UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ESSAYS IN DEVELOPMENT ECONOMICS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

SAINI DAS

Norman, Oklahoma

2021

ESSAYS IN DEVELOPMENT ECONOMICS

A DISSERTATION APPROVED FOR THE

DEPARTMENT OF ECONOMICS

BY THE COMMITTEE CONSISTING OF

Dr. Joan Hamory, Chair

Dr. Daniel Hicks, Co-Chair

Dr. Kevin Kuruc

Dr. Pallab Ghosh

Dr. Kirsten de Beurs

Acknowledgments

It did not start out easy. Nevertheless, the writing of this dissertation was made a fulfilling experience by a number of individuals. To them, I owe my deep gratitude. I thank my advisors Joan Hamory and Daniel Hicks for their passionate guidance along the way. I thank Joan for showing me the value of rigorous and meticulous research, for helpful feedback in my later projects, and giving me the opportunity to work alongside her in collaborative projects. To Dan, I owe my first foray into the world of development, for patiently molding (and sometimes steering me away from!) half-formed ideas and being an enthusiastic supporter of my career in economics. I also thank my professors in the University of Calcutta who stimulated in me the idea of graduate school.

I have been fortunate to be surrounded by some amazing women trailblazers- my advisor Joan, my sister Roshni, my aunts Rita and Rina. Thank you for pushing the envelope, so I can live by example. I especially want to thank my sister, for being by my side all the way to graduate school. I would not be here today without her showing the beacon. I also want to thank my entire extended family for coming together to support each other during this very tumultuous past year. Their perseverance was my strength.

Each of my committee members played a significant role in the development of my chapters through conversations and feedback. I thank Kevin Kuruc, for many a lively discussion that sparked new ideas. Pallab Ghosh, for feedback that enhanced my projects and Kirsten de Beurs for showing new avenues for research. I especially want to thank my fiancé Robert Press for being my sounding board during the endless hours spent in coding. But mostly, I

thank him for “toning down my crazy”.

Finally, I want to thank my parents Pranab and Chhabi Das, for their tireless support, unwavering trust, and an undaunted motivation to inspire and support me at every turn. I would not be where I am or who I am without their constant enthusiasm for the last couple decades.

Contents

List of Figures

List of Tables

Abstract

My dissertation chapters study the impact of environmental factors on outcomes of human functioning. The first chapter studies the impact of heat stress during a primary school-leaving exam in Indonesia on outcomes in later life. The Ebtanas are a national standardized test that students must take to gain entry to secondary school. Using individual-level data on test scores, I first show that cognitive performance during the test is affected by heat stress. Impacts are heterogeneous across different subjects, with math and science being the most heavily impacted. Next, I show that disruptive weather conditions during the Ebtanas have compounding negative effects on a wide range of long-term achievements such as adult educational attainment, labor market participation and entry to the marriage market. A 1℃ increase in temperature in the month of exam leads to 1.53% fewer years of education, 2% fewer hours worked and a 2% higher probability of being married by 18 for women. These findings stress that even examinations conducted during early or mid adolescence may have impacts that persist through adulthood.

The second chapter studies the repercussions of an agricultural productivity shock for labor market outcomes and inequality in India. I show that the increase in productivity had heterogeneous impacts on technological diffusion and local labor market outcomes. In wheat growing areas, the productivity increase was followed by investments in labor-saving technology, demonstrated by an increase in the use of tractors. Rice areas in contrast, invested more heavily in labor-enhancing technology such as fertilizers, creating new opportunities

for application of labor. These shifts exacerbated inequality in wheat districts while reducing inequality in rice districts. I show that these results are robust to fixed effects and instrumental variables strategy. These findings demonstrate that driven by differences in environmental factor endowments, a productivity shock can have heterogeneous impacts on agricultural labor markets and inequality.

Chapter 3 examines the efficacy of genealogically constructed networks in sharing risk under aggregate versus idiosyncratic income shocks in the context of geographically split-off families in Indonesia. While informal transfers are effective in sharing aggregate risks, they are ineffective in an idiosyncratic shock. Plausible reasons include a higher probability of repayment as well as greater economies of scale from resource pooling. We show that idiosyncratic shocks induce households to make long-term and costly changes to their household structure. We demonstrate this in the formation of new split-off families over time who reside outside the district. Our findings further reveal that controlling for shocks to members in a family network is an important source of omitted variable bias in empirical estimations of the impact of shocks on informal transfers.

Chapter 1

The Heat is On: The Long-term impact of Heat Stress during Primary School Leaving Examinations in Indonesia

1.1 Introduction

The relationship between environmental conditions and human functioning has received considerable scrutiny in academic circles recently (Dell *et al.* (2014); Deschenes (2014)). A branch of this literature has specifically focused on the effect of temperature on individual productivity and human capital production. Much of this effect is derived from contemporaneous weather variations and are thus temporary in nature (Graff Zivin *et al.* (2018); Garg *et al.* (2018); LoPalo (2019)). Less research has focused on long-term persistence of these adverse outcomes well into adulthood. Identifying to what extent random weather realizations induce long term changes to an individual's human capital accumulation is crucial from a public policy perspective directed towards adaptive behavior.

The idea that an environmental condition such as heat stress can lead to temporary disruptions in cognitive functioning is not new. In a laboratory setting, even minor fluctuations in temperature in the range of 2-5 degrees Celsius over normal core body temperature has been shown to induce non-trivial changes in physiology such as mental confusion, memory loss and loss in reaction time (Hancock (1986); Enander and Hygge (1990)). In economics, the few studies that document the direct cost of cognitive disruption in high stakes examinations primarily focus at the level of college admissions (Ebenstein *et al.* (2016); Park (2017); Graff Zivin *et al.* (2018)). The long term toll of such disturbances in high-stakes exams conducted earlier in life is unknown.

This is vital given that typically, less developed countries have a very low rate of graduation from high school. According to a recent report for Indonesia, only 34.6% of the population have completed education at the higher secondary level (World Bank (2018)). Of this, an even smaller proportion would have actually taken the examination for entry to college. As such, for over 65% of the population, decisions about adult human capital formation is path-dependent on circumstances (including examinations) occurring much earlier in life. Ideally, in this scenario, evaluating the ramifications of heat stress during a standardized national test occurring earlier in adolescence would yield more broadly generalizable estimates.

This paper addresses this need by studying the impact of heat stress during a primary school-leaving exam in Indonesia on outcomes in later life. The Ebtanas are a national standardized test in Indonesia that students must take to gain entry to secondary school. Taken in early to middle childhood, a typical student is between 12 or 13 years of age. Each student is tested on a range of subjects such as mathematics, science, social science, language and religion, each of which require varying levels of cognitive capability. Using rich data on self reported test scores from the Indonesian Family Life Survey (IFLS), I evaluate the sensitivity to different levels of temperature and humidity for each of these subjects.

I show that, similar to previous studies, and consistent across different specifications, math and science test scores are most impacted by high temperatures. Furthermore, these coefficients are larger when humidity, as quantified in heat stress, is taken into account. For instance, a 1℃ increase in temperature in the month of exam leads to a 0.6% decline in math scores, 0.8% decline in science scores, and a 0.5% decline in language scores. I find heterogeneous impacts across men and women, with men being more susceptible to heat stress. I also find that impacts across high and low performing students differ- with high and low performing students being relatively less impacted than average performing students. It is conceivable that students performing at the threshold score of entry to secondary school are the ones who are the most vulnerable to random variations in the testing environment.

One of the advantages of using a panel survey such as IFLS is the ability to follow a subsample of respondents into adulthood and observe realizations across a range of outcomes. To the extent that cognitive performance in the Ebtanas and long-term achievements are related, environmental stressors that cause random disruptions to an individual's cognitive functioning may have a long-lasting impact. I exploit variations in the temperature during the month that the student takes the exam to explain long run achievement using both reduced form and two stage least squares estimation strategies.

To causally identify this effect, temperature treatments across students must be uncorrelated with other factors. To this end, a rich set of controls for observable student characteristics and unobservable variables such as seasonality fixed effects, location fixed effects and rigor of the test through cohort fixed effects are included. Moreover, given the low propensity of students to take the test in a school that is in a different sub district than their residence, the concern of endogenous sorting of high-performing students across temperature assignments is alleviated.

In general, I find that while performance in the Ebtanas does not translate directly into lower labor market returns, it is a significant predictor of educational achievement and labor force participation. For instance, in results using a reduced form model, each 1℃

increase in temperature is associated with 0.092 fewer years of education which is a 1.12% decline in an average individual's total educational achievement. Alongside this, the labor market outcomes for men indicate a lowered probability to be in the labor force or a higher probability to engage in household work for women and a 2.62% decline in the mean number of hours worked for every ℃ rise in temperature during their exam month.

In exploring the heterogeneity of these impacts across men and women, it is interesting to note that while men incur greater losses in the short term as a result of heat stress, women report larger declines in human capital over the long run. In particular, I detect important marriage market impacts for women. High temperatures in the month of exam raise the probability of women being married by 18, an increase in dowry given and an increase in the spousal age gap. Given that the average age of marriage for women in the sample is 13.13 years, it is not unlikely that a poor performance in the Ebtanas may abruptly increase the incidence of an early marriage as well as qualitatively poor marriages.

The question of returns to education has received considerable focus in the literature due to the omitted variable bias of ability as well as other factors that affect both the education decision and earnings (Angrist and Keueger (1991);Duflo (2001)). In a 2sls setting, the random variations in temperature in the month of exam may be used to instrument for academic achievement. Instrumenting for math scores, for which I have the largest sample size, I interpret the long-term returns of a 1-point increase in score. I find that a 1 point increase in math score leads to individuals working 2.33% more hours in the labor market. However the impact on wages is insignificant plausibly due to an insufficiently powered instrument.

In exploiting the variations in temperature during the month of the exam, it may be argued that after controlling for the contemporaneous effect of cognitive impairment, alternative mechanisms such as a bad agricultural season may influence test performance through lowered income and health. However, this is implausible as over 80% of students take the

Ebtanas during the months of May and June, while the agricultural season runs from November to February. Also, to the extent that chronic heat stress over the years may adversely affect students over time, this setting may be used to uncover long term learning impacts of heat exposure during preceding school year. I test the sensitivity of the scores to temperatures in preceding seasons, months and years to rule out these channels as a significant predictor of test scores.

This paper contributes to several literatures. It adds to the growing literature on the impact of heat stress in less than ideal testing conditions of less developed countries. While the impact of temperature in developed countries has been studied before, the validity of these estimates, in the context of less developed countries, is questionable for a variety of reasons. First, geographically, almost all low income countries straddle the equator. Being close to tropics, typically, these regions exhibit a much lower range of temperature variations. While some developed countries may face temperature fluctuations of up to 10C in a single day, low-income countries experience this range over the course of a year. Second, it may be expected that adaptive behavior of people residing here may be attuned to prior knowledge about infrastructure (such as the lack of air conditioners) and expectation about weather conditions. For instance, students may take extra measures to wear breathable clothing and stay hydrated, given their prior knowledge of weather in summer. As such the temperature-cognition sensitivity of respondents in this region is different from the rest of the world.

It also contributes to the literature that focuses on identifying the economic returns to academic achievement amongst early adolescents (Ozier (2015);Duflo *et al.* (2010)). Earlier work on the impact of environmental stressors during high-stakes examinations conducted for college admissions naturally exclude a big portion of the population who drop out prior to reaching the college entrance exam. The findings in this study indicate that while, scoring an additional point in the exams may have modest impact on the returns in the labor market, the impact on educational attainment and subsequent quality of life is substantial. Given that the cohort considered here is in their twenties, combined human accumulation over the lifetime may further strengthen these effects. This serves to highlight the role of early adolescent intervention as potentially having long-lasting impacts.

Lastly, this study explores the temperature-cognition and the temperature-later life relationship across a wider range of outcomes than previously studied by using a household panel survey where a sub-sample of respondents are followed to maturity. In doing so, it identifies several horizons of human functioning which may potentially be affected by rising temperatures. First, I identify which subjects students are more likely to underperform in, due to heat stress. Next, I identify the cost of these random environmental disruptions over the long term beyond the direct labor market impact. The rest of the paper proceeds as follows- Section 2 provides background on the Indonesian context, Section 3 discusses the data. Sections 4 and 5 present the empirical approach and results for the short run and long run analyses respectively. Section 6 presents results on chronic heat stress and Section 7 concludes.

1.2 Background

1.2.1 Education in Indonesia

Education in Indonesia is subdivided into elementary school (Sekolah Dasar) for grades 1-6, junior high school (Sekolah Menengah Pertama, or SMP) for grades 7-9 and high school for grades 10-12. At the end of each level, students must appear in a National exam and obtain a passing grade to graduate onto the next level. While enrollment in elementary school is around 95%, only 78% of those go on to enroll in middle school of whom 58% graduate on to high school. [1]

Starting in 1980, the Indonesian system of educational testing was centrally standardized

[1] ASEAN State of Education Report 2013

as the Ebtanas (for a few core subjects) and Ebta (for non-Ebtanas subjects), in order to compare the quality of elementary, middle and high school graduates across schools in different provinces. [2] The test was typically administered at the end of the school year. Students were tested on a number of core/Ebtanas subjects such as Mathematics, Indonesian language and Moral code. The curriculum and minimum standards for graduation would be set centrally by the Directorate of Primary and Secondary Education.

The Ebtanas were replaced by the National Exam (Ujian Nasional or UNAS) in 2005 continuing to the present, as standards determining graduation strengthened across the country. It was organized by the Center for Education Assessment (Puspendik) and the National Education Standards Agency (BSNP). Despite these changes in administrative control, the purpose of the UNAS and Ebtanas remained the same during this period. However, following the change to UNAS, rich official data on exact dates for testing as well as cut-off scores became available from the Ministry of Education and Culture. Hence, for a subset of the study period (2005-15) I am able to explain scores as a function of daily variations in climate.

Data from the IFLS records for each respondent, the exact month and year of their Ebtanas/UNAS exam at the end of their elementary education (Sekolah Dasar). Since elementary education in Indonesia lasts around 6 years starting at 6 years old, test-takers are typically 12 or 13 years old. Students may graduate and/or enroll in junior high school (Sekolah Menengah Pertama, or SMP) provided they attain the minimum passing grade in the UNAS/Ebtanas.

[2] This discussion follows from the information in the webpage of the Ministry of Education and Culture in Indonesia which along with the Ministry of Religious Affairs determines national standards for education. https://puspendik.kemdikbud.go.id/ujian-nasional-un

1.2.2 Temperature, Humidity and Cognitive Performance

The link between heat stress and lowered labor productivity has been widely studied.[3] Heat stress is generated by a buildup of body heat due to either internal muscle use, such as exercise or environmental factors such as hot days. An increase of body temperature by 1-2 degree Centigrade manifests into mild to severe physiological distress such as mental confusion, memory loss and even heat stroke. Specifically, in relation to cognition, any increase in core body temperature is disruptive to attention and memory (Hancock (1986); Enander and Hygge (1990)).[4] Furthermore, the responsiveness to different tasks such as language versus oral or creative is different at different levels of heat stress (Ramsey (1995)).[5]

In addition to higher temperature, high humidity impedes the ability of the body to lower its core temperature via sweating. As such, the inability of the human physiology to respond efficiently on a humid day will cause greater stress and discomfort. In this study, I use distinguish between hot days versus hot and humid days using the well-established measure of dry bulb T_{db} and wet bulb T_{wb} temperatures respectively (Barreca (2012); Geruso and Spears (2018); LoPalo (2019)). Theoretically, keeping the value of the T_{db} constant on a day with 100% humidity, T_{wb} is equal to T_{db} and is lower for lower values of humidity. Given, that responsiveness of physiological processes are more accurately represented when humidity is taken into account, cognitive processes are also expected to be more associated with T_{wb} over T_{db}. I confirm this in the empirical exercise that follows.

Lastly, there is a concern that chronic exposure to high temperatures through childhood, may alter development of physiological systems in a way that reduces environmental interactions as adults (Riniolo and Schmidt (2006)). While transitory exposure to environmental

[3]For survey see Riniolo and Schmidt (2006).

[4]The core body temperature range is 97F (36.1C) to 99F (37.2C). For outside temperatures, however, this range constitutes being unusually high. For human beings to maintain normal functioning the ambient outside temperature should be around 22C or 72F (Westerterp-Plantenga *et al.* (2002))

[5]This particular study by Ramsey (1995) with 160 individuals records responses to a myriad range of tasks such as tracking, reaction time, vigilance, eye-hand coordination, mental, psychomotor, sensory, time estimation, monitoring, cognitive, complex and dual tasks.

stimuli has been much studied in laboratory settings, the literature on chronic heat exposure through childhood and their impact as adults is relatively sparse. This is a scenario, I attempt to resolve econometrically in the long term analysis results.

