Towson University
Department of Economics
Working Paper Series

Working Paper No. 2020-04

Healthcare Reform and Gender Specific Infant Mortality in Rural Nepal

By Juergen Jung and Vinish Shrestha

May 2022

Healthcare Reform and Gender Specific Infant Mortality in Rural Nepal*

May 20, 2022

Abstract

We estimate to what extent a large scale health care reform disproportionately affects the mortality rate of boys in the context of a developing country with cultural preferences favoring boys. We use arguably exogenous variations due to a health care reform—the National Health Policy—which was implemented in Nepal in 1991 along with data from the Nepal Living Standard Survey 1996 and estimate that improved quality of primary health care facilities (by one standard deviation) reduces the mortality rate of infant boys by 3.43 percentage points but does not affect the mortality rate of infant girls. Our analysis points to societal gender preferences for sons and the consequent neglect of daughters' health as potential drivers of some of the observed differences in mortality between genders and highlights the important role of cultural norms in shaping the outcomes of large scale health care reforms.

JEL: C35, I23, I10, I18

Keywords: Infant and child mortality; gender specific health investment; health inequality by gender; access to health care in developing countries.

*This paper was previously circulated under the title "Healthcare Reform and Gender Specific Child Investment in Developing Countries." We would like to thank Rajan Bishwakarma, Eric Edmonds, Swarnim Wagle and participants at the economic research seminar at Towson University for helpful comments. We would also like to thank Avidit Acharya, Apoorva Lal and Rashesh Shrestha for helping us resolve some data issues. All remaining errors are ours.

1 Introduction

In many developing countries, strong societal or cultural norms based on gender, caste, and ethnicity significantly affect the human capital development of children, especially if a child belongs to a marginalized sub-group. For instance, differences in nutritional investments by gender (Chen, Huq and d'Souza, 1981; Jayachandran and Kuziemko, 2011; Jayachandran and Pande, 2017), sex-selective abortion, and—even more extreme—infanticide have all contributed to an increase in the gender gap in measurable outcomes such as health and education and in many cases contributed to an overall gender imbalance in the population (Sen, 1990, 1992; Jayachandran, 2015).

While it is often suggested that policies aimed at increasing access to services such as healthcare and education reduce inequality between disadvantaged and advantaged groups, many of these policies are deployed homogeneously, without regard to existing gender biases or prevalent cultural norms. Therefore, it is important to ask whether such policies potentially widen the gap in human capital investments between genders in environments where males are favored due to both cultural (societal preference bias) and economic (labor market bias) factors.

In this paper we study whether a health reform—implemented to improve access and quality of health care—can inadvertently induce child health investment incentives that disproportionately benefit boys relative to girls in a setting where both "societal preferences" and labor markets favor males. We use an arguably exogenous health policy reform in Nepal—the National Health Plan (NHP) of 1991—and data from the Nepal Living Standard Survey (NLSS 1996) in combination with other data sources to address this question.

The NHP was implemented in 1991 with the main goal of improving access to health care, particularly in the rural parts of the country. The reform expanded preventive, promotive, and curative health care services. Priority was given to reducing infant and child mortality, which despite having decreased dramatically in the 1980s, was still high in the late 1980s at about 10 percent.[1] To achieve these goals the NHP established local health centers (health posts). A health post typically employs a health assistant or a senior auxiliary health worker and comprises about six employees. They serve as primary point of contact for basic health care services and referrals. Figure 1(a) shows that following the NHP the number of health posts increased significantly between 1991–1994. A similar increase in health care spending is shown in Figure 1(b). The reform also established district as well as zonal hospitals.[2] Zonal hospitals are equipped with higher quality infrastructure than district hospitals and they also provide specialized services related to pediatrics, gynecology, general medicine as well as ear, nose, and throat treatments (Nepal, 2007). The NHP is primarily a supply side reform with the government focusing on building up basic health care infrastructure.

[1]Compare data from the World Bank at: https://data.worldbank.org/indicator/SP.DYN.IMRT.IN?locations=NP

[2]Nepal comprises 14 zones that are further divided into 75 districts.

Still, health care is not free and health posts are partly financed via user charges. In general, the Nepalese health care system to a large extent is funded by out-of-pocket health expenditures. In 2003, well after the implementation of the NHP in 1991, out-of-pocket healthcare spending was as high as 62.5 percent of the total health expenditure with poorer households spending a smaller share of their income on health and only in 2007 did the government pass a provision to abolish user charges at health posts (Nepal, 2007).

Our main focus is is to analyze the effect of the reform at the primary care level by exploiting across-district variation in the number and quality of newly built health posts in combination with birth information of children born in years before and after the reform in a difference-in-differences setup. The NLSS contains two crucial pieces of information in support of this estimation strategy: (i) the date of construction of a health post and (ii) a number of variables that describe the quality of fully established health posts in survey year 1996. These quality measures include information such as the availability of doctors, hospital beds, delivery care, etc.

We first use the date of establishment to calculate the number of newly built health posts (per 1000 people) since the reform and relate it to a child's birth date. This allows us to estimate the effect of newly established health posts on infant mortality. Second, we use exploratory factor analysis and construct a Core-Infrastructure-Quality variable and a Reproductive-Care-Quality variable to assess whether districts with high quality health posts benefited more from the reform (in the form of reduced infant mortality rates) than districts with health posts of lesser quality. By including the quantity and quality measures simultaneously into a regression framework we can identify the quantity vs. quality effects of the NHP reform.

Our main results suggest that while the number of newly constructed health posts in a given district has no effect on infant or child mortality outcomes, districts with higher quality health posts after the reform have significantly lower mortality rates of boys but not girls. Specifically, a one standard deviation increase in the quality measure describing the infrastructure status of a health post (e.g., availability of beds, refrigerators, electricity, and doctors), reduces boys' infant mortality rate (Ages 0–1) by about 3.43 percentage points. We highlight that the overall quality of health posts was poor prior to the reform; although the reform improved the quality of health posts constructed after 1991 in some areas, the quality of many health posts remained poor even after the reform. As such, quality rather than quantity of health post is the determining factor in lowering the mortality rates of boys.

The validity of these findings rests on the identification assumption that in absence of the reform, there would be no systematic difference in child mortality between districts that are highly affected by the reform and districts that are not strongly affected by the reform. We show that although more health posts were established in areas that lacked health facilities, the allocation of health posts is not systematically related to child mortality outcomes prior to the reform. Furthermore, we use an event study design and show that trends in child mortality rates prior to the reform are not different in districts that are more affected by the reform compared to districts that are less affected. This is true across both genders. However, child mortality outcomes improved for sons in districts that received higher quality health posts following the reform. Additionally, we conduct placebo exercises that also support the validity of our identification strategy.

Literature Review. Differences in child investments by gender are well-documented in the literature, especially for South Asian developing countries. A recent review by Rose (2017) points out that among other forms of discrimination, girls receive less adequate nutrition and healthcare in many Asian countries. These differences in child investments are often attributed to discrimination in labor markets that significantly lower the expected rates of return of education and health investments in girls. Reports by the United Nations and the IMF point to significant gender gaps in labor market outcomes that are especially large in South Asian developing countries (e.g., Elborgh-Woytek et al., 2013 and Nations, 2015).[3] Since children are important sources of income for aging parents in developing countries, these labor market effects are important.

The second main factor contributing to child investment differences is rooted in "societal preferences" for sons which is caused by social institutions, laws, cultural norms, and traditions (Pande, 2003; Morrisson and Jütting, 2005; Barcellos, Carvalho and Lleras-Muney, 2014; Lin, Liu and Qian, 2014; Echávarri and Husillos, 2016). Although many of the studies on gender biased health investments focus on India and China (e.g., Behrman and Deolalikar, 1990; Rose, 2000; Young-Choi and Lee, 2006; Chakravarty, 2010; Jayachandran and Pande, 2017; Bharadwaj et al., 2020), differences in child investments by gender are large in many other developing countries as well (e.g., Bharadwaj and Lakdawala, 2013).[4]

This study makes several contributions to the literature. First, our paper builds on the literature studying the impact of health reforms in developing countries, much of which is focused on whether a reform lowers the out-of-pocket expenses for health care (Dow and Schmeer, 2003; Thornton et al., 2010; Matsushima, Yamada and Shimamura, 2020; Bernal, Carpio and Klein, 2017). While the literature seems in agreement that lowering access cost will increase utilization of health care in general, it is not clear whether higher utilization will lead to lower infant or child mortality rates. For instance, after removing user fees for facility-based delivery services, McKinnon et al. (2015) report slight reductions in neonatal mortality while Fitzpatrick (2018) does not find evidence for decreases in child or maternal mortality. The author even finds some evidence for increases in mortality. In other words, higher demand could lead to lower quality of health care due to under staffing issues, waiting lines, or absence of medicines. Different from these studies we focus on the gender specific effects of improvements in health care quality as opposed to the overall effects of lowering out of pocket expenses through expansion of insurance.

[3] A more recent report by the International Labor Organization documents that between 1995–2015 the global female labor force participation gap in Southeast Asia has actually increased (ILO, 2016).

[4] There is even some evidence that parents in developed countries favor boys (e.g., Dahl and Moretti, 2008).

Second, our paper contributes to the literature evaluating the effect of institutional reforms on the gender-gap in health outcomes in developing countries with strong son preferences. Grown, Gupta and Pande (2005) argue that improving infrastructure in developing countries—especially transportation, water, and sanitation—reduces the gender gap in health outcomes.[5] Some national poverty reduction policies—such as India's rural employment guarantee scheme (NREGS) which was passed in 2005—seem to have gender neutral effects with respect to education outcomes of children between the age of 13–16 (Shah and Steinberg, 2020). In contrast, Hazarika (2000), Asfaw, Klasen and Lamanna (2007) and Tanaka (2014) find a significant difference in the allocation of health care which favors boys in Pakistan, India, and South Africa, respectively. If the gender of the child is revealed before birth, this type of discrimination even extends to prenatal care as shown in Bharadwaj and Lakdawala (2013) for India, China, Bangladesh, and Pakistan—countries known to have son preference. Research by Oster (2009) points to non-monotonic effects that can explain the contradictory results in the literature. She shows that an increase in health investments benefits boys when the original investment level is low but once investment levels rise, girls begin to benefit as well from reforms that increase access to health care and the health disparities narrow. While Oster (2009) explores the non-monotone effect on vaccinations via health camps, we investigate the differential effects at the primary (health posts) level of care and not only evaluate the effects of access to care but also quality of care by using the NHP as a natural experiment.

Third, we highlight the importance of acknowledging cultural aspects when implementing new health policies that focus on more equitable outcomes in cultural settings with strong son preferences. In this regard, we contribute to the small but growing studies advocating for the need of heterogeneous policies that considers societal norms in enhancing the well-being of a general populace (e.g., Collier, 2017; Ashraf et al., 2019).

The study is organized as follows. Section 2 introduces the data and Section 3 discusses the National Health Policy reform. Section 4 introduces the empirical model and identification strategy and Section 5 presents the main results. We conclude in Section 6.[6]

[5] There are many studies for the U.S. suggesting that health policies that favor disadvantaged groups such as Medicaid do reduce differences in health outcomes between the target group and the general population (e.g., Buchmueller, Ham and Shore-Sheppard (2015) and Goodman-Bacon (2021)).

[6] The Online Appendix contains details about exploratory factor analysis, details about robustness exercises, results based on a theoretical model, as well as additional figures and tables.

2 Data

2.1 Data Sources

2.1.1 The Nepal Living Standard Survey (NLSS 1996)

Our main data source is the NLSS 1996, a detailed, nationally representative household survey conducted by the Nepal Bureau of Statistics in collaboration with the World Bank. The NLSS 1996 is a stratified sample with the ward being the primary sampling unit.[7] The survey samples 3,388 households across the country and is considered nationally representative.

Infant and Child Mortality Rates. The NLSS is preferred to other available data sets because it includes both, a detailed questionnaire about the history of maternity of all women who gave birth and a rural community module with detailed information about all available health facilities in the sampled communities.[8]

More specifically, for every child born, the mother of the child is asked to report the child's date of birth (i.e., month and year), the child's gender, the current mortality status of the child (alive or dead), and the highest level of schooling completed (regardless of the mortality status). In case a child has died, the mother reports how long the child has lived in terms of years, months, and days. This allows us to track the child's mortality status using the child's timing of birth, even from a cross-sectional data set. We then construct indicators for infant mortality (deaths within a year of birth) and under 3 child mortality using the time of birth and death (month and year), which are used as the primary dependent variables in the study.

Additionally, we calculate infant and under 3 child mortality rates for descriptive analysis. The infant mortality rate in a specific year t is the number of children less than one year of age that is reported dead divided by the number of all births reported in the same year. Similarly, child mortality rates of 0–3 year olds in year t is constructed as the number of children less than 3 years of age that is reported dead divided by the number of all births reported in years $t-2$, $t-1$, and t.

Health Posts and Quality Measures. The rural community module of the NLSS provides detailed information about the administrative and operational characteristics of the available local health care facilities. The questionnaire first asks about the type of health care facility (sub-health post, health post, hospital, private clinic) available in the Village Development Committee (VDC) of the primary sampling unit. In 1991 Nepal had close to 4,000 Village Development Committees (VDCs).[9] It then asks for the number of years that the facility has been operational, which allows us

[7] A ward is the smallest administrative unit in the 1991 Population Census.

[8] We do not use the more popular Demographic Health Survey (DHS) from year 1996 because it lacks community level files and therefore does not have information about local health care facilities.

[9] These VDCs are divided into several wards—the smallest administrative division in the country—depending on population. On average, a VDC has 9 wards.

to calculate the establishment year of the facility. To focus on addition of new health posts, we only consider sub-health posts and health posts established following the reform (termed together as health post in this study). In addition, detailed information concerning the quality of the health post is reported: 1) the provision of delivery care, 2) the provision of antenatal care, 3) the availability of beds, 4) whether a doctor is present, 5) whether a health worker is present, 6) the availability of treatment for malaria, 7) hours the health post remains open in a week, 8) access to electricity, and 9) if the health post has a refrigerator.

Demographic Information. Additionally, the NLSS provides a detailed set of demographic variables such as ethnicity (Brahmins, Chhetris and other caste), religion (Hindu, Buddhist), father's and mother's education, mother's district of birth, and mother's community of residence. Additional variables such as the ideal number of sons versus daughters,[10] household savings, access to mode of transportation, and vaccination status are used to discuss the main findings. As our analysis focuses on the effects of the NHP in rural areas of Nepal, we restrict the sample to 7,361 children born between 1980–1996 to mothers living in rural areas. The designation of urban versus rural region is part of the sampling design and about 80 percent of the sample respondents in the 1996 survey reside in rural areas.

2.1.2 Geographic Data

Since the reform also improves access to health care at district and zonal hospitals (tertiary care level), we include geographic data as indirect controls for possible changes in tertiary health care. Figure 2(a) depicts the zonal and district level divisions of Nepal and Figure 2(b) shows the location of zonal hospitals including an elevation scale.[11]

Unfortunately we do not have access to data about the quality of zonal hospitals during the implementation period of the reform in the early 1990s. We therefore cannot directly evaluate the effects of the reform at the tertiary care level or control for hospital quality effects directly. However, we can control for some of the tertiary-level effects using hospital distance measures such as the log of distance from a household's VDC to the closest zonal hospital;[12] (*ii*) the elevation of a household's VDC in relation to the nearest hospital, and (*iii*) the distance from a household's VDC to the the closest river.

To obtain the elevation measure we first locate the latitude and longitude of the center of the VDC where a child's mother lives. These location points (in latitudes and longitudes) are con-

[10]Respondents are asked about what they believe to be the ideal number of boys in their family. This variable is used as a son preference measure.

[11]Nepal is divided into 14 zones and 75 districts that are grouped into five development regions. Additionally, in 1991 there were close to 4,000 Village Development Committees (VDCs). These VDCs are divided into several wards—the smallest administrative division in the country—depending on population. On average, a VDC has 9 wards.

[12]We construct the measure of shortest distance to zonal hospitals from the VDC centroid.

sequently used together with the Elevation API of the Google Maps Platform to construct VDC specific elevation measures.[13] We then use a simple algorithm to calculate the shortest distance from the VDC centroid to the closest zonal hospital using the Haversine formula, which calculates the shortest direct distance on a sphere ("as-the-crow-flies"). Geographically, Nepal lies in the central Himalayan region and over a short distance of 193 km from the northern to southern border, the elevation drops dramatically from 8,800 meters to 70 meters (above sea-level). Such drastic changes in elevation create extensive variation in travel cost.

Next, given the country's location in the Hindu Kush Himalayan region, some 6,000 rivers flow through the country that can be categorized into ten larger river basins. These river basins mainly acquire annual river flow through monsoon precipitation between April–September. In addition, snow melt contributes to the pre-monsoon river flow of four major river basins—Mahakali, Karnali, Gandaki and Koshi—that originate from the Himalayas. The rivers in the mountains are deep, narrow and steep with flow through gradients of 40 meters per kilometer of distance (Devkota, 2014). Larger rivers exacerbate travel due to lack of bridges and thus increase the cost of visiting zonal hospitals both before and after the NHP. Due to the river gradient and steepness, water transportation is extremely risky in Nepal in general.[14] We therefore use the shape file of rivers in Nepal from the Humanitarian Data Exchange repository and construct a spatial map of river flow across the country as shown in Figure 2(c). We again apply the Haversine formula and use the mother's VDC location along with the river paths to calculate the shortest distance of the VDC to the closest river.

2.1.3 Additional Data

Two waves of the Nepal Housing and Population Census (1981, 1991) at the district level provide additional control variables including population estimates and literacy rates by gender. In addition, the Education Statistics 1993 provide data on schooling infrastructure, the total number of school teachers, the number of trained teachers, and the number of female teachers per school in 1993. We use these data to assess whether our results are driven by Nepal's schooling boom in the early 1990s—mainly a demand-driven phenomenon.[15]

[13] The Elevation API provides an interface to estimate the elevation of a given location using Digital Elevation Models facilitated by NASA's Shuttle Radar Topography Mission (SRTM3) at 3 arc-second (approximately 90 meters). Elevation models are divided into tiles with each covering one degree of latitude and longitude. Several data points that fall within a sample tile are averaged to reduce noise. Elevation data are affected by land terrain including mountains, which is relevant to our study. The algorithm we used is available upon request from the authors. The Elevation API is very detailed and is often used to develop biking and hiking applications. For more information go to: https://developers.google.com/maps/documentation/elevation/intro#Introduction

[14] Typically, people cross rivers using *Tuin*, which is a simple wooden structure hanging on a wire rope which is affixed to land structures (usually cliffs).

[15] A detailed discussion regarding the schooling boom is provided in the Online Appendix.

2.2 Descriptive Statistics

The descriptive statistics in Table 1 are based on a representative rural sample of the NLSS 1996. We focus on mothers as they are the responders to questions about child birth and child mortality. The average age of mothers in the sample is 34.7 years. The marriage rate is high at 95 percent. The majority of mothers, 87.4 percent, reports being Hindu and about 34 percent belong to high castes such as the Chhetri caste (20.5 percent) and the Brahmin caste (14 percent). These statistics resemble the country wide caste statistics. Only 10 percent of mothers report that they are able to read and 9 percent report being able to write. The country's overall literacy rate in 1981 among 10–19 year olds was 19 percent for females and 48 percent for males.[16]

2.2.1 Child Mortality Rates and Gender Preferences

About 88 percent of children born to women who were between 15–70 years old when the survey was conducted are reported to be alive in the survey year. The mortality rate is high for age group 0–1 (infants) at over 14 percent in 1980 with an average of 8.9 percent over the entire period from 1980–1996 as reported in Table 1. Figure 3(a) plots estimated year dummy coefficients from a specification that regresses an indicator for infant mortality on year indicators and district of birth fixed effects. The figure shows that sons' mortality decreased following the reform (omitted years 1991), while such a trend is not evident in daughters' mortality outcomes.

Mortality rates decrease considerably with age as highlighted in Figure 3(b) for both males and females. This is consistent with earlier findings that show that children less than a year old are the most vulnerable. It is reassuring that the year and gender specific mortality rates calculated from the NLSS 1996 are very similar to mortality rates published by the World Bank.[17]

The NLSS survey directly asks 15–49 year old women about their fertility preferences with respect to gender. Figure 4 shows the responses of mothers in the NLSS 1995–1996 rural sample when asked about the "ideal number" of sons and daughters. The figure depicts a strong gender preference for boys. The women in the sample prefer a lower number of girls—the mode of the ideal number of boys is 2, whereas the mode for girls is 1.

2.2.2 Health Posts and Quality Measures

In our sample the average number of total health posts in a district in 1989 is 1.6 which subsequently doubles to 3.18 by 1996. Figure 1(a) shows that a significant share of this increase occurs after the implementation of the NHP in 1991. One health posts approximately serves 3,000 people in NLSS 1996 sample. The average distance to health posts is 78 minutes (travel time on foot).

[16] Calculated by the authors using Census 1981 data.

[17] Compare https://data.worldbank.org/indicator/SP.DYN.IMRT.IN?locations=NP

We aggregate the quality measures of health posts constructed in years 1991 to 1996 (post reform) at the district level—Nepal has 75 districts—and show the result of this aggregation in Figure 5. The figure shows the proportion of health posts in each district offering a certain type of infrastructure or care. We then use exploratory factory analysis (EFA) to reduce the number of variables associated with the district level quality of health posts built following the reform and retain two quality measures, one strongly associated with a health post's core infrastructure and the second strongly associated with the availability of services for reproductive care. More specifically, the availability of doctors, hospital beds for inpatient care, refrigerators to store medicine, provision of electricity, and weekly opening hours of health posts heavily load onto factor 1 (which is referred as **Core-Infrastructure-Quality**) whereas antenatal care, delivery care and malarial treatment load onto factor 2 (which is referred to as **Reproductive-Care-Quality**). The advantage of EFA over alternative dimension reducing methods such as Principal Component Analysis (PCA) is that the factors are directly related to specific types of quality measures as opposed to measures solely constructed based on capturing maximum variation in as few variables as possible. Therefore, EFA factor scores can be interpreted more easily.[18] These quality measures represent the average quality of all health posts constructed following the reform in the district where the child was born.

