Technical Change and Gender Wage Inequality: Long-Run Effects of India's Green Revolution*

Version: January 13, 2017

Anthony Louis D'Agostino†

New technologies can have distributional consequences that narrow or widen existing inequalities. In this paper, I estimate the effect of India's Green Revolution on the gender wage gap in village labor markets. Widely considered the country's most important episode of agricultural technical change, productivity gains from the Green Revolution were due largely to the introduction of high-yielding variety seeds (HYVs). Their sensitivity to water conditions advantaged locations with reliable irrigation access, which I exploit by constructing a novel dataset of historical groundwater resources as a source of exogenous variation in Green Revolution adoption. Using data from 1956-1987 in a difference-in-differences framework, I estimate that groundwater-rich districts experienced large and significant increases in HYV adoption, crop revenues, and cropping intensity. These productivity gains affected labor markets, with male wages rising, but female wages declining, jointly raising the male wage premium an average 17 percent. Additionally, census and microdata results show women substituted away from wage work, and increased their time in unpaid own-farm work and home production. These outcomes are most likely driven by the sharp yield gains achieved with wheat HYVs. As a male labor-biased crop, the expansion of wheat production contributed to declines in female wage labor participation. These findings document how new technology shocks can generate winners and losers, and offers an example of productivity growth that exacerbated gender disparities.

JEL Codes: J16, J43, O12, Q16, Q25
Keywords: Gender inequality, technical change, agricultural labor, Green Revolution

*I thank Supreet Kaur, Wolfram Schlenker, and Kiki Pop-Eleches for their advice and encouragement. I thank Emily Breza, Ram Fishman, Eyal Frank, Shruthi Jayaram, Amir Jina, Anouch Missirian, Suresh Naidu, Matt Neidell, James Rising, Rodrigo Soares, Anja Tolonen, Anna Tompsett, and seminar participants at Columbia University, Cornell University, and the Interdisciplinary PhD Workshop in Sustainable Development for helpful comments. I thank Andrew Foster, Yoshifumi Usami, and ICRISAT for data assistance. This research was funded in part by grants from The Endeavor Foundation and the Columbia University Center for Development Economics and Policy. All errors are mine.

†School of International and Public Affairs, Columbia University. 420 West 118th Street, New York, NY 10027. ald2187@columbia.edu. http://www.columbia.edu/~ald2187

1 Introduction

Technological change in agriculture can induce virtuous economic growth.[1] Through increases in agricultural productivity, more food can be produced by fewer people, freeing surplus labor for higher-productivity employment outside of agriculture (Lewis, 1954; Johnston and Mellor, 1961; Gollin et al., 2002).[2] While this process drives agriculture's decline as a share of the economy, the productivity gains brought by technical change still affect workers who remain in the sector. Of particular interest is what these gains mean for gender equality. Up to half, and sometimes more, of all agricultural labor in developing countries is supplied by women (FAO, 2011). Additionally, sizable gender wage gaps persist in these contexts, especially in South Asia (UN Women, 2015).

Whether technology improves gender inequality is partly a consequence of how it changes the nature and quantity of work. Consider classifying tasks as primarily brawn-intensive, or primarily brain-intensive. Since men have comparative advantage in brawn-based tasks, new technologies which increase the relative value of brains would reduce gender wage disparities (Galor and Weil, 1996; Welch, 2000; Bacolod and Blum, 2010; Rendall, 2010, 2013).[3] On the other hand, labor-augmenting technologies that do not have an effect on the task composition of agricultural work, can increase male labor demand and gender inequality. Similar results arise if labor-saving technologies displace tasks women tend to perform (Goldin, 1994).

This paper examines the effect of a new agricultural technology on gender inequality using India's Green Revolution, an appropriate empirical setting for three reasons. First, the Green Revolution is arguably the largest episode of agricultural technical change that India has experienced in recent decades. Introduced in the mid-1960s, its practices were visible throughout much of the country within just a few years. By 1970, it directly affected more than 340 million people engaged in agriculture, and nearly the entire country through changes in food prices and availability (Hazell, 2009). At such a scale, its equilibrium effects reached multiple factor markets, including village labor markets which are the focus of this paper. Second, the initial research and development

[1] Johnston (1970); Timmer (1988, 2002); Byerlee et al. (2009); Barrett et al. (2010) review the large literature on agriculture's contribution to economic development.

[2] Engel's law, that food demand is income inelastic, constrains the agriculture sector's growth potential.

[3] Pitt et al. (2012) use the brain-brawn framework to explain gender differentials in schooling investments and occupational choice in Bangladesh.

leading to the new technology was conducted overseas, and therefore presents an example of exogenous technical change suitable for causal inference.[4] Lastly, the wide range of climatic, geological, and topographical characteristics found throughout India, each plausibly influencing the returns to adoption, provide exogenous sources of identifying variation in seed suitability.

This paper's empirical strategy leverages the crucial importance of irrigation access in farmers' adoption of high-yielding variety seeds (HYVs). The HYVs were the central innovation behind the Green Revolution's productivity gains, and were more water-sensitive than the traditional varieties that preceeded them. Crops that received too much or too little irrigation produced lower yields (Foster and Rosenzweig, 1995, 2010). As the HYVs were bred specifically to be responsive to higher fertilizer amounts, irrigation access was also necessary in ensuring fertilizer nutrient absorption (Palmer, 1972). Consequently, farms with control over irrigation exercised comparative advantage in HYV adoption, with those residing above groundwater aquifers the most advantaged. By investing in private tubewells and pumpsets, a farmer could exercise maximum control over irrigation quantity and timing.[5] Such farmers were at a significant advantage to their peers who were beholden to the vagaries of unpredictable rainfall, or the availability of irrigation from surface water infrastructure with its own hydrological and political economy uncertainties.

I consequently exploit the quasi-random assignment of groundwater aquifers as a source of identifying variation in Green Revolution suitability. Using an aquifer's geological characteristics, I construct a plausibly exogenous treatment intensity measure, which unlike alternative groundwater availability measures like depth to water or flow rate, is not influenced by human activity. These geological conditions generate variation in the cost of extracting irrigation water, and consequently the profitability of undertaking such investments. While these geological conditions are fixed, it is possible that observed results could be driven by population sorting that arises after the Green Revolution's start, which I rule out as an unlikely confound given the various mobility frictions in land markets and social insurance networks that dampen rural migration in India (Bardhan, 1973;

[4] Alternative settings where adoption is exogenous have relied on experimental designs, such as in Duflo et al. (2011); Carter et al. (2014); Beaman et al. (2013); Fink et al. (2014). Various barriers limit the scale of such experiments, which are frequently too small to have noticeable equilibrium effects on labor markets.

[5] While private irrigation provision was feasible before the Green Revolution, the returns to digging a tubewell greatly increased with the HYVs' arrival, triggering substantial irrigation investment.

Fernando, 2016; Munshi and Rosenzweig, 2016).

Using 30 years of district-level data in a difference-in-differences framework, I find that groundwater access played a substantial role in determining agricultural decisions and outcomes. Groundwater-rich districts held a persistent 21 percentage point lead in the share of planted area cropped with HYVs, stemming from both farmer-specific decisions and the government policies, like fertilizer and electricity subsidies, which were implemented to support them.[6] As a result, these districts experienced 20 percent higher unit area revenues, and changes to cultivation patterns that replaced crops without an HYV option (e.g., chickpeas, groundnut) with those that did. Usage of complementary inputs, like irrigation and phosphorous fertilizer, also significantly increased.

These productivity gains had persistent effects on village labor markets, affecting both wages and participation rates. Using newly digitized village-level daily wage data, I estimate the Green Revolution generated an average 7 percent wage increase for male laborers, while women's wages registered a relative decline. These opposing trends led to a 17 percent increase in the male wage premium. This finding is robust to including flexible time trends, adding fully-interacted baseline characteristics, dropping the Green Revolution's most successful states, and performing various spatio-temporal averages of the raw data. As well, these results are not driven exclusively by men and women sorting into different tasks, or from working in different periods of the agricultural season. Subsampling wage results to only crop-related activities, which excludes tasks like well-digging and carrying loads which could plausibly advantage men, only slightly attenuates the male premium estimate. An even more restrictive test that subsamples to time periods and agricultural tasks for villages in which both male and female wages are observed, still generates a male wage premium in treated locations, indicating that task-based or temporal sorting cannot entirely explain the widening gender wage gap.

Labor participation effects also differed by gender. While extensive margin male wage labor employment did not respond to the new technology, women's full-time participation rates in village agricultural labor markets fell an average 5 percentage points between 1961 and 1981. Women partially substituted wage labor with own-farm cultivation and home production. According to

[6] Identifying the relative contribution from individual farmers or governments is not essential for interpreting the paper's main results.

time-use survey reports, there is little evidence this substitution pattern increased women's leisure. Women who did not participate full-time in wage labor spent a larger share of their day performing unpaid work, like grinding grain, cooking and cleaning, and collecting fuel. Conditional on spending a positive amount of time in paid agricultural labor, there is no difference in the workday length between women in treated and untreated villages.

I build a conceptual model that expands on the Acemoglu and Autor (2011) framework of technical change and wage inequality, which offers several testable channels that link the two. The key labor market outcomes appear to be driven by a crop substitution mechanism that interacts with the existing gender division of agricultural labor. Of the HYV-eligible crops, yields for wheat were the most responsive to the new seeds. Yield gains were immediate and farmers responded by expanding its cultivation. Within a decade of the Green Revolution's start, the wheat share of total cropped area in treated districts had risen nearly 10 percentage points, representing a 100 percent increase over the 1965 sample mean. The wheat share continued rising through to the end of the sample, while the share of area under rice showed modest gains that only began in the mid-1970s. These changes magnified the existing gender division of labor. As a 'plough-positive crop' (Alesina et al., 2013), a larger share of tasks required for wheat production are performed by men. Whereas women's roles in transplanting and weeding makes them central figures in rice production (Paris, 1998), these tasks comprise a much smaller share of labor required in wheat production. Following earlier work positing an inverse relationship between wheat production and female labor participation (Bardhan, 1984; Mbiti, 2008), I find corroborating evidence from cross-sectional and panel methods that districts where wheat share increased, also experienced reductions in female wage labor participation.[7] Results from plot-level data on total labor demanded support this claim, with men's share of total worker-days employed on wheat plots 9 percentage points higher than on rice plots. Furthermore, this data suggests that the total number of worker-days required to cultivate a unit area of wheat is less than for rice. If anti-egalitarian norms prevail, then limited work opportunities would first be delegated to men (Fortin, 2005), which offers a complementary

[7] Boserup (1970) first claimed that shifting and intensive cultivation had different effects on the gender division of labor, and the relative comparison of rice- and wheat-growing areas is discussed in Miller (1982); Dyson and Moore (1983).

explanation for the negative female labor demand shock.

This paper comes closest in spirit to a small number of papers that explicitly model the effect of technical change on gender wage inequality (Black and Spitz-Oener, 2010; Lindley, 2012; Rendall, 2013), but differs in focusing exclusively on a single sector, in a developing economy, and on a single technology whose adoption can be analyzed as a natural experiment. While factor-biased technical change has served as a framework for understanding developed economy labor market trends since the 1990s (e.g., Krueger, 1993; Acemoglu, 1998, 2002; Card and DiNardo, 2002; Acemoglu, 2003; Autor et al., 2003, 2008), its empirical application for explaining changes in the gender wage gap has been a recent development, and this paper the first to do so in an agricultural context.

This paper also contributes to a literature on the Green Revolution in India, and is the first to provide causal evidence of its effects on the wage laborers who represent a large share of the rural workforce. In contrast, earlier work has concentrated on producers, and used the introduction of HYV seeds as a setting to understand technology adoption and learning (Besley and Case, 1993; Foster and Rosenzweig, 1995; Munshi, 2004), and the role of income shocks in shaping household investments (Rosenzweig, 1982, 1990). Additionally, whereas earlier debates on the Green Revolution's contribution to inequality concentrated on regional differences, and income gaps between large-scale farmers and smallholders (Prahladachar, 1983; Dhanagare, 1987; Freebairn, 1995), this is the first paper that demonstrates its role in exacerbating gender inequality.

Similarly, this paper demonstrates that growth in the agriculture sector will not always translate to improvements in women's employment opportunities, and provides a nuanced contribution to the development-empowerment view of development creating gains for gender equality (e.g., Boserup, 1970; Goldin, 1990; Duflo, 2012; Jayachandran, 2015). While this result is consistent with the downward-sloping portion of the U-shaped pattern of female labor force participation (FLFP), such reductions are conventionally attributed to contractions in the agriculture sector and employment growth in manufacturing dominated by male workers (Goldin, 1990, 1994; Mammen and Paxson, 2000; Heath and Jayachandran, 2016). Manufacturing growth is not an essential precursor to declines in women's employment, as developments in how agricultural work is performed

which magnify men's comparative advantage can produce comparable results.[8] This paper therefore offers a counterpoint to the setting in Qian (2008), in which women's comparative advantage in tea-picking generated income gains under post-Mao reforms which increased tea's profitability. Since women's earnings are a determinant of investments in children's health and schooling (Rosenzweig and Schultz, 1982; Thomas, 1990; Duflo, 2003; Qian, 2008), the employment setbacks women laborers faced from the Green Revolution may have not only worsened their own intra-household bargaining power, but also generated inter-generational effects.

The remainder of the paper is organized as follows. Section 2 provides an overview of Indian agriculture and the Green Revolution. Section 3 develops a conceptual model to guide the empirical analysis. Sections 4 and 5 describes the datasets and the empirical strategy. Results are summarized in Section 6, including alternative explanations and robustness checks, while Section 7 concludes.

2 Overview of Indian Agriculture and the Green Revolution

The empirical setting for this paper is rural India, from the late-1950s through the 1980s. Over this period the rural population is largely engaged in agriculture as cultivators, laborers hiring out their labor for wages, or in the combination of cultivation and wage labor. Since farm labor requirements often exceed a household's labor endowment, agricultural laborers, accounting for 15 percent of the rural population (1961), are hired to perform tasks like ploughing, sowing, and harvesting. Agricultural output encompasses a wide range of crops, though rice and wheat were two of the most important, and in 1961 jointly accounted for more than 40 percent of all foodgrain-producing land. Rice is grown across the country and throughout the year, while wheat is planted primarily in the north between October and December, and harvested from April through May (Krishna Kumar et al., 2004).[9]

Women play a substantial role in agricultural wage labor. Their labor participation rates

[8] Manufacturing plays a small role in this paper's labor markets, respectively accounting for less than 3 and 1 percent of rural male and female employment.

[9] The Indian Council of Agricultural Research publishes a Crop Calendar of Major Crops which summarizes by state the agricultural calendar for major crops.

vary widely by location, and while generally higher in rice- than in wheat-growing areas (Boserup, 1970), they on average represented 45 percent of all agricultural laborers (1961). The activities they perform, and their relative share of total labor performed, is in large part circumscribed by crop choice. In rice cultivation systems, where 'control-intensive tasks' like sowing, transplanting, and weeding command a large share of total labor, women's labor is more indispensable (Paris, 1998). Figure 1a highlights this using plot-level data to compare the male and female labor shares differentiated by task for rice-cropped and wheat-cropped plots.[10] Whereas sowing and weeding account for nearly 25 percent of total female labor in rice production, these activities represent less than 8 percent in wheat production. The division of labor also has temporal implications; the dominance of harvest and post-harvest activities in wheat production concentrates women's agricultural wage opportunities into fewer months than under a cycle of rice production. Importantly, the female labor share of all rice plot worker-days is 40 percent, but drops below 30 percent for wheat. Figure 1b shows that for the country in aggregate, the gender division of labor at the level of agricultural task appears relatively stable, based on official employment data collected in 1983 and 2004.[11]

Technical Change in Indian Agriculture: The Green Revolution

The 1966 introduction of high-yielding variety seeds (HYVs) marked the Green Revolution's arrival to India.[12] Initially available for rice, wheat, maize, sorghum, and pearl millet, the new seeds were the product of international agricultural research to increase crop productivity and strengthen food security in the face of droughts that had especially affected developing countries in the 1950s. Multiple varieties for each crop were distributed, such as Rojo 64A and Sonora 64 for wheat and IR-8, Jaya, Padma, and Pankaj for rice.[13] Since the initial variety breeding efforts were conducted

[10] These results are calculated using the 1999 round of the Rural Economic and Demographic Survey, explained in detail in Section 4.

[11] These results are calculated using the 1983 (38th Round) and 2004 (60th Round) of the National Sample Survey's Employment and Unemployment survey.

[12] Over the 1950s and 1960s, HYVs were released throughout much of Latin America and Asia. India's first bulk federal seed purchase of HYVs was made in 1966, though adoption in limited quantities began earlier in the 1960s (Chakravarti, 1973).

[13] Since variety-specific adoption data over time is not available, this paper treats HYVs for any single crop as a homogeneous set.

exclusively overseas, with wheat varieties developed in Mexico and rice varieties in the Philippines, the introduction of HYVs to India presents a clean setting to study exogenous technical change (Foster and Rosenzweig, 1996).[14]

HYVs impacted agriculture in three key ways. First, these "dwarf-varieties" were selected for shorter crop stalks that could support more grain without breaking or falling over (lodging), which in traditional, taller-stalked varieties limited potential yields (Evenson and Gollin, 2003). Second, they matured faster, in some cases reducing the time between planting and harvest by more than 40 days. This enabled farmers to cultivate two and potentially three crop cycles per year (Pingali, 2012). Third, they were highly responsive to fertilizer. Extension outreach programs trained farmers in appropriate fertilizer application in order to achieve optimal yields, contributing to rapid increases in total fertilizer usage. These effects collectively raised yields, farm output and labor productivity, and spurred wage gains for laborers (Cleaver, 1972), particularly in areas with inelastic labor supply.

HYV adoption was rapid, and particularly so for wheat.[15] Figure 2 plots the HYV share of planted area for all of India for each crop, and within a decade more than 70 percent of all wheat was planted with HYV seeds.[16] In contrast, rice required two decades before achieving 50 percent HYV coverage. Several factors contributed to this disparity. The rice HYVs were particularly vulnerable to viruses, bacterial disease, and pests. Early vintages suffered from undesirable cooking properties, like stickiness, with some varieties producing rice that tasted 'chalky' (Swaminathan, 1969; Chakravarti, 1973). In some locations, state governments set procurement prices for HYV rice below that of traditional rice, disincentivizing producers to adopt HYV rice. Farmers in areas where wheat could be grown often realized lower profit margins for HYV rice than achievable with HYV wheat (Roy, 1971). This may partly have been due to stronger social learning in wheat-growing areas where heterogeneity in growing conditions was more limited than in primarily rice-growing

[14] Following their introduction, domestic agricultural research centers made further HYV seed improvements by selecting for traits appropriate to local agronomic conditions and dietary preferences. I therefore focus only on the 1966 release to avoid the endogeneity of state-varying agricultural R&D capacity and investments.

[15] Farmers became aware of the seeds through demonstration plots (Radhakrishnan, 1969), and were able to access them through both national governmental bodies, like the National Seeds Corporation, and state-level organizations.

[16] By 1967, 90 percent of the sample districts in this paper's analysis were planting some amount of HYVs, rising to 98 percent by 1970. It is unlikely that differential availability across districts is a major driver of adoption patterns.

locations (Munshi, 2004).

Controlled irrigation was crucial in achieving maximum HYV yields (Sengupta and Ghosh, 1968), and farmers with access to groundwater resources were particularly advantaged. By drilling a well and purchasing a pumpset, a farmer could exert complete control over how much and when their crops were irrigated (Chinnappa, 1977).[17,18] Control over water supply was necessary for crop absorption of fertilizer nutrients, and the absence of irrigation resulted in yield losses for HYVs (Palmer, 1972). Groundwater also serves as insurance, by reducing farmers' risk exposure to negative rainfall shocks. If rainfall does not arrive at the right time of the crop cycle, or is insufficient in amount, then groundwater supplies can be drawn upon to compensate the difference.[19] The Green Revolution consequently prompted a wave of tubewell installations and pumpset purchases, as seen in Figure A1a. Groundwater irrigation was a significant improvement to the alternatives of surface water irrigation, or reliance on rainfall. For example, farmers dependent on canal irrigation networks could not count on water being available outside the primary rice-growing season (Sengupta and Ghosh, 1968), which limited their ability to capitalize on HYVs' faster maturation and potential for multiple cropping.

3 Conceptual Model

Consider a closed village economy comprised of J price-taking producers, each possessing 1 unit of land, and I households, each consisting of one man (m) and one woman (f). Producers allocate land z between two crops, a and b, both with output which is Cobb-Douglas in land and aggregate labor. Producers employ laborers of both genders, whose output is combined in crop-specific, constant elasticity of substitution (CES) labor aggregators $\widehat{L_a}$ and $\widehat{L_b}$. Accordingly, $\widehat{L_s} = (\phi_{sf} L_{sf}^{\rho} + \phi_{sm} L_{sm}^{\rho})^{\frac{1}{\rho}}, \quad \forall s \in \{a, b\}$, where ϕ's are exogenous labor productivity parameters that capture

[17] Sekhri (2011) explains the physical process by which pumpsets extract underground water from tubewells.

