Unintended Environment and Health Consequences of Distortionary Fertilizer Subsidies*

Gowthami Venkateswaran† Kathy Baylis‡ Andrew Hultgren†
Hemant Pullabhotla§

January, 2025

Abstract

Governments worldwide subsidize agricultural inputs to support farmers and increase food production. Selective subsidies that result in exceptionally low prices encourage farmers to deviate from optimal application levels and result in the overuse of fertilizers. This paper examines the unintended environmental and health consequences of increased fertilizer use driven by selective subsidy reforms. In 2010, India implemented a fertilizer subsidy change favoring nitrogen, which led to lower prices relative to phosphorus and potassium fertilizers. Leveraging the timing of this policy and exploiting exogenous variation in pre-determined geographic characteristics such as soil texture and river flow direction, I find significant effects of the subsidy on nitrogen pollution in downstream water bodies and infant mortality in rural areas. For every 1 % percent increase in nitrate levels, I find a 1.6 % increase in rural infant mortality rates.

JEL: Q18, H23, Q53, Q58, I18
Keywords: Fertilizer subsidies, Externalities, Water Quality, Health

*Preliminary Draft (please do not cite)
†University of Illinois Urbana-Champaign. gv4@illinois.edu
‡University of California Santa Barbara
§Deakin University

1 Introduction

Input subsidies are a common policy tool used by governments worldwide to reduce production costs and increase agricultural productivity. Among agricultural input subsidies, fertilizer subsidies are the most common and financially substantial. However, maintaining fertilizer subsidies over the long term can be fiscally challenging, prompting some governments to consider reforms or even complete removal (Gautam, 2015). In some cases, governments choose to phase out subsidies gradually rather than eliminate them all at once. This selective subsidy removal can create price distortions and affect fertilizer use patterns, often in unintended ways (Gulati & Banerjee, 2015). While a large body of literature has focused on the impacts of fertilizer subsidies when they are first introduced (Holden, 2019), less attention has been paid to the consequences of selective subsidy removals. This gap is significant because subsidies often make fertilizers more affordable, leading to overuse. Excessive fertilizer application increases nutrient runoff, contaminating water bodies and creating health risks for downstream populations.

In many developing countries, where water regulations are often insufficient, vulnerable communities are particularly at risk. Despite the significant consequences, the relationship between runoff-induced water pollution and public health in these contexts remains understudied. I examine this issue in the context of fertilizer subsidies in India where a reform favored nitrogen and altered fertilizer use patterns.

In this paper, I examine the effect of a selective subsidy reform favoring nitrogen fertilizers on unintended environmental and health outcomes in India. I study whether policy-induced increases in nitrogen use contributes to greater runoff and nitrate pollution, and whether this, in turn, affects infant mortality in downstream populations. Between 2005 and 2010, the subsidy burden of the Indian government increased by 500% as global prices increased substantially (Ravinutala, 2016). In response, the government introduced the Nutrient-Based Subsidy (NBS) program in 2010, which reduced price support for most major fertilizers, including phosphorus and potassium, but excluded nitrogen [1]. Nitrogen prices were left untouched since policymakers considered the decontrol of the urea sector to be politically sensitive (Kishore et al., 2021). As a result phosphorus and potassium fertilizer prices in-

[1] Prior to 2010, all major fertilizers N, P, and K based-fertilizers were controlled by the government. The government set the Maximum Retail Prices (MRP) for N, P, and K-based fertilizers. After 2010, the government moved away from product-based subsidy to nutrient-content-based subsidy for P and K-based fertilizers. See section 2 for more details on the policy.

creased while urea remained relatively cheaper under continued government price controls. This policy change led to a significant increase in nitrogen fertilizer use among farmers, with total nitrogen usage tripling between 2010 to 2015 (Gulati & Banerjee, 2015). I leverage this NBS policy as a natural experiment to examine the impact of relative increases in nitrogen subsidy on nitrate pollution in downstream river bodies and infant mortality.

India provides a compelling case to study fertilizer subsidies, nitrogen runoff, water pollution, and health for several reasons. First, the Indian government has heavily subsidized inputs, particularly fertilizers, since the Green Revolution. Input subsidies to power, fertilizer, irrigation, and credit comprised 1.5% of India's GDP in 2017 (Ramaswami, 2019). As a result of these subsidies, India is now the second-highest consumer of fertilizers globally, only second to China. Second, India's population is over 1.2 billion, and millions live in rural areas and rely on river water for bathing and drinking. Unlike developed countries, regulations on water pollution remain limited and less developed, making these populations even more vulnerable. The World Health Organization (WHO) reports that three out of every 1000 children under five years died due to water pollution (WHO, 2004). Finally, India is one of the few developing countries that has detailed data on fertilizer use, nitrate measures in water bodies, and infant mortality measures, thereby allowing me to study the implications of increased nitrogen use on water quality and infant mortality.

There are two main challenges in identifying the causal relationship between fertilizer use and infant mortality. First, nitrogen use and nitrogen runoff are endogenous. Second is the fact that fertilizer use can reduce infant mortality through other channels such as increased yields, increased incomes, and better health care investments.

To deal with the first challenge, we exploit exogenous variation in predetermined geographic characteristics to deal with both the endogeneity of fertilizer use and fertilizer runoff. In particular, I use the variation in the percentage of clay soils. I use the percentage of clay composition since it does not change easily with management practices. Clay soils are relevant in this context for two main reasons. 1) Soils with higher clay content are generally more productive because they can hold moisture and nutrients longer than coarser soils (Burke et al., 2019). For this reason, studies also show that the marginal returns to fertilizer application are generally higher in clay soils.[2] The identifying strategy relies on the assumption that the amount of clay

[2]Soils with a clay content greater than 70 percent may not be ideal as excessive clay can hinder water drainage and aeration, potentially reducing crop productivity. My data does not contain areas with extremely high clay content.

affects fertilizer use decisions since clay soils tend to have higher marginal returns to fertilizer application. 2) Clay soils contribute to more surface water runoff. Soils with a relatively high percentage of clay will have smaller pore space and therefore lower infiltration rates. Consequently, soils with higher levels of clay may result in greater surface runoff than sandy or silty soil.[3]

For the second challenge, confounding issues may arise due to the potential positive effects of increased fertilizer use on infant mortality through enhanced agricultural productivity as suggested by Bharadwaj et al. (2020) and Gollin and Udry (2021). Fertilizer use can enhance crop yields, and improve incomes and overall farmer welfare which may led to better food security and improvements in infant health. To address this issue, I employ an upstream-downstream specification by exploiting the flow direction of river bodies. Rather than focusing on local soil characteristics near the infant populations, I trace the watershed upstream and use the soil characteristics of these upstream areas to isolate the effects of nitrate water contamination on downstream populations.

The introduction of fertilizer subsidies during the Green Revolution was instrumental in boosting agricultural yields, improved nutrition, and economic output (Bharadwaj et al., 2020; Gollin & Udry, 2021; von Der Goltz et al., 2020). However, recent evidence suggests that the benefits of increased fertilizer use have begun to plateau, as higher application rates do not translate into productivity gains anymore (Itin-Shwartz, 2024; Mueller et al., 2017; Wuepper et al., 2020). A major consequence of prolonged fertilizer subsidies has been the excessive use of fertilizers, with India leading this growing issue in South Asia. Nitrogen fertilizers, such as urea are especially overused, due to their heavy subsidization (Gulati & Banerjee, 2015; Kishore et al., 2021). For example, urea has remained consistently about four times cheaper for Indian farmers than international prices and has not changed in prices for over a decade, leading to environmental and health challenges in the region (Cassou et al., 2017; Huang et al., 2017; Kurdi et al., 2020).

Water contamination is one of the most significant consequences of excessive fertilizer use, as nutrient runoff from agricultural lands pollutes surface water bodies and groundwater. Globally, only about 35% of nitrogen applied as fertilizer is absorbed by crops; the rest leaches into water systems, leading to degraded water quality through issues such as algal blooms, hypoxic zones, and nitrate contamination of drinking

[3]http://www.faculty.luther.edu/ bernatzr/RainfallRunoff/comet/hydro/basic/Runoff/print_version/04-soilproperties.htm

sources(Lassaletta et al., 2014; Zhang et al., 2018). These nitrogen-related contaminants pose severe health risks for downstream populations (Damania et al., 2019; James et al., 2005). For infants in particular, elevated nitrate levels in drinking water can lead to methemoglobinemia, also called the blue baby syndrome, a potentially fatal condition that impairs oxygen transport in the blood (Knobeloch et al., 2000)

Water contamination in developing countries from agricultural runoff is particularly concerning, as millions of people rely on untreated river water for drinking, bathing, and irrigation. The lack of effective regulatory oversight and water treatment infrastructure makes these communities even more vulnerable to the harmful effects of water pollution. Greenstone and Hanna (2014) show that, unlike air pollution regulations, water pollution regulations remain weak in India. Unlike in developed countries with stronger regulations, the absence of proper monitoring and management in these regions allows the impacts of polluted water to go unchecked, leading to long-term public health challenges. Nitrate contamination from agricultural runoff is one of the most common chemical pollutants in groundwater aquifers worldwide (Mateo-Sagasta et al., 2017), posing severe health risks, particularly for infants who may develop methemoglobinemia.

To measure the effect of the policy on nitrogen use, I combine district-level fertilizer consumption data between 2000 and 2015 with soil data. I first show that the policy does lead to increased nitrogen consumption, focusing on the periods before and after the introduction of the policy. Then, using data from around 500 water monitoring stations with data on nitrate levels, I study the changes in nitrate levels in water stations with high levels of clay in their upstream watershed catchment area. Finally, using household data from approximately 10,000 clusters, I study the impacts of the policy on infant mortality. In my main results, I find that DHS clusters located close to water monitoring stations with high fractions of clay soil in their upstream areas face increased risks of infant mortality. The main reduced form regressions are conditional on state-by-year fixed effects, absorbing any differential trends across states over time and DHS cluster fixed effects, accounting for cross-sectional time-invariant characteristics. Even-study style regressions indicate stable pre-trends validating the identification strategy. I focus on infant deaths within one year of birth to ensure a precise exposure period, avoiding the complications of long-term differential effects. I find that a 50% increase in nitrates results in 0.03 additional deaths in rural DHS clusters with high upstream clay in the post-policy period.

I conduct a series of robustness checks to ensure the validity of the results of this study. First, instead of categorizing clay soils into binary high and low variables, I adopt a specification where the upstream clay percentage is continuous. Second, I conduct falsification tests by using the main specification on other water quality indicators available in the data. These include measures such as pH, temperature, conductivity, dissolved oxygen, and fecal coliform. The policy should theoretically have no impact on any of the other water quality measures and as expected, I do not find significant impacts on these measures. Next, there is no consensus on what the correct distance is to use for an upstream watershed specification. I use the entire watershed retrieved for each station in the main equation. However, I also run a version that contains only a 100km upstream watershed area to account for potential pollution decay. Additionally, I also conduct placebo tests using the downstream watershed catchment area for all monitoring stations. Clay characteristics below the water stations should not have an impact on the pollution measures in the station. As expected, I did not find any changes in nitrate pollution with this model. When estimating infant mortality rates in the DHS clusters, I use a buffer of 20km around each cluster to locate nearby water stations in the main specification. I also test other buffer kilometers and find that results diminish as the infant clusters are located further from the water stations.

This paper makes three main contributions to the literature. First, it contributes to the literature on the effects of pollution on health. Much of the literature in economics focuses on air pollution, largely due to better data availability on air quality (Graff Zivin & Neidell, 2013). Majority of these studies place a focus on developed nations like the US (Currie & Neidell, 2005; Currie et al., 2009; Deryugina et al., 2019; Knittel et al., 2016; Schlenker & Walker, 2016). Among the literature that studies water pollution, the majority of them place a focus on measuring the impacts of regulatory measures such as the Clean Water Act in the US (Cai et al., 2016; Currie et al., 2013; Keiser & Shapiro, 2018). However, pollution sources and regulatory conditions differ significantly in developing countries. In India, while air pollution reduction policies have made some progress in lowering air pollutants, regulations on water pollution remain limited and less developed (Greenstone & Hanna, 2014).

While a few studies consider the impact of water pollution on health and human capital through improved sanitation channels (Gamper-Rabindran et al., 2010; T. Garg et al., 2018; Motohashi, 2024), industrial waste (Do et al., 2018; Ebenstein, 2012; Hagerty & Tiwari, 2022), pesticides (Dias et al., 2023; Lai, 2017; Skidmore et al., 2023), insecticides (Taylor, 2021), algal blooms (Jones, 2019; Taylor & Heal, 2023), there are relatively fewer studies that look at agricultural runoff which mostly can be attributed to difficulties in accessing data on runoff.

The closest work that complements this study is by Brainerd and Menon (2014). They study the effects of overall river fertilizer pollutants on infant health in India and find significant negative impacts. They exploit seasonal prenatal exposure to agrichemicals to identify the impact of agrichemical contamination on various measures of child health. I address a critical aspect that might remain unaddressed in this paper, the potential confounding effect of seasonal factors on the relationship between fertilizer pollution and infant mortality. By focusing specifically on the NBS policy, I disentangle the effects of fertilizer pollution from seasonal variation.[4]

Second, the paper also contributes to the broader understanding of the implications of the negative consequences when governments phase out input subsidies and cause relative price differences among inputs. There is a large body of literature studying agricultural subsidies, Holden (2019) provides an exhaustive review. While a growing body of literature explores the effect of subsidy-driven distortions on input use, productivity, and farmer welfare(Adamopoulos & Restuccia, 2014; Chakraborty et al., 2024; Donovan, 2021; S. Garg & Saxena, 2023; Hsieh & Klenow, 2009; Kurdi et al., 2020; Restuccia et al., 2008), less is known about the effects of partial phase-outs of subsidies on health outcomes. Environmental and human capital concerns are yet to be taken into policy consideration, although populations in developing countries are particularly vulnerable to water source pollution (Itin-Shwartz, 2024)

Third, the paper also contributes to improvements in measuring downstream pollution. Most studies conduct the upstream-downstream framework at some administrative level. This would involve mapping upstream districts to a given district, for example. While studying water pollution, researchers in the US can rely on the National Hydrography Dataset to retrieve watersheds and river networks. However, mapping precise upstream watersheds are still a challenge in developing country settings. There are few papers that have contributed to improved spatial computational methods in developing country settings (T. Garg et al., 2018; Hagerty & Tiwari, 2022). Using an open-source global watershed mapping platform developed by Heberger (2022), I enhance the spatial analysis by mapping upstream watersheds for each water monitoring station and running the analysis at the water station level allowing for precise identification of watersheds contributing to nitrate runoff.

[4] Another related paper is Zaveri et al. (2020) where they look at the effects of nitrate pollution on long-term impacts on adults.

Finally, despite the wide criticisms of the NBS program, there are no causal estimates of its impacts on the environment and health. I provide important and timely evidence on a policy that is still in place and frequently discussed by the government, policymakers (Gulati & Banerjee, 2015) and media [5] [6] [7].

The remainder of the paper proceeds as follows. Section 2 provides background details on the policy and discusses the expected effects of this policy on nitrogen use. Section 3 describes the data, and section 4 elaborates on the identification strategy. I report results in section 5 and section 6 concludes the paper.