2.2.3 Distance to Hospitals

Table 1 also reports the variables attributed to control for geographical characteristics of VDCs. The average elevation of a mother's residential location in the sample is 884 meters. There is substantial variation in the elevation as shown by the large standard deviation of 872 meters in Table 1. The average distance from a centroid of a VDC to the closest zonal hospital is 45.73 kilometers, which is 25.41 miles. Although this distance is considered relatively short in developed countries, in the context of a developing country where the majority of travel is done on foot (especially in the 1990s in Nepal) and with additional costs imposed by the undulating landscape, such a distance posits a very high cost of both time and effort.[19]

[18]The detailed methodology of the EFA procedure and the associated scree plots that are used to determine how many factors to include in the baseline regressions are presented in the Online Appendix. We also conduct three robustness checks to ensure that our chosen selection method of factors associated with the observed quality measures is not driving the results. First, we aggregate the quality measures at the primary sampling unit level (as opposed to district level). Second, we individually include the quality variables directly as control variables. Third, we use Principal Component Analysis (PCA) as an alternative dimension reduction method. Our results hold under these alternative specifications. The estimation tables based on the three robustness checks are available in the Online Appendix.

[19]Even without considering the extreme differences in elevation in Nepal, at a rate of 3 miles per hour as given by Naismith's rule, it takes an average of 8.47 hours to reach the closest zonal hospital. See Pharoah (2017).

3 The National Health Policy of 1991

In the 1980s the population of Nepal suffered from bad health outcomes due to poor socioeconomic characteristics (e.g., low levels of education, high poverty rates, prevalent caste and gender based discrimination) and a lack of political commitment to improve accessibility to health care. Inadequate health infrastructure, especially the low number of health facilities and widespread under staffing issues, contributed to high mortality rates of 93 deaths per $1,000$ births for infants and 133 deaths per $1,000$ births for children under five in 1991. At that point Nepal had one hospital per $168,000$ people, one hospital bed for every 4,000 people, one doctor per 92,000 people, and one health post per 24,000 people in rural areas (Shrestha and Pathak, 2012).

The National Health Policy (NHP) of 1991 extended and improved access to health care across Nepal through establishing (i) sub-health posts at the VDC level (village development committee), (ii) health posts or primary health centers at the electoral constituency level (*ilaka*), (iii) district hospitals at the district level, (iv) a zonal hospital in each zone, (v) regional hospitals, and (vi) central hospitals at the central level (Rai et al., 2001).[20] The NHP prioritized preventive, promotive, and curative health services to improve national health standards and reduce infant and child mortality (Nepal Law Commission, 1991). This effort resulted in significant increases in health care spending after the reform as shown in Figure 1(b).

At the primary care level, a sub-health post or health post serves as the first point of contact to health care in rural areas. The establishment of sub-health posts and health posts surges following the NHP as shown in Figure 1(a). The sub-health posts were established in phases after the NHP and according to the objective of the reform all village development committees (VDCs) of the country were to be equipped with at least one sub-health post. Although very minimal in nature, the establishment of sub-health posts includes the assignment of one village health worker, one maternal, and one child health worker. According to the policy, one of the health posts located in each of the 205 election constituencies was to be upgraded to a primary health care center. Since the majority of health units established following the policy were sub-health-posts and health posts, we collectively refer to these units as health posts from here on. The reform also provisioned organization of mobile teams to provide specialist services in remote areas. Although Figure 1(a) shows an increasing trend in the number of health posts since the 1950s, there is a considerable break in trend in 1991 following the NHP.

Table 2 describes the allocation of health posts across districts following the NHP. The relationship between the number of health posts in 1989 and the number of new health posts after the NHP is negative as shown in Column (1). Additionally, the coefficient of (log of) population size—determined by the sum of the population of all the sampled wards of a district—is close to one, suggesting that more populated areas were prioritized. Districts without a health post in 1989

[20] Refer to Figure 2(a) for a brief overview of the administrative division of Nepal.

received an average of one additional health post following the NHP compared to districts that already had a health post in 1989 (see Column 2). As the program was specifically designed to expand health care in rural areas, the results from Columns (1) and (2) confirm that health posts were indeed established in under-served districts. Column (3) shows that the number of electoral constituencies in a district is strongly and positively correlated with newly established health posts, indicating that the establishment of actual health posts followed the schedule. Furthermore, Column (4) indicates that the allocation of health posts is not systematically related to the child mortality rate prior to the program. The infant mortality rate coefficient is close to zero and statistically insignificant at the conventional levels.

Next, in Columns (5)–(6) and (7)–(8) we use the quality measures of the health posts that were established after the NHP as the dependent variables. Column (5) documents a negative relationship between the quantity of health posts constructed between 1991–1996 and the Core-Infrastructure-Quality measure of newly established health posts, which indicates a quality-quantity trade-off. This is consistent with qualitative studies suggesting that 30–50 percent of health posts lacked health staff (Agarwal, 1998) and further highlights the importance of the quality of health posts when analyzing the effects of the program. Furthermore, Column (6) indicates that the Core-Infrastructure-Quality measure of health posts established following the NHP is not correlated with child mortality before the reform.

Additionally, Table 3 shows that if a health post was built in a household's VDC after the NHP, the probability of living closer to a health post (within an hour) increases by 14.4 percentage points. This estimate is robust to adding district fixed effects. Adding controls for the presence of health posts prior to the NHP does not alter the magnitude. In fact, health posts established before the NHP in 1991 are not good predictors of the average distance of a household to a health post in 1996.

Even with improved access to health care, the quality of health care provided by the health posts was very basic and substantial variation exists in the quality measures of health posts in 1996 as shown in Figure 5. While antenatal care, delivery care, and treatment for malaria—criteria aligned with preventive care of NHP—are quite common across the district, the core measures of infrastructure such as the availability of doctors, hospital beds, electricity, hours open (in a week), and refrigerators are severely lacking. In summary, only 12 districts included health posts with beds and only 17 districts included health posts with doctors. Figure 6 shows the average predicted factor scores of health posts for each district obtained from EFA. Not surprisingly, the factors show a similar geographic distribution as the variables they are based on. It is interesting to see that variables pertinent to Core-Infrastructure-Quality are dispersed across the country and not just concentrated in one region.

The health reform also introduces a referral system at the primary level (health posts) to direct the rural population to zonal and regional hospitals. Zonal hospitals mainly serve the purpose of providing curative health care services. Nepal operates 11 zonal hospitals under the supervision of the Ministry of Health, which were established following the NHP. The zonal hospitals were provisioned to meet the program's (NHP) objective of improving curative care and the zonal hospitals included specialized services related to pediatrics, gynecology, general medicine, eye, ear, nose and throat care, and dental care.[21] Although almost every zone has one zonal hospital, access to the hospital is constrained by geographic factors such as the household's distance to the hospital, the elevation difference of the individual's location and the location of the zonal hospital, and the distance to the closest river. The location of mother's area of residence (i.e., her VDC) in the sample along with elevation is shown in Figure 2(b). The VDCs in the NLSS are more concentrated in the south, which is consistent with the higher population density in the southern regions.

4 Estimation Strategy

We focus on the effects of primary health care (basic health care) via health posts on child mortality. We use across-district variation in the number of health posts and their associated quality measures between cohorts born before and after the NHP to describe infant mortality. The formal expression of the infant-mortality model is:

$$\begin{aligned} \text{Mortality}_{imbd} &= \alpha + \eta \times (\Delta\text{HP}_d \times \text{Post}_{imbd}) + \delta \times (\text{Core-Infrastructure-Quality}_d \times \text{Post}_{imbd}) \\ &+ \beta \times (\text{Reproductive-Care-Quality}_d \times \text{Post}_{imbd}) \\ &+ \psi_d + \omega_b + \gamma \times X_{imbd} + \theta \times G_{imd} + \varepsilon_{imbd}, \end{aligned} \tag{1}$$

where Mortality_{imbd} indicates whether child i born to mother m in year b in district d has died before reaching age 1. The unit of observation is a child or child-record with associated information about the child's birth year b, gender, and mother's attributes. Every child record appears only once in our data. We also present a child-mortality version of this model where the Mortality_{imbd} indicator variable switches to one if a child has died before reaching the age of 3.

Variable $\Delta\text{HP}_d = \text{HP}_{1996,d} - \text{HP}_{1990,d}$ represents the change in the number of health posts per 1,000 people in district d from the year before the reform, 1990, to the survey year 1996. It measures the number of new health posts created due to the reform. Alternatively, in some specifications ΔHP_d is replaced by the number of constituencies in the district in 1990, since the scheduled allocation of new health facilities depended on the number of constituencies (per district). This is shown in Column (3) of Table 2. Using the number of constituencies as an alternative measure for the impact of the reform allows us to distinguish between the policy makers' intent to build new health posts (the law was written so that each VDC should have at least one) vs. the actual health

[21] Nepal also operates 6 regional hospitals (in Pokhara, Hetauda, Surkhet, Birgunj, Ghorahi, and Dang). See https://www.mohp.gov.np/eng/# for the list of hospitals.

posts being built. If the results between the two specifications differ significantly, one could argue that some unobserved conditions on the ground determine the number of actual health posts being built, rather than the reform. This would highlight a potential endogeneity problem.[22]

The factor score Core-Infrastructure-Quality$_d$ measures the quality of the infrastructure of a health post (i.e., beds, doctors), and Reproductive-Care-Quality$_d$ is the factor score associated with the quality/availability of child care services at a health post. Both are only observed in year 1996. Note that the reform in 1991 had a strong focus on improving and building new health posts. Since the quality measures are aggregates at the district level and we only use health posts built between 1991–1996 to construct the quality measures, these quality measures are directly attributable to the reform.

The health post and quality measures are interacted with an indicator variable labeled Post$_{imbd}$, which equals one if child i is born after the NHP (i.e., birth year $b \geq 1992$). The parsimonious specification also includes ψ_d and ω_b, representing a child's district of birth and year of birth fixed effects, respectively. The coefficients of interest are η, δ, and β. We estimate equation 1 using a linear probability model and the heteroskedasticity-robust standard errors are clustered at the district level. All estimations are performed separately for boys and girls.

In alternative specifications we add a set of geographic control variables (G) that include the shortest log-distance to a zonal hospital (SD Hospital), elevation (and elevation squared), and shortest distance to the nearest river (and shortest distance squared) interacted with pre- and post reform dummies, respectively. All geographic measures are measured with respect to a child's mother's VDC centroid location. Additionally, variable X_{imbd} is a vector that controls for the mother's predetermined characteristics such as ethnicity, religion, and age as well as child's birth order dummies.

The identification of equation 1 is based on the assumption that in the absence of the NHP, the trend in infant (child) mortality would not be systematically different between districts that receive either fewer or lower quality health posts and districts with either more or higher quality health posts. As previously mentioned, the main purpose of the NHP was to expand health services in the rural parts of the country, which is validated by the first stage results in Table 2. This suggests that the NHP was more of a supply-side policy rather than being demand-driven.[23] Moreover, Columns (4) and (6) in Table 2 provide suggestive evidence that neither the allocation of health posts built following the NHP nor the quality of health posts are systematically correlated with prior child mortality outcomes.

[22] Duflo (2001) employs a similar strategy to evaluate the effects of a large scale schooling reform in Indonesia by using the scheduled number of schools built following the reform.

[23] Other factors such as higher education and better employment opportunities can increase the demand for health services. If the establishment of health posts following the NHP were demand-driven, we would expect more health posts to be established in districts that already had a relatively high number of health posts.

5 Results

In this section we present reduced form estimates that describe the relationship between infant (child) mortality rates and the availability of health posts and the quality of health posts. We specifically focus on differences in mortality rates by gender.

5.1 Primary Health Care Effects: Child Mortality and the Number and Quality of Health Posts

Tables 4 and 5 present our estimates of the effect of the NHP at the local level on mortality outcomes for children of age 0–1 and 0–3 for boys and girls separately. The odd numbered columns include an intensity measure of health posts in a district, which is defined as the number of newly built health posts following the reform (per 1,000 people), whereas the even numbered columns include the number of electoral constituencies in a district per 1,000 people. These measures are interacted with a post policy birth year dummy variable. All columns include the interaction between the health post quality measures and the post NHP indicator variable. Columns 1, 2, 5 and 6 show results from the parsimonious model specification while columns 3, 4, 7 and 8 add geographic controls (interacted with the child's year of birth dummies) and individual characteristics (ethnicity, religion, mother's age, and child birth order dummies).

Table 4 shows the results for both, infant mortality (0–1 year olds) and child mortality (0–3 year olds) for males. Columns 1 and 3 show that although the interaction between the number of health posts (per 1,000 people) and the post policy birth indicator is negative, the coefficients are statistically insignificant at any conventional levels.

On the other hand, the interaction term coefficients between the Core-Infrastructure-Quality factor variable with the post policy birth indicator (births after 1991) are not only negative across all columns but also statistically significant at the 1 percent level. These coefficients do not change when adding geographic and individual level controls in Columns 3, 4, 7 and 8. The results that additional health posts do not reduce mortality if they lack adequate infrastructure is not surprising. Specifically, the interaction coefficient between Core-Infrastructure-Quality and the post indicator variable in Column (1) suggests that a unit increase in the Core-Infrastructure-Quality variable will reduce infant mortality of boys by 4.2 percentage points. In other words, an increase in the score value of Core-Infrastructure-Quality by one standard deviation reduces child mortality by 3.43 percentage points.[24]

Similarly, the interaction term coefficients between the Reproductive-Care-Quality factor variable with the post policy birth indicator (births after 1991) is negative and statistically signifi-

[24] This is calculated as the product of the standard deviation of Core-Infrastructure-Quality, which is 0.819 as shown in Table 1, and the point estimate of the interaction coefficient in Table 4, which is -0.0422.

cant at the 5 percent level when the dependent variable measures infant mortality (year 0–1) in columns 1–4 of Table 4. The Reproductive-Care-Quality coefficient is about half the size of the Core-Infrastructure-Quality coefficient. When using child mortality (years 0–3) as the dependent variable, the Reproductive-Care-Quality variable becomes smaller and are no longer statistically significant at the conventional levels. It is probably not surprising that the quality of certain services related to childbirth or care of a pregnant mother has less of an effect on the group including older children between 2–3 years of age.

Table 5 is structured similarly to Table 4 but presents the mortality outcomes of girls. The striking difference between the results of girls is that in addition to the number of health posts, both quality measures, Core-Infrastructure-Quality and Reproductive-Care-Quality, indicate no significant negative effect on girls' infant (0–1) or child mortality (0–3). In fact, the interaction coefficient between the quality measures and the post policy birth indicator tends to be positive, indicating that daughters' mortality may have increased following the reform. However, these coefficients are imprecisely estimated and not statistically significant.

Put together, the results from Tables 4 and 5 suggest that the number of health posts does not affect child mortality outcomes but improvements in health care quality offered via better equipped health posts do reduce the mortality rates of boys but not of girls.

The validity of the findings presented in Tables 4 and 5 is governed by the assumption that in the absence of the NHP, the child mortality rates in districts with better quality health posts would not be systematically different from districts with poor quality health posts. One possibility is that although more health posts were placed in districts with low numbers of health posts prior to the NHP as shown in Table 2, the districts that received higher quality health posts might already have experienced a decreasing trend in mortality before the program.[25] However, Column (6) in Table 2 indicates that variation in Core-Infrastructure-Quality cannot be explained by child mortality outcomes prior to the policy. Although the identification assumption cannot be fully tested, we conduct several validation exercises to help understand the validity of the identification strategy.

5.2 Falsification Exercise

First, we present a falsification exercise where we use the same explanatory variables as depicted in equation 1 but only consider births that took place before the NHP (from 1980–1990). We then artificially choose the policy year as 1986 and treat children born between 1980–1985 as the placebo control group and children born between 1986–1990 as the placebo treatment group. The interactions between the placebo variable Post, which now is an indicator reflecting the artificial policy year, and HP_d (health posts per 1000 people) along with quality measures allow us to track

[25] Figure 6 shows that districts with higher Core-Infrastructure-Quality scores are dispersed and not concentrated in a specific region. This suggests that the quality of health posts is not systematically related to regions.

whether there exists any systematic relationship between the allocation or quality of health posts following the NHP with the mortality trends of infants (0-1 years of age) or children under 3 years of age in the 1980s. Since the NHP primarily targets reducing infant and child mortality as a policy objective, the quality and allocation of health posts in the post NHP period could be driven by pre-policy trends in child mortality, particularly in the 1980s. The results from the falsification exercise show that the interaction term coefficients between Core-Infrastructure-Quality and the Post indicator are close to zero and statistically insignificant at the conventional levels for both boys (in Table 6) and girls (Table 7). We do not detect any significant evidence of preexisting trends in child mortality between districts that received high quality health posts compared to districts with low quality health posts in the 1980s.

5.3 Event Study

In order to further assess the validity of our identification strategy pertaining to primary care, we estimate an event study specification expressed as:

$$\begin{aligned} \text{Mortality}_{imbd} = \alpha &+ \sum_{b=1980}^{1996} I_{b \neq 1991} \times \eta_b \times (\Delta\text{HP}_d \times Y_{ib}) \\ &+ \sum_{b=1980}^{1996} I_{b \neq 1991} \times \delta_b \times (\text{Core-Infrastructure-Quality}_d \times Y_{ib}) \\ &+ \sum_{b=1980}^{1996} I_{b \neq 1991} \times \beta_b \times (\text{Reproductive-Care-Quality}_d \times Y_{ib}) \\ &+ \psi_d + \omega_b + \gamma \times X_{mbd} + \theta \times G_{imd} + \varepsilon_{imbd}, \end{aligned} \tag{2}$$

where all variables are identical to equation 1, except that Y_{ib} is an indicator variable equal to one if the birth year of a child i is b for $b \in \{1980, 1981, ..., 1990, 1992, ...1996\}$ and the omitted category are children born in year 1991 indicated by $I_b \neq 1991$. The coefficients of interest are η_b, δ_b, and β_b.

If there exist any pre-policy differences in mortality trends across districts, then such trends will be captured by the coefficients pertaining to children's year of birth before 1991. As a necessary but not sufficient condition for the validity of our identification strategy, the coefficient estimates of the interaction terms of health posts and quality measures (represented by the two factor variables) should be close to zero and statistically insignificant at any conventional levels prior to the reform.[26]

The estimation results are shown in Figures 7(a)–7(d) separately for boys and girls. The interaction terms pertaining to Core-Infrastructure-Quality (δ_b in equation 2) are plotted along with their respective confidence bounds obtained from the standard errors clustered at the district level. The results pertaining to boys in panels 7(a) and 7(b) show that prior to the reform in 1991 the interac-

[26]Note that the event study exercise can only provide suggestive evidence regarding the validity of the identification used. In this spirit, we conduct additional robustness checks in Section 5.4.

tion coefficients between the year of birth indicators and the Core-Infrastructure-Quality fluctuate around zero. However, for births that took place after the NHP, we observe a break in trend as the interaction term coefficients decrease, suggesting a decrease in infant and child mortality of boys. These effects are similar in magnitude between birth years 1992–1996.

Strikingly, the event study graphs in Figures 7(c) and 7(d) for girls show no significant effects. The interaction term coefficients fluctuate around zero and are not significantly different from zero for birth years before and after the NHP. Consistent with Tables 6 and 7, these figures provide no evidence of systematic differences in child mortality trends prior to the reform across districts that received higher quality health posts compared to districts that received lower quality health posts.

5.4 Robustness and heterogeneous effects

In this section we add a series of statistical checks to highlight the robustness of our results. We first add additional controls to our main specification in equation 1. In order to keep the exposition short, we only show the main estimator of interest which is the effect of Core-Infrastructure-Quality interacted with Post (parameter δ in expression 1) on infant mortality, separately by gender along with the 95 percent confidence bounds.

The first entry at the bottom of Figure 8(a) is labeled "1. Preferred" and shows our preferred benchmark estimate from column 4 in Table 4. In Row 2 we add interaction terms between child birth year dummies and the household's distance to the closest paved road (as reported in the survey) in addition to interaction terms of child birth year dummies and a distance measure to the nearest school. These terms control for variation based on infrastructure that may have been available at the time of the child's birth and which could have been built concurrently with the health care reform. In Row 3, we limit the sample to include only the first and second born children as birth order related to gender of these children is more likely to be exogenous and not dependent on past history of births. In Row 4, we add district specific linear time trends by interacting district of birth dummies with a child's year of birth. In Row 5, we use VDC specific fixed effects. Row 6 adds district-wise change in the number of schools between 1990 and 1996 (per 1000 people) interacted with child birth year dummies to account for the schooling boom in the 1990s. In Row 7, we use the household sampling weights of the NLSS 1996. Our main results are robust with respect to these specification changes or changes in the samples.