[18] This is of course predicated on groundwater availability. Problems like water scarcity and salinization caused by over-extraction became documented problems in the 1980s, and would not have affected investment decisions in the 1960s and 1970s.

[19] If aquifer recharge fro excess rainfall in positive rainfall years, balanced withdrawals, then groundwater could be a renewable resource. Rodell et al. (2009) combine satellite data and hydrological models to estimate withdrawal rates in northwest India, and find evidence for unsustainable consumption patterns that could eventually result in local water stress.

underlying differences between men and women, as well as their relative efficiencies producing a and b. The sum of land under crop production is $z_{aj} + z_{bj} \leq 1$ for $j \in \{1, \ldots, J\}$. Assume that no land rental markets exist, so that unallocated land generates no returns to land-holders. Output is respectively $y_a = A_a(t) z_a^{\alpha} \widehat{L_a}^{\beta}$ and $y_b = A_b(t) z_b^{\alpha} \widehat{L_b}^{\beta}$, where A_a and A_b are factor-neutral technology parameters with values that are non-decreasing in time and exogenously determined. Their output combines to produce a common consumption good, calories for example, and a is the numeraire good. Labor markets are perfectly competitive and all factors receive their marginal product.

In this framework, drawing on Acemoglu and Autor (2011), exogenous technical change may operate through two channels. First, factor-neutral change impacts A_s, leaving the within-crop ratio of marginal labor productivities unchanged. Conversely, factor-biased change affects male and female productivity differently. Assuming all productivity parameters are a function of some exogenous input x, this means that for either s crop, $\frac{\partial \phi_{sm}}{\partial x} \lesseqgtr \frac{\partial \phi_{sf}}{\partial x} > 0$, with the bias of technical change consistent with the larger term.

Producers decide how much land to allocate between the two crops, and over how much male and female labor to hire. They therefore solve the following profit maximization problem,

$$\max_{\substack{\{z_s, L_{sg}\} \\ \forall s \in \{a,b\}, \forall g \in \{m,f\}}} \quad \pi = y_a + p y_b - \sum_s \sum_g w_{sg} L_{sg} \qquad \text{s.t. } 0 \leq z_a + z_b \leq 1 \tag{1}$$

which results in six first-order conditions. First, the wage ratios for each crop can be solved as follows,

$$\frac{w_{sm}}{w_{sf}} = \frac{\phi_{sm}}{\phi_{sf}} \left(\frac{L^*_{sf}}{L^*_{sm}} \right)^{1-\rho}, \ \forall s \in \{a, b\} \tag{2}$$

Since there are no returns to unused land, the constraint in (1) binds, and leads to the following

expression for optimal land share,

$$z_b^* = \left(1 + \left(\frac{pA_b\widehat{L_b^*}^{\beta}}{A_a\widehat{L_a^*}^{\beta}}\right)^{\frac{1}{\alpha-1}}\right)^{-1} \tag{3}$$

Household utility is well-behaved and concave in the consumption good c, whose allocation is unobservable to the econometrician. Initially consider that both men and women experience the same disutility of work per unit time, which is additively separable between the two. Each adult works $h_{gi} = 1 - l_{gi},\ \forall g \in \{m, f\},\ i \in \{1, \ldots, I\}$, since total time endowment is 1 unit. Households solve the following problem,

$$\begin{aligned} \max_{c, l_f, l_m} \quad & \ln(c) - (1 - l_f)^{\delta} - (1 - l_m)^{\delta} \\ & \text{s.t.} \ \ c \leq w_m h_m + w_f h_f + R \end{aligned} \tag{4}$$

where $\delta > 1$, R is non-wage income, and interior solutions for both male and female labor are assumed. Gender-specific wages must equalize across crops in equilibrium, because of free labor movement. Workers therefore decide only on total labor supply and are indifferent to which crop they work with.

In a one-period model the first constraint binds. The efficient household labor supply ratio is

$$\frac{h_f^*}{h_m^*} = \left(\frac{w_f}{w_m}\right)^{\frac{1}{\delta-1}} \tag{5}$$

with the usual result of labor ratios equaling the wage ratio if $\delta = 2$.

Proposition 3.1. *Both factor-neutral and factor-biased technical change that generates relatively larger productivity gains for crop* b *increase the share of area under crop* b.

This follows directly from (3), which results in $\frac{\partial z_b^*}{\partial (A_b/A_a)} > 0$ since $\alpha < 1$ and the optimal labor demand expressions are a function of the gender-specific productivity parameters and market wages. Differentiating z_b^* with respect to either ϕ_{bf} or ϕ_{bm} also results in a positive expression.

In contrast, holding constant all technology parameters related to b, increases in A_a, ϕ_{af}, or ϕ_{am} reduce crop b's optimal land share.

Proposition 3.2. *Gender-biased technical change, for either crop, increases the biased gender's relative labor share employed in producing that crop.*

Holding the wage ratio constant in (2), the first derivative of the gender labor ratio with respect to the gender productivity parameter ratio is strictly positive. Furthermore, as the gender labor ratio is a function of only relative wages and same-crop productivity parameters, growth in gender-specific productivities employed to a does not affect the gender labor ratio employed to $a \neq b$.

Proposition 3.3. *Male-biased technical change, in either crop, increases the male wage premium.*

Combining (2) with (5) and taking the derivative with respect to the productivity ratio results in $\frac{\partial\left(w_m/w_f\right)}{\partial\left(\frac{\phi_{sm}}{\phi_{sf}}\right)} = \frac{\delta-1}{\delta-\rho}\left(\frac{\phi_{sm}}{\phi_{sf}}\right)^{\frac{\rho-1}{\delta-\rho}} > 0$, since $\rho < 1$ and $\delta > 1$. Conversely, the male wage premium reduces when technical change is female-biased.

Proposition 3.4. *Factor-neutral technical change does not impact gender wage ratios.*

The four labor demand first-order conditions of (1) show that growth in factor-neutral A_s raises wage levels for both men and women growing s. Since wage growth is comparable for men and women, the effects of factor-neutral productivity growth is cancelled out when wages are expressed as ratios as seen in (2).

4 Data

This section summarizes all datasets used in the analysis, with Table 1 reporting key summary statistics. I first describe data collected at the district-level, which is the administrative unit analagous to a US county.[20] Summaries of individual-level data follow, before describing how groundwater and weather data was constructed. Further details are provided in the online data appendix.

[20] The average district population in the 1961 census was 1.4 million.

4.1 District-Level Data

Agricultural Laborer Wages

Laborer wages are taken from the Agricultural Wages in India (AWI), an annual volume compiled by the Ministry of Agriculture's Directorate of Economics and Statistics. Average wages by gender and by month are reported by agricultural operation for a selected number of reporting centers (villages) per district. Similar to Indian agricultural statistics, a 'year' spans from July through June of the subsequent calendar year. Reported values are the sum of cash wages and in-kind wages converted to a cash equivalent.

States were given discretion over how to collect this data, resulting in inter-state differences in labor classification. In general, Andhra Pradesh, Assam, Gujarat, Himachal Pradesh, Karnataka, Kerala, Maharashtra, Orissa, and West Bengal reported wages by 'field labor' and 'other labor' categories. Bihar, Madhya Pradesh, Punjab, Tamil Nadu, and Uttar Pradesh reported wages by 'ploughing', 'sowing', 'weeding', 'reaping and harvesting', and 'other labor'. As per the Ministry of Agriculture's standard pro-forma for data collection, 'other labor' includes wages for "coolies employed for watering the fields, load-carriers, coolies, well-diggers, laborers cleaning silt from water-ways, embankments, etc." Whichever classification scheme was chosen, states largely adhered to it for all years in the sample.

At the time of writing, no available data source provides national coverage of male and female wages before and after 1966.[21] I address this by digitizing wage records from 1958-1971, which are merged with Usami (2014) who digitized records for 1973-1995. Due to resource constraints, I digitize records only for the months of July and January, supplementing this when necessary to ensure at least one concurrent, operation×month record is included for both male and female wages. These supplemental records increase the sample of female wages from states like Punjab, Uttar Pradesh, and Bihar who report female wages less frequently, consistent with comparatively low female labor participation rates. AWI data is *de facto* unbalanced as the composition of

[21] The World Bank India Agriculture and Climate dataset includes male wages for 1956-1987, but not female wages. The ICRISAT Meso-scale Apportioned dataset contains both male and female wages, but includes no observations before 1966. The first round of the National Sample Survey Employment Unemployment survey that is accessible to researchers was conducted in 1983.

reporting centers changes over time. Furthermore, wages for month m may be available in year t, but not in $t+1$. I address this with a series of robustness checks and comparisons against other datasets containing AWI values to ensure my results are not due to data collection artifacts.[22] Real wages, where used, are constructed using state-level Consumer Price Index for Agricultural Laborers (CPIAL) data (Besley and Burgess, 2000). To ensure consistency with the agricultural data, I apportion data to 1961 boundaries using guidelines developed by Kumar and Somanathan (2015).

Agricultural, Demographic, and Employment Data

Annual data on agricultural inputs and outputs is taken from the World Bank India Agriculture and Climate dataset. This is a balanced panel of 271 districts from 13 states for 1956-1987, and covers the majority of agricultural lands; sample districts in 1966 accounted for 88 percent of India's total rice production area.[23] To ensure temporal comparability, as substantial redistricting occurred over this period, the data providers apportioned all statistics to 1961 district boundaries. Data is available by crop for planted area under traditional and high-yielding varieties, irrigated area, and total output, and includes statistics on fertilizer usage, tractors, and bullocks. Area under HYV cultivation is available for wheat, rice, maize, sorghum, and pearl millet, henceforth referred to as the 'HYV-eligible crops', with all HYV measures used in the analysis a composite of these five crops. I construct a unit area revenue measure by taking the crop-weighted output per hectare priced at a district's average farm harvest price for 1960-1965 for each crop. I merge the World Bank data with the district-level population census for 1961, 1971, and 1981, which includes population and employment counts by broad occupation categories (e.g., cultivators, agricultural laborers, manufacturing employees), and is disaggregated by gender and by urban/rural sectors (Vanneman and Barnes, 2000).[24]

[22] New reporting centers are occasionally added to the sample, and so robustness checks are performed that subsample to only those which reported wage data before 1960. This subset of locations therefore would provide at least six years of pre-1966 data, against which post-1966 values can be compared. Reporting centers which are late entrants to the sample, in some cases after 1966, both provide less information about pre-treatment wage values and can bias coefficients of interest.

[23] This is according to statistics collected in the Reserve Bank of India's "Handbook of Statistics on Indian Economy."

[24] Population census data is also fixed at 1961 district boundaries.

4.2 Individual/Household-Level Data

Since district-level data can only provide a coarse overview of the Green Revolution's labor market effects, I use individual-level data from the 1970/71 and 1982/83 rounds of the National Council of Applied Economic Research's Rural Economic & Demographic Survey (REDS), a nationally-representative household panel collected from more than 4,000 households for around 240 villages across all major states.[25] The REDS panel includes 33,000 individual-years of employment data, and because is organized by household allows for analyzing labor decisions at the household-level. Women's time-use data collected in the 1983 Demographic Questionnaire provides information about time spent in various work and non-work activities. I test mechanisms with the 1999/00 round which includes plot-level data on crop choice, inputs, and labor demanded by worker gender, and is the first REDS round to provide labor allocation data at this level of disaggregation. As production processes likely shifted between the period of interest and the 1999 data, I run a series of checks on subsets of farms which have less access to inputs and infrastructure, as an approximation for the type of agriculture conducted early into the Green Revolution.

4.3 Groundwater Coverage

I construct a novel groundwater database by georeferencing and digitizing the Geological Survey of India's 1969 Geohydrological Map of India.[26] This map resulted from the first national assessment of hydrological resources and offers the most detailed overview of the availability of groundwater at the start of the Green Revolution. The map classifies groundwater formations as unconsolidated, semi-consolidated, or consolidated, based on lithology and tectonics. I focus exclusively on unconsolidated formations which likely provided the most accessible and inexpensive formation for extracting groundwater.[27] For district-level analyses, I construct a continuous groundwater variable spanning [0,1] by calculating the share of district area (using 1961 administrative boundaries) atop

[25] Households moving to urban areas after the initial round, or other villages not included in REDS collection, would be dropped from the sample.

[26] Rud (2012) uses similar, state-level data from the 1982 National Atlas of India to estimate the effect of electrification on industrial growth.

[27] Unconsolidated formations are further sub-classified into thicknesses (in meters) of $d \geq 150$, $150 \geq d \geq 100$, $100 \geq d$, and piedmont zones exhibiting large seasonal fluctuations in water storage. My approach therefore estimates the average effect across these categories.

any unconsolidated formation. The resulting district shares are illustrated in Figure 3. As villages are geographically compact, I use village-level GPS coordinates to construct a binary groundwater variable for each REDS village, coding them as one if located above an unconsolidated formation and zero otherwise. REDS village locations are marked in Figure 3 as red stars. Since a village may not reside atop an unconsolidated aquifer, but still access groundwater resources through a canal system or other diversion structure, a set of robustness checks employs an alternative groundwater measure which is constructed as the share of area covering an unconsolidated aquifer within an r kilometer radius of the village, for r varying from 1 to 50.

4.4 Weather Controls

In all district panel models involving agricultural outcomes or agricultural wages, I control for the influence of weather and include temperature and precipitation variables. I use India Meteorological Department daily temperature data at $1^\circ \times 1^\circ$ resolution (Srivastava et al., 2009), available for 1951-2013, and compute growing degree days using single sine-wave interpolation (Schlenker and Roberts, 2009). Daily values are then aggregated to season totals. This approach captures nonlinearities in crop response to extreme temperatures and often outperforms models using average temperature in out-of-sample predictions. I use monthly precipitation data from the Climatic Research Unit (University of East Anglia) TS 3.23 at $0.5^\circ \times 0.5^\circ$ resolution, available for 1901-present, and sum by season. The respective ranges for *kharif* and *rabi* seasons of June 1-September 30 and October 1-February 28 are fixed, constant across space and years, and consistent with earlier work (Krishna Kumar et al., 2004; Guiteras, 2009; Auffhammer et al., 2012). District values for both temperature and preciptation are computed using the pixel which contains the district's centroid. The set of weather variables included in panel specifications are the contemporaneous and one year lagged values of *kharif* rainfall in quadratic, *rabi* rainfall in quadratic, *kharif* growing degree day sums for [10°C, 24°C) and [32°C, $+\infty$), and *rabi* growing degree day sums for ($-\infty$, 10°C), [25°C, 31°C), and [32°C, $+\infty$).

5 Empirical Tests

I estimate the Green Revolution's effect on the gender wage gap by exploiting the quasi-random placement of groundwater aquifers, which while beneficial to agricultural production prior to 1966, gained substantial importance when the irrigation-sensitive HYV seeds were introduced. I later provide supporting evidence by plotting pre-trends of the share of cropped area under irrigation as evidence of the Green Revolution's role in irrigation expansion.

This section details the difference-in-differences (DID) estimation procedure used on district-level data spanning before and after 1966, and then describes the analagous procedure for individual-level analyses. While the 1970 REDS round was conducted after 1966, I follow earlier work in regarding it as representative of labor market and household activity early into the Green Revolution, and a suitable baseline for the 1982 round in which households had an additional 12 years to learn about and adopt the new technology. Consequently, the DID specification compares locations receiving the Green Revolution treatment (i.e., districts with better groundwater endowment, and REDS villages atop groundwater) with control group locations, for before and after 1966 (or 1970 to 1982 with REDS households).

5.1 District-Level Estimation

The main district-level specification uses a DID estimator of the form,

$$y_{dt} = \sum_{\substack{\tau \in [1956,1987] \\ \tau \neq 1965}} \beta_\tau \left(\text{Groundwater}_d \,\% \times \mathbb{1}\{t = \tau\}_t \right) + \boldsymbol{X}_{dt}^{'} \xi + \gamma_d + \phi_t + \epsilon_{dt} \tag{6}$$

for district d in year t, to estimate the effects of groundwater coverage on labor (e.g., log wages, labor participation rates) and agricultural outcomes (e.g., log yields, log revenue per unit area, area share irrigated). Wage models add month and labor operation fixed effects. In annually-varying models, 1965 is always the omitted year. For efficiency gains, select specifications replace year dummies with pooled intervals. The generalized version of this DID model replaces the time-interacted groundwater variables with an indicator function assuming unity from 1966, interacted

with groundwater coverage. Time-varying controls include the set of contemporaneous and lagged weather variables described in Section 4.4. Since the groundwater measure is static and fixed at 1969 values based on the geohydrological map, Green Revolution effects are estimated off of cross-sectional variation whose marginal effect on an outcome of interest is anticipated to significantly change in 1966 with HYV availability. The main identifying assumption underlying this model is that the parallel trends assumption holds (i.e., the evolution of treated district outcomes would have been similar to those of untreated districts in the absence of treatment) such that $\mathbb{E}[(\text{Groundwater }\%_d \times \phi_t)\epsilon_{dt}|\mathbf{X}'_{dt}, \gamma_d, \phi_t] = 0$. Graphical evidence on pre-trends is presented in Section 6 to support this assertion.

Furthermore, adoption of Green Revolution practices encompassed infrastructure investments (e.g., drilling of tubewells and purchase of pumpsets) and more intensive application of other inputs like fertilizers and pesticides. Since these are endogenously co-determined, they would serve as bad controls and are not considered as covariates. I instead employ the groundwater measure as a proxy for Green Revolution intensity and show in Section 6.1 strongly, positive covariance between groundwater and input intensities for quantities relevant to the Green Revolution. While the direct interpretation of the β_τ's is the marginal effect of groundwater coverage after 1966, estimates can be more broadly contextualized as marginal effects of Green Revolution intensity.

An expansion to the generalized DID estimator includes a vector of time-invariant characteristics, $\mathbf{S}_d$, fully interacted with year dummies. This allows non-groundwater variables like soil and topographic conditions, as well as pre-treatment baseline values of variables like female literacy or labor force participation rates to have flexible, time-varying effects on outcomes. This results in the following specification

$$y_{dt} = \beta\left(\text{Groundwater}_d\,\% \times \text{Post}_t\right) + \sum_{\substack{\tau\in[1956,1987]\\ \tau\neq 1965}} \Gamma_\tau \left(\boldsymbol{S}'_d \times \mathbb{1}\{t=\tau\}_t\right) + \boldsymbol{X}'_{dt}\xi + \gamma_d + \phi_t + f(\rho_d, \phi_t) + \epsilon_{dt} \quad (7)$$

which also allows for flexibly-specified time trends. Time-invariant local characteristics are absorbed

by district fixed effects, but trends or more idiosyncratic movements in wage determinants may drive changes to the outcome variable. I therefore perform robustness checks that include state $\times$ year fixed effects to more persuasively isolate marginal effects as due to district-level variation in Green Revolution suitability.

Standard errors in all district-level models are clustered by National Sample Survey region, an administrative unit used in the 1983 National Sample Survey that clubs together districts of comparable topography and geography within a state.[28] This allows for arbitrary correlation in errors across all district-year pairs for districts within a region, and results in $K = 53$ clusters. As districts vary in size, observations are weighted by 1961 gross cropped area. Estimates are then interpreted as the marginal effect for a hectare of cropped land.

5.2 Estimation Procedure for Household/Individual-Level Data

Models analyzing REDS data use the following specification,

$$y_{ihvt} = \beta^R \left(\mathbb{1}\{\text{Groundwater}_v\} \times \phi_t\right) + \boldsymbol{X}'_{ihvt}\xi^R + \gamma_v^R + \phi_t^R + \epsilon_{ihvt}^R \tag{8}$$

for individual i from household h in village v for year t, where R superscripts differentiate parameters from those estimated in district-level specifications. Observations are weighted with sampling probability weights calculated by NCAER.[29] Standard errors are clustered by village.

Pre-1966 baseline data is not available, so the groundwater variable in panel models is interacted with a year dummy instead of with Post_t, and because it remains time-invariant, drops out in panel models. $\boldsymbol{X}'$ includes controls for age, religion, education, and household size. Since twelve years separate the 1970 and 1982 rounds, long enough for large-scale reorganization of household composition and characteristics, I rely on village fixed effects instead of household fixed effects, which are plausibly less defensible. When estimated as a village-level panel, village fixed effects absorb time-invariant characteristics of labor markets and households across space. Cross-

[28] A lookup table that matches districts to regions is available upon request.