2 Background

2.1 Nutrient-Based Subsidy

India has subsidized fertilizers since the Green Revolution, allowing farmers to access fertilizers at affordable prices. However, between 2005 and 2010, the subsidy burden of the Indian government increased by 500% as international prices increased substantially (Ravinutala, 2016). This pushed the government to introduce changes to the subsidy structure to alleviate the financial burden and encourage a more balanced use of fertilizers.

In April 2010, the Indian government implemented a significant change in how fertilizer subsidies were handled, transitioning from a product-based subsidy to a nutrient-based subsidy system. This new approach focused on providing subsidies based on the nutrient content of fertilizers rather than providing fixed prices for specific fertilizer products. Under this scheme, subsidies are based on the content of key nutrients: nitrogen (N), phosphorus (P), potassium (K) and sulphur (S).[8] The government announces a fixed subsidy rate in rupees per kilogram for each nutrient every year, determined based on international prices and inventory levels. These per-kg rates are then converted into per-ton subsidies for each product based on their nutrient composition. [9] For instance, Diammonium Phosphate (DAP) is the widely

[5]https://www.downtoearth.org.in/agriculture/cacp-recommends-centre-to-bring-urea-under-nbs-regime-to-check-overuse-89907

[6]https://economictimes.indiatimes.com/industry/indl-goods/svs/chem-/-fertilisers/profitability-for-phosphatic-fertiliser-players-to-improve-at-current-nutrient-based-subsidy-rates-report/articleshow/97122078.cms?from=mdr

[7]https://indianexpress.com/article/explained/explained-economics/deregulating-non-subsidised-fertilisers-9447888/

[8]NBS is available for DAP, MOP, MAP, TSP and 12 other grades of complex fertilizers

[9]The top three major nutrients used in India are nitrogen (N), phosphorus (P) and potassium (K). Urea is the main source of nitrogen and is the most widely used fertilizer. Diammonium Phosphate

used product for phosphorus. Before 2010, the governemnt set a fixed price for DAP but after the NBS reform, DAP is subsidized based on its nutrient composition which is 18-46-0-0 representing the ratio of N:P:K:S.

Using the 2023-2024 rabi season subsidy rates for the product DAP, I show an example of how the subsidies are calculated. First, the nutrient composition of DAP is 18-46-0-0. The subsidy rates for N is 47.02 Rs/kg, P is 20.82 Rs/kg and since DAP doesn't contain any K or S, the subsidy rate per metric ton (1000 kgs) of DAP are calculated as the following: Nitrogen subsidy: 18% of 1000 kg = 180 kg, 180*47.02 = Rs. 8463. Phosphorus subsidy: 46% of 1000 kg = 460 kg, 460*20.82 = Rs. 9577.20. Adding 8463.60 + 9577.20, the subsidy for DAP is Rs. 18,040 per metric ton.

Following the announcement of the nutrient subsidies each season, manufacturers can set their own prices for fertilizers after taking the government-announced subsidy rates. The government transfers the subsidies directly to fertilizer companies based on actual sales made by retailers. Farmers pay the final subsidized price for each of the products. In recent years, these sales are recorded through point-of-sale devices at retail shops to ensure fertilizer availability to farmers at the right prices. Since manufacturers set prices for P and K fertilizers after decontrol under the NBS policy, the prices for those fertilizers increased significantly. Prices of P and K based fertilizers went up by over 150% and 255% respectively.

However, urea, the most widely used fertilizer in India, was exempt from the NBS scheme. Urea is the only fertilizer that is currently sold at a government-fixed retail price. The price of urea has remained at almost at the same level for more than 15 years. Urea prices were left untouched since policymakers considered the decontrol of the entire urea sector to be highly sensitive (Kishore et al., 2021).

This led to a price distortion with high relative subsidies for N and reduced subsidies for P&K. Due to this price increase, the share of nitrogen usage increased notably within just three years of the policy implementation. The share of nitrogen increased by 9.5% from 58.8% to 68.4%, suggesting that the imbalance in fertilizer use is policy-induced (Ansari & Sheereen, 2022). Figure 2 shows farmer-reported prices from the cost of cultivation survey data, demonstrating that the NBS policy resulted in P and K price increases relative to N prices.

Gulati and Banerjee (2015) argue that excluding urea from the NBS scheme was a significant oversight in India's fertilizer policy reform. They cite the Planning

(DAP) is the most widely used phosphate fertilizer and Muriate of Potash (MOP) is the most widely used potassium based fertilizer.

Commission's twelfth plan document, which acknowledges that "NBS roll-out was seriously flawed since urea was kept out of its ambit. Urea prices remain controlled with only a 10 percent rise at the time of adoption of the NBS in 2010. Meanwhile, prices of decontrolled products doubled" (Planning Commission, 2013, p. 14). For more information on the policy see Gulati and Banerjee (2015).

2.2 Clay soil characteristics

Increased response to fertilizer applications I rely on the fraction of clay available in the districts and upstream watersheds as a source of variation for fertilizer application and nutrient runoff. The soil science literature notes that clay-rich soils have distinct properties that enhance both nutrient retention and water-holding capacity, making clay soils particularly responsive to fertilizer applications. Unlike sandy soils, which have larger particles and lower surface area, clay soils consist of finer particles with higher surface area and cation exchange capacity allowing better nutrient and water retention. This feature supports greater crop growth and increases efficiency in nutrient uptake (Zhu et al., 2023).

Increased runoff A second feature of clay, which fortunately also addresses the concern of increased runoff, is their low infiltration rate due to smaller pore spaces. Clay soils are composed of fine particles with minimal pore scape, resulting in slower water infiltration and a tendency for higher surface runoff. The high water and nutrient retention capacity of clay soils mean that they are ideal for holding nutrients necessary for crop growth, but this benefit comes with trade-offs in terms of water management. As water moves slowly through clay, it can lead to water logging.

In agricultural regions with high clay content, the combination of nutrient-holding capacity and slower drainage can enhance fertilizer efficiency, as few nutrients are lost to the deeper soil layers. However, this same characteristic makes clay soil particularly prone to surface water runoff, especially during heavy rainfall. This is significant in my study context, as my identification relies on the exogenous property of clay soils in increased fertilizer efficiency and runoff to study nitrate pollution in neighboring water bodies [10].

[10] More details on soil properties can be found here (University of Wisconsin-Madison Division of Extension, n.d.)

2.3 Negative Externality of Excess Nitrogen Fertilizer Use on Water Pollution and Health

The increased use of urea under the NBS program in India may cause unintended negative externalities on river water quality due to the increased amount of nitrates entering nearby water bodies. The application of nitrogen-based fertilizers like urea introduces inorganic nitrogen into the soil where it undergoes a series of chemical transformations. Initially, nitrogen is decomposed to ammonia, which is then oxidized into nitrates and nitrites. While nitrates are absorbed by plants during their growth, excess nitrates that are not utilized by crops can leach into groundwater or runoff into neighboring water bodies. This phenomenon is exacerbated during periods of heavy rainfall, accelerating the washing of nitrates into rivers and lakes.

India's regulatory framework on water pollution is outdated and lacks focus on nitrate contamination. The Central Pollution Control Board (CPCB) was established in 1974 as part of the Water Act of 1974. However, this legislation was primarily targeted at reducing industrial pollution sources and extending sewage plants, with little attention to the regulation of nitrate runoff. In fact, even with strong regulations, nitrate contamination has been shown to be one of the few water pollutants that is increasing globally. The US Environmental Protection Agency considers nutrient pollution to be one of the most widespread, costly and challenging environmental issues today (Taylor & Heal, 2023). Greenstone and Hanna (2014) show that even non-nitrate water quality regulations had no impact due to weak implementation.

Excess nitrogen and phosphorus, primarily from fertilizer, leach into water bodies, feeding the growth of harmful algal blooms in a process known as eutrophication (Nixon, 1995). These blooms can create hypoxic or "dead zones" where aquatic life struggles to survive in due to oxygen depletion. The consequences of nitrate contamination are not limited to the environment, they pose serious risks to human health particularly for populations dependent on surface and groundwater for drinking and other purposes (Li et al., 2021; Liu et al., 2014). Studies have linked high intake of nitrates with the incidence of cancer, blood diseases and other illnesses (Knobeloch et al., 2000; Liu et al., 2014). High nitrate levels in drinking water can lead to methemoglobinemia, commonly known as blue baby syndrome. Methemoglobinemia reduces blood's capacity to transport oxygen. Brainerd and Menon (2014) and Zaveri et al. (2020) establish causal effects of nitrate pollution on birth outcomes, showing that early life exposure to nitrogen can increase the likelihood of infant deaths.

The challenges of nitrogen pollution in water bodies are further complicated because it is nearly impossible to remove nitrates from water at home. The treatment procedure is expensive and not feasible in most homes in India.[11] Installing advanced water treatment systems in rural or low-income areas is often not practical due to a lack of infrastructure and high maintenance costs associated with these technologies.

3 Data

3.1 Fertilizer Use

To estimate the difference in nitrogen use after and before the policy, I need detailed data on fertilizer use. I use the district-level fertilizer consumption data available from the district-level database (DLD) from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). The apportioned database from ICRISAT includes details on districts between 1990 upto 2017. This dataset provides information on the total use of nitrogen, phosphorus and potassium measured in tons, as well as per hectare usage. I also use data files that include season-wise crop area and production at the district-level.

I also use detailed farmer-level data on input decisions for all the main crops in India. This data comes from the Cost of Cultivation Surveys (CCS) [12]. The surveys collect data on input use, crop specific input prices and output prices received by farmers. Data is collected at the plot level in a few villages in selected tehsils all over India. I use four rounds of data collected annually from 2008-2009 to 2018-2019. In each round, the same set of farmers are followed for all planting seasons for three consecutive years. The data covers all major crops, including rice, wheat, maize and corn.

3.2 River Water Quality

To estimate water quality, I use the river water quality data from the National Water Quality Monitoring Programme (NWMP). [13] This dataset includes more than 1100 river monitoring stations. However, only half of these stations record nitrate data.

[11] Home treatment options are limited since methods such as installation of carbon filters or boiling do not reduce nitrate levels in water. The process of reverse osmosis can remove nitrates; however, it is a significantly expensive technology.

[12] The Cost of Cultivation Surveys are conducted by the Ministry of Agriculture and Farmers Welfare, Government of India

[13] NWMP is collected and managed by India's Central Pollution Control Board.

These stations also come with geospatial information. [14] For this study, I focus on the period from 2007 to 2016, as earlier periods do not have enough monitoring stations recording the main outcome variable nitrates.

The NWMP reports multiple water quality indicators, including nitrate-nitrite levels, fecal coliform, pH levels, dissolved oxygen (DO), temperature and conductivity. My main indicator of nitrogen pollution is the level of nitrate-nitrites. I use the cumulative nitrate-nitrite levels as the leading indicator to evaluate the measure of nitrogen pollution in water bodies. Higher nitrate-nitrite levels indicate greater greater nitrogen contamination. The safety thresholds for nitrate-nitrite levels in water is generally considered to be 10 ml/L [15]

In addition to nitrate-nitrite levels, I examine other indicators of water pollution such as dissoved oxgyen, fecal coliforms, temperature and pH. In figure 4, I plot all the monitoring stations in India that track this data.

To identify the direction of flow of the rivers which is key to my identification strategy, I use the Global Watershed API developed by Heberger (2022) to map the upstream watersheds. This tool provides all the river tributaries and the origins upstream of a given water monitoring station, as well as their corresponding watershed polygons. The use of the Global Watershed API represents a significant data enhancement in my paper. While most studies in this context often conduct analyses at the district level, my approach is more precise since I map each watershed upstream of each water quality monitoring station.

Part of the main identification strategy of the study relies on the measurement of nitrate pollution upstream of a given water monitoring station. These measurements require precise location of the water stations, however, some of the water stations do not have reliable latitude and longitude information. There I discard stations that have an upstream watershed area less than 100 kilometers in the analysis.

3.3 Soil Characteristics

I leverage exogenous variation in soil characteristics to determine fertilizer application. To identify the underlying soil type, I use data from the Harmonized World Soil Database (HWSD) v2.0, released by the Food and Agriculture Organization (FAO). The HWSD v2.0 is a high resolution global soil inventory, providing detailed information on various soil properties in a 30 arc-second (approximately 1km) raster format

[14] Source: https://cpcb.nic.in/wqm/WQMN_list.pdf

[15] https://www.atsdr.cdc.gov/csem/nitrate-nitrite/standards.html

worldwide. The HWSD v2.0 provides additional data for seven distinct soil layers, with the top layer representing about 0-20cm depth. Since I study the implications of fertilizer runoff, I focus on the topsoil from this raster.

I collect soil details at two scales: district-level and watershed-level. For the district-level analysis, I compute district-level fractions of clay, sand and silt using the 2011 district boundaries shapefile of India. In addition to the district-level analysis, I use a more granular version for the water quality monitoring station level analysis. I use shapefiles of upstream watershed boundaries for each water monitoring station to calculate fractions of clay, sand and silt for each upstream watershed boundary.

I exploit exogenous variation in soil characteristics. Soil details are collected from the Harmonized World Soil Database released by the Food and Agriculture Organization (FAO). This data set is a 30 arc-second raster containing details on various soil properties across the globe. I compute the district-level fractions of clayey soil and sandy soil using the 2011 district boundaries of India.

Figure **??** plots the average nitrogen consumption at the district level before the policy, along with the percentage of clay soils in each district. The plot shows that both high clay and low clay districts exhibit high levels of nitrogen use prior to the policy reform.

3.4 Health Outcomes

Infant birth and mortality variables come from the National Family Health Survey - IV (NFHS-4) 2015-2016. The NFHS is the Demographic Health Survey (DHS) version of India. This fourth round of NFHS collected data from January 2015 to December 2016, covering around 600,000 households nationwide. The NFHS collected the following four sets of questionnaires: a household questionnaire, a woman's questionnaire, a man's questionnaire and a biomarker survey.

I specifically focus on the women's questionnaire from the NFHS-4 in this study. These data contain detailed information on the reproductive history, such as the year and month of delivery of every child born for all women in the sample between the ages of 15 and 49. The reproductive histories also include data on when a child was born, their gender, whether they were twins and the birth orders. The data provides information on whether every child born is currently alive, and if not, the age of death is recorded. I choose infant mortality as the primary health outcome since it has been established that high levels of nitrates in the water can affect infants significantly through methemoglobinaemia.

The NFHS 4 also provides the latitude and longitude coordinates for around 28000 sample clusters. The locations of these clusters are displaced slightly for confidentiality reasons. Urban clusters are displaced up to 2 kilometers and rural clusters are displaced up to 5 kilometers, with less than 1% of the rural clusters displaced up to 10 kilometers. To account for this spatial measurement error while using the water quality station data, I draw a 20km buffer around each cluster location and calculate the average nitrate measures in the water stations within this buffer.

Since birth histories are available for all the women in the sample, I am able to construct a pseudo-panel of infant mortality, similar to Do et al. (2018). I construct a cluster-year panel of total number of births, deaths within the first month of life and deaths within a year of birth, see figure A13.Due to data limitations in earlier years, I exclude data prior to 2000 from my analysis.