In Figure 8(b) we keep the econometric specifications from expressions 1 but change the underlying sample to explore heterogeneous effects of the reform by specific subgroups. We again only show the main estimator of interest which is the effect of Core-Infrastructure-Quality interacted with Post (parameter δ in expression 1) on infant mortality. Row 1—labeled "1. Preferred"—again shows our preferred benchmark estimate from column 4 in Table 4 using the core sample as a reference point. In Row 2, we restrict the sample to observations from households belonging to the

Brahmin and Chhetri castes. These are both so called "high" castes whose members have above average education and income.[27] Row 3 includes people from castes other than Brahmin and Chhetri. The effects of better quality primary health care following the reform are more concentrated among children from caste groups other than Brahmin and Chhetri. The marginal benefits of the reform are likely to be higher for children belonging to "other caste groups" as the pre-reform infant mortality rate is noticeably higher for these children compared to the Brahmin and Chhetri caste groups (0.11 vs. 0.067). Row 4 includes children born to mothers who report that their ideal number of children is less than or equal to 3 (median number), whereas Row 5 only includes children born to mothers who report that their ideal number of children is larger than 3. Comparing the two we find that the effects of the program are similar across groups with different family size preferences. In Row 6, we restrict the sample to households whose primary travel mode is "by foot," while Row 7 includes households with access to other means of transportation such bikes, cars, or buses. Finally, in Row 8, we only include members of the Hindu religion, which is by far the largest religious group in our sample (87.5 percent).[28]

Additional robustness checks reveal that our results also hold with respect to the addition of controls for concurrent development projects, the schooling boom in the early 1990s, and alternative health post quality measures such as (*i*) aggregating health quality measures at the village (VDC) rather than the district level, (*ii*) using health quality measures such as availability of doctors and beds in health posts directly as opposed to quality measures obtained from dimension reduction methods, and (*iii*) health quality measures based on principal component analysis (PCA). These results are available in the Online Appendix.

6 Conclusion

In many developing countries households tend to invest more in sons than daughters due to a combination of *preference biases* and *labor market biases*. There exists well documented evidence about significant differences in nutritional investments into children by gender, sex-selective abortion, and skewed birth sex ratios. In situations where cultural norms dictate gender preferences, a homogeneously implemented health reform aimed at increasing access to healthcare, may disproportionately benefit boys compared to girls. Using the National Health Policy (NHP) of Nepal in 1991, we explore the effects of a country wide health reform on child health outcomes across gender. Our outcome variables are infant and child mortality outcomes constructed using data from

[27]Compare the discussions based on India in Mandelbaum (1988) and Pande and Astone (2007). The latter finds a higher degree of son preference in low caste women in a rural sample of the National Family Health Survey, India (NFHS-1), 1992–1993. The authors attribute their finding to "Sanskritization," where lower castes imitate higher castes.

[28]The next two religious groups in our sample in terms of size are Buddhists with 5.4 percent and Muslims with 3.8 percent.

the Nepal Living Standard Survey (1996).

We find that although the reform has increased the number of health post at the local (and primary health care) level, the aggregate number of established health posts in a mother's district has no effect on her children's mortality outcomes. However, improvements in the quality of newly established health posts attributed to core infrastructure such as beds and doctors, do reduce the mortality rates of boys but not of girls.

We document that such a disproportionate gender effect of the reform is concentrated among children in households from caste groups other than Brahmin and Chhetri. This finding can be explained by considerably higher pre-reform infant mortality rates among children from these other caste groups (11%) compared to children belonging to Brahmin and Chhetri households (6.7%). After the implementation of the reform in 1991 infant mortality of sons from other caste groups improved substantially (from 11% to 6.9%) but no such improvements are seen among daughters (infant mortality rate of 9.8% both pre- and post-reform). While the reform may have abridged the gap in sons' infant mortality between the more dominant vs. marginalized caste groups, it was ineffective in reducing caste-based health inequity among daughters.

We attribute the results from this study to two channels. First, we show that women prefer to have more sons compared to daughters, indicating a strict gender preference which could result in the observed differences in mortality outcomes between boys and girls. The NLSS survey, based on a 1995–1996 rural sample, reveals a strong gender preference for boys as the women in the sample prefer a lower number of girls—the mode of the ideal number of boys is 2, whereas the mode for girls is 1.

Second, households tend to neglect daughters' health on average as demonstrated by the higher immunization rate of boys. Additional auxiliary analysis show that sons are about 4.3 percentage points more likely to be immunized compared to daughters.[29] For biological reasons, boys are more vulnerable to disease than girls at very young ages. The marginal benefit of visiting a health post could therefore be higher for boys than for girls. Although the biological advantage among girls is well-defined in the medical literature, several studies highlight that the biological advantage of girls erodes in income poor environments and if societal norms favor sons (Chowdhury et al., 2017; Jayachandran, 2015; Sawyer, 2012; Costa et al., 2017). These studies provide evidence of strong gender bias at birth that lead to skewed sex ratios—especially in India and China—as prominently highlighted by Sen (1990). While most studies focus on India and China, gender bias and son preference exist in many parts of Southern Asia (Costa et al., 2017), including Nepal as we can directly see from the NLSS. Such biased preferences, favoring boys over girls, could explain differences in human capital investments across gender.

[29]Households with strong son preferences are also more inclined to borrow in an effort to increase the level of investment into their sons. We demonstrate this economic channel more formally in a theoretical model in the Online Appendix.

What is important to understand from a policy perspective is that societal settings and cultural norms that marginalize sub-groups based on gender, caste, and ethnicity should be carefully considered when designing health reforms in developing countries if the goal of the reform is to enhance the well being of the general populace. If these effects are ignored, we demonstrate that large scale institutional reforms can in fact increase observed disparities in outcomes.

Finally, although our analysis is focused on evaluating the effects of the reform at the primary care level, the NHP also operates at the tertiary care level and allocates the construction and improvements of district level and zonal hospitals. Our preliminary analysis shows that the reduction in distance to zonal hospital following the reform is associated with a reduction in infant (child) mortality among sons. However, these findings are based on the assumption that the NHP has improved the quality of hospitals. Due to lack of data regarding hospital quality in the 1990s, we cannot directly identify this channel and therefore leave the investigation of tertiary care effects of the NHP for future research.

Declaration of Interests

Funding: The study was not supported by any funding source.
Conflict of Interest: The authors declare that they have no conflict of interest.

References

Agarwal, Kokila. 1998. "Reproductive Health Case Study Nepal." The Policy Project, The Futures Group International in collaboration with Research Triangle Institute (RTI) The Centre for Development and Population Activities (CEDPA). 12

Asfaw, Abay, Stephan Klasen and Francesca Lamanna. 2007. "Intra-Household Gender Disparities in Children's Medical Care Before Death in India." *Institute for the Study of Labor (IZA) Discussion Paper* (2586). 5

Ashraf, Nava, Natalie Bau, Nathan Nunn and Alessandra Voena. 2019. "Bride Price and Female Education." *Journal of Political Economy* 128(2):591–641. 5

Barcellos, S.H., L.S. Carvalho and A. Lleras-Muney. 2014. "Child Gender and Parental Investments in India: Are Boys and Girls Treated Differently?" *American Economic Journal: Applied Economics* 6(1A):157–189. 4

Behrman, Jere and Anil Deolalikar. 1990. "The Intrahousehold Demand for Nutrients in Rural South India: Individual Estimates, Fixed Effects, and Permanent Income." *Journal of Human Resources* 25(4):665–696. 4

Bernal, Noelia, Miguel A. Carpio and Tobias J. Klein. 2017. "The Effects of Access to Health Insurance: Evidence from a Regression Discontinuity Design in Peru." *Journal of Public Economics* 154:122–136. 4

Bharadwaj, Prashant, James Fenske, Rinchan Ali Mirza and Namrata Kala. 2020. "The Green Revolution and Infant Mortality in India." *Journal of Health Economics* forthcoming. 4

Bharadwaj, Prashant and Leah K. Lakdawala. 2013. "Discrimination Begins in the Womb: Evidence of Sex-Selective Prenatal Investments." *Journal of Human Resources* 48(1):71–113. 4, 5

Buchmueller, Thomas, John Ham and Lara Shore-Sheppard. 2015. "The Medicaid Program." NBER Working Paper No. 21425. 5

Carney, Stephen and Min Bahadur Bista. 2009. "Community Schooling in Nepal: A Genealogy of Education Reform Since 1990." *Comparative Education Review* 53(2):189–211. 52

Cattell, Raymond B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research* 1(2):245–276. 46

Chakravarty, Abhishek. 2010. "Supply Shocks and Gender Bias in Child Health Investments: Evidence from the ICDS Programme in India." *The B. E. Journal of Economic Analysis & Policy* 10(1):1–26. 4

Chen, Lincoln C, Emdadul Huq and Stan d'Souza. 1981. "Sex Bias in the Family Allocation of Food and Health Care in Rural Bangladesh." *Population and Development Review* pp. 55–70. 2

Chowdhury, Ranadip, Sunita Taneja, Sarmila Mazumder, Nita Bhandari and Tor A Strand. 2017. "Gender Differences in Infant Survival: A Secondary Data Analysis in Rural North India." *BMJ open* 7:e014179. 20

Collier, Paul. 2017. "Culture, Politics, and Economic Development." *Annual Review of Political Science* 20(1):111–125. 5

Costa, Janaína Calu, Fernando C Wehrmeister, Aluísio JD Barros and Cesar G Victora. 2017. "Gender Bias in Careseeking Practices in 57 Low–and Middle–Income Countries." *Journal of Global Health* 7(1). 20

Costello, Anna B and Jason W Osborne. 2005. "Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis." *Practical Assessment, Research & Evaluation* 10(7):1–9. 46

Dahl, Gordon, B. and Enrico Moretti. 2008. "The Demand for Sons." *The Review of Economic Studies* 75:1085–1120. 4

Del Boca, Daniela, Christopher Flinn and Matthew Wiswall. 2013. "Household Choices and Child Development." *Review of Economic Studies* 81(1):137–185. 63

Devkota, Rohini P. 2014. "Climate change: trends and people's perception in Nepal." *Journal of Environmental Protection* 5(04):255. 8

Dow, William H. and Kammi K. Schmeer. 2003. "Health insurance and child mortality in Costa Rica." *Social Science & Medicine* 57(6):975–986. 4

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91(4):795–813. 14

Echávarri, R. and J. Husillos. 2016. "The Missing Link Between Parents' Preferences and Daughters' Survival: The Moderator Effect of Societal Discrimination." *World Development* 78:372–385. 4

Elborgh-Woytek, Katrin, Monique Newiak, Kalpana Kochhar, Stefania Fabrizio, Kangni R. Kpodar, Philippe Wingender, Benedict J. Clements and Gerd Schwartz. 2013. "Women, Work, and the Economy : Macroeconomic Gains from Gender Equity." *IMF Staff Discussion Notes* (13/10). 4

Fitzpatrick, Anne. 2018. "The Price of Labor: Evaluating the Impact of Eliminating User Fees on Maternal and Infant Health Outcomes." *AEA Papers and Proceedings* 108:412–415. 4

Goodman-Bacon, Andrew. 2021. "The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes." *American Economic REview* 111(8):2550–2593. 5

Grown, Caren, Geeta Rao Gupta and Rohini Pande. 2005. "Taking Action to Improve Women's Health through Gender Equality and Women's Empowerment." *The Lancet* 365(9458):541–543. 5

Hall, Robert E. and Charles I. Jones. 2007. "The Value of Life and the Rise in Health Spending." *Quarterly Journal of Economics* 122(1):39–72. 66, 67, 69, 70, 71

Hazarika, Gautam. 2000. "Gender Differences in Children's Nutrition and Access to Health Care in Pakistan." *The Journal of Development Studies* 37(1):73–92. 5

ILO. 2016. *Women at Work Trends 2016*. International Labour Office - Geneva: International Labor Organization. 4

Jayachandran, Seema. 2015. "The Roots of Gender Inequality in Developing Countries." *Annual Review of Economics* 7(1):63–88. 2, 20

Jayachandran, Seema and Ilyana Kuziemko. 2011. "Why do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India." *The Quarterly Journal of Economics* 126(3):1485–1538. 2

Jayachandran, Seema and Rohini Pande. 2017. "Why Are Indian Children So Short? The Role of Birth Order and Son Preference." *American Economic Review* 107(9):2600–2629. 2, 4

Kaiser, Henry F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20(1):141–151. 46

Lin, Ming-Jen, Jin-Tan Liu and Nancy Qian. 2014. "More Missing Women, Fewer Dying Girls: The Impact of Sex-selective Abortion on Sex at Birth and Relative Female Mortality in Taiwan." *Journal of the European Economic Association* 12(4):899–926. 4

Mandelbaum, David G. 1988. *Women's Seclusion and Men's Honor: Sex Roles in North India, Bangladesh, and Pakistan*. University of Arizona Press. 19

Matsushima, Midori, Hiroyuki Yamada and Yasuharu Shimamura. 2020. "Analysis on Demand- and Supply-Side Responses during the Expansion of Health Insurance Coverage in Vietnam: Challenges and Policy Implications toward Universal Health Coverage." *Review of Development Economics* 24(1):144–166. 4

McKinnon, Britt, Sam Harper, Jay S Kaufman and Yves Bergevin. 2015. "Removing user fees for facility-based delivery services: a difference-in-differences evaluation from ten sub-Saharan African countries." *Health Policy and Planning* 30(4):432–441. ZSCC: 0000134. 4

Morrisson, Christian and Johannes P. Jütting. 2005. "Women's Discrimination in Developing Countries: A New Data Set for Better Policies." *World Development* 33(7):1065 – 1081. 4

Nations, United. 2015. *The World's Women 2015: Trends and Statistics*. Number Sales No. E.15.XVII.8. New York: United Nations, Department of Economic and Social Affairs, Statistics Division. 4

Nepal Law Commission. 1991. "National Health Policy, 2048 (1991)." Ministry of Health. 11

Nepal, WCO. 2007. "Health System in Nepal: Challenges and Strategic Options.". 2, 3

Oster, Emily. 2009. "Does Increased Access Increase Equality? Gender and Child Health Investments in India." *Journal of Development Economics* 89(1):62–76. 5

Pande, Rohini P. 2003. "Selective Gender Differences in Childhood Nutrition and Immunization in Rural India: The Role of Siblings." *Demography* 40(3):395–418. 4

Pande, Rohini P. and Nan Marie Astone. 2007. "Explaining Son Preference in Rural India: The Independent Role of Structural versus Individual Factors." *Population Research and Policy Review* 26(1):1–29. 19

Pharoah, Paul. 2017. "Development and Validation of a New Walking Pace Function Using Crowd-Sourced GPS Data." *bioRxiv* p. 188417. 10

Rai, Shiba K., Ganesh Rai, Kazuko Hirai, Ayako Abe and Yoshimi Ohno. 2001. "The Health System in Nepal—An Introduction." *Environmental Health and Preventive Medicine* 6(1):1–8. 11

Rose, Elaina. 2000. "Gender Bias, Credit Constraints and Time Allocation in Rural India." *The Economic Journal* 110(465):738–758. 4

Rose, Elaina. 2017. Child Gender and the Family. In *The Oxford Handbook of Women and the Economy*, ed. Saul D. Hoffman Susan L. Averett, Laura M. Argys. Oxford University Press chapter 11, pp. 259–280. 4

Sawyer, Cheryl Chriss. 2012. "Child Mortality Estimation: Estimating Sex Differences in Childhood Mortality Since the 1970s." *PLoS Medicine* 9(8):e1001287. 20

Sen, Amartya. 1990. "More than 100 Million Women are Missing." *The New York Review of Books* 37(20):61–66. 2, 20

Sen, Amartya. 1992. "Missing Women." *BMJ: British Medical Journal* 304(6827):587. 2

Shah, Arjun. 1996. *Health Budget and Expenditure: A Trend Analysis*. Kathmandu: GTZ (German Technical Cooperation. 27

Shah, Manisha and Bryce Millett Steinberg. 2020. "Workfare and Human Capital Investment: Evidence from India." *Journal of Human Resources* forthcoming. 5

Shrestha, Ishwar B and Laxmi R Pathak. 2012. "Review of National Health Policy 1991.". 11

Shrestha, Vinish. 2019. "Can Basic Maternal Literacy Skills Improve Infant Health Outcomes? Evidence from the Education Act in Nepal.". 52

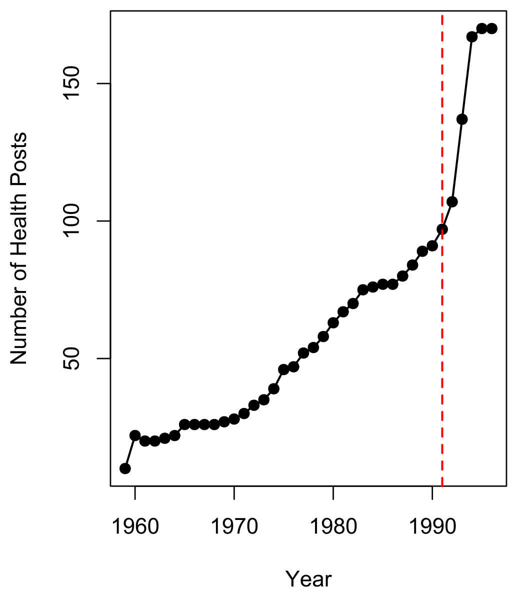
Shrestha, Vinish and Rashesh Shrestha. 2017. "Multigenerational Effects of Education Reform: Mother's Education and Children's Human Capital in Nepal.". 52

Tanaka, Shinsuke. 2014. "Does Abolishing User Fees Lead to Improved Health Status? Evidence from Post-apartheid South Africa." *American Economic Journal: Economic Policy* 6(3):282–312. 5

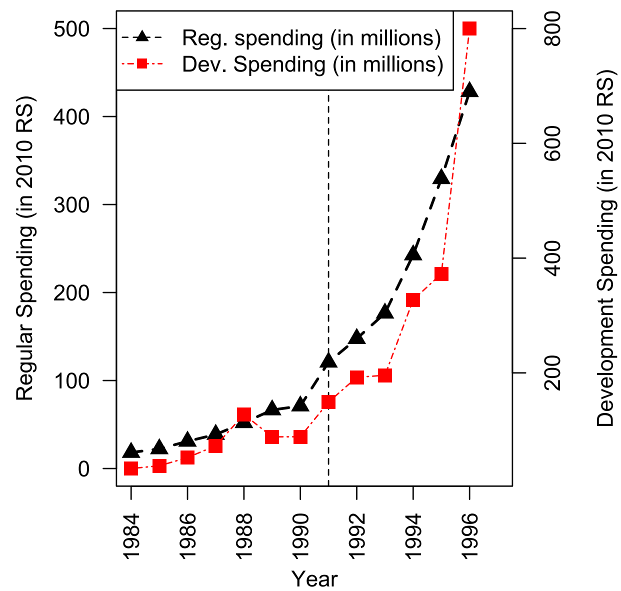
Thornton, Rebecca L., Laurel E. Hatt, Erica M. Field, Mursaleena Islam, Freddy Solís Diaz and Martha Azucena González. 2010. "Social Security Health Insurance for the Informal Sector in Nicaragua: A Randomized Evaluation." *Health Economics* 19(S1):181–206. 4

YoungChoi, Jin and Sang-Hyop Lee. 2006. "Does Prenatal Care Increase Access to Child Immunization? Gender Bias among Children in India." *Social Science & Medicine* 63(1):107–117. 4

Figures



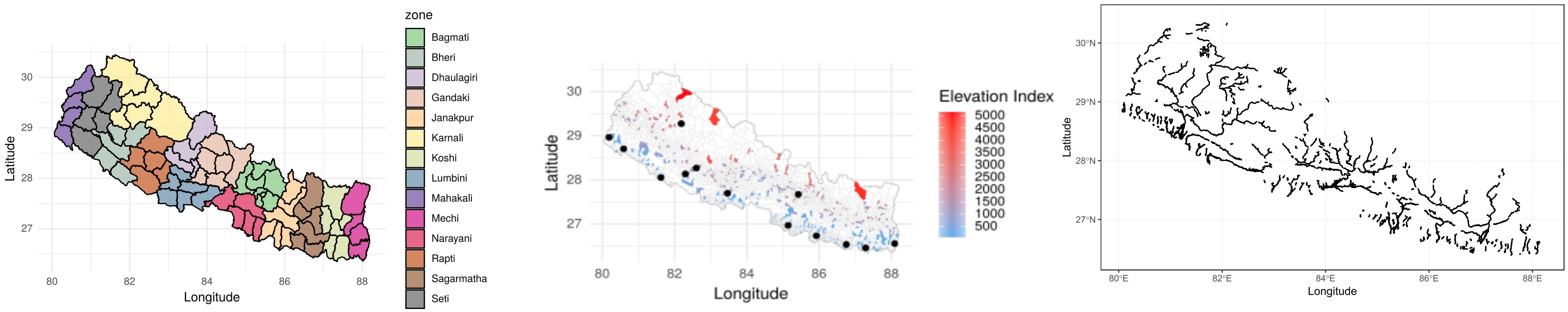
(a) Total Number of Health Posts



(b) Health Expenditure

Figure 1: **Number of Health Posts and Health Expenditures Over Time**

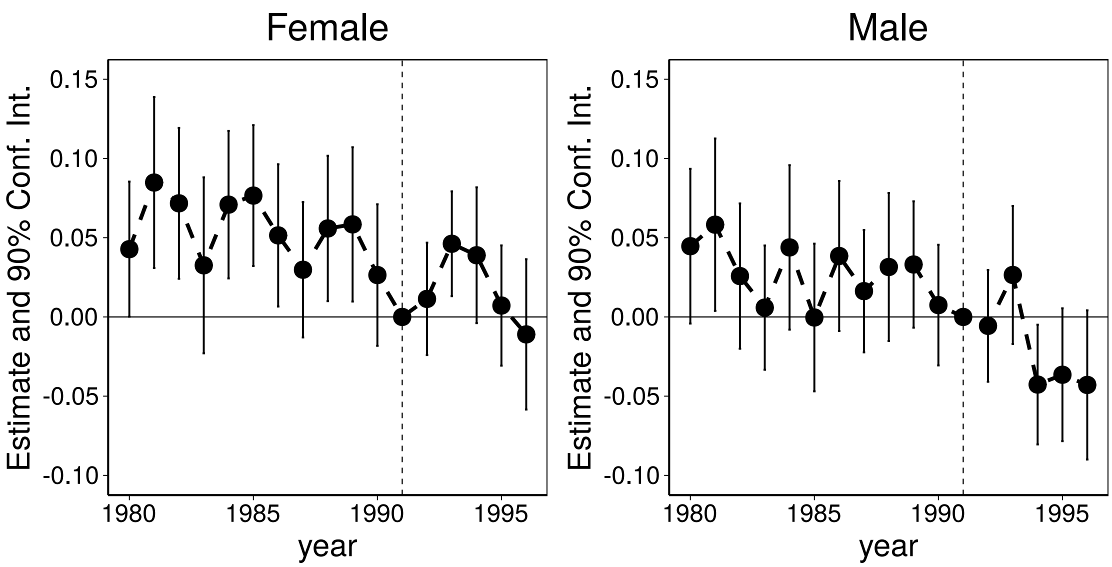
Notes: **Panel (a)** shows the total number of health posts annually between 1959–1996.
Panel (b) show health spending time trends. Regular spending is defined as re-occurring expenditure to maintain a certain level of health care. Development spending are categorized as one-time expenditures or investments.
The data source for Panels (a) is the NLSS 1996 (household and community files) and the source for Panel (b) is Shah, 1996.