1.3 Data

1.3.1 Weather Data

Geographically gridded reanalysis data is obtained from the Princeton Meterological Forcing Dataset (PMFD) for the period 1980-2015.[6] There are several advantages of using such reanalysis datasets- First, it combines several sources of high-resolution data collected by satellites and weather station data into a single platform. For example, the PMFD combines data on weather variables from the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR), Climactic Research Unit (CRU) and Global Precipitation Climatology Project (GPCP). Second, it offers varying ranges of temporal coverage up to sub-daily and spatial coverage, upto 1 degree. Lastly, the PMFD is more appropriate for use in studies that primarily deal with land-surface phenomenon.[7] It corrects for two sources of known biases- One, from using atmospheric data to model land surface phenomenon and another, from combing data collected by many different satellites. These biases have often been known to create noise in statistical analysis. Having a spatially and temporally consistent dataset reduces several of these concerns.

The next couple steps closely follows the methodology underlined in Geruso and Spears (2018). To calculate wet bulb temperature, I use average monthly dry bulb temperature as well as three other variables- specific humidity, pressure and precipitation.The three variables

[6] This discussion follows from Sheffield *et al.* (2006)).

[7] See for example Auffhammer *et al.* (2013).Heft-Neal *et al.* (2017) that shows the use of above surface weather temperature as an improvement over air surface temperature when modelling responses of individuals.

of interest are reported daily at a spatial extent of 0.25 x 0.25 degrees. [8]Using specific humidity and pressure, first relative humidity is calculated. These values of relative humidity and dry bulb temperature are then plugged into the Schull's identity for calculation of wet bulb temperature. To match this data to test data, I aggregate daily dry bulb and wet bulb temperature data up to the month.

I build on a highly geographically dis-aggregated dataset (available online) that contains latitude-longitude centroids and BPS/Census codes for 2007 sub district boundaries. [9][10] sub districts are the third lowest level of administration- after Provinces (first) and Regency/City (second) but above villages. Furthermore, sub districts are also the lowest level for which the individual survey data is publicly available. Since, these sub district boundaries have changed over time, I merge the 2007 BPS codes with the IFLS crosswalk to get geographically consistent sub districts over time. Each sub district is then matched to the four nearest latitude-longitude grid points. The temperature across these points, weighted by the inverse distance function is calculated and assigned back to the centroid of the respective sub district. Lastly, I merge in individual data.

1.3.2 Indonesia Family Life Survey

This study uses the four most recent waves of the Indonesia Family Life Survey, conducted in 1997, 2000, 2007 and 2014, for which the exact month and year that the Ebtanas is

[8]At the equator, a 1^o degree distance between latitudes and longitudes is approximately equal to 69.171 miles. Therefore, a 0.25 X 0.25 grid measures up to 17.29 X 17.29 sq. miles of area. The area of this grid differs as we move further away from the equator. Since the country of Indonesia spans only 6 degrees to the north of the equator and 11 degrees to the south, the grid points considered here may be assumed to be all of the same size.

[9]The dataset for the geocodes is available from a github source for free at https://github.com/hidesys/IFLS_GIS. The programmer uses Google Geocoding API to collect latitudes and longitudes which represent the centroids of over 3500 sub districts/kecamatans. To this, I added a further 100 geocodes to encompass a few more households. Also included are the 2007 Census/BPS codes of kecamatans for matching with IFLS wave 4.

[10]The Central Statistical Organisation of Indonesia (BPS) changed the denotation of Kecamatans from sub districts to districts in 2014. Here, I stick to the nomenclature used by IFLS.

taken by the student is available. Test score data and detailed characteristics on over 13,000 students are used in the first stage analysis. Table 1, Panel A summarizes the demographic characteristics of the test takers across the four waves. The average test taker is 12.67 years old. Age at exam is computed using birth year, which typically has some reporting bias, as respondents usually round up their ages to the closest year when reporting it.

An important determinant of educational achievement is parents' education. The mean educational achievement of mothers is higher than fathers by 0.54 years. While here it is reported as a continuous variable, the impact of parental education on student achievement is largely non-linear.[11] This is especially true for father's education. For instance, I observe a discrete jump in the likelihood to enroll in secondary education if the student has a father with at least one year of secondary education. For mothers, any level of education leads to greater educational achievement in students at an increasing rate. As such, the regression results include dummies for parental educational achievement.

One of the advantages of using a panel survey is the ability to follow through a sub-sample of the test takers to maturity. For around 8000 of the respondents, Panel D reports a select number of long term outcomes.

While in the actual exam women marginally outperform men in all the subjects, the average educational achievement for women is 7.87 years and is lower than that of men by 2.5 years. Given, that primary school involves around 7 years of education, this difference consists of the difference between receiving any secondary schooling at all, for women and not. As such, the performance in the Ebtanas may potentially lead to a discrete culmination of education for women. Labor market figures bolster this hypothesis by indicating that women on average are more involved in self employment.

Also reported is the age of marriage, as poor performance may hazard an early entry into the marriage market. However, this figure must be interpreted with caution as due to

[11] Results available upon request.

the survey design, which I explain section in later sections, I only observe the most recent marriages for women as opposed to the first marriage for men.

1.4 Academic Performance and Temperature in the Short Run

1.4.1 Empirical Strategy

Weather conditions have been shown to be associated with several aspects of individual productivity. In this section, I exploit variation in temperature in the month of exam, in the year that each respective student appeared for their Ebtanas. The threat to causal identification in this context is that temperature assignment may not be random with respect to students. That is, if low performing students as well as poor school infrastructure were located in regions that experienced more heat stress. Province interacted with month fixed effects account for any systematic differences like these that may be correlated with both the temperature and test score variable.[12] I further include a dummy for urban location to account for differences in school infrastructure such as the possible presence of air conditioners.

Moreover, while unlikely, given that over 82.5% of the students in the sample take the exam in the same sub district that they reside in, high performing students may endogenously sort themselves into schools in sub districts with a more ambient climate.[13] To a large extent endogenous sorting across temperature assignments is mitigated by the rigidity of the exam

[12]Note that the prefered specification uses a province by month fixed effects which are coarser controls than district or subdistrict by month fixed effects. This choice is due to sample size limitations. In the current sample, on average 52 people take the test in a typical district. This implies that in a given exam year (May-August), around 14 people will be taking the test in that district. Hence, there is not enough variation at levels lower than provinces. Moreover, to the extent that these fixed effects might also control for policy changes, as sub districts are essentially municipal level bodies under the administrative legislation of provinces, systematic differences in policies and infrastructure are more likely to occur at the provincial level rather than the district or sub district level.

[13]This number rises to over 90.2% of the students in the sample who take the exam in the same district that they reside in and 95.9% in the province of residence.

schedule set years in advance and the fact that it is arbitrated by the Ministry of Education. As such, temperature realization in a given month is as good as random.

Formally, the following two variants of the short run model is estimated at the individual year-month sub district level.

$$y_{sdpmt} = \beta_{0sdpmt} + \beta_1 Temp_{dpmt} + X_{sdp}\Gamma + \theta U_{sdp} + C_t + \theta_{mp} + \epsilon_{sdpmt} \tag{1.1}$$

Here, y_{sdpmt} is the Ebtanas score for student s in sub district d and province p in the month m of year t in Ebtanas core subjects that include Math, Science, Social science, Bahasa or Indonesian language, English and the cumulative of all these scores. $Temp_{dpmt}$ is the dry and wet bulb temperatures in the subdistrict d and province p in the month m of year t .$\beta_1$estimates the impact of a 1 C increase in temperature in the month of the exam.

Time-invariant omitted variables may bias the OLS estimates, overstating the actual impact. For example, students with less educated parents may mechanically perform worse irrespective of weather. The household survey structure of the data allows me to control for a rich set of individual characteristics like age, sex and dummies for parental educational attainment to improve precision. X_{sdp} is an array of these individual controls.

Lastly, to account for time-variant factors such as the rigor of the test that may similarly affect students in the same cross-section, or seasonality effects from taking the exam at different agricultural cycles of the year, I control for cohort C_t and province by month θ_{mp} fixed effects respectively. U_{sdp} is a binary control for residence in an urban heat island and ϵ_{sdpmt} represent standard errors clustered at subdistrict by year level.

1.4.2 Results

Linear Model estimates

Estimates of the form of equation (1) are reported in Table 2 and Table 3. In columns 1 and 5 of Table 2, I corroborate the findings from previous studies (LoPalo (2019), Geruso and Spears (2018)) that human functioning, or in this case, cognitive performance is sensitive to wet bulb temperature T_{wb}. The biggest realizations of an effect are in math and science scores. This is related to the conjecture that the region of the human brain responsible for mathematical computing is also the region most sensitive to heat stress. Conversely, religion studies is least affected by temperature.

Columns 2, 3, and 4 replicate the wet bulb results with increasingly finer geographic fixed effects. The regression coefficients show the effect of a 1C rise in T_{wb}. In particular for science scores, conditional on province by month seasonality effects, I find the highest cognitive toll. An 1C increase in T_{wb} in the month of exam, leads to a 0.08 point decline in science score. Given that the mean science score is 7.1, this constitutes a decline of about 1.13%.

The above estimates are based on the assumption that the effect of temperature is linear at every degree of increase, which is an assumption I relax in the next section. While demonstrably, a non-linear model has larger point estimates at higher temperatures, due to the smaller sample size, standard errors are larger as well. As such, the linear model is a good approximation of these results.

The non-linear impact of Wet bulb temperature on test scores

Building on the hypothesis that the effect of temperature is non-linear at the extremes, degree day bins over given temperature ranges are created (Deschênes and Greenstone (2011)). Given that in the current setting, the annual range in the variation of temperature is a little over 10 degrees, these bins are much tighter than in the cited studies.[14] The following variant of equation 1 is estimated-

$$y_{sdpmt} = \beta_{0sdpmt} + \sum_{j} \beta_j DDTemp_{dpmt} + X_{sdp}\varpi + \gamma U_{sdp} + C_t + \theta_{mp} + \lambda_{sdpmt} \tag{1.2}$$

$DDTemp_{dpmt}$ is the number of days falling in temperature bin j in the month m of year t. β_j represents the impact of an extra day falling in 15 different 0.5 degree centigrade temperature bins defined as (<21), [21-21.5), [21.5-22), [22-22.5), [22.5-23), [23-23.5), [23.5-24), [24-24.5), [24.5-25), [25-25.5), [25.5-26), [26-26.5), [26.5-27), [27-27.5) and [27.5-) against the omitted category of [21.5-22) for wet bulb and [22-22.5) for dry bulb temperature. In equation (2) a simplified linear model is estimated where $Temp_{dpmt}$ is the dry and wet bulb temperatures in the sub district d and province p in the month m of year t.

Note that by definition, T_{wb} is lower in value and range than T_{db}(see Figure 2). The value of T_{wb} ranges between 18.75 to 28.16 in the present sample, with a mean of 24.9. Corresponding to this, the minimum and maximum values of T_{db} is 21.67 and 29.43 respectively, with a mean of 26.73. As such, in regressions of equation (1) the omitted category for T_{wb} is set at [21.5-22) and the omitted category for T_{db} is set at [22-22.5). Tests for joint significance of all the temperature bins in these regressions indicate significance at 5% levels in most cases.

Figure 3 plots the regression coefficients across all the T_{wb} bins. Values of T_{wb} that lie within the range [24.5-25), corresponding to the mean value of T_{wb}, leads to a 0.04 point decline in mean math score with an increasing rate of decline with every incremental bin. This constitutes a decline of about 0.5% from the mean math score. As I discuss later, it is worthwhile to note that these estimates are very similar to the ones obtained from a linear

[14] Annual variation in temperature ranges between 7-10 C in Indonesia, which straddles the equator on both sides. In most developed countries, that lie much further away from the equator, temperature varies by this amount over an entire day, several times a year.

model. Estimates at the lower tail of the temperature distribution such as ones (<21), also have a negative impact on scores. Although not the focus of this paper, this is an interesting finding as there is some physiological evidence that lower than ambient temperatures are similarly disruptive to cognitive behavior.

Heterogeneity in impacts by level of proficiency

It is conceivable that the impacts on cognitive performance is driven by certain subgroups of the population. Table 4 further explores the heterogeneity of these results with respect to men and women. Column 1 replicates the second column from Table 3. The effect of heat stress on cognitive function is higher for women than for men across all subjects, with a higher impact in subjects with higher cognitive load such as math and science.

I also test whether students at different levels of ability exhibit different levels of sensitivity at the same level of heat stress. For this, I run regressions of the form of equation (1) at different quantiles of student ability, as indicated by their math scores. Figure 4 reports these coefficients at a moderate and extreme level of T_{wb} relative to the omitted category of [21.5-22) . At moderate levels of heat stress in panel A, the high ability students are most affected followed by median ability students. However, at high levels of heat stress, students in the middle of the distribution are more likely to be affected by heat stress. It is noteworthy that in both scenarios, students in the lowest decile who are least probable to graduate to secondary levels are also affected the least by heat. These results indicate that students in the middle of the ability distribution, who may barely reach the requisite scores to graduate, face extremely high costs of random fluctuations in weather.

Robustness and Reporting Bias in Test scores

One of the drawbacks of using survey reported test score data is that respondents may manipulate reported scores around the threshold that gains them entry to secondary school

(Ozier (2015)). Incidentally, in IFLS a subset of the respondents report bringing the score card with them to the interview session. The interviewer in these cases report that the score card was seen by them. Manipulation around the threshold score may lead to non-classical measurement error. I observe limited evidence of score manipulation in Figure 5. Respondents tend to round up the reported scores on a subject by subject basis, most likely due to a recall problem, but do not systematically manipulate the scores around certain thresholds. Hence, as a robustness of the first stage estimates, in Table 5, I run the regression models using the sub-sample of test takers who bring the report cards to the interview. These estimates are very similar to the ones reported earlier.

1.5 Long-run consequences of Exposure to High Temperature and Humidity

1.5.1 Empirical Strategy

To the extent that ability to perform in the Ebtanas directly determines the enrollment in secondary school, a higher test score is likely to be associated with several later life outcomes. Using temperature in the month-year of the exam as an instrument for academic achievement, the impact on long term outcomes is examined. I focus on a sub-sample of around 8000 students who are at least 20 years old and have, as such, realizations of long term outcomes in the education, labor and marriage market.[15]

To satisfy the exclusion restriction in this set-up, I begin with a reduced form relationship between the instrument which is the temperature during month-year of exam, and long-term variables of interest. The assumption in this model is that, controlling for other relevant factors, temperature variation during the month-year of the exam will affect long-term outcomes

[15]This restriction implies that students in this sample must have taken their Ebtanas prior to 2005, so as to be at least 20 years of age in IFLS 5. Only IFLS 5 round is used for these set of results.

only via the channel of performance in the exam. As argued in the previous section, this is a plausible assumption. Similar to the short-term regressions before, the relevant controls include time-invariant factors such as age, sex, parental education, an urban dummy and time-variant factors such as cohort and province by month fixed effects.

Following is the reduced form model estimated at the individual sub district province level-

$$L_{sdp} = \gamma_{0sdp} + \gamma_1 Temp_{dpmt} + X_{sdp}\Pi + \delta U_{sdp} + \kappa_{mp} + \epsilon_{sdp} \tag{1.3}$$

where, L_{sdp} are long term educational outcomes such as enrollment in secondary schooling, educational attainment and grade repetition for student s in sub district d and province p. Labor market outcomes relate to earnings and occupational choice while marriage outcomes relate to qualitative measures of marriage. $Temp_{dpmt}$ is the dry and wet bulb temperatures in the sub district d in month m of year t. X_{sdp} is an array of individual controls, while U_{sdp} is a binary control for residence in an urban area. κ_{mp} are seasonality fixed effects and ϵ_{sdp} represents clustered standard errors .

Provided that the exclusion restriction is satisfied, I next estimate a 2SLS model for sufficiency. Using wet bulb temperature as an instrument for math scores, which shows the most cognitive toll from adverse weather conditions as well as yields the maximum sample points, the first stage is estimated as in equation (4). The predicted values from the first stage is then plugged into the second stage in equation (5).

$$MathScore_{sdpmt} = \zeta_{0sdpmt} + \zeta_1 Temp_{dpmt} + X_{sdp}\Delta + \lambda U_{sdp} + m_t + C_t + e_{sdpmt} \tag{1.4}$$

$$L_{sdp} = \rho_{0sdp} + \rho_1 \widehat{MathScore}_{sdpmt} + X_{sdp}\Theta + \tau U_{sdp} + m_t + C_t + \upsilon_{sdp} \tag{1.5}$$

where, $MathScore_{sdpmt}$ is the Ebtanas score in Math for each student-month-year cell in sub district d and province p. $Temp_{dpmt}$is the wet bulb temperature in the month of the year that the exam was taken. L_{sdp} are the long term outcomes while X_{sdp} consists of individual controls and U_{sdp} represents an urban dummy. m_t andC_t are fixed effects for month and Cohort respectively. e_{sdpmt} and $\upsilon_{sdp}$are robust standard error terms.

1.5.2 Reduced Form Results

Random variations in the level of temperature during the month that the student takes her Ebtanas exam may affect long-term outcomes, after controlling for observable student characteristics. Specifications that include cohort and province fixed effects account for systematic differences that affect both student quality and scores obtained. As such, the long-term effect of the temperature in the month-year of the exam will work through the sole mechanism of its effect on the Ebtanas score.