3.5 Weather data

To account for the influence of precipitation on fertilizer runoff into watersheds, I use the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). CHIRPS provides data from 1981 to near-preset and uses both satellite imagery as well as in-situ station data to create a gridded rainfall product. Using the upstream watershed shapefiles of each water monitoring station, I retrieve the monthly total rainfall in the upstream watershed for each water quality monitoring station. Using the total rainfall upstream might be important since some stations have a smaller upstream watershed, and some others have bigger upstream watersheds. I then construct annual rainfall measured in millimeters as well as monsoon rainfall.

4 Estimation Framework

I aim to estimate the effect of the subsidy on river water pollution and infant mortality in the surrounding populations. There are two main challenges. First, fertilizer application and run-off are endogenous. I need plausibly exogenous variation in nutrient exposure to estimate the causal effects on health outcomes. Second, the increased use of nitrogen can also directly affect child health through increased agricultural productivity.

To address the first problem, the endogeneity of nitrogen use, I rely on exogenous variation from predetermined soil characteristics. Specifically, I leverage the fact that areas with a higher percentage of clay in their soil composition tend to benefit more from fertilizer use (Burke et al., 2019). This relationship comes from the various properties of clay soil. Clay soils have a higher cation exchange capacity, allowing for better retention of nitrogen in plant-available forms. It also helps that clay soils also lead to increased surface runoff. Although they retain water and nutrients well, their low permeability causes the soil to become saturated more quickly, leading to excess nitrogen being carried into river bodies, allowing me to isolate the effect of fertilizer use on nitrate pollution while ensuring the variation in runoff is also through clay soils.

For the second challenge, I focus on the effects of increased nitrogen use in the upstream areas on infant mortality in downstream regions. Since part of the identification strategy relies on measuring the soil type *upstream* of a given river monitoring station, I determine the upstream watershed for each monitoring station by using the flow direction of the rivers. I use the Global Watershed API developed by Heberger (2022) to map the upstream watershed boundaries. This tool provides comprehensive spatial mappings for all the river tributaries and their origins upstream of a given monitor, along with the corresponding watershed polygons. [16] In Figure 5, I explain the identification strategy by highlighting two water monitoring stations. For each station, I map the upstream river network and retrieve the corresponding upstream watershed.

4.1 Differences-In-Differences Design

4.1.1 Nitrogen Use

I use a difference-in-difference approach to estimate the impact of the policy change on environmental and infant mortality outcomes. The policy change in 2010 led fertilizer manufacturers to increase prices for P and K, resulting in farmers shifting toward greater nitrogen use due to its lower price. I treat the sudden price increase of P and K after 2010 as a shock that induced farmers to shift toward greater nitrogen use. I compare districts with higher clay content to those with lower clay content before and after the 2010 policy change. The main identifying assumption in this design is that both sets of districts would have seen their fertilizer use develop along parallel trends in the absence of the NBS policy.

I first estimate the following event-study style regression specified in equation 1:

[16] https://mghydro.com/watersheds/

$$\text{Y}_{ist} = \sum_{k \neq 2009} \beta_k \text{ High Clay}_i \cdot \mathbf{1}\{k = t\} + \phi_d + \gamma_{st} + \epsilon_{dt} \tag{1}$$

The main outcome of interest is the usage of nitrogen, phosphorus and potassium fertilizers in kgs per hectare, in district i, state s and time period t. I denote the treatment variable as High Clay, a dummy variable with the value 1 for districts with above median clay levels and zero otherwise. I omit 2009 as the baseline year, as the policy was launched in early 2010. The coefficients on the interaction term, β_k, recover the change in fertilizer use following the policy change. Each coefficient provides an estimate for the difference between the high and low clay districts before and after the policy. I should expect to see no systematic difference between districts with high and low levels of clay before 2009 to be consistent with the identifying assumption of parallel trends on the counterfactuals. If the policy change in fertilizer prices resulted in consumption shifts, then I should expect to see the coefficients diverge from 0 immediately after 2009. The differences could continue to diverge further over time as prices continued to remain high for P and K fertilizers.

My comparison of high versus low clay districts will be able to recover a lower bound of the effects following the policy change. This is because the districts that are classified into low clay content may still be affected by the policy since their baseline nitrogen use is not zero. I am interested in the residual variation that is not explained by time-invariant characteristics at the district level. I include district-fixed effects to account for district-level observable and unobservable characteristics that are constant throughout my sample. I also add state-by-year fixed effects to account for time-varying factors that differ across states.

Following the event-study style regressions, I also estimate aggregated versions of equation 1 to summarize average treatment effects. I define a post-NBS policy dummy that is equal to one after 2010. Following is my main regression specification:

$$\text{Fertilizer}_{ist} = \beta_1 \text{High Clay}_i + \beta_2 \text{Post}_t + \beta_3 (\text{High Clay}_i \times \text{Post}_t) + \mu_i + \eta_{st} + \epsilon_{iwt} \tag{2}$$

where $Fertilizer_{ist}$ is the N, P and K consumption per hectare at district i from state s in year t; HC_i, stands for High Clay and is a dummy variable that indicates whether the fraction of clayey soil in the district is above the median. Figure A12 in the appendix shows this binary treatment variable spatially. *Post* is a dummy for the post-policy period 2010 and after. The main estimate of interest is β_3, which is an interaction between $HighClay$ and the $Post$ period. μ and η are district and state-year-fixed effects that account for any cross-sectional time-invariant characteristics and any differential trends across states over time. I use state-by-year fixed effects in the main model because there are states that launched individual programs and welfare subsidies during different time points.

4.1.2 Nitrate Pollution in Water Bodies

Next, Similar to the specification in equation 1, I use a similar reduced-form specification to estimate nitrate pollution levels using water quality measures at monitoring stations. However, I use an upstream-downstream specification to estimate pollution effects to address endogeneity concerns. Instead of using clay levels at each water station level, I use clay levels upstream of each monitoring station by using the river flow direction. I compare water monitoring stations with higher clay content in their upstream watersheds to those with lower clay content before and after the policy change. The main identifying assumption in this design is that both groups of monitoring stations would have nitrate-nitrite levels develop along parallel trends in the absence of the policy. The model also includes total precipitation in the upstream watersheds for each water monitoring station as a control to account for the potential influence of rainfall on nitrogen runoff from agricultural fields.

I estimate the following event-study style regression specification:

$$\text{Nitrates}_{ist} = \sum_{k \neq 2009} \beta_k \text{ Upstream Clay}_i \cdot \mathbf{1}\{k = t\} + \phi_d + \gamma_{st} + \epsilon_{dt} \tag{3}$$

My main outcome of interest in this setting is the maximum nitrate-nitrite levels, $Nitrates_{ist}$ in water station i located in state s in year t. Additionally, I also estimate the same model for other water quality indicators such as fecal coliform rates, pH levels, dissolved oxygen content, temperature and conductivity. I omit 2009 as the baseline year, as the policy was launched in 2010. I am interested in the coefficient on the interaction term, β_k to get the changes in nitrate levels after the policy. I should expect no systematic difference between water monitoring stations with high and low levels of clay in their upstream watersheds to be consistent with the identifying assumption of parallel trends on the counterfactuals. If the policy resulted in increased pollution levels, I should expect to see the coefficients increase from 0 immediately after 2009. The differences should sustain over a few years particularly if nitrogen levels were up for a few years.

Similar to the previous specification, here too, my comparison of water monitoring stations with high versus low clay in their upstream watersheds will recover a lower bound of the effects of the policy change. This is because water stations that are classified into low clay groups may still be affected by the policy since their baseline nitrate levels are not zero. To retrieve the residual variation, I include station-level fixed effects to account for time-invariant observable and unobservable factors across water monitoring stations and state-by-year fixed effects to account for all the differential trends across the states over time.

I estimate aggregated versions of equation 3 to estimate average treatment effects. I define a post-NBS policy dummy that is equal to one after 2010.

$$\begin{aligned}\text{Nitrates}_{ist} = \beta_1 \text{Upstream Clay}_i + \beta_2 \text{Post}_t + \beta_3 \text{Upstream Clay}_i \times \text{Post}_t \\ + \beta_4 \text{Upstream Precipitation} + \mu_i + \eta_{st} + \epsilon_{it}\end{aligned} \tag{4}$$

where $Nitrates_{ist}$ measures nitrate pollution, represented by maximum nitrate-nitrite levels at monitoring station i located in state s in year t. *Upstream Clay* is a dummy variable that indicates if the water station has high clay content in its upstream watershed. I also look at other water quality indicators such as maximum fecal coliform, pH levels, dissolved oxygen (DO) and biochemical oxygen demand (BOD). I include station-level and state-by-year fixed effects to account for any differential trends across the states over time. I am interested in the estimate β_3 which is the interaction between *High Clay Upstream* and the *Post* period. μ_i and η_{st} are water station and state-year fixed effects.

4.1.3 Infant Mortality

Reduced Form Specification To estimate infant mortality, I use the exact reduced-form specification in equation 3. I use an upstream-downstream specification to estimate mortality rates at DHS clusters. I first map the closest water stations to the clusters and compute the average nitrates in the upstream watersheds of these stations. I then compare clusters with higher clay content upstream to those with lower clay content before and after the policy change. The key identifying assumption here is that both groups of DHS clusters would have mortality rates develop along parallel trends in the absence of the policy. Following is the event-study-like specification:

$$\text{Infant Mortality}_{ist} = \sum_{k \neq 2009} \beta_k \text{ Upstream Clay}_i \cdot \mathbf{1}\{k = t\} + \phi_d + \gamma_{st} + \epsilon_{dt} \quad (5)$$

My main outcome variable in this setting is the number of infant deaths within one year after birth at DHS cluster i in state s and year t. I omit 2009 as the baseline year. The coefficient of interest is β_k which provides the changes in mortality levels after the policy. For the identifying assumption of parallel trends to hold true, there should be no difference between clusters with high levels of clay in their upstream region and clusters with low levels of clay in their upstream region. The estimates will most likely recover a lower bound since DHS clusters that are classified into low clay groups may still be affected by the policy. I also include DHS cluster-level fixed effects and state-by-year fixed effects.

$$\begin{aligned}\text{Infant Mortality}_{ist} = \beta_1 \text{Upstream Clay}_i + \beta_2 \text{Post}_t + \beta_3 \text{Upstream Clay}_i \times \text{Post}_t \\ + \mu_i + \eta_{st} + \epsilon_{it}\end{aligned} \quad (6)$$

In equation 6, $Infant\ Mortality_{it}$ is the number of infant deaths within one year of birth and 5 years at cluster i from state s in year t. μ refers to DHS cluster fixed-effects and η refers to time fixed effects.

Instrumental Variable Design Estimation of equation 6 yields the reduced-form impact of the NBS policy on infant mortality, but the model does not provide estimates for the effect of nitrates. To address this, I employ an IV design, exploiting upstream clay levels to predict nitrate levels downstream. I use upstream clay levels interacted with post-NBS indicator as an instrument for nitrates to examine the effects of the NBS policy on infant mortality. Higher levels of clay in the upstream areas of the water quality monitoring station increase the risk of increased nitrate runoff through excessive nitrogen application and surface runoff.

The IV design builds on the key assumption of exclusion restriction. Here Upstream Clay interacted Post must affect infant health only through the channel of nitrate pollution after controlling for precipitation, DHS cluster fixed effects and state-year fixed effects. I adopt the upstream-downstream specification described in figure 6 to avoid local effects of soil quality on health. Higher levels of clay soils might affect the agricultural yield of crops, farmer welfare and thereby increased access to nutrition and health investments. To address this concern, I use upstream clay levels as an instrument for downstream nitrate contamination. I also include furthers tests regarding the exclusion restriction.

In the first stage of the IV design, a water monitoring station with high clay content in its upstream area is expected to experience higher levels of nitrogen runoff and, therefore, contain more nitrates. As shown in figure 3, there is substantial variation in clay percentages across the different districts of India, similarly, upstream watersheds of the water monitoring stations also contain variation in the percentages of clay. As expected, I show in table 5 that for stations with high levels of upstream clay, nitrates increase post-policy by 1.3 mg/l.

I adopt the following two-stage least square regressions where regressions 7 and 8 are second and first stage regressions, respectively.

$$\text{Infant Mortality}_{ist} = \beta_0 + \beta_1 \text{Nitrates}_{ist} + \beta_2 \text{Precipitation} + \mu_i + \eta_{st} + \epsilon_{it} \quad (7)$$

$$\text{Nitrates}_{ist} = \beta_0 + \beta_1 (\text{Upstream Clay}_i \times \text{Post}_t) + \beta_2 \text{Precipitation} + \mu_i + \eta_{st} + \epsilon_{it} \quad (8)$$

where $Infant\ Mortality_{ist}$ is the number of infant deaths per cluster i in state s and year t. $Nitrates_{ist}$ is the maximum nitrate-nitrite levels measured in mg/l at station i in state s in year t. I construct a time-variant instrument for the panel data analysis by interacting the time-invariant upstream clay levels for station i with a post-NBS indicator that takes the value 1 after 2010, when the policy was introduced. Monitoring station fixed effects are included to control for any time-invariant characteristics of each monitoring station. I also include state-year fixed effects to account for trends across states over time. Standard errors are clustered at the station level since variation in nitrates are observed at this level.

5 Results

5.1 Summary of nitrogen use after the policy reform

Figure 1 presents estimates of N:P and N:K ratios before and after the NBS policy. I run event-study like regressions with fertilizer ratio on year dummies. Standard errors are clustered at district level and the model includes district fixed effects. The recommended ratio is typically 4:2:1 (National Academy of Agricultural Sciences, 2009) with variations based on underlying soil types. I see an increase in both ratios following the policy implementation, suggesting that farmers may have increased their overall nitrogen use relative to phosphorus and potassium. This shift is counter to the government's stated objectives of reducing budget costs and promoting balanced nitrogen application.

5.2 Effects on Nitrogen Use

In this section, I first show that the 2010 fertilizer subsidy change impacts nitrogen use. I show that districts with clay levels above the median experience an increase in nitrogen consumption following 2010. This serves as an important foundation for the rest of analysis.

Figure 2 shows the increase in prices for phosphorus and potassium fertilizers after 2010. There is a slight increase in nitrogen prices because some non-urea fertilizers also contain nitrogen. Similar to S. Garg and Saxena (2023) I run a regression with log fertilizer consumption on year dummies and including district fixed effects. Figure 2 shows a decline in the consumption of P and K after the policy change.

Fig 7 visualizes the main results from the event study style regressions, focusing on fertilizer use. I show a significant and sustained increase in nitrogen consumption in districts characterized by high clay content following the 2010 policy change. The upward trend in nitrogen use continues at a steady rate until 2016. Notably, I do not find significant changes in the consumption patterns of phosphorus and potassium fertilizers despite the increase of prices of these fertilizers. The stability in the consumption of P and K might appear counterintuitive but most farmers prioritize maintaining established fertilization practices to safeguard yields, despite increased input costs.

Table 3 presents the main findings using the baseline specification in equation 2 conditioned on state-year fixed effects and district-fixed effects. The findings show a significant impact of the policy on nitrogen consumption in districts with high clay content. Dependent variables are listed in the columns. Standard errors are clustered at the district level. The estimate for nitrogen is 15.37, suggesting that districts with high clay content experienced an average increase in nitrogen use of 15.37 kgs/ha compared to pre-policy years. Confidence intervals exclude zero at 95% level for nitrogen use. Given that the mean nitrogen use is 84 kgs, the results translate to a 18% increase in nitrogen use post-policy. Columns 2, 4 and 6 report outcomes with clay as continuous variable instead of categorical high and low clay levels.