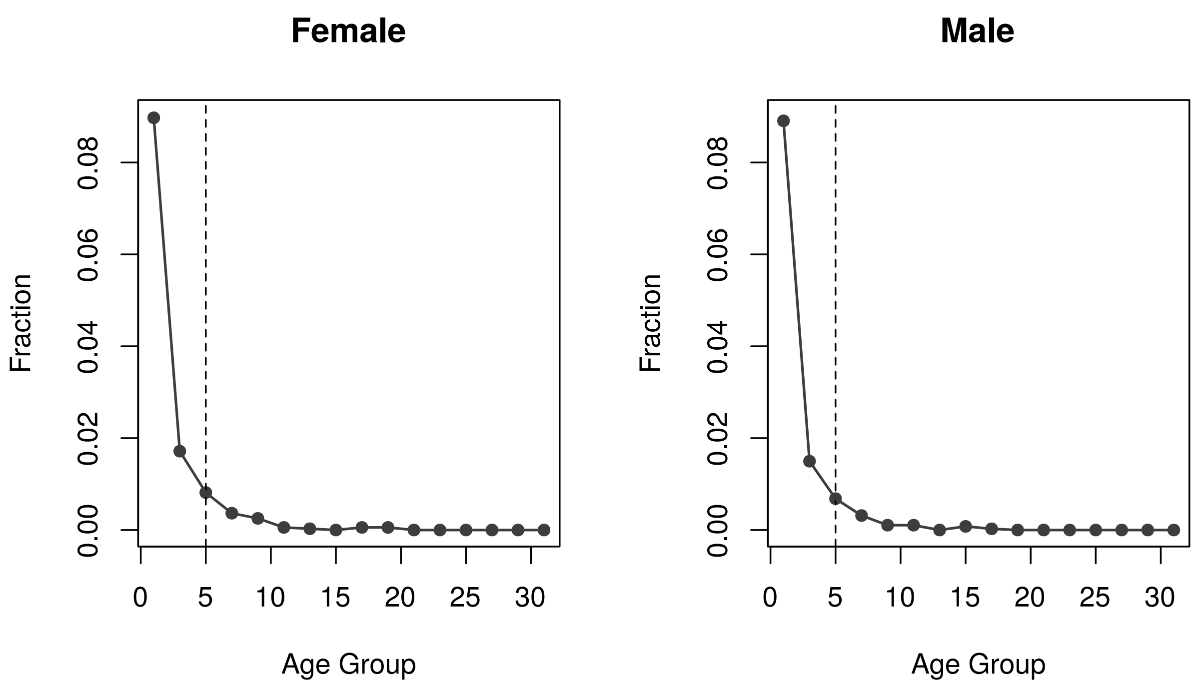
(a) Zones (b) Elevation (c) Rivers

Figure 2: **Zonal Divisions, VDC Elevation, Location of Zonal Hospitals, and Rivers**

Notes: **Panel (a):** Zonal divisions are indicated by colors in the legend. Nepal has 75 districts. The administrative area of districts comprise a number of village development committees (VDCs). As of 1991 there are approximately 4,000 VDCs in Nepal. These VDCs are divided into wards, which are the smallest administrative unit of the country. VDCs are also grouped to form *ilaka* which are electoral constituencies. Each *ilaka* consists of 4–5 VDCs. Data Source: GADM Maps and Data.
Panel (b): Zonal hospitals from right to left: 1) Mechi ZH (Bhadrapur, Jhapa), 2) Koshi ZH (Biratnagar), 3) Sagarmatha ZH (Rajbiraj, Saptari), 4) Janakpur ZH (Janakpur), 5) Bhaktapur (Central Hospital), 6) Mahendra (Central Hospital), 7) Lumbini ZH (Butwal, Lumbini), 8) Dhaulagiri ZH (Baglung), 9) Rapti ZH (Tulsipur, Dang), 10) Karnali ZH (Jumla), 11) Bheri ZH (Nepalgunj), 12) Seti ZH (Dhangadhi), 13) Mahakali ZH (Kanchanpur). Here, ZH stands for zonal hospital. Additionally, the two central hospitals are also included (5 and 6).
Panel (c): The names of the large river basins from left to right are: 1) Mahakali Basin; 2) Karnali Basin; 3) Babai Basin; 4) West Rapti; 5) Gandaki Basin; 6) Bagmati Basin; 7) Kamala Basin; 8) Koshi Basin; 9) Kankai Basin; and 10) Other Southern Rivers.
Data Sources for Panels (b) and (c): Google Maps API, The Humanitarian Data repository, NLSS 1996.



(a) Time Trend in Mortality by Gender



(b) Mortality by Age

Figure 3: **Time Trend and Age Trend in Mortality Rates by Gender**

Notes: **Panel (a)** plots the coefficients on year indicators obtained from a regression of infant mortality on year indicators (omitted year 1991) and district of birth fixed effects and plots the estimated coefficients of year indicators.
Panel (b) shows the mortality rates of females and males by age group.
The data source for both panels is the NLSS 1996 (household and community files).

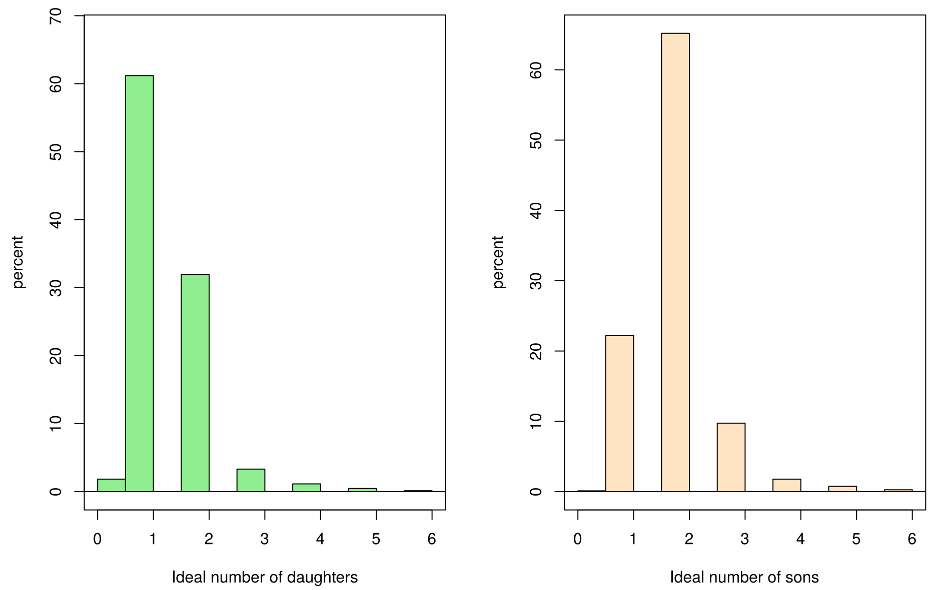


Figure 4: **Ideal Number of Children**

Notes: Data is from the Nepal Living Standard Survey (NLSS 1996) and based on authors' calculations. The figure shows preferences inclined towards sons; higher proportion of respondents report ideally wanting more sons compared to daughters. For example, over 60% of the respondents report wanting two sons but this number decreases to 30% for daughters.

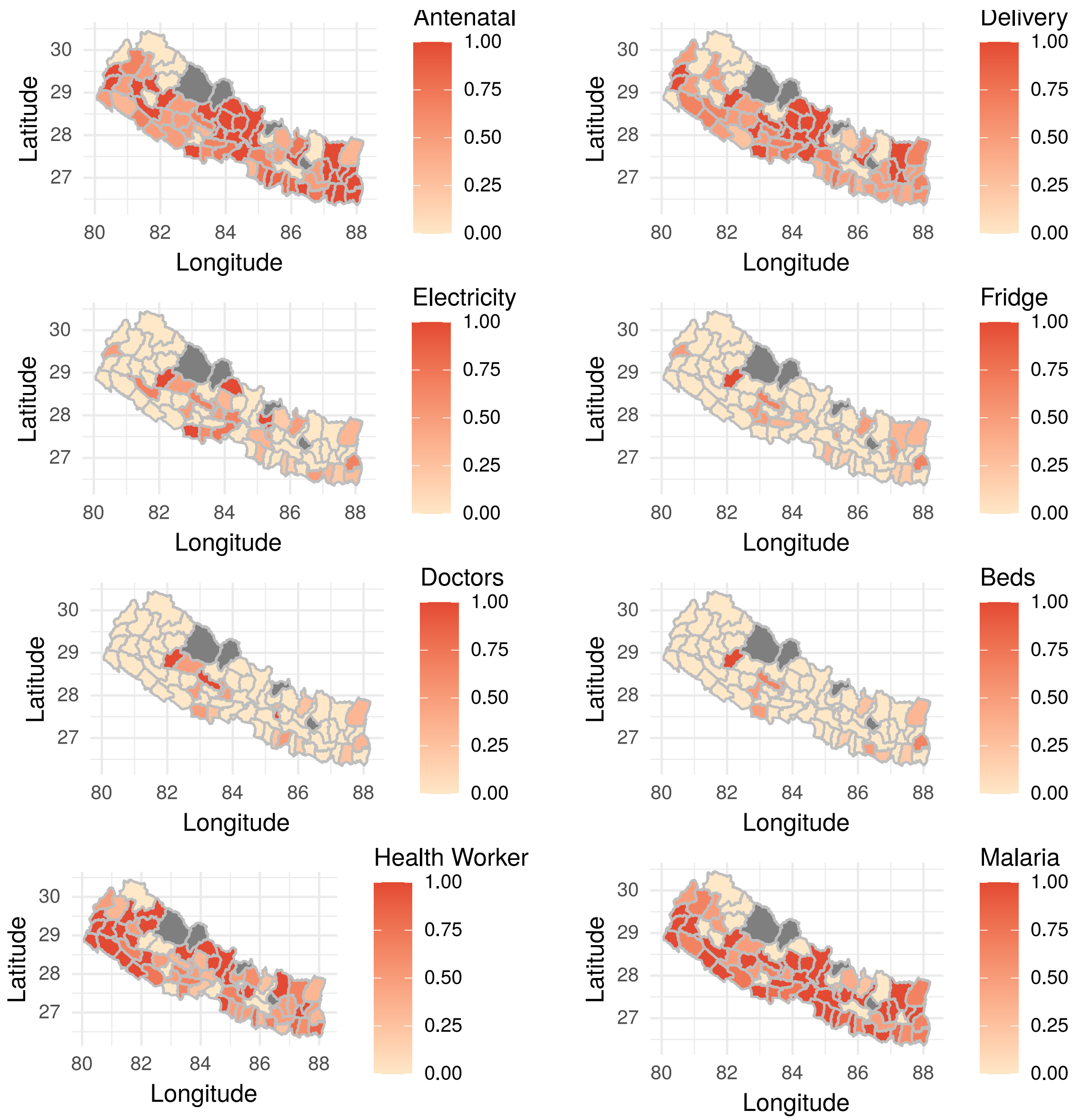
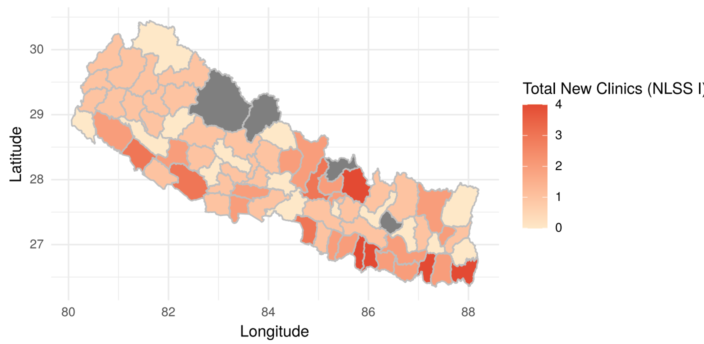
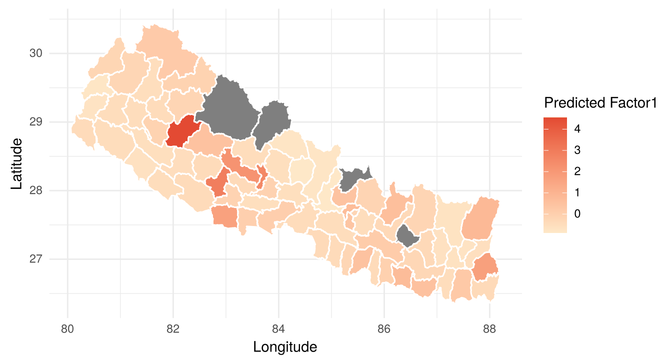


Figure 5: **Quality Measures of Newly Constructed Health Posts between 1991–1996**

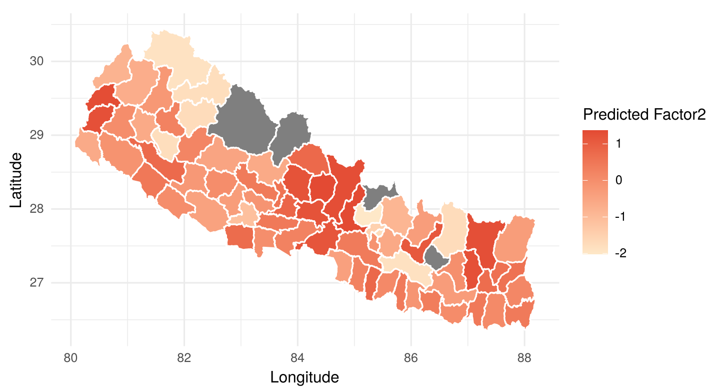
Notes: For each district these figures present the proportion of newly constructed health posts between 1991–1996 that are equipped with specific core infrastructure or offer a specific care service. If a district reports that 100 percent of its health posts offer antenatal care in the top-left panel, we color it red. If a district reports that none of its health posts offer antenatal care, we color it beige. The districts in gray lack data.
These variables are used to estimate factor scores using Exploratory Factor Analysis (EFA) described in Section 2. The two extracted factors are: **Core-Infrastructure-Quality** which refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** which refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts.
Data Source: NLSS 1996 (Community and Household Files).



(a) Number of Newly Constructed Health Posts (1991–1996)



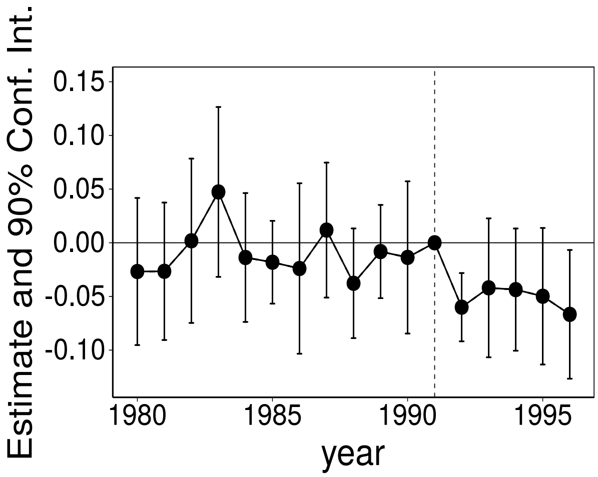
(b) Factor 1: Core-Infrastructure-Quality (HPs built between 1991–1996)



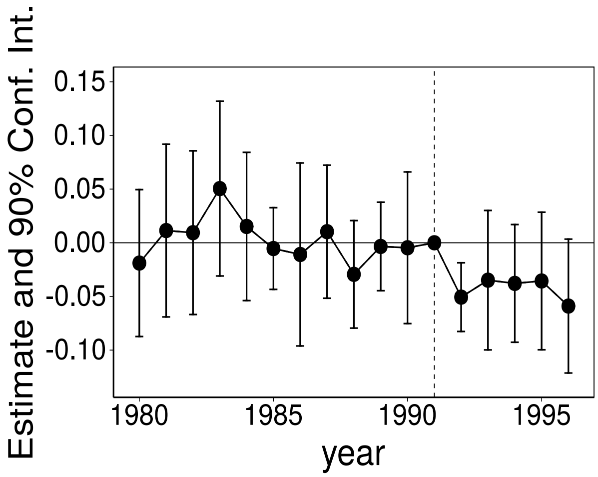
(c) Factor 2: Reproductive-Care-Quality (HPs built between 1991–1996)

Figure 6: **Number of Newly Constructed Health Posts and their Quality Based on EFA**

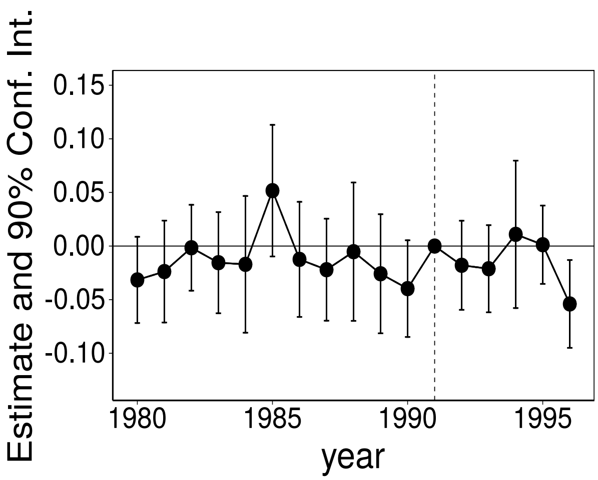
Notes: Factor scores are predicted using Exploratory Factor Analysis (EFA) described in Section 2. The districts in gray lack data. **Core-Infrastructure-Quality** (Factor 1) refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** (Factor 2) refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts.
Data Source: NLSS 1996.



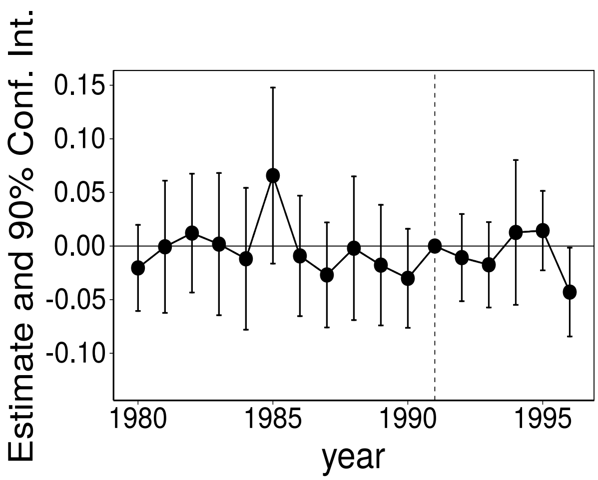
(a) Infant Mortality (Ages 0 to 1, Boys)



(b) Child Mortality (Ages 0 to 3, Boys)



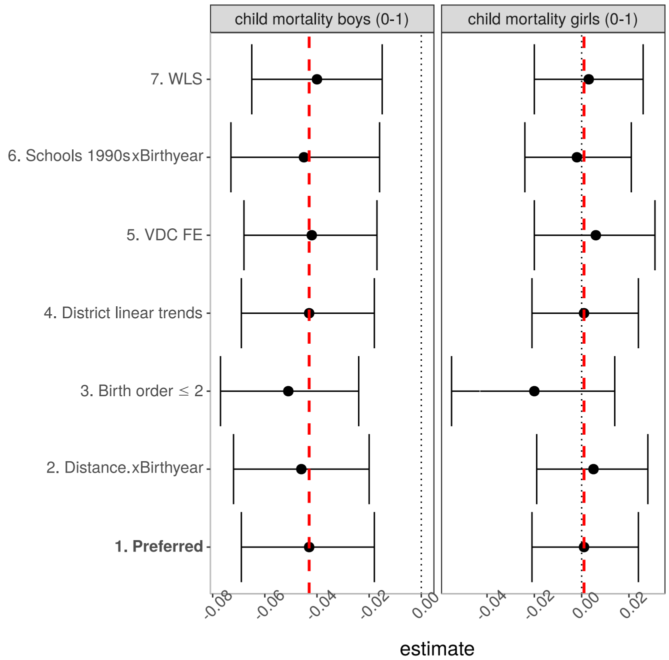
(c) Infant Mortality (Ages 0 to 1, Girls)



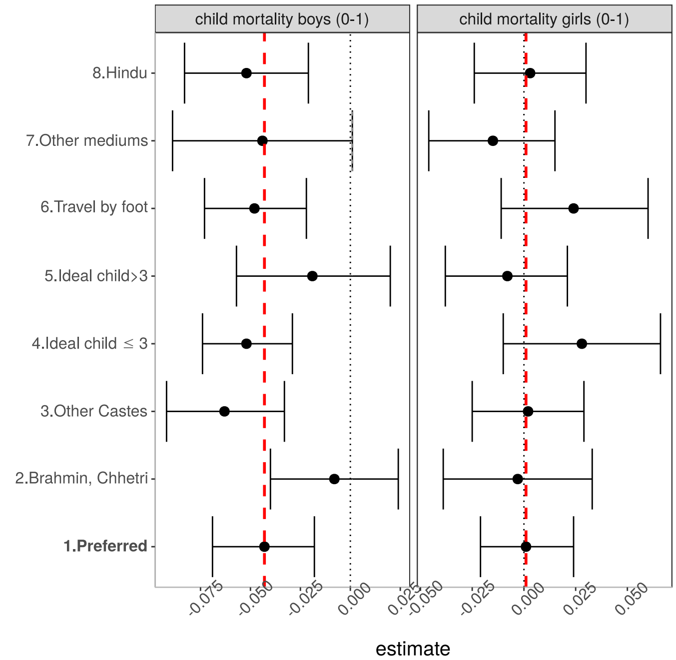
(d) Child Mortality (Ages 0 to 3, Girls)

Figure 7: **Event Study – Core-Infrastructure-Quality of Health Posts**

Notes: The sub-figures plot the estimates of δ_b (Core-Infrastructure-Quality) from equation 2 for $b \in \{1980, 1981, ..., 1990, 1992, ...1996\}$ with year 1991 as omitted category. The specification used controls for the interaction between child birth year dummies and the following variables: i) new health posts constructed following the reform (per 1,000 people), ii) Reproductive-Care-Quality, iii) log of distance to the shortest zonal hospital, iv) shortest distance to a river and its square term, and iv) VDC elevation and its square term. Additionally, the specifications control for ethnicity and religion dummies (Hindu, Buddhist) of the household as well as mother's age. The standard errors are clustered at the district level and the error bar represents the 90% confidence interval calculated using the clustered standard errors.
Data Source: NLSS 1996.



