Unintended Environment and Health Consequences of Distortionary Fertilizer Subsidies

Gowthami Venkateswaran[*] Andrew Hultgren[*]
Hemant Pullabhotla[†] Kathy Baylis[‡]

April 2025

Abstract

Governments worldwide subsidize agricultural inputs to support farmers and increase food production. While subsidies can encourage technology adoption, they may drive the overuse of some chemical inputs, exacerbating associated environmental and health impacts. This paper examines the unintended environmental and health consequences of increased fertilizer use driven by selective subsidy reforms in India. 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, we find significant effects of the subsidy on nitrogen pollution in downstream water bodies and infant mortality in downstream rural areas. For every 1% percent increase in nitrate levels, we find a 1.6% increase in rural infant mortality rates. These findings have broad implications for India's politically sensitive nitrogen fertilizer subsidies.

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

[*]University of Illinois Urbana-Champaign. gv4@illinois.edu
[†]Deakin University
[‡]University of California Santa Barbara

1 Introduction

Pollution exposure is the largest environmental risk factor in premature deaths worldwide (Fuller et al., 2022). Causal estimation of the effects of pollution on mortality has largely focused on developed economies, where regulatory enforcement has been more stringent and regulations often require economic (Chay & Greenstone, 2003; Currie & Neidell, 2005; Schlenker & Walker, 2016; Keiser & Shapiro, 2018; Keiser et al., 2023).[1] However, most of the world's population lives in developing economies, where marginal effects of water pollution may differ due to differences in information constraints, behavioral norms, or higher baseline pollution levels (Jalan & Somanathan, 2008; Arceo et al., 2016; Hsiang et al., 2019). Importantly, enforcement of water pollution regulations in the developing world has been limited, and associated mortality effects of water pollution are not well known (Greenstone & Hanna, 2014; Greenstone & Jack, 2015; Olmstead & Zheng, 2021). Here, we estimate the causal effect of nitrate water pollution on infant mortality in India, the world's most populous country.[2]

India presents a compelling setting in which to study water pollution and health for several reasons. First, over half of India's 1.2 billion people live in rural areas where reliance on river water for bathing and drinking is common and water quality regulations are inconsistently enforced. Second, the Indian government has heavily subsidized fertilizers since the Green Revolution, leading to overuse and runoff to rivers. Fertilizer subsidies represent 58% of all federal agricultural support and fully 0.4% of India's GDP (Ramaswami, 2019), leading India to be the second-highest consumer of fertilizers globally, after China. Third, India is one of the few developing countries that has detailed data on measures of water pollution, fertilizer use and infant mortality, thereby allowing us to measure treatment and outcomes at high resolution.

We use a large relative price decrease for nitrogen fertilizers under India's 2010 fertilizer subsidy reform as our shock to nitrogen use. Because both nitrate pollution trends and subsidy reform timing may be endogenous, we employ plausibly exogenous variation in predetermined soil characteristics and river flow direction to instrument for post-reform changes in river nitrogen pollution. We use a recently-developed

[1] For recent reviews, see Keiser et al., 2019; Currie and Walker, 2019; Aldy et al., 2022.

[2] We focus on infant mortality because infants are particularly susceptible to nitrate pollution. 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; Mateo-Sagasta et al., 2017).

hydrological model to identify highly-resolved upstream runoff catchment basins for each water quality monitor in our data, enabling a highly localized (versus district level) unit of analysis. We find that higher nitrate concentrations in river bodies increase infant mortality, with the effect driven by rural areas where river water exposure is more likely. In particular, rural infants located downstream of regions with high clay soils experience a statistically significant 1.6% increase in mortality rates for each 1% increase in nitrate pollution.

India has a long history of fertilizer subsidies. Between 2005 and 2010, the fertilizer subsidy burden faced by the Indian government increased by 500% as global prices increased, stressing central government budgets (Ravinutala, 2016). In response, the government introduced the Nutrient-Based Subsidy (NBS) program in 2010, which reduced price supports for most fertilizers, including phosphorus and potassium, but excluded nitrogen. Nitrogen prices were left untouched because policymakers considered liberalization of urea prices (the primary source of nitrogen) to be politically sensitive (Kishore et al., 2021). As a result, phosphorus and potassium fertilizer prices increased while urea became relatively cheaper under continued government price controls. This policy change has been associated with a significant increase in nitrogen fertilizer use among farmers (Gulati & Banerjee, 2015). We leverage the NBS policy as a natural experiment to examine the impact of this relative decrease in nitrogen fertilizer prices on nitrate river water pollution and infant mortality.

We face two main challenges to identifying the causal relationship between fertilizer use and infant mortality. First, post-policy changes in nitrogen use and runoff may be endogenous. We exploit variation in predetermined geographic characteristics – in particular, soil clay content – as an exogenous shifter of post-policy nitrogen. Soil clay content shifts both nitrogen fertilizer application and runoff. Fertilizer application shifts because soils with higher clay content hold moisture and nutrients longer than coarser soils, increasing plant nutrient uptake and thus the marginal return to nitrogen fertilizer (Burke et al., 2019).[3] Surface water runoff shifts because soils with a relatively high percentage of clay have smaller pore spaces, leading to lower water infiltration rates and greater retention of applied nutrients near the soil surface. Consequently, soils with higher levels of clay exhibit greater surface runoff than sandy or silty soils.[4]

[3]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. Our data does not contain areas with such extremely high levels of clay content.

[4]Note that greater retention of fertilizer near plant roots and higher rates of fertilizer runoff in

A second key identification challenge is that fertilizer use can affect infant mortality through increased yields and incomes (Bharadwaj et al., 2020), which may lead to better food security and improvements in infant health. This effect makes estimation of the impact of nitrogen pollution difficult as exposed communities may also experience offsetting local fertilizer benefits. To address this issue, we employ an upstream-downstream specification by exploiting the flow direction of river bodies. Rather than focusing on local soil characteristics near infant populations, we trace the watershed upstream and use the soil characteristics of these upstream areas to isolate the effects of nitrate water contamination on downstream populations. Importantly, we use an open-source global watershed mapping platform (Heberger, 2022) to map upstream watersheds for each water monitoring station. This allows us to conduct our analysis at the water station level, enabling precise identification of watersheds contributing to nitrate runoff and precise matching of local pollution exposure to local populations. To our knowledge, this represents the first developing-country analysis of water pollution mortality effects at this level of geographic precision.

Our analysis begins by measuring the effect of the NBS policy on nitrogen fertilizer use. We combine district-level fertilizer consumption data between 2000 and 2015 with a rich soil dataset from the Harmonized World Soil Database, a high resolution soil inventory providing detailed information on dozens of soil properties in a 30 arc-second (approximately 1km) grid. Using both event study and difference-in-differences specifications over plausibly exogenous variation in soil clay content, we confirm that districts with high-clay soils – where returns to nitrogen are higher – exhibit significantly increased nitrogen fertilizer consumption after the NBS policy is implemented relative to lower-clay districts within the same state. The effect we estimate is economically meaningful, representing an 18% increase over the mean nitrogen fertilizer consumption the data.

Next, we use data from 1,100 river water quality monitoring stations to study changes in nitrate water pollution levels driven by the NBS. We pair these data with detailed hydrological data from the Global Watershed API (Heberger, 2022) to map the upstream watershed of each river water quality monitoring station. This hydrological dataset provides all the river tributaries and origins upstream of a given water monitoring station, as well as their corresponding watersheds. These watersheds usefully define the land area within which the flow of water runoff will eventually

high-clay soils are not mutually inconsistent: high clay soils exhibit less infiltration of fertilizer into deep soils and groundwater (away from crop root zones). Thus more of the fertilizer is "available" both to the plant root zone as well as for runoff.

reach a given water station, enabling us to measure the soil clay content "upstream" (within the watershed) of each individual water quality station. Using upstream soil clay content as a source of exogenous variation in post-policy fertilizer application and runoff, we find that stations with high-clay upstream soils exhibit a large and statistically significant 5.6 mg/L increase in river nitrate pollution levels relative to lower upstream-clay stations. This increase is substantial: it represents over half of the 10 mg/L safe limit of nitrates in drinking waters.[5] Again, results hold for both difference-in-differences and event study estimators.

Finally, using household data from approximately 10,000 clusters from the Indian wave of the Demographic and Health Surveys (DHS), we study the impacts of river water nitrate pollution from the NBS policy on infant mortality. Using upstream soil clay content as an instrument for post-period changes in local river nitrate pollution, we 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 after the NBS. This effect is driven by rural DHS clusters, where populations are more likely to rely on river water for drinking and bathing. Our IV regression results are conditioned on state-by-year fixed effects that absorb any differential trends across states over time and DHS cluster fixed effects, that control for cross-sectional time-invariant local characteristics. We focus on infant deaths within one year of birth to ensure a precise exposure period, avoiding the complications of long-term differential effects. We find that each 1 mg/L increase in nitrates results in 0.03 additional deaths in LATE rural DHS clusters, which implies an elasticity of 1.6% for rural infant mortality to nitrate river water pollution.

We conduct a series of falsification and robustness checks to probe the validity of our results. First, we estimate the first-stage specification for other water quality indicators available in the data that should be unaffected by the NBS. These include measures such as pH, temperature, conductivity, dissolved oxygen, and fecal coliform, and we find no effects. Next, while we use the entire watershed retrieved for each station in the main specification, we also run a robustness test that contains only upstream areas within 100 km of water stations to account for potential pollution decay over space and find qualitatively similar effects on nitrate pollution, with a slightly smaller point estimate. Crucially, we also conduct placebo tests on the full IV estimates using the downstream watershed catchment area for all monitoring stations, and find no effects on nitrate pollution or infant mortality. Finally, when linking DHS

[5] U.S. EPA threshold defined in 56 FR 3526.

clusters to nearby water stations, we use a buffer of 20km around each cluster in our main specification. We test robustness to other buffer distances and find that results diminish as the DHS clusters are located further from the water stations, but remain qualititively similar.

Our findings have broad relevance. Agricultural input subsidies are a common policy tool used by governments worldwide to increase agricultural productivity and rural incomes (Byerlee et al., 2008). 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. Our results document an important unintended consequence of one such selective subsidy removal.

Additionally, excessive fertilizer use and associated pollution of water bodies is a widespread issue around the world (Mateo-Sagasta et al., 2017). The introduction of fertilizer subsidies during the Green Revolution was instrumental in boosting agricultural yields, improved nutrition, and economic output in many regions of the world (Bharadwaj et al., 2020; von Der Goltz et al., 2020).[6] However, a major consequence of prolonged fertilizer subsidies has been the excessive use of fertilizers (Huang et al., 2017; Cassou et al., 2017; Kurdi et al., 2020), with India as a prime example (Kishore et al., 2021; Gulati & Banerjee, 2015). 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 40% 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 sources (Galloway et al., 2003; Conant et al., 2013; Mateo-Sagasta et al., 2017). Our findings highlight that these nitrogen-related contaminants pose important health risks for downstream populations (James et al., 2005; Damania et al., 2019).

[6] Recent evidence suggests that the benefits of increased fertilizer use have begun to plateau, as higher application rates translate into reduced productivity gains (Wuepper et al., 2020; Mueller et al., 2017; Itin-Shwartz, 2024).

This paper makes three main contributions to the literature. First, it contributes to the literature on the human health effects of pollution. Much of this literature focuses on air pollution (e.g., Chay & Greenstone, 2003; Currie & Neidell, 2005; Currie et al., 2009; Knittel et al., 2016; Schlenker & Walker, 2016; Deryugina et al., 2019; Anderson, 2020; Graff Zivin & Neidell, 2013), largely in developed nations like the U.S. A smaller literature studies water pollution, again largely focused in the U.S. (Currie et al., 2013; Keiser & Shapiro, 2018; Keiser et al., 2019). However, pollution sources, levels, and regulatory conditions differ significantly in developing countries.[7] While air pollution reduction policies have made some progress in lowering air pollutants in both India and China, regulations of water pollution remain limited and weakly enforced (Greenstone & Hanna, 2014).

Second, the paper contributes to the broader understanding of the negative consequences of selective subsidies that distort relative input prices (see Holden (2019) for a review). While a growing body of literature explores the effect of subsidy-driven distortions on input use, productivity, and farmer welfare (Adamopoulos & Restuccia, 2014; Donovan, 2021; Hsieh & Klenow, 2009; Restuccia et al., 2008; Chakraborty et al., 2024; S. Garg & Saxena, 2023; Kurdi et al., 2020), less is known about the effects of distortionary subsidies on health outcomes. In the Indian context, the environmental and human capital damages of fertilizer subsidies have yet to be taken into policy consideration, even though populations there are particularly vulnerable to water source pollution (Fuller et al., 2022).