[29] The sampling probability weights correct for the survey design which oversampled villages participating in the Intensive Agriculture Development Program and Intensive Agriculture Area Program schemes, and high-income households.

sectional models are required when variables of interest are available for only a single year, and follow

$$y_{ihv} = \beta^R \left(\mathbb{1}\{\text{Groundwater}_v\}\right) + \boldsymbol{X}'_{ihv}\xi^R + \epsilon^R_{ihv} \tag{9}$$

where village fixed effects are dropped because of perfect collinearity with groundwater.

6 Results

The results are organized as follows. Section 6.1 provides evidence of groundwater's suitability as a Green Revolution proxy, loosely analagous to an instrumental variables 'first stage' results. Sections 6.2 and 6.3 present results for agricultural wage and labor outcomes. Section 6.4 describes a crop substitution mechanism that is consistent with the agriculture, wage, and labor outcome results. Section 6.5 considers alternative explanations, while 6.6 presents results from robustness checks.

6.1 Verifying Groundwater's Suitability as a Green Revolution Proxy

A credible Green Revolution proxy must satisfy several requirements. First, it should be predictive of HYV adoption. Second, given the input complementarities associated with the new technology, it should positively covary with other inputs like irrigation and fertilizer. Lastly, the proxy should be predictive of yield and unit area revenue increases.

Figure 4 addresses the first requirement and plots the β_τ's from equation (6) in which the HYV share of total cropped area of HYV-eligible crops, $\frac{\sum^C \text{HYV Area}^c_{dt}}{\sum^C \text{Total Area}^c_{dt}} \mid c \in \{\text{HYV eligible crops}\}$, is regressed on the time-interacted groundwater variable. More groundwater coverage translated to higher HYV adoption rates over the whole period, gradually increasing from 1966 and stabilizing in the early 1980s at a relative difference of ≈2.3 percentage points for each 10 percentage points of groundwater coverage.

The Green Revolution additionally impacted many aspects of agricultural production, as shown in Table 2. First, the share of crop area receiving irrigation increased by 13 percentage

points ($p < 0.01$) for districts fully covering groundwater aquifers, as seen in (A1). Columns (A2)-(A7) estimate the effect on crop shares of total cropped area. Treated farmers substituted away from crops like groundnut, barley, and pulses, for which HYV options did not exist (A7), towards more intensive cultivation of rice and wheat, the dominant HYV-eligible crops.[30] There is little evidence the Green Revolution provoked expansion of cultivated land. On the contrary, the point estimate of (A8), measuring the relative difference in agricultural land expansion growth rates, suggests slower extensification.[31] This suggests the Green Revolution was unlikely to impact labor demand through increases in the manual labor tasks associated with land clearance, preparation, and irrigation construction, all of which would likely have required more male than female labor. Instead, farmers capitalized on HYVs' faster maturation to produce more harvest cycles per year (A9), generating a 6.7 pp ($p = 0.01$) increase in cropping intensity. Lastly, growth in phosphorous fertilizer usage registered a 59 percent increase ($p = 0.03$) in fully groundwater districts.

Output gains for treated districts were large, as seen in Panel B.[32] Wheat and rice yield growth rates were respectively 20 percent ($p = 0.06$) and 16 percent ($p = 0.05$) higher for treated districts after 1966, contributing the majority of the observed 21 percent ($p < 0.01$) unit area revenue increase. Traditional crops like sorghum (jowar) and pearl millet (bajra) enjoyed higher yield growth in Green Revolution areas, but are imprecisely estimated. Maize yield growth declined by 30 percent and appears to be driven by a reduction in the maize planting share of treated districts where yields prior to 1966 were among the highest in the country.

6.2 Agricultural Laborer Wage Response to Green Revolution

The preceding section demonstrated that groundwater-rich districts experienced significantly larger gains in crop output and unit area revenues as a result of the Green Revolution technology package. This section details the consequent wage responses for the male and female agricultural wage workers

[30] The pattern of more intensive rice and wheat production has been observed in other countries adopting Green Revolution practices (Pinstrup-Andersen and Hazell, 1985).

[31] This result is consistent with the Borlaug hypothesis that claims increases in crop yields would reduce deforestation pressures. This finding is not restricted to India, as Pinstrup-Andersen and Hazell (1985) finds this result to be common among Green Revolution adopting countries. Stevenson et al. (2013) estimate that in the absence of the Green Revolution, global cropland would have increased an additional 2 percent by 2004, which is a smaller effect size than claimed in earlier studies.

[32] Profit estimates cannot be constructed for this period because of inadequate cost of cultivation data.

who provided the labor to realize those gains.

Before providing the main results, I first provide a basic summary of the wage data in Table 3 to facilitate interpretation of subsequent results. Panel A demonstrates an average 24 percent male wage premium across all reporting villages in the Agricultural Wages in India dataset, which is not very responsive to the addition of village fixed effects or an interaction term with the groundwater measure. The REDS data includes individual-level wage data only for 1970, respectively presented in Panels B and C for agricultural and non-agricultural wages, the latter drawn primarily from work in fishing and logging, mining, sales, and crafts and trades. There are substantially more workers receiving agricultural wages than non-agricultural, and at a lower overall average value. The male premium ranges from 32-40 percent, once village fixed effects are included. Furthermore, it should be noted that there is relatively limited within-village wage variation by sex. For example, the inclusion of village fixed effects in (B2) explains nearly 60 percent of all variation. This value rises to 84 (75) percent when regressing wages for women (men) only on village fixed effects. Columns (B3,C3) include controls for worker's education level, marital status, age, and age^2, but have little effect on the estimated male premium. Lastly, running the specification shown in (A3) and restricting the data to 1970, produces an estimate of 0.31 ($p < 0.01$) which is comparable to the estimate from REDS.

Figure 5 plots the time-varying coefficients of groundwater coverage on log nominal wages, estimated separately by sex with 1965 the omitted year. Male laborers in treated districts experienced a relative wage increase that peaked at 17 percent in the early 1970s, gradually decaying to 1965 levels by the early 1980s. This finding corroborates case study reports of male wage increases prompted by both higher labor demands (greater yields to be harvested and threshed), the improved bargaining position opposite landowners that laborers found themselves in as a result of labor supply deficits, and labor organizing movements that agitated for higher wages (Frankel, 1971). Female wages, on the other hand, underwent a steady decline from 1967 onwards, with growth rates dipping to 20 percent below 1965 in the late 1970s.[33]

[33] Women in groundwater-rich districts have overall higher wages than women elsewhere over the 1958-1987 period, with each 10 percentage point increase in groundwater coverage translating to a 3.5 percent ($p < 0.01$) increase in daily wages. The analagous value for male workers is 1.8 percent ($p = 0.11$).

I then estimate the male premium in Figure 5 by pooling male and female wages using a triple difference estimator,

$$\text{log wages}_{iodst} = \sum_{\substack{\tau \in [1956,1987] \\ \tau \neq 1965}} \Big[\beta^0_\tau \left(\text{Groundwater}_d \times \mathbb{1}(\text{Male}) \times \phi_t\right) + \beta^1_\tau \left(\mathbb{1}(\text{Male}) \times \phi_t\right) \tag{10}$$

$$+ \beta^2_\tau \left(\text{Groundwater}_d \times \phi_t\right)\Big] + \pi \left(\text{Groundwater}_d \times \mathbb{1}(\text{Male})\right) + \boldsymbol{X}'_{iodst}\xi$$

$$+\alpha \mathbb{1}(\text{Male}) + \gamma_d + \delta_o + \phi_t + \epsilon_{iodst}$$

which modifies Equation 6 by adding indices for a worker of gender i performing agricultural operation o. I use an indicator for male workers to simplify interpretation, such that an increase (decrease) in β^0_τ's, the coefficients of interest, widens (narrows) the wage gap. Figure 6 plots the results and confirms male-biased wage growth in Green Revolution districts, peaking in the late 1970s and beginning a decline in the early 1980s. Taking an estimate of the average post-period treatment effect leads to a male wage premium semi-elasticity of 0.16 ($p < 0.01$).

6.3 Impact of Green Revolution on Non-Wage Labor Outcomes

Groundwater-rich locations experienced better agricultural outcomes, but only male laborers received the wage increases induced by productivity gains. This section investigates how occupational outcomes were impacted and whether compositional changes in the laborer workforce might explain the rising wage gap. Lastly, I analyze women's time-use data to identify whether workday length was a margin of adjustment.

Effect on Occupational Choice

I first use population census data to estimate treatment effects on occupational choice, focusing on the rural population share engaged in agricultural labor or cultivation, with estimates presented

in Table 4 relative to the omitted 1961 × groundwater variable.[34,35,36] Despite substantial wage premiums in treated districts, the extensive-margin male agricultural laborer response is inelastic, as in column (1). Female laborer share was more responsive, with participation lower by 3 percentage points in 1971 and 5 percentage points in 1981 (both $p < 0.01$) in treated districts.

Column (4) indicates that women exiting, or reducing their participation in, agricultural labor substituted into own-farm cultivation which can consist of a combination of own-farm labor and supervision of hired labor. Examining other industries identified in the census suggests this was the primary form of labor movement, as other sectors were too small to have absorbed the totality of female laborers.

Census data can only provide a coarse analysis of labor outcomes, so I turn to microdata for granularity. Table A1 presents extensive- (Panel A) and intensive- (Panel B) margin responses with the 1970 and 1982 REDS rounds. The top panel displays estimates using linear probability models for the effect on occupation outcomes, based on a respondent's primary occupation.[37] The point estimate in (A3) is positive, though noisily estimated, indicating that women in groundwater-rich villages were less likely in 1982 (than in 1970) to report being an agricultural laborer, consistent with district-level results. Results are stronger when subsampled to households with at least one member identifying as a wage laborer. In (2), treatment for non-household head males living with household heads who are agricultural laborers implied a 26 percentage point increase in the likelihood of also being a laborer. Responses for women move in the opposite direction; women in households where at least one male is an agricultural laborer were 22 percentage points less likely to be a laborer in 1982.[38] This distinction is important because women working without their husbands or other male relatives face stigma and restrictions, social or physical, on their outside

[34] I use totals based on the sum of individuals identifying in either occupation as a main (primary) or marginal (secondary) worker. While the determination if specific to the census year, respondents are categorized as main workers if they work more than 183 days in the previous year in that sector.

[35] Cultivators are workers engaged in any aspect of own-farm production which also includes labor supervision.

[36] Occupation data is not disaggregated by age group, but was collected across all ages. Consequently the very young comprise a non-negligible share of the 'non-working' population. Since children under 10 years old work in agriculture, all children cannot be correctly removed from the non-worker population. Models using occupation shares constructed using only the population aged 15 and above produce similar results to those in Table 4.

[37] The 1970 round collected individual 'secondary occupation' data whereas the 1982 round did not. I therefore focus only on primary occupation status.

[38] A similar, albeit weaker effect is estimated when conditioning on the household head, instead of any male household member, identifying as a laborer.

working options. For example, in the 1970 REDS round, 87 percent of women who identified as an agricultural laborer lived in households where at least one male was also an agricultural laborer (76 percent) or resided in households with no males older than 15 (11 percent). As well, women working on family farms may free up male labor to pursue outside wage opportunities (Desai and Jain, 1994), and is a potentially efficient labor allocation in the presence of a male wage premium.

Intensive-margin results for agricultural and non-agricultural wage laborers are shown in Panel B. Unconditional days worked for the year are shown in (1)-(2) and (5)-(6), with remaining columns conditioned on positive days worked. All estimates are noisy and unable to reject the null of no change in the number of days laborers worked between 1970 and 1982. In agriculture (non-agriculture), employed women report working an average of 161 (131) days, both below the 183 day cutoff employed by the census to categorize 'main workers'. Consequently, some of the difference in occupation results between census and REDS data may be attributable to differences in how occupation affiliation is determined between REDS and the population census.

Tests of Composition Effects in Laborer Markets

In the absence of worker productivity data, I examine worker characteristics which are plausibly related to productivity, but find little support that compositional differences from the REDS data for 1970 and 1982 can explain increased inequality. The sample in Table 5 is the subset of workers identifying as agricultural laborers or reporting a positive number of days performing agricultural laborer work ('1(Ag Wage)'), displayed alongside the remainder of adults in the sample aged 15-60 ('All Others').[39] Linear probability models are used for binary outcomes like illiteracy and whether anyone in the household has an education above primary-level (B1-B2).

There are some noticeable differences in worker characteristics between groundwater and non-groundwater villages that emerge over time. While treated location laborers are increasingly from households that are poorer, and lack any family member with an above-primary education, they also are increasingly literate. This reflects a broader development of declining illiteracy rates in groundwater locations, with a 16 percentage point decline by 1982 ($p < 0.01$). That there are no

[39] About 7.6 percent of the combined male and female sample not reporting agricultural labor as their primary activity stated a positive number of days worked as a laborer.

significant differences in the triple interaction between men and women in treated locations, casts doubt on the extent to which male labor-capital complementarities confers an advantage through a human capital channel, for example in being able to read instructions to operate a new piece of agricultural equipment. Additionally, but not presented, I find no evidence for systematic changes in workforce composition according to land-holdings, using either land ownership as a continuous variable, or with landlessness as a binary outcome.

Treatment Effects on Women's Time-Use

Daily wages for female laborers in groundwater-rich locations may have mechanically declined because of shorter workdays. I test this using the REDS 1982 Demographic Questionnaire and find a workday explanation to be unlikely.[40] In this survey module, women provided an hourly accounting of their activity over a representative day for each of the three agricultural seasons.[41] Table 6 offers results for the main agricultural season, taken as the harvest period (November) for areas primarily growing rice and the sowing period (October) for wheat, but results across all three seasons are largely comparable. Models in Panel A are subset to women reporting positive time performing the described tasks, while Panels B and C include the full, unconditional sample. While women in treated locations both work fewer hours (A1) and are less likely to spend any time working (B1), treated village wage laborers (A2) did not spend significantly fewer hours per day in wage labor; the point estimate of -0.09 hours, is only slightly greater than 1 percent of the sample average workday length of 7.4 hours.[42] In lieu of agricultural and paid work, women in groundwater-rich villages spent substantially more time in domestic, unpaid activities like food preparation, cooking, cleaning, fuel and water collection, and child-care (Panel C). Notably, there is no treatment effect on reported leisure hours (C5), indicating that within-day intensive margin labor supply differences across households are largely driven by differences in labor activity, not in total time expenditure.

[40] Comparable data was not collected in the REDS 1970 round.

[41] Male time-use data was not collected in this survey instrument, ruling out male-female workday length comparisons.

[42] The wage labor category does not differentiate between agricultural and non-agricultural work, but 79 percent of women who reported positive time in this category lived with heads of household reporting farming, fishing, and hunting as a primary occupation.

6.4 Wheat Substitution and Negative Female Labor Demand

The Green Revolution's effect on yield growth was uneven across crops, as observed earlier. Wheat by and large registered the largest gains, partly because of pre-existing infrastructure advantages in areas already growing wheat before 1966, but primarily because the HYV seeds for wheat were simply more productive than for other crops. Figure 7a compares the average district-level wheat/rice yield ratio over five-year periods for before (*x-axis*) and after 1966 (*y-axis*). The majority of observations lie above the 45° line and indicate relatively stronger wheat yield growth, both in locations where pre-1966 wheat yields were lower than rice yields (left of 1) and higher (right of 1). Inclusion in this plot requires having grown both crops before and after 1966, but examining all districts that grew either rice *or* wheat (before and after 1966) generates a similar pattern, as pictured in Figure 7b. These results are unlikely to be driven by groundwater access alone, since subsetting to districts with below-median groundwater ($<$ 20 percent) produces the same pattern.

In response, farmers in treated areas consequently intensified wheat production, increasing the area share under wheat as seen in Figure 7c. By 1982, wheat's share in treated districts averaged 13.3 percentage points ($p < 0.01$) higher than in 1965. The stable pre-trend indicates this response was not driven by an underlying secular trend. In contrast, the share of cropped area under rice did not start rising until the mid-1970s, reaching only 2.4 percentage points ($p = 0.227$) above 1965-levels in 1982 (Figure 7d).

Wheat's rising prominence affected labor markets not only through the wage channel, but also in having a gender-biased effect on labor demand. Wheat production has historically been understood as male labor-biased, particularly in comparison to rice, since a larger share of all work relies on upper body strength and there is less demand for the sowing and weeding which differentiate rice production as more female labor demanding (Boserup, 1970; Bardhan, 1984).

Evidentiary support for this claim has often been limited by data availability. I test this claim using household-level data from the REDS 1999 round. Data on the number of crop-specific worker-days employed for each agricultural operation (e.g., sowing/transplanting, irrigating, etc.) for each worker-type (e.g., casual laborer, permanent laborer, family laborer) were collected from 4,612 cultivating households. I use this data to test whether wheat is more male-labor intensive by

running the following specification, for a unit of analysis at the level of p plot operated by household h,

$$\frac{\text{worker-days}^m_{ph}}{\text{worker-days}_{ph}} = \beta_1 \mathbb{1}(\text{Wheat})_{ph} + \beta_2 \mathbb{1}(\text{Other})_{ph} + \boldsymbol{X}'_{ph}\xi + \gamma_h + \epsilon_{ph}$$

which controls for plot size and land ownership status, and includes household fixed effects. Observations are weighted using household sampling weights and standard errors are clustered by village.

Table 7 presents results with β_1 the coefficient of interest, interpreted as effect of growing wheat relative to rice (omitted) on the male share of total labor demanded. Column (1) indicates that wheat production involves an 8.8 percentage point ($p < 0.026$) higher intensity of male labor than rice production. Since this may be driven by differences in female labor supply across households growing rice and growing wheat, (2) subsamples to only households growing both rice and wheat and leads to a 6.6 percentage point estimate ($p < 0.014$). Since the production conditions facing the average farm in 1999 may be unlike those in the 1960s and 1970s, it is not immediately clear that these recent labor arrangements are informative about labor demand decisions made during the Green Revolution. To address this, I take various subsamples that better approximate the operating conditions of that time period. In Panel A, (3) drops households not using HYVs for any crop while (4) drops households in Punjab, Haryana, and Uttar Pradesh which were the locations where the Green Revolution was most successful. In Panel B, (1-3) respectively subsamples to households from villages not atop a groundwater aquifer, households not applying irrigation to another crop, and households that neither owned not rented mechanized farm equipment (e.g., tractor, disc harrow, power tiller, etc.). These results all corroborate relatively greater male labor share in wheat production than for rice. In this sample, rice trails wheat as the second most important crop for the *rabi* season, supporting the implicit assumption in these models that rice is the appropriate counterfactual of what would be planted had wheat been less successful. As an additional check, all non-wheat crops are pooled into a single omitted category against which the marginal effect of wheat on male labor share is estimated, and produces always positive effects of slightly smaller magnitudes than when benchmarked against rice alone, shown in Table A3.

Differences in total labor requirements across crops also appear to be an important factor. Consider Figure 8 which illustrates the distribution of total worker-days allocated to the production of a unit acre of rice and wheat. For nearly every major operation, the amount of work involved in rice cultivation, at least as measured in total worker-days, exceeds that for wheat. While these figures do not control for other factors which may be correlated with crop choice or location, restricting the sample to only households that do not use HYVs, or farms which neither possess nor hired mechanized equipment, produce very similar results. Additionally, focusing only on households that in the sample year grew rice *and* wheat reveals comparable labor demands for all activities aside from transplanting/sowing and weeding, both for which rice required an additional 7 worker-days. As a result, more intensive wheat production can plausibly be linked to lower overall labor requirements. Under conditions of male under-employment, available work may more likely be given to male workers, and mechanically generating an opportunity deficit for women laborers.

In Section 6.3, I estimated declining female wage labor participation in treated areas, using both district-level and individual-level data. Similarly, if the shift towards wheat disproportionately favored male labor, then presumably women would cut-back their participation rates. Table 8 offers supporting evidence using both the pre-period 1961 cross-section alone (Panel A) and the decadal panel from 1961-1981 (Panel B). In both sets of results, a higher wheat share is correlated with lower female wage labor participation, estimated at a drop of 2.1 percentage points for each 10 percentage point increase in wheat share. Since the 1971 census deviated significantly in how employment status was captured, with particularly large effects for women who tend to work fewer days per year to begin with, I also run a panel using 1961 and 1981 data and find similar results.

6.5 Alternative Channels

In this section I consider alternative mechanisms that could potentially drive the observed wage and labor responses, and reason why they offer a less compelling explanation than the preferred crop substitution channel.