5.3 Effects on Nitrate Pollution

Now that I have shown an increase in nitrogen consumption after the policy change, I estimate the impact of this increase in nitrogen use on nitrogen pollution in downstream river stations. I find that stations with high levels of clay in their upstream watersheds experience higher nitrate levels after that policy.

I present the main findings using the baseline specification from equation 4 in table 3. To account for the potential influence of rainfall on nitrogen runoff, I include precipitation controls in the model. Specifically, I control for both total annual precipitation (in mm) and total precipitation during the monsoon months in the upstream watershed region for each water station. The model is conditioned on state-year fixed effects and water station fixed effects. Standard errors are clustered at the station level.

I find that post policy, stations that have high levels of clay in their upstream watershed regions experience an increase in nitrate-nitrite levels by 5.5 mg/l on average. This increase is particularly alarming, considering that the safe limit for nitrate-nitrogen in drinking water is 10 mg/L. In this case, an increase of 5.5 mg/l in nitrates could potentially push many stations close to, or even over the safety threshold for drinking water.

This increase is statistically significant and confidence intervals exclude zero at 95% level. The dependent variable mean is around 3.08 mg/L, meaning that the observed increase in nitrates represents a staggering 80 percent rise in nitrate-nitrite concentrations.

Figure A11 presents results from the event-study version of the model. Pre-trends appear stable, suggesting that there were no significant differences in nitrate levels between stations with high levels and low levels of clay in their upstream watersheds, before the NBS program launch in 2010. Following the policy change, nitrate contamination increases in high-clay areas. This increase persists over time and is consistently positive and nearly significant in most post-policy years. Such a huge increase, persisting over time can lead to severe environmental disruptions and health consequences. Nitrate-nitrite data became available for a larger number of monitoring stations only after 2006, which is just three years before the implementation of the NBS program in 2010. As a result, my analysis has a relatively shorter pre-period compared to the post-period.

5.3.1 Falsification Tests on Water Quality Indicators

To further validate the robustness of my results from the water quality regressions and to ensure that the estimated effects are indeed attributable to the fertilizer policy change, I conduct a series of falsification tests using other water quality indicators that should theoretically be unaffected by the policy and my regression specification. I use these tests to rule out the possibility that the observed changes in nitrate-nitrite levels are due to other unrelated factors affecting water quality in general.

I conduct falsification tests to examine the effects of the policy using other water quality indicators that are unrelated to the policy. I test my framework on five other water quality indicators unrelated to nitrogen pollution: fecal coliform, temperature, dissolved oxygen, pH and conductivity. Using the same design and model specification, I compare these water quality measures before and after the 2010 policy implementation in stations with high levels of clay in their upstream watersheds versus stations with low percentages of clay. I find no statistically significant differences in any of the indicators. Results are reported in table 3. This lack of effect on unrelated water quality measures strengthens the interpretation of my main results, suggesting that the observed increase in nitrate-nitrite levels is indeed a consequence of the fertilizer policy change. I also run even study plots for these indicators to check for pre-trends. Pre-trends are stable and are reported in figure A11. Panel A presents regression results where the entire upstream region for each water station is considered. In Panel B, the upstream region is capped at 100 km to account for pollution decay over distance, serving as an additional robustness check for the results.

Table 4 reports estimates using clay levels downstream of water monitoring stations. Downstream Clay is an indicator variable to denote water stations with clay levels above the median in the area below the water monitoring station. Standard errors are clustered at the water station level.

5.4 Effects on Infant Mortality

The impact of agricultural policies on public health can be complex. Having established a clear link between the 2010 fertilizer policy change and increased nitrate-nitrite levels in water bodies, I now estimate the impacts of this policy on infant health outcomes. I specifically focus on infant mortality rates, defined as deaths within the first year of life. I use these measures for several reasons. First, it is well-established that nitrates affect infants the most, so much so that many countries set the nitrate threshold in drinking water based on the level needed to prevent methenoglobenimia. Second, infant mortality rates are less likely to be confounded by other long-term lifestyle factors or cumulative exposure issues.

Table 6 report the reduced-form and instrumental variable approach findings based on equation 6 and equation 7. All specifications are conditioned on state-year fixed effects and DHS cluster fixed effects with standard errors clustered at the cluster level. The results show a huge difference in how the policy affected urban and rural areas. In rural clusters with high levels of clay in their upstream areas, I find an increase in infant mortality rates post policy. The dependent variable mean is 0.08, this increase represents a substantial relative change in infant mortality. Interestingly, the coefficients for infant mortality in urban areas, while statistically insignificant are negative. Urban areas might not be affected due to better water management and access to alternative sources of drinking water.

5.5 Rainfall Heterogeneity

In this section, I examine if rainfall amplifies the impact of the policy on nitrate levels in rivers. Agricultural runoff, especially from nitrogen-based fertilizers, is known to increase with higher precipitation which facilitates nutrient movement into nearby water bodies (Wang et al., 2023). This effect may be particularly relevant during monsoon months when rainfall is intense and runoff potential is heightened.

To test this relationship between nitrogen runoff and precipitation, in addition to the main difference-in-difference design, I use a triple-difference approach. First, I calculate monthly precipitation data at the upstream watershed level for each water monitoring station and then aggregate the rainfall data annually and during the monsoon months of June, July, August and September.

I then look at the interaction between a dummy variable that indicates post-NBS years and the measure of rainfall upstream from a water monitoring station. In the specification, I categorize rainfall into quartiles across years and monitoring stations, omitting the lowest quartile as the reference. Results, shown in Column 1 in Table 8 indicate that nitrate concentrations are highest in the second quartile of rainfall. Interestingly, the effect diminishes in the higher quartiles, with results that are slightly lower and statistically insignificant. This pattern suggests that while moderate rainfall in the second quartile facilitates fertilizer runoff, heavier rainfall in the upper quartiles may dilute nitrate concentrations in water bodies, reducing detectable pollution levels.

5.6 Robustness Checks

I evaluate and validate the robustness of the main results, I through many sensitivity analyses and alternative specifications.

Alternative Instrumental Variable Approach

As a robustness check, I employ both a reduced-form difference-in-differences design and an instrumental variables approach to estimate the effects of the NBS policy on infant mortality. Ordinary least squares estimates may be biased due to the endogeneity of nitrate levels in river bodies near each cluster.

For the IV design, I use the percentage of clay upstream of each DHS cluster, interacted with a post-NBS policy indicator as the instrument. High clay content increases the likelihood of higher nitrogen use and consequently higher nitrate pollution in downstream regions. As shown in table 6 estimates from both specifications show similar results, however, only the IV estimates are statistically significant.

Clay soil categorization

I explore whether an alternative definition of the soil variable affects my findings. Instead of categorizing districts into high clay and low clay groups, I use the fraction of clay as a continuous variable in the specification accounting for state-year fixed effects and district fixed effects. This allows me to examine how varying levels of clay content affect my main outcomes. Results are robust and are reported in table 3. This addresses potential concerns about arbitrary cutoffs in the binary high and low clay categorization I use in the main analysis.

Upstream watershed boundary

There is no consensus on how far upstream is upstream for pollution mapping. If the distance is larger, then the mapping might not be accurate due to decay in pollution measures. Studies generally deal with this issue by defining various distance ranges and check if the results are still robust. I use the entire watershed identified from the watershed mapper in my main specification. However, I also restrict the upstream area to only 100km radius and rerun the same specifications and find similar results. Pable B of table 3 reports results of nitrate pollution. The table also reports results for other water quality measures and find no impacts.

Infant Mortality - DHS cluster buffers To estimate the effect of upstream nitrate pollution on infant mortality, I create 20km buffers around each infant cluster and calculate the average nitrate levels from water monitoring stations located within these buffers in my main specification. Given that the DHS infant clusters are spatially jittered by approximately 10km for confidentiality, I consider the 20km buffer as a reasonable buffer for my analysis. However, to ensure robustness, I run the same specification using different buffers of 10, 20, 30 and 40km. Results in table 7 show that the estimated effect of nitrate pollution on infant mortality is strongest and statistically significant within the 10 and 20km buffers for rural DHS clusters. Beyond 20km, as I include DHS clusters that are further away from these surface water stations, the effect diminishes and loses significance, suggesting that the proximity to contaminated water sources is key.

Placebo tests using downstream clay To validate the estimates on nitrate pollution, I conduct placebo tests using clay levels from the downstream regions of each water monitoring station. Since downstream clay levels should not influence nitrate levels in upstream water sources, I should not observe any increase in nitrates associated with these downstream soil variables. Consistent with the expectations, results in table 4 show no significant effect of downstream clay levels on nitrate concentrations in upstream stations, supporting the validity of the identification strategy.

6 Conclusion

The need for enhancing agricultural productivity, particularly through increased fertilizer use is a priority for many governments. These policies aim to protect farmers, ensuring they have access to inputs that can improve yields. However, poorly designed subsidy programs can have far-reaching consequences that go beyond their intended goals. In this paper, I examine how India's Nutrient-Based Subsidy program, launched in 2010, skewed fertilizer use towards nitrogen and explore the unintended environmental and public health impacts of this policy. Using detailed datasets on district-level fertilizer use, water quality indicator measures, and geospatial information on child mortality constructed from the demographic and health surveys, I show that the fertilizer subsidy change has negative externalities on river pollution and human health. To address potential endogeneity concerns on nitrogen use, I use the timing of the policy and predetermined exogenous variation in soil physical texture and river flow direction.

Using a difference-in-difference approach, I conduct three key analyses. First, I estimate changes in nitrogen use as a result of the policy. I show that districts with high clay content, where the returns to nitrogen fertilizer are greater, saw a significant increase in their nitrogen use - about 18% on average after the policy. This result highlights how the policy disproportionately incentivized nitrogen use in these regions, exacerbating an already imbalanced fertilizer application system that favored nitrogen over phosphorus and potassium.

Second, I examine whether this increase in nitrogen use results in higher nitrate levels in nearby water bodies. Results from approximately 500 water-quality stations show an increase in nitrate-nitrite levels in water monitoring stations located downstream from watersheds with high clay content. This result suggests that the subsidy reform not only altered fertilizer application rates but also had significant impacts on water pollution. In fact, the nitrate levels nearly doubled in these regions. To rule out other potential causes, I perform falsification tests by looking at water quality indicators that should not be affected by the policy such as fecal coliform levels, temperature, conductivity, and pH. These tests show no significant changes, further supporting my results that the policy-driven increase in nitrogen use is the primary driver behind the rise in nitrate contamination.

Third, I examine the link between these nitrate increases and infant mortality rates. By constructing a pseudo-panel of approximately 10000 clusters from the DHS survey data, I show that rural areas downstream of high-clay regions experienced increased infant mortality rates in the years following the policy change, suggesting that the environmental pollution had tangible public health consequences. Using an instrumental variable approach, I find that a 1.3 mg/l increase in nitrate levels leads to 0.03 additional infant deaths in rural DHS clusters with high upstream clay in the post-policy period.

My results underscore the externalities that can be generated from policies that favor one type of input over others. In this case, the government's decision to maintain low urea prices while reducing the subsidies on P and K disrupted the balance of fertilizer use even further, leading to nitrate pollution and public health costs. Although the government's decision to reduce subsidies for phosphorus and potassium was primarily motivated by the need to cut the overall budget spent on fertilizer subsidies, this shift led to increased nitrogen use. Despite increased nitrogen use, yields do not increase substantially post-policy. Not only did it increase environmental burdens, but it also did not contribute to increased food production, which may have mitigated some of the societal costs. The NBS policy provides a good example of how hard-to-reverse environmental and health costs must be accounted for when evaluating the introduction of policies.

Many governments face political and economic pressure to reduce subsidies, and they often do so in a phased-out manner, as was the case in India. The NBS policy was introduced to reduce the financial burden of fertilizer subsidies, but political concerns prevented any changes to the highly subsidized urea, leading to an imbalance that was difficult to reverse. This case shows the risks of sudden, one-sided policy shifts that do not fully account for their broader social and environmental impacts.

References

Adamopoulos, T., & Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, *104*(6), 1667–1697. https://doi.org/10.1257/aer.104.6.1667

Ansari, A. N., & Sheereen, P. (2022). An Analysis of Fertiliser Subsidies in India. *Saudi Journal of Economics and Finance*, *6*(12), 406–412. https://doi.org/10.36348/sjef.2022.v06i12.001

Bharadwaj, P., Fenske, J., Kala, N., & Mirza, R. A. (2020). The Green revolution and infant mortality in India. *Journal of Health Economics*, *71*, 102314. https://doi.org/10.1016/j.jhealeco.2020.102314

Brainerd, E., & Menon, N. (2014). Seasonal effects of water quality: The hidden costs of the Green Revolution to infant and child health in India. *Journal of Development Economics*, *107*, 49–64. https://doi.org/10.1016/j.jdeveco.2013.11.004

Burke, W. J., Frossard, E., Kabwe, S., & Jayne, T. S. (2019). Understanding fertilizer adoption and effectiveness on maize in zambia. *Food Policy*, *86*, 101721. https://doi.org/https://doi.org/10.1016/j.foodpol.2019.05.004

Cai, H., Chen, Y., & Gong, Q. (2016). Polluting thy neighbor: Unintended consequences of china s pollution reduction mandates. *Journal of Environmental Economics and Management*, *76*, 86–104.

Cassou, E., Tran, D. N., Nguyen, T. H., Dinh, T. X., Nguyen, C. V., Cao, B. T., Jaffee, S., & Ru, J. (2017). An overview of agricultural pollution in vietnam: Summary report 2017.

Chakraborty, P., Chopra, A., & Contractor, L. (2024). The equilibrium impact of agricultural support prices and input subsidies.

Currie, J., Graff Zivin, J., Meckel, K., Neidell, M., & Schlenker, W. (2013). Something in the water: Contaminated drinking water and infant health. *Canadian Journal of Economics/Revue canadienne d'économique*, *46*(3), 791–810.

Currie, J., & Neidell, M. (2005). Air pollution and infant health: What can we learn from california's recent experience? *The Quarterly Journal of Economics*, *120*(3), 1003–1030.

Currie, J., Neidell, M., & Schmieder, J. F. (2009). Air pollution and infant health: Lessons from new jersey. *Journal of health economics*, *28*(3), 688–703.

Damania, R., Desbureaux, S., Rodella, A.-S., Russ, J., & Zaveri, E. (2019). *Quality unknown: The invisible water crisis*. World Bank Publications.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, *109*(12), 4178–4219.

Dias, M., Rocha, R., & Soares, R. R. (2023). Down the River: Glyphosate Use in Agriculture and Birth Outcomes of Surrounding Populations. *Review of Economic Studies*, *90*(6), 2943–2981. https://doi.org/10.1093/restud/rdad011

Do, Q.-T., Joshi, S., & Stolper, S. (2018). Can environmental policy reduce infant mortality? Evidence from the Ganga Pollution Cases. *Journal of Development Economics*, *133*, 306–325. https://doi.org/10.1016/j.jdeveco.2018.03.001

Donovan, K. (2021). The equilibrium impact of agricultural risk on intermediate inputs and aggregate productivity. *The Review of Economic Studies*, *88*(5), 2275–2307.

Ebenstein, A. (2012). The consequences of industrialization: Evidence from water pollution and digestive cancers in china. *Review of Economics and Statistics*, *94*(1), 186–201.

