The Curse of Plenty: Early Childhood Roots of the Rise in Chronic Disease*

[PRELIMINARY DRAFT]

Sheetal Sekhri† and Gauri Kartini Shastry‡

February 2019

Abstract

Chronic diseases are leading drivers of the global disease burden. Recent literature has documented long-term health benefits of positive income and nutrition shocks in early childhood. Dietary changes in environments with widespread poverty, however, can undermine the salutary effects of income shocks. We study the Green Revolution in India and its expansion in historically groundwater-endowed districts to shed light on this possibility. We find that areas where the Green Revolution increased crop yields the most saw an increase in diabetes incidence in cohorts born after the introduction of high-yield varieties. The increase in morbidity is observed in men but not women. We also find evidence of increased mortality among these same men 7 years later. Investigating potential mechanisms, we find evidence of dietary changes that could explain these results: large significant increases in total calories and fat consumed along with small, insignificant changes in protein consumption.

*Ramiro Burga provided superlative research assistance.

†Department of Economics, University of Virginia, email:ssekhri@virginia.edu

‡Wellesley College, 106 Central Street, Department of Economics, Wellesley, MA 02481, email: gshastry@wellesley.edu

1 Introduction

Chronic diseases are on the rise globally. According to the World Health Organization, chronic diseases, such as diabetes and heart disease, accounted for 60% of all deaths in 2001 and were projected to account for even more by 2020 (WHO 2002). In the U.S., 60% of adults have at least one chronic disease, making these diseases substantial contributors to growing health care costs.[1] Once more prevalent in the developed world, these diseases can no longer be considered "rich country diseases" - not only do the majority of chronic disease-related deaths occur in the developing world today, but rates of disease have changed so much in the past 40 years that middle-income countries, including India, now have the highest rates in the world (WHO 2016). In fact, economic growth in many developing countries has led to a double burden of malnutrition, with high rates of under-nutrition, such as stunting or micronutrient deficiencies, existing alongside growing disease burdens from chronic conditions related to over-nutrition, such as diabetes and heart disease.[2] In India, more than 35 percent of children under 5 are stunted and almost 60% are anemic, while at the same time, the rate of obesity among men and women aged 15-49 doubled from approximately 10% to 20%, from 2005 to 2015 (NFHS-4 2017). Potential causes of this global rise in chronic conditions include changes in diet and lifestyle (Popkin 2001; Swinburn et al. 2011; Yach et al. 2006). This paper studies whether childhood exposure to high-yield crops during the Green Revolution in India have contributed to this rise in chronic disease, possibly due to changes in diet.

The Green Revolution arrived in India in 1966, with the introduction of high-yield varieties of rice and wheat by the Indian government. The resulting growth in agricultural output allayed concerns among policy-makers about famine and food security, reducing food prices and increasing calorie availability. In this paper, however, we contribute to recent work suggesting that the impact of the Green Revolution on health in India

[1]https://www.cdc.gov/chronicdisease/resources/infographic/chronic-diseases.htm

[2]https://www.who.int/nutrition/double-burden-malnutrition/en/. A recent articulation of this challenge separates out micronutrient deficiencies, resulting in "triple" burden (Pinstrup-Andersen 2007).

has not been fully understood. For example, while Brainerd and Menon (2014) find that fertilizer use driven by the Green Revolution increased infant and child mortality, Bharadwaj et al. (2018) find that the Green Revolution reduced infant mortality.

We study the long-term impact of increased crop yields stemming from the Green Revolution in India on chronic diseases such as diabetes and heart disease. We exploit historical geographical variation in the thickness of underground aquifers; the crop varieties introduced in the Green Revolution required substantial irrigation (as well as fertilizer) in order to generate high yields. Using an event-study model as well as a trend break specification that accounts for pre-existing trends, we first document that adoption of high-yield varieties, fertilizer use and production of rice and wheat increased faster in districts with thicker aquifers after 1966. Next, we provide robust evidence of an increase in diabetes incidence and suggestive evidence of increased prevalence of heart disease for men born in these districts after 1966 using the India Human Development Survey from 2005. The event-study specification provides support for the identifying assumptions. Using a second wave of the survey that revisited the same households six to seven years later, we find an increase in mortality for men in these same cohorts. We provide suggestive evidence that calorie and fat consumption, but not protein consumption, are greater in households with an adult male born after 1966 in these districts and rule out differential survival and migration as possible explanations for the increase in chronic disease. Consistent with these suggestive results, Pingali (2015, 2012) and Gómez et al. (2013) note that dietary diversity may have fallen during the Green Revolution, and argue that new agricultural policy must take into account the dual burden of malnutrition.

Our study is motivated by the fetal origins literature that tests the hypothesis that improved access to nutrition in-utero has long-term benefits on health and well-being.[3] We relate most closely to the subsequent literature documenting long-term effects of income shocks during childhood (see Almond et al. 2018 for a review of the broader

[3] The hypothesis is best attributed to Barker (1990) and has spawned an active research area within economics (see Almond and Currie 2011 for a review).

literature on the long reach of the childhood environment). Papers that study welfare policies in the US consistently find positive impacts on future health outcomes (Aizer et al. 2016; Hoynes et al. 2016). A larger number of papers find positive long-term impacts on other adult outcomes, such as educational attainment, earnings and other indicators of socio-economic status (Chetty et al. 2016; Gould et al. 2011; Lavy et al. 2016; Løken et al. 2012).

In developing country settings, a number of papers focus on long-term effects of early-life health. Currie and Vogl, in their review of this literature, argue that the i) impacts may be larger in developing countries since shocks are more common, baseline health is lower, and parents may be less able to compensate for such shocks, and ii) it may be harder to study this question in settings where the bias from mortality selection is greater (2013). While many papers focus on variation in nutrition around the time of birth (see, for example, Adhvaryu et al. 2019; Almond and Mazumder 2011; Field et al. 2009; Maccini and Yang 2009; McEniry and Palloni 2010; Shah and Steinberg 2017), there are fewer papers looking at long-term impacts of early childhood (postnatal) interventions (Hoddinott et al. 2008 is a notable exception). Our paper also relates closely to the work studying the impact of famines and droughts, specifically on long-term health (Dinkelman 2017; Fung 2009; St Clair et al. 2005), and also on height and schooling attainment (Almond et al. 2007; Chen and Zhou 2007; Dercon and Porter 2014; Gørgens et al. 2012; Meng and Qian 2009; Umana-Aponte et al. 2011).[4]

Our paper contributes to the literature in several ways. First, only a few previous papers have studied the consequences of positive and persistent income shocks in a developing country. Much of the focus has been on negative shocks. Relative to identification strategies that focus on extreme events, such as famines, our study of the impact of greater access to nutrition is arguably more generalizable to the experience of economic growth in many developing countries. In addition, extreme events are more likely to lead

[4]There is also a large literature on the impact of other shocks to child health, such as disease, pollution and war, on long-term outcomes (please see Currie and Vogl 2013 for a review).

to biases from endogenous migration and mortality selection. Second, most, if not all, papers in this literature find positive impacts of increased calorie and macro-nutrient availability. We are one of the first to connect childhood income shocks to changes in diet to explain an unexpected long-term impact.[5] A central contribution of our study is that it connects the early childhood literature to the literature on the dual burden of under- and over-nutrition. We highlight long-term consequences of childhood exposure to calorie-rich low-nutrient diets, and emphasize the importance of dietary diversification. Finally, we contribute to the literature on the impact of the Green Revolution.

India is by no means alone in facing this dual burden of malnutrition and obesity - China is another example (Popkin 2001) - and the Green Revolution is only one contributing factor. While the emergence of "rich country diseases" is, in some ways, a consequence of economic growth, globalization has exacerbated the problem. A similarly dramatic rise in obesity and diabetes in Mexico has led to suggestions that free trade agreements like NAFTA cause obesity due to increases in calorie-rich low-nutrient food from the U.S. Indicative of possible changes in dietary composition, NAFTA likely increased local prices of fruits and vegetables that were exported to the U.S., while reducing prices of processed foods imported into Mexico.[6] One policy implication of our paper is to highlight that public health programs and interventions need to consider the problems of dietary diversity and the long-term consequences of changes in diet, in addition to the more traditional concerns of under-nutrition.

The paper is organized as follows. Section 2 provides background on the Green Revolution in India, while Section 3 provides a conceptual framework for how the Green Revolution may have affected dietary diversity. Section 4 describes the data. Section 5 details the estimation strategy and Section 6 describes the results on crop-yields and

[5] The study design in Hoddinott et al. (2008) - comparing a protein-rich supplement to a caloric supplement with no protein - suggests the importance of examining diet composition. The authors find positive long-term effects of the protein-rich supplement but since the two supplements also differed in the number of calories, they are unable to differentiate between the effect of increased calories and increased protein.

[6] https://www.nytimes.com/2017/12/11/health/obesity-mexico-nafta.html

adult health. Section 7 provides suggestive evidence supporting changes in diet as a mechanism and ruling out alternative explanations. Section 8 concludes.

2 The Green Revolution

We identify the impact of positive income and nutrition shocks by exploiting temporal and geographic variation in adoption of high-yield crop varieties during the Green Revolution. In this section, we provide background on the Green Revolution necessary to understand this variation.

2.1 Timeline

After Independence in 1947, the Indian government was concerned about stagnant food production and the country's increasing reliance on food imports. Various programs were designed and implemented to boost agricultural production but were ultimately unsuccessful.[7] Indigenous crop varieties were not responsive to commonly used agricultural methods, such as fertilizer. In 1963, the Indian Agriculture Research Institute invited Norman Borlaugh to India. Borlaugh, who had been successfully breeding wheat in Mexico since 1944, determined that hybrid wheat of semi-dwarf varieties would be most suitable for India.[8] Hence in 1966, the Indian government introduced the dwarf variety of wheat grown in Mexico as well as a similar high-yield variety of rice grown in the Philippines, developed by the International Rice Research Institute (Sen et al. 1974). With the aid of philanthropic donations, the government purchased 18,000 tons of wheat, enough to cover 1 million hectares. One concern was that the hybrid wheat

[7] Approximately 80 million metric tons of food was produced each year from 1958 to 1964. Favorable weather conditions led to an increase to 89 million metric tons in 1964, but it dropped back to 72-74 million metric tons in 1965 and 1966. Imports rose from 3 million metric tons in 1958 to 10 million in 1966.

[8] Norman Borlaugh received the Nobel Peace prize in 1970 for his contribution to wheat breeding. Wheat breeding resulted in significant increases in yield and produced strains resistant to common fungal diseases.

required large amounts of fertilizer and irrigation to achieve high yields. Borlaugh proved to be right and the semi-dwarf varieties succeeded in India and responded very well to high doses of commercial fertilizers. Production increased by as much as 20 times in the first five years. Subsequently, food production increased dramatically and around 62 percent of cereal production was attributed to high-yield varieties by 1975 (Herdt et al. 1985). Sekhri (2014) calculates and plots the area of land farmed with high-yield varieties each year, demonstrating the dramatic rise starting in 1966 (see Appendix Figure A6 in the online appendix to Sekhri 2014, reproduced as Figure 2 below).