Third, the paper also contributes to improvements in measuring downstream pollution. Most studies outside the U.S. conduct an upstream-downstream framework at some aggregated level (e.g. Greenstone & Hanna, 2014; Dias et al., 2023), for example describing one political region as upstream of another based on the flow direction of major rivers. Unlike the U.S. context (Keiser & Shapiro, 2018), mapping precise upstream watersheds is still a challenge in developing country settings, with some recent exceptions (T. Garg et al., 2018; Hagerty & Tiwari, 2022). Using an open-source global watershed mapping platform developed by Heberger (2022), we enhance the spatial analysis of watersheds in India by mapping upstream watersheds for each water monitoring station. This mapping employs detailed gridded terrain data in

[7] Outside the U.S., studies of water pollution and human health have examined sanitation and human waste (Gamper-Rabindran et al., 2010; T. Garg et al., 2018), industrial waste (Ebenstein, 2012; Do et al., 2018), and pesticides (Lai, 2017; Skidmore et al., 2023; Dias et al., 2023). None of these examines runoff of fertilizers, which are heavily subsidized in many parts of the developing world (Huang et al., 2017). An exception is Brainerd and Menon (2014), which we discuss below.

combination with computational efficiencies to produce highly localized definitions of the catchment basin of each water monitoring station in our data. This detailed mapping allows us to conduct our analysis at the water station level instead of at an aggregate political unit, allowing for precise identification of watersheds contributing to nitrate runoff.

The closest work that complements this study is by Brainerd and Menon (2014). They study the effects of overall river agricultural pollutants on infant health in India and find significant negative impacts, exploiting seasonal prenatal exposure to agri-chemicals to identify impacts on various measures of child health. Our study differs from Brainerd and Menon (2014) in several important ways. First, the Brainerd and Menon (2014) measure of agricultural pollution is constructed from statewide averages of high concentration indicators across multiple pollutants, making the results difficult to interpret.[8] In contrast, we link river monitor-specific measures of nitrate and nitrite pollution directly with the DHS villages near them, measuring a well-defined treatment at each village with a clearly interpretable marginal effect. Additionally, Brainerd and Menon (2014) estimate health effects in the state where fertilizer is applied. Thus their estimates reflect both potential health gains from enhanced crop productivities and household incomes as well as potential losses from chemical exposure (Bharadwaj et al., 2020). In our analysis, the upstream-downstream element of our specification allows us to isolate the health effects of nitrate pollution by focusing on communities downstream of those where fertilizers have been applied. Finally, Brainerd and Menon (2014) rely for identification on differences in wheat and rice planting seasons or on differential trends in wheat and rice cropped areas across predominantly wheat states (in Northern India) and rice states (in Southern and Eastern India). However, differences in planting seasons are correlated with seasonalities in health outcomes, and may confound the effect of temperature and rainfall on infant health (McMahon & Gray, 2021). Further, differential trends cropped area trends are correlated with multiple determinants of health outcomes, such as differences in women's labour-force participation (Carranza, 2014). In contrast, we employ plausibly exogenous variation in clay soil distributions, water monitor catchment basin

[8] Brainerd and Menon (2014) measure treatment as an average of seven dummy indicators for statewide levels of different water pollutants, where indicator i takes a value of 1 if the average level of pollutant i across monitors within the state-month exceeds the U.S. EPA safe threshold for that pollutant. The dummy values of the seven pollutant averages are then themselves averaged, which defines the measure of treatment in a state-month. These indicators include measures for fluoride and chromium, which can indicate fertilizer runoff but are also associated with mining, brick kilns, tanneries, and pesticides (Sankhla & Kumar, 2018; Karunanidhi et al., 2021).

geographies, and river flow direction along with a specific policy shock to measure within-state variation in nitrate pollution and health.

Finally, we note that despite the widespread criticisms of the NBS program in India (e.g., Patel, 2016; Praveen, 2021; Mohapatra, 2022), there are no causal estimates of its impacts on environmental and health outcomes. We provide important and timely evidence on a policy that is still in place and frequently discussed by the Indian government, policymakers, and media (Gulati & Banerjee, 2015; Shagun, 2023; Ghosal, 2023; Damodaran, 2024).

The remainder of the paper proceeds as follows. Section 2 provides background details on the NBS policy and discusses its implications for nitrogen use. Section 3 describes the data, and Section 4 elaborates on the identification strategy. We report results in Section 5 and Section 6 concludes.

2 Background

2.1 Nutrient-Based Subsidy

India's fertilizer market has been government-regulated since the mid-1950s. Following the Green Revolution in the late 1960s, the government began subsidizing fertilizers in 1977, allowing farmers to access fertilizers at affordable prices. However, between 2005 and 2010, the Indian government's subsidy burden increased by 500% driven by a substantial increase in international prices (Ravinutala, 2016). At the height of the global food price spike in 2008-09, fertilizer subsidies accounted for 65% of the total government agricultural subsidy expenditure (Gulati & Banerjee, 2015), prompting the government to introduce changes to the subsidy structure to alleviate the fiscal burden and encourage a more balanced use of fertilizers.

In April 2010, the Indian government significantly changed how fertilizer subsidies were handled, transitioning from a product-based subsidy to the nutrient-based subsidy (NBS) system still in place today.[9] This new approach focuses on subsidizing fertilizers based on their content of key nutrients – nitrogen (N), phosphorus (P), potassium (K) and sulphur (S) – rather than setting fixed prices for specific fertilizer

[9]There have been earlier changes to the fertilizer subsidy policy previously as well, but these tended to be short-lived. For instance, a similar increase in the fiscal burden in 1991 led to the deregulation of predominantly phosphorus (P) and potassium (K) fertilizers, while nitrogen remained subsidized. However, the subsidies for P and K were reintroduced within a year. The time period used in our analysis does not include any other policy changes, except for the 2010 policy change.

products.[10] However, one key exception to this policy was urea, the most widely used fertilizer in India and the primary source of nitrogen for Indian farmers, which continued to be subsidized at a product-specific subsidy rate. The result of this partial deregulation of the fertilizer subsidy policy was a marked increase in the *relative* price of P and K compared to N in the farmers' fertilizer mix.

In practice, the NBS policy is implemented in the following manner. Each year, typically before the start of the two main agricultural seasons (the monsoon or *kharif* season and the winter, dry season or *rabi*), the government announces a fixed subsidy rate in rupees per kilogram for each nutrient, determined based on international prices, inventory levels, and expected demand. These per-kg rates are then converted into per-ton subsidies for each product based on that product's nutrient composition. For instance, Diammonium Phosphate (DAP) is a widely used product for phosphorus. Before 2010, the government set a fixed price for DAP; but since the NBS reform, DAP has been subsidized in India based on its nutrient composition . Using the 2023-2024 rabi season subsidy rates for the product DAP, we show an example of how the subsidies are calculated. First, the nutrient composition of DAP is 18-46-0-0 representing the ratio of N:P:K:S. The 2023-2024 subsidy rate for N was 47.02 Rs/kg and for P was 20.82 Rs/kg (since DAP does not contain any K or S, we ignore them in this example). The subsidy rate per metric ton (1,000 kg) of DAP is then calculated as the following:

Nitrogen subsidy: 18% of 1,000 kg = 180 kg,	$180 \times 47.02 =$ Rs. 8,463
Phosphorus subsidy: 46% of 1,000 kg = 460 kg,	$460 \times 20.82 =$ Rs. 9,577
Total DAP subsidy per metric ton:	Rs.18,040

Following the announcement of the nutrient subsidies each season, manufacturers can set their own maximum retail prices (MRP) for fertilizers after taking the government-announced subsidy rates into account. 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 after decentralization under the NBS

[10] 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 (DAP) is the most widely used phosphate fertilizer and Muriate of Potash (MOP) is the most widely used potassium based fertilizer. NBS values are available for Diammonium phosphate (DAP), Muriate of Potash (MOP), Monoammonium phosphate (MAP), Triple superphosphate (TSP) and 12 other grades of complex fertilizers.

policy, manufacturers set prices for P and K fertilizers, the prices for those fertilizers increased by over 150% and 255% respectively.

Urea, unlike other products, is the only fertilizer sold at a government-fixed retail price. It accounts for nearly 83% of the nitrogen inputs consumed by Indian farmers in 2022-23 (Pawar et al., 2025), making it the most prominent fertilizer among farmers. The price of urea has remained almost at the same level for more than 15 years. Urea prices were left untouched since policymakers considered the decentralization of the entire urea sector to be highly sensitive (Kishore et al., 2021).

Thus, the introduction of the NBS policy led to substantial differential price distortions across inputs, with persistently high subsidies for N (via urea) and new reduced subsidies for P and K. After this relative price decrease for N, the share of nitrogen usage increased notably within just three years of the policy implementation. Figure 1(a) shows farmer-reported prices from the cost of cultivation survey data, demonstrating that phosphorus and potassium prices increased relative to nitrogen prices following introduction of the NBS policy. Panel (b) shows the relative change in fertilizer application rates compared to 2009, at the district-level. We observe that after the policy change, nitrogen use increases, while both phosphorous and potassium use fall.

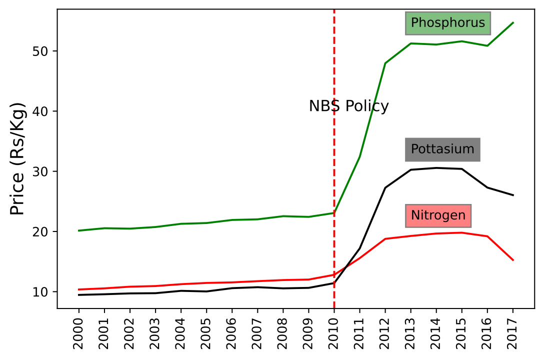
The descriptive patterns we see in Figure 1 echo what previous researchers and government policy documents have noted. For instance, Gulati and Banerjee (2015) argue that excluding urea from the NBS scheme was a significant oversight in India's fertilizer policy reform. They cite the Planning Commission's 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).[11]

2.2 Nitrate Runoff, Water Pollution and Health

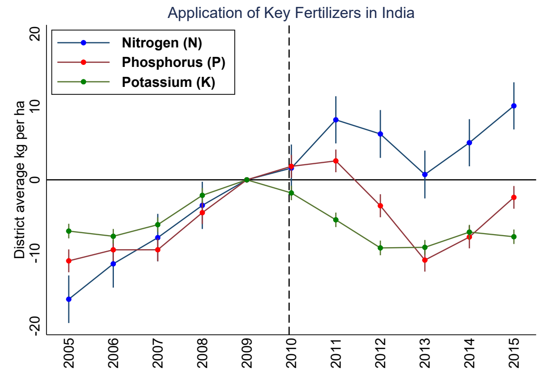
The increased use of urea under the NBS program in India may cause unintended negative externalities on river water quality due to the increased amount of nitrates entering nearby water bodies. First, the application of nitrogen-based fertilizers like urea introduces inorganic nitrogen into the soil 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,

[11]For more information on the policy, see Gulati and Banerjee (2015).

Figure 1. Trends in fertilizer prices and change in levels of usage after NBS policy



(a) Farmer retail price trends



(b) Relative changes in district-level fertilizer application rates

Notes: Panel (a) shows average annual farmer-reported retail prices (INR/kg) for N, P, and K fertilizers using data from the Cost of Cultivation surveys (Section 3). Panel (b) plots the coefficients on year dummies from a district-level panel regression with district fixed effects and shows change in N,P and K application (kg/ha) relative to 2009.

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. Second, the degree to which a plant will be able to absorb nitrogen will be a function of other nutrient availability (von Liebig, 1843). Thus, if a plant does not have sufficient phosphorous, it will not be able to process the nitrogen efficiently, potentially leading to greater nitrogen runoff.