Gendered Selection Across Tasks or Timing

Gendered sorting into peak and off-peak season labor, or across agricultural operations with different prevailing wages, may contribute to wage divergence. To address this, I use the operation-specific (o) wage ratio, $\frac{w_{odtm}}{w_{odtf}}$, as the dependent variable which implicitly subsamples to month $\times$ operation $\times$ location wage records reported for men and women.[43,44] For efficiency gains I further pool estimates into three-year intervals. Figure A2 indicates that a significant male premium persists under these restrictive conditions. In a generalized DID setup, this estimated premium is 9.3 percent (p=0.023) which is a smaller than under the triple difference specification, thereby limiting but not ruling out a sorting contribution to rising wage inequality. For example, men and women may be simultaneously engaged in harvesting, but differentially into higher- and lower-value crops which is a level of task disaggregation not available in any of the available datasets. Alternatively tasks may be divisible, with gendered division into sub-tasks.

Income Effects

Income effects may have contributed to women's withdrawal from wage labor, though in isolation would have led to wage increases in the absence of a demand shock. I examine this first at the district-level and then with REDS. Using unit area crop revenues as a proxy for agricultural income, I do not find statistically significant support that locations experiencing the largest revenue increases also saw reductions in FLFP. Using the 1970 and 1982 REDS rounds, Figure 9a uses binned scatterplots to illustrate the relationship between log total household expenditure, the running variable, and binary occupational outcomes.[45] In 9a which relates expenditure to agricultural wage labor participation, participation for both sexes smoothly declines across expenditure deciles, with male supply slightly more elastic. Whereas men in richer households are more likely to be self-

[43] Operations for which male and female wages may be simultaneously reported vary by time-frame. Before 1972, these operations are field labor, reaping & harvesting, sowing, weeding, and other labor. From 1973, only field labor and harvesting are available for both sexes.

[44] This presumes that wage availability in a given month is indicative of that task being performed in the stated time, which Kurosaki and Usami (2016) call into question given the occasional reporting of task wages outside the timeframes customary for a given location.

[45] All models partial out village $\times$ year fixed effects, household size, education attainment, marital status, and total land ownership.

employed farmers (9b), household expenditure cannot predict women's likelihood. In both figures, female participation rates are lower than male across the entire expenditure support. Assuming only the effect of a supply shock from higher-income females leaving the workforce, then women's wages would have had to increase or the FLFP of lower-income women would have risen. Figure 9d compares the extensive margin elasticities for women in 1970 and 1982, with all pre-log household expenditures deflated to 1973/74 Indian rupees, and there is little evidence of the clockwise pivoting of these curves over time that would be consistent with an income effect explanation.

Investments in Farm Machinery

I do find support for the Green Revolution spurring investment in improved farm equipment. Figure 10 indicates a steady increase in tractor purchases that accelerates in the late-1970s. Tractors are imperfect substitutes for bullocks, and their growth appears to be related to the reduction in bullocks.[46] However, the gradual rise in tractor usage appears to be inconsistent with the immediate post-1966 wage response. As well, the most immediate effect of introducing tractors would be to displace ploughmen, presumably depressing male wages. While it is *ex ante* ambiguous whether tractors are net labor-displacing or labor-augmenting (Billings and Singh, 1970), to be consistent with the reduction in female labor, the effect would have to be both labor-displacing and biased towards female-dominated tasks. In contrast, Mencher and Saradamoni (1982) claim that mechanization primarily displaced traditionally male tasks.

Manufacturing Sector Spillover Effects

Labor demand competition from non-agricultural sectors could plausibly have driven male laborer wage growth, but there is little evidence to support this.[47] First, Figure A3 demonstrates that in groundwater-rich locations, and in general, manufacturing accounts for a small share of total rural male employment, and therefore is unlikely to have influence over the significantly larger wage labor sector. Second, census employment data shows that manufacturing employment growth for

[46] Many of the operations performed by bullocks, such as ploughing, tilling, and transport can be performed by tractors in less time.

[47] The manufacturing share of female labor is substantially lower than for men.

men and women was slower in Green Revolution locations, consistent with (Foster and Rosenzweig, 2004) who find higher activity in the tradeables sector where agricultural wages were low.

6.6 Robustness Checks

This section describes robustness checks that address plausible threats to identification, and test estimate stability under alternative specifications or samples of the data.

Threats of State-Specific Effects

There are two primary concerns about the role of states. First, high spatial correlation in groundwater endowments as observed in Figure 3 suggests that time-varying, state-specific influences like input subsidy programs or infrastructure investments are potential confounds.[48] I address this by including state × year fixed effects, which are more flexible than state linear or quadratic trends. In columns (2) of Table 11. This dampens the point estimate for the log wheat yield model from 0.198 to 0.116, but also shrinks the standard errors and is significant at the 5 percent level. Including these fixed effects increases the point estimates for the log rice yield and log gross unit area revenue. In columns (3) of Table 12, adding state × year dummies reduces the male wage premium estimates, which remain significant at the 5 percent level.

Second, Punjab and Haryana are widely considered the most successful Green Revolution states. They had better irrigation facilities to begin with and had already been planting wheat before the HYVs were introduced. To what extent are the findings due exclusively to these two states? Their removal from the sample negatively affects the point estimates in the agricultural outcomes models, halving the treatment effect on rice. Regardless, the revenue effects remain substantial at 17 percent (dropping Punjab) and 15 percent (dropping Punjab and Haryana). Dropping them from the wage sample in Table 12 does not much impact the point estimates.

[48] Any time-invariant state-level feature is absorbed in the district fixed effects for all district-level panel models.

Sensitivity Tests of Male Wage Premium to Model Specification and Wage Construction

Agricultural wages are published at the level of gender $\times$ month $\times$ operation $\times$ village (reporting center) $\times$ year. Earlier results were from a district fixed effects specification, whereas the data permits including village fixed effects. One ensuing complication is village composition is more unbalanced; a village may report in fewer than 20 percent of all years. Columns (2) in Table 12 show this addition has little effect. A related feature is that some villages do not begin reporting until right before 1966 or even afterwards. As a result, these locations offer no pre-treatment information and possibly bias the male wage premium estimate. To address this, Panel B includes only those villages which reported wage data before 1960, and therefore offer a more robust set of pre-treatment observations. For each column, the estimate size with this subsample increases, though it should be noted that Uttar Pradesh drops entirely from the sample in (7). Results are also not driven by a single agricultural operation; following a leave-one-out procedure always generates a significant male wage premium in the range of values shown in Table 12.

Unbalancedness also arises at the observation unit. As an example, a gender $\times$ operation wage may be listed in month m for year t, but not for $t+1$. This creates the adverse effect of some locations reporting more observations per year than other locations. To partially address this, I perform two types of averaging. The first involves averaging over all relevant wages to calculate an average gender $\times$ district $\times$ year wage. The second replicates this process at village-level. Both of these processes obviate concerns that results are driven by overweighting from locations with more frequent reporting. The respective results are in columns (3) and (4) of Table A4 and can be compared against baseline values in (1) and (2). For both the full sample and the truncated sample of villages reporting wage data before 1966, the averaging procedure decreases the male wage premium point estimate, but in all specifications retains statistical significance at less than 1 percent.

Conley Spatial-Adjusted Standard Errors

The main specification allows for arbitrary serial correlation of errors across observations for districts from the same National Sample Survey region. This approach assumes no correlation in the error structure for districts in different regions, even though they may be geographically proximate. Spatially-corrected Conley standard errors relax this assumption, and allow for error correlation between locations within a given distance $\times$ time lag from each other, using a Bartlett cutoff kernel (Conley, 1999; Hsiang, 2010). The top panels of Table 9 and 10 respectively present estimated standard errors for models with log unit area revenue and log daily wage as dependent variables when clustered by district, region, and state. As the spatial scale of clustering grows, standard errors on the specified treatment interaction widen, yet remain significant at the 1 percent level.[49] The lower panels demonstrate how standard errors respond to changes in distance and lag length cutoffs, and remain highly significant at below the 1 percent level for all distance $\times$ lag combinations.

Randomization Inference

What is the likelihood that estimated effects are a product of chance or model misspecification? I test this by employing *randomization inference* and generating groundwater 'pseudo-samples' which randomly re-assign district-level groundwater coverage values among districts in the same state. Under this approach, all samples possess the same within-state distribution of groundwater values as in the true data sample, but are randomly shuffled in each iteration.[50] Since my specifications use a time-invariant groundwater variable, the re-assigned values are held constant throughout the entire time period. Consider reassignment for only two districts $i, j,\ i \neq j$ and any periods $t,\ t+\Delta$. If G_i, G_j are the true groundwater coverage values, reassignment implies the treatment permutation of $G_j^{'} = G_i, G_i^{'} = G_j$ for all $t,\ t+\Delta$ periods. This provides an exact test statistic that avoids relying on asymptotic theory and distributional assumptions (Gerber and Green, 2012; Young, 2016). Figures 11a and 11b respectively depict the β sampling distribution from 10,000 within-

[49] In the absence of block-bootstrapping, the estimated standard errors in columns (3) are likely too conservative given the small number of clusters (Cameron and Miller, 2015), but is presented without adjustment for purposes of comparison.

[50] This procedure is also referred to as *permutation inference*. The Green Revolution's simultaneous introduction across the entire country prevents random assignment in the time dimension.

state groundwater permutations for models with log gross unit area revenue (treatment variable: $\mathbb{1}\{t \geq 1966\} \times \text{Groundwater}^r$ %) and log wage (treatment variable: $\mathbb{1}\{t \geq 1966\} \times \text{Groundwater}^r$ % $\times$ $\mathbb{1}$ { Male }) as the dependent variable, where r superscripts denote randomly assigned groundwater values. In Figure 11a, the 'true' β of 0.214 lies at the 95th percentile. The sampling distribution is centered at 0.15 instead of zero, largely because of limited variation of within-state groundwater values. For example, a model consisting only of state fixed effects explains more than two-thirds of the cross-sectional variation in groundwater values. Permutations will therefore re-assign positive groundwater values across districts within a state, and have a limited null effect for states where cross-district groundwater variation is small. By comparison, consider an experiment where half the observations belong to a control group, and the other half the treatment group. For some number of treatment re-assignments, all true control group members will be treated and treatment group members untreated. This configuration creates the mirror opposite of the true treatment assignment, and therefore should generate the strongest null result, assuming the treatment is responsible for observed effects. In regards to re-assignment of groundwater values, the absence of mirror opposite treatment assignment positively shifts the sampling distribution. The distribution of pseudo-sample male wage premium estimates in Figure 11b is centered at 0 and has wider tails. The true effect size is at the 96th percentile of this distribution.

Effects from Other Geophysical Characteristics

I consider the possibility that omitted variables, unaccounted for in the specification which includes state $\times$ year fixed effects, may be positively biasing groundwater estimates, especially since groundwater volumes may be a function of surface water recharge. In Table 11, columns (3) and (4) cumulatively add variables capturing geophysical characteristics which are fully-interacted with year dummies, providing maximal flexibility in their effect. Column (3) addresses the potential role that surface water irrigation may have played in exploiting the HYVs, and allows for the possibility that effects are not driven by groundwater access, but rather a broader set of irrigation sources. In this specification, the share of district area within 20 kilometers of a named river is

included as an explanatory variable that is interacted with the full set of year fixed effects.[51] It is unlikely that irrigation projects are diverting river water that far away from source, and therefore this captures a likely upper bound on the share of irrigable land. Including these variables has no effect on the groundwater estimate. In (4), I add 26 soil and slope dummies coded in the World Bank dataset to incorporate non-hydrological characteristics that could influence HYV adoption and output productivity. Their inclusion dampens the treatment effect magnitude on wheat yields and crop revenue, but increases it on rice yields. In each of these models, estimated treatment effects when including soil and slope dummies are within one standard deviation of the benchmark model estimate.

7 Conclusion

In this paper I have estimated the Green Revolution's effect on gender wage inequality in Indian agricultural labor markets. Using plausibly exogenous access to groundwater resources to proxy for Green Revolution intensity, I find that areas rich with groundwater were better suited to exploiting the newly available high-yielding varieties and experienced significant growth in yields, crop revenues, and input intensity. Using annual, district-level wage data I find the consequent labor demand shock was male-biased, and widened pre-existing gender wage inequality. Multiple, individual-level datasets on labor allocation, time-use, and earnings provide support for a mechanism that links wheat production to women's labor outcomes. Since wheat is comparatively male labor-intensive, the rising concentration of wheat production in response to significant wheat yield growth depressed female labor demand, and reduced women's wages and employment shares. I find limited empirical support for alternative explanations that would generate results of the correct sign and timing, from plausible channels like manufacturing employment growth, in-migration, changes in the laborer workforce composition, or investments in farm mechanization.

This paper offers three key implications. First, technical change can create gender-biased employment effects without differentially modifying women's and men's labor productivity. Instead, by

[51] These values are constructed using the India inland water shapefile originally sourced from Digital Chart of the World and available at http://www.diva-gis.org/datadown.

increasing the productivity of a pre-existing gender-biased technology more than a gender-neutral one, producers shift towards the former with the outcome approximating gender-biased technical change. Second, women's substitution of paid labor with unpaid cultivation and domestic work shows that the declines in female labor force participation conventionally associated with manufacturing growth in the feminization U-shaped hypothesis can result from changes in agriculture alone without appeal to manufacturing. In this context, there is limited growth in firm employment, yet women's participation still registered a sizable reduction. Lastly, since expenditures on children's health and education may respond to the division of income between husband and wife, this setting casts technological change as potentially contributing to longer-term, non-labor consequences. Examining how the distributional effects of the Green Revolution may have led to changes in human capital investments or fertility behavior are important and promising areas for future research.

These findings are particularly salient given the current promotion of Green Revolution policies and technologies throughout sub-Saharan Africa, where indigenous crops had largely been passed over in the first-generation development of HYVs (Pingali, 2012; Dawson et al., 2016). One potential means of reducing the likelihood of productivity gains exacerbating inequalities would be to target R&D efforts on crops that more intensively employ women, or are more likely to be farmed by women. While Doss (2002) finds limited support using survey data from Ghana for a clean differentiation of 'male' and 'female' crops, she offers evidence that staple crops are more likely to be grown by women. Assuming fixed gender roles that are not responsive to yield improvements, a tenuous claim given evidence from Burkina Faso (Doss, 1999), gains targeted to staples would more likely reduce gender inequality. In striving for progress on Sustainable Development Goals 1 (ending poverty), 2 (achieving food security), and 5 (achieving gender equality), initiatives like the Alliance for a Green Revolution in Africa will likely encounter trade-offs, as satisfaction of some goals may undermine others. Further research is needed to determine if the technologies and practices that are most effective in strengthening food security come at the expense of women's outcomes, as in India's Green Revolution. If such is the case, then multiple policy instruments may be necessary to simultaneously increase agricultural productivity and further promote women's economic empowerment.

References

Acemoglu, Daron, "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality," *The Quarterly Journal of Economics*, 1998, *113* (4), 1055–1089.

_ , "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 2002, *40*, 7–72.

_ , "Labor- and Capital-Augmenting Technical Change," *Journal of the European Economic Association*, 2003, *1* (1), 1–37.

_ **and David Autor**, "Skills, Tasks and Technologies: Implications for Employment and Earnings," in "Handbook of Labor Economics," Elsevier B.V., 2011, pp. 1043–1171.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

Auffhammer, Maximilian, V Ramanathan, and Jeffrey R Vincent, "Climate Change, the Monsoon, and Rice Yield in India," *Climatic Change*, 2012, *111* (2), 411–424.

Autor, David H, Frank Levy, and Richard J Murname, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118* (4), 1279–1333.

_ **, Lawrence F Katz, and Melissa S Kearney**, "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 2008, *90* (2), 300–323.

Bacolod, Marigee P and Bernardo S Blum, "Two Sides of the Same Coin: US "Residual" Inequality and the Gender Gap," *The Journal of Human Resources*, 2010, *45* (1), 197–242.

Bardhan, Pranab K, "Size, Productivity, and Returns to Scale: An Analysis of Farm-Level Data in Indian Agriculture," *Journal of Political Economy*, 1973, *81* (6), 1370–1386.

_ , *Land, Labor, and Rural Poverty: Essays in Development Economics*, New York: Columbia University Press, 1984.

Barrett, Christopher B, Michael R Carter, and C Peter Timmer, "A Century-Long Perspective on Agricultural Development," *American Journal of Agricultural Economics*, 2010, *92* (2), 447–468.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry, "Profitability of Fertilizer: Experimental Evidence from Female Rice Farmers in Mali," *The American Economic Review*, 2013, *103* (3), 381–386.

Besley, Timothy and Anne Case, "Modeling Technology Adoption in Developing Countries," *The American Economic Review*, 1993, *83* (2), 396–402.

_ **and Robin Burgess**, "Land Reform, Poverty Reduction, and Growth: Evidence from India," *The Quarterly Journal of Economics*, 2000, *115* (2), 389–430.

Billings, Martin H and Arjan Singh, "Mechanisation and the Wheat Revolution: Effects on Female Labour in Punjab," *Economic and Political Weekly*, 1970, *5* (52), A169–A174.

Black, Sandra E and Alexandra Spitz-Oener, "Explaining Women's Success: Technological Change and the Skill Content of Women's Work," *The Review of Economics and Statistics*, 2010, *92* (1), 187–194.

Boserup, Ester, *Women's Role in Economic Development*, Earthscan, 1970.

Byerlee, Derek, Alain de Janvry, and Elisabeth Sadoulet, "Agriculture for Development: Toward a New Paradigm," *Annual Review of Resource Economics*, October 2009, *1* (1), 15–31.

Cameron, A Colin and Douglas J Miller, "A Practitioner's Guide to Cluster-Robust Inference," *The Journal of Human Resources*, 2015, *50* (2), 317–372.

Card, David and John E DiNardo, "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 2002, *20* (4), 733–783.

Carter, Michael R, Rachid Laajaj, and Dean Yang, "Subsidies and the Persistence of Technology Adoption: Field Experimental Evidence from Mozambique," *NBER Working Paper Series*, 2014, pp. 1–45.

Chakravarti, A K, "Green Revolution in India," *Annals of the Association of American Geographers*, 1973, *63* (3), 319–330.

Chinnappa, B Nanjamma, "Adoption of the New Technology in North Arcot District," in B H Farmer, ed., *Green Revolution?*, Boulder, CO: Westview Press, 1977, pp. 92–123.

Cleaver, Harry M, "The Contradictions of the Green Revolution," *The American Economic Review*, 1972, *62* (1/2), 177–186.

Conley, Timothy G, "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92*, 1–45.

Dawson, Neil, Adrian Martin, and Thomas Sikor, "Green Revolution in Sub-Saharan Africa: Implications of Imposed Innovation for the Wellbeing of Rural Smallholders," *World Development*, 2016, *78*, 204–218.

Desai, Sonalde and Devaki Jain, "Maternal Employment and Changes in Family Dynamics: The Social Context of Women's Work in Rural South India," *Population and Development Review*, 1994, *20* (1), 115–136.

Dhanagare, D N, "Green Revolution and Social Inequalities in Rural India," *Economic and Political Weekly*, 1987, *22* (19/21), AN137–AN139–AN141–AN144.

Doss, Cheryl R, "Twenty-Five Years of Research on Women Farmers in Africa: Lessons and Implications for Agricultural Research Institutions," Technical Report 99-02, CIMMYT, Mexico D.F. 1999.

_ , "Men's Crops? Women's Crops? The Gender Patterns of Cropping in Ghana," *World Development*, 2002, *30* (11), 1987–2000.

Duflo, Esther, "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa," *The World Bank Economic Review*, 2003, *17* (1), 1–25.

_ , "Women Empowerment and Economic Development," *Journal of Economic Literature*, 2012, *50* (4), 1051–1079.

_ **, Michael Kremer, and Jonathan Robinson**, "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, 2011, *101* (6), 2350–2390.

Dyson, Tim and Mick Moore, "On Kinship Structure, Female Autonomy, and Demographic Behavior in India," *Population and Development Review*, 1983, *9* (1), 35–60.

Evenson, R E and D. Gollin, "Assessing the Impact of the Green Revolution, 1960 to 2000," *Science*, 2003, *300*, 758–762.

FAO, "The State of Food and Agriculture: 2010-11," Technical Report, Rome, Italy 2011.

Fernando, A Nilesh, "Shackled to the Soil: The Long-Term Effects of Inherited Land on Labor Mobility and Consumption," 2016, pp. 1–83.

Fink, Gunther, B Kelsey Jack, and Felix Masiye, "Seasonal Credit Constraints and Agricultural Labor Supply: Evidence from Zambia," *NBER Working Paper Series*, 2014, pp. 1–44.

Fortin, N M, "Gender Role Attitudes and the Labour-Market Outcomes of Women Across OECD Countries," *Oxford Review of Economic Policy*, 2005, *21* (3), 416–438.

Foster, Andrew D and Mark R Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 1995, *103* (6), 1176–1209.

_ **and** _ , "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution," *The American Economic Review*, 1996, *86* (4), 931–953.