2.2 Spatial Variation in Expansion

As noted above, these high-yield varieties were very sensitive to water availability. Adequate groundwater enabled farmers to irrigate their farms in times of low rainfall. Reliance on groundwater started increasing manifold, while at the same time, surface water became less important. In fact, the Green Revolution is also called the "pump revolution" by some scholars. The cultivated area under high-yield varieties increased substantially in regions with the greatest endowments of groundwater (Sekhri 2014). Punjab, Haryana, Western Uttar Pradesh, and Tamil Nadu sit on top of such aquifers and became model Green Revolution states. States with thick aquifers, such as West Bengal and Uttar Pradesh, became major producers of grains in India.

The thickness of aquifers was determined prehistorically. Jain et al (2007) document the age of various aquifers, their rock composition, and their spatial coverage in India. The earliest Indian aquifer formations date back to 3500 million years ago. The Pre-Cambrian formations are from this period. The late Pre-Cambrian formations are 600 to 1400 million years old. A tertiary system of aquifers in Northern India was formed from Himalayan upheavals. The youngest aquifers at the foothills of the Himalayas are of the Pleistocene age. Most of central India is under the Deccan traps, from the late Cretaceous to early Eocene eras. The thickness of these traps ranges from 10-50 meters in the eastern part of the country to 2000 meters in the west (SDVGHAL 1997). These different types of formations are responsible for different hydrological characteristics of the various aquifers. Much of the Deccan Plateau has sporadic aquifers, whereas most of the Indo-Gangetic plain has thicker aquifers. We discuss the geographic variation we exploit further in subsection 4.1 below.

3 Expected Effects of High-Yield Staples

There are a number of mechanisms through which the expansion of high-yield staple crops such as wheat and rice could affect long run outcomes. The most direct consequence of the Green Revolution, the increase in available food, could reduce hunger and certain forms of malnutrition during childhood. At the same time, farmers benefit from higher output and income, suggesting the possibility of other mechanisms, such as reductions in parental stress (Baranov et al. 2017; Evans and Garthwaite 2014) or increases in other types of child investments.

Most relevant to our paper is the impact of the Green Revolution on dietary diversity. While few papers have established causal linkages between the increase in high yield varieties and dietary diversity, there are a number of potential channels (Pingali 2012). The relative price of more nutritious foods to staples is likely to rise due to the increased supply of rice and wheat. Farmers may also switch out of growing non-rice foods, both because of the increased return to growing staples and because of the changes in farming practices necessary for the new high-yield varieties, exacerbating the change in relative prices and availability. For example, Cagauan and Arce (1992) note that farmers in the Philippines stopped harvesting wild leafy vegetables and fish from rice paddies due to the increased need for chemical pesticides.

The change in the relative prices of different nutrients would likely induce families to alter the composition of calories consumed. Depending on whether these households exhibit Giffen behavior with respect to staples such as rice and wheat, as Jensen and Miller (2008) found for poor households near subsistence in China, it is ambiguous as to how the Green Revolution may have impacted dietary diversity. Long-term consequences of this altered diet may be coming directly from early life nutrition for exposed individuals, but may also come indirectly via adult diet through habit-formation and the evolution of tastes (Atkin 2013). The latter is, in some ways, a more intuitive pathway from childhood diet to adult health, given the large literature connecting diet with metabolic diseases such as heart disease and diabetes.

The long-term impacts of early life nutrition depend on the types of malnutrition we consider. Insufficient intake of various micro-nutrients such as iron and vitamin D can have serious health consequences and also reduce educational attainment or earnings; an increase in (relatively nutrient-poor) staples consumption by itself would not improve these outcomes. More generally, it is now believed that the human body responds to poor nutrition in-utero or during early childhood by developing a "thrifty phenotype" which allows it to make the most out of the calories available. If the future nutrition environment is substantially more nutrient-rich, this can lead to greater incidence of diseases such as high blood pressure, type II diabetes, and heart disease (Barker 1990). In a very closely related paper, Hoynes et al. (2016) document that the introduction of the Food Stamp Program in the U.S. led to a reduction in metabolic diseases; one important difference in our study is that, in addition to a positive income shock, the Green Revolution may also have affected relative prices of various nutrients. Martorell et al. (2009), following up on the same nutrition intervention as Hoddinott et al. 2008 mentioned above, find that improved nutrition during childhood did not increase the risk of heart disease, but it is worth noting that the nutrition intervention compared a supplement providing 91 kcal energy and 6.4 g protein per 100 mL to the control group that received a supplement providing 33 kcal per 100 mL from sugar.

In sum, the long-run impact of the Green Revolution on adult health is ambiguous.

An improved nutrition environment during early childhood would be expected to improve adult health, unless changes in diet counter those benefits. Our findings on increased incidence of diabetes and greater consumption of total calories and fat but not protein is consistent with an increase in staples consumption relative to more nutrient-rich foods during the Green Revolution, leading to subsequent habit-formation.

4 Data

4.1 Aquifer, climate and agricultural data

The historical data on agricultural production and investment is from the India Agriculture and Climate Data Set (IAC), assembled by the World Bank.[9] The data cover the 1956-1987 time period and span 271 districts in 13 major Indian states. The covered states include Andhra Pradesh, Bihar, Gujarat, Haryana, Karnataka, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal.[10] The data provide information on area under high-yield and conventional varieties of various crops, use of inputs such as fertilizers, edaphic variables like soil type and quality, and geographical variables like altitude and distance from the sea.

The data also reports the depth of underground aquifers for each district. These variables originate from the Water Resources Plates produced by the Government of India's National Atlas of India (WRP-NAI). These maps show contours of water table depths. The three reported binary variables in the IAC indicate depth greater than 150 meters, between 100 and 150 meters, and less than 100 meters. These three indicators are reported for only 124 districts. We used the WRP-NAI to confirm these classifications and verify that the remaining districts have only sporadic aquifers. The various plates

[9] The World Bank collated the dataset drawing on data assembled by James McKinsey Jr. and Robert Evenson of the Yale Growth Center.

[10] All splits in districts have been accounted for in the data and the boundaries correspond to 1961 boundaries.

of the atlas are depicted in Appendix Figure A1. We matched average annual rainfall and temperature from University of Delaware data to this dataset.

Using these data, we classify Indian districts into four categories based on the thickness of available aquifers - thickest or most well-endowed (aquifers greater than 150 meters), medium thick or well-endowed (aquifer thickness between 100 and 150 meters), thick or endowed (aquifer thickness up to 100 meters), and sporadic or not-endowed (very little groundwater availability). We restrict the sample to the first three categories. Specifically, we do not compare areas with and without aquifers as these might be very different from each other, geologically. Our empirical analysis relies on comparing areas with different water endowments over time. Figure 1 presents a map of Indian districts, shaded by aquifer thickness. The 124 districts we use in all of our analysis (the ones with aquifers of any thickness) are concentrated in a the north and on the east coast of the sub-continent, but among those 124 districts, there are districts of all three aquifer classifications around the entire country. Approximately 24 percent of the districts have thick aquifers, 59 percent have medium thick aquifers and 18 percent have the thickest aquifers. Panel A of Table 1 presents summary statistics for the agricultural outcomes we use from the IAC by type of district from 1965, the year before the Green Revolution. Fertilizer use and rice and wheat production did differ, although we also show that changes in these outcomes from 1964 to 1965 did not differ significantly across the three types of districts.

4.2 Household health data

The data on chronic diseases comes from two waves of the India-Human Development Survey (IHDS) conducted in 2005 and subsequently in 2011-2012. The IHDS is a nationally representative survey which interviewed 41,554 households located in 33 states and union territories, covering 375 districts in India. The first round of interviews were completed during 2004 and 2005. The survey collected information on income, consumption, employment, health, and different aspects of gender and family relationships from both male and female respondents. We use responses to whether the individual was ever diagnosed by a doctor with heart disease, diabetes, cancer, or asthma. Data on a total of 215,753 individuals from these households were collected. A second round of the IHDS re-interviewed most of these households in 2011 and 2012 (83 percent of the original sample were re-interviewed, split households located in the same village were interviewed and 2134 new households were resampled, producing a total of 42,152 households). This allows us to construct a panel of households over time. For detailed information on the study design, please see Desai et al. (2005, 2015). Panel B of Table 1 reports summary statistics for the main variables used in the analysis for individuals born prior to 1966. There are some significant differences between the types of districts, motivating our use of a trend-break model.

4.3 Household food diaries

The data on calorie, protein, and fat consumption is computed from the Household Consumption Expenditure Surveys conducted by the National Sample Survey Organization (NSSO). The "thick" round of the survey is conducted every 5 years and is nationally representative. We use the wave conducted in 1999-2000 (the 55th round), 5 years prior to the IHDS. The survey data comprises all expenditure incurred by the household including consumption out of home-grown produce (imputed at producer prices) and out of in-kind wages, gifts, loans, free collections, all imputed at prevailing local retail prices. We use the reported quantities of specific food items consumed, convert them into calorie intake based on standard food conversion tables, and aggregate across all consumed food items.[11] Following Deaton and Drèze (2009), we use the "nutritive value of Indian foods" published by the National Institute of Nutrition (Gopalan et al. 1980) for the conversion of food items into caloric equivalents as the conversion factors have remained stable over

[11] The survey covers over 300 food items. The IHDS data also has food consumption. However, we do not rely on that as it does not provide as exhaustive coverage as the NSSO data.

time. We then used the methodology used in Subramanian and Deaton (1996) to convert the food intake into per-capita calorie, protein, and fat intake in adult equivalents. We confine ourselves to the rural sample as the Green Revolution was by and large a rural phenomenon. In the 55th round, 6046 rural villages were covered, surveying 71,385 households and 374,856 individuals within these households. Information on consumption of food was collected independently for two different reference periods of 7 days and 30 days from the same households. We use the 30 day reference period.[12]

5 Empirical Model

Our identification strategy relies on comparing cohorts born after 1966 to those born before 1966 (the year high-yield varieties were introduced in India) across areas of different historically-determined groundwater endowments.[13] We begin with an event-study specification that flexibly estimates time-varying effects of the introduction of the Green Revolution across different types of districts, conditional on year and district fixed effects.

The event-study approach has important advantages. This specification relaxes the assumption of a one-time change, allowing the effect to grow over time, diminish over time or even change non-monotonically. We consider all children born after the introduction of high-yield varieties to be 'treated' and all children born before to be 'untreated.' One challenge with this approach is that children born right before the Green Revolution will also have benefited from improved nutrition during childhood, albeit at a later age.[14] The event-study approach allows us to estimate impacts for each cohort separately.

To interpret the post-Green Revolution coefficients as causal, we rely on the identify-

[12]The survey was conducted in the following major states of India: Andhra Pradesh, Assam, Bihar, Gujarat, Haryana, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Orissa, Punjab, Rajasthan, Tamil Nadu, Uttar Pradesh, and West Bengal. It was also conducted in the Union Territories.

[13]This strategy of comparing areas with different natural resource endowments has been used to examine reductions in poverty in India (Sekhri 2014), to estimate returns to oil production in the U.S. (Michaels 2010), and to study adaptation of agriculture to climate shocks (Hornbeck and Keskin 2014).