(a) Robustness w.r.t. Additional Control Variables



(b) Heterogeneous Effects based on Alternative Subsamples

Figure 8: **Robustness and Heterogeneous Effects**

Notes: Subplot (*a*) shows point estimates of the main parameter of interest (δ in expression 1) using infant mortality (0–1) as the dependent variable. **Row 1** shows the preferred estimate from column 4, Table 4. **Row 2** includes the interaction between child birth year dummies and the household's i) distance to the closest paved road, and ii) distance to the nearest school. **Row 3** only includes children of birth order two or below. **Row 4** adds district specific linear time trends. **Row 5** uses the VDC fixed effects instead of the district fixed effects. **Row 6** adds the interaction between increases in primary schools in the 1990s and birth year dummies, while **Row 7** presents the weighted least square estimate using sampling weights. The vertical red lines represent the preferred estimates, while the black dotted lines are positioned at zero. The 95% confidence intervals, based on the clustered standard errors at the district level, are represented by the error bars.
Subplot (*b*) **Row 1** shows the preferred estimate. **Row 2** restricts the sample to Brahmin and Chhetri households, while **Row 3** includes people from castes other than Brahmin and Chhetri. **Row 4** only includes children born to mothers who report ideal number of children as 3 or below (3 is the sample median), whereas **Row 5** uses the sub-sample of mothers with ideal children above 3. **Row 6** restricts the sample to households whose primary mode of transportation to amenities (e.g., paved roads, market places) is by foot, while **Row 7** includes households using other means of transportation (e.g., bike, bus, car). **Row 8** only includes people of Hindu religion. The vertical red lines represent the preferred estimates. The 95% confidence intervals, based on the clustered standard errors at the district level, are represented by the error bars.

Tables

Table 1: **Summary Statistics of Main Sample**

Statistic	N	Mean	St. Dev.
Child Alive	7,361	0.880	0.325
Mortality (Ages 0 to 1)	7,361	0.089	0.285
Mortality (Ages 0 to 3)	7,361	0.105	0.307
Mortality (Ages 4 to 9)	7,361	0.013	0.112
Birth Order	7,361	3.324	2.132
Child Gender (1=Female)	7,361	0.483	0.500
Mother's Age	7,361	34.681	9.097
Married	7,361	0.952	0.213
Never Married	7,361	0.001	0.029
Separated	7,361	0.004	0.064
Widower	7,361	0.043	0.203
Hindu	7,361	0.875	0.331
Buddhist	7,361	0.054	0.226
Muslim	7,361	0.038	0.192
Other Religion	7,361	0.011	0.106
Religion Missing	7,361	0.021	0.144
Chhetri	7,361	0.205	0.403
Brahmin	7,361	0.139	0.346
Mother Can Read	7,361	0.099	0.298
Mother Can Write	7,361	0.089	0.285
Change in Primary Schools (1981-1991)	7,361	0.143	0.080
Health Post Distance (in mins)	7,361	78.679	100.471
School Distance (in mins)	7,355	26.261	62.946
Total Health Posts (1989)	7,361	1.475	1.079
Total Health Posts (1996)	7,361	3.131	1.504
Factor 1 (Health Post Quality Measure 1)	7,361	−0.042	0.819
Factor 2 (Health Post Quality Measure 2)	7,361	0.054	0.807
Number of Health Posts (per 1,000 people)	7,361	0.007	0.007
Male Literacy Rate (ages 10-19, 1981)	7,361	48.004	12.138
Female Literacy Rate (ages 10-19, 1981)	7,361	19.773	10.130
Elevation (in meters)	7,361	882.514	868.450
Shortest Distance to Hospital (km)	7,361	45.495	27.860
Shortest Distance to River (km)	7,361	5.026	4.313

Notes: The source of data is the Nepal Living Standard Survey (1996). Elevation data is based on data obtained from the Elevation API of the Google Maps Platform. Distances are based on authors' calculations.

Table 2: **Allocation of Health Posts and Quality Measures**

	Dependent variable:							
	Health post (1991-1996)				Core-Infrastructure Quality		Reproductive-Care-Quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health posts (in 1989)	-0.603^{***}							
	(0.118)							
No health post (in 1989)		1.040^{***}						
		(0.266)						
Electoral constituencies (in 1991)			0.410^{***}					
			(0.112)					
Child mortality (ages 0-1 in 1989)				0.005		0.015		-0.032^{*}
				(0.019)		(0.018)		(0.017)
Health posts (btw 1991-1996)					-0.381^{***}		-0.003	
					(0.108)		(0.113)	
log(population)	0.785^{***}	0.624^{***}	0.167	0.462^{***}	0.238^{**}	0.072	0.123	0.102
	(0.119)	(0.115)	(0.135)	(0.119)	(0.117)	(0.115)	(0.123)	(0.109)
Observations	71	71	71	71	71	71	71	71
R^2	0.409	0.330	0.315	0.181	0.159	0.015	0.017	0.066

Notes: Columns (1) , (2), (3) and (4) use the number of health posts constructed between 1991–1996 as the dependent variable, whereas Columns (5) to (8) use the quality measures of the health posts constructed after the reform from 1991–1996.

The quality measures are based on two factors calculated using exploratory factor analysis. **Core-Infrastructure-Quality** refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 3: **Probability of Living Less Than an Hour from a Health Post**

	Living within an hour from a health post				
	(1)	(2)	(3)	(4)	(5)
Health post built after NHP	0.1288***		0.1368*	0.1635***	0.1618***
	(0.0485)		(0.0740)	(0.0566)	(0.0552)
Health post built before NHP		-0.0615	0.0592	0.0223	0.0309
		(0.0530)	(0.0718)	(0.0606)	(0.0613)
Observations	2,089	2,089	2,089	2,089	2,089
R^2	0.01508	0.00422	0.34166	0.15395	0.16798
District FE			✓		
Household Controls				✓	✓
Geographic Controls					✓

Notes: The dependent variable is an indicator for whether the reported distance to the closest health facility in the survey year (1996) is within an hour from the respondent's household. The household controls include ethnicity dummies and religious affiliation (Hindu, Buddhist) of the household head, whereas geographic controls include the distance to the nearest river from the VDC centroid and VDC elevation (along with their squared terms). Standard errors clustered at the district level are presented in parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4: **Primary Health Care Effects of Health Posts on Child Mortality (Boys)**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	-0.0100		-0.0055		0.0023		0.0059	
	(0.0084)		(0.0100)		(0.0109)		(0.0117)	
Constituencies (per 1000 ppl) ×Post		-0.1389		-1.233		0.0081		-1.135
		(0.2462)		(0.7681)		(0.3137)		(0.9080)
Core-Infrastructure-Quality×Post	-0.0422***	-0.0396***	-0.0434***	-0.0403***	-0.0432***	-0.0437***	-0.0450***	-0.0452***
	(0.0136)	(0.0130)	(0.0129)	(0.0124)	(0.0155)	(0.0148)	(0.0147)	(0.0141)
Reproductive-Care-Quality×Post	-0.0250**	-0.0226**	-0.0223*	-0.0192	-0.0184	-0.0190	-0.0166	-0.0161
	(0.0111)	(0.0107)	(0.0116)	(0.0119)	(0.0135)	(0.0129)	(0.0131)	(0.0133)
Observations	3,806	3,806	3,806	3,806	3,806	3,806	3,806	3,806
R^2	0.06559	0.06547	0.10224	0.10244	0.07286	0.07285	0.10645	0.10659
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls			✓	✓			✓	✓
Geographic Controls			✓	✓			✓	✓

Notes: The sample is restricted to boys born between 1980 and 1996. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the NHP in 1991 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people (which denotes the scheduled allocation), interacted with the Post dummy.
Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.
All specifications include the child's district of birth and year of birth fixed effects. The specifications with individual controls include ethnicity dummies, religious affiliation (Hindu, Buddhist), mother's age, and the child's birth order dummies. Geographic controls include elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. Standard errors clustered at the district level are presented in the parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 5: **Primary Health Care Effects of Health Posts on Child Mortality (Girls)**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	0.0235		0.0185		0.0205		0.0171	
	(0.0164)		(0.0166)		(0.0141)		(0.0144)	
Constituencies (per 1000 ppl) ×Post		-0.7039		1.481		-0.9145		1.158
		(0.9574)		(1.250)		(1.103)		(1.430)
Core-Infrastructure-Quality×Post	0.0077	0.0022	0.0011	-0.0053	0.0057	0.0010	-0.0011	-0.0067
	(0.0106)	(0.0101)	(0.0115)	(0.0113)	(0.0114)	(0.0109)	(0.0124)	(0.0121)
Reproductive-Care-Quality×Post	0.0224	0.0165	0.0121	0.0056	0.0342**	0.0290*	0.0209	0.0151
	(0.0164)	(0.0161)	(0.0134)	(0.0141)	(0.0165)	(0.0160)	(0.0139)	(0.0147)
Observations	3,555	3,555	3,555	3,555	3,555	3,555	3,555	3,555
R^2	0.05770	0.05713	0.09310	0.09299	0.06048	0.06018	0.09345	0.09332
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls			✓	✓			✓	✓
Geographic Controls			✓	✓			✓	✓

Notes: The sample is restricted to girls born between 1980 and 1996. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the NHP in 1991 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people (which denotes the scheduled allocation), interacted with the Post dummy.
Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.
All specifications include the child's district of birth and year of birth fixed effects. The specifications with individual controls include ethnicity dummies, religious affiliation (Hindu, Buddhist), mother's age, and the child's birth order dummies. Geographic controls include elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. Standard errors clustered at the district level are presented in the parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 6: **Primary Health Care Effects of Health Posts on Child Mortality (Boys) – Falsification Exercise**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	0.0324*		0.0275	0.0275	0.0256		0.0213	0.0213
	(0.0179)		(0.0199)	(0.0199)	(0.0172)		(0.0196)	(0.0196)
Constituencies (per 1000 ppl) ×Post		-0.5681				-0.5299		
		(0.4878)				(0.4616)		
Core-Infrastructure-Quality×Post	0.0123	0.0039	0.0122	0.0122	0.0071	0.0005	0.0070	0.0070
	(0.0186)	(0.0167)	(0.0210)	(0.0210)	(0.0176)	(0.0158)	(0.0207)	(0.0207)
Reproductive-Care-Quality×Post	-0.0057	-0.0118	0.0067	0.0067	-0.0186	-0.0235	-0.0132	-0.0132
	(0.0224)	(0.0237)	(0.0294)	(0.0294)	(0.0199)	(0.0212)	(0.0249)	(0.0249)
Observations	2,248	2,248	2,248	2,248	2,248	2,248	2,248	2,248
R^2	0.07326	0.07225	0.12988	0.12988	0.08479	0.08427	0.13599	0.13599
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls			✓	✓			✓	✓
Geographic Controls			✓	✓			✓	✓

Notes: The sample is restricted to boys born between 1980 and 1990 (prior to NHP) and 1986 is used as the artificial policy year. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the artificial policy in 1986 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people (which denotes the scheduled allocation), interacted with the Post dummy.
Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.
All specifications include the child's district of birth and year of birth fixed effects. The specifications with individual controls include ethnicity dummies, religious affiliation (Hindu, Buddhist), mother's age, and the child's birth order dummies. Geographic controls include elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. Standard errors clustered at the district level are presented in the parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 7: **Primary Health Care Effects of Health Posts on Child Mortality (Girls) – Falsification Exercise**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	0.0089		0.0106	0.0106	0.0177		0.0278	0.0278
	(0.0249)		(0.0246)	(0.0246)	(0.0240)		(0.0250)	(0.0250)
Constituencies (per 1000 ppl) ×Post		1.199				1.263		
		(1.034)				(1.179)		
Core-Infrastructure-Quality×Post	0.0083	0.0060	0.0163	0.0163	0.0042	-0.0003	0.0147	0.0147
	(0.0199)	(0.0187)	(0.0192)	(0.0192)	(0.0218)	(0.0205)	(0.0209)	(0.0209)
Reproductive-Care-Quality×Post	0.0137	0.0116	0.0044	0.0044	-0.0030	-0.0072	-0.0089	-0.0089
	(0.0174)	(0.0180)	(0.0192)	(0.0192)	(0.0217)	(0.0220)	(0.0231)	(0.0231)
Observations	2,123	2,123	2,123	2,123	2,123	2,123	2,123	2,123
R^2	0.07350	0.07368	0.12974	0.12974	0.07164	0.07160	0.13265	0.13265
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls			✓	✓			✓	✓
Geographic Controls			✓	✓			✓	✓

Notes: The sample is restricted to girls born between 1980 and 1990 (prior to NHP) and 1986 is used as the artificial policy year. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the artificial policy in 1986 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people (which denotes the scheduled allocation), interacted with the Post dummy.
Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.
All specifications include the child's district of birth and year of birth fixed effects. The specifications with individual controls include ethnicity dummies, religious affiliation (Hindu, Buddhist), mother's age, and the child's birth order dummies. Geographic controls include elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. Standard errors clustered at the district level are presented in the parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Online Appendix:
Healthcare Reform and Gender Specific Child Investment in Rural Nepal

Contents

Online Appendix 42

A Exploratory Factor Analysis (EFA)

We chose EFA as the factor extraction method since we are interested in interpreting the factor variables constructed from the quality measures of health posts. EFA is based on the notion of latent construct, which is used to form the quality measures. We use EFA to construct two factors that can be interpreted directly: *i*) **Core-Infrastructure-Quality** (Factor 1) which depends heavily on the variables measuring the availability of beds, refrigerators, electricity and doctors; and *ii*) **Reproductive-Care-Quality** (Factor 2) which heavily loads on variables measuring the availability of antenatal care, delivery services, and malaria treatment.

The factor model can be written as:

$$X = \mu + \lambda \times f + e \tag{A.1}$$

where, X is a p element vector of traits that are observed, λ is the loading matrix of dimension $p \times m$, f is the score of m elements (the common latent factors), and e is the error term. Vector X comprises observed variables that are loading onto the latent factors (of dimension m). Each element of X is sampled from a population with a mean of

$$\mu = \begin{vmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{vmatrix} = \text{population mean of observed variables}$$

The score f in equation A.1 are the common latent factors (of dimension m, where $m < p$) so that:

$$f = \begin{vmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{vmatrix} = \text{unobserved latent factors.}$$

From a regression standpoint, the factor model can be thought of as a set of regressions, where $f_i,\ i = \{1,2,..m\}$ are the common factors and e_i are the specific factors that describe the outcome variables collected in vector X:

$$\begin{aligned}
X_1 &= \mu_1 + \lambda_{11} f_1 + \lambda_{12} f_2 + \lambda_{13} f_3 + \ldots + \lambda_{1m} f_m + e_1, \\
X_2 &= \mu_2 + \lambda_{21} f_1 + \lambda_{22} f_2 + \lambda_{23} f_3 + \ldots + \lambda_{2m} f_m + e_2, \\
X_3 &= \mu_3 + \lambda_{31} f_1 + \lambda_{32} f_2 + \lambda_{33} f_3 + \ldots + \lambda_{3m} f_m + e_3, \\
&\vdots \\
X_p &= \mu_p + \lambda_{p1} f_1 + \lambda_{p2} f_2 + \lambda_{p3} f_3 + \ldots + \lambda_{pn} f_m + e_p.
\end{aligned}$$

The elements of the factor loading matrix can be collected in a matrix of $p \times m$ as:

$$\lambda = \begin{vmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \lambda_{31} & \lambda_{32} & \cdots & \lambda_{3m} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{vmatrix}.$$

The following assumptions are needed to identify the factors:

1. $E[f_i] = 0$ and $E[e_i] = 0$ so that $E[X_i] = \mu_i$; $i = \{1, 2, ..., p\}$
2. $var(f_i) = 1$; $var(e_i) = \Psi_i$; $i = \{1, 2, ..., p\}$
3. The common factors are uncorrelated with one another: $cov(f_i, f_j) = 0$; $i! = j$.
4. The specific factors are uncorrelated with each other: $cov(e_i, e_j) = 0$; $i! = j$.
5. The specific factors are uncorrelated with common factors: $cov(f_i, e_j) = 0$; $i = \{1, 2, ..., p\}$.

Under the above assumptions, the correlation matrix for X can be written as:

$$\sum = \lambda\lambda' + \Psi, \tag{A.2}$$

where $\sum$ is the variance-covariance matrix, λ is the factor loading matrix, and Ψ is the a diagonal matrix consisting of specific variances. The model assumes that common factors are uncorrelated with one another, the specific factors are uncorrelated with one another, and the specific factors are uncorrelated with the common factors.

The maximum likelihood estimation (MLE) assumes that the data are sampled from a multivariate normal distribution with mean vector μ and variance covariance matrix of $\sum$ in equation A.2. The MLE estimates for μ, the factor loading λ, and the specific variance Ψ are determined by maximizing the following log likelihood function:

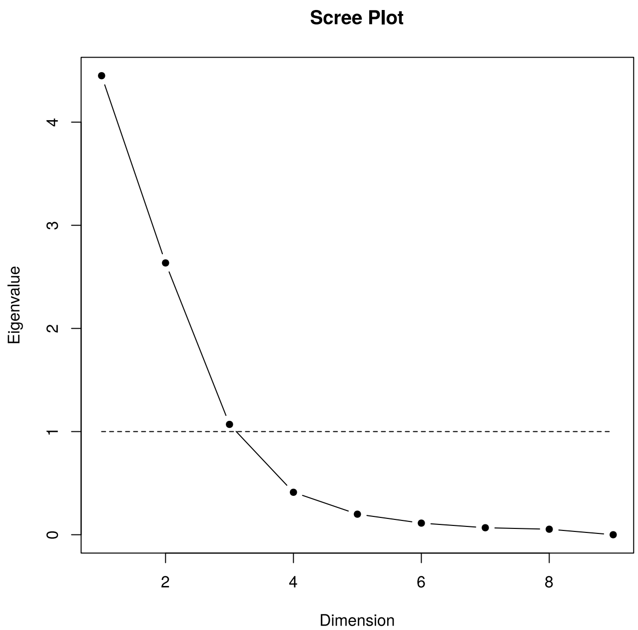
$$L(\mu, \lambda, \Psi) = -\frac{n \times p}{2} \log(2\pi) - \frac{n}{2} \log\left(|\lambda\lambda' + \Psi|\right) - \frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)(\lambda\lambda' + \Psi)(X_i - \mu) \tag{A.3}$$

The estimates $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\Psi}$ maximizes the equation A.3. In equation A.2, the model remains unchanged if replaced by $M\lambda$, where M is some orthogonal matrix, which is also known as rotation matrix. The oblique rotation method is used for the purpose of improving the interpretation of factors. Although both varimax and oblique rotation increases the interpretation of results by clearly differentiating the intensity of loading of each variable onto a factor (i.e., a particular variable will tend to have large or small loading on each factor), unlike varimax rotation this method allows factors to be correlated with one another. The following decisions summarize the exploratory factor analysis used in this study:

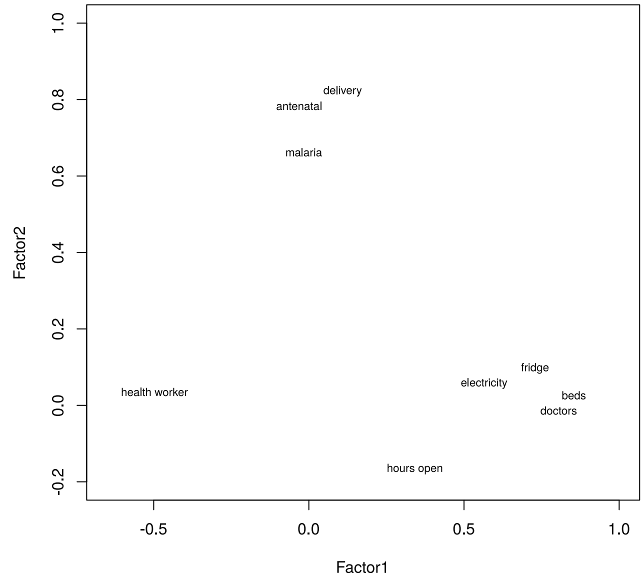
- Use the maximum likelihood method for estimation.
- Two factors are retained based on eigenvalues and the scree plot test.
- Oblique rotation method is favored to varimax.
- Using the factor loading, the factor scores are obtained by estimating least square regression. $\hat{f} = (\hat{\lambda}'\hat{\lambda})^{-1}\hat{\lambda}'(X_i - \hat{\mu})$, where $\hat{f}$ is $m \times 1$ vector.

