0.071)	(0.076)	(0.034)	(0.019)	(0.002)
College opening	-0.038	0.332	1.176	0.169	-0.130	0.050***
	(0.558)	(0.684)	(0.737)	(0.218)	(0.171)	(0.016)
Spillover	0.021	-0.010	-0.023	-0.135	-0.001	0.024***
	(0.165)	(0.202)	(0.218)	(0.120)	(0.050)	(0.005)
Observations	2,228	2,228	2,228	1,153	9,725	6,423
Mean Dep. Var.	11.481	18.042	19.487	2.017	5.054	0.689

Table 6: Impact of colleges on health status

This table presents the results on mobidity and mortality for adults. All columns are in units of percentage of population. The first four columns are indicators for whether the respondent ever had that condition in the last one year. Columns 5 and 6 are respectively the infant mortality rate and the adult mortality rate measured using the DHS. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Test for Parallel pre-Trends					
	(1)	(2)	(3)	(4)	(5)	(6)
	Health Workers	Health Facilities	HF: Urban	HF: Rural	HF: Public	HF: Private
1990	0.0	0.0	0.0	0.0	0.0	0.0
	(.)	(.)	(.)	(.)	(.)	(.)
1998	-275.3	-146.8	-8.3	-138.5	-49.0	-97.8
	(539.4)	(215.5)	(89.1)	(154.4)	(34.1)	(203.4)
2005	367.9	65.4	24.1	41.4	-31.9	94.1
	(538.7)	(215.2)	(89.0)	(154.2)	(34.1)	(203.1)
2013	1791.5***	97.5	185.6**	-88.1	-18.5	116.0
	(538.7)	(215.2)	(89.0)	(154.2)	(34.1)	(203.1)
Observations	628	628	628	628	628	628

Table 7: Pre-trends in outcome variables

This table present the trends in outcome variables between districts where a college first opened in 2000-10 and districts where a college first opened in 2010-20. Each column shows the trend in a different outcome variable. The estimates are of the parameters β_τ in Equation 2.