Nitrate contamination poses severe risks to both human health and the environment (Li et al., 2021; Liu et al., 2014). High nitrate levels in drinking water can lead to methemoglobinemia, commonly known as blue baby syndrome. Methemoglobinemia reduces the blood's capacity to transport oxygen. In the developing-country context, Brainerd and Menon (2014) find maternal exposure to poor water quality (including nitrate pollution) associated with adverse birth outcomes, while Zaveri et al. (2020) find early life exposure to nitrate in the first three years of life is associated with shorter adult height attained among women. Studies have also linked high intake of nitrates with the incidence of cancer, blood diseases and other illnesses (Liu et al., 2014; Knobeloch et al., 2000). 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 due to oxygen depletion. Here, we focus on infant health impacts of nitrate pollution.

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.[12] 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.

2.3 Clay soil characteristics

Our research design exploits the fraction of clay soils in the districts and upstream watersheds as a source of variation for fertilizer application and nutrient runoff. Here, we describe the physical characteristics of clay-rich soils, which motivate their use in our empirical specification. We seek to provide an accessible summary of the soil science literature; see Churchman (2018), Wang and Li (2019), and Musei et al. (2024) for detailed discussions.

Clay-rich soils have distinct properties that enhance both nutrient retention and water-holding capacity, making clay soils particularly responsive to fertilizer applications (Churchman, 2018; Cambouris et al., 2016). Clay soils consist of finer particles than sand and therefore have much greater surface area relative to their volume. They also carry more negatively charged surface, which increases the soil's cation exchange capacity (CEC). A higher CEC increases a soil's ability to temporarily hold positively charged nutrient ions on particle surfaces and release them gradually to plants, rather than allowing them to be washed away immediately. As a result, finer-textured soils typically retain water and many nutrients more effectively than coarse sandy soils, where large pore spaces allow water to drain rapidly and reduce the time over which nutrients remain available in the root zone.

These properties help explain why fertilizer can be more productive on clay-rich soils. Because clay soils generally hold moisture for longer periods, crops are less likely to face short-run water stress after fertilizer application, and nutrients remain in the biologically active root zone for longer. In contrast, sandy soils drain quickly,

[12] 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.

so both water and soluble nutrients are more likely to move below the root zone before plants can fully use them. As a result, clay-rich soil can increase the marginal return to fertilizer application.

This mechanism is also consistent with plot-level evidence from the Cost of Cultivation Surveys.[13] In translog-style production-function estimates with farmer-crop and state-crop-year fixed effects, nitrogen has a positive average association with both yield and revenue, and the estimated return to nitrogen is higher on plots with greater clay content, as reflected in a positive $N \times \text{clay}$ interaction (Appendix Tables A.1 and A.2).

Clay-rich soils also affect the pathway through which potential fertilizer runoff reaches water bodies. The key hydrologic mechanism is that clay soils have smaller pores and therefore lower infiltration rates, particularly when soils are already wet or compacted.[14] In sandy soils these pores are larger, so rainfall can infiltrate more quickly into the soil profile. In clay-rich soils, by contrast, water enters more slowly. When rainfall intensity exceeds the infiltration capacity of the soil, excess water flows over the land surface as runoff. This makes clay-rich areas more prone to transporting fertilizer residues and dissolved nutrients into nearby streams, rivers, and other surface water bodies (Bombino et al., 2019; Wang & Li, 2019).

Taken together, these two linked features of clay-rich soils - greater agronomic responsiveness to fertilizer and greater propensity for surface runoff - motivate our use of upstream clay content as a shifter of downstream pollution exposure.

2.4 Implications for empirical design

The nutrient uptake and runoff characteristics of clay soils are significant in our study context. Our identification strategy relies on the pre-determined, exogenous distribution of clay soils in India, yielding exogenous variation in fertilizer efficacy and runoff, and thus exogenous variation in nitrate pollution in spatially-connected water bodies (Figure 2). As described in Section 4, we exploit this variation, along with exogenous variation in river flow direction, in our empirical design. Importantly, the upstream-downstream design isolates the pollution channel rather than local agricultural income effects: upstream soil texture affects the likelihood that fertilizer-derived nitrogen reaches downstream populations through connected water bodies, while the direct production benefits of fertilizer accrue primarily where the fertilizer is applied.

[13] More details on the Cost of Cultivation Surveys are provided in Section 3.1

[14] "Pore space" refers to the voids between soil particles through which air and water move

3 Data

3.1 Fertilizer Usage

We measure nitrogen use before and after the policy using district-level fertilizer consumption data available from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) District-Level Database (DLD). The ICRISAT-DLD provides annual district-level totals of nitrogen (N), phosphorus (P), and potassium (K) consumption (in tons) and per-hectare use for 1990–2017 (Figure 1, Panel B).

In addition to district aggregates, we analyze farmer-level input choices using the Cost of Cultivation Surveys (CCS).[15] The CCS collect detailed plot-level information on input use, crop-specific input prices and output prices for a rotating sample of farmers followed across all planting seasons for three consecutive years within a wave. We use annual CCS waves from 2008–2009 through 2018–2019, covering major crops including rice, wheat, cotton, sugarcane, and maize.

3.2 River Water Quality

We measure surface water quality using station-level data from the National Water Quality Monitoring Programme (NWMP).[16] The NWMP reports multiple indicators, including nitrate and nitrite, fecal coliform, pH, dissolved oxygen (DO), temperature, and conductivity, for more than 1,100 river monitoring stations (Figure 2, panel A). Our data are constrained in that nitrate-nitrite data became available for a larger number of monitoring stations only after 2006, three years before the implementation of the NBS program in 2010. Thus, we focus our analysis on 2007–2014, the period with sufficient spatial coverage of nitrate/nitrite water pollution measures.[17]

Our first outcome of interest is nitrogen pollution, measured as the cumulative concentration of nitrate-nitrite (mg/L) at the station-year level. Higher nitrate-nitrite levels correspond to more nitrogen contamination. For reference, a commonly used safety threshold for nitrate-nitrite levels in water is 10 mg/L.[18] In addition to nitrate-nitrite levels, we examine other indicators of water pollution such as dissolved oxygen,

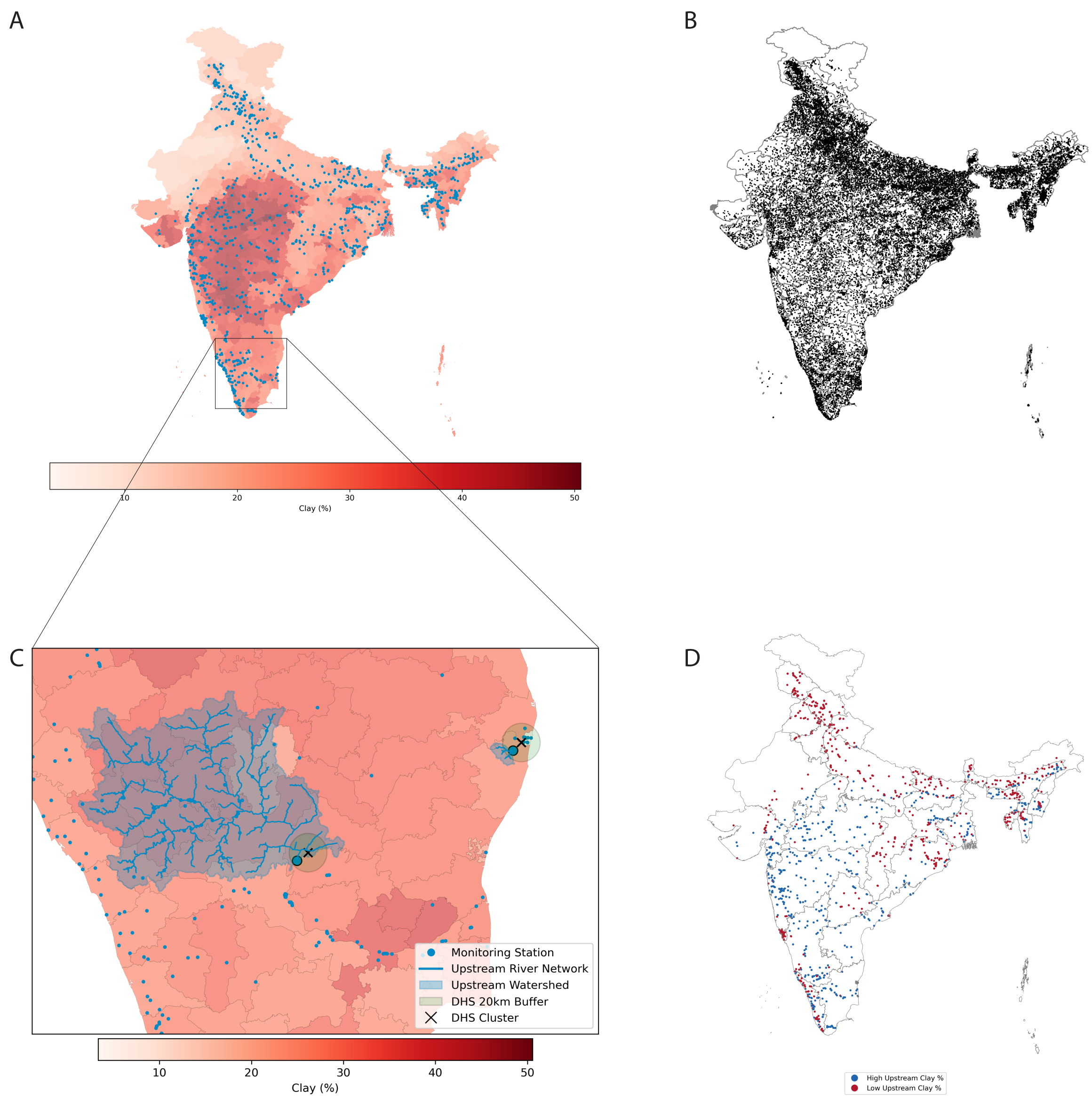
[15] The Cost of Cultivation Surveys are conducted by the Ministry of Agriculture and Farmers Welfare, Government of India.

[16] NWMP is collected and managed by India's Central Pollution Control Board. The list of stations and their geographic coordinates is sourced from https://cpcb.nic.in/wqm/WQMN_list.pdf.

[17] Our analysis period ends in 2014, which is the last complete year of infant mortality reported in the data.

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

Figure 2. Illustration of the data.



Notes: Panel (A) shows the distribution of soil clay content (background) and water monitoring stations (blue circles) throughout India, and Panel (B) shows the coverage of DHS clusters. Panel (C) illustrates the calculation of clay content for upstream watersheds. Two example DHS clusters (black x's) are shown, each with one example water monitoring station within 20km. One station has a large upstream watershed (light blue shape, center) while the other has a small upstream watershed (upper right). Average upstream clay content for each water station is calculated over each watershed area. Panel (D) shows the distribution of high- versus low-clay water monitoring stations in blue and red respectively, with "high" clay stations as those with above median clay content.

fecal coliform, temperature and pH. Panel A of Figure 2 maps the NWMP monitoring network for stations reporting nitrate-nitrite levels.

3.3 Upstream-Downstream Watersheds

A key component of our identification strategy is that pollution observed at a monitoring station reflects upstream activities. We therefore construct hydrologically consistent upstream exposure measures for each water quality monitoring station using the Global Watershed API developed by Heberger (2022). For any point on a river network this hydrological model returns flow direction, river tributary structure, upstream origins, and the associated watershed polygons. Using this API we construct exposure to upstream watersheds for each water monitoring station in our data (Figure 2, Panel C).

The use of the Global Watershed API represents a significant data enhancement relative to earlier research on fertilizer pollution. Prior work in this context often aggregates exposure within administrative boundaries (e.g., districts or counties) (T. Garg et al., 2018; Skidmore et al., 2023) or uses fixed-radius buffers around stations (Cai et al., 2016). Such strategies can result in measurement error through the misattribution of downstream or cross-basin exposure to a station and mask substantial variation in the area that actually drains to a station. In contrast, our watershed-based approach addresses such limitation by (a) restricting exposure to *upstream* areas only, aligning the exposure set to the hydrologically appropriate drainage polygons, and allowing watershed size and shape to vary by station. These choices reduce measurement error and sharpen the link between agricultural inputs and observed nutrients in surface water.

Our watershed approach requires precise locations of the water stations. However, some of the water stations do not have reliable latitude and longitude information. We quality-check CPCB station coordinates, snap stations to the nearest river flowline when minor offsets occur, and discard stations with unreliable geolocations. To avoid spurious catchments (e.g., canal segments or tiny headwater polygons), we exclude stations whose delineated upstream watershed area is smaller than 100 km^2.