Gamper-Rabindran, S., Khan, S., & Timmins, C. (2010). The impact of piped water provision on infant mortality in brazil: A quantile panel data approach. *Journal of Development Economics*, *92*(2), 188–200.

Garg, S., & Saxena, S. (2023). Distributional effects of indian agricultural interventions [Working Paper].

Garg, T., Hamilton, S. E., Hochard, J. P., Kresch, E. P., & Talbot, J. (2018). (not so) gently down the stream: River pollution and health in indonesia. *Journal of Environmental Economics and Management*, *92*, 35–53.

Gautam, M. (2015). Agricultural subsidies: Resurging interest in a perennial debate. *Indian Journal of Agricultural Economics*, *70*(1), 83–105.

Gollin, D., & Udry, C. (2021). Heterogeneity, measurement error, and misallocation: Evidence from african agriculture. *Journal of Political Economy*, *129*(1), 1–80.

Graff Zivin, J., & Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, *51*(3), 689–730. https://doi.org/10.1257/jel.51.3.689

Greenstone, M., & Hanna, R. (2014). Environmental Regulations, Air and Water Pollution, and Infant Mortality in India. *American Economic Review*, *104*(10), 3038–3072. https://doi.org/10.1257/aer.104.10.3038

Gulati, A., & Banerjee, P. (2015). *Rationalising fertiliser subsidy in india: Key issues and policy options* (tech. rep.). Working paper.

Hagerty, N., & Tiwari, A. (2022). Industrial water pollution and agricultural production in india. *Working Paper*.

Heberger, M. (2022). Global Watersheds [Web Application].

Holden, S. T. (2019). Economics of farm input subsidies in africa. *Annual Review of Resource Economics*, *11*(1), 501–522.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, *124*(4), 1403–1448.

Huang, J., Gulati, A., & Gregory, I. (2017). Fertilizer subsidies-which way forward?

Itin-Shwartz, B. (2024). Survey based assessment of sustainable agricultural practices: Evidence from indian plots. *Agribusiness*, *40*(2), 416–457.

James, C., Fisher, J., Russell, V., Collings, S., & Moss, B. (2005). Nitrate availability and hydrophyte species richness in shallow lakes. *Freshwater biology*, *50*(6).

Jones, B. A. (2019). Infant health impacts of freshwater algal blooms: Evidence from an invasive species natural experiment. *Journal of Environmental Economics and Management*, *96*, 36–59. https://doi.org/https://doi.org/10.1016/j.jeem.2019.05.002

Keiser, D. A., & Shapiro, J. S. (2018). Consequences of the Clean Water Act and the Demand for Water Quality. *Quarterly Journal of Economics*, *134*(1), 349–396.

Kishore, A., Alvi, M., & Krupnik, T. J. (2021). Development of balanced nutrient management innovations in South Asia: Perspectives from Bangladesh, India, Nepal, and Sri Lanka. *Global Food Security*, *28*, 100464. https://doi.org/10.1016/j.gfs.2020.100464

Knittel, C. R., Miller, D. L., & Sanders, N. J. (2016). Caution, drivers! children present: Traffic, pollution, and infant health. *Review of Economics and Statistics*, *98*(2), 350–366.

Knobeloch, L., Salna, B., Hogan, A., Postle, J., & Anderson, H. (2000). Blue babies and nitrate-contaminated well water. *Environmental health perspectives*, *108*(7), 675–678.

Kurdi, S., Mahmoud, M., Abay, K. A., & Breisinger, C. (2020). *Too much of a good thing? evidence that fertilizer subsidies lead to overapplication in egypt* (Vol. 27). Intl Food Policy Res Inst.

Lai, W. (2017). Pesticide use and health outcomes: Evidence from agricultural water pollution in china. *Journal of environmental economics and management*, *86*, 93–120.

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., & Garnier, J. (2014). 50 year trends in nitrogen use efficiency of world cropping systems: The relationship between yield and nitrogen input to cropland. *Environmental Research Letters*, *9*(10), 105011.

Li, D., Zhai, Y., Lei, Y., Li, J., Teng, Y., Lu, H., Xia, X., Yue, W., & Yang, J. (2021). Spatiotemporal evolution of groundwater nitrate nitrogen levels and potential human health risks in the songnen plain, northeast china. *Ecotoxicology and Environmental Safety*, *208*, 111524.

Liu, C.-W., Sung, Y., Chen, B.-C., & Lai, H.-Y. (2014). Effects of nitrogen fertilizers on the growth and nitrate content of lettuce (lactuca sativa l.) *International journal of environmental research and public health*, *11*(4), 4427–4440.

Mateo-Sagasta, J., Zadeh, S. M., Turral, H., & Burke, J. (2017). *Water pollution from agriculture: A global review. executive summary*. Rome, Italy: FAO Colombo, Sri Lanka: International Water Management.

Motohashi, K. (2024). Unintended consequences of sanitation investment: Negative externalities on water quality and health in india. *Working Paper*.

Mueller, N. D., Lassaletta, L., Runck, B. C., Billen, G., Garnier, J., & Gerber, J. S. (2017). Declining spatial efficiency of global cropland nitrogen allocation. *Global Biogeochemical Cycles*, *31*(2), 245–257.

National Academy of Agricultural Sciences. (2009). *Crop response and nutrient ratio* (Policy Paper No. 42). National Academy of Agricultural Sciences. New Delhi.

Nixon, S. W. (1995). Coastal marine eutrophication: A definition, social causes, and future concerns. *Ophelia*, *41*(1), 199–219. https://doi.org/10.1080/00785236.1995.10422044

Ramaswami, B. (2019). *Agricultural subsidies* (tech. rep.) (Study Prepared for the XV Finance Commission, New Delhi).

Ravinutala, S. (2016). *Redesigning India's Urea Policy* [MPA/ID Thesis]. John F. Kennedy School of Government, Harvard University.

Restuccia, D., Yang, D. T., & Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, *55*(2), 234–250.

Schlenker, W., & Walker, W. R. (2016). Airports, air pollution, and contemporaneous health. *The Review of Economic Studies*, *83*(2), 768–809.

Skidmore, M. E., Sims, K. M., & Gibbs, H. K. (2023). Agricultural intensification and childhood cancer in brazil. *Proceedings of the National Academy of Sciences*, *120*(45), e2306003120. https://doi.org/10.1073/pnas.2306003120

Taylor, C. A. (2021). Cicadian rhythm: Insecticides, infant health and long-term outcomes. *Center for Environmental Economics and Policy Working Paper*, (9), 1–45.

Taylor, C. A., & Heal, G. (2023). Fertilizer and algal blooms: A satellite approach to assessing water quality. In P. Antràs & D. Zilberman (Eds.), *Risks in agricultural supply chains* (pp. 83–106). University of Chicago Press.

University of Wisconsin-Madison Division of Extension. (n.d.). The important role of soil texture on water [Accessed on November 10, 2024]. *CropsSoils*.

von Der Goltz, J., Dar, A., Fishman, R., Mueller, N. D., Barnwal, P., & McCord, G. C. (2020). Health impacts of the green revolution: Evidence from 600,000 births across the developing world. *Journal of health economics*, *74*, 102373.

Wang, R., Zhang, J., Cai, C., & Wang, S. (2023). Mechanism of nitrogen loss driven by soil and water erosion in water source areas. *Journal of Forestry Research*, *34*(6), 1985–1995.

WHO. (2004). *Global health observatory data repository technical report* (tech. rep.). World Health Organization. http://refhub.elsevier.com/S0304-3878(18)30205-0/sref53

Wuepper, D., Le Clech, S., Zilberman, D., Mueller, N., & Finger, R. (2020). Countries influence the trade-off between crop yields and nitrogen pollution. *Nature Food*, *1*(11), 713–719.

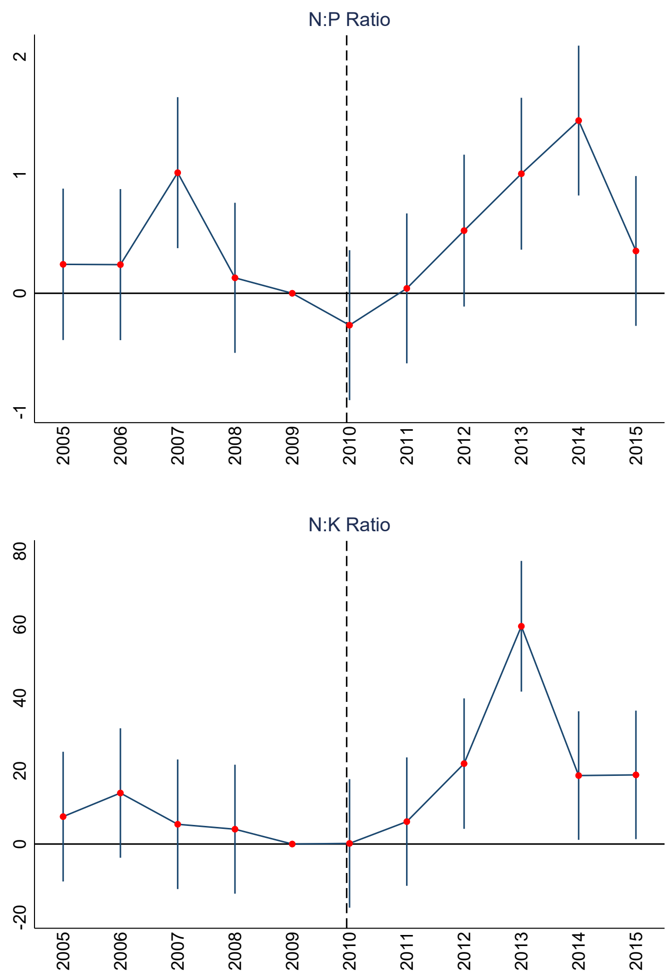
Zaveri, E. D., Russ, J. D., Desbureaux, S. G., Damania, R., Rodella, A.-S., & Ribeiro Paiva De Souza, G. (2020). The nitrogen legacy: The long-term effects of water pollution on human capital [Available at SSRN: https://ssrn.com/abstract=3533562]. *World Bank Policy Research Working Paper*, (No. 9143).

Zhang, B.-l., Cui, B.-h., Zhang, S.-m., Wu, Q.-y., & Yao, L. (2018). Source apportionment of nitrogen and phosphorus from non-point source pollution in nansi lake basin, china. *Environmental Science and Pollution Research*, *25*, 19101–19113.

Zhu, X., Ros, G. H., Xu, M., Cai, Z., Sun, N., Duan, Y., & de Vries, W. (2023). Long-term impacts of mineral and organic fertilizer inputs on nitrogen use efficiency for different cropping systems and site conditions in southern china. *European Journal of Agronomy*, *146*, 126797. https://doi.org/https://doi.org/10.1016/j.eja.2023.126797

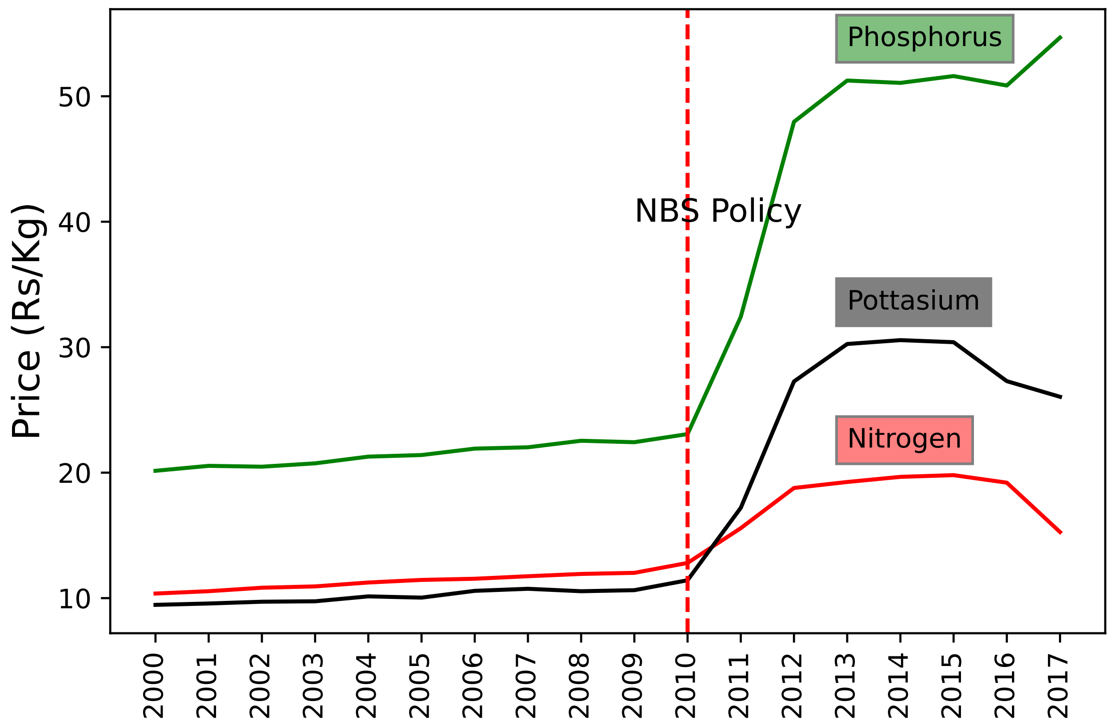
7 Figures

Figure 1. Nitrogen-Phosphorus and Nitrogen-Potassium ratios pre and post NBS policy



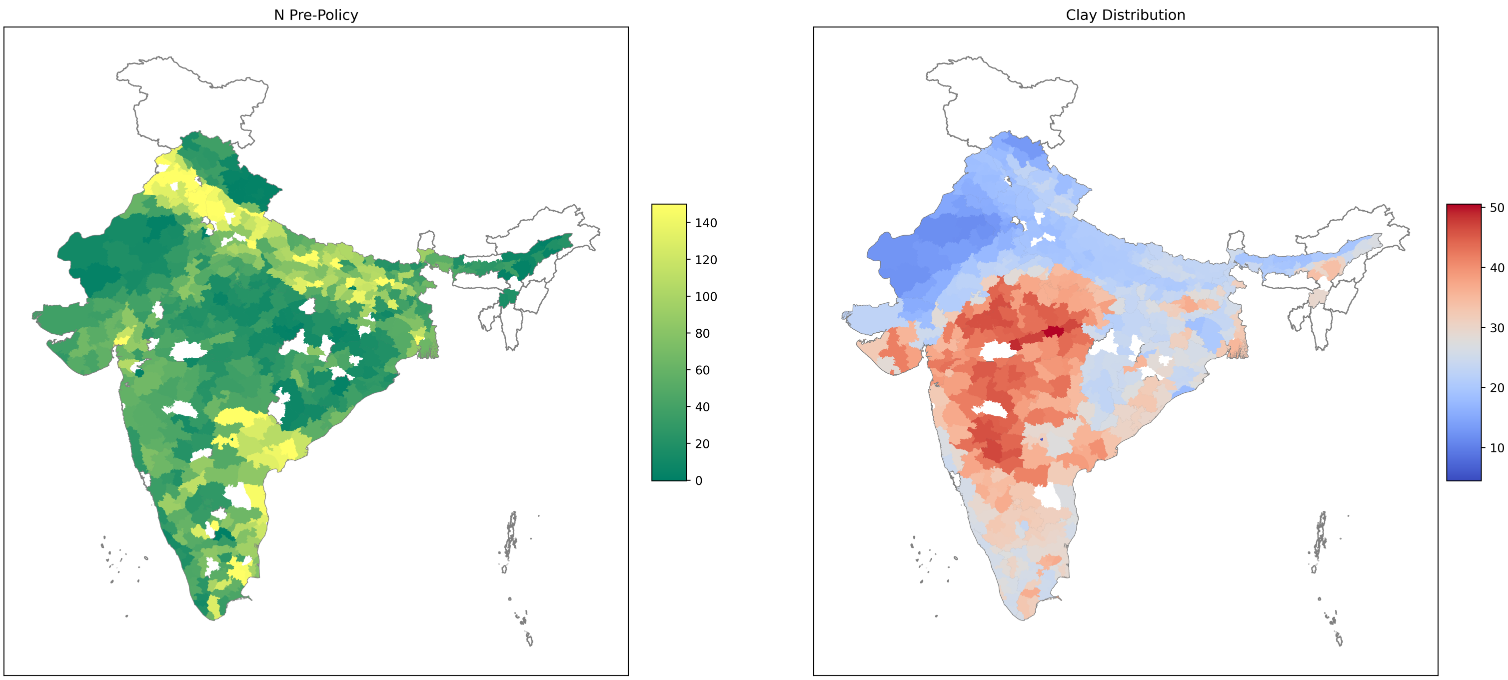
Notes: The figures shows N:P ratios and N:K ratios before and after the policy using district level dataset. Y axis plots coefficients from an event-study like regression where N:P and N:K are the outcome variables regressed on year dummies. The specification includes clustered standard errors and district fixed effects.