8 Figures

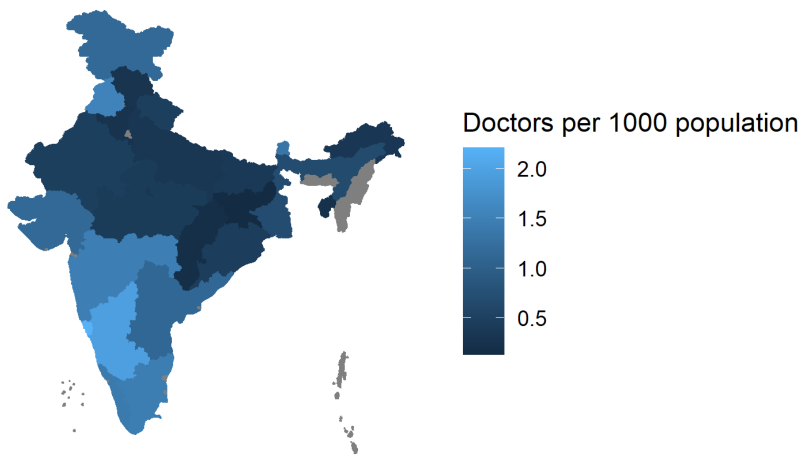
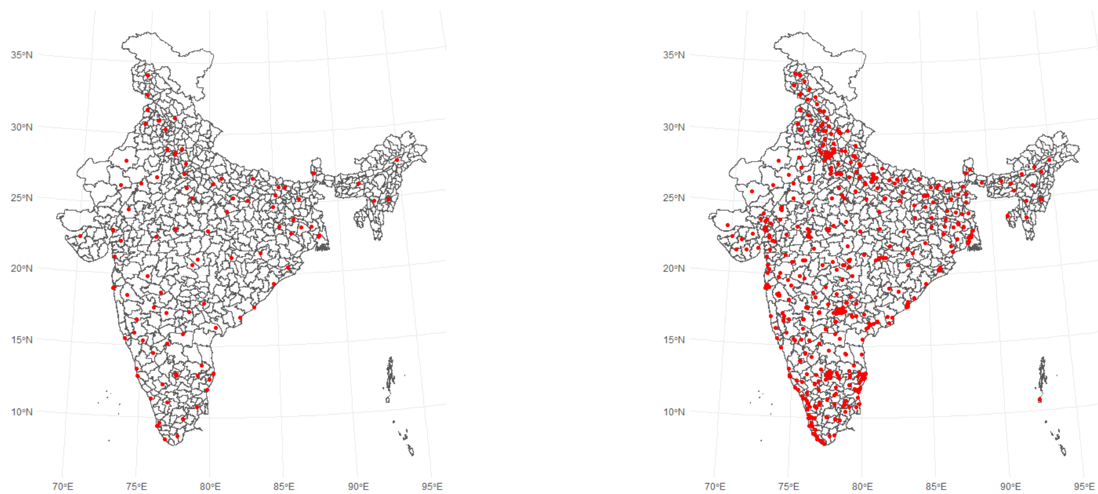
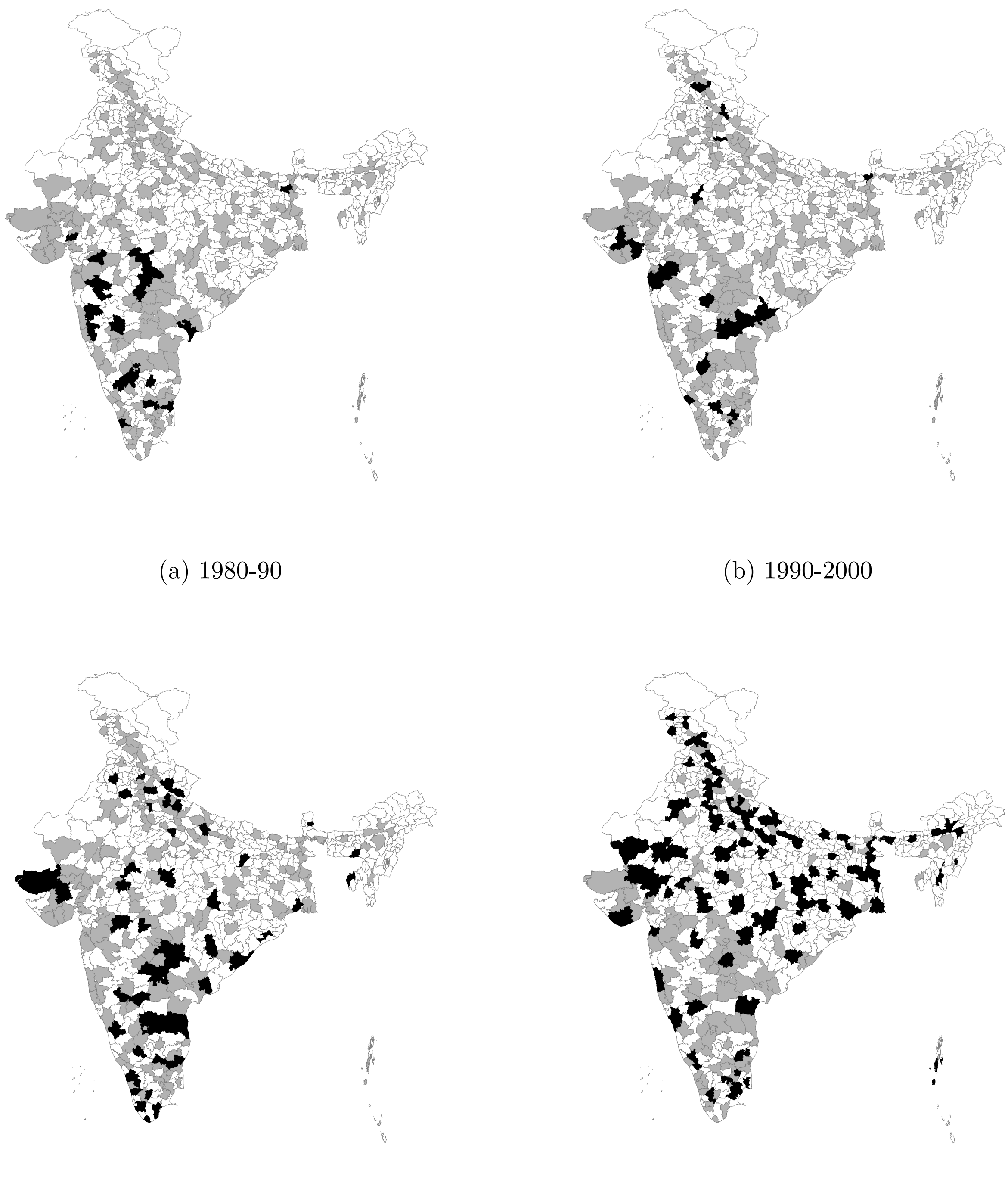


Figure 1: Doctors per 1000 population across states



(a) 1980 (b) 2020

Figure 2: Location of medical colleges across the country.