3.4 Soil Characteristics

The second source of exogenous variation that we leverage for identification is differences in soil characteristics that affect the marginal productivity of fertilizer and its runoff (Section 2.3). To identify the underlying soil type, we use data from the

Harmonized World Soil Database (HWSD) v2.0, from 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 worldwide. We focus on the topsoil layer (0–20 cm), which is most relevant for nitrogen application and runoff. For each spatial unit, we compute the fractional shares of clay, sand, and silt.

We extract soil information at two scales. First, using 2011 district boundaries, we compute district-level averages of soil fractions for our district-based analysis. Second, for the station-level water quality analysis, we compute the same fractions within each delineated upstream watershed polygon for each NWMP station. Figure B.1 overlays pre-policy district-level nitrogen consumption with the district-level share of clay soils; both high- and low-clay districts exhibit high nitrogen use prior to the reform.

3.5 Health Outcomes

We study infant mortality using the 2015–2016 National Family Health Survey (NFHS-4), India's Demographic and Health Survey (DHS) round. NFHS-4 collected data from January 2015 to December 2016, surveying approximately 600,000 households nationwide using four instruments: household, women's, and men's questionnaires and biomarker data for a sub-sample of individuals.

Here we focus on the women's questionnaire from the NFHS-4 which records 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 the child's gender, twin status, birth order and survival status. The data indicate whether each child born is currently alive; if not, the age of recorded death. Our primary health outcome is infant mortality (death within 12 months of birth), given the established vulnerability of infants to nitrate exposure through methemoglobinemia.

The NFHS-4 also provides latitude and longitude coordinates for around 28,000 sample clusters (Figure 2, Panel B). 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 limit spatial misclassification when linking a cluster to water quality, we compute the average nitrate–nitrite concentration for all stations within a 20 km buffer around each NFHS-IV cluster.

Following Do et al. (2018), we then construct a cluster-by-year pseudo-panel of births, neonatal deaths (within the first month), and infant deaths (within 12 months). Due to limited earlier coverage of nitrate in NWMP, we exclude pre-2007 births from the analysis. Figure B.6 shows the annual distribution of births and deaths for the NFHS-IV sample.

3.6 Weather data

To account for the influence of precipitation on fertilizer runoff into watersheds, we use monthly rainfall from 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. For each NWMP station, we overlay its delineated upstream watershed and compute the monthly total rainfall upstream. We then aggregate to total annual precipitation (mm) and total monsoon-season precipitation. Using the total rainfall upstream might be important since some stations have a smaller upstream watershed, and some others have bigger upstream watersheds. Our approach allows for heterogeneous exposure based on watershed size and shape, which vary across stations.

4 Estimation Framework

There are two main challenges to estimating the effect of the change in fertilizer subsidies on nitrogen runoff and child mortality. First, fertilizer application and runoff are endogenous. We 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, we rely on exogenous variation from predetermined soil characteristics. Specifically, we leverage two aspects of clay soils that might affect nitrogen runoff. First, 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. Second, clay soils also lead to increased surface runoff. See Section 2.3 for a detailed discussion of both of these aspects.

For the second challenge, we focus on the effects of increased nitrogen use in upstream areas on infant mortality in downstream areas. The differences in the size and shape of the upstream basins, along with their difference in soil composition, give us plausibly exogenous variation in fertilizer use and associated runoff.

4.1 Differences-In-Differences Design

4.1.1 Nitrogen Use

We 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 relative price.

We first provide evidence that nitrogen use increased in high-clay districts relative to low clay districts after the NBS policy was implemented in 2010. Because estimates employ state-year fixed effects, our identifying assumption in this design is that within-state variation in predetermined district-level clay content is as good as random with respect to post-policy farmer actions.

We first estimate the following event-study style regression:

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

The main outcome of interest is the usage of nitrogen, phosphorus and potassium fertilizers in kg per hectare, in district i, state s and time period t. We denote the treatment variable as High Clay, an indicator variable with the value 1 for districts with above median clay levels (over the national distribution) and zero otherwise. We 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. If parallel trends hold in the pre-period, we should expect to see no systematic difference between districts with high and low levels of clay before 2009. If the policy change in fertilizer prices resulted in a shift in fertilizer application, we should expect to see the coefficients diverge from 0 immediately after 2009.

We include district-fixed effects to account for district-level observable and unobservable characteristics that are constant throughout the sample. We also add state-by-year fixed effects to account for time-varying factors that differ across states (this is important in the Indian context, where states have a high degree of independence). Our comparison of high- versus low-clay districts will produce a lower bound of the effects following the policy change because the districts that are classified into low-clay content may still be affected by the policy .

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

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

where $Fertilizer_{ist}$ is the N, P and K consumption per hectare at district i from state s in year t; and $HighClay_i$ is a dummy variable that indicates whether the fraction of clayey soil in the district is above the median. Figure B.5 in the appendix shows the spatial distribution of this binary treatment variable. $Post$ 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_i and η_{st} are district and state-year-fixed effects that account for any cross-sectional time-invariant characteristics and any differential trends across states over time. As noted above, we include state-by-year fixed effects in the main model because there are states that launched individual programs and welfare subsidies during different time points.

We also test our underlying assumption that farmers face differential marginal productivity of N, P and K on high versus low clay soils by estimating a production function using the detailed input data from the cost of cultivation data, for each of the 5 largest Indian crops; rice, wheat, maize, sugarcane and cotton. See Appendix A for details.

4.1.2 Nitrate Pollution in Water Bodies

Next, similar to the specification in equation 1, we use a reduced-form specification to estimate nitrate pollution levels using water quality measures at monitoring stations. To address concerns of endogeneity, instead of using clay levels at each water station, we focus on clay levels upstream of each monitoring station. Effectively, we 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 within-state variation in predetermined, upstream clay content is as good as random.

We estimate the following event-study style regression specification:

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

In this step, our main outcome of interest is the maximum nitrate-nitrite levels, $Nitrates_{ist}$ in water station i located in state s in year t. Additionally, we also estimate the same model for other water quality indicators such as fecal coliform rates, pH levels, dissolved oxygen content, temperature and conductivity which should be unchanged by the policy. We omit 2009 as the baseline year, as the policy was launched in 2010. Our main coefficient of interest is the coefficients on the interaction terms, β_k that measure changes in nitrate levels after relative to the baseline 2009 period.

As with the previous specification, here too, our 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, we 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.

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

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

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. We use the same specification to examine the impacts on the placebo outcomes: maximum fecal coliform, pH levels, dissolved oxygen (DO) and biochemical oxygen demand (BOD). As above, we include station and state-by-year fixed effects (μ_i and η_{st}) to account for any differential trends across the states over time. Our coefficient of interest is β_3 which is the interaction between *High Clay Upstream* and the *Post* period.

4.1.3 Infant Mortality

Reduced Form Specification Our last step is to estimate the effect of the policy on infant mortality. We first map the closest water stations to the clusters and compute the average nitrates in the upstream watersheds of these stations. We then compare clusters with higher clay content upstream to those with lower clay content before and after the policy change. We then test whether infant mortality changed differentially after the policy in clusters linked to stations with greater upstream clay content. As before, the identifying assumption is that within-state variation in predetermined, upstream clay content is as good as random.

We estimate the following reduced-form event-study-style specification:

$$\text{Infant Mortality}_{cst} = \sum_{k \neq 2009} \beta_k \left(\text{Upstream Clay}_{j(c)} \times \mathbf{1}\{t = k\}\right) + \alpha_c + \gamma_{st} + \varepsilon_{cst}, \quad (5)$$

where Infant Mortality$_{cst}$ denotes infant mortality in DHS cluster c, located in state s, in year t, and $j(c)$ denotes the monitoring stations assigned to cluster c. We omit 2009 as the baseline year. Analogous to the water quality regression, we include DHS cluster-level fixed effects, α_c and state-by-year fixed effects γ_{st} when examining the impacts on infant mortality. We expect these estimates to recover a lower bound since DHS clusters that are classified into low clay groups may still be affected by the policy.

Instrumental Variable Design Estimation of equation 5 yields the reduced-form impact of the NBS policy on infant mortality, but the model does not provide causal estimates for the effect of nitrates themselves on mortality. To address this, we employ an IV design, exploiting upstream clay levels to predict nitrate levels downstream. We 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 key exclusion restriction assumption underlying our IV strategy is that Upstream Clay $\times$ Post must affect infant health only through the channel of downstream nitrate pollution after controlling for DHS cluster fixed effects and state-year fixed effects. We adopt the upstream-downstream specification depicted in Panel C of Figure

2 to avoid local effects of soil quality on health.[19]

We estimate the following two-stage least square regressions where regressions 6 and 7 are the first and second stage regressions, respectively.

$$\text{Nitrates}_{jst} = \pi \left(\text{Upstream Clay}_j \times \text{Post}_t\right) + \theta \text{Precipitation}_{jst} + \mu_j + \eta_{st} + u_{jst}, \quad (6)$$

$$\text{Infant Mortality}_{cst} = \beta \, \widehat{\text{Nitrates}}_{j(c)st} + \delta \text{Precipitation}_{cst} + \alpha_c + \eta_{st} + \varepsilon_{cst}, \quad (7)$$

where Nitrates_{jst} denotes nitrate concentration measured at station j in state s and year t, and $\widehat{\text{Nitrates}}_{j(c)st}$ is the average of the fitted nitrate exposure for the stations assigned to cluster c. Monitoring station fixed effects, μ_j, absorb time-invariant station characteristics, while state-year fixed effects, η_{st}, absorb state-specific shocks over time. Standard errors are clustered at the monitoring station level.

5 Results

We first demonstrate the effect of the NBS policy on fertilizer use in high versus low clay districts. Second, we estimate the effect of the policy change on nitrate pollution measured downstream of high versus low clay watersheds. Third, we estimate the effect of both the policy and (instrumented) nitrates on infant mortality at the cluster level. We then include a series of robustness tests and placebo tests for each regression. In the following, we discuss the results for each of the above steps in turn.

5.1 Effects of NBS on Nitrogen Use

We first ask whether the NBS affected nitrogen use. We test whether districts with clay levels above the median experience an increase in nitrogen consumption following the NBS in 2010 (Fig 3). Using event-study regressions on fertilizer use, controlling for state-year and district fixed effects, we find 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 in high clay districts continues at a steady rate until 2016. Notably, we do not find significant changes in the use of phosphorus and potassium fertilizers in high versus low clay districts after the NBS, suggesting that not only did the rate of nitrogen application increase, but so did the

[19] As discussed previously, higher levels of clay soils might affect local agricultural yields, local farmer welfare, and thereby increase local access to nutrition and health investments. To address this concern, we use upstream clay levels as an instrument for downstream nitrate contamination.

Table 1. 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

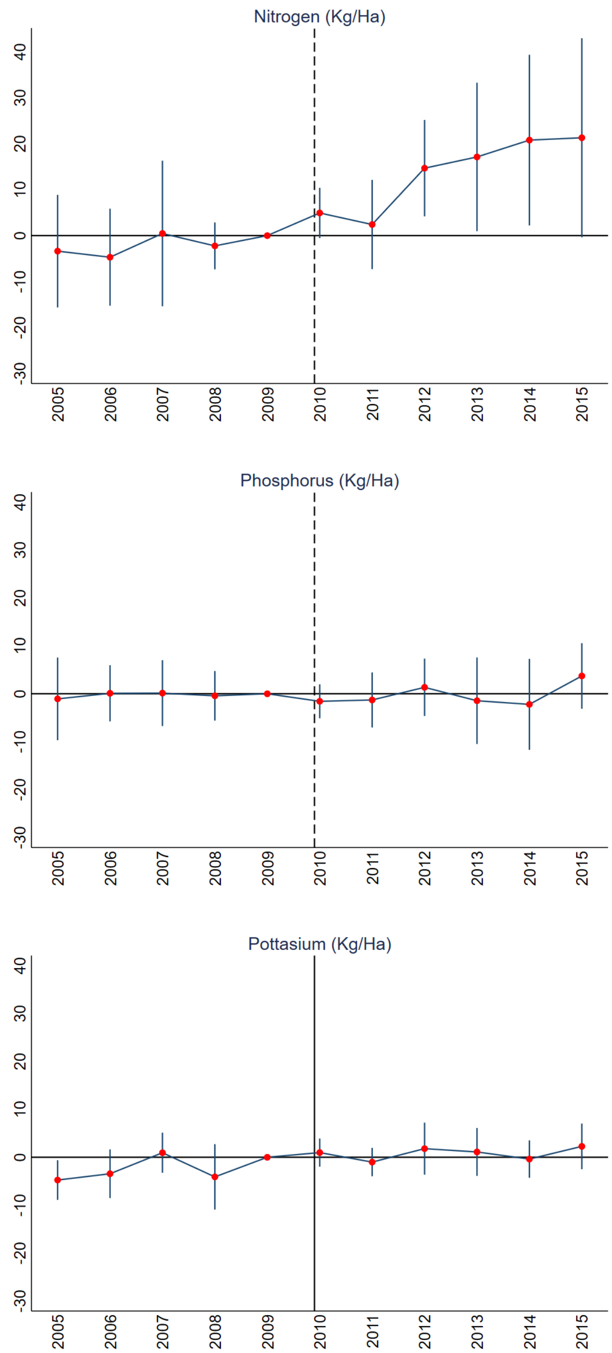
* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

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.

ratio of nitrogen to potassium and phosphorus.