I begin with the reduced form results of the effect of T_{wb} in Tables 6, 7, 8 and 9. These results are focused on a sub-sample of students over 20 years of age with realizations of outcomes in education, labor and marriage market.

Impact on Educational Attainment

Using variants of equation (3), estimated reduced form coefficients for several standards of educational attainment are reported in Table 6. These include years of education obtained and dummies for enrollment in secondary, post-secondary education and a dummy for grade repetition.

A 1C rise in T_{wb} leads to a 0.126 point decline in years of education for the average person. Given that the average level of education is 8.2 years, this roughly translates to a 1.53% decline in an individual's educational achievement. Moving along, Columns 2 and 3 indicate that this effect is mostly driven by women. A 1C rise in temperature leads to a 1.84% less educational achievement in women, while leaving an insignificant impact for men.

It must be noted here that on average men receive 2.5 more years of education than women beyond the primary level (see Table 2). For women therefore, the heterogeneous impact of high temperatures during the Ebtanas amount up to the difference between getting and not getting secondary education.

Next I turn to the probability of enrollment in secondary education. Again, 1C rise in temperature lowers the probability of enrolling by 2.3% for an average person and by 2.6% for women in particular. Again, this impact is not reiterated amongst men. This impact is substantial on aggregate, since only 34% of women go on to enroll in secondary education at all. As such, variations in temperature are an important factor that explains the decision to enroll for women.

Rates of secondary completion, conditional on enrollment is significantly affected by random variations in temperature for men. As noted earlier, the hazard to cognitive performance from high temperatures is different for different students. For instance, it is more likely that the marginal cost from heat stress to the median scoring student, whose decision to go to the next level is affected by whether or not they get a qualifying score, is much higher than the marginal cost to a high-scoring student, who is more likely to go to complete their education anyways. It is plausible that a simple reduced form specification fails to account for this difference.

Lastly, both men and women face an increased probability of grade repetition as a result of high temperatures during Ebtanas.

Impact on Marital outcomes

Table 7 depicts the impact of temperature during the month of the exam on some qualitative outcomes for marriage, after controlling for relevant socioeconomic characteristics. Adverse weather has been shown to increase the hazard rate of marriage for women in other settings (Corno *et al.* (2020)). To the extent that weather conditions lead to poor test performance, expectations about unfavorable labor market outcomes may drive qualitatively poor marriages.

This is more so, given that the average age of marriage for a respondent is 13.13 years, which is also close to the threshold age for entry into secondary school. Consequently, I explore several markers of qualitative indicators for marriage such as age at marriage, spousal age gap and dowry for both men and women.[16]

It is noteworthy here that the respondents for the module on marriage and fertility are men rather than women. While typically, this module is asked to women in household surveys in across the world, the prevalence of polygamy in Indonesia necessitates that these questions be answered by men for a complete marital history of the family.[17] For example, in the current sample, 92% of the men have a single spouse, 5.69% have two spouses and the rest have more than two spouses. For each of these male respondents, I keep information pertinent to the first marriage which is most likely to affected by temperature during the primary exam. Using spousal identification information, I match the women from the household roster to the male respondents.

Due to this survey design, while I am able to parse out the information related to the first marriage for men, for whom I have the complete marital history, for women, I only have information pertaining to their most recent marriage. This is evident in Table 2. The average age of marriage for women in the sample at 13.84 years is higher than the age of marriage for men at 12.54 years. A simple distribution of the age of marriage for women reveals bimodal characteristics. This implies that while on average women get married earlier for their first marriage, many women may get remarried at a later age.

Consequently, the results for women's age of marriage must be interpreted with caution.

[16]These are only a few of many possible indicators for marriage quality. Others such as Hicks and Hicks (2013) have looked at a wider range like self-reported health, attitudes towards children and decision-making power.

[17]For instance the Kenyan Life Panel Surveys and the Demographic and Health Surveys rely on the woman's questionnaire module for marital history.

For an average respondent, a higher T_{wb} raises the probability of being married before the age of 18 by 1.5%. Also, value of dowry received and the spousal age gap is observed to go up as a result of higher T_{wb}.

Impact on Wages and Employment

Educational performance is a direct determinant of how an individual performs in the labor market. Reduced form estimates of the effect of temperature in the month of exam on long-term labor outcomes such as labor force participation and indicators for reported primary activity is shown in Table 7. It may be expected that while both men and women are affected by high T_{wb}, women arguably face a greater risk of leaving the work force while men face a greater risk in terms of hours worked.

A high value of T_{wb} is associated with a lowered probability of labor force participation in both men and women. This measure is calculated as a dummy based on reported primary activity during the past week if the activity was working or helping to get work or job search, conditional on the person being in the labor force. On average, a 1C rise in T_{wb}, lowers the probability of working by 1%. This hazard is even higher for women in Column 3. Women are twice as less likely to participate in the labor force for every 1C rise in T_{wb}.

It is conceivable that the lower labor force participation of women is driven by a greater likelihood to work in housekeeping jobs. I show that the probability for a woman to be in a housekeeping activity as their primary activity rises by 2% for every 1C rise in T_{wb}. They are also shown to be less likely to be in paid work as a result of high temperatures.

In Table 9 results for hours worked and wages are shown. While the association of number of hours worked with T_{wb} is negative in general, it is insignificant. This is across both measures of hours worked- hours per week, which refers to a general annual measure, and hours last week which is based on recall of the week before the survey. Since, men in general are more directly involved in the labor market, these effects are more pronounced for men. In columns 2 and 3, this heterogeneity with respect to men and women is explored. The results for men indicate that a 1C rise in T_{wb}, leads to a loss 1.59 hours of work in the last week recall period. This translates to a decline of 2% of the mean hours worked by each male respondent. Impact on measures of income and productivity calculated as wages earned last month and wages earned hourly respectively, are also shown. These results are more prominent for men than for women.

1.5.3 Instrumental Variable results

Tables A.1, A.2, A.3 and A.4 show the returns to education in the three outcome groups discussed above. Over the years, while a few subjects being tested in the Ebatnas have changed, math is one of the core subjects for which I have the most consistent data. Moreover, from a cognitive toll perspective, random variations in weather affect these scores more significantly than other scores. Hence, T_{wb} in the month of exam is used as an instrument for the endogenous score variable.

A 1 point increase in math score, on a scale from 1-10 leads to 0.528 more years of education, which is a 6.42% increase in the average educational achievement. A higher score is also associated with a greater probability of enrollment in secondary education, secondary education completion and a lower probability of repeating a grade. Outcomes in labor and marriage market are also shown to improve as a result of higher test scores, and closely reiterate the results in the reduced form model. A 1 point increase in math score leads to a 0.1% higher probability of being in the labor force, a 0.1% lower probability of primarily working in household chores and a 0.084% higher probability of working for pay. Similarly, a higher score increases the average age of marriage for men by 0.6 years, reduces the probability of marriage by 18 and significantly reduces the amount of dowry exchanged.

It must be mentioned here that beyond labor force participation, hours worked in the labor market as well as wages earned therein are also negatively affected by high temperatures.

In an IV setting, this is predictive of labor market returns to academic achievement. I show that every additional point scored in the exam, leads to an increase in the hours worked by 2.33%, total monthly wages earned by 15% and hourly wages earned by 14%. However, these results are indicative and not significant in the IV form, plausibly due to insufficient power in the instrument.

An important caveat emerges from the above analysis. Random variations in the performance of the Ebtanas directly determine individual decision making in early life. Given, that the cohort considered here is relatively young, it is entirely plausible that, human capital accumulated over time may either strengthen or weaken these effects over the individual's lifetime. Nonetheless, the persistence of these impacts over the present term, highlights the need for public policy directed towards adaptive mechanisms.

1.6 Chronic Heat Stress

Theoretically, beyond contemporaneous cognitive impairment, there are a couple ways that heat stress may have a chronic impact on long-term learning. Empirically, this may be tested by including temperatures preceding the month of exam as additional controls. However, given the low variation in monthly temperatures in a cross-section, these estimates are likely to be imprecise. Moreover, it must be noted that, these specifications may fail to capture the full extent of chronic heat stress vis-a-vis physiological distress, as the marginal vulnerability for different individuals may be different.

A drier than usual agricultural season is one such candidate. A dry agricultural season (August-October) is usually associated with delayed planting, thereby leading to shorter germination periods for crops and a poor harvest. As such, a dry agricultural season may impact students' family income and health, thereby leading to poor performance in the Ebtanas. However, in the case of Indonesia, this is unlikely since, the main agricultural season spans the months of November to February. This does not coincide with the timing when majority of the students take the exam, which is in May or June. Consequently, it is unlikely that students may be facing agricultural distress and heat stress at the same time.

To the extent, that the dry agricultural season coincides with the hot months of the school year, the effect of a drier than usual season may persist over the next few months. I attempt to address this in Table A.5. Test scores are regressed on monthly dry season temperature just preceding the month-year that the student takes the exam, after controlling for the effect of exam-month temperature. September temperatures, which is generally the hottest month in the season has a significant negative impact on test scores, signaling a chronic learning impact from heat. However, this effect disappears for math scores, which is known to draw the heaviest cognitive impairment from contemporaneous heat stress.

Alternatively, heat stress may have a chronic learning impact over time via hot temperatures during the hot months of preceding school years. Table A.6 shows that dry season temperatures in the year immediately preceding the exam year has the largest impacts on learning.

1.7 Discussion

Rising global temperatures over the next few decades, have left less developed countries, that are largely situated around the tropics, the most vulnerable to climatic fluctuations. The IPCC 5th Assessment report predicts in different climate model scenarios that relative to per-industrial levels, temperatures in the later part of the 21st century will exceed by 2C. In terms of human systems and human functioning, this poses a huge risk, given the sensitivity of these systems to temperature.

This paper serves to identify several of these vulnerabilities as a function of cognitive impairment. Hot temperatures while contemporaneously affecting test performance, has farreaching impacts on long-term human capital accumulation. Moreover, these impacts are realized early on in life- along dimensions such as occupational choice and early marriage which potentially lower the future quality of life. In doing so, this paper makes the case against high-stakes examinations around the world, where exam dates are rigidly set and students do not get the choice of picking a date.

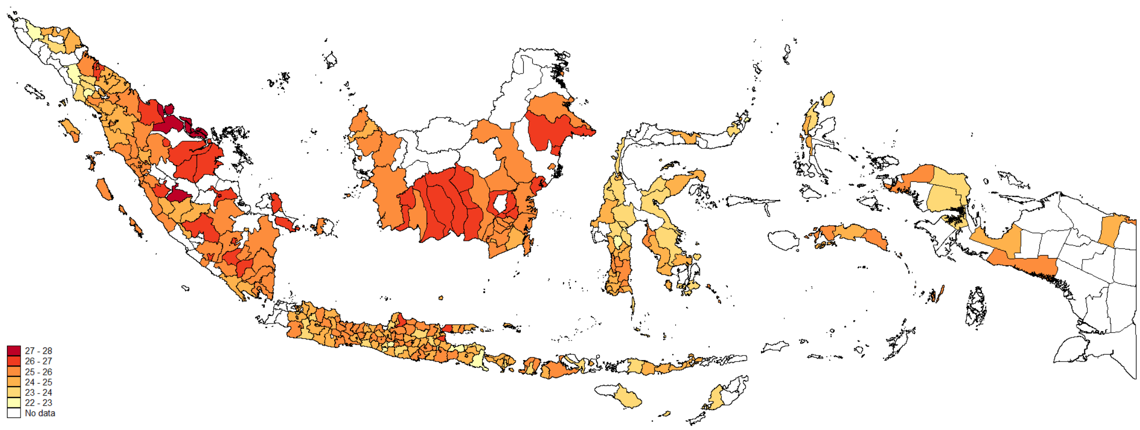
Tto mitigate these adverse effects, public policy may be informed to affect adaptive decision making. While in the context of less developed countries, expensive measures such as installation of air conditioners in classrooms may not be feasible, other low cost policy directives may be adopted at different levels of administration. To start, the dates and times of the exams are set by the Ministry of Education way in advance. Drawing from the results in this study, there are sufficiently large payoffs to setting the date of these examinations in the cooler months of the year, such as between October to February as opposed to the summer months of May-June. However, this obviously would put the academic cycles in a different time period than academic cycles in the rest of the world, potentially putting students from less developed countries at a disadvantage with their peers.

As such, over the medium run, other strategies may be adopted. For instance, exam timings may be moved to earlier in the day when heat stress is lowest rather than later in the day. Qualifying testing centers may be required to have adequate natural ventilation, seating arrangements away from direct sunlight and provision for drinking water. Adaptive measures at relatively low cost may be undertaken by students themselves to mitigate these negative effects if sufficient information and awareness is transmitted to the public. For example, students may be encouraged to wear lighter clothing and carry hydrating sources of drinks with them.

The Indonesian system of Education has undergone some changes in recent times. Starting in 2021, the Ministry of Education and Culture proposed the scrapping of the National Exam/Ebtanas as a tool to determine entry to secondary school, and replacing it with a minimum competency assessment and character survey to be held in the middle of the academic year rather than at the end. However, the implementation of these recommendations, as in the past, is questionable- given the opposition from several fronts. Insofar as that this paper identifies other horizons of potential gains, such educational policies are likely to improve outcomes in Indonesia and elsewhere, where school-leaving examinations are heavily relied upon as a token of merit.

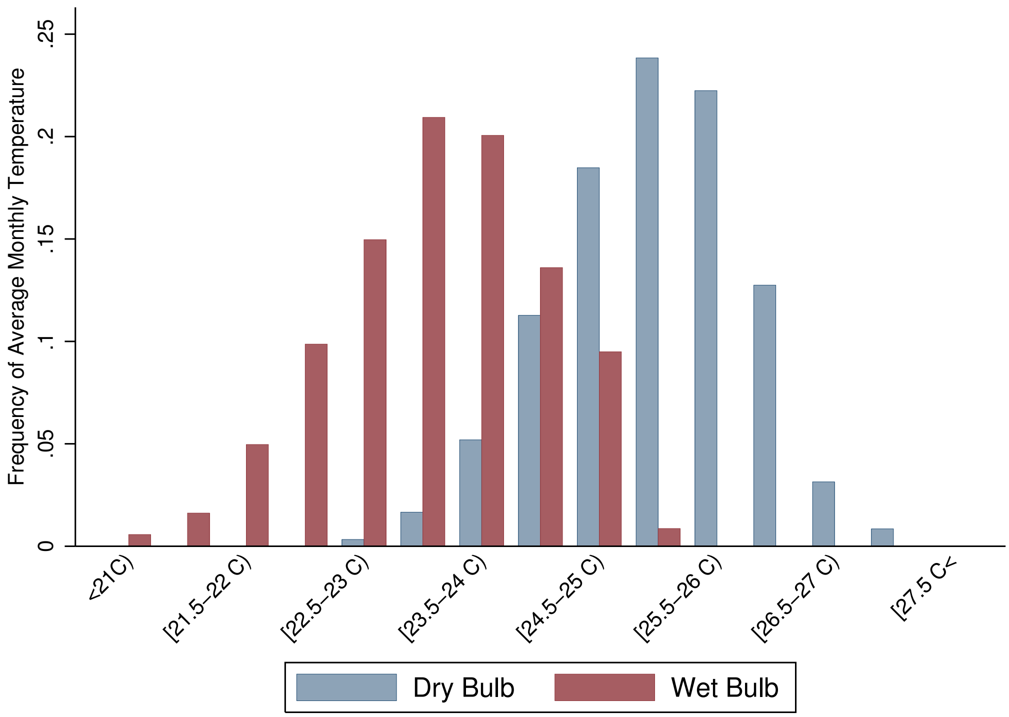
Figures

Figure 1.1: Variation in Wet bulb temperature by geographic coverage of test takers