[14]If anything, this should lead us to underestimate the impact of the Green Revolution.

ing assumption that secular trends in the outcomes are unrelated to aquifer thickness.[15] The stagnant agricultural production prior to 1966 across India helps support our identifying assumption, though it is possible that districts with thicker aquifers were trending differentially prior to the introduction of high-yield varieties. Our event-study allows us to examine pre-existing differential trends.

Following the event-study model, we also estimate a trend break model to address these concerns about identification. In addition, we take a conservative approach including district-specific time trends to deal with possible differences in pre-existing trends.[16] These approaches together enable us to identify the dietary consequences of the Green Revolution.

We begin by estimating the following event-study model:

$$\begin{aligned} Y_{it} = \; & \Sigma_{l=1962}^{1971} \beta_l^{MT} W_i^{MT} * \tau_l + \beta_{1972-87}^{MT} W_i^{MT} * DV(1972-87) \\ & + \; \Sigma_{l=1962}^{1971} \; \beta_l^{T} W_i^{T} * \tau_l + \beta_{1972-87}^{T} W_i^{T} * DV(1972-87) \\ & + \beta_1 X_{it} + \tau_t + d_i + \varepsilon_{it} \end{aligned} \tag{1}$$

where Y_{it} is an outcome of interest in district i in year t. W_i^{MT} is an indicator for whether the district has medium thick aquifers, between 100 and 150 meters, while W_i^T is an indicator for whether the district has the thickest aquifers, greater than 150 meters. The omitted category is districts with aquifers thinner than 100 meters (districts with sporadic aquifers are excluded from the sample). τ_t and d_i indicate the inclusion of year and district fixed effects to control for year-specific shocks and time-invariant district characteristics, respectively. X_{it} are time-varying district characteristics. β_l^{MT} and β_l^T are streams of coefficients estimated for each year. We group the years before 1961

[15] Recall that we omit districts with sporadic aquifers, ensuring that we are only comparing districts with some access to groundwater.

[16] Burgess and Pande (2005) use trend break models to study the consequences of rural banks for poverty in India. This strategy is similar to their approach.

together as the omitted reference category and group the years after 1972 together as well; DV(1972-87) is a dummy variable indicating this group of years. Robust standard errors are clustered at the district level for all regressions.

Finally, in order to address any concerns over differential pre-trends, we estimate a trend-break model as follows:

$$Y_{it} = \gamma_1^{FT}\ W_i^{FT}\ (T_t - 1951) + \gamma_{post}^{FT}\ W_i^{FT}\ * (T_t - 1966) * post_t\ + \gamma_2^{T}\ W_i^{T}\ (T_t - 1951)$$
$$+\gamma_{post}^{T}\ W_i^{T}\ * (T_t - 1966) * post_t + \gamma_1\ W_i^{FT}\ * post_t + \gamma_2\ W_i^{T} * post_t$$
$$+\gamma_3\ X_{it} + \tau_t + d_i + \varepsilon_{it} \qquad (2)$$

where T_t is a linear time variable, $post_t$ is an indicator that takes on the value 1 after 1966 and 0 before, and the other variables are as defined above. We estimate an overall linear trend from 1951 to 1981 (1956 to 1987 in the agricultural data) for medium thick and the thickest aquifer districts, relative to the reference districts with thick aquifers (γ_1^{FT} and γ_2^{T}). γ_{post}^{T} and γ_{post}^{MT} are the change in these trends post 1966. γ_1 and γ_2 are estimates of the intercept shifts post-1966 for the medium thick and thickest aquifer districts, respectively. This specification allows us to test for a trend-break in 1966. In some specifications, we also include district-specific trends to account for linearly time-varying unobservable differences across districts.

6 Results

6.1 Adoption of High-Yield Varieties

We estimate equation (1) for the area planted with high-yield varieties, aggregate fertilizer use, and agricultural production of rice and wheat and of conventional crops, separately, and report the coefficients on the interaction terms in Figures 2 to 5, respectively. Recall that these results control for district and year fixed effects and are estimated relative to districts with aquifers less than 100 meters thick. In addition, we control for annual average rainfall and temperature.

As is evident in Figure 2, no land was devoted to planting high-yield crop varieties prior to 1965, but starting in 1966 we see small increases in acreage, relative to districts with aquifers thinner than 100 meters, that grow substantially over time. Even though the confidence intervals overlap, the coefficients for the last post period (1972-1987) differ significantly between the thickest aquifer-districts and the medium thick aquifer districts (p-value = 0.04). Figure 3 plots the coefficients from the event-study model with aggregate fertilizer use as the dependent variable. Again, the differential increases in fertilizer use in districts with higher groundwater endowments after the Green Revolution is evident. Note, however, that some of the coefficients in the years prior to the Green Revolution are significant as well. It does appear that the trends change around 1966, highlighting the need to estimate the trend-break model, specification (2). The differences between the thickest and medium thick aquifer districts for a given year are generally significant in the post-years but not in the pre-years.

Figure 4 repeats this analysis with rice and wheat production as the dependent variable. The differential increases in grain production in the post period are statistically significant from zero for both types of district and significantly different from each other as well. Again, there appear to be differential pre-trends, motivating the use of a trend-break model and district-specific trends. Finally, Figure 5 plots the coefficients from estimating equation (1) using production of conventional crops (jowar, bajra, maize and barley), crops for which new varieties were not introduced during the Green Revolution. No pattern emerges from changes in the production of these crops.

In Table 2, we report estimates from the trend-break model, equation (2), for these agricultural outcomes. Each column is a separate regression; the bottom two rows indicate the overall difference in trend between districts with medium thick aquifers relative to districts with thinner aquifers over the entire time period. The top two rows report the change in this differential trend after the Green Revolution for the two types of districts. The middle two rows report any discrete jumps in the outcomes after the Green Revolution. Column (1) demonstrates that the area planted with high-yield varieties was exhibiting no differential trend prior to the Green Revolution, as one would expect, but then starts to increase faster in districts with more abundant aquifers afterwards. The change in trend after 1966 is statistically significant at the 5 percent level and 1 percent level for districts with medium thick and the thickest aquifers, respectively. Considering the magnitudes of the coefficients: The area planted with high-yield varieties increased by 31,000-41,000 hectares right after 1966 in groundwater-rich districts, but then continued to rise faster by 3000-7000 hectares per year.

Column (2) demonstrates that fertilizer use also experienced a statistically significant increase in trend post 1966 in ground-water rich districts. Note that, as we saw in Figure 3, the pre-existing trends are also significant at the 10 percent level. Column (3) shows that rice and wheat production was also exhibiting a positive trend even prior to the Green Revolution. In water-abundant areas, rice and wheat production was growing faster even with conventional varieties. But there was a sharp and statistically significant trend break (increase) in 1966. The yearly increase in production more than doubled in both types of districts. Since both the intercept shifts and the trend break relate to the Green Revolution, we provide an F-statistic testing the joint significance of the top four coefficients in each column at the bottom of the table. We reject the null hypothesis of no impact in the first three columns. We also find that the trend break differs between the two types of water-abundant districts in the first three columns (p-values 0.008, 0.002, 0.064). In Column (4), we demonstrate that the conventional crops do not exhibit these patterns.

The statistically significant trend breaks, conditional on pre-trends in rice and wheat production and fertilizer use, aid us in our identification. In an alternative specification, we include district-specific linear trends to account for any confounding bias emerging from these pre-trends. Table 3 shows that the results hardly change at all; note that including district-specific trends precludes us from estimating differential overall trends by type of district.

These patterns highlight that access to groundwater played an instrumental role in adoption of high-yield crop varieties in Indian agriculture and subsequently increased agricultural output, but only for rice and wheat. Farm profitability for high-yield varieties was substantially greater than conventional varieties (Foster and Rosenzweig 1995). Foster and Rosenzweig (1995) also show that irrigation assets significantly influenced adoption of high-yield varieties. All of this evidence indicates that the income profile of ground-water rich districts likely changed differentially after 1966. We now turn to examining the long-term effects for those born in this time of boom and abundance.

6.2 Long-term Health

Table 4 estimates equation (2) for chronic diseases, following the format of Table 2. We estimate cohort trends for those born between 1951 and 1981, an intercept shift in 1966 and a change in trend after 1966, separately for districts with medium thick and the thickest aquifers, relative to districts with thick aquifers. We report the results separately for males and females. We focus on heart disease (Columns 1 and 3) and diabetes (Columns 2 and 4), since these are most closely related to changes in nutrition.[17] All specifications include district and birth-year fixed effects. As above, standard errors are clustered by district.

We begin by looking at the overall cohort trends from 1951 to 1981 for men (the bottom two coefficients reported in Columns 1 and 2); differences in these trends for groundwater-rich districts are negative and usually significant for both heart disease and

[17] We also look at diseases that are less related to changes in nutrition, such as cancer and asthma, as a robustness check, described below.

diabetes prevalence. Note that this is not simply the effect of age, since we include birth-year fixed effects. For cohorts born prior to the Green Revolution, younger men in water-abundant districts are healthier than men of the same cohorts in other districts, relative to the health differences among older men. However, we then see a trend reversal for both outcomes. First, the intercept shifts in 1966 for both types of districts are positive and significant, except for heart disease in the thickest aquifer districts. In other words, heart disease and diabetes prevalence are higher for cohorts born after 1966 in these districts, by around 1.3-2 percentage points, even conditional on district and birth-year fixed effects and differential cohort trends. Second, the differential change in trend is positive and significant for medium thick aquifers for heart disease at the 10 percent level and for the thickest aquifers for diabetes at the 1 percent level. The F-statistic presented at the bottom of the table tests the joint significance for all four post-Green Revolution coefficients; we reject the null hypothesis that prevalence of heart disease and diabetes were unaffected by the Green Revolution for men. Table 5 demonstrates that these results are virtually unchanged by the addition of district-specific trends.

In Columns (3) and (4) of Table 4 (and Table 5), we estimate these models for women and find that women do not experience these same patterns in heart disease and diabetes prevalence after the Green Revolution. One possible explanation for this difference is biological differences in the age gradients of heart disease and diabetes prevalence for women versus men (recall that our sample is between the ages of 24 and 54, with those born after the Green Revolution younger than 39). Alternatively, it is important to note that there is widespread disparity between resource outlays for male and female children in India. If female children are treated differently than male children when rice and wheat become cheaper, they may develop different eating habits that could affect their health in adulthood.

Figures 6 and 7 plot event-study estimates from equation (1) for heart disease and diabetes for men. We omit confidence intervals in the interest of readability, but reproduce the figures with confidence intervals in the online Appendix. We include district and birth-year fixed effects as well as district-specific trends. Relative to district-specific trends, we see no indication of differential trends in the three types of districts in the years prior to the Green Revolution. In the post period, we see no evidence of any differences for heart disease prevalence, but find that diabetes prevalence rises in the districts with thickest aquifers relative to the districts with just thick aquifers, the omitted category. An F-test of all post 1966 coefficients across both types of districts produces a p-value of 0.044. An F-test of the joint significance of all pre-1966 coefficients produces a p-value of 0.28, supporting our identifying assumptions.