Table 1 presents the main findings using the baseline specification in equation 2 that regresses fertilizer on clay soil content and the NBS, controlling for state-year fixed effects and district-fixed effects. The findings show a significant impact of the policy on annual nitrogen application per hectare 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 after the NBS compared to pre-policy years. Given that the mean nitrogen use is 84 kgs, the results translate to a 18% increase in nitrogen use in high clay districts post-policy. Columns 2, 4 and 6 report outcomes with the percent of clayey soils as a continuous variable instead of categorical high and low clay levels.

Figure 3. 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.

5.2 Effect of NBS on Nitrate Pollution

We next estimate the impact of the NBS on nitrate pollution measured at downstream water quality monitoring stations. Specifically, we compare nitrate levels measured downstream of high versus low clay watersheds. The results of the baseline specification from equation 4 are given in table 2.

We observe a jump in nitrate pollution downstream of high-clay watersheds after the NBS. We find that post policy, stations downstream from high clay watershed regions experience an increase in nitrate-nitrite levels of an average of 5.5 mg/l, almost tripling the prior level of 3.08 mg/L. The magnitude of increase is particularly alarming given 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.

Figure 4 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 was introduced 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.

Table 2. Impact of NBS Policy on Water Quality Measures

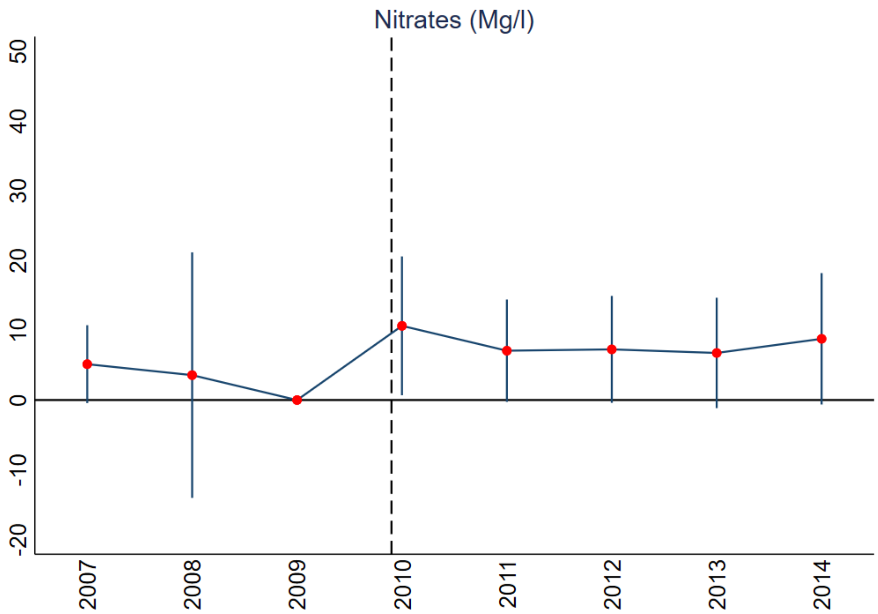
	(1) Nitrates	(2) Fecal Coli.	(3) Temperature	(4) DO	(5) pH	(6) Conductivity
Panel A: Upstream Clay						
Upstream Clay x Post	5.574**	-473018.3	0.0956	0.183	-0.157	163.1
	(2.220)	(412726.4)	(0.451)	(0.221)	(0.125)	(254.6)
Panel B: Upstream Clay upto 100 km						
Upstream Clay x Post	2.571*	-1537558.6	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.

Figure 4. 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 indicators and year dummies. Bars are 95% confidence intervals. Standard errors are clustered at the district level. The model includes water station fixed effects and state-year fixed effects.

5.3 Effects on Infant Mortality

Having established a link between the 2010 fertilizer policy change and nitrate-nitrite levels in water bodies, we now estimate the impacts of this policy on infant health outcomes. We specifically focus on infant mortality rates, defined as deaths within the first year of life. Medical literature has found that infants are particularly susceptible to nitrates, 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. Given that rural children are much more likely to rely on river water for drinking and bathing than urban children, and therefore might be more exposed, we also split our sample into rural versus urban clusters.

We use our measure of the percent clay content of soil in the upstream watershed interacted with the policy timing as an intent to treat estimate and as an instrument for nitrate pollution. For the IV, we use the percentage of clay upstream of each DHS cluster, interacted with a post-NBS policy indicator as the instrument for the average nitrate levels in nearby test sites. Our instruments are strong at all conventional levels, with F-stats between 24 and 80.

Table 3. 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	3.194	2.717	3.869
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 4. 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 4 report the reduced-form and instrumental variable approach findings based on equation 6 and equation 7. All specifications are conditioned on state-by-year fixed effects and DHS cluster fixed effects with standard errors clustered at the cluster level. Overall, we observe that a one mg/L increase in nitrates at nearby water stations results in 0.03 additional infant deaths. The results are even more striking for rural areas, with rural mortality increasing to 0.05 deaths. Given the dependent variable mean in rural areas is 0.09, this increase represents a substantial relative change in infant mortality and implies an elasticity of infant mortality to river water nitrates of 1.6%.[20] The coefficients for infant mortality in urban areas are statistically insignificant. Urban areas might not be affected by the increase in river water nitrites due to better water management and access to alternative sources of drinking water. The reduced form estimates are in the same direction as the IV estimates but are not statistically significantly different from zero at conventional levels.

5.4 Robustness Checks

5.4.1 Testing nitrogen-clay complementarities

One critical assumption that underpins our approach is that farmers will use more nitrogen on clayey soils. We test this assumption by using the cost of cultivation data to estimate a translog-like production function for six common crops in India that are grown on both clayey and less clayey soils: rice, wheat, maize, soy, sugarcane and cotton. We include both farmer and state by year fixed effect, to control for state-specific policies such as minimum support prices. We interact nitrogen with soil clay content, phosphorous and potasium (results are available in Appendix A). At the mean level of inputs, we find that N has a positive marginal product for all crops, and a higher marginal product in high clay soils. This is consistent with our observation that for the same input price, farmers in higher clay regions use more nitrogen fertilizer.

5.4.2 Placebo Tests on other Water Quality Indicators

Next, as a placebo test, we estimate whether other water quality indicators were affected by being downstream from high-clay watersheds after the NBS. We use our

[20] Calculated as $\frac{0.05/0.09}{1/2.7} = 1.6\%$, where 2.7 is the mean level of water station nitrates near urban DHS clusters.

baseline specification and evaluate changes in five other water quality indicators unrelated to nitrogen pollution: fecal coliform, temperature, dissolved oxygen, pH and conductivity. We find no statistically significant differences coming from high versus low-clay watersheds for any of the other five indicators. Results are reported in Table 2. This lack of effect on unrelated water quality measures suggests that the observed increase in nitrate-nitrite levels is indeed a consequence of the change in fertilizer policy.

As a second placebo test, we assess whether *upstream* nitrates or other waterborne pollutants are affected by *downstream* clay soil levels (testing the opposite river flow direction). If our results are picking up a general within-state watershed effect that changed after 2010 instead of one specifically associated with upstream nitrogen use, we might expect that downstream soil quality might be correlated with nitrate levels. Table 6 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. We find no effect of downstream soil clay levels on nitrate or other pollutant levels before and after the nitrogen policy change. We further use downstream clay as a placebo instrument for local nitrate pollution in our infant mortality IV, and find no effects (Table 7).

5.4.3 Clay soil categorization

We explore whether an alternative definition of the soil variable affects our findings. Instead of categorizing districts into high clay and low clay groups, we use the fraction of clay as a continuous variable in the specification accounting for state-year fixed effects and district fixed effects. This approach allows us to examine how varying levels of clay content affect my main outcomes. Results are robust and are reported in Table 2. However IV results for infant mortality are underpowered,[21] likely due to nonlinearities in how soil clay drives nitrate runoff.

5.4.4 Upstream watershed boundary

We next test whether our results are sensitive to the spatial assumptions we make when generating our data. There is no consensus on how far upstream is upstream for

[21] Results available upon request.

pollution mapping. If the distance is larger, 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. We use the entire watershed identified from the watershed mapper in my main specification. However, we also restrict the upstream area to only 100km radius and rerun the same specifications and find similar results. Panel B of Table 2 reports results of nitrate pollution. The table also reports results for other water quality measures and find no impacts.

5.4.5 Infant Mortality - DHS cluster buffers

To estimate the effect of upstream nitrate pollution on infant mortality, we create 20km buffers around each infant cluster and calculate the average nitrate levels from water monitoring stations located within these buffers in our main specification. Given that the DHS infant clusters are spatially jittered by approximately 5–10km for confidentiality, we consider the 20km buffer as a reasonable buffer for our analysis. However, to ensure robustness, we run the same specification using different buffers of 10, 20, 30 and 40km. Results in Table 5 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 we include surface water stations that are farther from DHS clusters, the effect diminishes and loses significance, suggesting that the proximity to contaminated water sources is key.

Table 5. 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 6. Placebo Test: Impact of NBS Policy on Water Quality Measures using Downstream Soil

	(1) Nitrates	(2) Fecal Coli.	(3) Temperature	(4) DO	(5) pH	(6) 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 7. Downstream Placebo: Impact on Infant Mortality (IV)

	(1) Full Sample	(2) Rural	(3) Urban
Nitrates	0.00297	-0.00409	0.0156
	(0.00806)	(0.0108)	(0.0114)
Observations	127267	88294	38971
Dep var mean	0.0959	0.111	0.0613

Standard errors in parentheses

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

Table reports placebo estimates using downstream soil properties as an instrument for upstream nitrate pollution, otherwise following our main IV specification. Downstream buffer depicted in Figure B.7. Standard errors are clustered at the DHS cluster level. Column 1 reports IV estimates for all DHS clusters, while columns 2 and 3 report estimates for rural and urban clusters.

6 Discussion and 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, we examine how India's Nutrient-Based Subsidy program, launched in 2010 skewed fertilizer use towards nitrogen and we 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, we show that the change to India's fertilizer subsidy that favoured nitrogen over phosphorous and potassium has negative externalities on river pollution and human health. To address potential endogeneity concerns on nitrogen use, we use the timing of the policy and predetermined exogenous variation in soil physical texture and river flow direction.

Using a difference-in-difference approach, we conduct three key analyses. First, we estimate changes in nitrogen use as a result of the policy. We find 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, we examine whether this increase in nitrogen use results in higher nitrate levels in nearby water bodies. Results from approximately 1,100 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, we 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, we examine the link between these nitrate increases and infant mortality rates. By constructing a pseudo-panel of approximately 10,000 clusters from the DHS survey data, we 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, we find that a one mg/L increase of river water nitrates leads to 0.03 additional infant deaths in DHS clusters with high upstream clay in the post-policy period. This effect is concentrated in rural clusters, where a one mg/L increase of nitrites leads to an additional 0.05 infant deaths.

Our results are all based on differences between high and low-clay upstream watersheds. To understand the overall effect of the policy change, we need to know whether the policy increased nitrogen use and nitrate pollution in net or only in relative terms. If, for example, the policy resulted in decreased nitrogen use and nitrate pollution in low-clay watersheds, the policy may have been a net benefit for health. However, when we examine nitrogen use in high- versus low-clay locations after the policy, we observe that nitrogen use increased in both sets of locations(Figure **??**).

If we assume that the policy did not affect nitrogen use in low-clay soils and downstream river water stations, that would mean that, based on the first stage of our IV regression, nitrates increase 1.3 mg/L downstream of high-clay soils, resulting in an additional 0.04 infant deaths, an increase of about 55% over baseline mortality levels. In rural areas, the policy increased infant deaths by 0.08 infant deaths (1.6 mg/L times 0.05 deaths per mg/L), almost doubling baseline mortality levels.