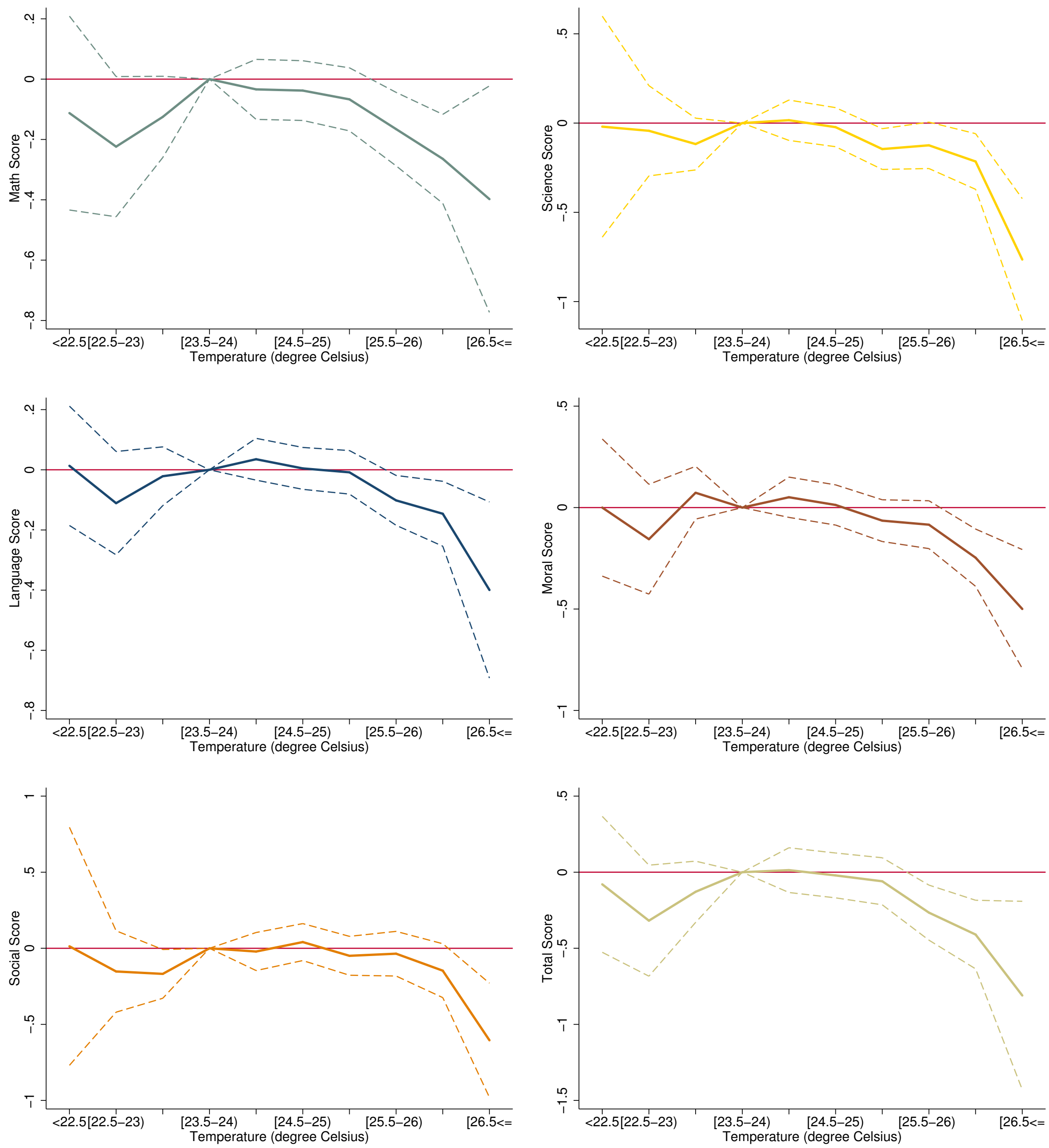
Note: Only sub districts which also report test score data are plotted here. No data does not indicate non-availability of weather variables. Weather data is obtained from the Princeton Meterological Forcing Dataset. Wet bulb temperature is given as a non-linear function of temperature and humidity, calculated using the Schull's identity. Test score data is obtained from the Indonesia Family Life Surveys (IFLS) available from 1997, 2000, 2007 and 2014. Regions in color correspond to the survey clusters in IFLS.

Figure 1.2: Variation in Wet bulb and Dry bulb temperature



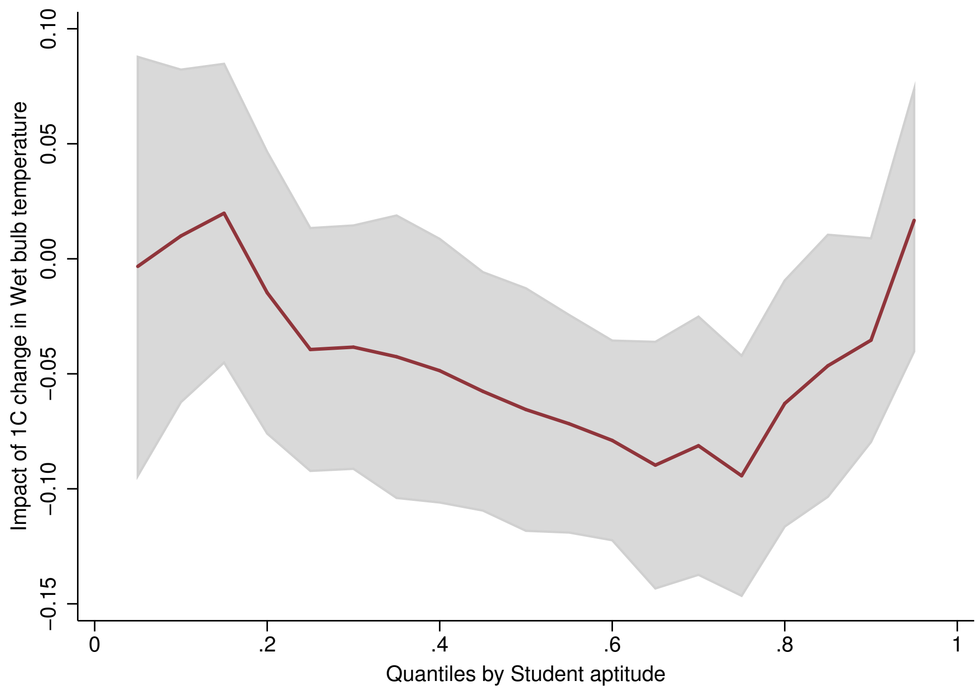
Note: Weather Data is from PMFD(Princeton Metereological Forcing Dataset). Wet bulb temperature is calculated following Schull's identity and uses dry bulb temperature, pressure, and humidity. Bar chart represents the average monthly by dry bulb or wet bulb for the study area. X axis denotes the bins for either dry or wet bulb temperature.

Figure 1.3: Non-linear effect of Wet bulb temperature on test scores by subject



Note: Points in the graph represent coefficients from the regression of subject level scores on wet bulb temperature. Estimates are reported at 90% confidence intervals, indicated by dotted lines.Independent variables are the number of days in a exam month that falls in the temperature bins listed along the X axis. Controls include student level characteristics, year, month, province and seasonality fixed effects.

Figure 1.4: Impact of temperature on students at various quantiles



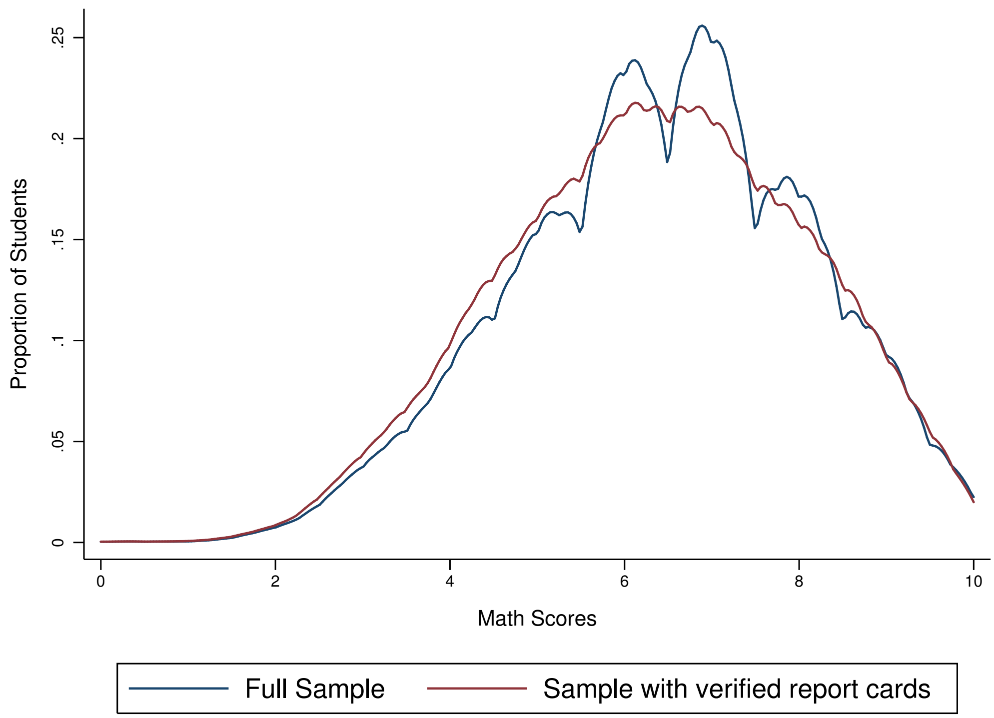
Note:
Points in the graph represent coefficients from the quantile regression of math scores on wet bulb temperature. Estimates are reported at 95% confidence intervals, indicated by shaded area. Controls include student level characteristics,year,month,province and seasonality fixed effects.

Table 1.1: Summary Statistics: Student Characteristics in IFLS and Weather Variables

	Obs	Mean	Std. Dev.
Panel A: Student Level Characteristics in IFLS			
Age at the time of Exam	14917	12.67	0.96
Sex (Male=1)	15395	0.49	0.50
Father's Years of Education:	11638	6.87	3.72
Mother's Years of Education	10989	7.41	3.98
Dummy for Urban Residence	15395	0.55	0.50

Note: Student level data is from four waves of the Indonesia Family Life Survey.

Figure 1.5: Reporting Bias in test scores



Note:
Verified sample consists of 78% of the students in the full sample. Comparison of the two samples suggest that respondents who did not bring their report cards to the interview center tend to round their scores up or down during recall. The plot does not suggest presence of systematic upward manipulation of scores by respondents.Estimates of the results are shown in Table 5.

Table 1.1: (contd) Summary Statistics: Short and Long Term Achievements by sex

	All	Men	Women
Panel B: Exam Level Characteristics			
Math Score (0-10 scale)	6.442	6.398	6.484
	(1.662)	(1.660)	(1.663)
Science Score (0-10 scale)	7.076	6.949	7.196
	(1.202)	(1.202)	(1.189)
Social Score (0-10 scale)	6.380	6.397	6.364
	(1.420)	(1.404)	(1.435)
Language Score (0-10 scale)	6.158	6.139	6.176
	(1.487)	(1.471)	(1.502)
Religion Score (0-10 scale)	7.220	7.165	7.272
	(1.243)	(1.233)	(1.250)
Panel C: Select Long-term Outcomes			
Years of Educational Attainment	8.223	10.34	7.873
	(4.218)	(2.776)	(4.311)
Hours Worked Last week	63.46	60.58	63.91
	(27.82)	(27.34)	(27.87)
Age at Marriage	13.13	12.54	13.85
	(5.840)	(5.985)	(5.580)

Note: Values for only select long-term outcomes are shown here. Age of marriage for men is lower than women as this question is based on men's recall in the IFLS Book 3A which has information on marital history. The survey design is due to the prevalence of polygamy. Therefore, while information on most men are available from their first marriage, for women married to them, this may not be the case.

Table 1.2: Short Term Effect of Temperature on Ebtanas Test Scores

	Dry Bulb Temperature	Wet Bulb Temperature			
	(1)	(2)	(3)	(4)	(5)
Mathematics	-0.055**	-0.055***	-0.036**	-0.057***	-0.049**
	(0.025)	(0.014)	(0.014)	(0.017)	(0.025)
Science	-0.082***	-0.042***	-0.014	-0.030	-0.082***
	(0.027)	(0.016)	(0.016)	(0.018)	(0.027)
Language	-0.047***	-0.038***	-0.022**	-0.028**	-0.047***
	(0.018)	(0.010)	(0.011)	(0.012)	(0.017)
Religion	-0.074***	-0.033***	-0.017	-0.034**	-0.073***
	(0.025)	(0.013)	(0.014)	(0.016)	(0.025)
Social Studies	-0.036	-0.049***	-0.024	-0.028**	-0.033
	(0.030)	(0.015)	(0.016)	(0.012)	(0.031)
Total Core score	-0.184**	-0.141***	-0.058	-0.097**	-0.189***
	(0.072)	(0.037)	(0.038)	(0.044)	(0.071)
Province FE		X			
District FE			X		
Subdistrict FE				X	
Province X Month FE	X				X

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and clustered at the subdistrict by year level. Dependent variables are the core Ebtanas subjects, listed along the first column. The number of observation N equals 12613, 8085, 12445, 7950 and 7821 for math, science, language, religion and social studies subjects respectively. Province by month fixed effects are local seasonality controls. Other controls include age,sex,parents education and an urban dummy.

Table 1.3: Heterogeneity in the Short Term Effect of Wet Bulb Temperature on Ebtanas Test Scores

	All (1)	Men (2)	Women (3)
Mathematics	-0.049**	-0.066**	-0.028
	(0.025)	(0.033)	(0.032)
Science	-0.082***	-0.088**	-0.070**
	(0.027)	(0.037)	(0.035)
Language	-0.047***	-0.057**	-0.034
	(0.017)	(0.023)	(0.023)
Religion	-0.073***	-0.123***	-0.023
	(0.025)	(0.034)	(0.032)
Social Studies	-0.033	-0.050	-0.014
	(0.031)	(0.042)	(0.040)
Total	-0.189***	-0.199**	-0.164*
	(0.071)	(0.097)	(0.091)
Province x Month FE	X	X	X

Note: *** $p<0.01$, ** $p<0.05$, * $p<0$.Standard errors are robust and clustered at the subdistrict by year level.Dependent variables are the core Ebtanas subjects, listed along the first column. Province by month fixed effects are local seasonality controls. Other controls include age,sex,parents education and an urban dummy.

Table 1.4: Robusness Check 1: Report Cards presented to Interviewer

	Full Sample	Sample restricted to respondents who showed report cards
	(1)	(2)
Mathematics	-0.049**	-0.044
	(0.025)	(0.027)
N	13,643	11,616
Science	-0.082***	-0.092***
	(0.027)	(0.029)
N	8,659	7,447
Language	-0.047***	-0.037*
	(0.017)	(0.019)
N	13,470	11,600
Religion	-0.073***	-0.082***
	(0.025)	(0.028)
N	8,535	7,319
Social Studies	-0.033	-0.029
	(0.031)	(0.033)
N	8,377	7,233
Total	-0.183***	-0.187**
	(0.060)	(0.077)
N	8,608	7,441

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis. Dependent variables are the core Ebtanas subjects, listed along the first column. Independent variable is the continuous wet bulb temperature in the month of exam. Results are the reported coefficients from the estimation of equation (1). They include month, cohort, province and seasonality fixed effects besides student level controls like age,sex,parents education and an urban dummy.

Table 1.5: Reduced Form Estimates of the Effect of Wet Bulb Temperature on Long Term Education Outcomes

	All (1)	Men (2)	Women (3)
Years of Education	-0.126*	-0.096	-0.145*
	(0.067)	(0.100)	(0.087)
N	6,237	2,887	3,350
R^2	0.124	0.107	0.195
Secondary Enrolment	-0.025**	-0.023	-0.026**
	(0.010)	(0.015)	(0.013)
N	6,249	2,891	3,358
R^2	0.168	0.166	0.198
Secondary Completion	-0.016*	-0.035**	0.001
	(0.009)	(0.015)	(0.012)
N	4,122	1,978	2,144
R^2	0.079	0.101	0.098
Grade Repitition	0.007*	0.009	0.004
	(0.004)	(0.009)	(0.002)
N	2,216	1,093	1,123
R^2	0.348	0.393	0.098

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis. Dependent variables are listed along the first column. Results include month, cohort, province and seasonality fixed effects besides student level controls like age,sex,parents education and an urban dummy.

Table 1.6: Reduced Form Estimates of the Effect of Wet Bulb Temperature on Long Term Marital Outcomes

	(1)
I(Married by 18<0)	
All	0.015*
	(0.009)
Men	0.009
	(0.013)
Women	0.022
	(0.014)
Spousal age Gap	0.551**
	(0.236)

Note: *** p<0.01, ** p<0.05, * p<0.1.Standard errors are robust and in parenthesis.Dependent variables are listed along the first column. Results include month, cohort, province and seasonality fixed effects. For full list of controls, see text.

Table 1.7: Reduced Form Estimates of the Effect of Wet Bulb Temperature on Long Term Labor Market Outcomes

	All	Men	Women
	(1)	(2)	(3)
Panel A:			
Labor Force Participation	-0.010	-0.006	-0.016
	(0.009)	(0.010)	(0.014)
N	6,133	2,838	3,295
R^2	0.238	0.057	0.045
Primary Activity: Housework	0.009	-0.001	0.020
	(0.008)	(0.005)	(0.014)
N	5,965	2,707	3,258
R^2	0.316	0.025	0.060
Primary Activity: Paid work	0.002	-0.013	-0.009
	(0.009)	(0.008)	(0.014)
N	6,133	2,838	3,295
R^2	0.094	0.106	0.094

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are robust and in parenthesis.Dependent variables are listed along the first column. Results include month, cohort, province and seasonality fixed effects besides student level controls like age,sex,parents education and an urban dummy.