Figure 2. Fertilizer Prices reported through Cost of Cultivation Surveys



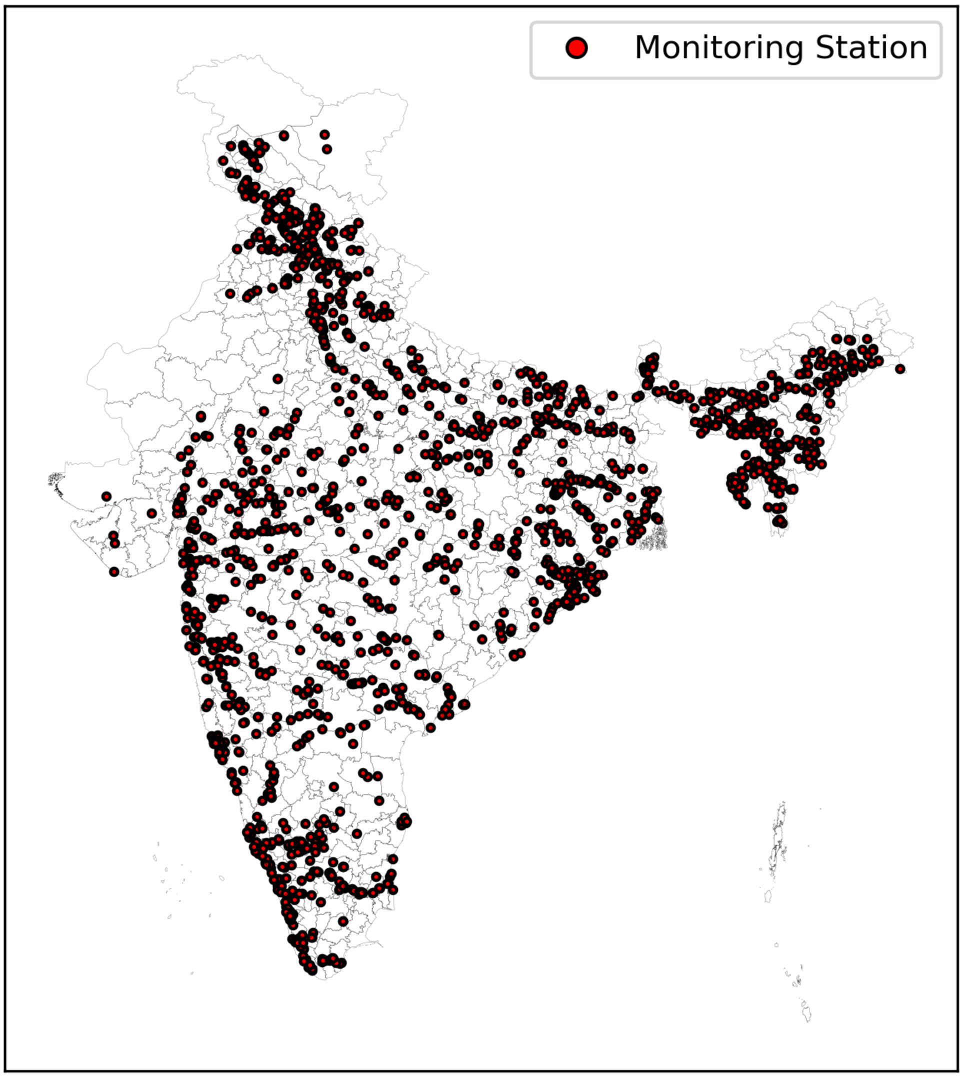
Notes: This figure shows farmer-reported fertilizer prices for the three main nutrients N, P and K. Data is retrieved from the Cost of Cultivation Surveys

Figure 3. Nitrogen Use and Clay Type by Districts



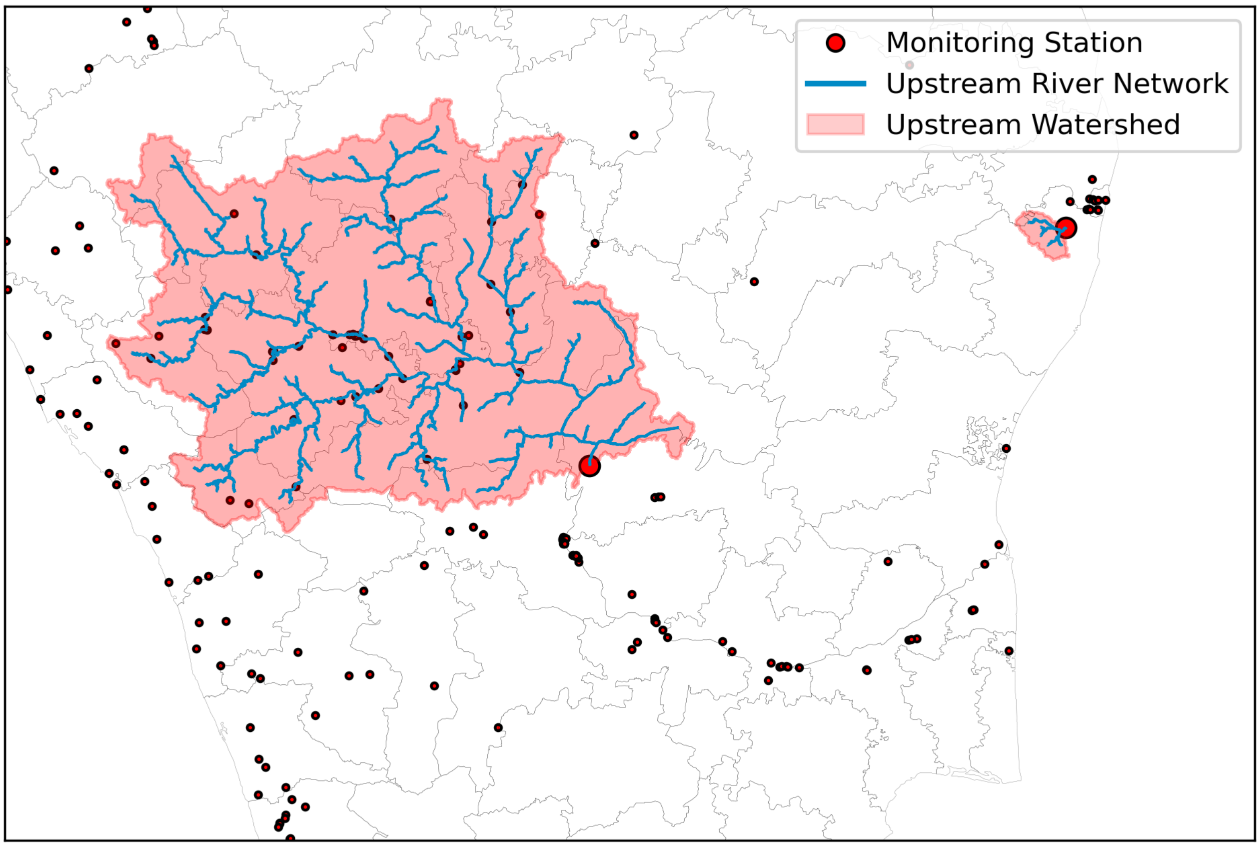
Notes: The left panel displays nitrogen consumption before the policy change in 2010. The right panel shows the percentage of clay soils on average for each district. Nitrogen data comes from ICRISAT and soil data comes from the FAO soil raster database.

Figure 4. Water Monitoring Stations



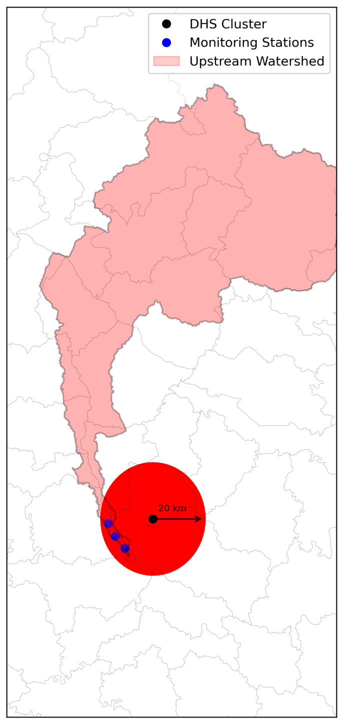
Notes: This map plots all the river monitoring stations in India. Data is from India's Central Pollution Control Board.

Figure 5. Illustration of Upstream Soil Mapping for Water Monitoring Stations



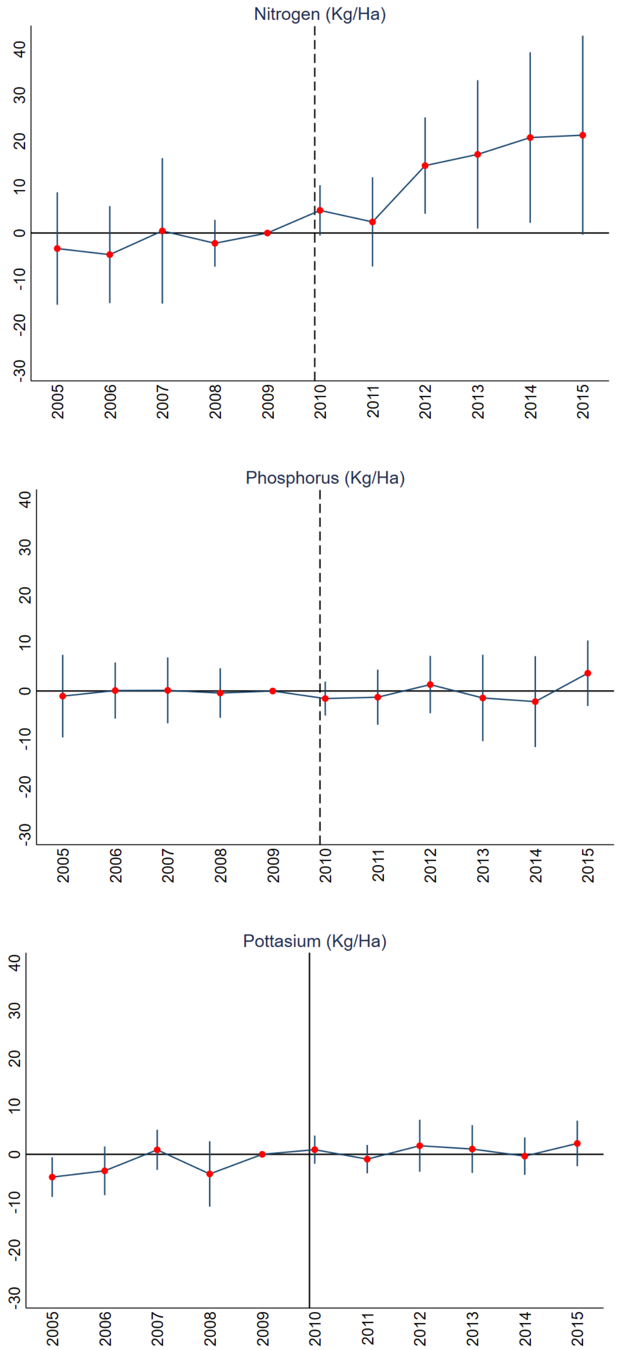
Notes: This figure illustrates the upstream-downstream framework used in the main specification to analyze the effect of upstream soil characteristics on water quality in downstream monitoring stations. I highlight two monitoring stations and their respective upstream river network and watershed area retrieved using the global watershed mapping API. This map displays district borders in grey lines and water quality monitoring station locations in red.

Figure 6. Illustration of Upstream Soil Mapping for DHS Clusters



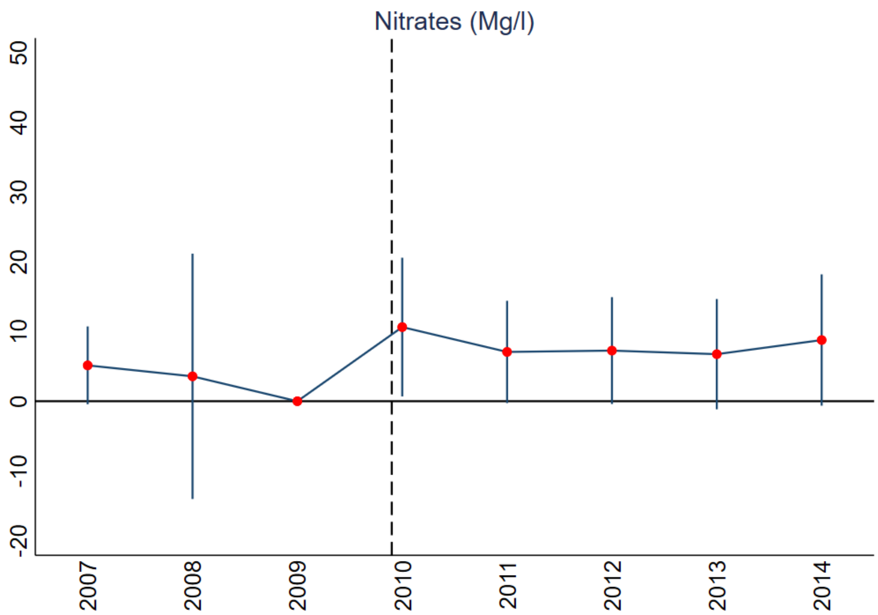
Notes: This figure illustrates the upstream analysis, which analyses the effect of upstream soil characteristics on infant mortality in a downstream DHS cluster. I highlight a dhs cluster and take a buffer of 20km. I then take all the water monitoring stations within this buffer and calculate the average clay levels of the upstream watershed of each of these stations.