Our 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 increased 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. 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

Aldy, J. E., Auffhammer, M., Cropper, M., Fraas, A., & Morgenstern, R. (2022). Looking back at 50 years of the clean air act. *Journal of Economic Literature*, *60*(1), 179–232.

Anderson, M. L. (2020). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*, *18*(4), 1886–1927.

Arceo, E., Hanna, R., & Oliva, P. (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city. *The Economic Journal*, *126*(591), 257–280.

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

Bombino, G., Denisi, P., Gómez, J. A., & Zema, D. A. (2019). Water infiltration and surface runoff in steep clayey soils of olive groves under different management practices. *Water*, *11*(2), 240. https://doi.org/10.3390/w11020240

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

Byerlee, D., De Janvry, A., Sadoulet, E., Townsend, R., & Klytchnikova, I. (2008). *World development report 2008: Agriculture for development*. World Bank, Washington D.C., USA.

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.

Cambouris, A. N., Ziadi, N., Perron, I., Alotaibi, K. D., St. Luce, M., & Tremblay, N. (2016). Corn yield components response to nitrogen fertilizer as a function of soil texture. *Canadian Journal of Soil Science*, *96*(4), 386–399. https://doi.org/10.1139/cjss-2015-0134

Carranza, E. (2014). Soil Endowments, Female Labor Force Participation, and the Demographic Deficit of Women in India. *American Economic Journal: Applied Economics*, *6*(4), 197–225. https://doi.org/10.1257/app.6.4.197

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.

Chay, K. Y., & Greenstone, M. (2003). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics*, *118*(3), 1121–1167.

Churchman, G. J. (2018). Game changer in soil science: Functional role of clay minerals in soil. *Journal of Plant Nutrition and Soil Science*, *181*(1), 99–103. https://doi.org/10.1002/jpln.201700605

Conant, R. T., Berdanier, A. B., & Grace, P. R. (2013). Patterns and trends in nitrogen use and nitrogen recovery efficiency in world agriculture. *Global Biogeochemical Cycles*, *27*(2), 558–566.

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.

Currie, J., & Walker, R. (2019). What do economists have to say about the clean air act 50 years after the establishment of the environmental protection agency? *Journal of Economic Perspectives*, *33*(4), 3–26.

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

Damodaran, H. (2024). Explained: Argument for deregulating non-subsidised fertilisers. *The Indian Express*. Retrieved July 12, 2024, from https://indianexpress.

com/article/explained/explained-economics/deregulating-non-subsidised-fertilisers-9447888/

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.

Fuller, R., Landrigan, P. J., Balakrishnan, K., Bathan, G., Bose-O'Reilly, S., Brauer, M., Caravanos, J., Chiles, T., Cohen, A., Corra, L., et al. (2022). Pollution and health: A progress update. *The Lancet Planetary Health*, *6*(6), e535–e547.

Galloway, J. N., Aber, J. D., Erisman, J. W., Seitzinger, S. P., Howarth, R. W., Cowling, E. B., & Cosby, B. J. (2003). The nitrogen cascade. *Bioscience*, *53*(4), 341–356.

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.

Ghosal, S. (2023). Profitability for phosphatic fertiliser players to improve at current nutrient-based subsidy rates: Report. *The Economic Times*. Retrieved January 19, 2023, from 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

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

Greenstone, M., & Jack, B. K. (2015). Envirodevonomics: A research agenda for an emerging field. *Journal of Economic Literature*, *53*(1), 5–42.

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.

Hsiang, S., Oliva, P., & Walker, R. (2019). The distribution of environmental damages. *Review of Environmental Economics and Policy*.

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.

Jalan, J., & Somanathan, E. (2008). The importance of being informed: Experimental evidence on demand for environmental quality. *Journal of development Economics*, *87*(1), 14–28.

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

Karunanidhi, D., Aravinthasamy, P., Subramani, T., Kumar, D., & Venkatesan, G. (2021). Chromium contamination in groundwater and sobol sensitivity model based human health risk evaluation from leather tanning industrial region of south india. *Environmental Research*, *199*, 111238.

Keiser, D. A., Kling, C. L., & Shapiro, J. S. (2019). The low but uncertain measured benefits of us water quality policy. *Proceedings of the national academy of sciences*, *116*(12), 5262–5269.

Keiser, D. A., Mazumder, B., Molitor, D., & Shapiro, J. S. (2023). Water works: Causes and consequences of safe drinking water in america. *National Center for Environmental Economics Seminar, Washington, DC, April*, *13*, 853–872.

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.

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.

McMahon, K., & Gray, C. (2021). Climate change, social vulnerability and child nutrition in south asia. *Global Environmental Change*, *71*, 102414.

Mohapatra, J. K. (2022). Biting the bullet: Direct benefit transfer of fertiliser subsidy. *IC Centre for Governance*, *171*, 171.

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.

Musei, S. K., Kuyah, S., Nyawira, S., Ng'ang'a, S. K., Karugu, W. N., Smucker, A., & Nkurunziza, L. (2024). Sandy soil reclamation technologies to improve crop productivity and soil health: A review. *Frontiers in Soil Science*, *4*, 1345895. https://doi.org/10.3389/fsoil.2024.1345895

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

Olmstead, S., & Zheng, J. (2021). Water pollution control in developing countries: Policy instruments and empirical evidence. *Review of Environmental Economics and Policy*, *15*(2), 261–280.

Patel, A. (2016). Subsidies to farm sector in india need to rationalize & its efficient use. *Int. J. Sci. Res. Sci. Eng. Technol*, *2*, 1–11.

Pawar, N. D., Singhal, K., Bhushan, C., Pregger, T., & Jochem, P. (2025). Decarbonization of urea production in india and its impact on water withdrawal and costs: A cost optimization approach. *Journal of Cleaner Production*, *486*, 144433.

Praveen, K. (2021). Indian fertiliser policy: Retrospect and prospect. In *Soil science: Fundamentals to recent advances* (pp. 765–779). Springer.

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.

Sankhla, M., & Kumar, R. (2018). Fluoride contamination of water in india and its impact on public health.

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

Shagun. (2023). Cacp recommends centre to bring urea under nbs regime to check overuse. *DownToEarth*. Retrieved June 8, 2023, from https://www.downtoearth.org.in/agriculture/cacp-recommends-centre-to-bring-urea-under-nbs-regime-to-check-overuse-89907

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

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.

von Liebig, J. F. (1843). *Chemistry in its application to agriculture and physiology*. JM Campbell.

Wang, Z.-H., & Li, S.-X. (2019). Nitrate n loss by leaching and surface runoff in agricultural land: A global issue (a review). *Advances in Agronomy*, *156*, 159–217. https://doi.org/10.1016/bs.agron.2019.01.007

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).

A Production function estimates

We estimate farmers' returns to fertilizer application and the extent to which marginal returns vary with clay soil content using plot-level panel data from the Cost of Cultivation Surveys. Tables A.1 and A.2 report estimates from a translog production function using log yield and log revenue as outcomes, respectively. For both, we run estimates separately for the six major crops in India - rice, wheat, maize, soybean, sugarcane and cotton (columns (2) - (7)) - as well as pooled results for all crops (column (1)). Nitrogen, phosphorus, and potassium enter as main effects, and the specifications include selected pairwise interactions between nitrogen and clay content, phosphorus, and potassium. All input variables are centered on their sample means, so the coefficient on nitrogen can be interpreted as the average marginal return to nitrogen when other inputs are evaluated at their mean levels. To keep the specification parsimonious and avoid instability from overfitting, we do not include the full set of other two-way and higher-order interactions. Our specifications include farmer-crop and state-crop-year fixed effects, accounting for constant plot-level attributes such as soil endowments and farmer ability, as well as state-specific policy changes that may affect local agricultural input and output markets.

Across all crops, we find that returns to nitrogen application remain positive (for both yield and revenue) and increase with clay content. These results complement the findings from soil science literature and provide additional support for our research design, leveraging soil clay content as a shifter for fertilizer use.

Table A.1. Translog production function estimates: Crop yield

	(1) All	(2) Rice	(3) Wheat	(4) Maize	(5) Soybean	(6) Sugarcane	(7) Cotton
ln_Nkg_cent	0.090***	0.106***	0.075***	0.093***	0.024	0.043***	0.120***
	(0.004)	(0.006)	(0.007)	(0.019)	(0.013)	(0.009)	(0.016)
ln_Pkg_cent	0.016***	0.017***	0.016***	0.035***	-0.002	0.014***	-0.003
	(0.002)	(0.003)	(0.004)	(0.009)	(0.010)	(0.003)	(0.009)
ln_Kkg_cent	0.024***	0.027***	0.007*	0.032***	-0.001	0.022***	0.034***
	(0.002)	(0.003)	(0.003)	(0.007)	(0.008)	(0.004)	(0.005)
ln_Nkg_cent × ln_clay_cent	0.009*	0.015*	0.010	0.040	-0.039	0.032*	0.086***
	(0.005)	(0.007)	(0.005)	(0.031)	(0.038)	(0.013)	(0.017)
ln_Nkg_cent × ln_Pkg_cent	0.004***	0.003*	0.003*	0.017**	0.001	-0.001	0.004
	(0.001)	(0.001)	(0.002)	(0.006)	(0.003)	(0.002)	(0.004)
ln_Nkg_cent × ln_Kkg_cent	0.001	0.004***	-0.001	-0.018***	0.006	0.004	-0.013***
	(0.001)	(0.001)	(0.002)	(0.005)	(0.004)	(0.002)	(0.004)
ln_area_cent	0.136***	0.164***	0.105***	0.144***	0.037**	0.068***	0.092***
	(0.005)	(0.006)	(0.009)	(0.023)	(0.014)	(0.013)	(0.019)
Constant	3.551***	3.557***	3.434***	3.064***	2.415***	6.581***	2.701***
	(0.002)	(0.003)	(0.003)	(0.012)	(0.011)	(0.006)	(0.012)
N	251,591	132,989	63,728	13,702	13,695	13,442	14,035
Adjusted R^2	0.92	0.69	0.79	0.74	0.57	0.61	0.47
Within-R^2	0.06	0.10	0.03	0.05	0.00	0.04	0.06
State-Crop-Year FE	X	X	X	X	X	X	X
Farmer-Crop FE	X	X	X	X	X	X	X

Notes: This table shows the coefficients on production function parameters for a translog production function estimated using crop-specific farmer-plot-level input and output data from the Cost of Cultivation Surveys. The data consist of plot-level panel data, with each farmer followed for three years in the survey. standard errors in parentheses, adjusted for clustering at the district-year level. Significance level at 10, 5, and 1 percent shown by *, **, and ***, respectively.

Table A.2. Translog production function estimates: Crop revenue

Table A.3. log(Revenue)

	(1) All	(2) Rice	(3) Wheat	(4) Maize	(5) Soybean	(6) Sugarcane	(7) Cotton
ln_Nrs_cent	0.044***	0.053***	0.039**	0.050*	0.025	0.031	0.100***
	(0.006)	(0.008)	(0.013)	(0.020)	(0.020)	(0.019)	(0.027)
ln_Prs_cent	0.024***	0.025***	0.017***	0.029	-0.002	0.022***	0.006
	(0.003)	(0.003)	(0.004)	(0.020)	(0.017)	(0.006)	(0.011)
ln_Krs_cent	0.025***	0.033***	0.004	0.041***	-0.001	0.020**	0.037***
	(0.002)	(0.003)	(0.003)	(0.011)	(0.008)	(0.007)	(0.007)
ln_Nrs_cent × ln_clay_cent	0.010	0.017*	0.002	0.020	0.037	0.052**	0.070***
	(0.007)	(0.008)	(0.005)	(0.031)	(0.058)	(0.019)	(0.020)
ln_Nrs_cent × ln_Prs_cent	-0.003*	-0.003	-0.001	0.004	0.000	-0.008	-0.014**
	(0.001)	(0.001)	(0.002)	(0.007)	(0.009)	(0.006)	(0.005)
ln_Nrs_cent × ln_Krs_cent	0.003*	0.002	0.001	-0.003	-0.003	0.008	-0.003
	(0.001)	(0.002)	(0.002)	(0.007)	(0.005)	(0.006)	(0.005)
ln_area_cent	0.086***	0.106***	0.063***	0.082***	0.030	0.060**	0.074**
	(0.007)	(0.008)	(0.014)	(0.024)	(0.027)	(0.021)	(0.028)
Constant	10.692***	10.634***	10.680***	10.114***	10.256***	11.983***	11.042***
	(0.003)	(0.004)	(0.004)	(0.019)	(0.018)	(0.006)	(0.013)
N	251,703	133,024	63,728	13,747	13,725	13,444	14,035
Adjusted R^2	0.77	0.72	0.83	0.63	0.48	0.58	0.52
Within-R^2	0.02	0.03	0.01	0.01	0.00	0.02	0.03
State-Crop-Year FE	X	X	X	X	X	X	X
Farmer-Crop FE	X	X	X	X	X	X	X

Notes: This table shows the coefficients on production function parameters for a translog production function estimated using crop-specific farmer-plot-level input and output data from the Cost of Cultivation Surveys. The data consist of plot-level panel data, with each farmer followed for three years in the survey. standard errors in parentheses, adjusted for clustering at the district-year level. Significance level at 10, 5, and 1 percent shown by *, **, and ***, respectively.