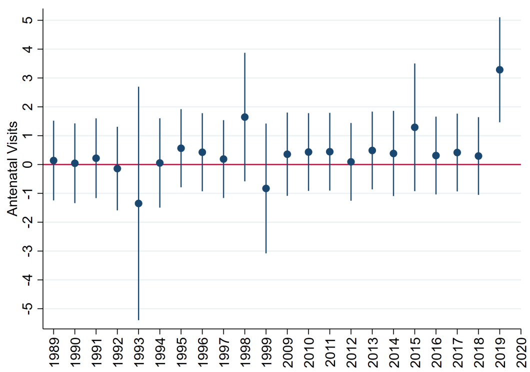
(a) 1980-90

(b) 1990-2000

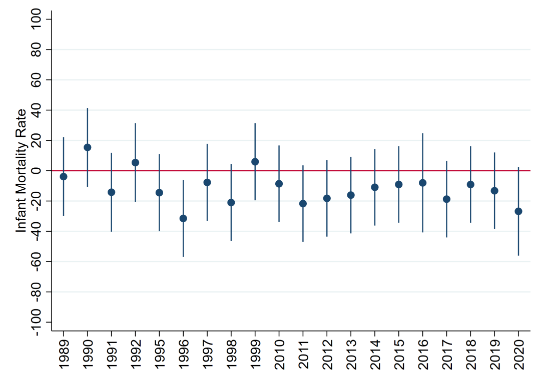
(c) 2000-10

(d) 2010-20

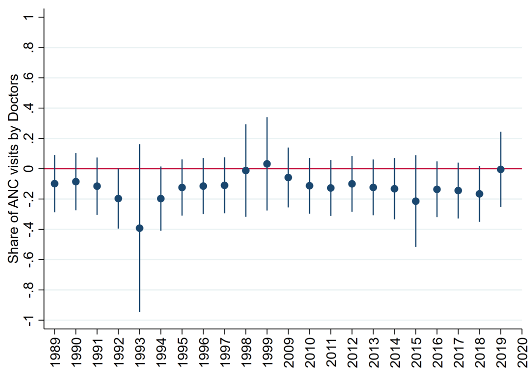
Figure 3: Districts by year of first opening of medical colleges



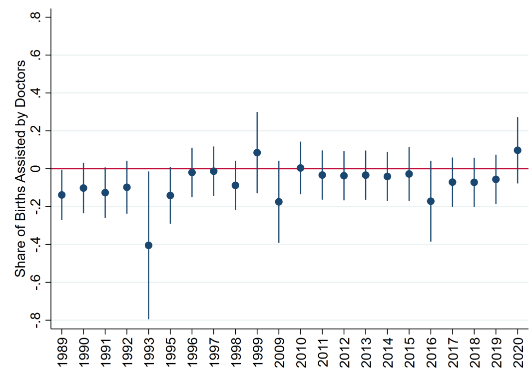
(a) ANC visits



(b) Infant mortality rate



(c) Share of ANC visits by doctors



(d) Share of deliveries assited by doctors

Figure 4: Parallel pre-trends in outcome variables

These figures present the trends in outcome variables between districts where a college first opened in 2000-10 and districts where a college first opened in 2010-20. The estimates plotted are the coefficients β_τ in Equation 2.

9 Appendix

	Dep. Var. : Health Services Uptake									
	ANC visits		First ANC		Birth in Facility		Tetanus Inj. (during preg)		Tetanus Inj. (before preg)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Graduate batches	-0.003	0.046***	0.003	0.005	-0.004***	-0.005***	-0.014***	-0.008***	0.005	0.003
	(0.006)	(0.010)	(0.003)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)	(0.010)	(0.018)
College opening	0.236***	-0.084	-0.009	0.042	-0.017***	-0.029***	-0.116***	-0.066***	0.062	-0.005
	(0.051)	(0.093)	(0.030)	(0.040)	(0.005)	(0.008)	(0.016)	(0.019)	(0.064)	(0.124)
Spillover	0.078***	0.114***	0.051***	0.038***	-0.016***	-0.019***	-0.084***	-0.031***	0.038	0.046
	(0.015)	(0.028)	(0.009)	(0.012)	(0.002)	(0.003)	(0.005)	(0.006)	(0.037)	(0.066)
Observations	8,608	7,089	8,398	6,968	8,950	7,547	8,608	7,090	4,351	2,723
Mean Dep. Var.	3.453	4.711	3.835	3.363	0.543	0.761	1.647	1.914	1.132	1.218