Figure 7. Event Study Plots of N, P and K use in High vs. Low Clay Districts



Notes: The figure shows the regression coefficients of fertilizer use for N, P and K on the interaction terms between upstream high and low clay levels and year dummies. The 95% confidence intervals are shown with dashed lines. Standard errors are clustered at the district level. The model includes district fixed effects and state-year fixed effects.

Figure 8. Maximum Nitrate-Nitrite Levels in Water Monitoring Stations in high vs low clay districts



Notes: The figure shows the regression coefficients of nitrate-nitrite levels in mg/l on the interaction terms between upstream clay levels and year dummies. The 95% confidence intervals are shown as the shaded region. Standard errors are clustered at the district level. The model includes water station fixed effects and state-year fixed effects.

8 Tables

Table 1. Yearly Fertilizer Consumption and Share

Year	Nitrogen		Phosphate		Potassium	
	Tons	%	Tons	%	Tons	%
2005	40541.51	61.18	16580.83	26.69	7696.33	11.81
2006	43927.33	62.56	17670.44	26.45	7437.82	10.66
2007	45898.67	62.81	17585.36	25.36	8384.78	11.51
2008	48161.48	59.81	20775.95	26.84	10578.99	13.03
2009	49899.38	58.58	23322.29	28.15	11631.98	13.27
2010	52996.01	58.57	25747.32	28.88	11246.96	12.54
2011	54815.70	61.90	25370.15	28.30	8243.23	9.80
2012	53821.25	65.25	21136.08	26.10	6458.92	8.66
2013	53633.78	67.74	17986.17	23.19	6612.28	9.07
2014	54315.87	65.51	19526.48	23.95	8085.15	10.22
2015	55642.46	64.63	22376.63	25.79	7541.08	9.26
2016	53589.92	64.04	21463.46	25.85	7990.75	9.79
2017	52838.46	63.75	21352.56	25.96	8589.55	10.29

Notes: Table reports annual fertilizer consumption for the major nutrients N, P and K from 2005 to 2017 using district-level fertilizer consumption data from ICRISAT.

Table 2. Impact of NBS Policy on Fertilizer Use

	Nitrogen		Phosphorus		Potassium	
	(1)	(2)	(3)	(4)	(5)	(6)
High Clay x Post	15.37**		-0.0241		3.031	
	(5.949)		(3.259)		(2.204)	
Clay x Post		0.881***		0.136		0.133
		(0.268)		(0.141)		(0.115)
Observations	4162	4162	4162	4162	4162	4162
Dep var mean	84.48	84.48	33.99	33.99	16.90	16.90
District FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: Table reports estimates of the Nutrient-Based Subsidy Reform on Fertilizer use in districts with high vs low levels of clay soil. High Clay districts refer to districts with clay levels above the median and is a dummy variable with the value 1 for high clay districts and 0 otherwise. Standard errors are clustered at the district level. All regressions include district fixed effects and state-year fixed effects. Columns 1, 3 and 5 of the table present estimates using the binary treatment variable while columns 2,4 and 6 present results with clay as continuous variable instead of categorical high and low clay levels.

Table 3. Impact of NBS Policy on Water Quality Measures

	(1) Nitrates	(2) Fecal Coli.	(3) Temperature	(4) DO	(5) pH	(6) Conductivity
Panel A: Upstream Clay						
Upstream Clay x Post	5.574**	-473018.3	0.0956	0.183	-0.157	163.1
	(2.220)	(412726.4)	(0.451)	(0.221)	(0.125)	(254.6)
Panel B: Upstream Clay upto 100 km						
Upstream Clay x Post	2.571*	-1537558.6	0.487	0.326	-0.0648	-487.1
	(1.346)	(1469486.4)	(0.479)	(0.273)	(0.0827)	(707.5)
Observations	2739	3615	4111	4128	4158	4080
Dep var mean	3.085	2568940.3	28.40	8.135	18.05	1289.1
Station FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: Table reports estimates of the Nutrient-Based Subsidy Reform on water quality measures from water monitoring stations. Sample consists of water quality measurements recorded at stations all over India from 2007 to 2014. Upstream Clay refer to water stations with clay levels above the median in their upstream regions and is a dummy variable with the value 1 for high clay upstream and 0 otherwise. Standard errors are clustered at the water station level. Column 1 reports coefficients for nitrate pollution, which is recorded as maximum nitrate-nitrite levels in (mg per L. Columns 2 - 6 include other water quality indicators such as maximum fecal coliform, maximum temperature, dissolved oxygen, pH and maximum conductivity. Panel A presents regression results where the entire upstream region for each water station is considered. In Panel B, the upstream region is capped at 100 km to account for pollution decay over distance, serving as a robustness check for results.

Table 4. Placebo Test: Impact of NBS Policy on Water Quality Measures using Downstream Soil

	(1)	(2)	(3)	(4)	(5)	(6)
	Nitrates	Fecal Coli.	Temperature	DO	pH	Conductivity
Downstream Clay x Post	1.161	-12171.4	-0.141	0.292	235.4	-242.9
	(1.704)	(249872.5)	(0.593)	(0.206)	(171.2)	(549.0)
Observations	2739	3615	4111	4128	4158	4080
Dep var mean	3.085	2568940.3	28.40	8.135	18.05	1289.1
Station FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Table reports placebo estimates using downstream soil properties. Sample consists of water quality measurements recorded at stations all over India from 2007 to 2014. Downstream Clay refer to water stations with clay levels above the median in the region below the water monitoring station and is a dummy variable with the value 1 for high downstream stations and 0 otherwise. Standard errors are clustered at the water station level. Column 1 reports coefficients for nitrate pollution, which is recorded as maximum nitrate-nitrite levels in (mg/l). Columns 2 - 6 include other water quality indicators such as maximum fecal coliform, maximum temperature, dissolved oxygen, pH and maximum conductivity.

Table 5. First-stage Regressions

	Full Sample	Rural	Urban
	(1)	(2)	(3)
Upstream Clay x Post[=1]	1.341***	1.607***	1.150***
	(0.149)	(0.247)	(0.233)
Constant	2.571***	1.994***	3.309***
	(0.0694)	(0.111)	(0.114)
Observations	31246	18318	12918
Dep var mean	0.0716	0.0897	0.0458
No. of clusters	6962	4298	2663
KP F-Stat	80.68	42.32	24.28
Cluster FE	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: Table reports first stage estimates. Sample consists of DHS clusters from 2007 to 2014. The dependent variable is Upstream Clay interacted with the post-NBS indicator of a reference DHS cluster. Upstream Clay refer to DHS clusters with clay levels above the median in their upstream watershed regions and is a dummy variable with the value 1 for high clay upstream and 0 otherwise. Standard errors are clustered at the DHS cluster level . The KP F-Stat is the Wald version of the Kleibergen and Paap (2006) statistic on the excluded instrumental variables. Column 1 includes the full sample of DHS clusters located within 20kms of a water monitoring station. Column 2 and 3 only uses the rural and urban clusters respectively.

Table 6. Impact of NBS Policy on Infant Mortality

	Reduced Form			IV		
	Full Sample	Rural	Urban	Full Sample	Rural	Urban
	(1)	(2)	(3)	(4)	(5)	(6)
Upstream Clay x Post	0.0106	0.0165	-0.00155			
	(0.0112)	(0.0142)	(0.0181)			
Nitrates				0.0329*	0.0516**	-0.0234
				(0.0191)	(0.0236)	(0.0314)
Observations	51198	31002	20188	31246	18318	12918
Dep var mean	0.0745	0.0926	0.0467	0.0716	0.0897	0.0458
KP F-Stat				80.68	42.32	24.28
Cluster FE	Yes	Yes	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: Table reports coefficient estimates from reduced form specifications and instrumental variables. Standard errors are clustered at the DHS cluster level. The first three columns report estimates for the full sample, rural and urban samples of DHS clusters under the reduced form specification. Columns 4, 5 and 6 show results from the IV specification for the full sample, rural and urban samples. The sample is limited to clusters that have at least one monitoring station within a 20km buffer. The KP F-Stat is the Wald version of the Kleibergen and Paap (2006) statistic.

Table 7. Impact of NBS Policy on Infant Mortality (Rural)

	(1) 10 km	(2) 20 km	(3) 30 km	(4) 40 km
Panel A: IV				
Nitrates	0.0920*	0.0516**	0.0159	0.0307
	(0.0494)	(0.0236)	(0.0212)	(0.0336)
Observations	6795	18318	30524	41694
Dep var mean	0.0833	0.0897	0.0886	0.0906
No. of clusters	1601	4298	6957	9046
KP F-Stat	21.60	42.32	48.96	11.18
Panel B: Reduced Form				
Upstream Clay x Post	0.0308	0.0165	-0.00426	-0.00737
	(0.0207)	(0.0142)	(0.0114)	(0.00986)
Observations	11673	31002	50973	67240
No. of clusters	1949	4963	7872	10043
Dep var mean	0.0834	0.0926	0.0942	0.0967
Cluster FE	Yes	Yes	Yes	Yes
State x Year FE	Yes	Yes	Yes	Yes

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

Notes: Table presents estimated coefficients for the effect of upstream nitrate pollution on infant mortality across different buffer distances around each DHS rural cluster. Standard errors are clustered at the cluster level. Column 1 consists of clusters that have at least one water station within 10 km. Column 2 extends the buffer to 20 km. Columns 3 and 4 extend the buffer to 30 and 40 km.

Table 8. Reduced-form heterogeneity results for Nitrate Pollution

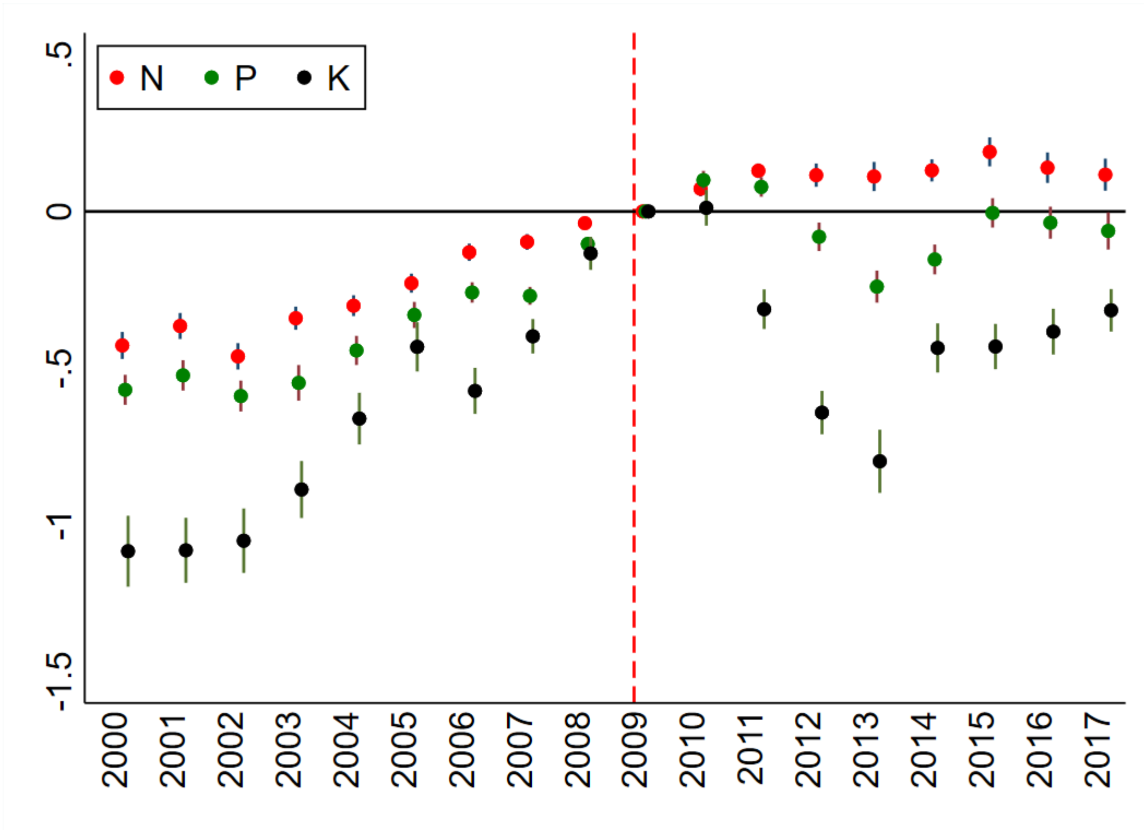
	(1)	(2)	(3)
Upstream Clay x Post	5.574**	0.436	1.257
	(2.220)	(2.906)	(2.495)
Upstream Clay x Post x Rain quartile 2		5.692**	5.047**
		(2.797)	(2.474)
Upstream Clay x Post x Rain quartile 3		3.525	2.722
		(2.236)	(1.817)
Upstream Clay x Post x Rain quartile 4		4.605*	3.208
		(2.702)	(2.245)
Observations	2736	2736	2736
Dep var mean	3.087	3.087	3.087
R square	0.472	0.476	0.477

Standard errors in parentheses

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

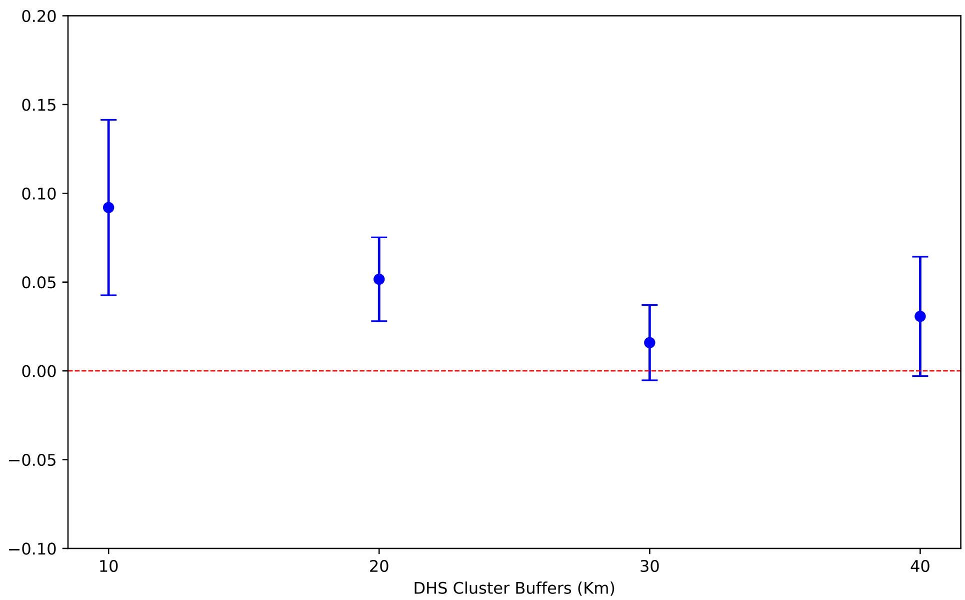
Notes: Table reports estimates from heterogeneity analysis. Standard errors are clustered at the water station level. Columns 1 reports results without rainfall interactions. Column 2 reports results with total rainfall. Column 3 reports results with monsoon rainfall. The lowest quartile is the baseline period and is dropped from the regression.