Results. Figure A.1 shows the scree plot that plots the number of factors along with the eigenvalues, factor loading, and predicted scores of two factors that were selected. It is customary to retain factors with eigenvalues greater than one. However, Costello and Osborne (2005) find that this method retains too many factors and suggest to only retain factors above the break point in the scree plot. Figure A.1(a) shows that the break-point falls at the 3^{rd} factor, at which the eigenvalue is only slightly above one. Hence, by using the conventional rule-of-thumb of retaining factors with eigenvalues above one (Kaiser, 1960—Kaiser's eigenvalue rule) as well as observing the scree plot (Cattell, 1966—Cattell's scree plot), we keep two factors that are well above one.

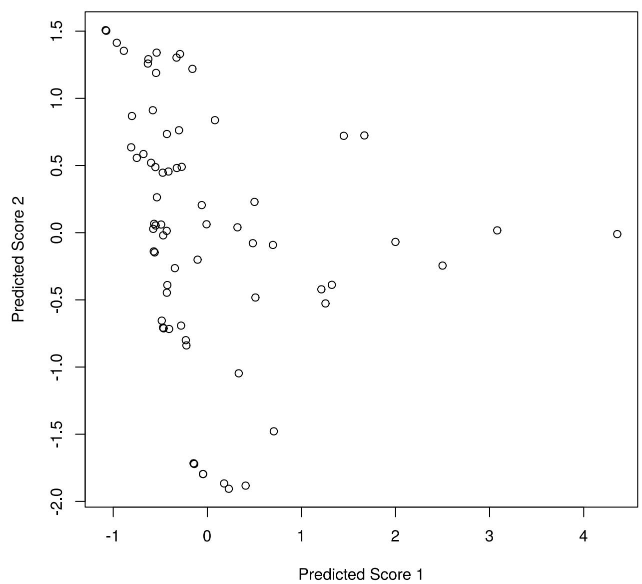
By using the *varimax* rotation—an orthogonal method of rotation that produces uncorrelated factors—the interpretation of the quality measures simplifies. Figure A.1(b) shows the factor loading of the quality measures following the *varimax* rotation. The availability of doctors, hospital beds for inpatient care, provision of electricity, hours open, and refrigerators to store medicine heavily load onto Factor 1 whereas antenatal care, malaria treatment and delivery care load onto Factor 2. Therefore Factor 1 (Core-Infrastructure-Quality) defines the core infrastructure status of the health post and Factor 2 (Reproductive-Care-Quality) refers to the quality of prenatal care and the provision of delivery options.



(a) Scree Plot Showing Eigenvalues



(b) Loadings



(c) Predicted Scores

Figure A.1: **Results from Factor Analysis using the Maximum Likelihood Method**

Notes: **Core-Infrastructure-Quality** (Factor 1) refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** (Factor 2) refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.

B Additional Robustness Exercises

B.1 Other Development Projects

One concern regarding the validity of our identification strategy is possible interference from other public development programs at the village level which could have a direct effect on child health. For example, the Water Resources Act—implemented in 1992—may have reduced infant mortality through the provision of cleaner water. This could bias our main estimates upward if the effects of this policy were more concentrated in districts that received higher quality health posts. However, this still does not explain the discrepancy in findings between boys and girls, as it is unlikely that a household discriminates across gender in terms of drinking water.

In the early 1990s, following the second democratic revolution, many development programs were initiated in Nepal. Figure B.1 depicts the total number of the major development projects by year. As shown, the projects increased in number after 1990. In order to account for projects that might influence child mortality outcomes in the communities sampled by the NLSS 1996, we add comprehensive control variables regarding concurrent development projects into the econometric specifications from Section 5.1. Community level files of the NLSS contain data that allow us to construct indicator variables for the following government or NGO development programs or groups: (*i*) water related development, (*ii*) women's professional development group initiatives, (*iii*) forest preservation projects, and (*iv*) the establishment of credit association groups. More specifically, we merge the community level data referring to development groups with the primary data set by ward (the primary sampling unit) and construct an indicator variable for a development group being in operation during the child's year of birth by using the group's year of establishment from the survey. We add these development group specific indicator variables in the model specification one by one and then re-estimate the econometric specification in equation 1. The new results are presented in Tables B.1 and B.2 for boys and girls, respectively. The interaction coefficients between the quality factors and the year of birth post reform are similar in magnitude as the coefficients in Tables 4 and 5.

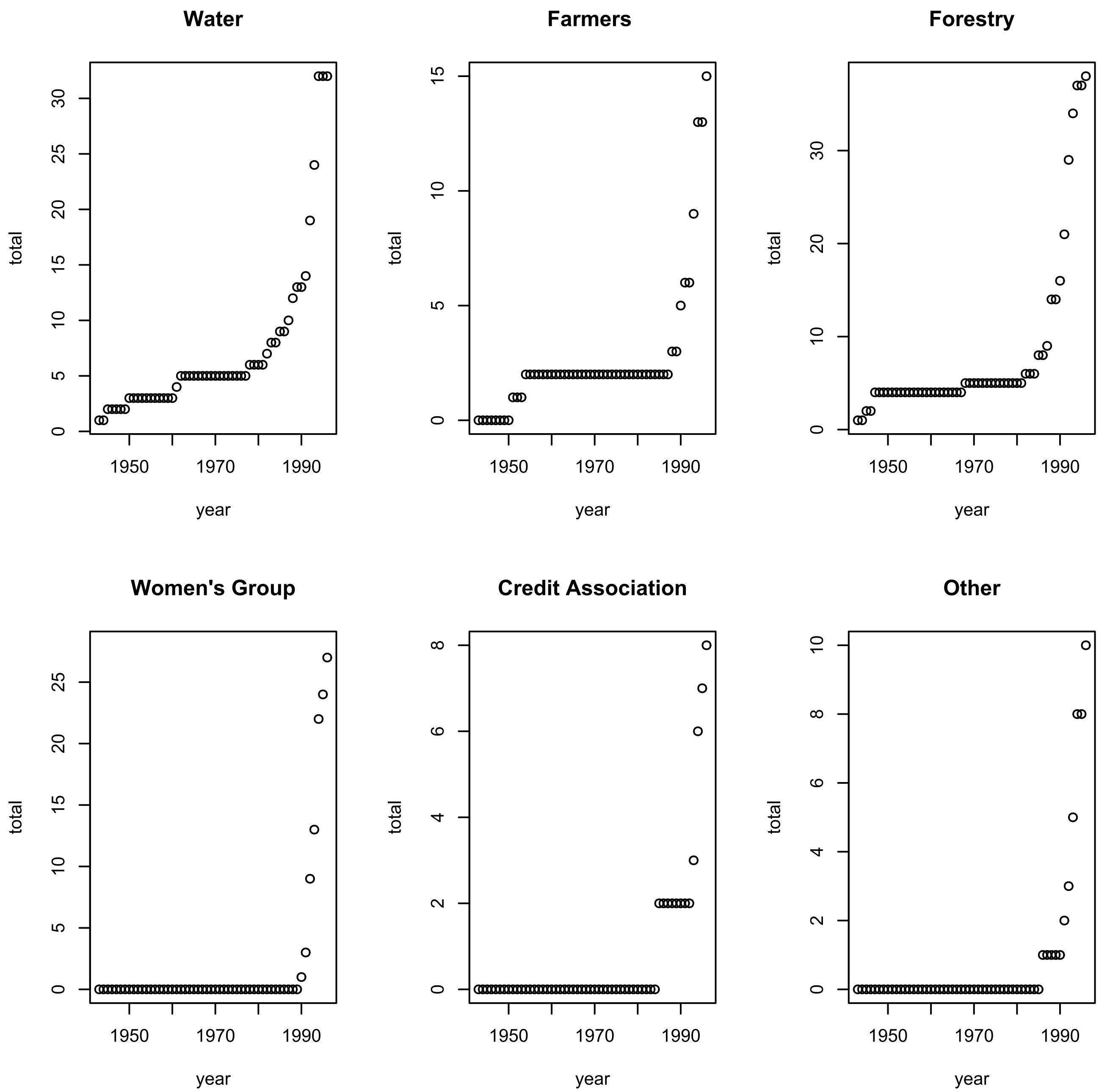


Figure B.1: **The Number of Development Groups at the Ward Level by Type and Year**

Notes: The data source is the NLSS 1996, community level files.

Table B.1: **Primary Health Care Effects of Health Posts on Child Mortality (Boys) – Controlling for Other Development Programs**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	-0.0055	-0.0051	-0.0053	-0.0062	0.0058	0.0062	0.0061	0.0052
	(0.0100)	(0.0100)	(0.0101)	(0.0101)	(0.0117)	(0.0117)	(0.0118)	(0.0118)
Core-Infrastructure-Quality×Post	-0.0434***	-0.0431***	-0.0427***	-0.0437***	-0.0450***	-0.0448***	-0.0442***	-0.0453***
	(0.0129)	(0.0129)	(0.0132)	(0.0129)	(0.0147)	(0.0147)	(0.0150)	(0.0147)
Reproductive-Care-Quality×Post	-0.0225*	-0.0228**	-0.0219*	-0.0226*	-0.0165	-0.0170	-0.0161	-0.0169
	(0.0117)	(0.0114)	(0.0118)	(0.0116)	(0.0133)	(0.0130)	(0.0133)	(0.0131)
Observations	3,806	3,806	3,806	3,806	3,806	3,806	3,806	3,806
R^2	0.10225	0.10228	0.10245	0.10258	0.10645	0.10647	0.10669	0.10678
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Geographic Controls	✓	✓	✓	✓	✓	✓	✓	✓
Water Group	✓				✓			
Women's Group		✓				✓		
Forest Group			✓				✓	
Credit Group				✓				✓

Notes: The sample is restricted to boys born between 1980 and 1996. This specification adds indicators for other public programs addressing the development of either (*i*) water, (*ii*) agriculture, (*iii*) forest sector, and (*iv*) lending and credit in the area of a mother's residence during the year of the child's birth. Columns 1-4 and 5-8 pertain to infant mortality and child mortality, respectively.

All specifications control for the interaction between the number of new health posts established between 1990 and 1996 per 1000 people, controls for quality measures of health posts interacted with the post reform indicator, ethnicity dummies (Brahmin, Chhetri), religion dummies (Hindu, Buddhist), mother's age dummies, child's year of birth fixed effects and district of birth fixed effects, and child birth order dummy variables.

Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.

Standard errors clustered at the district level are presented in parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table B.2: **Primary Health Care Effects of Health Posts on Child Mortality (Girls) – Controlling for Other Development Programs**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	0.0180	0.0178	0.0185	0.0181	0.0163	0.0164	0.0170	0.0169
	(0.0166)	(0.0162)	(0.0166)	(0.0167)	(0.0145)	(0.0141)	(0.0144)	(0.0145)
Core-Infrastructure-Quality×Post	0.0007	0.0008	0.0011	0.0012	-0.0016	-0.0014	-0.0014	-0.0010
	(0.0114)	(0.0115)	(0.0119)	(0.0115)	(0.0124)	(0.0124)	(0.0127)	(0.0124)
Reproductive-Care-Quality×Post	0.0133	0.0125	0.0121	0.0120	0.0228	0.0213	0.0206	0.0208
	(0.0136)	(0.0134)	(0.0135)	(0.0133)	(0.0140)	(0.0139)	(0.0140)	(0.0138)
Observations	3,555	3,555	3,555	3,555	3,555	3,555	3,555	3,555
R^2	0.09339	0.09326	0.09310	0.09327	0.09412	0.09361	0.09347	0.09353
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Geographic Controls	✓	✓	✓	✓	✓	✓	✓	✓
Water Group	✓				✓			
Women's Group		✓				✓		
Forest Group			✓				✓	
Credit Group				✓				✓

Notes: The sample is restricted to girls born between 1980 and 1996. The table is structured similarly to Table B.1. The standard errors are clustered at the district level and presented in parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

B.2 Schooling Boom in Early 1990s

The number of schools increased quite dramatically in the early 1990s. Following the democratization movement of 1990, education was liberalized as the country moved towards a community based approach of schooling, which created unchecked growth in private schooling (Carney and Bista, 2009; Shrestha, 2019; Shrestha and Shrestha, 2017). The boom in schooling infrastructure was mainly a demand-driven phenomenon—both private and public schools were established in regions with more educated populace. This is shown in Table B.3 where the number of private schools per capita in 1993 and the number of primary schools per capita are regressed on education variables such as male and female literacy rates in 1981 and enrollment rates in 1977. In Columns (1)–(3), prior literacy and enrollment rates are positively correlated with the number of private schools per capita in 1993. This suggests that more private schools were built in districts that already had relatively better educational outcomes. It is important to note that variables representing the number of health posts and quality factors are close to zero and statistically insignificant at the conventional levels. This indicates that school construction in the 1990s was not systematically correlated with the intensity of the health reform.

Table B.3: **School Construction Pattern in the Early 1990s**

	Dependent variable:								
	Private Schools (1993) per capita			Secondary Schools (1993) per capita			Primary Schools (1993) per capita		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male Literacy (1981)	0.002***			0.00001**			0.00001		
	(0.0004)			(0.00001)			(0.00002)		
Female Literacy (1981)		0.003***			0.00001			−0.00002	
		(0.001)			(0.00001)			(0.00002)	
Secondary Enrollment (1977)			0.577***			0.001			−0.005
			(0.091)			(0.001)			(0.004)
Health posts (per capita)	0.060	0.059	0.101	−0.002	−0.002	−0.001	−0.002	−0.002	−0.002
	(0.103)	(0.103)	(0.098)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)
Core-Infrastructure-Quality	−0.003	−0.006	0.003	−0.0002	−0.0002	−0.0001	−0.001	−0.001	−0.001
	(0.010)	(0.010)	(0.009)	(0.0001)	(0.0002)	(0.0002)	(0.0004)	(0.0004)	(0.0004)
Reproductive-Care-Quality	0.002	−0.002	0.003	−0.0001	−0.0001	−0.0001	−0.0005	−0.0003	−0.0003
	(0.007)	(0.007)	(0.007)	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0003)	(0.0003)
Observations	72	72	72	73	73	73	73	73	73
R^2	0.340	0.345	0.397	0.078	0.046	0.029	0.046	0.048	0.065
Note:	*p<0.1; **p<0.05; ***p<0.01								

Notes: The dependent variables used are the district-specific number of respective schools per person in 1993. Population estimates from the Census 1991 are used for this purpose. Data for schools are extracted from the Education Statistics of Nepal (1993), literacy rates are from the 1981 Census, and the enrollment rate is obtained from the study conducted by the UNESCO titled Primary Education – A Sub-Sector Study.

Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

B.3 Alternate Quality Measures for Health Posts

B.3.1 Health Post Quality Aggregation at PSU Level

So far, while evaluating the effect of the reform at the primary (health post) level, we have used the factor scores that are obtained from MLE and estimated the effects of health post quality on infant mortality at the district level. However, quality data are available at the community level (the primary sampling unit is a ward) and the question arises whether aggregation of quality measures at the district level are driving our main results. To assess this issue, we replicate the effects of the NHP at the primary level by constructing factor scores at the PSU (primary sampling unit) level. The results are presented in Tables B.4 and B.5, respectively. The findings suggest that quality measures represented by Core-Infrastructure-Quality reduce boys' infant and child mortality outcomes but have no effect on the mortality of girls. These findings are consistent with the main results presented in Tables 4 and 5.

Table B.4: **National Health Policy and Child Mortality (Boys) – EFA at the VDC (Village) Level**

	Dependent variable:							
	Infant Mortality (Boys) Ages 0-1				Child Mortality (Boys) Ages 0-3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health posts (per 1000 ppl) *Post	0.0059		0.0082		0.0210		0.0246	
	(0.0104)		(0.0118)		(0.0138)		(0.0151)	
Constituencies (per 1000 ppl)*Post		−0.0294		−0.1932		−0.0595		−0.1661
		(2.2890)		(2.7769)		(2.6997)		(3.4728)
Quality Measure 1*Post	−0.0167*	−0.0173*	−0.0163*	−0.0170*	−0.0214**	−0.0233***	−0.0207**	−0.0230***
	(0.0088)	(0.0089)	(0.0087)	(0.0088)	(0.0089)	(0.0089)	(0.0091)	(0.0089)
Quality Measure 2*Post	−0.0021	−0.0024	−0.0031	−0.0031	0.0016	0.0004	−0.0002	−0.0008
	(0.0106)	(0.0107)	(0.0111)	(0.0111)	(0.0123)	(0.0126)	(0.0130)	(0.0132)
Other Controls	N	N	Y	Y	N	N	Y	Y
Birth year FE	Y	Y	Y	Y	Y	Y	Y	Y
VDC FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	3,404	3,404	3,404	3,404	3,404	3,404	3,404	3,404
R^2	0.1425	0.1425	0.1427	0.1426	0.1424	0.1419	0.1426	0.1420
Note:	*p<0.1; **p<0.05; ***p<0.01							

Notes: The sample is restricted to boys born between 1980 and 1996. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the NHP in 1991 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people interacted with the Post dummy. In addition, we control for ethnicity (Brahmin, Chhetri), religious affiliation (Hindu, Buddhist), mother's age, child's year of birth fixed effects, and district of birth fixed effects. All specifications include Core-Infrastructure-Quality and 2 obtained from the Exploratory Factor Analysis and their interactions with the Post dummy.

Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.

Specifications (3), (4), (7) and (8) add the interaction between the post reform indicator and (*i*) new schools constructed between 1990–1996 and (*ii*) the literacy rate in 1981. Bootstrapped standard errors from 199 replications, clustered at the district level, are presented in parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table B.5: **National Health Policy and Child Mortality (Girls) – EFA at the VDC (village) Level**

	Dependent variable:							
	Infant Mortality (Girls) Ages 0-1				Child Mortality (Girls) Ages 0-3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health posts (per 1000 ppl)*Post	0.0205		0.0230		0.0171		0.0189	
	(0.0188)		(0.0190)		(0.0188)		(0.0190)	
Constituencies (per 1000 ppl)*Post		−0.7316		−0.1957		−0.6904		−0.0757
		(3.6608)		(3.8657)		(3.6608)		(3.8657)
Quality Measure 1*Post	0.0037	0.0021	0.0037	0.0019	0.0083	0.0070	0.0082	0.0067
	(0.0094)	(0.0093)	(0.0097)	(0.0096)	(0.0094)	(0.0093)	(0.0097)	(0.0096)
Quality Measure 2*Post	0.0109	0.0085	0.0086	0.0062	0.0130	0.0111	0.0107	0.0087
	(0.0146)	(0.0146)	(0.0148)	(0.0150)	(0.0146)	(0.0146)	(0.0148)	(0.0150)
Other Controls	N	N	Y	Y	N	N	Y	Y
Birth year FE	Y	Y	Y	Y	Y	Y	Y	Y
VDC FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	3,188	3,188	3,188	3,188	3,188	3,188	3,188	3,188
R^2	0.1200	0.1196	0.1207	0.1201	0.1249	0.1247	0.1255	0.1252

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: The sample is restricted to girls born between 1980 and 1996. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the NHP in 1991 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people interacted with the Post dummy. In addition, we control for ethnicity (Brahmin, Chhetri), religious affiliation (Hindu, Buddhist), mother's age, child's year of birth fixed effects, and district of birth fixed effects. All specifications include Core-Infrastructure-Quality and Reproductive-Care-Quality and their interactions with the Post dummy.

Core-Infrastructure-Quality refers to the core infrastructure such as the availability of doctors, hospital beds, refrigerators, electricity, and weekly opening hours of health posts and **Reproductive-Care-Quality** refers to variables measuring the availability of antenatal care, delivery care and malarial treatment at the health posts. Both quality measures are obtained via exploratory factor analysis.