_ **and** _ , "Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970–2000," *Economic Development and Cultural Change*, 2004, *52* (3), 509–542.

_ **and** _ , "Microeconomics of Technology Adoption," *Annual Review of Economics*, 2010, *2* (1), 395–424.

Frankel, Francine R, *India's Green Revolution: Economic Gains and Political Costs*, Princeton, NJ: Princeton University Press, 1971.

Freebairn, Donald K, "Did the Green Revolution Concentrate Incomes? A Quantitative Study of Research Reports," *World Development*, 1995, *23* (2), 265–279.

Galor, Oded and David N Weil, "The Gender Gap, Fertility, and Growth," *The American Economic Review*, 1996, *86* (3), 374–387.

Gerber, Alan S and Donald P Green, *Field Experiments*, New York: W. W. Norton & Company, 2012.

Goldin, Claudia, *Understanding the Gender Gap* An Economic History of American Women, New York: Oxford University Press, 1990.

_ , "The U-Shaped Female Labor Force Function in Economic Development and Economic History," *NBER Working Paper Series*, 1994, pp. 1–49.

Gollin, Douglas, Stephen Parente, and Richard Rogerson, "The Role of Agriculture in Development," *The American Economic Review*, 2002, *92* (2), 160–164.

Guiteras, Raymond, "The Impact of Climate Change on Indian Agriculture," 2009.

Hazell, Peter B R, "The Asian Green Revolution," *IFPRI Discussion Paper*, 2009, pp. 1–40.

Heath, Rachel and Seema Jayachandran, "The Causes and Consequences of Increased Female Education and Labor Force Participation in Developing Countries," *NBER Working Paper Series*, 2016, pp. 1–28.

Hsiang, Solomon M, "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America," *Proceedings of the National Academy of Sciences*, 2010, *107* (35), 15367–15372.

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

Johnston, Bruce F, "Agriculture and Structural Transformation in Developing Countries: A Survey of Research," *Journal of Economic Literature*, 1970, *8* (2), 369–404.

_ **and John W Mellor**, "The Role of Agriculture in Economic Development," *The American Economic Review*, 1961, *51* (4), 566–593.

Krueger, Alan B, "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," *The Quarterly Journal of Economics*, 1993, *108* (1), 33–60.

Kumar, Hemanshu and Rohini Somanathan, "State and District Boundary Changes in India: 1961-2001," *SSRN Electronic Journal*, 2015, pp. 1–61.

Kumar, K Krishna, K Rupa Kumar, R G Ashrit, N R Deshpande, and James Hansen, "Climate impacts on Indian agriculture," *International Journal of Climatology*, 2004, *24* (11), 1375–1393.

Kurosaki, Takashi and Yoshifumi Usami, "A Note on the Reliability of Agricultural Wage Data in India: Reconciliation of Monthly AWI Data for District-Level Analysis," *Review of Agrarian Studies*, 2016, *6* (1), 6–38.

Lewis, W Arthur, "Economic Development with Unlimited Supplies of Labor," *The Manchester School*, 1954, *22* (2), 139–191.

Lindley, Joanne, "The Gender Dimension of Technical Change and the Role of Task Inputs," *Labour Economics*, 2012, *19* (4), 516–526.

Mammen, Kristin and Christina Paxson, "Women's Work and Economic Development," *Journal of Economic Perspectives*, 2000, *14* (4), 141–164.

Mbiti, Isaac M, "Monsoon Wedding? The Effect of Female Labor Demand on Marriage Markets in India," 2008, pp. 1–45.

Mencher, Joan P and K Saradamoni, "Muddy Feet, Dirty Hands: Rice Production and Female Agricultural Labour," *Economic and Political Weekly*, 1982, *17* (52), A149–A153–A155–A167.

Miller, Barbara D, "Female Labor Participation and Female Seclusion in Rural India: A Regional View," *Economic Development and Cultural Change*, 1982, *30* (4), 777–794.

Munshi, Kaivan, "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics*, 2004, *73* (1), 185–213.

_ **and Mark Rosenzweig**, "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *American Economic Review*, January 2016, *106* (01), 46–98.

Palmer, Ingrid, "Science and Agricultural Production," United Nations Research Institute for Social Development Geneva 1972.

Paris, Thelma R, "Technology and Policy Needs of Poor Women in Asian Rice Farming," *Gender, Technology and Development*, 1998, *2* (2), 187–218.

Pingali, P L, "Green Revolution: Impacts, Limits, and the Path Ahead," *Proceedings of the National Academy of Sciences of the United States of America*, 2012, *109* (31), 12302–12308.

Pinstrup-Andersen, Per and Peter B Hazell, "The Impact of the Green Revolution and Prospects for the Future," *Food Reviews International*, 1985, *1* (1), 1–26.

Pitt, Mark M, Mark R Rosenzweig, and Mohammad Nazmul Hassan, "Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy," *American Economic Review*, 2012, *102* (7), 3531–3560.

Prahladachar, M, "Income Distribution Effects of the Green Revolution in India: A Review of Empirical Evidence," *World Development*, 1983, *11* (11), 927–944.

Qian, Nancy, "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance," *The Quarterly Journal of Economics*, 2008, *123* (3), 1251–1285.

Radhakrishnan, P N, "Management of HYVP," *Economic and Political Weekly*, 1969, *4* (4), 249–251.

Rendall, Michelle, "Brain Versus Brawn: The Realization of Women's Comparative Advantage," *Institute for Empirical Research in Economics Working Paper Series*, 2010, pp. 1–27.

_ , "Structural Change in Developing Countries: Has it Decreased Gender Inequality?," *World Development*, 2013, *45* (C), 1–16.

Rodell, Matthew, Isabella Velicogna, and James S Famiglietti, "Satellite-Based Estimates of Groundwater Depletion in India," *Nature*, 2009, *460* (7258), 999–1002.

Rosenzweig, Mark R, "Educational Subsidy, Agricultural Development, and Fertility Change," *The Quarterly Journal of Economics*, 1982, *97* (1), 67–88.

_ , "Population Growth and Human Capital Investments: Theory and Evidence," *Journal of Political Economy*, 1990, *98* (5), S38–S70.

_ **and T Paul Schultz**, "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India," *The American Economic Review*, 1982, *72* (4), 803–815.

Roy, Shyamal, "Profitability of HYV Paddy Cultivation," *Economic and Political Weekly*, 1971, *6* (26), A75–A78.

Rud, Juan Pablo, "Electricity Provision and Industrial Development: Evidence from India," *Journal of Development Economics*, 2012, *97* (2), 352–367.

Schlenker, Wolfram and Michael J Roberts, "Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields Under Climate Change," *Proceedings of the National Academy of Sciences*, 2009, *106* (37), 11594–11598.

Sekhri, Sheetal, "Public Provision and Protection of Natural Resources: Groundwater Irrigation in Rural India," *American Economic Journal: Applied Economics*, October 2011, *3* (4), 29–55.

Sengupta, S and M G Ghosh, "HYVP for Rice: Performance in a Bengal District," *Economic and Political Weekly*, 1968, *3* (43), A26–A28.

Srivastava, A K, M Rajeevan, and S R Kshirsagar, "Development of a High Resolution Daily Gridded Temperature Data Set (1969-2005) for the Indian Region," *Atmospheric Science*

Letters, 2009, *10* (4), 249–254.

Stevenson, James R, Nelson Villoria, Derek Byerlee, Timothy Kelley, and Mywish Maredia, "Green Revolution Research Saved an Estimated 18 to 27 Million Hectares from Being Brought into Agricultural Production," *Proceedings of the National Academy of Sciences*, 2013, *110* (21), 8363–8368.

Swaminathan, M S, "Scientific Implications of HYV Programme," *Economic and Political Weekly*, 1969, *4* (1/2), 67–69–71–73–75.

Thomas, Duncan, "Intra-Household Resource Allocation: An Inferential Approach," *The Journal of Human Resources*, 1990, *25* (4), 635–664.

Timmer, C Peter, "The Agricultural Transformation," in H Chenery and T N Srinivasan, eds., *Handbook of Development Economics*, Elsevier BV, 1988, pp. 275–331.

_ , "Agriculture and Economic Development," in B Gardner and G Rausser, eds., *Handbook of Agricultural Economics*, Elsevier Science B.V., May 2002, pp. 1488–1546.

UN Women, "Progress of the World's Women 2015-2016," Technical Report 2015.

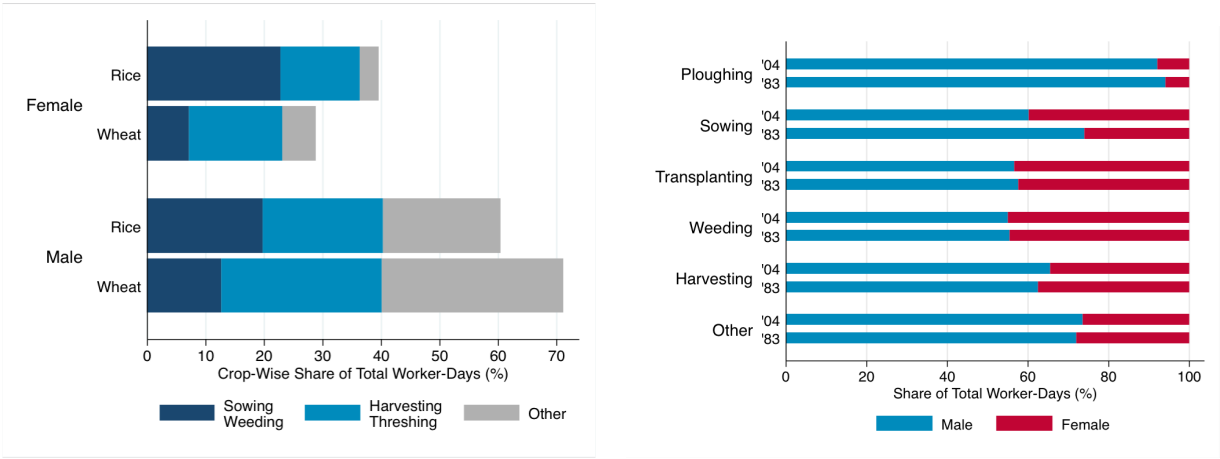
Usami, Yoshifumi, "Agricultural Wages in India, 1973-74 to 1994-95," *Statistics on Indian Economy and Society*, 2014.

Vanneman, Reeve and Douglas Barnes, "Indian District Data, 1961-1991: Machine-Readable Data File and Codebook," 2000.

Welch, Finis, "Growth in Women's Relative Wages and in Inequality Among Men: One Phenomenon or Two?," *American Economic Review*, 2000, *90* (2), 444–449.

Young, Alwyn, "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," 2016, pp. 1–74.

Figure 1: Gender Division of Agricultural Labor by Activity



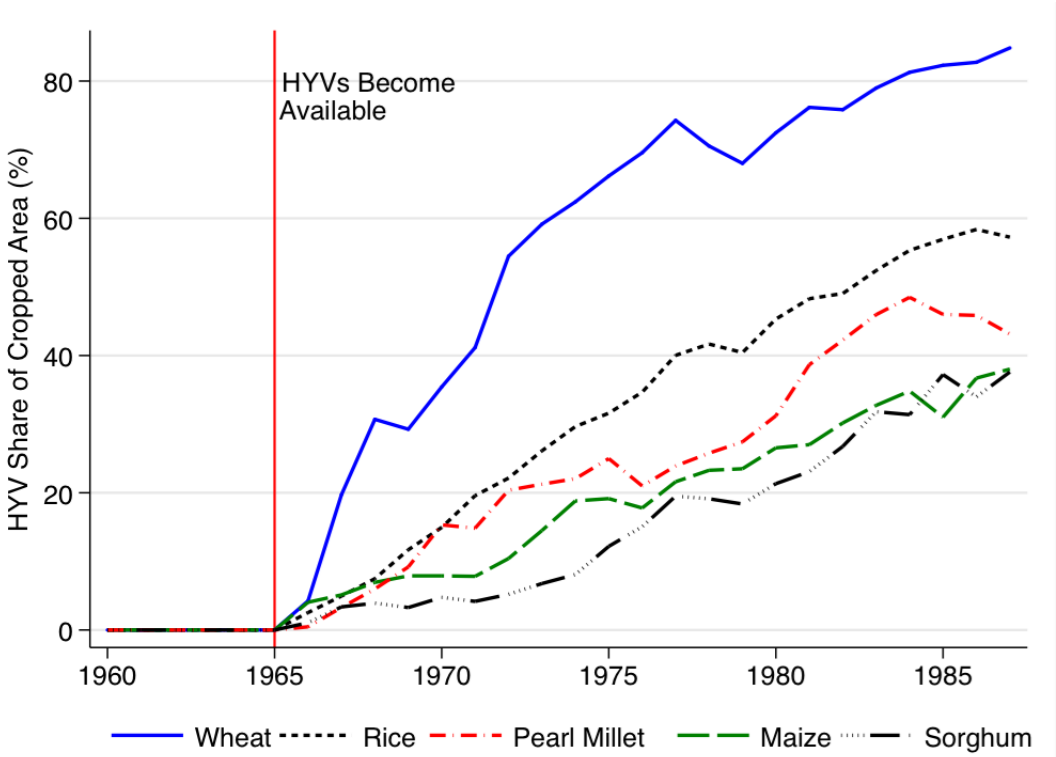
(a) Activity Division by Gender and Crop

(b) Activity Division by Gender

Notes: In (a), total worker-days are summed for each activity across all plots growing rice or wheat in the REDS 1999 sample. 'Other' activities include land preparation, fertilizer application, irrigation management, and an unspecified other category. In (b), total worker-days are summed across each operation for all observations in the National Sample Survey's Employment Unemployment Survey for 1983 (Round 38) and 2004 (Round 60). 'Other' activities are unspecified in the original data.

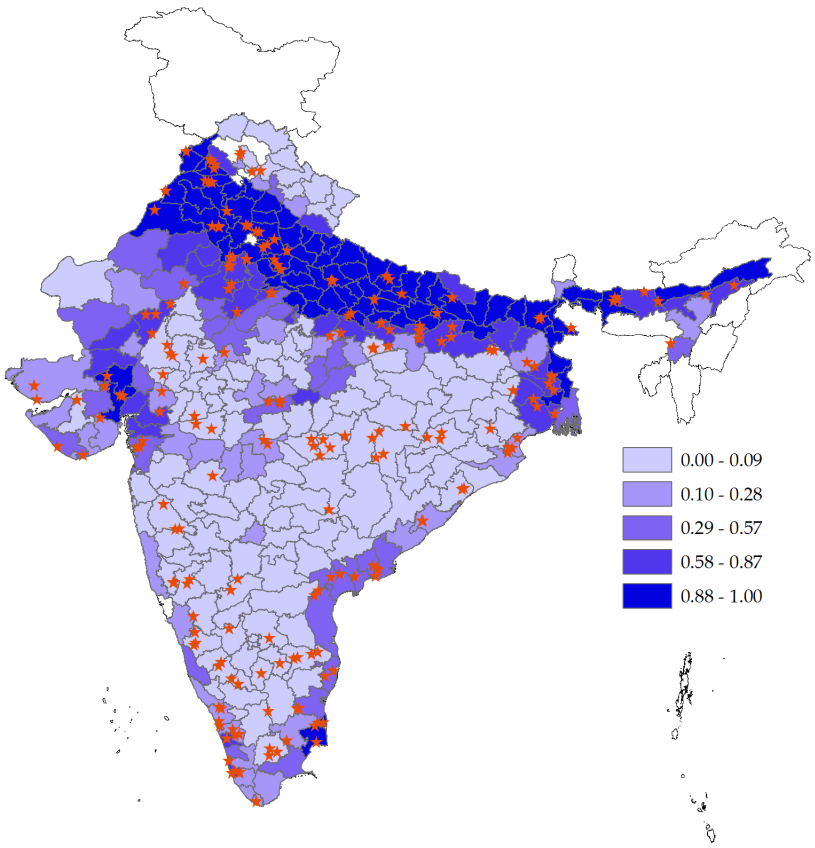
Data: REDS 1999; National Sample Survey Employment Unemployment Survey Rounds 38 (1983), 60 (2004)

Figure 2: Crop-Wise HYV Share of Gross Cropped Area



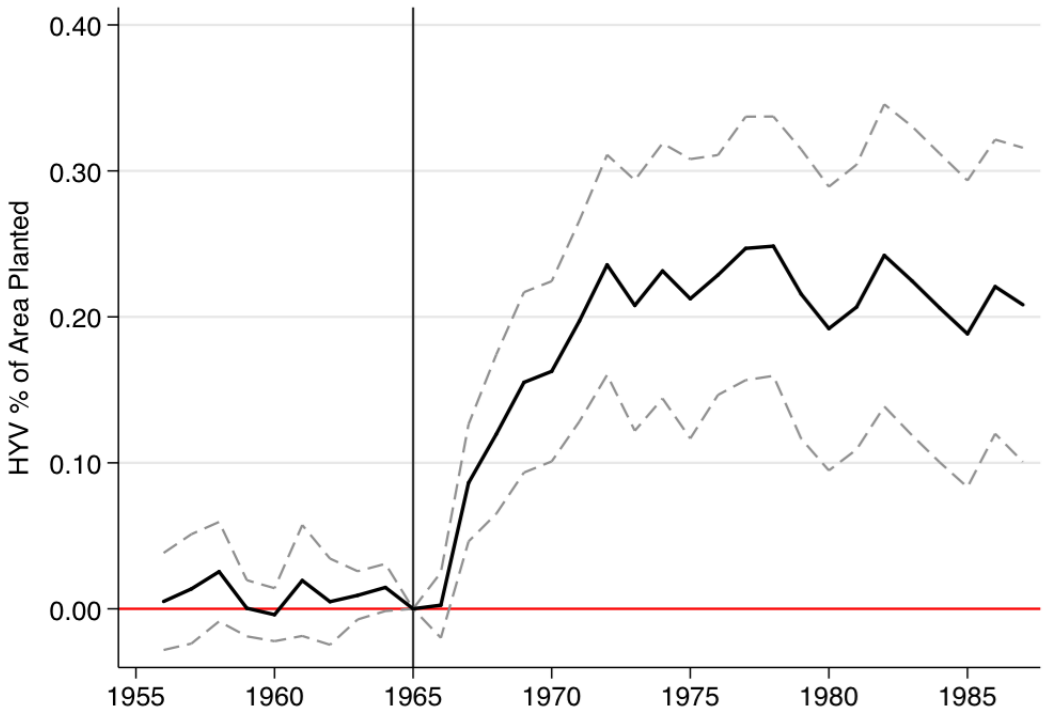
Notes: The crop-specific fraction of area cropped with high-yielding variety seeds (HYVs) is plotted on the *y-axis*. Values are totals for all of India.
Data: Ministry of Agriculture, Government of India

Figure 3: District Area Share Covering Unconsolidated Aquifers



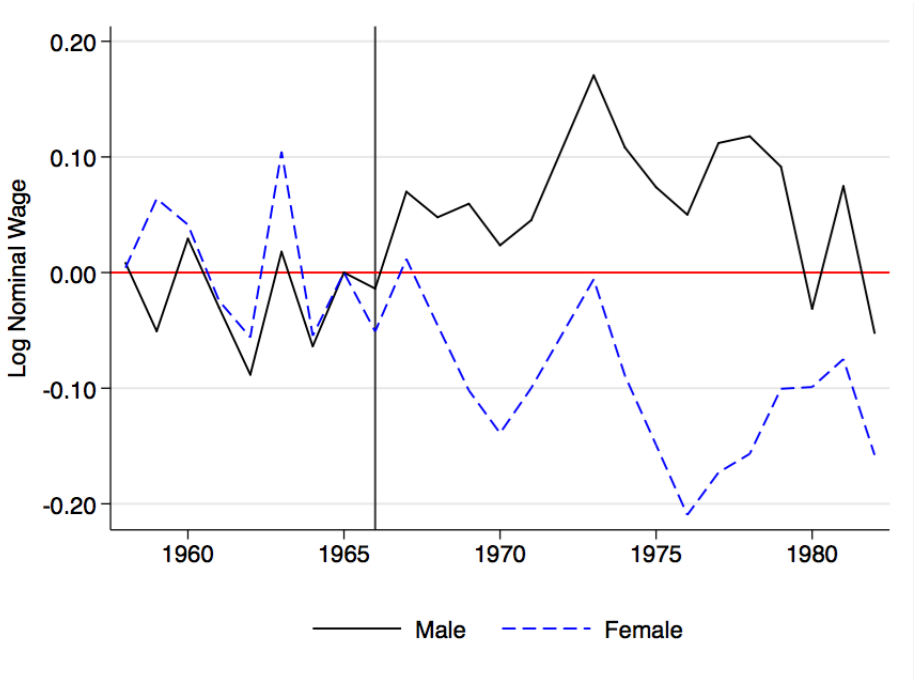
Notes: District groundwater coverage is a continuous measure and is computed as the share of district area covering unconsolidated groundwater formations using 1961 Census of India district boundaries. REDS villages are marked as red stars.
Data: Geohydrological Map of India 1969