Figure A9. Log Consumption of Fertilizers at District Level



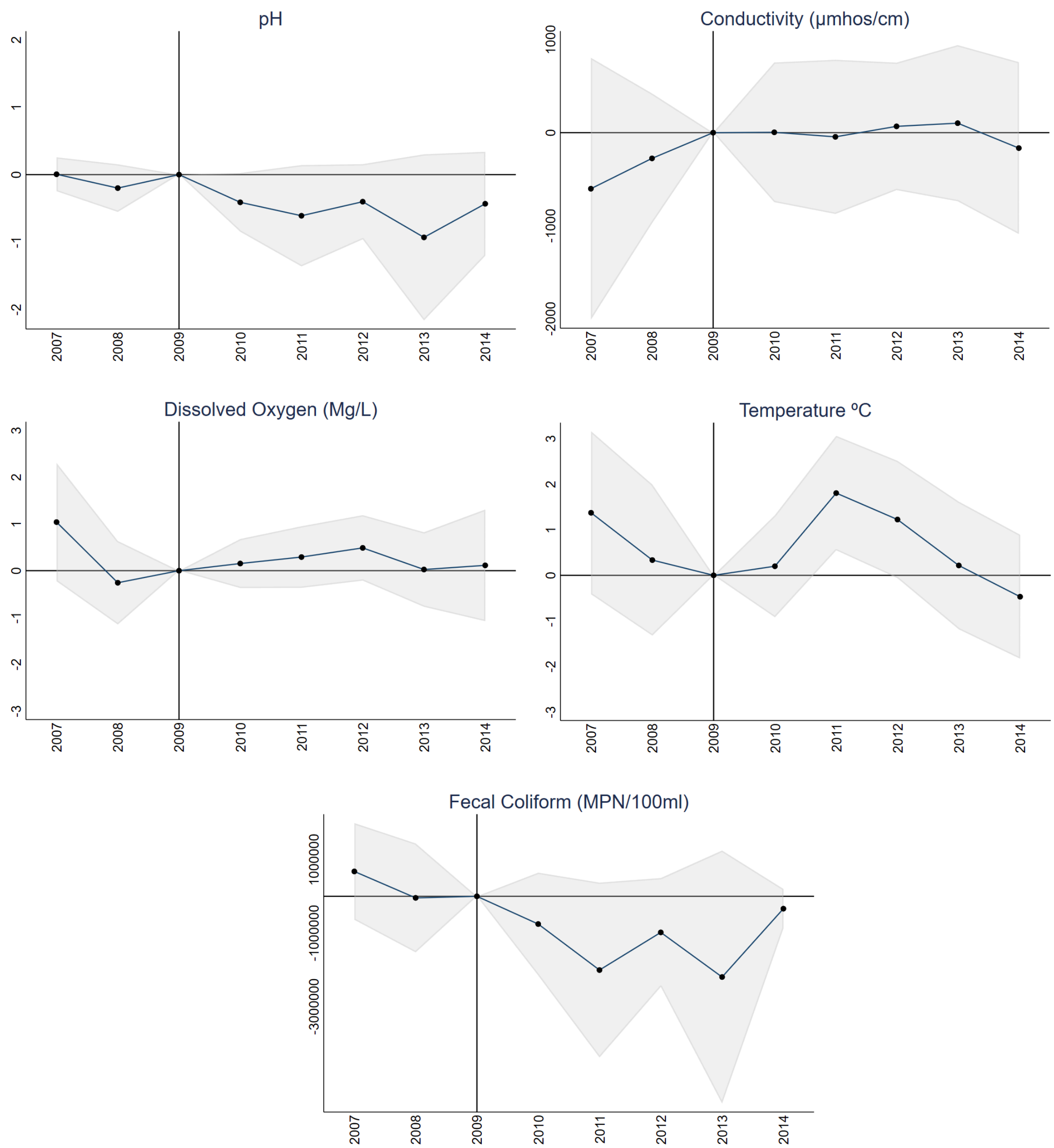
Notes: This figure shows the estimated coefficients from an event-study style regression with log fertilizer consumption of N, P and K at the district level as the outcome variable. Log fertilizer consumption is regressed on year dummies with district fixed effects. Time period 2009 (t = -1) is omitted.

Figure A10. Infant Mortality Estimates



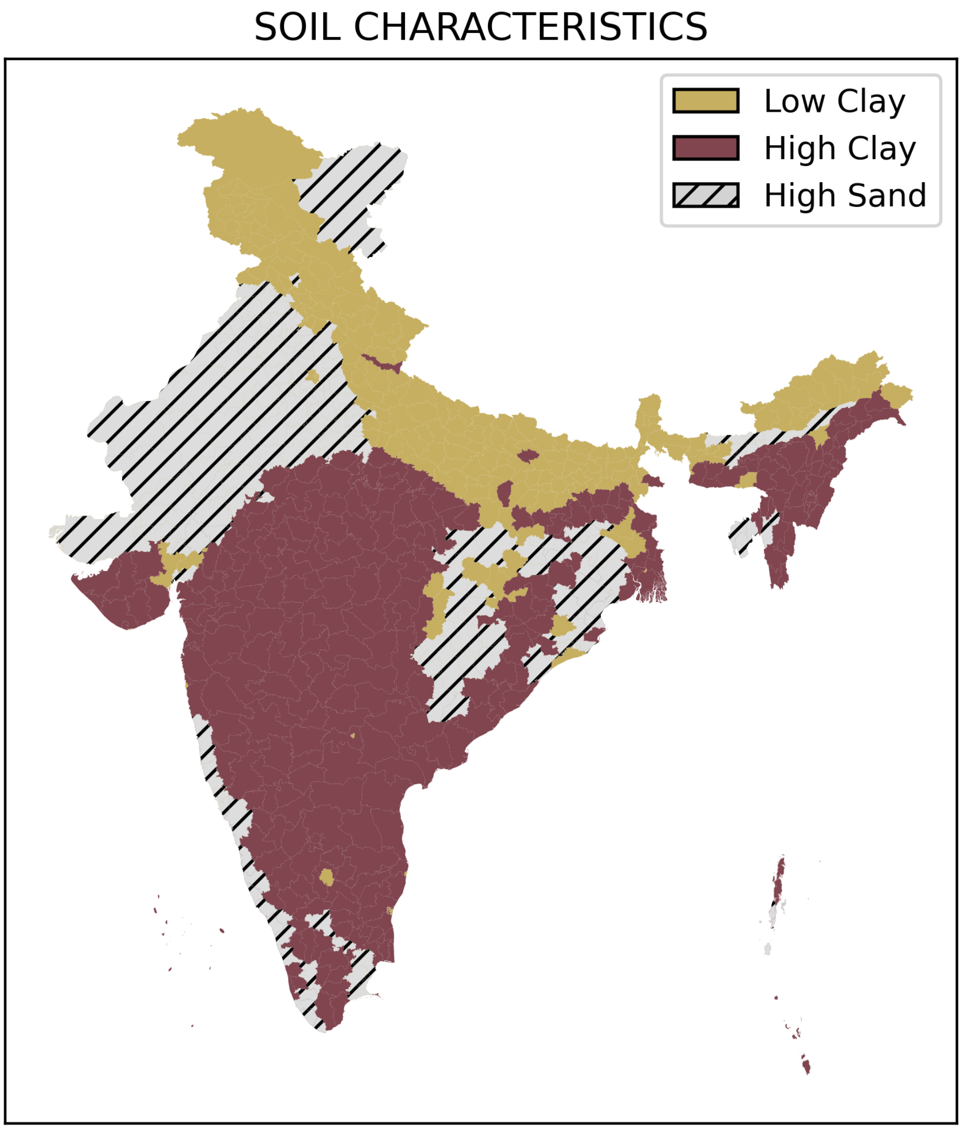
Notes: The figure reports coefficients for infant mortality in rural DHS clusters from the IV regression with different buffers.

Figure A11. Water Quality Indicators in High vs Low Clay Districts



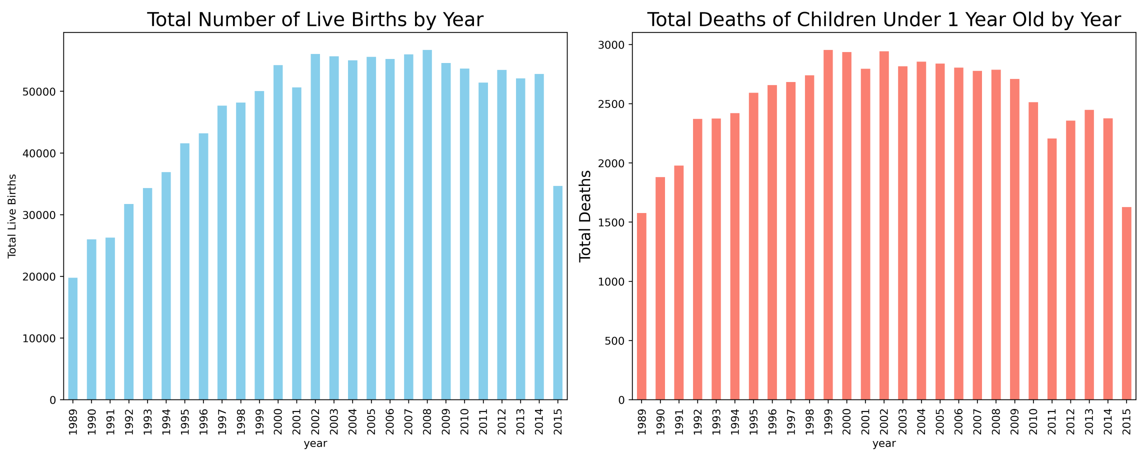
Notes: The figure shows the regression coefficients of various water quality indicators from the station-level data. The coefficients plotted are the interaction terms between upstream clay levels and year dummies. The 95% confidence intervals are shown as the shaded region. Standard errors are clustered at the district level. The model includes water station fixed effects and state-year fixed effects.

Figure A12. Treatment Definition



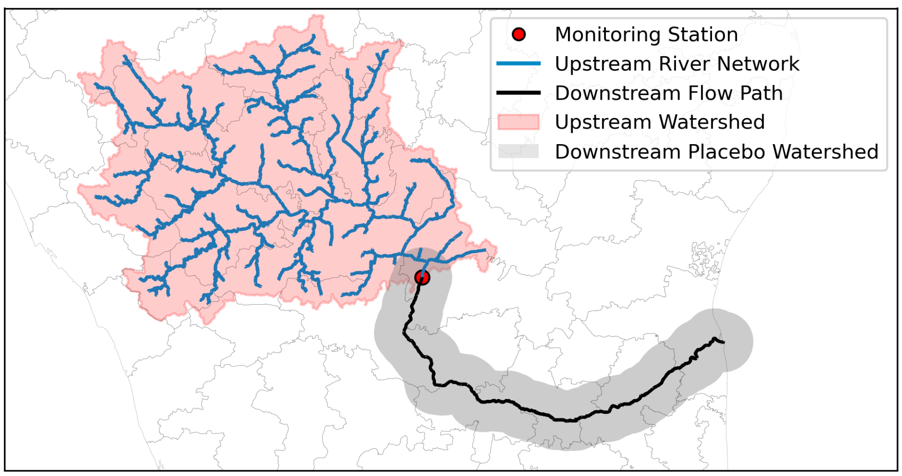
Notes: This figure shows the binary treatment at the district level in the difference in difference and event study regression designs. I classify districts into high clay or low clay based on the median clay content. The main model also does not include districts with high levels of sandy soils.

Figure A13. Total infant births and deaths by year



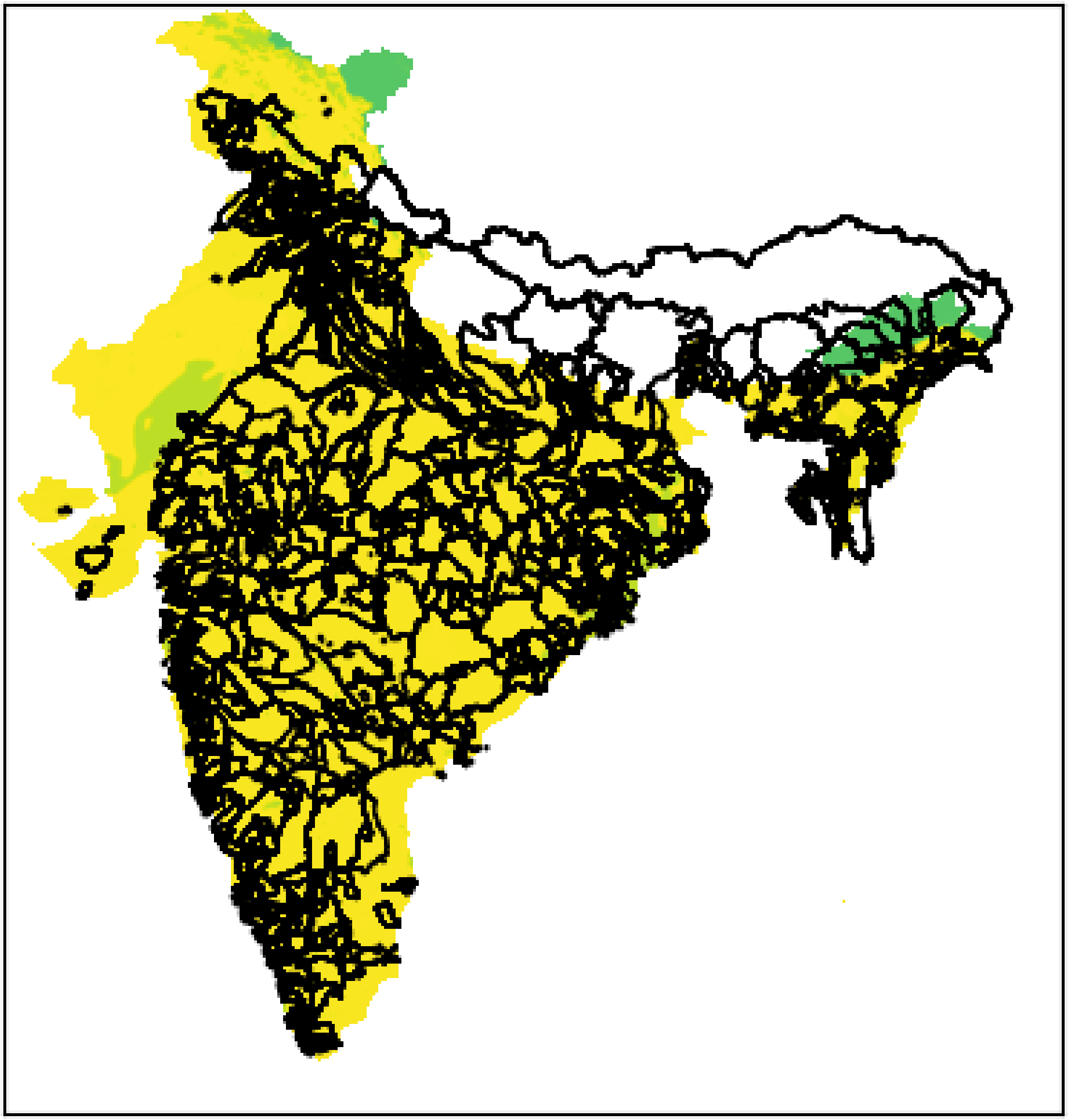
Notes: This figure shows the total number of live births and infant deaths within one year of birth as reported by mothers in DHS clusters.

Figure A14. Downstream flow path



Notes: This figure shows the downstream flow path for the highlighted water monitoring station.

Figure A15. Upstream polygons



Notes: This figure shows the upstream polygons for all the water monitoring stations in the data.