Specifications (3), (4), (7) and (8) add the interaction between the post reform indicator and (i) new schools constructed between 1990–1996 and (ii) the literacy rate in 1981. Bootstrapped standard errors from 199 replications, clustered at the district level, are presented in parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

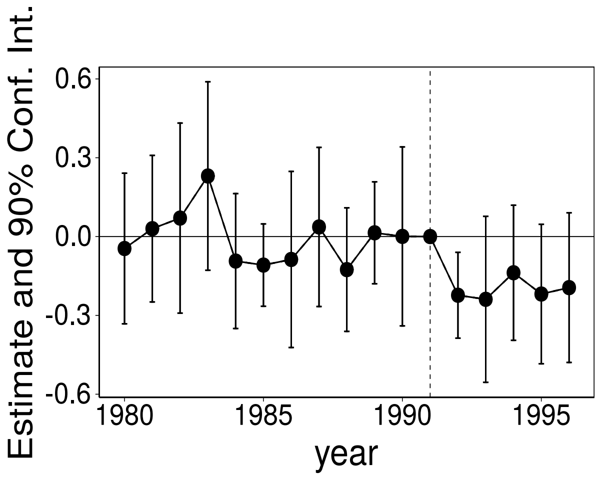
B.3.2 Raw Quality Measures

Next, instead of using the factor scores as independent variables, we directly use the quality measures such as the district-specific proportion of health posts with doctors and hospital beds (that load heavily on Core-Infrastructure-Quality) to corroborate the main findings. The results are presented in Tables B.6 and Figure B.2 and suggest that the availability of doctors and hospital beds improve child mortality of boys but have no effect on girls—the interaction coefficients pertaining to quality measures including the availability of doctors and beds are statistically significant at the conventional levels.

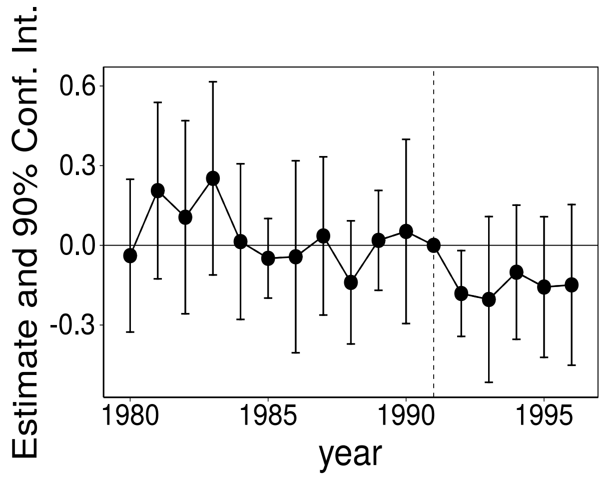
Table B.6: **Primary Health Care Effects of Health Posts on Child Mortality (Boys) Based on Raw Quality Variables of Health Posts**

	0 to 1 (boys)		0 to 3 (boys)		0 to 1 (girls)		0 to 3 (girls)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000)×Post	4.23×10^{-5}	-0.0005	0.0122	0.0115	0.0177	0.0206	0.0160	0.0205
	(0.0088)	(0.0098)	(0.0109)	(0.0117)	(0.0162)	(0.0164)	(0.0140)	(0.0144)
Prop. of HP with beds×Post	-0.2039***		-0.1984***		-0.0093		-0.0300	
	(0.0546)		(0.0614)		(0.0465)		(0.0478)	
Prop. HP with doctors×Post		-0.1079**		-0.1069*		0.0342		0.0423
		(0.0486)		(0.0541)		(0.0480)		(0.0535)
Observations	3,806	3,806	3,806	3,806	3,555	3,555	3,555	3,555
R^2	0.10268	0.10097	0.10655	0.10517	0.09311	0.09319	0.09350	0.09357
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓	✓	✓	✓
Geographic Controls	✓	✓	✓	✓	✓	✓	✓	✓

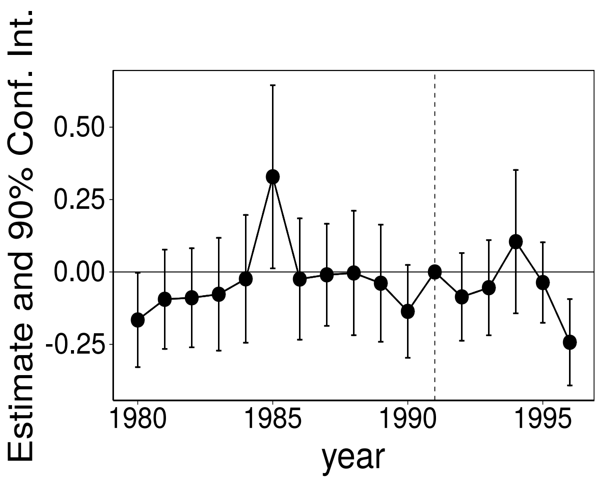
Notes: The sample is restricted to boys born between 1980 and 1996. These set of results differ from the findings shown in Table 4 such that component measures of Core-Infrastructure-Quality are used rather than the factor score variable. Columns 1–2 (5-6) and 3–4 (7-8) pertain to infant mortality and child mortality, respectively. All specifications control for the interaction between the number of new health posts established between 1990 and 1996 per 1000 people and the post reform dummy, ethnicity dummies (Brahmin, Chhetri), religion dummies (Hindu, Buddhist), mother's age dummies, child's year of birth fixed effects and district of birth fixed effects. Additionally, all specifications control for geographic variables including elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.



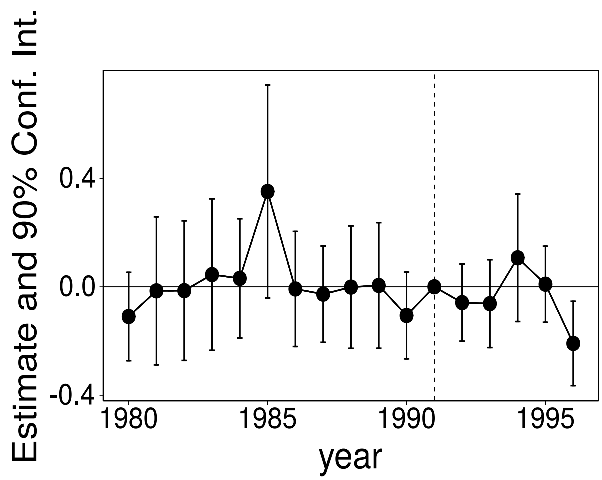
(a) Child Mortality (Ages 0 to 1, Boys)



(b) Child Mortality (Ages 0 to 3, Boys)



(c) Child Mortality (Ages 0 to 1, Girls)



(d) Child Mortality (Ages 0 to 3, Girls)

Figure B.2: **Event Study – Proportion of Health Posts with Beds**

Notes: The figure replicates Figure 7 but plots the interaction between child's year of birth and the proportion of health posts with beds – the health infrastructure variable that primarily loads onto the Core-Infrastructure-Quality variable. The error bar represents the 90% confidence interval calculated using such bootstrapped standard errors, respectively. Data Source: NLSS 1996.

B.3.3 Quality Measures based on Principal Components Analysis

In this section we use an alternative dimension reduction method, principal components analysis (PCA), as a robustness check to ensure that our preferred method of EFA does not drive our main result. The results are presented in Tables B.7 and Figure B.8 and suggest that the alternatively constructed quality measures improve child mortality of boys but have no effect on girls—the interaction coefficients pertaining to the quality measures based on PCA are statistically significant at the conventional levels.

Table B.7: **Primary Health Care Effects of Health Posts on Child Mortality (Boys)**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	-0.0127		-0.0076		-0.0011		0.0029	
	(0.0089)		(0.0105)		(0.0112)		(0.0120)	
Constituencies (per 1000 ppl) ×Post		-0.1998		-1.323*		-0.0515		-1.221
		(0.2304)		(0.7479)		(0.2962)		(0.8800)
Quality Measure 1×Post	-0.0376***	-0.0343***	-0.0383***	-0.0351***	-0.0413***	-0.0410***	-0.0425***	-0.0423***
	(0.0122)	(0.0115)	(0.0119)	(0.0111)	(0.0144)	(0.0137)	(0.0139)	(0.0132)
Quality Measure 2×Post	-0.0199**	-0.0174*	-0.0169	-0.0139	-0.0132	-0.0130	-0.0110	-0.0100
	(0.0098)	(0.0095)	(0.0104)	(0.0106)	(0.0118)	(0.0113)	(0.0117)	(0.0119)
Observations	3,806	3,806	3,806	3,806	3,806	3,806	3,806	3,806
R^2	0.06559	0.06541	0.10214	0.10235	0.07305	0.07305	0.10656	0.10675
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls			✓	✓			✓	✓
Geographic Controls			✓	✓			✓	✓

Notes: The sample is restricted to boys born between 1980 and 1996. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the NHP in 1991 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people (which denotes the scheduled allocation), interacted with the Post dummy.
Quality Measure 1 and **Quality Measure** 2 refer to the first two principal components from the implementation of PCA.
All specifications include the child's district of birth and year of birth fixed effects. The specifications with individual controls include ethnicity dummies, religious affiliation (Hindu, Buddhist), mother's age, and the child's birth order dummies. Geographic controls include elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. Standard errors clustered at the district level are presented in the parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table B.8: **Primary Health Care Effects of Health Posts on Child Mortality (Girls)**

	child mortality 0 to 1				child mortality 0 to 3			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Health Posts (per 1000 ppl)×Post	0.0240		0.0182		0.0204		0.0166	
	(0.0163)		(0.0169)		(0.0140)		(0.0146)	
Constituencies (per 1000 ppl) ×Post		-0.6690		1.473		-0.8605		1.149
		(0.9339)		(1.243)		(1.085)		(1.427)
Quality Measure 1×Post	0.0051	-0.0007	-0.0004	-0.0061	0.0018	-0.0032	-0.0028	-0.0078
	(0.0098)	(0.0095)	(0.0107)	(0.0101)	(0.0106)	(0.0102)	(0.0118)	(0.0112)
Quality Measure 2×Post	0.0206	0.0154	0.0110	0.0057	0.0308**	0.0262*	0.0187	0.0141
	(0.0152)	(0.0147)	(0.0119)	(0.0125)	(0.0151)	(0.0145)	(0.0124)	(0.0129)
Observations	3,555	3,555	3,555	3,555	3,555	3,555	3,555	3,555
R^2	0.05772	0.05714	0.09310	0.09300	0.06042	0.06012	0.09341	0.09329
Birth Year and District FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual Controls			✓	✓			✓	✓
Geographic Controls			✓	✓			✓	✓

Notes: The sample is restricted to girls born between 1980 and 1996. Columns 1–4 and 5–8 pertain to infant mortality and child mortality, respectively. The odd numbered specifications control for the interaction between an indicator variable Post which equals one if a child is born after the NHP in 1991 and the number of new health posts (per 1000 people) established between 1990–1996. Alternately, the even numbered specifications control for the number of constituencies per 1000 people (which denotes the scheduled allocation), interacted with the Post dummy.
Quality Measure 1 and **Quality Measure** 2 refer to the first two principal components from the implementation of PCA.
All specifications include the child's district of birth and year of birth fixed effects. The specifications with individual controls include ethnicity dummies, religious affiliation (Hindu, Buddhist), mother's age, and the child's birth order dummies. Geographic controls include elevation, the shortest distance to the zonal hospital, the shortest river distance (and their squared terms) interacted with pre-reform and post-reform dummies. Standard errors clustered at the district level are presented in the parenthesis. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

C Model of Child Investment Based on Gender Specific Labor Market Returns

This section develops a stylized two-period model to build intuition and testable hypotheses for gender specific investments into child health. Households are assumed to live for two periods. In period one households value consumption c and leisure t_l of parents and in period two they value consumption z of children using utility function

$$u\left(c,\overbrace{1-t_w-\phi t_h}^{t_l}\right)+\delta v(z), \tag{C.1}$$

where $\frac{\partial u}{\partial c}>0, \frac{\partial^2 u}{\partial c^2}<0, \frac{\partial u}{\partial t_l}>0, \frac{\partial^2 u}{\partial t_l^2}<0, v'>0, v''<0$, ϕ is a healthy time cost factor, and δ is a time preference factor. Parameter ϕ captures the cost involved in accessing the closest health post or hospital and comprises costs such as travel time, wait time, etc. that add to the actual health care time t_h. In period one the household is endowed with one unit of time and decides to spend it either working t_w, as leisure t_l, or as time investment into their child t_h. The time spent on a child is not used as consumption good but is purely an investment made by parents in period one (Del Boca, Flinn and Wiswall, 2013). In addition,parents make health-investment decisions for their child in period one. The production of child health involves two inputs—medical care m and time spent care giving or accessing health care for the child t_h. The production function of the health capital of the child is

$$h=H(m,t_h), \tag{C.2}$$

where $\frac{\partial H}{\partial m}>0, \frac{\partial^2 H}{\partial m^2}<0, \frac{\partial H}{\partial t_h}>0, \frac{\partial^2 H}{\partial t_h^2}<0$. For example, m comprises postnatal visits, immunizations, doctor visits etc. and $\phi \times t_h$ is the total time associated with such visits including care giving time. However, unscientific care, which is still practiced in parts of many developing countries, are not included in m. The price for medical services is p_m.

Income in the first period depends on the time parents decide to work t_w and the market wage rate w. At the end of period 1, parents die. In the second period, the child (now an adult) starts reaping benefits from the health-investment provided by her parents in period 1 and also receives the market return R from household savings s. The household maximization problem can be written as

$$\max_{(c,z,t_l,t_w,t_h,m)} u(c,t_l)+\delta v(z) \text{ s.t.}$$

$$\begin{aligned} c + p_m \times m + s &= w \times t_w, \\ z &= R^i \times h + R \times s, \\ t_l + t_w + \phi t_h &= 1, \text{ and } (C.2), \end{aligned} \tag{C.3}$$

where R^i is a gender specific return on the investment made in the child and subscript $i \in \{s,d\}$ indicates sons s and daughters d. Gender bias is not incorporated in the utility function but health returns differ across gender due to labor market and cultural biases. We assume that that all prices are exogenous and $R^s > R^d$.

The latter is shown to be true for all ages over 18 years of age pertaining to wages in both non-agricultural and agricultural sectors. Figure C.1 provides the life-cycle trend of participation in wage-employment across non-agriculture and agriculture sectors by gender. The figure demonstrates a stark difference in wage-related activities by gender, specifically for non- agriculture sector, where women's participation is below the 5 percent level.

Solving the model results in a system of five equations in five unknowns c, z, t_w, t_h, and m:

$$u_c = \frac{\delta v'}{\left(\frac{1}{R}\right)}, \tag{C.4}$$

$$u_c = \frac{u_{t_l}}{w}, \tag{C.5}$$

$$u_c = \frac{\left(u_c \frac{R^i}{R}\right) H_m}{p_m}, \tag{C.6}$$

$$u_c = \frac{\left(u_c \frac{R^i}{R}\right) H_{t_h}}{\phi w}, \tag{C.7}$$

$$c + p_m m + \frac{z}{R} = w t_w + \frac{R^i H(m, t_h)}{R}, \tag{C.8}$$

where the prime notation indicates derivatives of functions and subscripts indicate partial derivatives. The first four optimality conditions describe the respective trade-off between consumption today c with consumption tomorrow z, with leisure t_l, with health time investment t_h, and with healthcare m. The left hand side of the four expressions is the return per dollar spent on first period consumption, which in equilibrium, has to equal the return per dollar spent on child consumption z, leisure time t_l, health time t_h and health care m. Expression ($C.8$) is the lifetime budget constraint.

Assuming $H(m, t_h) = A m^{\alpha_1} t_h^{\alpha_2}$ and $\alpha_1 + \alpha_2 < 1$ we can solve expressions ($C.6$) and ($C.7$) for optimal healthcare services

$$m_i^* = \left[\alpha_1 A \left(\frac{R^i}{R}\right) \left(\frac{1}{w} \frac{\alpha_2}{\alpha_1}\right)^{\alpha_2}\right]^{\frac{1}{1-\alpha_1-\alpha_2}} \times (\phi)^{\frac{-\alpha_2}{1-\alpha_1-\alpha_2}} \times (p_m)^{\frac{\alpha_2-1}{1-\alpha_1-\alpha_2}} \text{ for } i \varepsilon \{s,d\}. \tag{C.9}$$

Using this closed form solution we state our first result.

Proposition 1. *If the investment return of sons is larger than the investment return of daughters, i.e. $R^s > R^d$, then a price change in healthcare services p_m, or a change in the utility cost of time investments into health ϕ, or a quality improvement of healthcare services A will lead to larger health care demand adjustments of medical services in families with sons so that* $\left|\frac{dm(p_m;R^s)}{dp_m}\right| > \left|\frac{dm(p_m;R^d)}{dp_m}\right|$, $\left|\frac{dm(\phi;R^s)}{d\phi}\right| > \left|\frac{dm(\phi;R^d)}{d\phi}\right|$, *and* $\left|\frac{dm(A;R^s)}{dA}\right| > \left|\frac{dm(A;R^d)}{dA}\right|$.

Proof of Proposition 1. Deriving expression (C.9) with respect to p_m and ϕ respectively we have

$$\frac{dm}{dp_m} = \left(\frac{\alpha_2 - 1}{1-\alpha_1-\alpha_2}\right)\left[\alpha_1 A\left(\frac{R^i}{R}\right)\left(\frac{1}{w}\frac{\alpha_2}{\alpha_1}\right)^{\alpha_2}\right]^{\frac{1}{1-\alpha_1-\alpha_2}} \times (\phi)^{\frac{-\alpha_2}{1-\alpha_1-\alpha_2}} \times (p_m)^{\frac{\alpha_1+2\alpha_2-2}{1-\alpha_1-\alpha_2}} < 0,$$
$$\frac{dm}{d\phi} = \left(\frac{-\alpha_2}{1-\alpha_1-\alpha_2}\right)\left[\alpha_1 A\left(\frac{R^i}{R}\right)\left(\frac{1}{w}\frac{\alpha_2}{\alpha_1}\right)^{\alpha_2}\right]^{\frac{1}{1-\alpha_1-\alpha_2}} \times (\phi)^{\frac{\alpha_1-1}{1-\alpha_1-\alpha_2}} \times (p_m)^{\frac{\alpha_2-1}{1-\alpha_1-\alpha_2}} < 0,$$
$$\frac{dm}{dA} = \left(\frac{-\alpha_2}{1-\alpha_1-\alpha_2}\right)\left[\alpha_1 A\left(\frac{R^i}{R}\right)\left(\frac{1}{w}\frac{\alpha_2}{\alpha_1}\right)^{\alpha_2}\right]^{\frac{1}{1-\alpha_1-\alpha_2}} \times (\phi)^{\frac{\alpha_1-1}{1-\alpha_1-\alpha_2}} \times (p_m)^{\frac{\alpha_2-1}{1-\alpha_1-\alpha_2}} > 0,$$

and further deriving these expressions with respect to R^i we get

$$\frac{d^2m}{dp_m \times dR^i} < 0, \frac{d^2m}{d\phi \times dR^i} < 0, \frac{d^2m}{dA \times dR^i} > 0$$

which establishes the result. ■

Proof. Substituting (C.9) into (C.4)–(C.8) we can calculate closed form solutions t_h^*, h^* and s^*. Deriving these solutions w.r.t. p_m, ϕ, A and R^i as in the proof of Proposition 1 establishes the result. ■

Additional results concern the effect of medical prices p_m, time cost factor ϕ, and health care technology (quality) A on healthy time investments t_h, child health h, and household savings s. The following proposition summarizes the results.

Proposition 2. *If $R^s > R^d$, then a price change in healthcare services p_m, or a change in the utility cost of time investments into health ϕ, or a change in the health care technology (quality) A lead to larger level adjustments of healthy time investments t_h, health h,and household savings s in families with sons.*

Proof. Substituting (C.9) into (C.4)–(C.8) we can calculate closed form solutions t_h^*, h^* and s^*. Deriving these solutions w.r.t. p_m, ϕ, A and R^i as in the proof of Proposition 1 establishes the result. ■

Next, we extend the base model above and allow for health dependent survival probabilities of children $\pi = \Pi(h)$ as well as an additional input of factor X into the health production of children

$h = H(m, t_h, X)$. Variable X could comprise food, clean water, etc. that can be purchased at market price p_X. This introduces an additional return channel for health investments through increased life expectancy of children directly in the utility function.[30] Preferences are written as

$$u(c, t_l) + \delta \times \Pi(H(m, t_h, X)) \times v(z, \bar{b}),$$

and the solution to the optimization system is

$$u_c = \frac{\delta \overbrace{\Pi(H(m, t_h, X))}^{\text{Survival effect}} v'}{\left(\frac{1}{R}\right)}, \tag{C.10}$$

$$u_c = \frac{u_{t_l}}{w}, \tag{C.11}$$

$$u_c = \frac{H_m \left(\overbrace{u_c \frac{R_s}{R}}^{\text{Market return}} + \overbrace{\delta \Pi' v}^{\text{Survival effect}} \right)}{p}, \tag{C.12}$$

$$u_c = \frac{H_{t_h} \left(\overbrace{u_c \frac{R_s}{R}}^{\text{Market return}} + \overbrace{\delta \Pi' v}^{\text{Survival effect}} \right)}{\phi w}, \tag{C.13}$$

$$u_c = \frac{H_X \left(\overbrace{u_c \frac{R_s}{R}}^{\text{Market return}} + \overbrace{\delta \Pi' v}^{\text{Survival effect}} \right)}{p_X}, \tag{C.14}$$

$$c + p_m m + p_x X + \frac{z}{R} = w \times t_w + \frac{R_s H(m, t_h, X)}{R}. \tag{C.15}$$

The Euler equation in expression ($C.10$) now includes the survival probability as a function of health investments. In addition, the right hand sides in expressions ($C.12$) and ($C.14$) show how increases in time investments into health t_h, healthcare spending m, and factor X will increase the utility return per dollar spent through two channels: (i) the market return channel, expression $u_c \frac{R_s}{R}$, measures the direct investment return of higher health levels through the gender specific market

[30]Note that in a standard CRRA utility function, the level of utility would be negative when the coefficient of relative risk aversion is larger than one. When survival rates do not depend on health, the level of period utility flow does not matter. However, when the survival probability of offspring becomes a choice, the level of utility is not irrelevant anymore. For instance, parents would prefer no children if the level of utility of children in the second period is negative. To avoid this problem, we follow Hall and Jones (2007) and add a positive constant $\bar{b}$ into the second period utility function of children so that $v(z, \bar{b}) > 0$ for any $z > 0$.

return R^i in the second period; (ii) the survival effect channel, expression $\delta\Pi' v$, measures the increase in utility of period two consumption z from a higher survival probability due to better health.