Figure 4: Time-Varying Effect of Groundwater Coverage on HYV Coverage



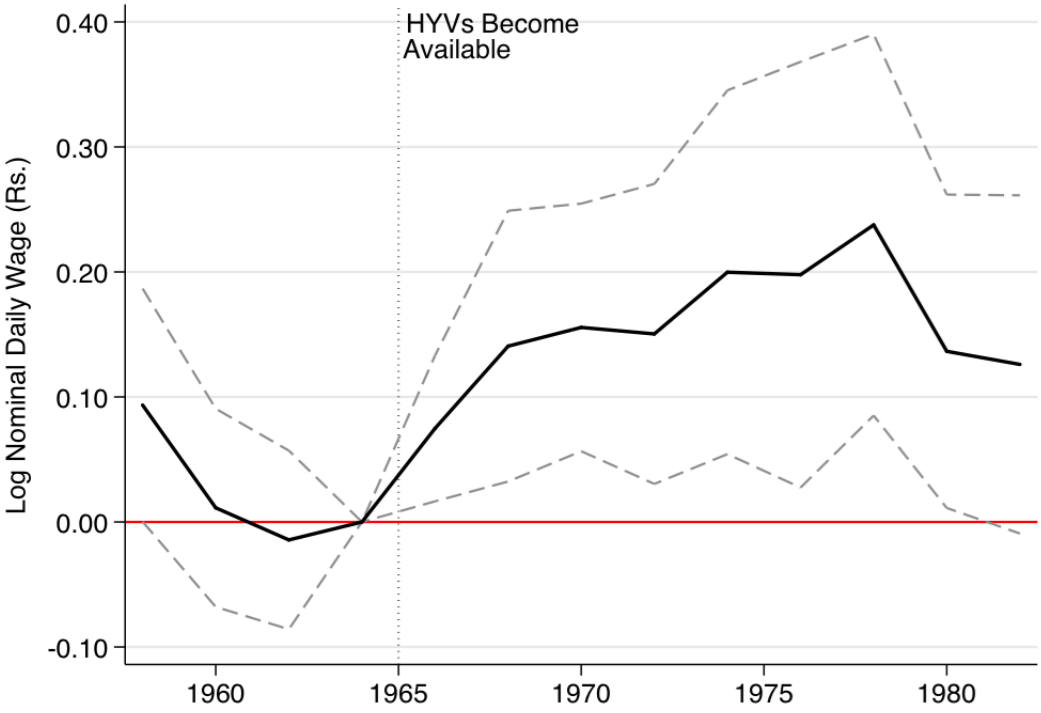
Notes: HYV coverage is defined as the combined share of rice, wheat, maize, sorghum, and pearl millet area planted with high-yielding varieties (HYVs). Plotted coefficients are year-varying effects of groundwater coverage, measured in percent [0,1], with the marginal effect interpreted in percentage point changes of HYV coverage. Model includes year and district fixed effects, and the set of current and lagged temperature and precipitation variables described in Section 4.4. Observations are weighted by 1961 gross cropped area and standard errors are clustered by 1983 National Sample Survey region. Dashed lines represent 95 percent confidence intervals. Results are for 271 districts from 1956-1987.
Data: World Bank Climate and Agriculture; Geohydrological Map of India 1969

Figure 5: Time-Varying Effect of Groundwater on Male and Female Agricultural Laborer Wages



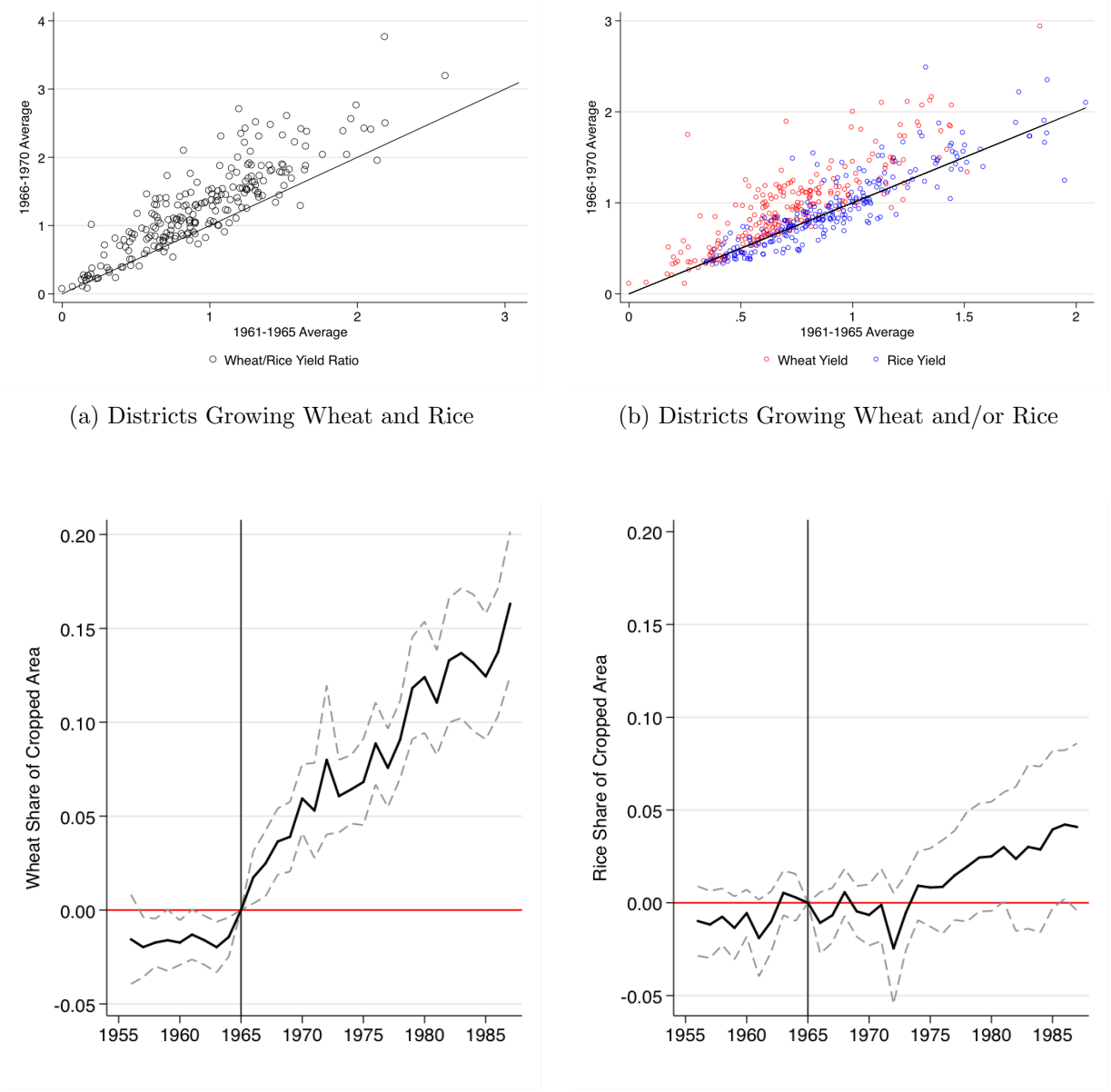
Notes: Coefficients from models separately regressing log nominal male and female wages on groundwater coverage interacted with year dummies. Both models include year, district, month of year, and agricultural operation fixed effects. Standard errors are clustered by 1983 National Sample Survey region.
Data: Agricultural Wages in India; Geohydrological Map of India 1969; Vanneman and Barnes (2000)

Figure 6: Time-Varying Effect of Groundwater on Male Laborer Wage Premium



Notes: Plotted coefficients are the marginal effect of the triple interaction of groundwater coverage, $\mathbb{1}$(male), and time dummies (pooled into two-year intervals) on log nominal wages, as specified in Equation (6.2). Model includes district, year, month of year, gender, and agricultural operation fixed effects. Standard errors are clustered by 1983 National Sample Survey region.
Data: Agricultural Wages in India; Geohydrological Map of India 1969; Vanneman and Barnes (2000)

Figure 7: Wheat and Rice Yield and Area Share Response



(a) Districts Growing Wheat and Rice

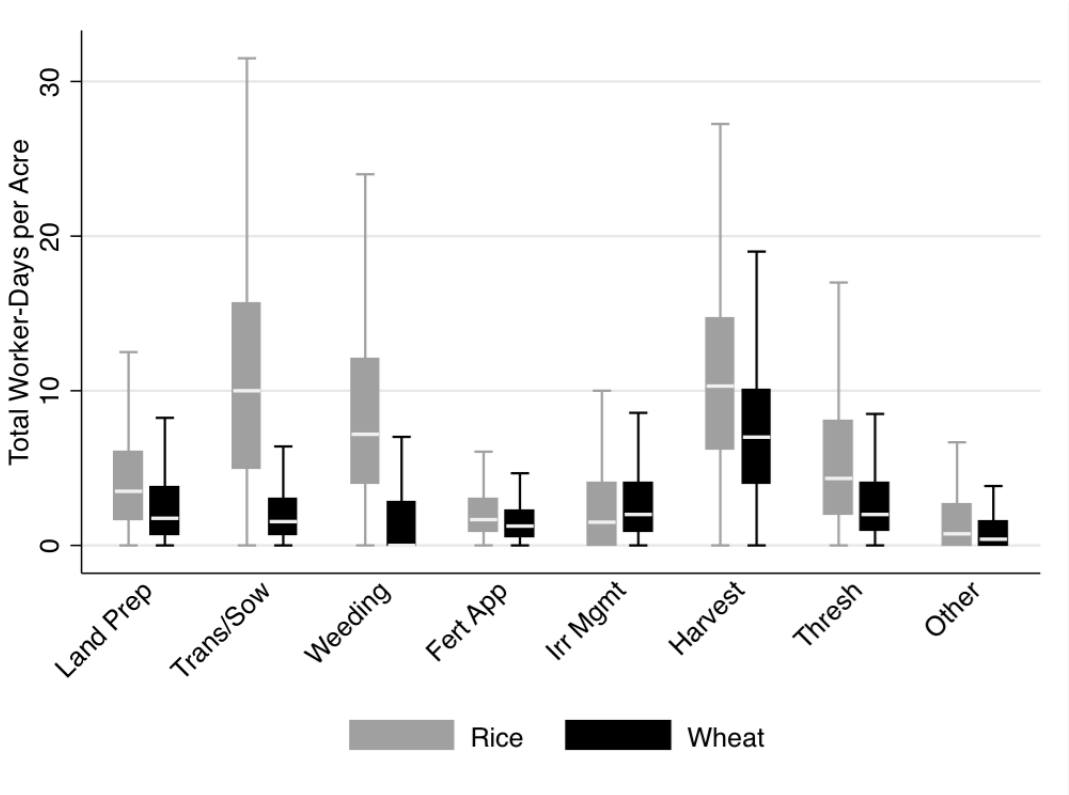
(b) Districts Growing Wheat and/or Rice

(c) Wheat Area Share

(d) Rice Area Share

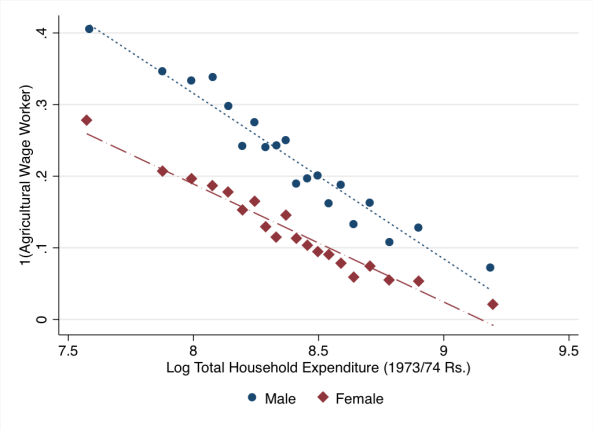
Notes: **Row 1**: Figures compare the 1961-1965 average (a) wheat/rice yield ratio, and (b) wheat and rice yields in tonnes per hectare (*x-axis*) against the same quantities averaged over 1966-1970 (*y-axis*). The black lines in both (a) and (b) denote 45°. **Row 2**: Area shares vary by district-year and are computed as the ratio of area cropped with wheat or rice to the sum area of all crops. Plotted coefficients are year-varying effects of groundwater coverage, measured in percent [0,1], with the marginal effect interpreted in percentage point changes of crop coverage. Models includes year and district fixed effects, and the set of current and lagged temperature and precipitation variables described in Section 4.4. Observations are weighted by 1961 gross cropped area and standard errors are clustered by 1983 National Sample Survey region. Dashed lines represent 95% confidence intervals. Results are for 271 districts from 1956-1987.
Data: World Bank Climate and Agriculture; Geohydrological Map of India 1969

Figure 8: Boxplots of Total Unit Acre Labor Requirements, by Agricultural Operation for Rice and Wheat

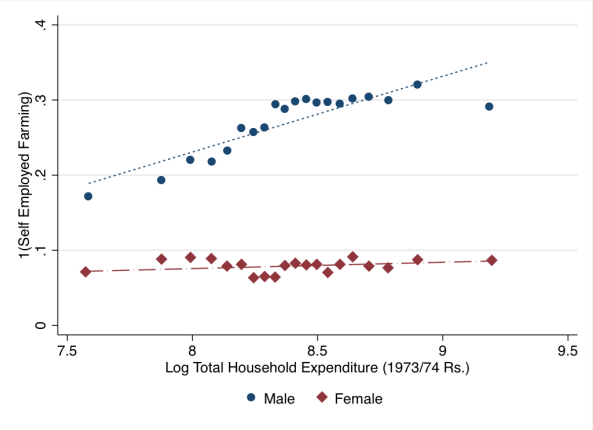


Notes: Each boxplot shows the distribution of total unit acre worker-days for rice and wheat plots, computed as the sum of male and female worker-days across casual, family, and permanent workers divided by plot size.
Data: REDS 1999

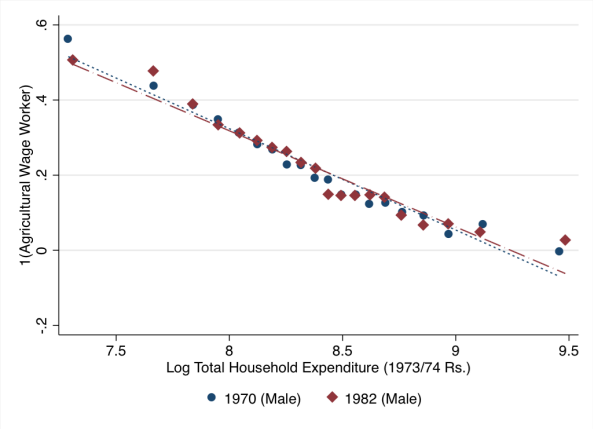
Figure 9: Relationship Between Household Expenditure and Extensive Margin Labor Supply



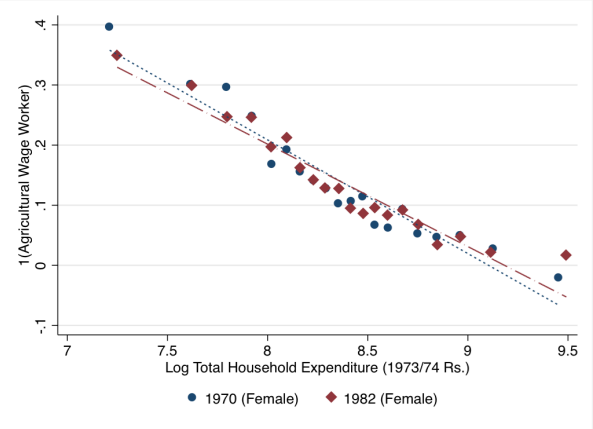
(a) Agricultural Wage Work



(b) Self-Employed Farming



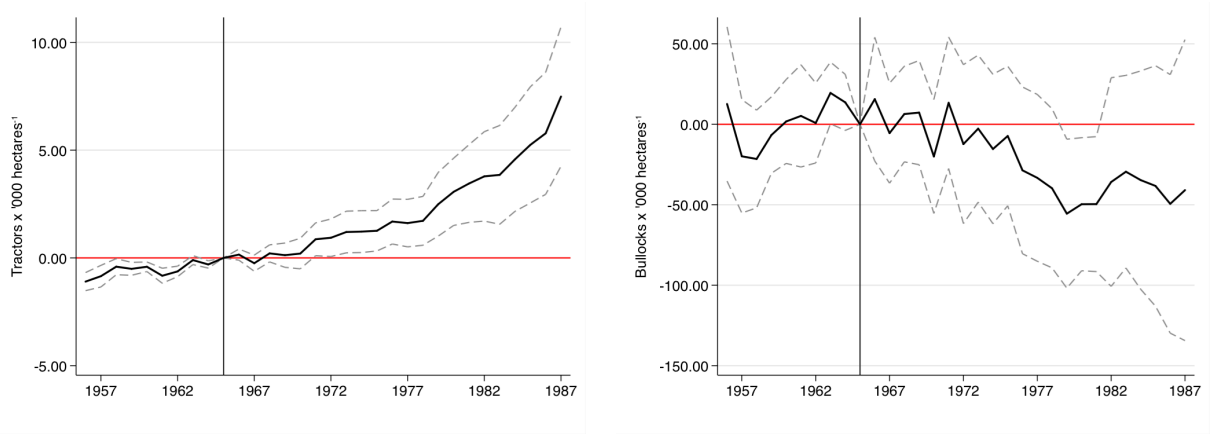
(c) Agricultural Wage Work, Males



(d) Agricultural Wage Work, Females

Notes: **Row 1:** Binned scatterplot results depicting relationship of log total household expenditure with a dummy dependent variable for, (a) agricultural wage laborer, and (b) self-employed farmer, according to respondent's usual activity status. **Row 2:** Similar to Row 1, but with separate response functions by year for (c) men and (d) women. Sample is subset to individuals aged 15-60. Expenditures are deflated to 1973/74 Indian rupees using the state-level consumer price index for agricultural laborers (Besley and Burgess, 2000). All models partial out village × year fixed effects, household size, education, marital status, and total land ownership.
Data: REDS 1970, 1982

Figure 10: Groundwater Coverage and Farm Equipment Capital Stock

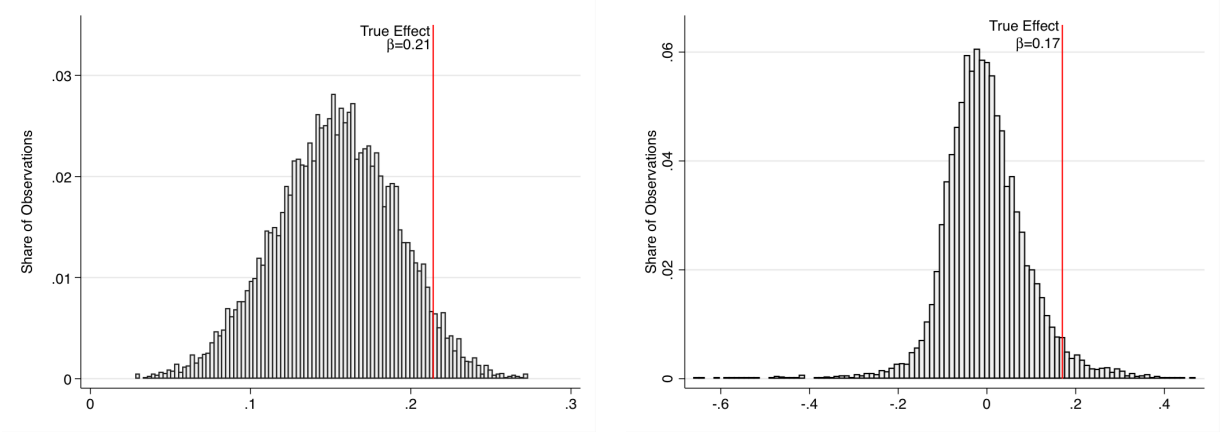


(a) Tractor Density

(b) Bullock Density

Notes: Both outcomes are measured as number of units per thousand cropped hectares. Plotted coefficients are year-varying effects of groundwater coverage, measured in percent [0,1]. Both models include year and district fixed effects, and the set of current and lagged temperature and precipitation variables described in Section 4.4. Observations are weighted by 1961 gross cropped area and standard errors are clustered by 1983 National Sample Survey region. Dashed lines represent 95 percent confidence intervals. Results are for 271 districts from 1956-1987.
Data: World Bank Climate and Agriculture; Geohydrological Map of India 1969