Table 1.8: Reduced Form Estimates of the Effect of Wet Bulb Temperature on Long Term Labor Market Outcomes

	All (1)	Men (2)	Women (3)
Panel B:			
Hours Worked last week	0.105	-1.115	1.198
	(0.497)	(0.692)	(0.731)
N	8,752	4,223	4,529
R^2	0.049	0.067	0.052
Hours Worked per Week	0.083	-0.804	0.841
	(0.448)	(0.605)	(0.673)
N	8,752	4,223	4,529
R^2	0.047	0.063	0.052
Log(Wages Earned	0.045	-0.001	0.103
Per Month)	(0.099)	(0.124)	(0.165)
N	3,068	1,871	1,197
R^2	0.071	0.089	0.112
Log(Hourly Wages	-0.032	-0.024	-0.052
Earned)	(0.038)	(0.047)	(0.063)
N	3,063	1,869	1,194
R^2	0.042	0.053	0.077

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis.Dependent variables are listed along the first column.Wages are calcualted in millions of Indonesian Rupiah. Results include month, cohort, province and seasonality fixed effects besides student level controls like age,sex,parents education and an urban dummy.

Table 1.9: Placebo Checks: Impact of temperature in months after exam

	Math	Science	Language	Religion	Social Studies	Total
	(1)	(2)	(3)	(4)	(5)	(6)
# days with T_{wb}>27.5	-0.028	-0.025	-0.038	-0.068*	-0.102**	-0.165
in month +1 of exam	(0.038)	(0.043)	(0.028)	(0.037)	(0.046)	(0.109)
# days with T_{wb}>27.5	-0.023	-0.081	0.046	-0.000	0.071	-0.118
in month +2 of exam	(0.049)	(0.053)	(0.037)	(0.046)	(0.057)	(0.133)
# days with T_{wb}>27.5	-0.003	0.025	-0.067**	-0.002	-0.004	0.079
in month +3 of exam	(0.035)	(0.043)	(0.026)	(0.038)	(0.045)	(0.107)
N	13,643	8,659	13,470	8,535	8,377	8,608
R^2	0.179	0.149	0.202	0.161	0.134	0.165

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis.

Chapter 2

Environmental Factor Endowments, Technological Diffusion and Inequality: Evidence from the Green Revolution in India

2.1 Introduction

Weather, terrain and the depth of the topsoil are few of the many geographical attributes that determine why different crops flourish in different environments. For instance, while paddy rice is geographically limited to the lowland tropical climate of South-East Asia, maize is grown extensively in the temperate, uneven terrain of Africa and Central America. Local farming systems generally incorporate these environmental factor endowments and factor limitations into their cultivation modes. As such, in modern times, the pattern of technological diffusion in agriculture is determined to a large extent by the different crop regions of the world.

In this study, I examine the impact of an agricultural productivity shock spurred by the introduction of High-Yield Variety (HYV) seeds that increased yields per hectare in rice and wheat growing crop regions of India. To the extent that differences in the environmental factor endowments of wheat and rice regions determine the pattern of technological adoption, the impact of productivity on labor and inequality outcomes may also differ.

Rice and wheat are the two major staple crops in India, together occupying over 37% of the net cultivable area in India. Starting in 1960s, HYV seeds of wheat and rice were introduced and widely adopted by predominantly wheat and rice producing districts.[1] Between 1965-85, while wheat districts moved to highly automated modes of production, the mechanization level in rice districts were more modest. I demonstrate this in the abrupt increase in the use of tractors and pumps in wheat districts. In rice districts, on the other hand, adoption of the new technical change was followed by a relative increase in the use of fertilizers, pesticides and chemical weeders, which are generally labor-intensive tasks.

I draw upon the differences in cultivation modes of rice and wheat in India, centered around the environmental factor endowments such as weather, soil and terrain, to explain these patterns. Environmental factor endowments regulate the importance of labor in the production process of rice and wheat. First, wheat is grown in a dry and resistant weather conditions. As opposed to this, rice is grown in an agro-ecologically delicate ecosystem that requires constant monitoring. This creates opportunities for labor in different steps of the rice production process- starting with transplantation from the nursery to farms, manual weeding, frequently checking soil acidity as farms remain waterlogged throughout the growth period to harvesting and threshing.

Next, the diversity of the soil conditions that surround rice ecosystems limit the scope

[1] The roll-out of HYV seeds has been extensively covered in the literature. Notable examples include Frankel (2015); Mohan and Evenson (1975); Murgai (1999); Sanyal (1983); Pingali (2007); Parayil (1992); Westley (2019)

of using a single variety of seed or technology to all farms.[2] As a result, in 1960 when the first imported rice seeds were introduced they failed to be widely adopted. Being of a single strain, the initial seeds were susceptible to local soil stress, posed from differences in soil alkanity or moisture. Since, this resulted in significantly lower productivity gains accruing to rice farmers, from the perspective of the rice farmers, there was little incentive for further investments in mechanization. Lastly, environmental factor endowments such as terrain plays an important role in mechanization. In India wheat is generally cultivated in highlands, which is better suited for the application of heavy machinery such as tractors and pumps. Rice, in contrast, is mostly grown in lowlands, unsuitable for the application of large labor-saving equipment.

The main results in this study pertain to the differential impact of the productivity increase on labor market outcomes. I show that these differences in the pattern of technical adoption, precipitated a divergence in factor accumulation of both regions. As more and more labor-intensive jobs were replaced by machines in wheat districts, it lowered the demand of labor, particularly for low-skilled laborers, essentially creating a displacement (of labor) effect from the agricultural sector. In the absence of simultaneous generation of tasks in which labor has a comparative advantage, I show that this effect was reflected in lower overall employment. Furthermore, I show that since the brunt of this displacement was borne by the now easily dispensable low-skilled agricultural workers, it also drove up inequality. The rice districts, on the other hand, experienced an increase in employment as well as a decline in inequality.

An important aspect to consider here is that if the divergence in mechanization was driven by systematic differences in other initial factor endowments of rice and wheat districts such as income and institutions, then the estimates of the effect of productivity would be biased. To account for factors which could potentially appear as unobservables, I combine data from

[2] Frankel (2015); Dasgupta (1977); Munshi (2004) discuss the lower success rates of the new rice seeds.

various sources on district level controls. Furthermore, using a number of robustness checks and instrumental variables I show that such factors did not play a role in the pattern of technological diffusion.

Earlier work on HYV adoption has shown to generate differences across a wide range of economic outcomes.[3]This study builds upon this extensive body of research in distinguishing between the experiences of the rice and wheat districts in terms of very different structural changes.[4] On the one hand, a movement towards more mechanization and lowered labor demand resulted in a significant displacement of low-skilled workers from the wheat regions. While on the other hand, employment across all dimensions of agricultural employment increased in rice districts which adopted a more labor-enhancing technology.

Primarily, this study makes two important contributions. First, it shows that geographical attributes surrounding different crop regions are important determinants of the patterns of productivity and technological diffusion. These environmental factor endowments guide the adoption of a labor-saving versus labor-enhancing mode of technology. In doing so, the increase in productivity has heterogenous impacts in different crop regions.

Second, it adds to the growing literature on the role of labor-saving technical change in making jobs previously performed by labor obsolete. Earlier work (Akerman *et al.* (2015); Bessen (2015); Graetz and Michaels (2018)) has shown that mechanization may deeply disrupt labor markets by contracting the demand for workers on one hand, and generating new tasks for labor, on the other. However, the impact of mechanization on agricultural labor demand has been discussed to a much less extent (exceptions include Olmstead and Rhode (2001); Eisner (2019); Manuelli and Seshadri (2014)). This study shows very similar results as other sectors, wherein the new technology had the dual impact of displacing workers from

[3]While, Bharadwaj *et al.* (2020) and von der Goltz *et al.* (2020) show that HYV adoption directly lowered infant mortality, Brainerd and Menon (2014) examine the long-term effect on child health in terms of sustained exposure to agrochemicals in water as a result of GR. Dasgupta (2018) studies the decline of dominant parties post-GR as a result of agricultural producers lobby

[4]Prahladachar (1983); Gollin *et al.* (2019); Moscona (2017)

the mechanized task in the wheat regions while setting off a mechanism that created the demand for labor in new tasks in rice regions.

Section 2 gives a historical overview of the Green Revolution. Sections 3 and 4 describe the data, methodology and the identification. Section 5 is a discussion of results, Section 6 explores the drivers of growth during the Green Revolution, Section 7 is results using an instrumental variable strategy and Section 8 concludes.

2.2 Agricultural Context and Background

2.2.1 Cultivation of Rice and Wheat in India

Rice in India is cultivated in the wetlands. Typically, farming involves several steps which are labor-intensive and involve careful planning and preparation to reduce crop stress. It starts with the pre-planting or land preparation stage, which involves tilling or plowing the soil to level it. This is done with the help of tractors or bullocks. Next, the seedling is transplanted from a seed bed to the wet field by hand. Once the rice is transplanted, the crops must be monitored for ample water while preserving the nutrient content of the soil. Rice farms usually remain flooded during their growth period which while it works in favor of preserving the nitrogen in soil, provides an active breeding ground for pests. Hence labor involvement in pesticide application and weeding activities is prevalent. Finally, at the end of the crop cycle, harvesting is done manually as well using simple tools such as sickles. Harvesting activities include cutting, stacking, handling, threshing etc. This is around the time that additional labor is hired, creating seasonal employment opportunities.

The labor-intensiveness of rice cultivation is also owed in a large way to the agro-ecological instability of rice cultivation. For example, the International Rice Research Institute (IRRI) classifies rice ecosystems into four major groups built around the availability and the ability to control the water in these regions- Irrigated Rice Eco System, Rainfed Upland Rice Eco

System, Rainfed Lowland Rice Eco System and Flood Prone Rice Eco System. Few countries in the world have such a diversity in rice ecosystems. This limits the capability of developing a universal technology for application to all rice farms. Best practices are often gathered by way of learning by doing and are only effective at the very localized level.

Table 1 describes the different characteristics of soil in rice and wheat regions. The top three soil types in order of their fertility are alluvial, medium black and red soils. In terms of area under alluvial soils, while wheat districts have an advantage, rice districts have a more even distribution across these three soil types. Other soil characteristics such as topsoil depth and soil acidity (a higher pH value is better in the sense that the plant can access more nutrients in the soil) also has minor differences on average but indicate that wheat districts may be slightly more fertile.

2.2.2 Green Revolution (1965-85)

The Green Revolution (henceforth to be denoted as GR), that took place in South Asia during the 1960s, marked a movement away from traditional modes of cultivation to modern and more sustainable methods and crops. In India, imported varieties of higher yielding (dwarf) seeds of rice and wheat were introduced to the districts which were climatically best suited for their cultivation. The need to adopt it in India, arose out of a severe food sufficiency problem. Recurrent droughts combined with heavy dependence on imports of food grains, created a need to reform and stabilize the agricultural sector. Thus, funded by the Rockefeller and Ford Foundations, dwarf varieties of wheat and rice were developed in the International Center for the Improvement of Corn and Wheat (CIMMYT) in Mexico and the International Rice Research Institute (IRRI) in Phillipines respectively. These new seeds were introduced as a pilot program to 15 districts under the Intensive Agriculture District Programme (IADP).

Imported wheat varieties were crossed with Indian varieties after their initial success

(Evensson et al. 1998; Evensson and Golin 2003; Gollin et al. 2005). This gave birth to the first generation of "High-Yield Variety" crops- Kalyan Sona and Sonalika. They were subsequently adopted across the wheat belt states of India. This region is characterized by a humid subtropical climate. In the north-west region, Punjab, Haryana and western Uttar Pradesh accounted for over 85% of all wheat production and came to be known as the "Bread Basket" of India. Other states in the Northeast like Bengal, Bihar and east Uttar Pradesh, and states in the center like Madhya Pradesh and Gujarat had scattered adoption of dwarf wheat seeds. The introduction of the imported rice varieties, were adopted across the East coast states of West Bengal, Orissa and several southern states like Tamil Nadu and Karnataka. Soon the first generation of Indian developed HYV rice varieties Padma and Jaya were also developed and rolled out to the rice growing districts.

In spite of some concerns related to the quality of the initial rice seeds, the documented increases in yield and productivity of both rice and wheat were substantial over the entire period.[5],[6] Between 1965 and 1970, after the new technology had been rolled out to the rest of the country, high yield wheat crops, imported from Mexico were showing a staggering rate of success in terms of increase in yields as well as rate of adoption.Table 2 sheds some light on the magnitude of HYV adoption over the period of study. It summarizes the trend in terms of area under cultivation and yield over five time period bins across the rice and wheat districts. Panel A shows the yield per hectare and area brought under HYV cultivation of rice over 1956-60 to 1981-85. As is evident, the yields as well as land brought under the new technology is gradually increasing over time. Starting in 1966, when it was first introduced,

[5]For example Dasgupta (1977) argues that the new rice seeds had three problems in particular that resulted in a slowdown of adoption and technical learning- First, the high turnover rate of seeds from frequent scientific trials. Second, the need for frequent restocking and the third, the overall lower quality of the seeds. In another instance, Frankel (1971) states that *"The new paddy varieties have shown less striking results. Important technical problems remain to be fully solved. Imported varieties show higher vulnerability to plant diseases than do local strains;crop duration... is commonly too long...and the coarse grain quality of the imported varieties compares unfavorably.....*"

[6](see for example Brown 1971, Randhwa 1971)

the proportion of land planted to HYV rice crops gradually rose from a 9.5 percent to 60 percent of the entire croplands under rice cultivation. Concurrent to technological adoption, the yields per hectare increased by over 53 percent in the entire timeline considered of which a gain of 9.8 percent was recorded in the first ten years of introduction. Panel B shows similar outcomes for wheat production, except that wheat recorded a much higher rate of growth. The adoption rate in terms of land brought under HYV cultivation increased to 77 percent during this time. Between 1956 to 80, the yield in tons per hectare of wheat rose by over 125 percent, of which a staggering 40 percent growth was recorded in the first ten years.

Despite an adoption increase, a slower rate of productivity increase meant that rice districts were lagging in certain aspects of the structural change, especially the adoption of labor substituting techniques. We show that this difference in productivity was widened especially during the later wave of GR (1975-85). Why the wheat districts took off in terms of mechanization, while the rice districts lagged can be summarized into two main reasons. Wheat districts, being traditionally located in drier climatic zones, were better suited for the application of mechanized modes of production like tractors and pumps. In contrast, rice is grown in the hot and wet subtropical regions of India as its cultivation requires standing water and heavy monsoon rains. However, this is very unsuitable for HYV cultivation due to rampant breeding of pests and weeds. Secondly, rice cultivation is generally susceptible to minor fluctuations in soil salinity and moisture from farm to farm. This withheld the smooth diffusion and integration of the rice technology from farm to farm.

Two crucial facts about the adoption of the new technology are yielded from the precluding discussion. First, the HYV seeds were developed abroad and thus unrelated to local economic conditions.Over the period of study they massively contributed to the increase in productivity. Second, the climate associated with cultivation of rice is distinct from that of wheat. This contributed to wheat and rice being adopted only to regions climatically suitable for them. Both these points allow us to exogenously demarcate the rice and wheat growing districts as well as obtain a variation in the rate of adoption.

2.3 Data Description

The data in this study is compiled from various sources and is detailed in the Appendix. The primary source is data on agricultural productivity, inputs and climatic conditions taken from World Bank India Agriculture and Climate Data for 1956 through 1985.[7] It is based on district-level data which is the second lowest administrative level in India, the lowest being a village. The number of districts covered were 271 in 13 states of India. This number is considerably lower than the current count of districts in India at 466 (Census 1991). The reason being, a number of districts split form their parent districts or parts of a district moved over to another district, over time. The World Bank data aggregates district boundaries back to their 1961 boundaries.

These 271 districts which account for over 85% of all agricultural land area and include all of the major rice and wheat producing states of India, are classified as either rice or wheat producing using the strategy discussed in the next section. Variables on agricultural investments include labor, tractors, bullocks and fertilizers used per hectare. Yield measures include quantities produced and area planted to all major food crops such as rice, wheat, maize, millet and sorghum. District controls for literacy and population are also derived from this data. The variables on weather are projections from a General Circulation Model (GCM) using station level meteorological data. These variables are informative of spatial variation in weather but not inter-temporal variation as inputs to the GCM such as soil characteristics are invariable over time.

Additional data on land use and land policy changes is obtained from Apportioned-

[7] This discussion follows from the data description prepared by Sanghvi, Kumar and McKinsey Jr , the findings of which are available in World Bank Technical Paper No. 402.

ICRISAT database. Data for all variables are reported at 1966 boundaries for consistent comparison. Here, I use the data on the number of farms in each district, according to size to show changes over time in land use and land policy changes.

2.4 Empirical Strategy

The primary identification is derived from exogenous agroclimatic conditions that directly governed the adoption of rice or wheat HYV crops. India is classified into 15 agroclimatic zones (Planning Commission Report 1989) at the state level, based on soil, temperature and precipitation attributes.[8] However, this dis-aggregation is not available at the level of the district.[9] Instead I use an alternative strategy. Evidence from several sources indicate that the initial adoption of the technology by farmers during the first wave of the Green Revolution (1965-70) was dependent on whether these farms previously cultivated rice or wheat. During the initial period when the comparative advantage of the wheat technology was not yet fully understood, due to the lack of a benchmark, the decision by farmers to adopt a new variety of seeds was based on whether the farm growing conditions were conducive towards the cultivation of one or the other crop. Any district which therefore adopted the new HYV wheat or rice seeds did so because they had been traditional cultivators of rice or

[8] These are the Western and Eastern Himalayan regions, Trans-Gangetic plains in the North; the Upper, Middle and Lower Gangetic plains in North-central and East; the Eastern, Western and Central plateau and hills region in the middle; the Southern plateau and hills region in the south; the Eastern and Western coast; Gujarat plains and hills; the Western dry region and the Islands region.