We are not able to derive closed form solutions for this version of the model, but we can derive results similar to Proposition 1 and Proposition 2 numerically. For the numerical solution we assume that preferences follow $\ln(c) + \ln(1 - t_w - \phi t_h) + \delta\Pi(h)\ln\left(z + \bar{b}\right)$, health is produced according to $H(m, t_h, X) = A\left(\alpha_1 (A_m m)^\rho + \alpha_2 t_h^\rho + (1 - \alpha_1 - \alpha_2) X^\rho\right)^{\frac{1}{\rho}}$, and the survival function, following Hall and Jones (2007), is $\Pi(h) = 1 - \frac{\theta}{1+h}$. Additionally we assume the following parameter values in the model: $\delta = 0.99$, $A = 1$, $A_m = 1$, $\alpha_1 = 0.33$, $\alpha_2 = 0.33, \rho = 0.01$, $\theta = 0.5$, $p_m = 0.3$, $p_x = 0.3$, $R_s = 1.025$, $R_d = 0.98 \times R_s$, $R = 1.04$, $\phi = 1$, $\bar{b} = 5.5$, and $w = 8$.

The second column of Figure C.2 shows the derivatives of m^*, t_h^*, h^*, and s^* w.r.t. medical prices ϕ by gender. The figures indicate that the level effects of changes in the time cost of health care is larger for sons. Similar results can be derived w.r.t. changes in medical technology (i.e., total factor productivity in the health production function A_m) or the price of healthcare p_m. We present these results in Figures C.3 and C.4, respectively.

Our numerical results suggest that a healthcare reform that leads to a reduction in healthcare prices, the time cost of accessing healthcare, or increases in the quality of healthcare all cause higher usage of healthcare services in families with sons because the returns of better health of sons are higher $R^s > R^d$ than the returns of better health of daughters.

The establishment of health clinics following the NHP, as shown in Figure 1(a), improves access to medical care and reduces the overall cost of obtaining medical care. For example, prior to the NHP a household might have had to travel to a location further away to see a doctor. Additional clinics and health posts established through the NHP will reduce the travel time and thereby reduce the cost of accessing healthcare. The model discussed in this section suggests that the household will now disproportionately increase health investments into sons as a reaction to the NHP due to differences in expected market returns of sons vs. daughters. This happens in the absence of any gender preference bias for sons within the family. If we add such a bias into our theoretical model, our results would be even stronger.

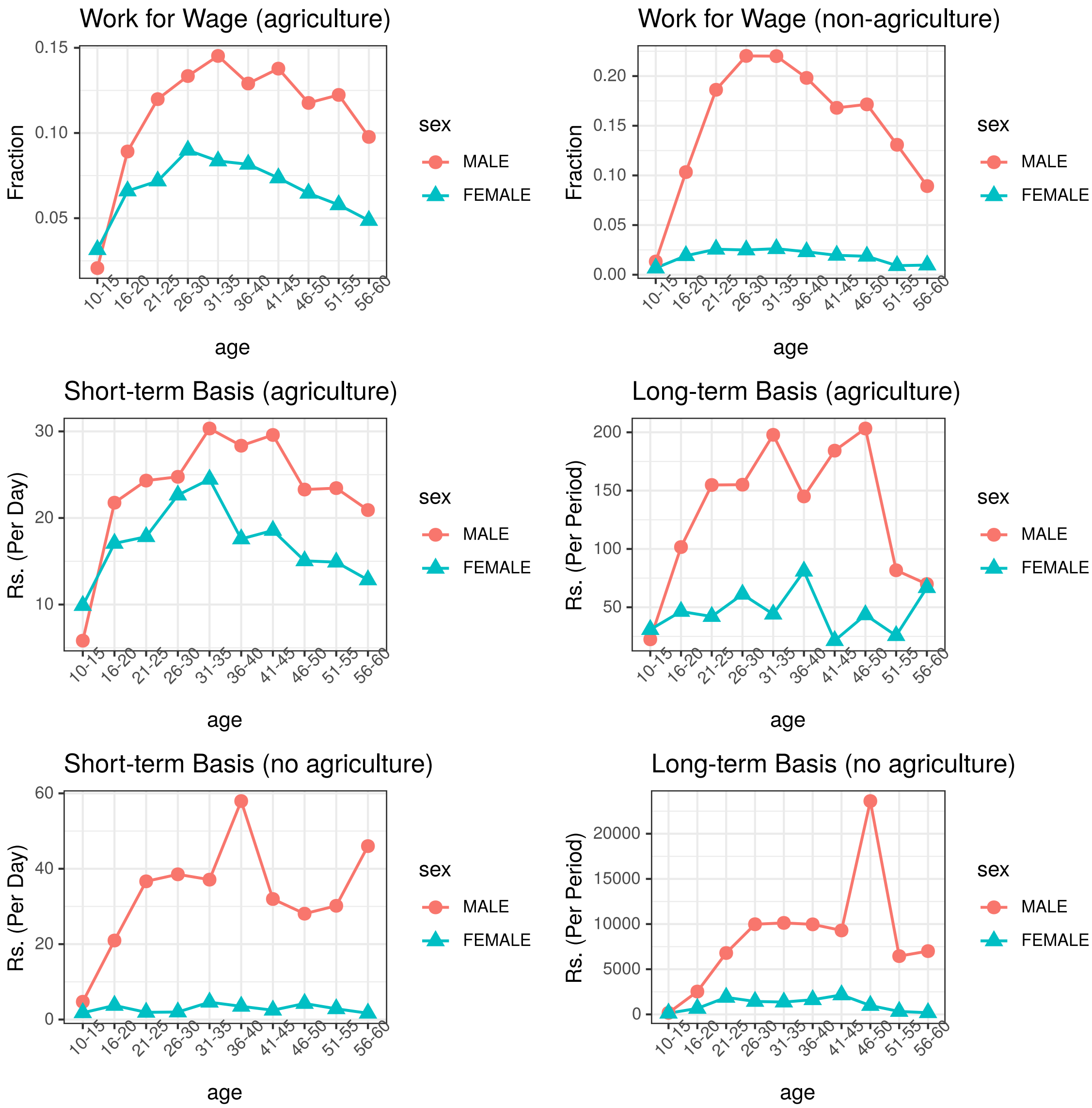


Figure C.1: **Wage Employment in Agriculture and Non-Agriculture Sector (by Gender)**

Notes: Wage related questions in the survey are asked for short term (per day) and long term work (per period, e.g., per year) for agriculture and non-agriculture work, separately. For calculation of average wages, the sample is restricted to people who reported working (either in agriculture, non-agriculture, or were self employed).
Data Source: Nepal Living Standard Survey (1996 and 2003). Authors' calculations.

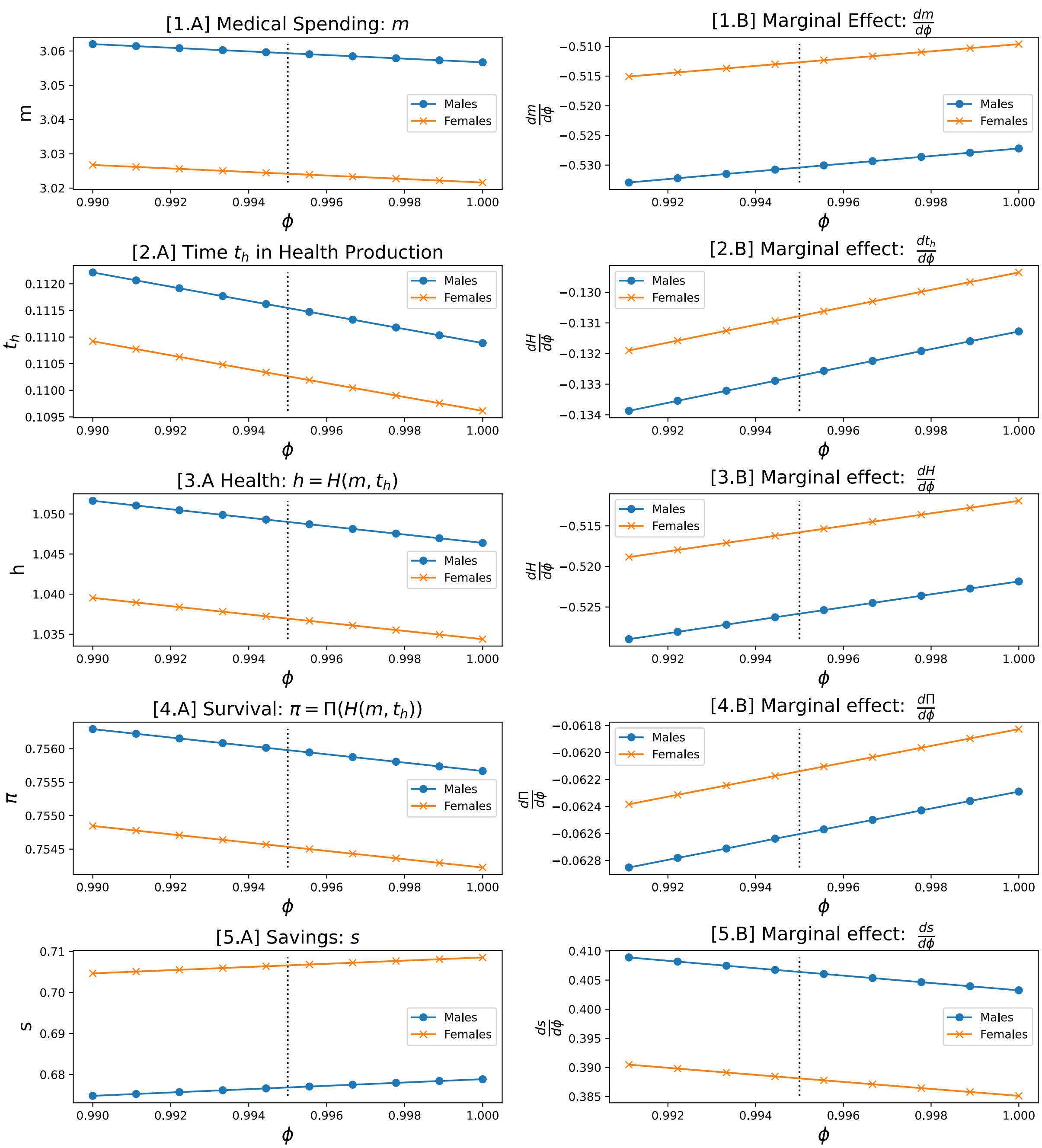


Figure C.2: **Effect of Time Costs of Health Care based on Theoretical Model**

Notes: The time cost associated with health care is denoted ϕ. For the numerical solution we assume that preferences follow $\ln(c) + \ln(1 - t_w - \phi t_h) + \delta\Pi(h)\ln\left(z + \bar{b}\right)$, health is produced according to $H(m, t_h, X) = A\left(\alpha_1 (A_m m)^\rho + \alpha_2 t_h^\rho + (1 - \alpha_1 - \alpha_2) X^\rho\right)^{\frac{1}{\rho}}$, and the survival function, following Hall and Jones (2007), is $\Pi(h) = 1 - \frac{\theta}{1+h}$. Additionally we assume the following parameter values in the model: $\delta = 0.99$, $A = 1$, $A_m = 1$, $\alpha_1 = 0.33$, $\alpha_2 = 0.33$,$\rho = 0.01$, $\theta = 0.5$, $p_m = 0.3$, $p_x = 0.3$, $R_s = 1.025$, $R_d = 0.98 \times R_s$, $R = 1.04$, $\phi = 1$, $\bar{b} = 5.5$, and $w = 8$.

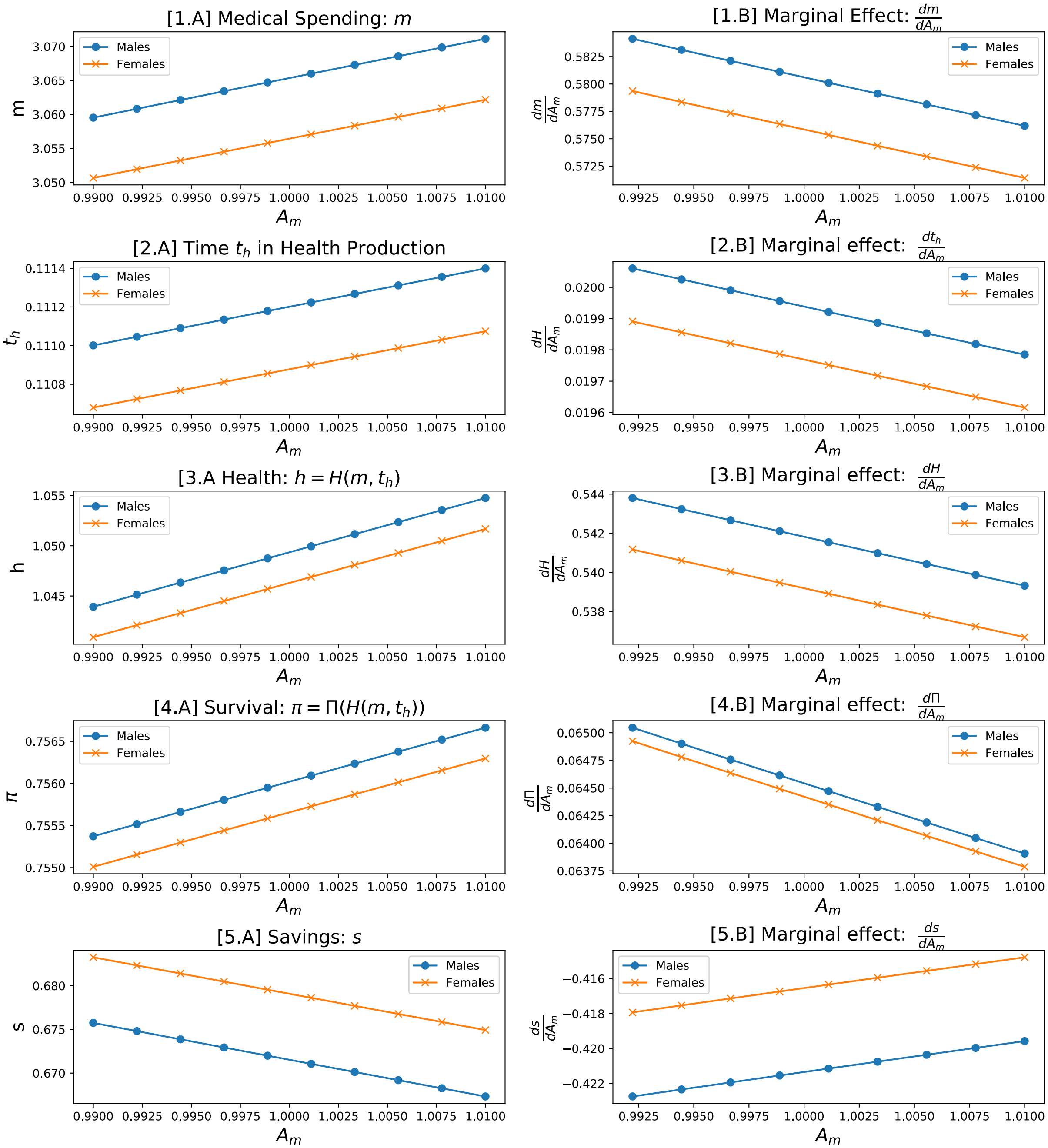


Figure C.3: **Change in healthcare TFP A_m**

Notes: For the numerical solution we assume that preferences follow $\ln(c) + \ln(1 - t_w - \phi t_h) + \delta \Pi(h) \ln\left(z + \bar{b}\right)$, health is produced according to $H(m, t_h, X) = A\left(\alpha_1 (A_m m)^\rho + \alpha_2 t_h^\rho + (1 - \alpha_1 - \alpha_2) X^\rho\right)^{\frac{1}{\rho}}$, and the survival function, following Hall and Jones (2007), is $\Pi(h) = 1 - \frac{\theta}{1+h}$. The following parameter values are used to plot the graphs: $\delta = 0.99$, $A = 1$, $A_m = 1$, $\alpha_1 = 0.33$, $\alpha_2 = 0.33$, $\rho = 0.01$, $\theta = 0.5$, $p_m = 0.3$, $p_x = 0.3$, $R_s = 1.025$, $R_d = 0.98 \times R_s$, $R = 1.04$, $\phi = 1$, $\bar{b} = 5.5$, and $w = 8$.

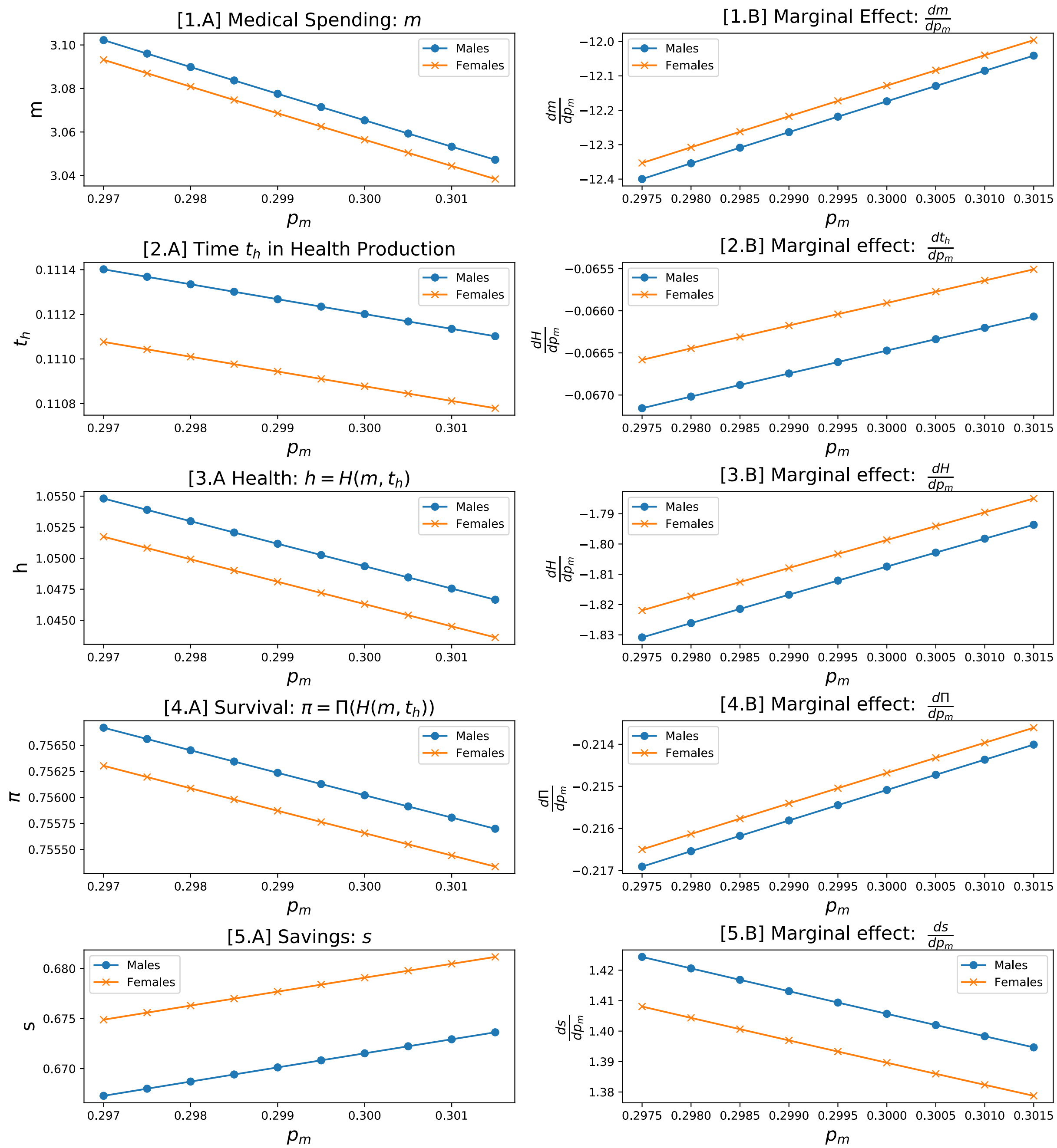


Figure C.4: **Change in Medical Price** p_m

Notes: For the numerical solution we assume that preferences follow $\ln(c)+\ln(1-t_w-\phi t_h)+\delta\Pi(h)\ln\left(z+\bar{b}\right)$, health is produced according to $H(m,t_h,X) = A\left(\alpha_1(A_m m)^\rho+\alpha_2 t_h^\rho+(1-\alpha_1-\alpha_2)X^\rho\right)^{\frac{1}{\rho}}$, and the survival function, following Hall and Jones (2007), is $\Pi(h)=1-\frac{\theta}{1+h}$. The following parameter values are used to plot the graphs: $\delta=0.99$, $A=1$, $A_m=1$, $\alpha_1=0.33$, $\alpha_2=0.33$, $\rho=0.01$, $\theta=0.5$, $p_m=0.3$, $p_x=0.3$, $R_s=1.025$, $R_d=0.98\times R_s$, $R=1.04$, $\phi=1$, $\bar{b}=5.5$, and $w=8$.