Figure 11: Sampling Distributions From Randomization Inference Tests



(a) y = Log Gross Revenue Ha^{-1} (b) y = Log Wage

Notes: Histograms plot the sampling distribution of the coefficient of interest from 10,000 permutations of within-state groundwater values for (a) '$\mathbb{1}$(t $\geq$ 1966) $\times$ Groundwater %', with log gross unit area revenue the dependent variable and (b) '$\mathbb{1}(t \geq 1966)$ $\times$ Groundwater % $\times$ $\mathbb{1}$(Male)', with log wage the dependent variable. Specification in (a) is comparable to that for Figure 4. Specification in (b) is comparable to that for Figure 6. Figure (b) truncates values below the 0.1 and above the 99.9 percentile levels.
Data: World Bank Climate and Agriculture; Agricultural Wages in India; Geohydrological Map of India 1969

Table 1: Summary Statistics

	Mean	Min	Max	SD	N
Agricultural Wages in India (1958-1983)					
Male Wage (Current Rs.)	4.35	0.18	30	3.18	21104
Female Wage (Current Rs.)	4.05	0.18	25	2.72	29291
1(Sowing)	0.27	0	1	0.44	50395
1(Weeding)	0.047	0	1	0.21	50395
1(Harvesting)	0.27	0	1	0.45	50395
1(Field Labor)	0.26	0	1	0.44	50395
1(Other Labor)	0.15	0	1	0.36	50395
World Bank India Agriculture and Climate (1956-1987)					
Rice Yield (t/ha)	1.03	0	24	0.69	8672
Wheat Yield (t/ha)	0.97	0	18.0	0.73	8672
% Rice Area	0.30	0	1.00	0.31	8672
% Wheat Area	0.14	0	0.73	0.15	8672
% HYV for Eligible Crops	0.23	0	1	0.26	8672
% Gross Area Irrigated	0.24	0	1	0.21	8655
Cropping Intensity	1.21	0.50	4.38	0.19	8654
Groundwater Coverage (%)	0.41	0	1.00	0.42	271
Population Census (1961, 1971, 1981)					
Male, % Agricultural Laborers	0.11	0.0016	0.31	0.062	889
Female, % Agricultural Laborers	0.092	0.00027	0.34	0.079	889
Male, % Cultivators	0.33	0.023	0.56	0.093	889
Female, % Cultivators	0.14	0.0019	0.71	0.13	889
Male, % Part-Time Workers	0.012	0.00014	0.10	0.010	889
Female, % Part-Time Workers	0.088	0.00030	0.42	0.067	889
Rural Economic and Demographic Survey (1970, 1982)					
1(Female)	0.48	0	1	0.50	33596
Age	32.7	15	60	13.3	33596
Household Size	8.39	1	43	4.40	33596
1(Illiterate)	0.53	0	1	0.50	33596
1(Hindu)	0.88	0	1	0.32	33596
1(Ag Laborer)	0.20	0	1	0.40	33596
1(Self-Employed Farmer)	0.18	0	1	0.38	33596
Days Earning Ag Wages	174.8	1	365	80.7	5973
Days Earning Non-Ag Wages	169.0	1	365	98.0	1051
1(Landless)	0.20	0	1	0.40	33596
Land Owned (ha.)	3.37	0	92	5.02	33596
Total Expenditure Per Capita (Rs. 1973/74)	692.6	60.1	8268.8	442.1	33596
1(Groundwater)	0.54	0	1	0.50	33596

Notes: Crop area shares are computed based on gross cropped area. The HYV share measure is the ratio of total area planted with HYVs to area planted with all varieties for the HYV-eligible crops (wheat, rice, maize, sorghum, and pearl-millet).
Data: World Bank Climate and Agriculture; Vanneman and Barnes (2000); REDS; Geohydrological Map of India 1969

Table 2: Effects of Groundwater Coverage on Selected Agricultural Measures

Panel A: Cropping Decisions

	% Total Cropped Area							Cropped Area		Log Fertilizer
	Irrigated	Wheat	Rice	Maize	Sorghum	Pearl Millet	Non-HYV Crops	Log Net	Intensity	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\mathbb{1}(t \geq 1966) \times$ Groundwater %	0.133***	0.105***	0.0206	-0.00930**	0.00645	-0.00991*	-0.112***	-0.0208	0.0665**	0.590**
	(0.0245)	(0.0117)	(0.0130)	(0.00358)	(0.0103)	(0.00582)	(0.0202)	(0.0200)	(0.0256)	(0.271)
Mean DV	0.24	0.13	0.27	0.04	0.13	0.10	0.32	6.28	1.20	6.45
N	8655	8672	8672	8672	8672	8672	8672	8654	8654	8409

Panel B: Cropping Outcomes

	Log Revenue Ha^{-1}	Log Yields					Log Farm Harvest Price				
		Wheat	Rice	Maize	Sorghum	Pearl Millet	Wheat	Rice	Maize	Sorghum	Pearl Millet
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\mathbb{1}(t \geq 1966) \times$ Groundwater %	0.214***	0.198*	0.157*	-0.294***	0.0439	0.112	-0.00798	0.0694	0.0392	0.0308	0.105
	(0.0515)	(0.105)	(0.0796)	(0.102)	(0.0980)	(0.113)	(0.0518)	(0.0651)	(0.0523)	(0.0675)	(0.0670)
Mean DV	3.73	-0.16	-0.06	-0.09	-0.72	-0.85	4.62	4.56	4.28	4.31	4.33
N	8668	7670	8208	7563	7118	6054	8672	8672	8672	7648	7648

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample covers 271 districts from 1956-1987. 'Groundwater %' is the share of district area covering unconsolidated groundwater aquifers and spans [0,1]. Wheat, rice, maize, sorghum, and pearl millet are the five HYV-eligible crops listed in the World Bank dataset. Columns (A2-A6) include cropped area for both HYVs and traditional varieties. Column (A7) is the share of total cropped area planted with crops not possessing an HYV option. Log net cropped area is measured in hectares ('000s) and captures the maximum area under cultivation for any single season. Log fertilizer is measured in tonnes ('000s) of phosphorus. Cropped area intensity (A9) is the ratio of gross to net cropped area and measures the number of cycles a parcel of land is cropped in a calendar year. (B7-B11) are measured in log Indian rupees per quintal. All models include year and district fixed effects and the set of contemporaneous and lagged weather variables. Observations are weighted by 1961 district-level gross cropped area. Standard errors are clustered by 1983 National Sample Survey region.

Data: World Bank Climate and Agriculture; Geohydrological Map of India 1969

Table 3: Gender Decomposition of Daily Wages Using AWI and REDS Data

Panel A: Log Agricultural Wages (AWI 1958-1983)				
	(1)	(2)	(3)	(4)
Male	0.240***	0.260***	0.234***	0.244***
	(0.0245)	(0.0252)	(0.0192)	(0.0264)
Male × 1(Groundwater)				-0.0383
				(0.0474)
Mean DV	1.27	1.27	1.27	1.27
Operation FE		X	X	X
Village FE			X	X
N	37938	37938	37938	37938
Adj. R^2	0.700	0.707	0.867	0.867

Panel B: Log Agricultural Wages (REDS 1970)				
	(1)	(2)	(3)	(4)
Male	0.422***	0.320***	0.322***	0.344***
	(0.0376)	(0.0255)	(0.0263)	(0.0316)
Male × 1(Groundwater)				-0.0564
				(0.0518)
Mean DV	0.74	0.74	0.74	0.74
Village FE		X	X	X
Controls			X	X
N	3082	3082	3082	3082
Adj. R^2	0.167	0.745	0.760	0.760

Panel C: Log Non-Agricultural Wages (REDS 1970)				
	(1)	(2)	(3)	(4)
Male	0.501***	0.382***	0.370***	0.404***
	(0.0906)	(0.0894)	(0.0948)	(0.0956)
Male × 1(Groundwater)				-0.0937
				(0.195)
Mean DV	0.84	0.84	0.84	0.84
Village FE		X	X	X
Controls			X	X
N	507	507	507	507
Adj. R^2	0.188	0.629	0.625	0.625

Notes: Models using AWI data (Panel A) include wages for all operations other than ploughing and 'other field labor,' and include year and month fixed effects. Observations are clustered by National Sample Survey region. Models using REDS 1970 data (Panels B,C) are weighted using sampling probability weights, with observations clustered by village. The included controls for these models are education, marital status, age, and age^2.
Data: Agricultural Wages in India; REDS 1970; Geohydrological Map of India 1969

Table 4: Effect of Groundwater Access on Occupational Affiliation

Dependent Variable: Share of Rural Population by Reported Primary Source of Earnings						
	% Ag Laborers		% Cultivators		% Part-Time	
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
1971 × Groundwater %	0.0075	-0.031***	0.0096*	0.12***	-0.00018	0.030**
	(0.0067)	(0.0070)	(0.0052)	(0.025)	(0.00069)	(0.013)
1981 × Groundwater %	0.0082	-0.050***	0.0032	0.082***	0.00015	0.023*
	(0.0077)	(0.0077)	(0.0056)	(0.025)	(0.00048)	(0.012)
DV Mean	0.12	0.10	0.32	0.11	0.01	0.07
N	880	880	880	880	880	880
Adj. R^2	0.75	0.57	0.83	0.74	0.17	0.57

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Percentage values in (1)-(4) are constructed using the total number of rural workers (main and marginal) by occupation and the total rural population. Values in (5)-(6) are computed using the population statistics of marginal workers. 'Groundwater %' is the share of district area covering unconsolidated groundwater aquifers and spans [0,1]. All estimates are relative to the omitted interaction of 1961 × Groundwater %. All models include district and year fixed effects, and the set of contemporaneous and lagged weather variables. Observations are weighted by 1961 district-level rural population. Standard errors are clustered by 1983 National Sample Survey region.
Data: Vanneman and Barnes (2000); Geohydrological Map of India 1969

Table 5: Tests for Selection on Observables of Agricultural Laborers

Panel A

	Age		1(Illiterate)	
	1(Ag Wage)	Others	1(Ag Wage)	Others
	(1)	(2)	(3)	(4)
1982 × 𝟙(Groundwater) × 𝟙(Male)	-1.890	-3.358*	-0.0476	-0.0182
	(3.796)	(1.797)	(0.108)	(0.0698)
1982 × 𝟙(Groundwater)	-0.287	-0.340	-0.104*	-0.162***
	(3.295)	(1.031)	(0.0595)	(0.0588)
Mean DV	34.88	32.20	0.69	0.55
N	8127	25469	7952	24389
Adj. R^2	0.0906	0.0405	0.267	0.378

Panel B

	1(HH Edu > Primary)		Log Total Exp PC	
	1(Ag Wage)	Others	1(Ag Wage)	Others
	(1)	(2)	(3)	(4)
1982 × 𝟙(Groundwater) × 𝟙(Male)	0.180*	-0.0272	-0.00516	0.0625
	(0.0983)	(0.0425)	(0.0778)	(0.0451)
1982 × 𝟙(Groundwater)	-0.291***	-0.0262	-0.160**	-0.130
	(0.0730)	(0.0655)	(0.0760)	(0.0937)
Mean DV	0.32	0.67	6.10	6.32
N	8061	25284	8127	25469
Adj. R^2	0.274	0.325	0.403	0.418

Notes: Dependent variable in each model is the individual-specific observable (*row 1*). Models are then subset to agricultural laborers ('1(Ag Wage)') or everyone else ('All Others') (*row 2*). Groundwater is a binary measure for whether village is located above an unconsolidated aquifer. Omitted category is 1970 × 𝟙(Groundwater). Respondents are coded as agricultural laborers ('1(Ag Wage)' in column headings) if they primarily identify as an agricultural laborer or reported a positive number of days worked over the research period, while 'All Others' includes the remainder of the population of sampled individuals aged 15-60. Respondents in (B1-B2) are coded as 1 if any household member has an education exceeding primary-level. Expenditure values in columns (B3-B4) are deflated to 1973/74 Indian rupees using the state-level consumer price index for agricultural laborers (Besley and Burgess, 2000). All models include village and year fixed effects, as well as all two-factor interactions from year, groundwater, and sex, that are not displayed. Observations are weighted by household-level sampling probability weights. Standard errors are clustered by village.
Data: REDS 1970, 1982; Geohydrological Map of India 1969

Table 6: Cross-Sectional Evidence Using Women's Time-Use Records

Panel A: Conditional Daily Hours Spent in Work Activities

	Total Work	Wage Labor	Agriculture	Other Work
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Groundwater)	-1.880***	-0.0896	-1.633***	-1.763**
	(0.317)	(0.566)	(0.479)	(0.707)
Mean DV	6.67	7.43	6.26	3.24
N	3780	840	1915	429
Adj. R^2	0.156	0.179	0.184	0.235

Panel B: Unconditional Daily Hours Spent in Work Activities

	Total Work	Wage Labor	Agriculture	Other Work
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Groundwater)	-2.724***	-1.230***	-1.759***	0.0396
	(0.351)	(0.313)	(0.326)	(0.112)
Mean DV	4.49	1.54	1.87	0.26
N	5333	5333	5333	5333
Adj. R^2	0.302	0.202	0.176	0.0314

Panel C: Unconditional Daily Hours Spent in Domestic Activities

	Household Work	Grinding Grain	Collection of		Leisure
			Fuel	Water	
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$(Groundwater)	1.407***	0.754***	0.325**	-0.0764	0.0336
	(0.301)	(0.186)	(0.134)	(0.111)	(0.265)
Mean DV	7.05	0.82	0.39	0.50	5.23
N	5333	5333	5333	5333	5333
Adj. R^2	0.167	0.246	0.0706	0.0570	0.0636

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Dependent variables are activity-specific number of hours spent during a representative day in the primary agricultural season (harvest time for rice-growing locations, sowing-time for wheat). Panel A models restrict the sample to women reporting positive time in the specified category, while Panels B and C include the full sample of women respondents. Total work (A1, B1) is the sum of time engaged in wage labor, agricultural work, cattle care, salaried employment, and 'other work', including in crafts and services, marketing, trading and business, and fishing. Household work consists of cleaning, washing, cooking, and food collection, and is mutually exclusive to other Panel C activities. Groundwater is a binary measure for whether village is located above an unconsolidated aquifer. All models control for age, age^2, caste, religion, years of schooling, number of living children, number of respondents in household, husband's age and age^2, and husband's years of schooling. Sampling weights used in all models. Standard errors are clustered at the level of village.
Data: REDS Demographic Questionnaire 1982; Geohydrological Map of India 1969

Table 7: Effect of Crop Choice on Male Share of Total Worker-Days

Dependent Variable: Male Share of Total Worker-Days

Panel A

	Full Sample (1)	1(Grows Rice & Wheat) (2)	1(Does Not Use HYVs) (3)	1(Not Original GR States) (4)
1(Wheat)	0.0878***	0.0655***	0.0843***	0.0916*
	(0.0258)	(0.0137)	(0.0106)	(0.0486)
DV Mean	0.64	0.74	0.66	0.58
Household FE	X	X	X	X
N	10726	3327	3843	7603
Adj. R^2	0.716	0.698	0.805	0.650

Panel B

	Household Lacks		
	Groundwater (1)	Any Irrigation (2)	Mechanized Machinery (3)
1(Wheat)	0.0550*	0.175*	0.178*
	(0.0295)	(0.0915)	(0.100)
DV Mean	0.54	0.56	0.58
Household FE	X	X	X
N	4798	1940	2497
Adj. R^2	0.657	0.601	0.614

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results from crop-level data for full sample of 4,612 households. Rice is the omitted crop category, and all non-rice, non-wheat crops are estimated as a separate, undisplayed category. Column (A3) restricts sample to households that not using HYVs on any plot. In (A4), households from Punjab, Haryana, and Uttar Pradesh are dropped from the sample. Column (B1) restricts the sample to households in villages not atop an unconsolidated aquifer, while (B2) subsets to households not irrigating any plot. The sample in (B3) consists only of households who neither possessed any mechanized farm equipment assets nor had a positive expenditure on renting mechanized equipment. All models control for plot size and land ownership status. All observations weighted by household sampling probability weight. Standard errors are clustered by village.
Data: REDS 1999; Geohydrological Map of India 1969

Table 8: Effect of Crop Share on Occupational Choice by Gender

Dependent Variable: Share of Rural Workforce by Gender				
Panel A: 1961 Cross-Section Only				
	% Ag Laborers		% Cultivators	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
% Wheat Area	-0.0888	-0.207**	0.0573	-0.300**
	(0.0586)	(0.0863)	(0.0818)	(0.131)
% Rice Area	0.0119	-0.0373	-0.0186	-0.0805
	(0.0278)	(0.0352)	(0.0368)	(0.0638)
DV Mean	0.10	0.09	0.37	0.21
N	269	269	269	269
Adj. R^2	0.0403	0.0933	0.00641	0.0753

Panel B: 1961, 1971, 1981 Panel				
	% Ag Laborers		% Cultivators	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
% Wheat Area	0.0113	-0.211***	-0.00931	0.161
	(0.0507)	(0.0447)	(0.0295)	(0.0964)
% Rice Area	0.0664*	-0.0228	-0.0120	0.0515
	(0.0336)	(0.0518)	(0.0230)	(0.0401)
DV Mean	0.12	0.11	0.34	0.13
N	807	807	807	807
Adj. R^2	0.640	0.456	0.784	0.732

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Omitted category is share of district area under all other crops. Panel models include year and district fixed effects. Employment shares are based on sum of 'main' and 'marginal' workers classified under each occupation, relative to total rural population by gender. All models weighted by 1961 gross cropped area. Standard errors are clustered by 1983 National Sample Survey region.
Data: World Bank Climate and Agriculture; Vanneman and Barnes (2000)

Table 9: Conley Spatial-Adjusted Standard Errors for Log Unit Area Revenue Model

Dependent Variable: Log Unit Area Revenue

	(1)	(2)	(3)
$\mathbb{1}(t \geq 1966) \times$ Groundwater %	0.214***	0.214***	0.214***
	(0.0309)	(0.0515)	(0.0649)
District Clustered	X		
Region Clustered		X	
State Clustered			X
# Clusters	271	53	13
N	8668	8668	8668

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Conley Spatial-Adjusted Standard Errors

Time Cutoff	**Distance Cutoff**					
	100 km	500 km	1000 km	1500 km	2000 km	2500 km
5 Years	(0.010)	(0.025)	(0.032)	(0.036)	(0.037)	(0.038)
10 Years	(0.011)	(0.025)	(0.032)	(0.036)	(0.037)	(0.038)
15 Years	(0.011)	(0.025)	(0.033)	(0.036)	(0.037)	(0.038)
20 Years	(0.010)	(0.025)	(0.032)	(0.036)	(0.037)	(0.038)
25 Years	(0.006)	(0.023)	(0.031)	(0.035)	(0.036)	(0.037)
30 Years	(.)	(0.022)	(0.031)	(0.035)	(0.036)	(0.036)
N	8668	8668	8668	8668	8668	8668

Notes: **Lower Panel**: Standard errors on the treatment interaction term with log unit area revenue as the dependent variable, for each distance cutoff $\times$ time cutoff combination using a Bartlett decay kernel. Distance cutoffs are based on district centroid coordinates using 1961 boundaries. All results are significant at the 1 percent level.
Data: World Bank Climate and Agriculture; Geohydrological Map of India 1969

Table 10: Conley Spatial-Adjusted Standard Errors for Male Wage Premium Model

Dependent Variable: Log Daily Wage (Rs.)			
	(1)	(2)	(3)
$\mathbb{1}(t \geq 1966) \times$ Groundwater % $\times$ $\mathbb{1}$ (Male)	0.167***	0.167***	0.167***
	(0.0365)	(0.0447)	(0.0471)
District Clustered	X		
Region Clustered		X	
State Clustered			X
# Clusters	213	52	13
N	44710	44710	44710

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Conley Spatial-Adjusted Standard Errors

Time Cutoff	**Distance Cutoff**					
	100 km	500 km	1000 km	1500 km	2000 km	2500 km
5 Years	(.)	(0.021)	(0.020)	(0.019)	(0.019)	(0.019)
10 Years	(0.019)	(0.021)	(0.021)	(0.020)	(0.020)	(0.019)
15 Years	(0.016)	(0.019)	(0.019)	(0.018)	(0.017)	(0.017)
20 Years	(0.011)	(0.015)	(0.014)	(0.013)	(0.012)	(0.012)
25 Years	(0.011)	(0.015)	(0.014)	(0.012)	(0.012)	(0.012)
30 Years	(0.013)	(0.016)	(0.016)	(0.014)	(0.014)	(0.014)
N	44710	44710	44710	44710	44710	44710

Notes: **Lower Panel**: Standard errors on the treatment interaction term with log daily wage as the dependent variable, for each distance cutoff $\times$ time cutoff combination using a Bartlett decay kernel. Distance cutoffs are based on district centroid coordinates using 1961 boundaries. All results are significant at the 1 percent level.
Data: Agricultural Wages in India; Geohydrological Map of India 1969

Table 11: Robustness Tests for Crop Productivity Models

Dependent Variable: Log Wheat Yield									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbb{1}(t \geq 1966) \times$ Groundwater %	0.198*	0.116**	0.194*	0.116	0.252**	0.185*	0.186*	0.236*	0.0834*
	(0.105)	(0.0472)	(0.105)	(0.0785)	(0.0992)	(0.107)	(0.109)	(0.129)	(0.0470)
Mean DV	-0.16	-0.16	-0.16	-0.16	-0.13	-0.19	-0.22	-0.30	-0.30
State × Year		X							X
20km River Buffer			X						
Soil/Slope Dummies				X					
AWI Districts Only					X				
Minus Punjab						X	X	X	X
Minus Haryana							X	X	X
Minus Uttar Pradesh								X	X
N	7670	7670	7670	7670	6103	7350	7162	5626	5626

Dependent Variable: Log Rice Yield									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbb{1}(t \geq 1966) \times$ Groundwater %	0.157*	0.246***	0.159*	0.178***	0.218**	0.0804	0.0600	0.0500	0.225**
	(0.0796)	(0.0803)	(0.0802)	(0.0630)	(0.104)	(0.0608)	(0.0624)	(0.0980)	(0.0898)
Mean DV	-0.06	-0.06	-0.06	-0.06	-0.03	-0.08	-0.10	-0.08	-0.08
State × Year		X							X
20km River Buffer			X						
Soil/Slope Dummies				X					
AWI Districts Only					X				
Minus Punjab						X	X	X	X
Minus Haryana							X	X	X
Minus Uttar Pradesh								X	X
N	8208	8208	8208	8208	6576	7888	7724	6188	6188

Dependent Variable: Log Gross Unit Hectare Revenue									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\mathbb{1}(t \geq 1966) \times$ Groundwater %	0.214***	0.231***	0.213***	0.192***	0.252***	0.174***	0.152***	0.124**	0.219***
	(0.0515)	(0.0577)	(0.0504)	(0.0543)	(0.0584)	(0.0450)	(0.0457)	(0.0555)	(0.0636)
Mean DV	3.73	3.73	3.73	3.73	3.72	3.71	3.70	3.65	3.65
State × Year		X							X
20km River Buffer			X						
Soil/Slope Dummies				X					
AWI Districts Only					X				
Minus Punjab						X	X	X	X
Minus Haryana							X	X	X
Minus Uttar Pradesh								X	X
N	8668	8668	8668	8668	7008	8348	8156	6620	6620

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Gross unit hectare revenue is a crop-weighted measure of farm productivity described in Section 4. Models in (3) include the share of district area within 20 kilometers of a named river (described in Section 6.6, interacted with a full set of year dummies. Models in (4) add 26 soil and slope dummies, each interacted with a full set of year dummies. Models in column (5) subsample to only districts from which AWI wage data was ever reported. All models include year and district fixed effects with observations weighted by 1961 gross cropped area. All models control for the set of current and lagged temperature and precipitation variables described in Section 4.4. Standard errors are clustered by 1983 National Sample Survey region. Results are for 271 districts from 1956-1987.