Overall, these findings indicate that the male cohorts born after the Green Revolution in areas where high-yield varieties took off experienced an increased likelihood of having heart disease and diabetes. At the time of the 2005 IHDS survey, the cohorts born after the Green Revolution were between the ages of 24 and 38. The IHDS was repeated in 2011-12 and 83 percent of the households were revisited. This gives us the opportunity to examine health outcomes for the same individuals over time. Specifically, we look at survival of these individuals between the ages of 24-38 and 30-44. We examine the probability that an individual who was interviewed in 2005 had died by the 2011-12 round of the survey. Table 6 reports estimates of the trend-break model using as the dependent variable, a dummy variable that takes the value 1 if the individual interviewed in 2005 had died by the time the household was surveyed again in 2011-12. We estimate these models separately for men and women and include district-specific trends in Columns (2) and (4). For the districts with the thickest aquifers, we find evidence of a negative overall trend, but a positive trend reversal after the Green Revolution for men. The F-test at the bottom of the table indicates that the 4 coefficients post the Green Revolution are jointly significant. In other words, there is a decline in the survival probability of males born immediately after the Green Revolution as they age from being 24 - 38 years to being 30 - 44. Our previous result indicated that these cohorts are more likely to be diabetic.

Note how similar the trend break coefficient in mortality for the thickest aquifer district (0.0046 from Table 6, Column 2) is to the trend break coefficient in diabetes prevalence for the same districts (-0.0042 from Table 5, Column 2) in magnitude.[18] We also find no trend break in the mortality of females (Columns 3 and 4 of Table 6, consistent with no change in their probability of having heart disease or diabetes in 2005.

7 Mechanisms

7.1 Changes in Diet

In this subsection, we present suggestive evidence supporting the mechanism we posit for the surprising increase in chronic disease for cohorts born after the positive income shock from the Green Revolution. Using household-level data on food consumption from the 55th round of the NSSO (1999-2000), we estimate the difference in per-capita consumption of macro-nutrients (total calories, protein and fat) in households with any adult male born after 1966 in districts with varying groundwater access.[19] We control for district fixed effects and household size in consumption units (taking into account how much individuals of different ages and genders tend to consume). Table 7 shows that, conditional on household size, households with an adult male born after 1966 tend to consume less calories per capita, but that this is reversed if the household lives in a district with medium thick or the thickest aquifers (Column 1). All coefficients are statistically significant at the 1 percent level. Column (2) shows that fat calories consumed follows the same pattern, although the reversal in districts with medium thick aquifers

[18] In results not shown, we do not discern any trend break in diabetes or heart disease prevalence in 2011 even for men. This is consistent with an increase in mortality if the healthier men from 2005 are the ones who survive. This suggests that the selection from survival may be biasing our estimates towards 0, even in the 2005 sample.

[19] We are unable to examine possible intermediary outcomes, such as weight, height, or obesity, in the IHDS because these were only measured for women in 2005 (and we find no effects for women). While these outcomes are measured for men in 2011, selection on survival biases us away from finding an effect, although as described above, we document this selection using the panel structure of the data.

is only marginally significant. In Column (3), we find that consumption of calories from protein does not increase significantly for these households. In Column (4), we examine differences in per-capita consumption of calories from nutrients other than protein and fat, that is, generally carbohydrates. We see a marginally significant increase in these calories in households with adult males born after 1966 in districts with thicker aquifers relative to districts with less thick aquifers. In Column (5), we look at the percent of calories consumed that come from protein and see corresponding statistically significant decreases. Given current nutritional guidelines linking carbohydrates to chronic diseases and the importance of protein consumption and dietary diversity, these results are illuminating. Note that these consumption differences are detected in 1999-2000, more than 30 years after the Green Revolution, suggesting that changes in diet persisted.

7.2 Alternative Explanations

Next we consider two alternative explanations for our findings. First, differential survival until 2005 could explain the differential increase in heart disease and diabetes if those who would have had these diseases had already died by 2005 in districts with less thick aquifers. Second, the Green Revolution might have lead to differential migration, which could explain our findings if migration patterns were related to both adoption of high-yield varieties and the probability of getting heart disease or diabetes. In this subsection, we provide evidence arguing against these two explanations.

7.2.1 Differential Survival

In order to examine the first alternative explanation, we use data on population by 5-year age group and gender from the 2001 Census of India, 4 years before the IHDS. If the results were driven by greater survival in districts with more groundwater access, we should also see more people in those cohorts. Given data limitations, we estimate this first using a traditional differences-in-differences model and then using a trend-break model. Table 8 presents the differences-in-differences estimates for men and women separately and for rural and urban areas separately as well. Specifically, the dependent variable is the number of people (in 1000s) in a district in a 5-year age group. A useful coincidence allows us to use this data in this manner because the 5-year age groups start at 20: 20-24, 25-29, etc. Individuals aged 35 and older were all born prior to the Green Revolution and cohorts 34 and younger were born after.

All regressions include district and age group fixed effects and Panel B includes district-specific trends (by 5-year birth cohorts). Considering the columns for men, we see that the coefficients are of inconsistent sign and usually not significant. The only statistically significant coefficients are actually of the opposite sign, possibly suggesting that even survival to age 35 is lower in districts with greater groundwater access, in line with our mortality results above. For women, we see some positive and significant coefficients but also some negative and significant coefficients. Given the lack of differential changes in chronic disease or mortality for women, we suspect these results may be more related to migration patterns. In Panel B, for example, it could be that more women stayed in rural areas in districts with medium thick aquifers instead of migrating to urban areas relative to districts with less thick aquifers. In Table 9, we estimate a trend-break model with district-specific trends, similar to Tables 3 and 5. Again, we see no evidence of greater survival of men in groundwater-rich districts which could have explained higher diabetes incidence. The results for women are even more mixed than in the differences-in-differences specification.

As noted in the introduction, Bharadwaj et al. (2018) find that the Green Revolution reduced infant mortality while Brainerd and Menon (2014) find that fertilizer use during the Green Revolution increased infant mortality. Related is the possibility that the composition of mothers who have children after the Green Revolution may have changed. Bharadwaj et al. (2018) show that the characteristics of women who give birth are unaffected by the Green Revolution, suggesting there is limited selection on this front.

7.2.2 Differential Migration

Next, we look at whether migration patterns could be driving our result. Migration patterns could bias our results if they are correlated with both aquifer thickness and the probability of getting these chronic diseases. For migration to *drive* our results, specifically, it would need to be that healthier people are moving from water-rich to water-poor districts (or alternatively that less healthy people are moving from water-poor to water-rich districts). This seems difficult to explain - our prior would be that water-rich districts would have been more attractive and moving costs would be more easily born by healthier people, resulting in the opposite relationship. We also note that migration was not very common in India during this time period (and even more recently). Nevertheless, we repeat the empirical analysis from Table 5 for non-migrant and migrant men in Appendix Table A1. The IHDS survey asks respondents, ”How many years ago did your family first come to this village/town/city?” and allows for a ”forever” option. We define those who respond ”forever” as non-migrants. Note that the variable is at the household-level; individuals born in the village may still be considered ’migrants’ if the household head was born in a different village. We believe this is the conservative designation since the household head’s decision to migrate may be related to their health or that of their relatives. Unfortunately, the IHDS does not ask migrants which district the respondent’s family came from. Reassuringly, we find that the same patterns in the non-migrant sample of males as reported in Table 5: evidence of a discrete jump and a trend break for diabetes incidence and marginal evidence for heart disease in Columns (1) and (2) of Appendix Table A1. We find similar patterns for migrant men (Columns 3 and 4), but the intercept shifts and trend breaks are not jointly significant. These tables provide suggestive evidence against this alternative explanation: the increase in chronic disease among cohorts born after the Green Revolution in water-abundant districts is not driven by in-migration of people more prone to chronic disease.

7.2.3 Other Diseases

Another approach to addressing more general concerns about differential trends in health across districts with varying aquifer thickness is to examine patterns in health conditions unrelated to diet and nutrition. Appendix Table A2 repeats the analysis of Table 5 using an indicator for having cancer or asthma as the dependent variable. If our results are driven by differential trends in health unrelated to the Green Revolution-caused increase in rice and wheat production, we might expect to see similar patterns for these health conditions as we saw for heart disease and diabetes. For example, the Green Revolution may have resulted in increased pollution which could have affected cancer rates and respiratory diseases. Specifically, increased air pollution due to rubble burning could have increased respiratory diseases. We find no evidence of changes in trend for these diseases (see Appendix Table A2). This also helps us rule out changes in diagnosis rates or health care access more broadly as an explanation for our main results.

8 Conclusion

In this paper, we examine the long-term impacts of the Green Revolution in India on adult health. We exploit geographic variation in historically-determined aquifer thickness that predicts adoption of high-yield crop varieties and subsequent growth in rice and wheat production. We find increases in the incidence of chronic diseases, specifically heart disease and diabetes, for men born after the start of the Green Revolution. These negative long-term impacts are surprising, given that most of the literature finds long-term health benefits of positive income and nutrition shocks during early childhood. We provide suggestive evidence of a possible mechanism: changes in dietary habits during childhood driven by the abundance, and lower relative prices, of rice and wheat persisted in adulthood, leading these men to consume more total calories, carbohydrates and fat but not more protein. We also provide evidence ruling out other alternative explanations for our results, such as differential survival or in-migration of those prone to chronic diseases, and differential trends in other health conditions.

These findings have implications for agricultural policy and welfare policy, not just for India, but other developing countries as well. As noted in the introduction, many developing countries are facing the dual burden of malnutrition and chronic diseases related to over-consumption. Policies that impact food production and distribution need to take into account dietary diversity and changes in diet that may have long-lasting negative effects.

References

Adhvaryu, A., J. Fenske, and A. Nyshadham (2019). Early life circumstance and adult mental health. *Journal of Political Economy*.

Aizer, A., S. Eli, J. Ferrie, and A. Lleras-Muney (2016). The long-run impact of cash transfers to poor families. *American Economic Review 106*(4), 935–71.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of economic perspectives 25*(3), 153–72.

Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature 56*(4), 1360–1446.

Almond, D., L. Edlund, H. Li, and J. Zhang (2007). Long-term effects of the 1959-1961 china famine: Mainland china and hong kong. Technical report, National Bureau of Economic Research.

Almond, D. and B. Mazumder (2011). Health capital and the prenatal environment: the effect of ramadan observance during pregnancy. *American Economic Journal: Applied Economics 3*(4), 56–85.

Atkin, D. (2013). Trade, tastes, and nutrition in india. *American economic review 103*(5), 1629–63.

Baranov, V., S. R. Bhalotra, P. Biroli, and J. Maselko (2017). Maternal depression, womens empowerment, and parental investment: Evidence from a large randomized control trial. Technical report, IZA Discussion Papers.

Barker, D. J. (1990). The fetal and infant origins of adult disease. *BMJ: British Medical Journal 301*(6761), 1111.

Bharadwaj, P., J. Fenske, R. A. Mirza, N. Kala, et al. (2018). The green revolution and infant mortality in india. Technical report, University of Warwick, Department of Economics.

Brainerd, E. and N. Menon (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.

Burgess, R. and R. Pande (2005). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review 95*(3), 780–795.

Cagauan, A. G. and R. G. Arce (1992). Overview of pesticide use in rice-fish farming in southeast asia. *Rice-fish research and development in Asia 3*(14,752), 217.