Table 8: Impact of colleges on take-up of healthcare by residence location

This table presents the results on different measures of healthcare take-up in urban and rural areas. The first two columns are the number of ANC visits made during pregnancy in urban and rural areas respectively. The next two columns are the month of the first ANC visit. The fifth and sixth columns are an indicator of whether the child was delivered in a facility. The last four columns are indicators for whether the mother received a tetanus injection. Within each district the mean of the dependent variable in urban and rural areas is separately calculated and used in the respective columns. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Share of Antenatal Care Visits											
	Doctor		Nurse/Anm		TBA		CHW		Aganwadi		Asha	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Graduate batches	0.000	0.000	-0.004***	-0.004***	0.000	-0.001	0.000	-0.000	-0.002*	-0.003**	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
College opening	0.013	0.012	-0.012	-0.021*	-0.005	0.013**	-0.002	-0.002	-0.009	0.009	-0.002	0.024**
	(0.009)	(0.011)	(0.009)	(0.012)	(0.006)	(0.007)	(0.002)	(0.003)	(0.009)	(0.011)	(0.008)	(0.010)
Spillover	-0.005**	0.007**	-0.009***	0.005	-0.005**	0.001	0.005***	0.003*	-0.015***	-0.005*	-0.014***	-0.006**
	(0.003)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)	(0.003)
Observations	8,609	7,091	8,609	7,091	8,609	7,091	4,713	4,459	6,774	5,857	6,774	5,857
Mean Dep. Var.	0.461	0.688	0.364	0.392	0.015	0.018	0.007	0.006	0.121	0.102	0.093	0.069

Table 9: Impact of colleges on share of ANC visits by provider type and residence location

This table presents the results on share of ANC visits assisted by different providers in urban and rural areas. Multiple providers may be present during the same visit. The providers in each pairs of columns are doctors, nurses/ANM, traditional birth attendant, community health worker, anganwadi worker, and ASHA worker respectively. Within each district the mean of the dependent variable in urban and rural areas is separately calculated and used in the respective columns. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Share of Deliveries Assisted											
	Doctor		Anm		TBA		Friends		Other		None	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Graduate batches	0.002***	0.001	-0.005***	-0.005***	0.003***	0.002**	0.006***	0.007***	0.001	0.000	0.001**	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
College opening	0.013*	0.004	-0.068***	-0.078***	-0.027***	0.007	0.013*	0.005	-0.001	-0.002	-0.002	-0.002
	(0.007)	(0.010)	(0.007)	(0.011)	(0.007)	(0.008)	(0.007)	(0.010)	(0.005)	(0.005)	(0.003)	(0.003)
Spillover	0.007***	0.002	-0.027***	-0.017***	-0.000	0.003	0.014***	0.003	-0.001	0.003**	0.002*	0.003***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	8,609	7,092	8,609	7,092	8,609	7,092	8,609	7,092	6,774	5,857	8,609	7,092
Mean Dep. Var.	0.378	0.573	0.434	0.580	0.258	0.142	0.384	0.241	0.040	0.027	0.017	0.008

Table 10: Impact of colleges on share of deliveries assisted by provider type and residence location

This table presents the results on share of deliveries assisted by different providers in urban and rural areas. Multiple providers may be present during the delivery. The providers in each pairs of column are doctors, nurses/ANM, traditional birth attendant, community health worker, anganwadi worker, and ASHA worker respectively. Within each district the mean of the dependent variable in urban and rural areas is separately calculated and used in the respective columns. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.

	Dep. Var. : Disease Incidence									
	Diarrhea		Fever		Cough		Diabete		Adult Mortality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Graduate batches	-0.000	-0.000	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
College opening	0.003	0.003	0.009	0.009	0.016**	0.016**	0.001	0.001	0.001***	0.001***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.008)	(0.002)	(0.002)	(0.000)	(0.000)
Spillover	0.000	0.000	0.001	0.001	-0.000	-0.000	-0.001	-0.001	0.000***	0.000***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)
Observations	2,163	2,163	2,163	2,163	2,163	2,163	1,135	1,135	6,364	6,364
Mean Dep. Var.	0.116	0.116	0.182	0.182	0.195	0.195	0.018	0.018	0.007	0.007

Table 11: Impact of colleges on health status by residence location

This table presents the results on morbidity and mortality in urban and rural areas. All columns are in units of percentage of population. The first eight columns are indicators for whether the respondent ever had that condition in the last one year. Columns 9-10 and 11-12 are respectively the infant mortality rate and the adult mortality rate measured using the DHS. Within each district the mean of the dependent variable in urban and rural areas is separately calculated and used in the respective columns. The coefficients estimated are the parameters in equation 1. All regressions include district and time fixed effects.