[9] There are a few other limitations in using Planning Commission classification for our analyses. First, a district may be mis-classified as rice or wheat producing since multiple districts in India have more than one climatic zones spanning it. Additionally, the differences between a few climatic zones may be subtle enough that at the district level, this difference will not translate to an abrupt change in crop cultivation practices. For example the middle gangetic plains and lower gangetic plains have very similar attributes.While the former receives annual rainfall between 1200-1500mm, the latter receives rainfall in the range of 1300-2100mm. Both these regions have hot temperatures and deltaic alluvial soil which is ideal for the cultivation of both rice and wheat. This makes it difficult to classify districts in these agro-climatic regions as either a predominantly rice or wheat producer. Second, some remote bordering districts may have agricultural practices which more resemble their across the border neighboring districts rather than their in-state districts. In this scenario, the district may learn from across the state farms and choose to adopt their cultivation practices. Classifying a district using the state data would introduce a bias in our estimation.

wheat respectively. It can thus be assumed that the adoption of HYV rice or wheat seeds in 1970 was highly correlated to agro-climatic conditions and is thus an exogenous determinant of the district being a major rice or wheat producer.

A district is classified as a rice(wheat) producer if the HYV adoption of rice(wheat) was greater than the HYV adoption of wheat(rice) in 1970.[10] Figure 1 maps out this classification at the end of the first wave and the end of the final wave of GR. A potential source of bias can arise if geographically close districts learn about the success of the wheat seeds and adopt by learning. I find that between 1970-85, about six districts switched from being a major producer of rice to a major producer of wheat demonstrating a case of technical spillover, as in Figure 1, panel B. I show in later sections, that the main findings are robust to the exclusion of these districts. This classification is further matched to data on district-wise, season-wise crop area production available from 1997-2010 for a subset of the districts in the sample as an additional check.[11] The strategy used here yields almost a perfect match with the classification of rice and wheat producers using official statistics. Both of these evidence suggest that there is limited mobility within a district with respect to the major crop cultivated before and after the technical change.

I first test the hypothesis that the increase in productivity lead to different outcomes in the rice and wheat districts. While the initial adoption was determined exogenously, further capital intensification over the entire period will be a function of increased productivity. Hence, empirically the following regression equation is estimated-

$$L_{ist} = \beta_1 + \beta_2 \cdot Y_{ist} + \gamma \cdot D_{ist} + \theta_1 \cdot S_{ist} + \theta_2 \cdot W_{ist} + \alpha_t + \epsilon_{ist} \tag{2.1}$$

Where L_{ist} is outcomes such as wage, employment and inequality in district i of state s

[10] At the state level, the states of Andhra Pradesh, Karnataka, Tamil Nadu, Bihar, Orissa and West Bengal are rice producers as over 95% of their districts are predominantly rice producing. Similarly, the states of Punjab, Haryana, Uttar Pradesh, Gujarat, Maharashtra and Rajasthan may be classified as wheat producers.

[11] This data was published by Ministry of Agriculture, Department of Economics and Statistics.

and year t, Y_{ist} is the log of yield per hectare of rice or wheat produced in district i and year t. D_{ist} is a vector of time-variant district controls such as the log of population density and literacy rate. S_{ist} constitutes time-invariant geographic controls like dummies for soil type and altitude which influence growing conditions. W_{ist} is a vector of controls for crop-season specific rainfall and temperature. α_t, represents time fixed effects and finally ϵ_{ist} is the error term clustered at the district level.

The independent variable is normalized and calculated as the yield of rice (wheat) per area of rice (wheat) cultivated and weighted by the share of rice (wheat) cultivation in the district in (2). The first term calculates the productivity increase at the intensive margin. It gives the direct increase in yields. The second term, calculates the increase in rice or wheat cropped areas as a proportion of cropped areas under all major food crops.[12]

$$Yield\,per\,hectare\,of\,rice(wheat) = \frac{Yield\,of\,rice(wheat)}{Area\,under\,rice(wheat)\,cultivation} \cdot \frac{Area\,under\,rice(wheat)\,cultivation}{Total\,Area\,under\,cultivation} \tag{2.2}$$

It must be noted, that since districts are of different sizes, the estimated effect from a percentage change in productivity at the intensive margin will be heterogeneous (and therefore biased) for larger districts if only a small number of farms adopt and others do not adopt at all. Again, larger districts may increase productivity at the extensive margin more rapidly than smaller districts. To account for these differences, the above measure is adopted. In all specifications, I refrain from using state level fixed effects as within-state variation of a district to change from being a rice to wheat producer is very limited. Additionally, since the changes in wages and employment are a result of systematic structural changes in rice and wheat districts, using a state fixed effect would absorb the very variation that I want to show. Nevertheless, to rule out concerns for a causal channel, in section 6, I examine

[12]This includes rice, wheat, maize, jowar (sorghum) and bajra (millet).The GR was expanded to cover all of these five major food crops and a few cash crops in India and the rest of the world. However, the progress of the coarser grains were less focused upon in India.

several concurrent factors that may potentially affect the crop yields as well as the outcome variables.

2.5 Main Results

2.5.1 Impact of Productivity on Employment and Wages

Variants of equation (1) are used to estimate the elasticity of labor market outcomes with respect to productivity in Table 2. The extent to which a percentage increase in rice or wheat productivity would impact these outcomes, may depend on the very nature of capital intensification in these regions. To capture these differences, I individually focus on rice and wheat districts in Panels B and C.

For an average district, while increase in wheat productivity lead to higher wages, an increase in rice productivity lead to greater employment generation. When disaggregated by rice and wheat regions, these differences are little more distinct. For instance, I find that for every percentage increase in yields, wages rose by 8% in rice districts and by 9% in wheat districts.[13] However, only rice districts observe a significant increase in employment.

In columns, 3 and 4, I take a closer look at the nature of employment generation in these regions as distinguished by measures of employment. The National Census definition states that agricultural laborers are employed on a daily wage basis. Cultivators or share croppers on the other hand, may be an employee/family worker paid in cash, kind, or in share of crop. That is to say, while agricultural laborers may suffer the cost of uncertainty of employment and therefore wages, cultivators often share in the risk of a poor harvest. The rapid industrialization of agricultural farming in wheat districts during the Green Revolution

[13] Here wages have been deflated by the price of staple crops. Average calories per kilogram of rice and wheat are 1508 and 3400 respectively. Average daily caloric requirement is 1900 for an Indian women and 2100 for Indian men. Mean of the dependent variable indicates that the real wage is 3.2 kilos of rice and 4.3 kilos of wheat, indicating that the wage difference is much higher than what appears at face value.

lead to small farms and cultivators being subsumed into larger farms. As such, while wheat districts saw some positive impact on employment generation of agricultural workers, the impact on employment generation on aggregate and on cultivators was insignificant.

In contrast, for every percentage increase in yield, rise districts experienced a significant increase in employment of both cultivators and agricultural workers. Even with the caveat that demand for agricultural workers generally spike at the end of harvest times, which is also when most agricultural surveys are conducted, a significant increase in cultivator employment indicates that the increase in aggregate employment of rice districts was not driven by seasonal changes.

I posit that the employment generation in rice districts is owed to the agroecological instability of rice cultivation which requires constant manual monitoring. For example, Appendix Table B.1 draws on evidence from the National Industrial Classification of jobs that form the basis of the Employment-Unemployment surveys to depict the emergence of several job classifications, centered around the rice production process, which did not exist prior to 1970. To some extent this might simply indicate advances in statistical data collection. However, the more likely reason for the introduction of the new job classifications is to reflect their emergence in real time. As such, it bolsters the hypothesis that the cultivation of rice involves relatively labor-intensive tasks.

2.5.2 Impact of Productivity on Welfare

To the extent that the gains of productivity created differences in employment and wages in rice and wheat districts that stimulated the growth of labor enhancing tasks in rice districts and the adoption of labor enhancing techniques in wheat districts, it motivates a discussion on welfare. If increased productivity lead to greater wage increases in wheat districts, while displacing cultivators, it is possible that it lead to greater inequality in wheat districts. On the other hand, it may be expected that the creation of new tasks in rice regions would lower inequality.

Table 3 computes elasticities for the change in various welfare measures with respect to productivity for each of these regions. In terms of inequality, I find that both regions saw a decline in the rural gini coefficient. However, rice districts saw a much greater decline. Wheat regions however, became richer on average during this period as shown by the decline in rural head count ratio. This indicates that while the additional employment in rice districts contributed to bridging some of the existing income gaps, it may have lead to greater entrenchment in agriculture.

2.6 Drivers of Growth in Rice versus Wheat districts

2.6.1 Differences in forms of Investments and Factor Accumulation

In discussing the factors that lead to the differential impact of productivity on labor market outcomes in rice and wheat regions, I begin with the primary differences in agricultural investments and capital accumulation in these regions that followed in the post-GR period. Wherein these factors pose a threat to the causal results above, I use robustness checks to rule out endogeneity concerns.

The income shock following the GR raised incomes over subsistence levels and enabled investments in capital goods. I use measures such as number of tractors and oil pumps and net area under irrigation to indicate capital intensification. Figure 3 demonstrates that investments in fixed capital goods took off right around 1965-70 which coincides with the first wave of GR. Moreover, the increasing rate of capital intensification, indicates an increasing returns to these labor saving techniques. Post 1970, however, there is an uptick in the level of investments in wheat districts. Net area under irrigation, plausibly due to development of new irrigation projects also rose in wheat districts relative to rice districts.

In contrast, rice districts invested more in working capital goods such as fertilizers, pesticides and chemical weeders. Figure 4 shows a similar pattern of uptake of investments in the second wave of GR as Figure 3, except with the rice districts investing at a greater rate. The panels indicate time series of usage of phosphorus, nitrogen and potassium, the three primary nutrients for plant growth. Nitrogen and potassium is used as a fertilizer while phosphorus is used a chemical weeding method. These investments were complementary to agricultural practices that are relatively more labor-intensive in rice districts.

Insofar as investments in labor enhancing techniques are less expensive, the reason for this divergence might very well stem from rice districts being poorer prior to the introduction of the new technical change, which will bias the OLS coefficients on employment in rice districts upwards. A common strategy to account for these omitted variables is to use neighboring districts, that is districts that share a common border and will have very similar economic conditions.

Between 1970-85, 20 districts switched from being a predominantly wheat to predominantly rice cultivator and 6 switched from being a predominantly rice to wheat cultivator. [14] Let us say that a district adopted wheat in 1970 solely based on geographical and not economic conditions, when switching to rice it should exhibit structural changes similar to a rice districts. As shown in Table A.2, districts that switched to rice saw a significant increase in wages and no impact on employment. On the other hand, districts that switched to wheat had a significant increase in wages and a significant decline across all measures of employment.

[14] While in general, the propensity of a district to switch from a major producer of a given crop is low, it is still possible as there is more than a single requirement for successfully cultivating either rice or wheat. If a wheat district in the alluvial soil region has irrigation, it can compensate for the lack of standing water for rice and switch. Similarly, a lowland district with a comparatively drier climate that previously cultivated rice will be able to switch to wheat production.

2.6.2 Differences in Patterns of Adoption

The increase in aggregate productivity of rice and wheat can be expressed in terms of changes at the intensive and extensive margins. Specifically, aggregate productivity may increase if either the yield per hectare of crop planted increases or the proportional area brought under cultivation of the crop increases. The intensive margin of productivity increase is calculated as $\frac{Yield\,of\,rice(wheat)}{Area\,under\,rice(wheat)\,cultivation}$ and the extensive margin is calculated as $\frac{Area\,under\,rice(wheat)\,cultivation}{Total\,area\,under\,cultivation}$. Aggregate productivity is calculated as a product of these two terms.

Empirically we test if the district rate of adoption given as the proportion of rice or wheat HYV cultivation is associated with the above measures of productivity. Table 4 presents these results after controlling for district level characteristics. In general, while productivity gains for rice districts were realized at the intensive margin, for wheat districts the majority of this realization was at the extensive margins. Columns 1 and 2 show the direct gain of adopting HYV technology. A 1 percent increase in proportion of land planted to HYV rice is associated with a 16 percent increase in rice yield. Similarly, a 1 percent increase in proportion of land planted to HYV wheat is associated with a 7 percent increase in yield. On the other hand, in Columns 3 and 4, indirect gains, or gains along the extensive margins are 22 percent and 33 percent for rice and wheat respectively.

The differences in the extensive margin can be further explained by several factors. First and foremost, is the change in land use patterns. The greater gains accruing to wheat districts at the extensive margin, may either be due to an increase in the area under wheat cultivation as a proportion of all land under cultivation (change in taste and preference effect) or new land being brought under cultivation (land use effect). In Table 5, Panel A, I show that both of these factors contributed to the extensive margin gains. Between pre and post adoption of the new technology, wheat as a proportion of total crops cultivated increased significantly in wheat over rice districts. The t-stat of 8, quantifies the difference in extensive margin due to the taste and preference effect. Furthermore, there was a significant increase in the total area under cultivation of wheat as well. I calculate the t-stat for this difference to around 3.67.

The significant changes in taste and preference in wheat districts can be attributed to the underlying diet patterns in these regions. While both rice and wheat are staple crops, the contribution of wheat to the overall diet of people living in wheat districts is comparatively lower than the contribution of rice to the diet of people in rice districts. Wheat consumption is often supplemented with the consumption of other cheaper and minor food grains such as maize, jowar (sorghum) and bajra (millet). With the introduction of high yielding seeds wheat seeds, and better yields it became less expensive to substitute away from these minor grains.

While it is not possible to deduce if land formerly allocated to other minor other crops were being substituted for wheat production or the new (possibly less fertile) land was being allocated to wheat production or a combination of both, this exercise is indicative of much larger changes in extensive margins for wheat than for rice. Even if it is a combination of both of these factors, given the higher economic rents of cultivating more fertile land, it can assumed that as more and more new land is brought under cultivation, the average gains along the intensive margin would decline as well. This further explains the differences in intensive margins to some extent. Since, rice districts were already operating at maximum capacity in terms of area under rice cultivation or the ratio of rice cultivation to all other crops, the only direction for improvement was along the intensive margin.

In the following section, I argue that the above findings are robust to differences in other confounding conditions.

2.6.3 Other potential drivers of Agricultural Investments

Patterns of Land Use and Land Policy Changes

A potential candidate for differential mechanization could be differences in farm sizes or operational land holdings. The period of study in this paper coincides with several land reform policies that were undertaken in India. These policies, aimed at redistribution of land from large landholders to small holders, were undertaken at the state level. To the extent that such a policy may benefit rice districts more than wheat districts, we would observe more benefits of productivity increase accruing to laborers, in the form of increased employment. In Table 5, Panel B, using ICRISAT data on a states that are predominantly rice and wheat producing states, I examine changes in land holdings in the pre and post period of GR introduction

I show that over time, distribution of farm sizes in both rice and wheat districts converged to medium sized holdings, indicating overall but not differential changes in land consolidation. This is not entirely surprising, given the findings of Besley and Burgess (2000) where they argue that the growth and poverty effects of land reforms was via improved tenurial contracts rather than via redistribution of land. While, changes in tenurial arrangements may have an independent effect on wages and employment, they would be a threat to the empirical strategy used here, only if accounting for them absorbs all variation and we end up with an insignificant effect of productivity. Hence, I attempt to control for changes in tenure contract.

I use two different strategies. First, tenurial policy changes were legislative measures generally undertaken at the state level. This indicates that within a state, patterns of tenure arrangement should remain unchanged. Controlling for state fixed effects, I show that the results remain unchanged (Table A.1). Second, if these changes were driven by a few large pro-worker rice states and pro-industry wheat states, the estimates would not be reliable. I drop the most pro-worker state of West Bengal and the most pro-industry state of Gujarat in Table A.3 and show that the previous results hold up.

2.7 Instrumental Variables Strategy: Agroclimatic Conditions

Inasmuch that the differences in regional factors determine district level productivity and therefore labor market outcomes beyond the ways discussed above, I attempt to tackle them using an instrumental variables approach in Table 6. Variations in agroclimatic conditions surrounding the cultivation of either rice or wheat are highly associated with realised agricultural yields during GR. Data from the Global Agro-Economic Zones (GAEZ) database, computes soil suitability for different regions (at the state level) of India.[15] The rice and wheat soil suitability indices in the instrument set Z_s, quantifies the extent to which prevailing soil conditions are conducive to the growth of the particular crop under given soil management conditions. Assuming that pre-1965 soil suitability conditions, computed at pre-1965 soil management conditions, while predicting the post-1965 productivity measure would be orthogonal to other confounders in this model, I use values of the indices at low-input and rain-fed conditions.