Chen, Y. and L.-A. Zhou (2007). The long-term health and economic consequences of the 1959–1961 famine in china. *Journal of health economics 26*(4), 659–681.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review 106*(4), 855–902.

Currie, J. and T. Vogl (2013). Early-life health and adult circumstance in developing countries. *Annu. Rev. Econ. 5*(1), 1–36.

Deaton, A. and J. Drèze (2009). Food and nutrition in india: facts and interpretations. *Economic and political weekly*, 42–65.

Dercon, S. and C. Porter (2014). Live aid revisited: long-term impacts of the 1984 ethiopian famine on children. *Journal of the European Economic Association 12*(4), 927–948.

Desai, S., A. Dubey, and R. Vanneman (2005). India human development survey-i (ihds-i) technical paper no. 1 design and data quality university of maryland and national council of applied economic research. *New Delhi*.

Desai, S., A. Dubey, and R. Vanneman (2015). India human development survey-ii (ihds-ii)[computer file]. university of maryland and national council of applied economic research, new delhi [producers], 2015. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor)*.

Dinkelman, T. (2017). Long-run health repercussions of drought shocks: Evidence from south african homelands. *The Economic Journal 127*(604), 1906–1939.

Evans, W. N. and C. L. Garthwaite (2014). Giving mom a break: The impact of higher eitc payments on maternal health. *American Economic Journal: Economic Policy 6*(2), 258–90.

Field, E., O. Robles, and M. Torero (2009). Iodine deficiency and schooling attainment in tanzania. *American Economic Journal: Applied Economics 1*(4), 140–69.

Foster, A. D. and M. R. Rosenzweig (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of political Economy 103*(6), 1176–1209.

Fung, W. (2009). Early childhood malnutrition and adult obesity: Evidence from the 1959-61 china famine. *Harvard University, Mimeograph*.

Gómez, M. I., C. B. Barrett, T. Raney, P. Pinstrup-Andersen, J. Meerman, A. Croppenstedt, B. Carisma, and B. Thompson (2013). Post-green revolution food systems and the triple burden of malnutrition. *Food Policy 42*, 129–138.

Gopalan, C., B. Rama Sastri, S. Balasubramanian, et al. (1980). Nutrition value of indian foods.

Gørgens, T., X. Meng, and R. Vaithianathan (2012). Stunting and selection effects of famine: A case study of the great chinese famine. *Journal of Development Economics 97*(1), 99–111.

Gould, E. D., V. Lavy, and M. D. Paserman (2011). Sixty years after the magic carpet ride: The long-run effect of the early childhood environment on social and economic outcomes. *The Review of Economic Studies 78*(3), 938–973.

Herdt, R. W., R. Barker, and B. Rose (1985). *The rice economy of Asia*, Volume 1. Resources for the Future.

Hoddinott, J., J. A. Maluccio, J. R. Behrman, R. Flores, and R. Martorell (2008). Effect of a nutrition intervention during early childhood on economic productivity in guatemalan adults. *The lancet 371*(9610), 411–416.

Hornbeck, R. and P. Keskin (2014). The historically evolving impact of the ogallala aquifer: Agricultural adaptation to groundwater and drought. *American Economic Journal: Applied Economics 6*(1), 190–219.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016). Long-run impacts of childhood access to the safety net. *American Economic Review 106*(4), 903–34.

Jensen, R. T. and N. H. Miller (2008). Giffen behavior and subsistence consumption. *American Economic Review 98*(4), 1553–77.

Lavy, V., A. Schlosser, and A. Shany (2016). Out of africa: Human capital consequences of in utero conditions. Technical report, National Bureau of Economic Research.

Løken, K. V., M. Mogstad, and M. Wiswall (2012). What linear estimators miss: The effects of family income on child outcomes. *American Economic Journal: Applied Economics 4*(2), 1–35.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review 99*(3), 1006–26.

Martorell, R., P. Melgar, J. A. Maluccio, A. D. Stein, and J. A. Rivera (2009). The nutrition intervention improved adult human capital and economic productivity. *The Journal of nutrition 140*(2), 411–414.

McEniry, M. and A. Palloni (2010). Early life exposures and the occurrence and timing of heart disease among the older adult puerto rican population. *Demography 47*(1), 23–43.

Meng, X. and N. Qian (2009). The long term consequences of famine on survivors: evidence from a unique natural experiment using china's great famine. Technical report, National Bureau of Economic Research.

Michaels, G. (2010). The long term consequences of resource-based specialisation. *The Economic Journal 121*(551), 31–57.

NFHS-4 (2017). India fact sheet, nfhs-4 (national family health survey-4). *International Institute for Population Studies*.

Pingali, P. (2015). Agricultural policy and nutrition outcomes–getting beyond the preoccupation with staple grains. *Food Security 7*(3), 583–591.

Pingali, P. L. (2012). Green revolution: impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences 109*(31), 12302–12308.

Pinstrup-Andersen, P. (2007). Agricultural research and policy for better health and nutrition in developing countries: a food systems approach. *Agricultural Economics 37*, 187–198.

Popkin, B. M. (2001). The nutrition transition and obesity in the developing world. *The Journal of nutrition 131*(3), 871S–873S.

SDVGHAL, B. (1997). Hydrogeological characteristics of deccan trap formations of india. *Hard Rock Hydrosystems* (241), 75.

Sekhri, S. (2014). Wells, water, and welfare: the impact of access to groundwater on rural poverty and conflict. *American Economic Journal: Applied Economics 6*(3), 76–102.

Sen, B. et al. (1974). *The green revolution in India. A perspective.* Wiley Eastern Private Ltd.

Shah, M. and B. M. Steinberg (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy 125*(2), 527–561.

St Clair, D., M. Xu, P. Wang, Y. Yu, Y. Fang, F. Zhang, X. Zheng, N. Gu, G. Feng, P. Sham, et al. (2005). Rates of adult schizophrenia following prenatal exposure to the chinese famine of 1959-1961. *Jama 294*(5), 557–562.

Subramanian, S. and A. Deaton (1996). The demand for food and calories. *Journal of political economy 104*(1), 133–162.

Swinburn, B. A., G. Sacks, K. D. Hall, K. McPherson, D. T. Finegood, M. L. Moodie, and S. L. Gortmaker (2011). The global obesity pandemic: shaped by global drivers and local environments. *The Lancet 378*(9793), 804–814.

Umana-Aponte, M. et al. (2011). *Long-term effects of a nutritional shock: the 1980 famine of Karamoja, Uganda.* Centre for Market and Public Organisation, University of Bristol.

WHO (2002). *The world health report 2002: reducing risks, promoting healthy life.* World Health Organization.

WHO (2016). *Global report on diabetes.* World Health Organization.

Yach, D., D. Stuckler, and K. D. Brownell (2006). Epidemiologic and economic consequences of the global epidemics of obesity and diabetes. *Nature medicine 12*(1), 62.