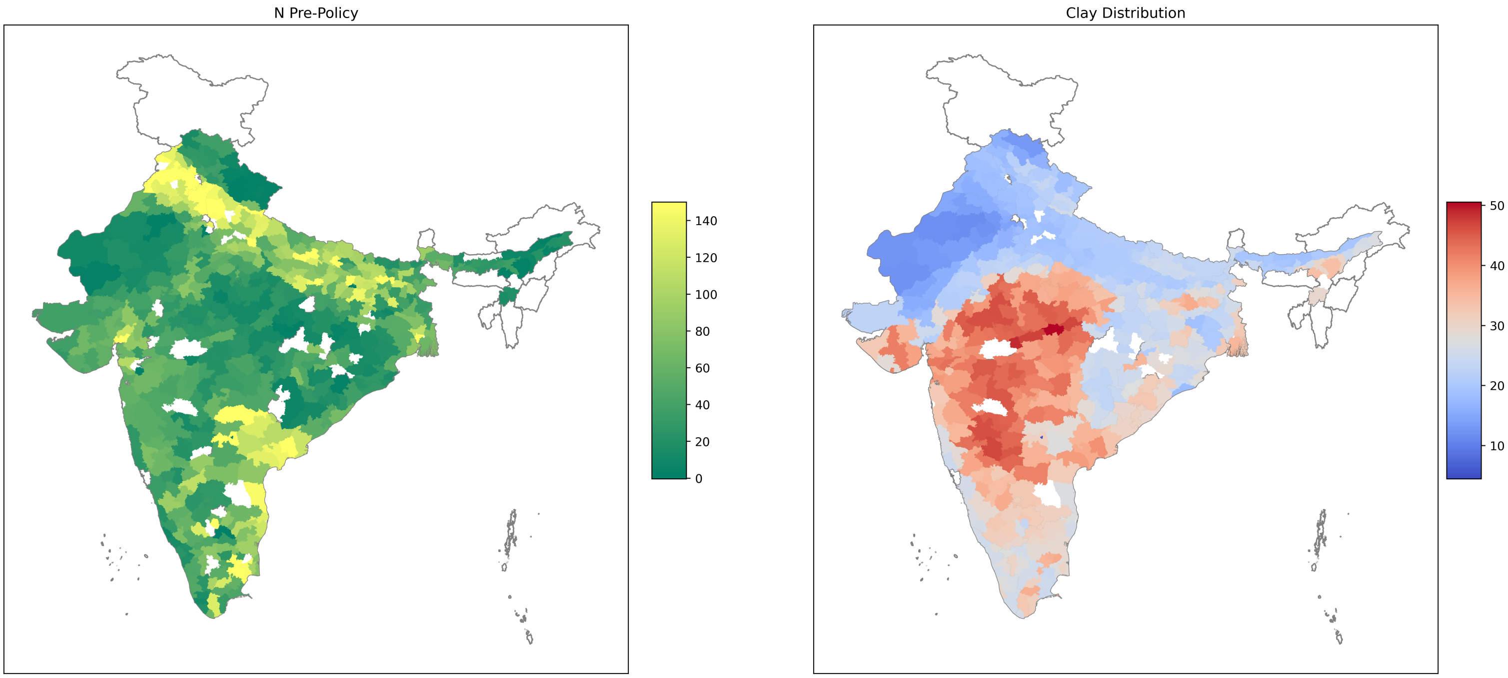
B Additional Figures and Tables

Table B.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

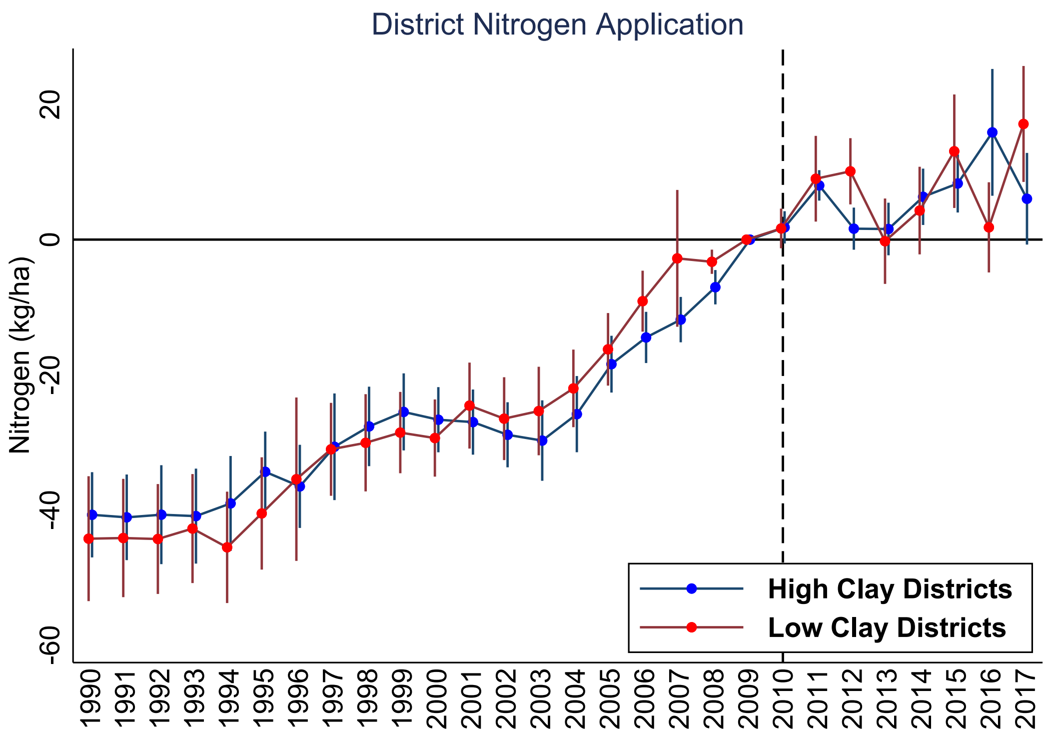
Notes: Table reports annual fertilizer consumption for the major nutrients N, P and K from 2005 to 2017 using district-level fertilizer consumption data from ICRISAT.

Figure B.1. Nitrogen Use and Clay Type by Districts



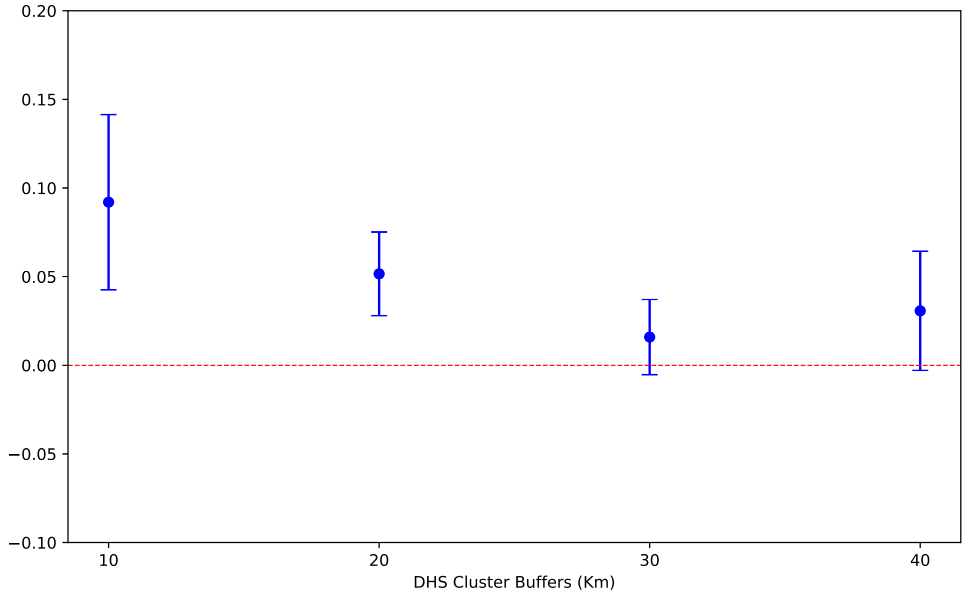
Notes: The left panel displays nitrogen consumption before the policy change in 2010. The right panel shows the percentage of clay soils on average for each district. Nitrogen data comes from ICRISAT and soil data comes from the FAO soil raster database.

Figure B.2. Nitrogen Use by District Clay Content



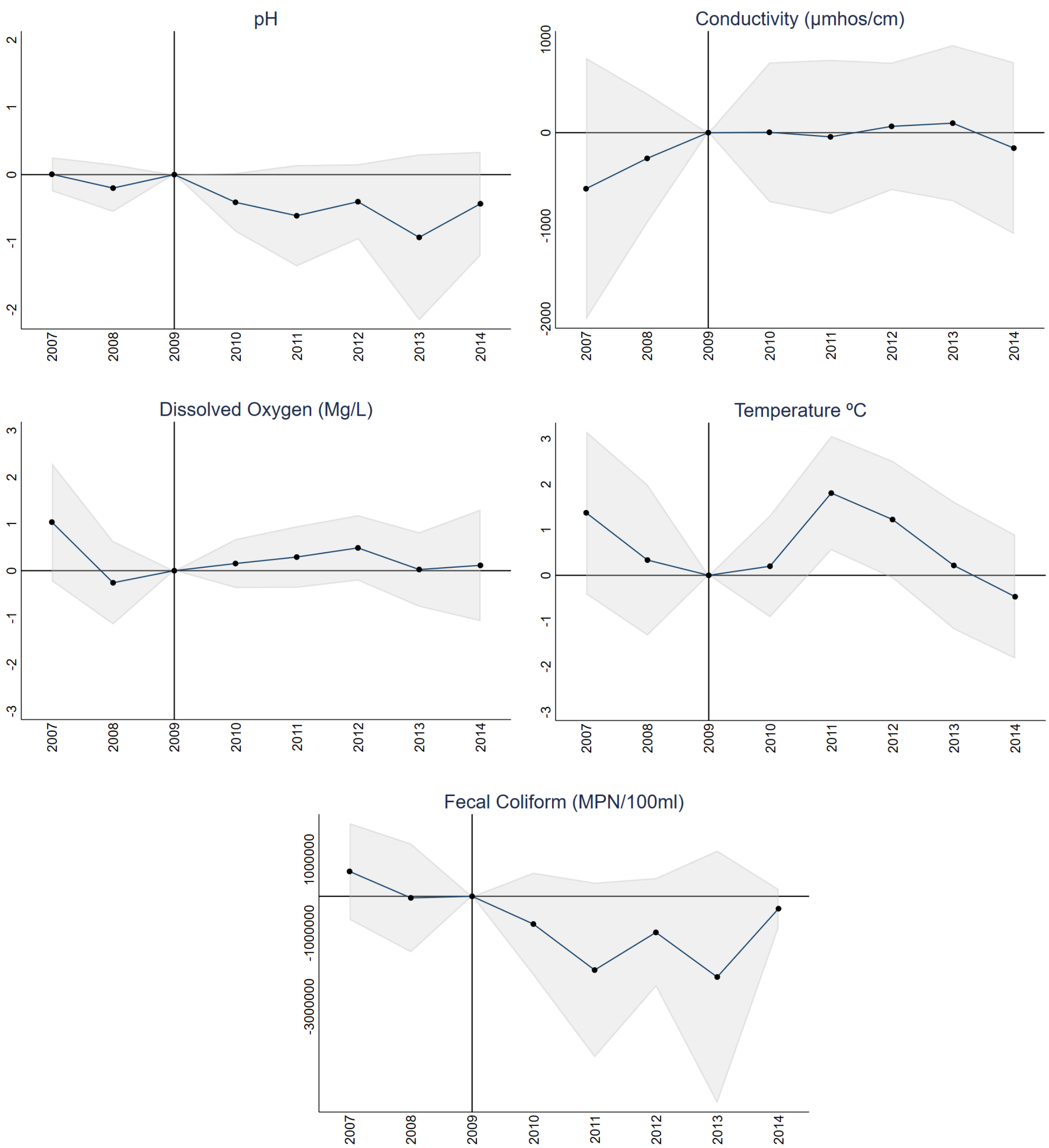
Notes: This figure displays the results of two regressions of district-level nitrogen usage on year fixed effects, estimated separately for high- and low-clay districts. Regressions also includes district fixed effects. 95% confidence intervals displayed are cluster-robust at the district level.