The estimated IV 2sls model is as follows-

$$Y_{ist} = b_1 + b_2 \cdot Z_s + \pi \cdot X_{ist} + \rho \cdot W_{ist} + \kappa_t + \nu_{ist} \tag{2.3}$$

$$L_{ist} = \gamma_1 + \gamma_2 \cdot \hat{Y}_{ist} + \eta \cdot X_{ist} + \lambda \cdot W_{ist} + \mu_t + \omega_{ist} \tag{2.4}$$

where, Z_s is a set of two instruments of the index of agroecological suitability for rice

[15] This approach has been used in Nunn and Qian (2011)

and wheat production in state s and $\hat{Y}_{ist}$ is predicted rice and wheat yields. In the second stage, predicted yields of rice and wheat in state s are used to capture the effect on labor outcomes L_{ist}. The rest of the controls remain the same.

Panel A, Table 6 indicates a close similarity of these results with those discussed earlier - I observe wage increases in both regions. Rice districts observe a significant increase in employment propelled by agricultural workers. An increase in wheat productivity, on the other hand leads to large declines in employment of agricultural laborers. I also observe an increase in employment of cultivators in both regions, however the magnitude is larger for rice districts relative to wheat districts.

2.8 Conclusion

The relationship between agricultural productivity and labor market outcomes is not a straightforward one. This paper identifies an important wrinkle in this relationship wherein different crop regions may experience different forms of the same technological advancement. These differences are rooted in their environmental factor endowments. Environmental factor endowments play a role in how the technology is adopted, thereby having important consequences for labor market and welfare outcomes.

Distinguishing between the experiences of different crop regions may be further used to explain differences in economic outcomes that go beyond the ones studied here, both in the context of a macro and a micro perspective. For instance, a technological advancement that relies heavily on labor enhancing methods, may have positive ramifications for female employment and conflict reduction. As long as there are positive returns associated with the additional labor, we can expect allocation of labor from domestic and conflict sectors. On the other hand, there may be macro implications as well, that is beyond the scope of this study. A labor saving technology may precipitate the migration of labor to other sectors or regions, encourage greater investments in infrastructure, and improve overall quality of life.

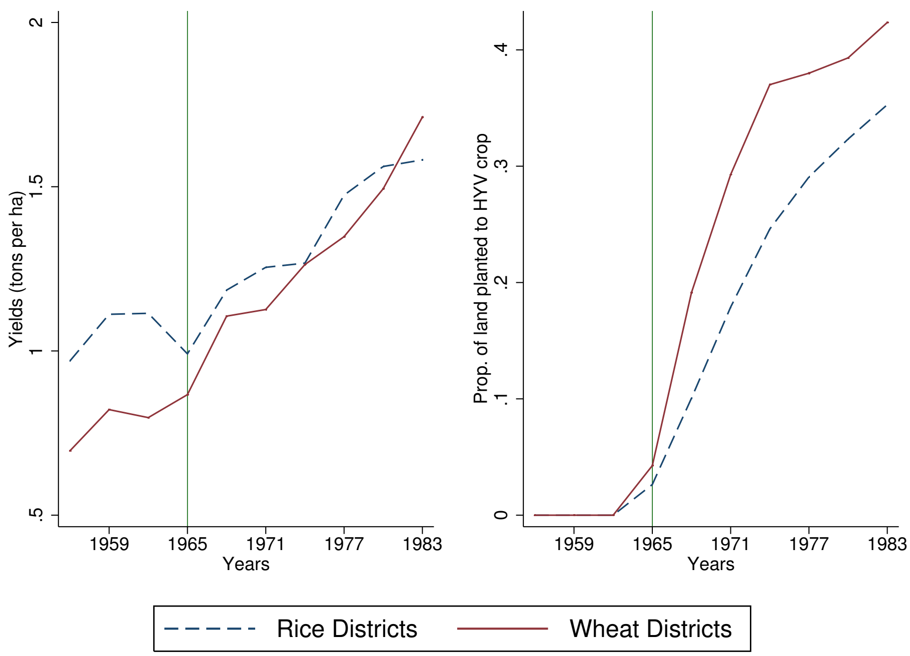
An important caveat of this study is the inability to observe the impact of productivity after the returns from investments have slowed down. It may be expected that while we observe welfare improvement in the short run for regions that adopted a labor enhancing technology, once decreasing returns to labor set in, these impacts will disappear. As such, this study paves the way for future research to study the long term impact of a productivity increase after the immediate gains have slowed down.

Much of the developing world is still largely agrarian. Of whom, a large number are at stages of technological growth that predate the outcomes observed here. Lessons drawn from the structural changes during the Green Revolution in India may go a long way in predicting the diffusion of the technology in these regions as well as caution against their disruptive impact in some regions. For instance, while I do observe increase in agricultural employment in rice regions which is welfare improving from a development perspective, I also find wheat regions becoming richer over time, which is desirable from a growth perspective. Policymakers considering the introduction of potential growth inducing policies have to often weigh the difficult option of sacrificing growth over development and may thus need to be aware of the differences in potential impact.

Moreover, they may need to be aware of the impacts of technological leapfrogging for regions that are unequipped to absorb the impact of such a shock. In such cases, technology may displace labor too rapidly, leading to "uneven growth". To safeguard against such disruption, a positive income shock may have to be accompanied by concurrent employment generation in the affected sectors to produce the desired results for development.

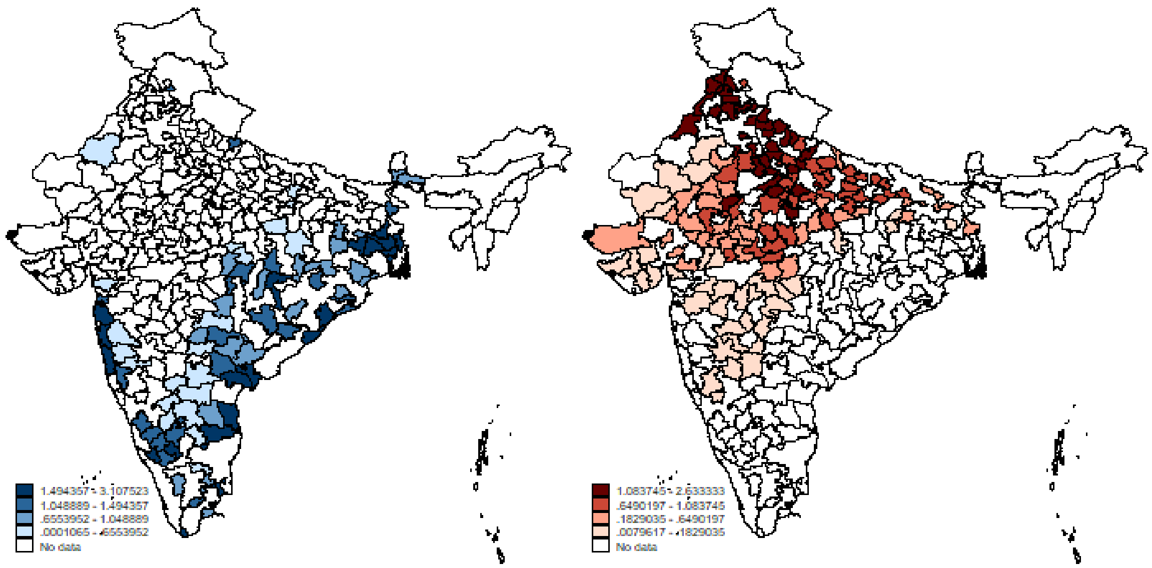
Figures

Figure 2.1: Growth in yields and rate of adoption and post Green Revolution



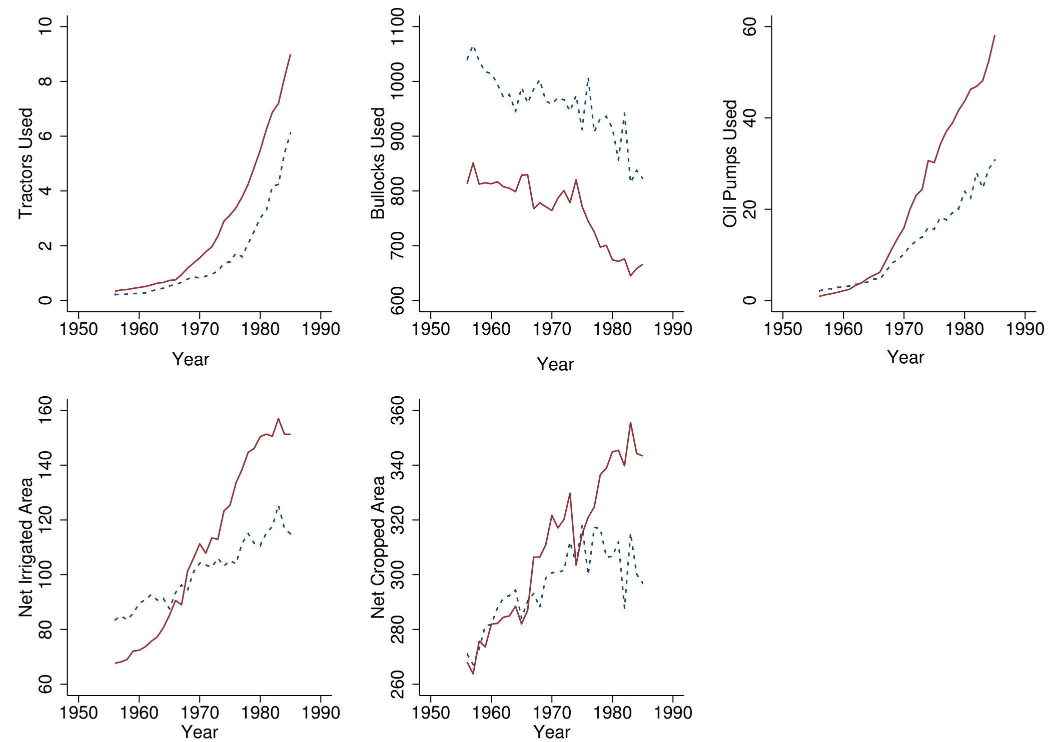
Source: Author's calculations from World Bank India Agriculture and Climate Dataset.

Figure 2.2: Variation in Yield in Rice and Wheat districts



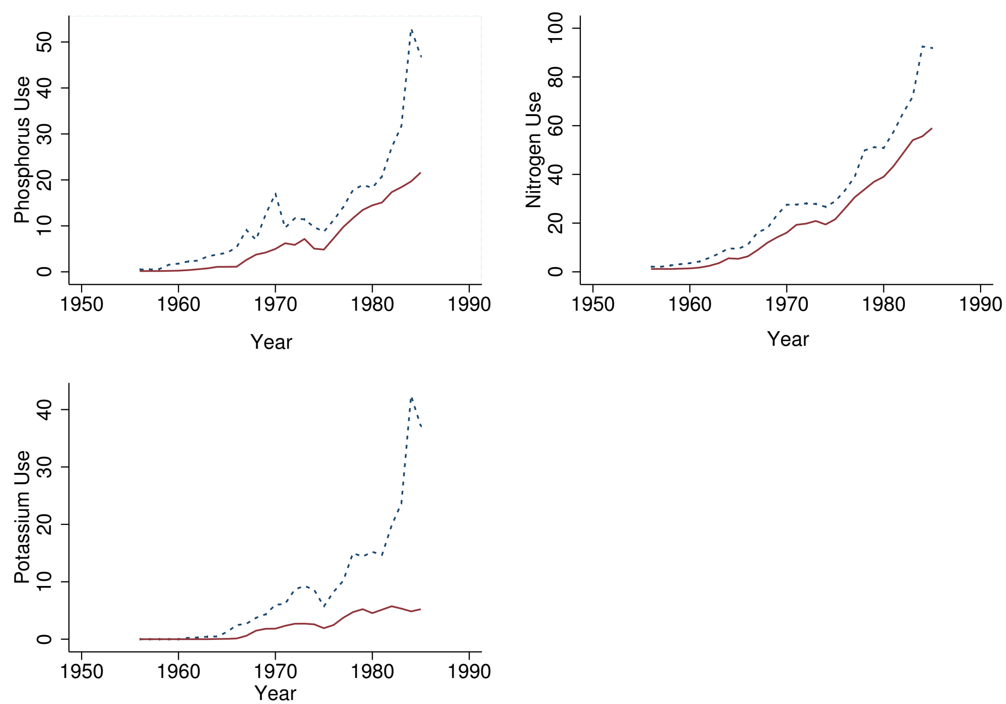
Note: The figures indicate variation in yields in rice (left) and wheat(right) regions in tons/ha at 1985 levels.

Figure 2.3: Capital Intensiveness of the new technical change



Note: Blue Dashed lines represent rice districts and Red solid lines represent wheat districts. Tractors, bullocks and oil pumps (in units per hectare) are averaged over district years separately for rice and wheat districts. Net Irrigated Area and Net cropped Area are plotted from variables directly available from World Bank India Agriculture and Climate Dataset.

Figure 2.4: Labor Intensiveness of the new technical change



Note: Blue Dashed lines represent rice districts and Red solid lines represent wheat districts. Tractors, bullocks and oil pumps (in units per hectare) are averaged over district years separately for rice and wheat districts. Fertilizer use is in terms of tons per hectare. Data is from World Bank India Agriculture and Climate Dataset.

Tables

Table 2.1: Summary Statistics

	Rice Districts			Wheat Districts		
	Observations	Mean	Std. Dev	Observations	Mean	Std.Dev
District Characteristics						
Yield of Rice/Wheat(in tons/ha)	3210	1.25	(0.73)	4920	1.12	(0.62)
Proportion of land planted to HYV rice/wheat	3210	0.24	(0.33)	4920	0.36	(0.41)
Price of rice/wheat (in Rs/quintal)	3210	4.61	(0.58)	4920	4.50	(0.59)
Log (Population Density)	2674	0.63	(0.83)	4920	0.62	(0.71)
Literacy Rate	3210	0.33	(0.11)	4920	0.30	(0.10)
Geographic Conditions						
Altitude (in Kilometers)	3210	0.31	(0.13)	4920	0.38	(0.14)
Kharif/Rabi[1] Season Temperature	3210	0.27	(0.01)	4920	0.19	(0.03)
Kharif/Rabi Season Precipitation	3210	0.22	(0.22)	4920	0.01	(0.01)

Soil Characteristics	Soil Types				
	Alluvial	Med. Black	Red	Saline	Laterite
% of Wheat districts with soil type	54.9	17.1	5.5	14.0	1.2
% of Rice districts with soil type	26.2	11.2	30.8	0	18.7
	Topsoil Depth (in cms)				
	0-25	25-50	50-100	100-300	>300
% of Wheat districts with topsoil type	0.04	0.05	0.28	0.08	0.54
% of Rice districts with topsoil type	0.04	0.23	0.27	0.16	0.29

Note: The basis of classification of rice and wheat districts is the level of rice versus wheat technical adoption. The Kharif season (July to October) corresponds to the sowing/growing season of rice, while the Rabi season corresponds to the sowing/growing season of Wheat (November to February). Temperature (in Celsius) and Precipitation variables have been scaled by a factor of 1/100 and 1/1000 repectively.The number of observations correspond to a district-year observation in the sample.

Table 1.2: Effect of Productivity Increase on Employment and Wages

	Dependent Variables			
	Log Daily Wages	Log Aggregate Employment^	Log # of Agricultural Laborers	Log # of Cultivators
	(1)	(2)	(3)	(4)
Panel A: All Districts				
Log Rice Yield (tons/ha)	0.0021	0.074***	0.150***	0.060***
	(0.011)	(0.015)	(0.019)	(0.018)
Log Wheat Yield (tons/ha)	0.031***	-0.0059	-0.022	0.012
	(0.011)	(0.013)	(0.017)	(0.016)
Mean of Dependent Variable	0.039	286,475	932	334
N	5667	5667	5667	5667
Adjusted R^2	0.571	0.390	0.528	0.337
Panel B: Sample restricted to Rice Districts				
Log Rice Yield (tons/ha)	0.078***	0.180***	0.240***	0.150**
	(0.029)	(0.058)	(0.062)	(0.070)
Mean of Dependent Variable	0.032	364,723	446	1101
N	2607	2607	2607	2607
Adjusted R^2	0.382	0.487	0.556	0.318
Panel C: Sample restricted to Wheat Districts				
Log Wheat Yield(tons/ha)	0.085***	0.039	0.140**	0.022
	(0.021)	(0.025)	(0.058)	(0.024)
Mean of Dependent Variable	0.043	235,422	261	822
N	4096	4096	4096	4096
Adjusted R^2	0.487	0.505	0.552	0.529
District controls	Y	Y	Y	Y
Geographic controls	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y

Note: Standard errors clustered at district level are shown in parenthesis. *p<.10 **p<.05 ***p<.001. All regressions include district level controls, soil controls, climate controls and time fixed effects from OLS regression specification. Aggregate Employment is the sum of cultivators and agricultural laborers weighted by the number of days worked in the state by farm workers. The Census defines agricultural laborers as those who work for wages on private/Government owned farms. Cultivators are defined as a person engaged as employer/family worker paid in cash or kind or share of crop.