Figure 1: Map of Indian Districts Shaded by Aquifer Thickness

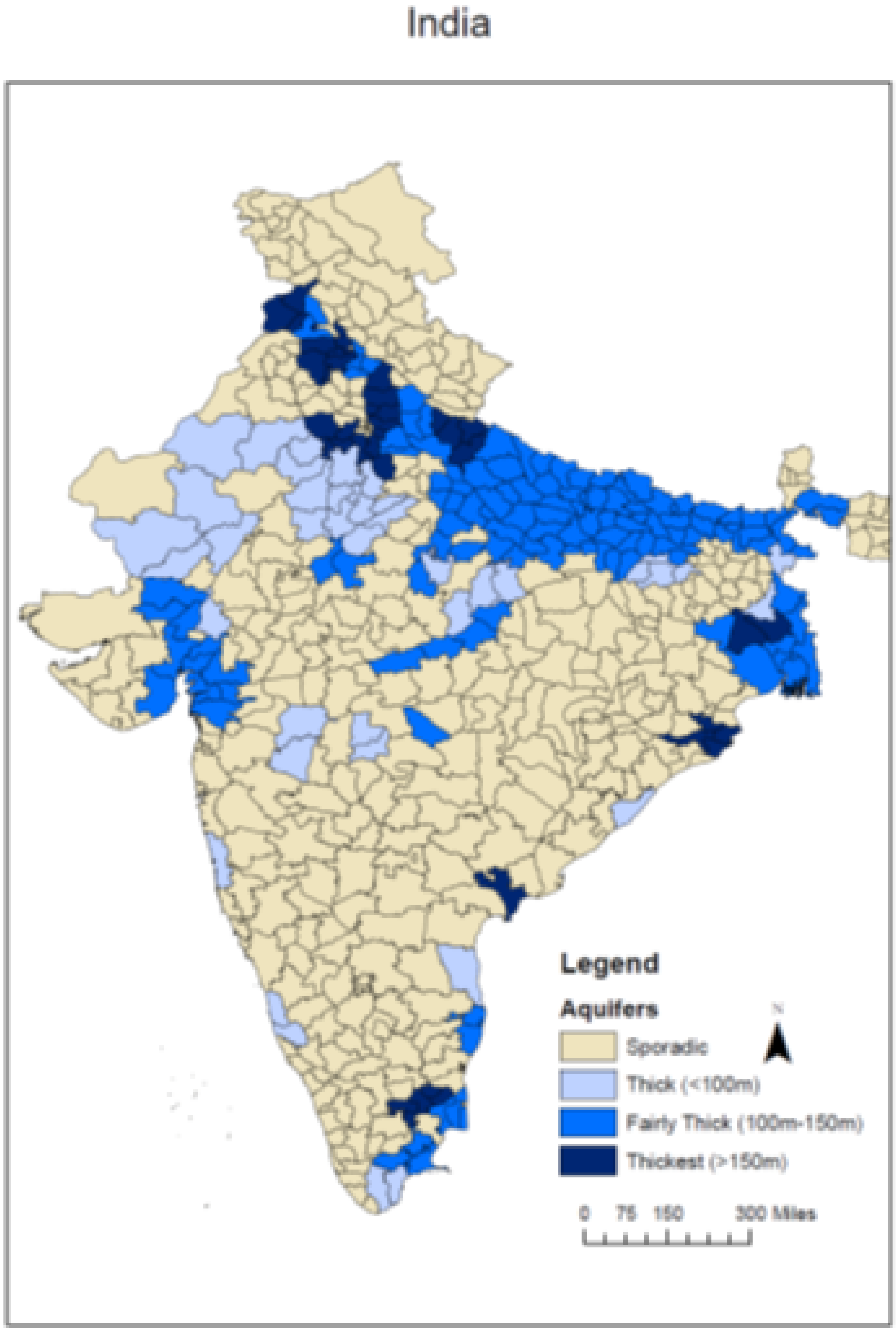


Figure 2: Event-study Estimates for Area Planted with High-Yield Varieties

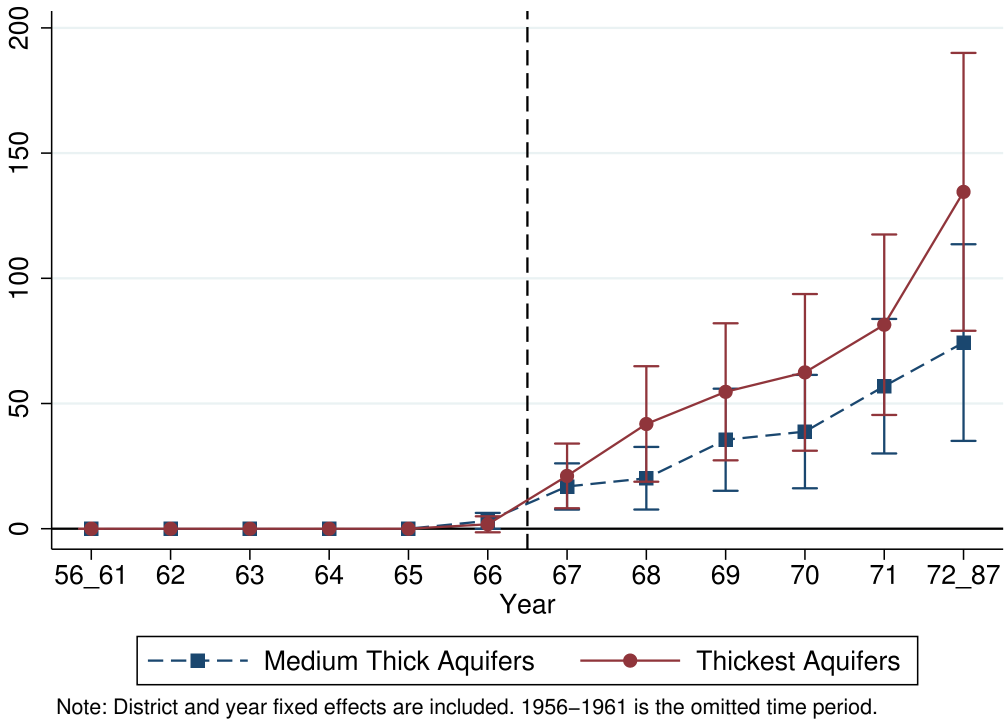


Figure 3: Event-study Estimates for Fertilizer Use

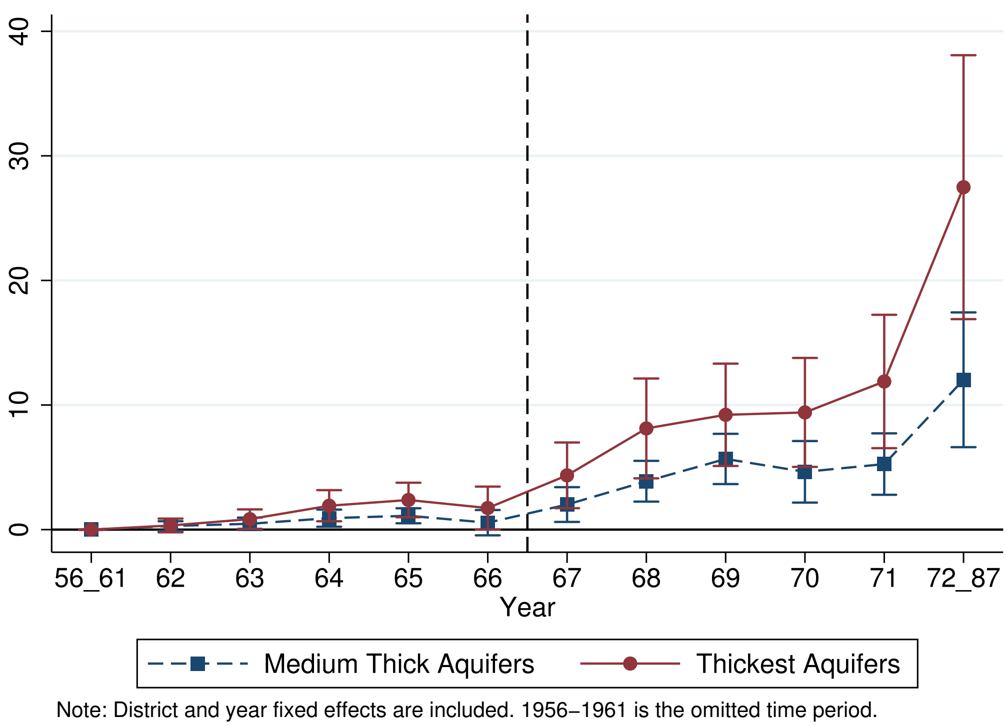
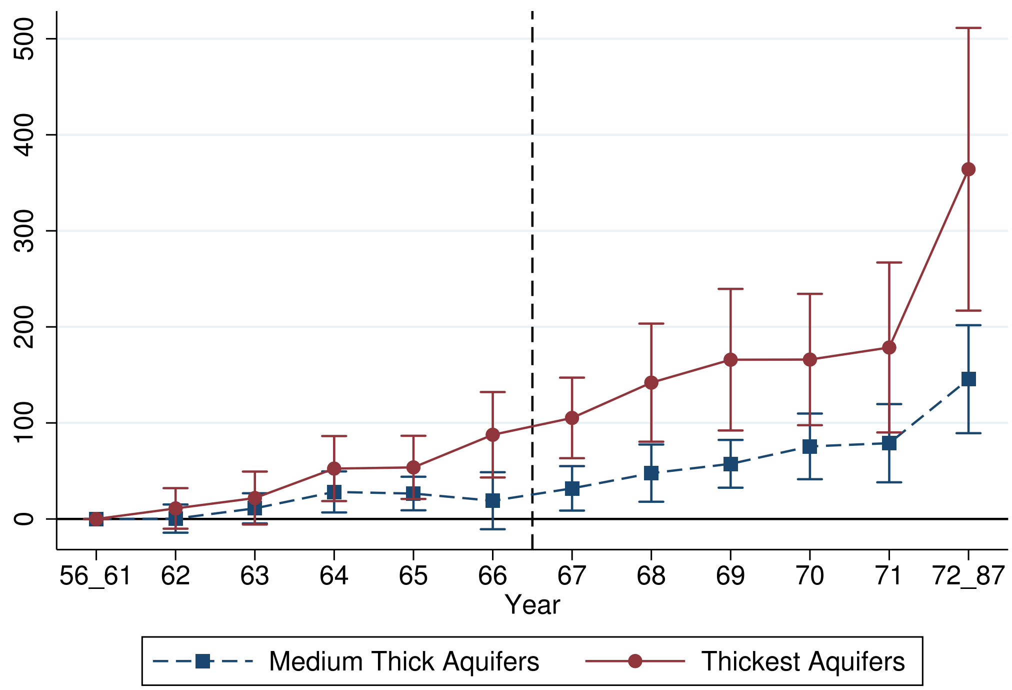
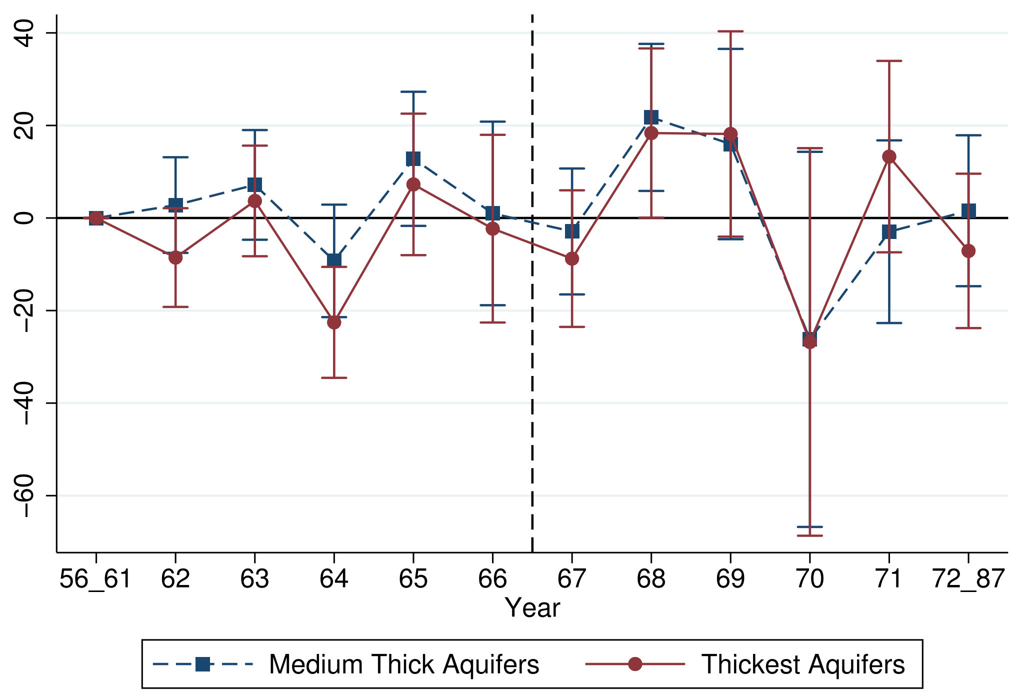


Figure 4: Event-study Estimates for Rice and Wheat Production



Note: District and year fixed effects are included. 1956–1961 is the omitted time period.

Figure 5: Event-study Estimates for Other Crops Production



Note: District and year fixed effects are included. 1956–1961 is the omitted time period.