Data: World Bank Climate and Agriculture; Geohydrological Map of India 1969

Table 12: Robustness Tests for Male Wage Premium Models

Dependent Variable: Log Nominal Wage

Panel A

	Full Sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}(t \geq 1966) \times$ Groundwater % $\times$ $\mathbb{1}$ (Male)	0.169***	0.111**	0.162***	0.163***	0.132***	0.144***	0.126**
	(0.0431)	(0.0514)	(0.0440)	(0.0480)	(0.0477)	(0.0482)	(0.0488)
Mean DV	1.19	1.19	1.15	1.17	1.15	1.14	1.14
State $\times$ Year		X					X
Time-Invariant Year-Interacted			X				
Minus Punjab				X	X	X	X
Minus Haryana					X	X	X
Minus Uttar Pradesh						X	X
N	48759	48759	44710	47525	46002	45074	45074

Dependent Variable: Log Nominal Wage

Panel B

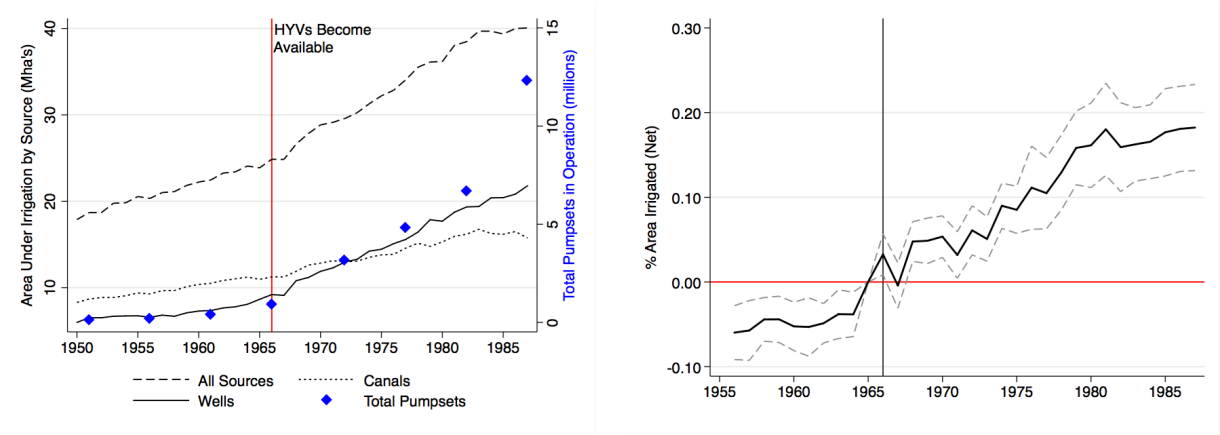
	1(Reports Before 1960)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}(t \geq 1966) \times$ Groundwater % $\times$ $\mathbb{1}$ (Male)	0.225***	0.129**	0.210***	0.231***	0.181***	0.181***	0.146**
	(0.0481)	(0.0636)	(0.0527)	(0.0540)	(0.0494)	(0.0494)	(0.0556)
Mean DV	1.12	1.12	1.09	1.08	1.05	1.05	1.05
State $\times$ Year		X					X
Time-Invariant Year-Interacted			X				
Minus Punjab				X	X	X	X
Minus Haryana					X	X	X
Minus Uttar Pradesh						X	X
N	28543	28543	26688	27465	26068	26068	26068

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: **Panel A**: Robustness checks for estimating the male wage premium in groundwater-rich districts, using the specification in Equation 7 . Column (1) is the benchmark result. In column (3), the 1961 value is interacted with a full set of year dummies for the following variables: rural female literacy rate, female labor force participation rate, male labor force participation rate, and log gross unit area revenue. **Panel B:** Models are comparable to those in Panel A, but subset to reporting centers that ever provided AWI wage values before 1960. All models include district, gender, month, operation, and year fixed effects in addition to those described. All models control for the set of current and lagged temperature and precipitation variables described in Section 4.4. Standard errors are clustered by 1983 National Sample Survey region.

Data: Agricultural Wages in India; Geohydrological Map of India 1969; Vanneman and Barnes (2000)

Figure A1: Effect of Green Revolution on Irrigation Coverage and Pumpset Adoption



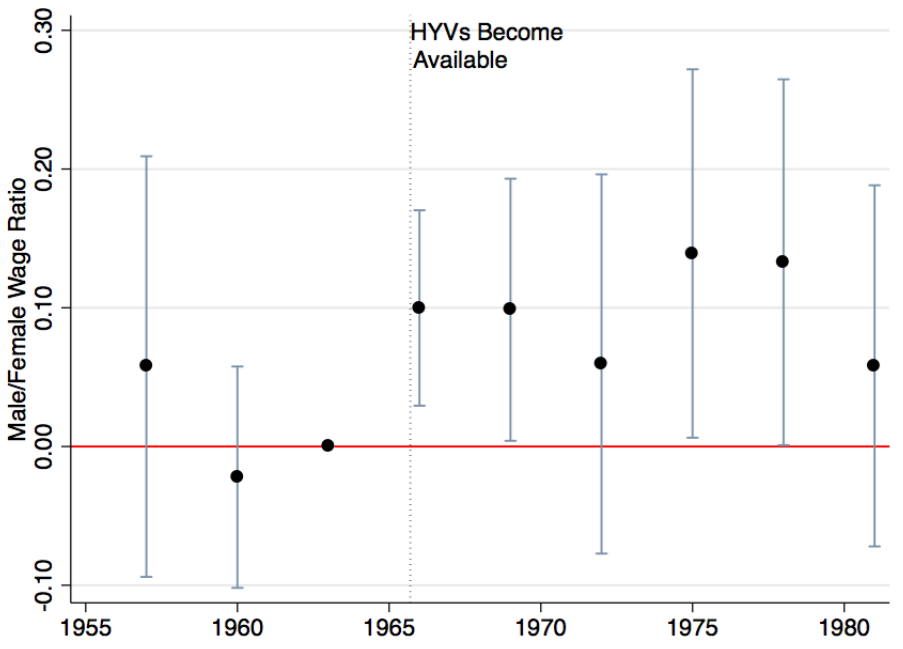
(a) Irrigation Coverage and Pumpset Census

(b) Irrigated Share of Net Cropped Area

Notes: In (a), 'all sources' includes canals, tanks, tubewells, other wells, and 'other sources.' Pumpset counts are total of diesel and electric pumpsets. In (b), irrigated share is the ratio of net irrigated area to net cropped area. Plotted coefficients are year-varying effects of groundwater coverage, measured in percent [0,1], with the marginal effect interpreted in percentage point changes of irrigation share. Model includes year, district, and state × year fixed effects, and the set of current and lagged temperature and precipitation variables described in Section 4.4. Observations are weighted by 1961 gross cropped area and standard errors are clustered by 1983 National Sample Survey region. Dashed lines represent 95 percent confidence intervals. Results are for 271 districts from 1956-1987.

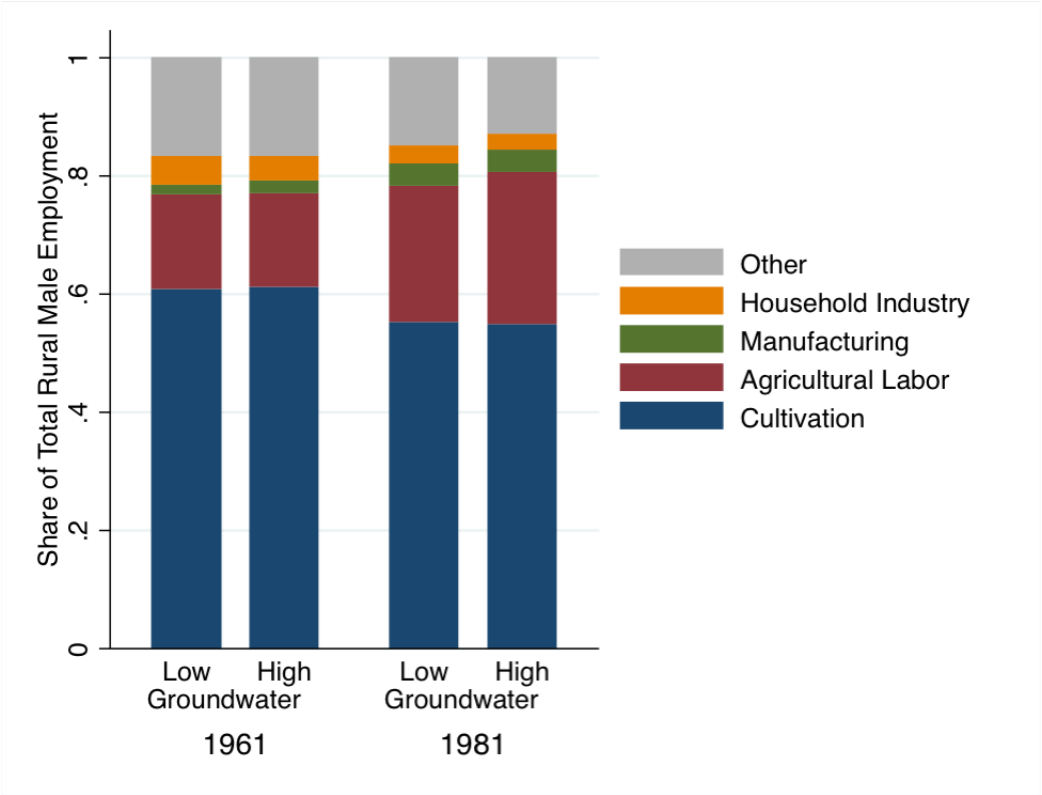
Data: (a) Ministry of Agriculture; Central Water Commission; Water Resources Information System Directorate; Ministry of Statistics and Programme Implementation, (b) World Bank Climate and Agriculture; Geohydrological Map of 1969

Figure A2: Effect of Groundwater on Male/Female Wage Ratio



Notes: Results from model regressing male/female laborer wage ratio on groundwater coverage interacted with dummies for pooled three-year intervals. Model includes year, district, month of year, and agricultural operation fixed effects. Confidence intervals shown are at the 95 percent level. Standard errors are clustered by 1983 National Sample Survey region.
Data: Agricultural Wages in India; Geohydrological Map of India 1969; Vanneman and Barnes (2000)

Figure A3: Male Employment Shares by Industry



Notes: Stacked bar plots represent the share of rural male employment in each industry for 1961 and 1981, using a cut-off of 20 percent groundwater coverage for categorization as 'low' or 'high.' Industries included in 'Other' are livestock-rearing, mining, construction, communications, transportation, and services.
Data: Vanneman and Barnes (2000); Geohydrological Map of India 1969

Table A1: Effect of Groundwater Coverage on Individual-Specific Labor Outcomes

Panel A: Extensive Margin Outcomes

	1(Ag Laborer)				1(Farmer)		1(Non-Earner)	
	Male		Female		Male	Female	Male	Female
	All	HoH AgLaborer	All	Male AgLaborer				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1982 × 𝟙(Groundwater)	-0.0124	0.256**	-0.0309	-0.223***	0.0655	0.0585*	-0.0503	0.0350
	(0.0480)	(0.117)	(0.0499)	(0.0813)	(0.0555)	(0.0347)	(0.0391)	(0.0534)
Mean DV	0.42	0.48	0.23	0.36	0.27	0.06	0.24	0.73
N	17526	4132	16066	7479	17526	16066	17526	16066
Adj. R^2	0.314	0.317	0.330	0.397	0.274	0.202	0.313	0.284

Panel B: Days Worked as Wage-Earner by Sector

	Agricultural				Non-Agricultural			
	Days		Days | Days > 0		Days		Days | Days > 0	
	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1982 × 𝟙(Groundwater)	-11.54	-6.460	-16.76	-2.942	3.009	-0.257	17.07	-76.35
	(10.76)	(8.775)	(19.24)	(21.96)	(4.111)	(2.406)	(32.65)	(55.19)
Mean DV	59.57	35.53	185.84	161.28	9.57	2.40	162.81	131.74
N	17526	16066	3772	2201	17526	16066	793	258
Adj. R^2	0.218	0.311	0.315	0.329	0.104	0.0629	0.605	0.692

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Panel A presents results from linear probability models regressing binary outcome variables on village-level groundwater coverage, while models in Panel B regress the number of days worked in wage-earning, non-salaried activities on groundwater. Models (B3-B4) and (B7-B8) are conditioned on observed positive labor supply in the stated agricultural or non-agricultural sector category. The sample for all models is restricted to individuals aged 15-60. All models include village fixed effects and control for respondent's age, age^2, marital status, household size, and religion. Observations are weighted by household-level sampling probability weights. Standard errors are clustered by village.
Data: REDS 1970, 1982; Geohydrological Map of India 1969

Table A2: Effect of Groundwater Access on Rural Migration and Population Growth

	Log Population				% Urban		Log Rural In-Migrants		% Rural In-Migrant	
	Rural		Urban							
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1971 × Groundwater %	0.00369	-0.0110	-0.0807***	-0.0681**	-0.0114***	-0.00824***	-0.252***	-0.0698*	-0.0269***	-0.0230**
	(0.0126)	(0.0174)	(0.0270)	(0.0286)	(0.00276)	(0.00263)	(0.0535)	(0.0397)	(0.00618)	(0.0100)
1981 × Groundwater %	0.0282	0.0166	0.00501	0.0277	-0.00978	-0.00578	-0.421***	-0.0570	-0.0379***	-0.0330**
	(0.0201)	(0.0248)	(0.0567)	(0.0604)	(0.00604)	(0.00654)	(0.0883)	(0.0439)	(0.00794)	(0.0130)
DV Mean	13.47	13.42	11.80	11.67	0.18	0.17	11.38	12.64	0.14	0.47
N	807	807	804	804	807	807	807	807	807	807
Adj. R^2	0.951	0.939	0.890	0.891	0.645	0.651	0.332	0.805	0.285	0.190

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: In-migrants from columns (7)-(10) include all individuals who were not born in the village or town in which they responded to the census. Percent values in columns (9)-(10) are constructed as a share of total male and female rural district population, with no age adjustments. 'Groundwater %' is the share of district area covering unconsolidated groundwater aquifers and spans [0,1]. All estimates are relative to the omitted interaction of 1961 × groundwater. All models include district and year fixed effects. Observations are weighted by 1961 gross cropped area. Standard errors are clustered by 1983 National Sample Survey region.

Data: Vanneman and Barnes (2000); Geohydrological Map of India 1969

Table A3: Effect of Wheat Cropping on Male Share of Total Worker-Days

Dependent Variable: Male Share of Total Worker-Days							
	Full Sample	1(Grows Rice & Wheat)	1(Does Not Use HYVs)	1(Not Original GR States)	Household Lacks		
					Groundwater	Any Irrigation	Mechanized Machinery
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1(Wheat)	0.0424***	0.0475***	0.0573***	0.0340**	0.0347*	0.0228	0.0469
	(0.00934)	(0.00987)	(0.01000)	(0.0154)	(0.0202)	(0.0168)	(0.0457)
DV Mean	0.64	0.74	0.66	0.58	0.54	0.56	0.58
Household FE	X	X	X	X	X	X	X
N	10724	3327	3843	7600	4795	1938	2497
Adj. R^2	0.699	0.677	0.790	0.633	0.655	0.431	0.532

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Results from crop-level data for full sample of 4,612 households. All crops other than wheat are pooled into a single, omitted category. Column (A3) restricts sample to households that not using HYVs on any plot. (A4) drops households from Punjab, Haryana, and Uttar Pradesh from the sample. (B1) restricts the sample to households in villages not atop an unconsolidated aquifer, while (B2) subsets to households not irrigating any plot. (B3) consists only of households who neither possessed any mechanized farm equipment assets nor had a positive expenditure on renting mechanized equipment. All models control for plot size and land ownership status. All observations weighted by household sampling probability weight. Standard errors are clustered by village.
Data: REDS 1999; Geohydrological Map of India 1969

Table A4: Effect of Spatially Averaging Wage Data on Estimates of Male Wage Premium Response

Panel A

	Full Sample			
			d,s,t Averaging	v,s,t Averaging
	(1)	(2)	(3)	(4)
$\mathbb{1}(t \geq 1966) \times$ Groundwater % $\times$ $\mathbb{1}$ (Male)	0.169***	0.165***	0.131***	0.132***
	(0.0430)	(0.0421)	(0.0361)	(0.0327)
Mean DV	1.19	1.19	0.95	0.95
District FE	X		X	
Village FE		X		X
Operation FE	X	X		
Month FE	X	X		
N	48759	48759	7475	8081

Panel B

	1(Reports Before 1960)			
			d,s,t Averaging	v,s,t Averaging
	(1)	(2)	(3)	(4)
$\mathbb{1}(t \geq 1966) \times$ Groundwater % $\times$ $\mathbb{1}$ (Male)	0.225***	0.216***	0.185***	0.193***
	(0.0480)	(0.0489)	(0.0364)	(0.0366)
Mean DV	1.12	1.12	0.86	0.87
District FE	X		X	
Village FE		X		X
Operation FE	X	X		
Month FE	X	X		
N	28543	28543	4500	4763

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: **Panel A**: Columns (1-2) are benchmark results for estimating the male wage premium in groundwater-rich districts and respectively include district and village fixed effects. In (3), all wage observations for a district $\times$ gender $\times$ year are averaged to a single value. (4) performs a similar averaging procedure by village. **Panel B:** Models are comparable to those in Panel A, but subset to locations that ever provided AWI wage values before 1960. All models include gender and year fixed effects in addition to those labeled. All models control for the set of current and lagged temperature and precipitation variables described in Section 4.4. Standard errors are clustered by 1983 National Sample Survey region.

Data: Agricultural Wages in India; Geohydrological Map of India 1969; Vanneman and Barnes (2000)