Figure B.3. Infant Mortality Estimates



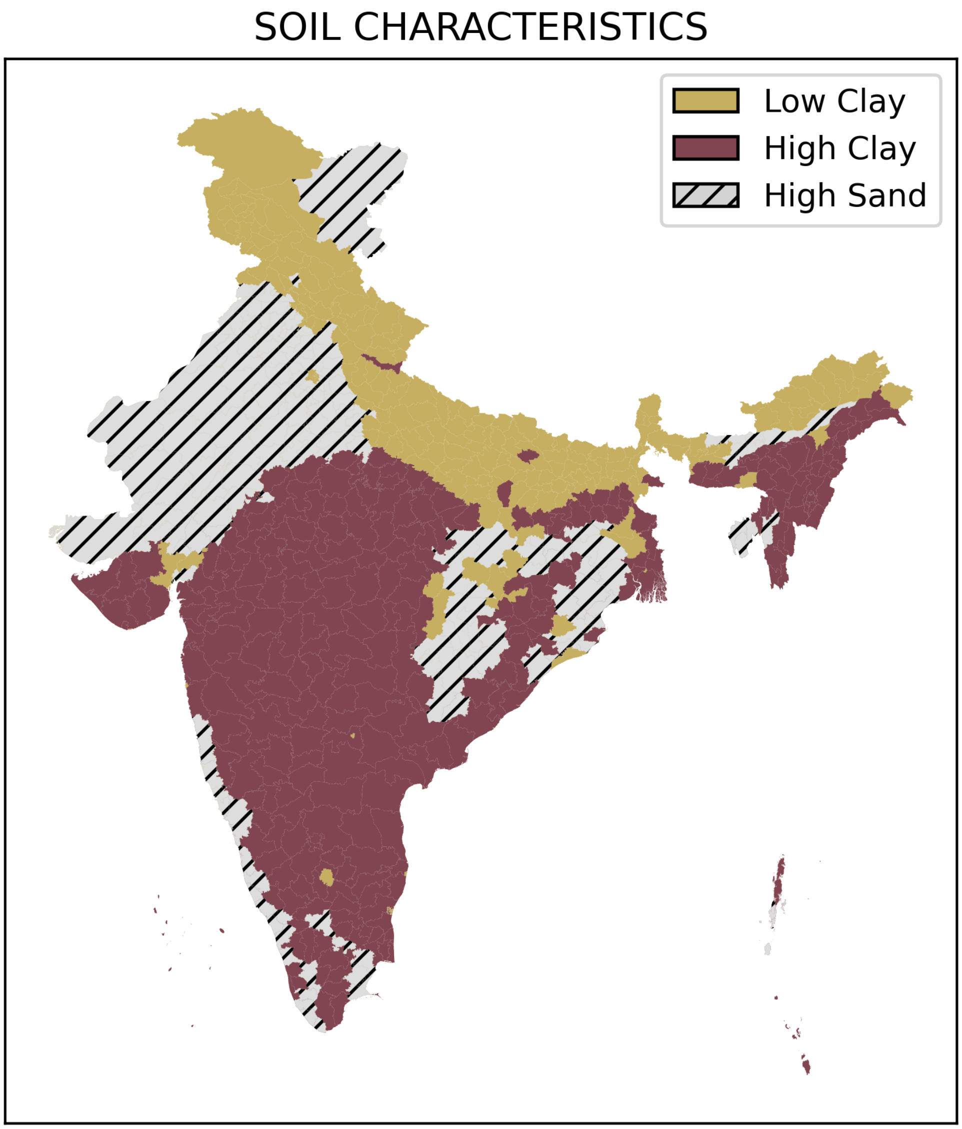
Notes: The figure reports coefficients for infant mortality in rural DHS clusters from the IV regression with different buffers.

Figure B.4. Water Quality Indicators in High vs Low Clay Districts



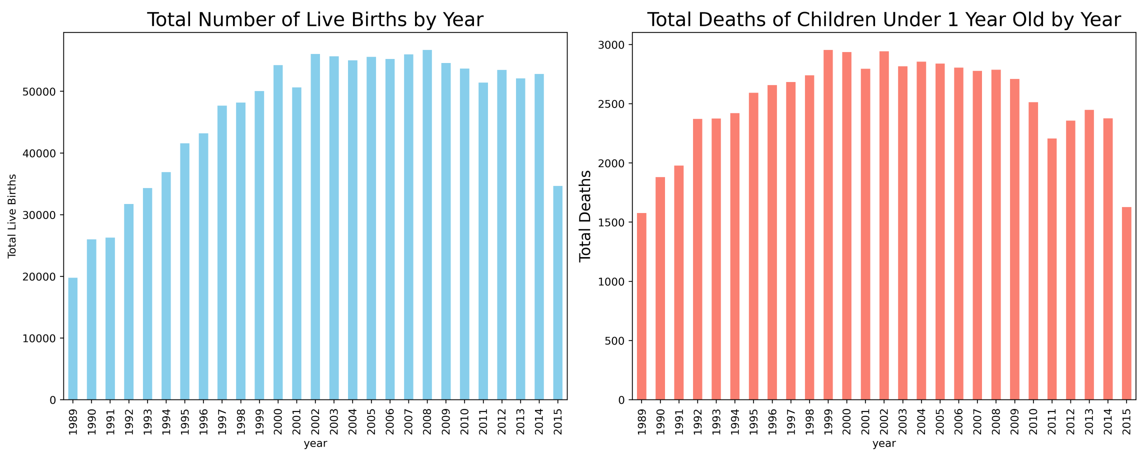
Notes: The figure shows the regression coefficients of various water quality indicators from the station-level data. The coefficients plotted are the interaction terms between upstream clay levels and year dummies. The 95% confidence intervals are shown as the shaded region. Standard errors are clustered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

Figure B.5. Treatment Definition



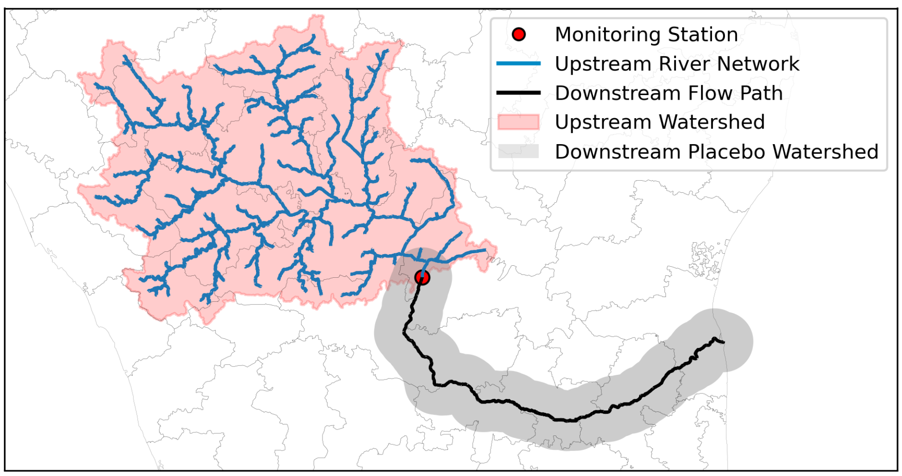
Notes: This figure shows the binary treatment at the district level in the difference in difference and event study regression designs. I classify districts into high clay or low clay based on the median clay content. The main model also does not include districts with high levels of sandy soils.

Figure B.6. Total infant births and deaths by year



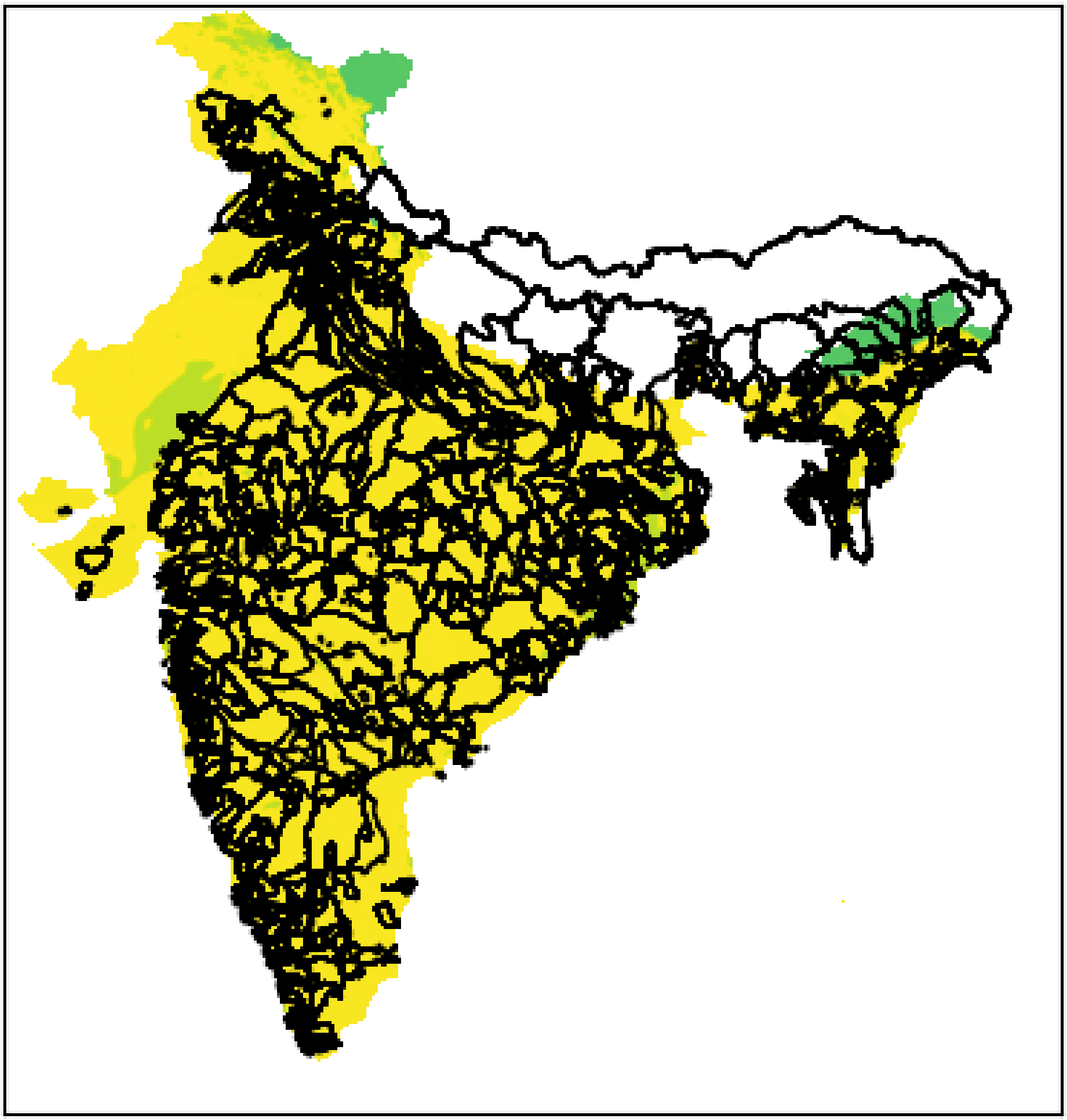
Notes: This figure shows the total number of live births and infant deaths within one year of birth as reported by mothers in DHS clusters.

Figure B.7. Downstream flow path



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

Figure B.8. Upstream polygons



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