Figure 6: Event-study Estimates for Impact on Heart Disease for Men

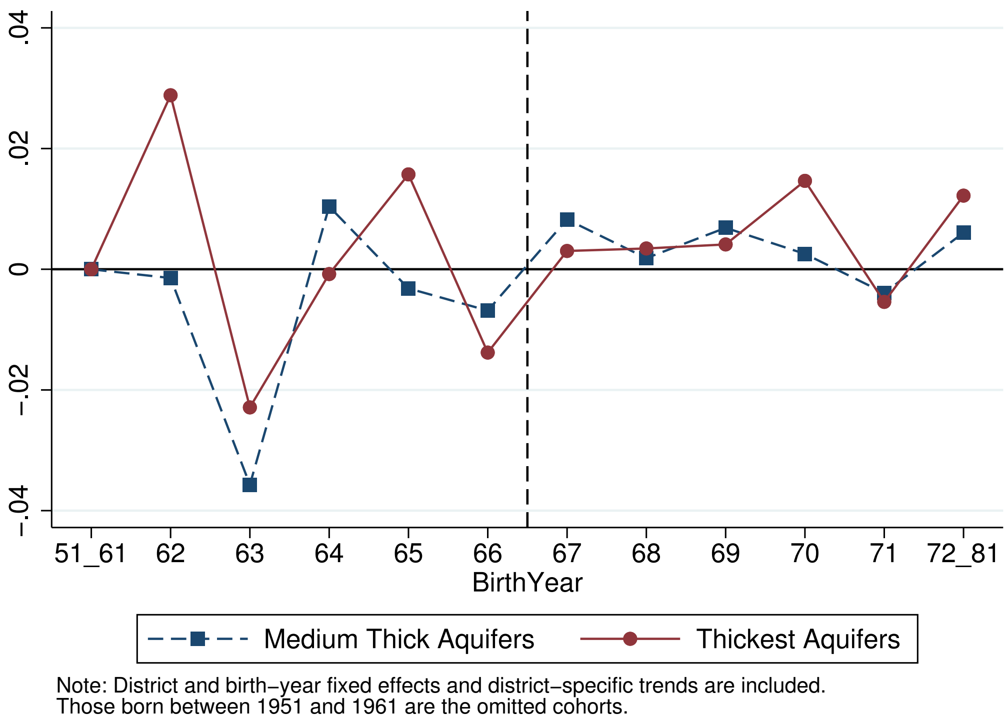


Figure 7: Event-study Estimates for Impact on Diabetes for Men

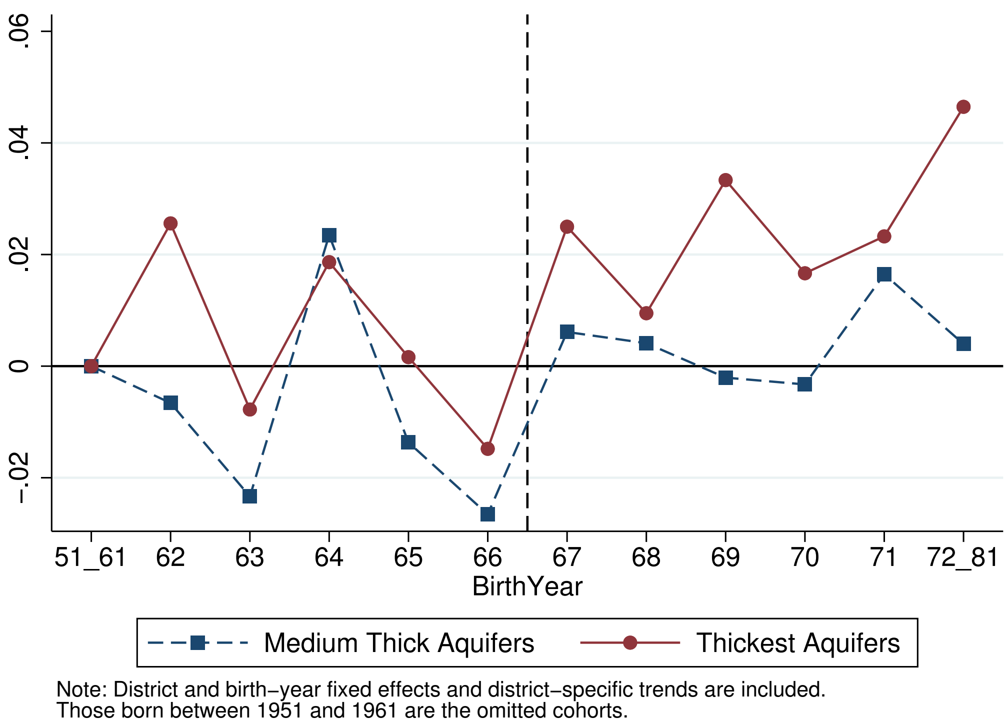


Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
				Medium		P-value Of
	N	All	Thick	Thick	Thickest	All Differences
Panel A: Agricultural Outcomes						
In 1965						
HYV area	123	0.00	0.00	0.00	0.00	.
Fertilizer Use	123	2.44	1.06	2.50	4.08	0.000
Rice & wheat production	123	171.33	90.14	184.92	236.34	0.001
Other crops production	123	48.49	53.60	47.56	42.53	0.657
Change from 1964 to 1965						
Fertilizer Use	123	-0.10	-0.30	-0.10	0.16	0.193
Rice & wheat production	123	-31.57	-30.98	-32.66	-29.74	0.949
Panel B: Demographic and Health Outcomes, Cohorts Born Before 1966						
Males						
Age	6088	45.753	45.834	45.676	45.855	0.353
Literate	6079	0.714	0.677	0.719	0.742	0.250
Married	6086	0.983	0.988	0.983	0.977	0.174
Heart disease	6088	0.009	0.007	0.008	0.014	0.143
Diabetes	6088	0.019	0.012	0.016	0.034	0.024
Cancer	6088	0.002	0.002	0.002	0.002	0.947
Asthma	6088	0.009	0.011	0.009	0.008	0.638
Females						
Age	5645	45.753	45.800	45.715	45.793	0.846
Literate	5627	0.407	0.310	0.426	0.472	0.002
Married	5643	0.995	0.996	0.994	0.995	0.568
Heart disease	5645	0.019	0.013	0.021	0.020	0.223
Diabetes	5645	0.018	0.012	0.019	0.021	0.289
Cancer	5645	0.002	0.003	0.002	-0.000	0.005
Asthma	5645	0.011	0.006	0.012	0.013	0.125

Table 2: Trend break estimates for agricultural outcomes

VARIABLES	(1) HYV area	(2) Fertilizer Use	(3) Rice & wheat production	(4) Other crops production
Change in trend: (1967-1987) trend				
x medium thick aquifers	2.78**	0.83***	6.75***	-0.21
	(1.22)	(0.20)	(1.89)	(1.26)
x thickest aquifers	7.15***	1.97***	17.00***	-0.37
	(1.65)	(0.38)	(5.41)	(1.43)
Intercept shift: Post-1966				
x medium thick aquifers	31.44***	-1.02	-11.06	1.82
	(9.38)	(0.71)	(10.67)	(5.34)
x thickest aquifers	40.83***	-2.18*	-14.28	16.00*
	(12.37)	(1.24)	(15.98)	(8.82)
Overall trend: (1956-1987) trend				
x medium thick aquifers	-0.00	0.08*	2.77***	0.05
	(0.16)	(0.05)	(0.85)	(0.80)
x thickest aquifers	-0.33	0.17*	7.30***	-0.93
	(0.23)	(0.09)	(2.04)	(0.84)
Observations	3,936	3,936	3,936	3,936
R-squared	0.78	0.78	0.82	0.71
F-test (p-value)	0.000	0.000	0.001	0.279

Note: Each column presents the results from estimating equation (2) using an agricultural outcome from the IAC as the dependent variable. The first two rows are estimates of the changes in trend post 1966, the second two rows indicate the coefficient on the intercept shift in 1966 and the last two rows present estimates of pre-existing trends. All columns include district and year fixed effects as well as controls for average annual rainfall and average annual temperature. Standard errors clustered at the district level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: Trend break estimates for agricultural outcomes with district trends

VARIABLES	(1) HYV area	(2) Fertilizer Use	(3) Rice & wheat production	(4) Other crops production
Change in trend: (1967-1987) trend				
x medium thick aquifers	2.72**	0.82***	6.49***	-0.21
	(1.17)	(0.19)	(1.80)	(1.29)
x thickest aquifers	6.92***	1.95***	16.44***	-0.31
	(1.60)	(0.38)	(5.35)	(1.45)
Intercept shift: Post-1966				
x medium thick aquifers	31.42***	-1.03	-11.30	1.80
	(9.42)	(0.72)	(10.56)	(5.44)
x thickest aquifers	36.52***	-2.53**	-25.64*	16.88*
	(11.43)	(1.26)	(15.04)	(9.08)
Observations	3,936	3,936	3,936	3,936
R-squared	0.95	0.95	0.95	0.76
F-test (p-value)	0.000	0.000	0.000	0.236

Note: Each column presents the results from estimating equation (2) using an agricultural outcome from the IAC as the dependent variable. The first two rows are estimates of the changes in trend post 1966 and the second two rows indicate the coefficient on the intercept shift in 1966. All columns include district and year fixed effects, district-specific trends in birth-year as well as controls for average annual rainfall and average annual temperature. Standard errors clustered at the district level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 4: Trend break estimates for long-term health

	(1)	(2)	(3)	(4)
	Males		Females	
VARIABLES	Heart disease	Diabetes	Heart disease	Diabetes
Change in trend: (1967-1981) trend				
x medium thick aquifers	0.0012*	0.0018	-0.0012	-0.00041
	(0.00062)	(0.0012)	(0.00076)	(0.00081)
x thickest aquifers	0.00072	0.0043***	0.00087	-0.00041
	(0.0011)	(0.0014)	(0.0010)	(0.0013)
Intercept shift: Post-1966				
x medium thick aquifers	0.013***	0.015**	0.0040	0.0038
	(0.0047)	(0.0074)	(0.0071)	(0.0062)
x thickest aquifers	0.0029	0.020**	0.0033	-0.00097
	(0.0073)	(0.0092)	(0.0070)	(0.0091)
Overall trend: (1951-1981) trend				
x medium thick aquifers	-0.0015***	-0.0021*	0.00036	-0.00018
	(0.00049)	(0.0012)	(0.00072)	(0.00079)
x thickest aquifers	-0.0011	-0.0048***	-0.00092	-0.00013
	(0.00099)	(0.0015)	(0.00097)	(0.0012)
Observations	15,159	15,159	14,948	14,948
R-squared	0.019	0.038	0.023	0.027
F-test (p-value)	0.025	0.040	0.136	0.856

Note: Each column presents the results from estimating equation (2) using a health outcome from the IHDS 2005 as the dependent variable. The first two rows are estimates of the changes in trend post 1966, the second two rows indicate the coefficient on the intercept shift in 1966 and the last two rows present estimates of pre-existing trends. The sample is restricted to individuals born between 1951 and 1981. All columns include district and birth-year fixed effects. Standard errors clustered at the district level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: Trend break estimates for long-term health with district trends

	(1)	(2)	(3)	(4)
	Males		Females	
VARIABLES	Heart disease	Diabetes	Heart disease	Diabetes
Change in trend: (1967-1981) trend				
x medium thick aquifers	0.0013**	0.0016	-0.0012	-0.00037
	(0.00061)	(0.0012)	(0.00079)	(0.00084)
x thickest aquifers	0.00088	0.0042***	0.00090	-0.00037
	(0.0011)	(0.0014)	(0.0011)	(0.0013)
Intercept shift: Post-1966				
x medium thick aquifers	0.013***	0.016**	0.0050	0.0031
	(0.0048)	(0.0075)	(0.0072)	(0.0064)
x thickest aquifers	0.0029	0.020**	0.0033	-0.0014
	(0.0074)	(0.0092)	(0.0072)	(0.0093)
Observations	15,159	15,159	14,948	14,948
R-squared	0.034	0.062	0.036	0.044
F-test (p-value)	0.016	0.040	0.127	0.905

Note: Each column presents the results from estimating equation (2) using a health outcome from the IHDS 2005 as the dependent variable. The first two rows are estimates of the changes in trend post 1966 and the second two rows indicate the coefficient on the intercept shift in 1966. The sample is restricted to individuals born between 1951 and 1981. All columns include district and birth-year fixed effects as well as district-specific trends. Standard errors clustered at the district level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Trend Break Estimates of Mortality from 2005 to 2011

	(1)	(2)	(3)	(4)
	Males		Females	
VARIABLES	Mortality	Mortality	Mortality	Mortality
Change in trend: (1967-1981) trend				
x medium thick aquifers	0.00044	0.00033	0.00097	0.00076
	(0.0017)	(0.0017)	(0.0013)	(0.0014)
x thickest aquifers	0.0047**	0.0046**	-0.0011	-0.0012
	(0.0020)	(0.0020)	(0.0015)	(0.0016)
Intercept shift: Post-1966				
x medium thick aquifers	0.0029	0.0027	-0.0026	-0.0032
	(0.016)	(0.016)	(0.0090)	(0.0091)
x thickest aquifers	0.025	0.024	-0.023*	-0.022*
	(0.019)	(0.019)	(0.013)	(0.013)
Overall trend: (1951-1981) trend				
x medium thick aquifers	0.00019		-0.0012	
	(0.0016)		(0.0012)	
x thickest aquifers	-0.0041**		0.0012	
	(0.0018)		(0.0014)	
Observations	15,159	15,159	14,948	14,948
R-squared	0.028	0.038	0.022	0.037
F-test (p-value)	0.050	0.055	0.370	0.422

Note: Each column presents the results from estimating equation (2) using an indicator for whether an individual surveyed in the IHDS 2005 had died before the IHDS 2011 survey as the dependent variable. The first two rows are estimates of the changes in trend post 1966, the second two rows indicate the coefficient on the intercept shift in 1966 and the last two rows present estimates of pre-existing trends. The sample is restricted to individuals born between 1951 and 1981. All columns include district and birth-year fixed effects. Columns (2) and (4) also include district-specific trends. Standard errors clustered at the district level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Green Revolution and Per Capita Calorie Consumption

	(1)	(2)	(3)	(4)	(5)
	Calories from				
				Total	Percent
VARIABLES	Total	Fat	Protein	-Protein-Fat	Protein
Any adult male born after 1966	-213***	-98.2**	-14.1**	-101***	0.00095***
	(53.3)	(46.2)	(6.28)	(30.8)	(0.00033)
x medium thick aquifers	157***	84.4*	5.88	67.0*	-0.0011***
	(60.0)	(49.6)	(6.87)	(35.1)	(0.00037)
x thickest aquifers	211***	117**	14.4	79.3*	-0.00095*
	(71.4)	(49.8)	(10.6)	(47.8)	(0.00051)
Observations	44,444	44,444	44,444	44,444	44,444
R-squared	0.041	0.022	0.052	0.064	0.295

Note: This table estimates the differential impact of having an adult male born after 1966 in groundwater-rich districts on calorie consumption from the NSSO from 1999-2000. All columns include district fixed effects and a control for household size in consumption units. Standard errors clustered at the district level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: DID Estimates of Population Size

	(1)	(2)	(3)	(4)	(5)	(6)
		Males			Females	
VARIABLES	Rural	Urban	Total	Rural	Urban	Total
Panel A						
Post x medium thick aquifers	1.85	1.50	3.35	3.96	1.39	5.34*
	(1.84)	(1.80)	(2.93)	(2.40)	(1.72)	(3.20)
Post x thickest aquifers	-0.20	1.96	1.76	-2.21	1.63	-0.58
	(2.17)	(1.87)	(3.50)	(2.75)	(1.79)	(3.88)
Observations	1,194	1,194	1,194	1,194	1,194	1,194
R-squared	0.964	0.961	0.963	0.942	0.953	0.948
Panel B						
Post x medium thick aquifers	-0.84	-0.93	-1.77	1.89**	-0.76*	1.13
	(0.81)	(0.59)	(1.16)	(0.96)	(0.41)	(1.18)
Post x thickest aquifers	-1.48	-1.87**	-3.35**	-0.88	-1.18**	-2.06
	(1.04)	(0.84)	(1.67)	(1.20)	(0.57)	(1.50)
Observations	1,194	1,194	1,194	1,194	1,194	1,194
R-squared	0.995	0.998	0.996	0.993	0.996	0.994

Note: This table presents differences-in-differences estimates using population size (in 1000s) from the 2001 Census, by gender and 5-year age group, as the dependent variable. The sample is restricted to individuals born between 1952 and 1981. All columns include district and age group fixed effects. Panel B also includes district-specific trends (by 5-year birth cohort). Standard errors clustered at the district level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Trend Break Estimates of Population Size

	(1)	(2)	(3)	(4)	(5)	(6)
		Males			Females	
VARIABLES	Rural	Urban	Total	Rural	Urban	Total
Change in trend: (1967-1981) trend						
x medium thick aquifers	-2.24*	0.10	-2.14	-2.15**	-0.80*	-2.94**
	(1.14)	(0.45)	(1.36)	(0.95)	(0.44)	(1.19)
x thickest aquifers	0.75	0.32	1.07	-0.62	-0.91*	-1.54
	(1.43)	(0.46)	(1.60)	(1.14)	(0.50)	(1.43)
Intercept shift: Post-1967						
x medium thick aquifers	0.28	-0.97	-0.70	2.96**	-0.36	2.60*
	(1.05)	(0.73)	(1.49)	(1.16)	(0.39)	(1.39)
x thickest aquifers	-1.86	-2.03**	-3.89*	-0.56	-0.73	-1.29
	(1.31)	(0.99)	(1.98)	(1.44)	(0.58)	(1.70)
Observations	1,194	1,194	1,194	1,194	1,194	1,194
R-squared	0.99	1.00	1.00	0.99	1.00	0.99
F-test	0.045	0.162	0.031	0.013	0.056	0.017

Note: Each column presents the results from estimating equation (2) using population size (in 1000s) from the 2001 Census, by gender and 5-year age group, as the dependent variable. The first two rows are estimates of the changes in trend post 1966 and the second two rows indicate the coefficient on the intercept shift in 1966. The sample is restricted to individuals born between 1952 and 1981. All columns include district and age group fixed effects as well as district-specific trends (by 5-year birth cohort). Standard errors clustered at the district level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

9 Appendix Figures and Tables

Figure A1: Water Resources Plates, National Atlas of India

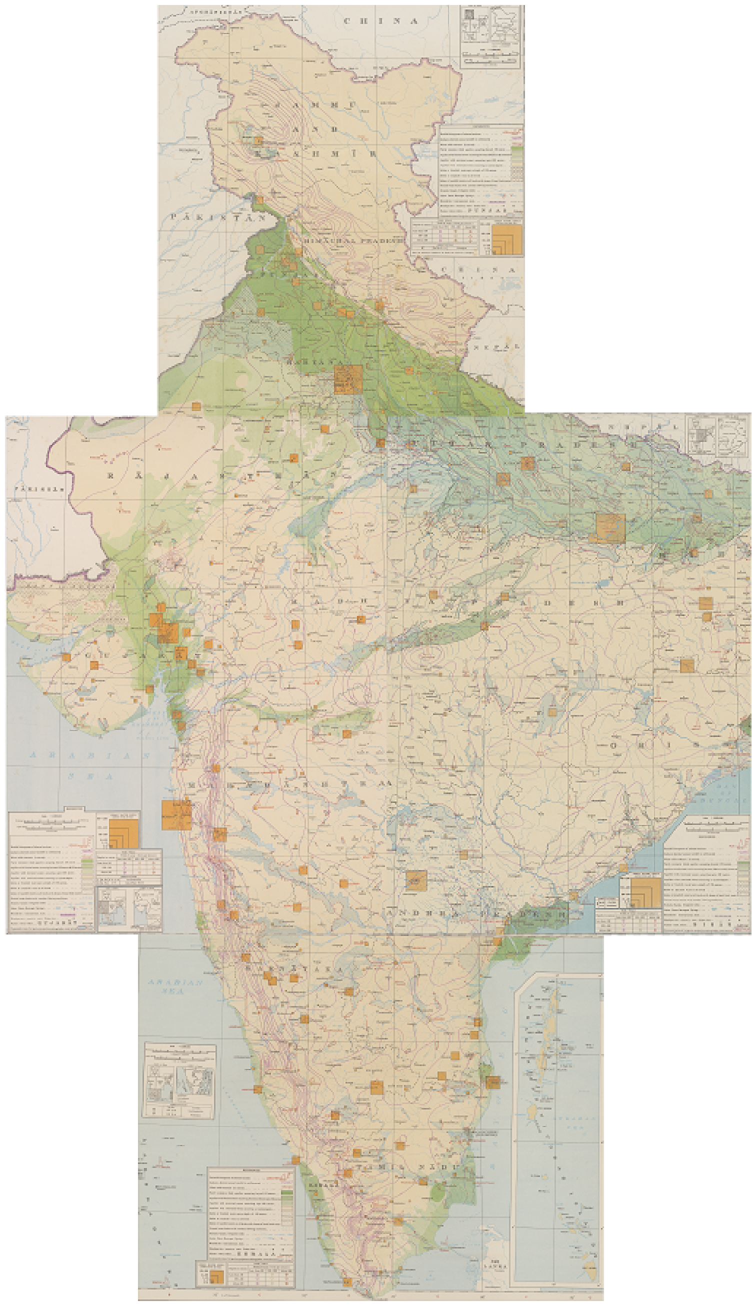


Table A1: Trend break estimates for long-term health, by migration status

	(1)	(2)	(3)	(4)
	Non-Migrant Men		Migrant Men	
VARIABLES	Heart disease	Diabetes	Heart disease	Diabetes
Change in trend: (1967-1981) trend				
x medium thick aquifers	0.0010	0.00093	0.0019*	0.0034
	(0.00068)	(0.0010)	(0.0011)	(0.0030)
x thickest aquifers	0.00028	0.0041***	0.0024**	0.0037
	(0.0013)	(0.0015)	(0.0011)	(0.0036)
Intercept shift: Post-1966				
x medium thick aquifers	0.013**	0.014*	0.011	0.013
	(0.0054)	(0.0071)	(0.013)	(0.025)
x thickest aquifers	0.00024	0.019*	0.0053	0.029
	(0.0089)	(0.010)	(0.010)	(0.032)
Observations	11,590	11,590	3,568	3,568
R-squared	0.042	0.067	0.068	0.117
F-test (p-value)	0.105	0.046	0.156	0.698

Note: Each column presents the results from estimating equation (2) using a health outcome from the IHDS 2005 as the dependent variable. The sample is restricted to non-migrant men in Columns (1) and (2) and to migrant men in Columns (3) and (4). The first two rows are estimates of the changes in trend post 1966 and the second two rows indicate the coefficient on the intercept shift in 1966. The sample is restricted to individuals born between 1951 and 1981. All columns include district and birth-year fixed effects as well as district-specific trends. Standard errors clustered at the district level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Trend break estimates for long-term health conditions unrelated to diet

	(1)	(2)	(3)	(4)
	Males		Females	
VARIABLES	Cancer	Asthma	Cancer	Asthma
Change in trend: (1967-1981) trend				
x medium thick aquifers	-0.00013	-0.00071	0.00041	-0.00034
	(0.00026)	(0.0011)	(0.00046)	(0.00090)
x thickest aquifers	0.00030	-0.0018	0.00034	-0.0015
	(0.00030)	(0.0011)	(0.00034)	(0.0011)
Intercept shift: Post-1966				
x medium thick aquifers	-0.00079	-0.0041	0.0015	-0.0066
	(0.0026)	(0.0049)	(0.0043)	(0.0062)
x thickest aquifers	0.0015	0.00041	0.0020	-0.013
	(0.0022)	(0.0072)	(0.0038)	(0.0095)
Observations	15,159	15,159	14,948	14,948
R-squared	0.034	0.029	0.022	0.034
F-test (p-value)	0.685	0.247	0.821	0.612

Note: Each column presents the results from estimating equation (2) using a health outcome from the IHDS 2005 as the dependent variable. The first two rows are estimates of the changes in trend post 1966 and the second two rows indicate the coefficient on the intercept shift in 1966. The sample is restricted to individuals born between 1951 and 1981. All columns include district and birth-year fixed effects as well as district-specific trends. Standard errors clustered at the district level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.