Table 1.3: Effect of Productivity Increase on Welfare measures

Dependent Variables	Rice Districts (1)	Wheat Districts (2)
ΔRural Gini Coefficient	-0.015***	-0.007**
(1972-1987)	(0.004)	(0.003)
ΔRural Head Count Ratio	0.617	-0.947**
(1972-1987)	(1.091)	(0.460)
ΔRural Per capita Expenditure	0.014	0.082
(1972-1987)	(0.544)	(0.489)
District Controls	Y	Y
Time Fixed Effects	Y	Y
Geographic Controls	Y	Y

Note: Data is from Banerjee and Iyer(2005). Standard errors, in parenthesis are clustered at the district level. *p<.10 **p<.05 ***p<.001 are significance levels. Dependent variables are listed across the top of the panels. The Gini coefficient is scaled by a factor of 100 to obtain results in percentage terms. All regressions include the full list of controls. See notes under Tables 2 for a list of district level controls included in the above regressions.

Table 1.4: How much can the adoption of HYV explain the Increase in Productivity?

	Intensive Margin		Extensive Margin		Aggregate	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Rice districts						
Log Rate of Rice Adoption	0.155***		0.221***		0.376***	
(in tons/ha)	(0.037)		(0.052)		(0.065)	
N	2041		2041		2041	
Adjusted R^2	0.55		0.67		0.57	
Panel B: Wheat districts						
Log Rate of Wheat Adoption		0.067***		0.325***		0.400***
(in tons/ha)		(0.018)		(0.067)		(0.065)
N		3216		3216		3218
Adjusted R^2		0.54		0.66		0.73
District Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls[1]	Yes	Yes	Yes	Yes	Yes	Yes
Time F.E	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors, in parenthesis are clustered at district level. *p<.10 **p<.05 ***p<.001 are significance levels. Dependent variables are listed across the top of the panels. The intensive margin of productivity increase is calculated as yield of crop per area of the crop cultivated. The extensive margin is calculated as the area of the crop per total area under cultivation. The aggregate productivity is calculated as a product of the intensive and extensive margins. The independent variable is the proportion of land planted to HYV crop for that district. Geographic controls include temperature,precipitation and soil dummies. Temperature (scaled by a factor of 1/100) and precipitation (scaled by a factor of 1/1000) variables have been averaged over the Kharif (July to October) and Rabi (November to March) growing seasons for rice and wheat respectively. Other controls include literacy rate and altitude (in kilometers).

Table 1.5: Pattern of Changes in Land Use and Land Holdings

		Rice districts			Wheat districts			
		Pre 1970	Post 1970	Difference	Pre 1970	Post 1970	Difference	t-stat
Panel A: Land Use and Crop substitution								
(Mean) Area under Rice or Wheat (ha)		199	215	16	73	111	38	4.93
(Mean) Total Area under cultivation (ha)		285	306	21	285	331	46	3.67
Ratio		0.7	0.8		0.25	0.3		
t-stat of Difference in Mean ratios		8						
Panel B: Land Consolidation			Pre	Post		Pre	Post	
Years used: 1970 (pre), 1980 (post)								
Mean # of Small farms	1-2 ha		67.34	72.94		27.48	51.42	
Mean # of Medium farms	3-10 ha		33.07	23.99		23.57	17.70	
Mean # of Large farms	>10 ha		9.32	5.15		6.79	3.19	

Note: Data in Panel B is from ICRISAT villages.

Table 1.6: IV estimates of the effect of Productivity Increase on Employment and Wages

	Dependent Variables			
	Log Daily Wages	Log Aggregate Employment^	Log # of Agricultural Laborers	Log # of Cultivators
	(1)	(2)	(3)	(4)
Panel A: Second Stage IV results				
Log Rice Yield (tons/ha)	0.140**	0.043	0.206**	0.101*
	(0.059)	(0.045)	(0.080)	(0.039)
Log Wheat Yield(tons/ha)	0.120***	-0.007	-0.070	0.089**
	(0.037)	(0.030)	(0.050)	(0.038)
N	5667	5667	5667	5667
R^2	0.260	0.233	0.403	0.193
Panel B: First Stage Results$^{\zeta}$	Log(Rice Yield)		Log(Wheat Yield)	
	(5)		(6)	
Crop suitability of Rice	0.221***		-0.451***	
	(0.008)		(0.008)	
Crop suitability of Wheat	-0.016**		0.182***	
	(0.005)		(0.005)	
N	5667			
Cragg Donald Wald F Stat$^{\phi}$	319.81			

Note: Notes- Standard errors clustered at district level are shown in parenthesis. *p<.10 **p<.05 ***p<.001. All regressions include district level controls, soil controls, climate controls and time fixed effects from OLS regression specification.Dependent variables for the first stage are endogenous variables of rice and wheat yield for a given district. Independent variables in the first stage is the set of instruments of agroecological index.Cragg Donald Wald F statistic is the test for weak instrument, where is used to indicate a value that exceeds the Stock and Yogo(2005)critical value at 10% maximal IV size. Âggregate Employment is the sum of cultivators and agricultural laborers weighted by the number of days worked in the state by farm workers. The Census defines agricultural laborers as those who work for wages on private/Government owned farms. Cultivators are defined as a person engaged as employer/family worker paid in cash or kind or share of crop.

Chapter 3

Family Ties and Coping: Evidence from Idiosyncratic and Aggregate shocks in Indonesia

3.1 Introduction

Households in low income countries often rely on informal strategies to smooth unpredictable shocks to income within the household (for a complete surveys see Dercon 2002). Outside the household unit, a significant strand of literature focuses on the role of transfers from close social links formed in village communities (Townsend 1994; Rosenzweig 1988), ethnic groups (Grimard 1997) or extended family (Altonji, 1992; Angelucci *et al.* 2010) as a viable coping mechanism. While most of these studies focus on the degree of resource sharing provided by the links, little is known about their formation as a response strategy. To the extent that these social links are involved in some form of resource pooling, their efficacy as well as evolution may differ by the type of shock considered.

Income may be subject to aggregate risks such as drought, pest, flood, etc. or idiosyncratic risks such as death, illness, or job loss. In general, while an aggregate shock affects everyone in the area, an idiosyncratic shock will only affect a single individual. This has a direct consequence for the degree of reversibility in adopted coping strategies in case of either type of shock. For example Kinsey *et al.* (1998) show that households respond to idiosyncratic health shocks by increasing their labor supply, depleting assets or diversifying into other occupations. On the other hand, in response to aggregate shocks like a drought, they are more likely to migrate away, allocate resources in favor of boys, increase the hazard into child marriage and even lower realized fertility (Kleemans, 2015; Corno *et al.*, 2020; McKenzie, 2003).

This study builds on the above two strands of literature to assess the role and evolution of extended family network to two of the most commonly occurring shocks – an idiosyncratic health shock and an aggregate shock from weather. Given that link formation is costly, while households may resort to innovatory measures of utilizing their existing links under an aggregate shock; a health shock encourages the formation of new links.[1] Using detailed longitudinal data from the third, fourth and fifth waves of the Indonesian Family Life Survey (IFLS), which tracks original household members as they "split-off" forming new households, we construct networks resembling a binary tree type structure as described in Ambrus *et al.* (2010). This structure, while the most typical and expansive of all, is infrequently observed due to survey data limitations.

The detailed nature of the IFLS allows us to observe various dimensions of risk sharing behavior across all links. In order to evaluate transmission via links, we focus on a particular instrument- private money transfers in the form of borrowing or lending.[2] Households with more links engage in more transactions to share risk, as do those connected to newer, wealth-

[1] For example (Smith *et al.*, 2002) show that while urban households gained new members, rural households lost household members during the Asian Financial crisis, as a strategy to diversify risk.

[2] Several studies have focused on instruments like livestock assets depletion, gifts, remittances etc. (Kazianga and Udry, 2004; Janzen *et al.*, 2018)). While these mechanisms are prominent in the context of Africa, the primary economic activities in Indonesia are small holder farming, cash cropping and exports.

ier households. Furthermore, if spending down wealth is a viable consumption smoothing tactic as demonstrated in the literature (Janzen *et al.*, 2018), it can be abated to the extent that family links are efficient in sharing the income shock. Indeed, we show that presence of split families in the network results in less sale of assets.

Nevertheless, a priori it is often unclear if the presence of a rich network is by itself beneficial. We show that the composition and characteristics of the network, such as split-offs located in urban areas, presence of more married families, and family size of these married family links are likely to influence the capacity of the network to assist. Furthermore, the optimal number of links in the network is an issue of much debate. For example, Murgai *et al.* (2002) states that the optimal group size is subject to the association cost, which rises with the number of members in the network, and extraction cost which rises with the expected level of insurance. The construction of the family structure, to the extent that it is observed within the IFLS, generates some insights into the efficiency gains from inclusion of an additional member. In general, risk sharing is increasing in the number of split-off families that a household is connected to.

Our main results indicate that while every additional link is effective in smoothing a weather shock, they are distinctive in their inability to smooth a health shock. There are several plausible reasons for this. First, an aggregate shock is likely to more immediate so that the benefits from resource pooling are likely to compensate for income loss. Second, aggregate shocks are temporary which increases the probability of repayment of informal transfers. Moreover, it is plausible that income loss associated with health shocks are considerably higher which further reduces full repayment potential. Lastly, given that an aggregate shock to a household does not happen in isolation, the incentive to resource pool is more pervasive.

We derive predictions for the potential of the network to engage in risk sharing under varying levels of shock exposure to network members. This finding is line with Fafchamps and Lund (2003) where they find that informal lending responds more to network shocks than individual household member shocks. Indeed, we find that in some cases the ability of the household to smooth their consumption through the network is more dependent on the ability of the network to engage.

Our findings highlight the importance of including the network of households linked to the reference household when looking at monetary transfers between households. Leaving out network shocks, especially in the case of aggregate shocks like drought, will constitute omitted variable bias leading to a bias towards zero on of the level of borrowing. Incorporating the network characteristics of a household also allows for a more complete explanation of the factors of borrowing and lending, decreasing the amount of variation in the residual.

Our results indicate the household decision to split off in response to a shock may be determined by expected income loss and subsequent expectations regarding resource sharing. This line of inquiry is closely related to the one adopted by Foster and Rosenzweig (2002) that relies on an ever-evolving family structure. They find that economic growth may directly affect the composition and number of households in a panel survey by increasing within household inequality in schooling, marriage and riskiness. In the Indonesian data, households are observed to merge together under aggreagate shocks when the necessity to resource pool is immediate, inexpensive and has economies of scale. On the other hand, households are more likely to split and migrate to other regions when income loss is expected to be higher than any economies of scale from co-residence.

Sections 2 and 3 describe the data and methodology respectively. Section 4 is a discussion of the main results. Finally Section 5 concludes.

3.2 Data Description

3.2.1 Aggregate and Idiosyncratic Shocks

Our primary source of data is the Indonesian Family Life Survey (IFLS), which is a longitudinal panel representing 83% of the country's population of Indonesia. We use data on household characteristics and informal transfers from the three most recent waves- the third wave (IFLS3) conducted in 2000-2001, the fourth wave (IFLS4) conducted in 2007-2008 and the fifth wave (IFLS5) conducted in 2014-2015. The survey includes over 7,244 original households from the first wave and 8,697 new households subsequently split-off from over the next 4 waves. These split-off households form the building block of the network measures described in the following section.

We combine this data with two different measures of shocks that affect household income. First is an idiosyncratic health shock, from the IFLS module on reported health issues.[3] We construct a dummy for being affected by any illness if the respondent experienced a chronic, geriatric or psychological distress. For a subset of this sample, we construct a dummy for being affected by illnesses that limit mobility if respondents reported they lost days of work due to it.

The second shock is an aggregate temperature shock. Recent studies (Donaldson and Storeygard 2016; Nakamura *et al.* 2020) have placed emphasis on the minimization of measurement error when using satellite data to calculate drought as opposed to survey reported drought or even weather station data. Remotely sensed land surface temperature data collected by MODIS (Moderate Resolution Imaging Spectroradiometer) and computed at the

[3] The IFLS questionnaire reports a range of health related problems which we subdivided into four main categories- Serious chronic conditions pertain to heart, liver, kidney, stomach or cancer related disorders. Respiratory illnesses refer to tuberculosis, asthma or lung disorder. Old age related conditions are diabetes, arthritis, cholesterol and memory related disease and psychological stress includes hypertension, emotional, nervous or psychiatric problems.

administrative level of a district is matched to households in the IFLS.[4] Average temperatures over the dry agricultural season (August to October) are calculated to denote severity of an aggregate shock.

3.2.2 Formation of Family Network in IFLS

Network Link Formation

The description about how family links are formed and surveyed between rounds necessitates a distinction between split-off family members and non co-resident family members. A split-off family is formed when a member leaves a parent household to form her own household, as the two children in Figure 1 do in the second period. These members are then tracked and interviewed using the same questionnaire as the original households and may include new members such a spouses. This indicates that while a split-off family is essentially treated as another respondent, non co-resident members, who are simply used as a reference in various questions, may exist beyond the respondents surveyed in the IFLS.

Figure 1 is a schematic representation of this categorization.[5] For the purposes of this paper, two household are linked if at least one individual is observed living in both housesholds in two periods. Due to the fact the IFLS has been going on for over 20 years, some household are seen indirectly connected, with a household in between. Direct links are formed between members who directly cohabited at some point. For example, links between father and son would be denoted as first-degree links. Second degree links are formed between members who share a family member who previously lived together. This would include links formed between son and grandfather or son and uncle, both of whom share father as a

[4] Appendix A.1 describes the MODIS satellite data in more detail. Here we primarily focus on the general link between weather and productivity. For comparison across studies such as Harrari and Ferrara (2013) that exclusively focus on an agricultural mechanism of weather shocks, we develop drought measures there.

[5] LaFave and Thomas (2017) schematically depict the formation of split-offs over each of the waves. We recreate their work in Figure 1 to delineate the observable parts of the network, which are used to form links and the unobservable parts.

shared member.

Since we have much more information on the household characteristics of split-offs than non co-residents, we define a network link as existing between households when they are connected to split off families. All split-off links present in the IFLS universe have been considered when constructing the network, beginning with the second wave, which was conducted in 1997. Spread over 20 years, we are able to observe families up to the third generation.

To add some figures to the perspective, 7724 households were interviewed in the first wave, of which 878 split-off in IFLS 2. A further 3,913 split-offs occured in IFLS3, 2,855 in IFLS 4, and 1,124 in IFLS5. This gives a total of 8,770 split-off family links in the IFLS universe, with every household having 1.06 links on average.

While these split-off households tracked by the IFLS are valuable, they can never capture the full extent of all connections that households have. Of the 22,256 households in our sample, 5,966 have no observed household links. It would be incorrect to assume that these households have no network of relatives from which to borrow. If we were to include those households without any considerations it would incorrectly bias the estimates on any included network parameters. These households do have significant differences in their construction. Non-connected households have older heads and more members, but both groups cover the same range of values.[6]

Of course, we do not perfectly observe all split-offs between each round either. This is because the split-offs may appear in the survey conditional on being located to one of the 13 out of 27 provinces, and not refusing to being tracked and interviewed. This would indicate that our constructed measure is an under-estimation on the number of counts and spread of the network. However both of these are minor concerns, as 83% of the population lives

[6]The differences between an isolated and connected household is not shown here but available upon request.

in these 13 provinces, and the re-contact rate at an average of over 92% is as high as many longitudinal waves in developed countries.

Our network link formation using a rich panel data has two advantages. First, we can observe binary tree type family models as stated in Ambrus *et al.* (2010). The binary tree model is the most realistic of all networks, however it is seldom represented in survey data. Second, we can assign shock data to all members in the network, which is an improvement over the approach adopted by Fafchamps and Lund (2003) where only shocks to core members in the network have been considered.

Household Characteristics of split-off links

From the network information and household data we construct measures of a household's network. First, these measures provide information about the network's ability to borrow and lend money. Included is the total number of household links, the number of links that split off since the last round of the IFLS, and the total number of links which have a married household head. Second is information about the level of shock a household's network is experiencing. Here we use the mean level of drought or health shock the network is experiencing, or the mean number of household members who are ill.

Due to the way in which the survey gathers data, a household's full number of links is bound to be under counted. Despite this, we argue adding in the network data will decrease bias in any model because otherwise there we fail to account for their respective "connectedness". This is particularly true with any aggregate level shock such as weather, since the shocks between connected households are highly correlated.

As households split-off, they are predictably different from the original household. Table 1 displays the characteristics of household broken down by if they are new split-offs or not. Split-off households are in general younger, unmarried with smaller families and have a greater likelihood of being in an urban area. These findings are not unexpected since most split-offs occur from children leaving their parent's homes. There may be some expectation that split-off households are created in a way to minimize risk from weather shocks, which is not indicated initially. We explore this in later sections. Split-offs also have fewer health shocks as well. We attribute this to several factors. First, those suffering health shocks are unlikely to leave a household. Second, the split-off household are younger and health problems develop with age. Lastly, split-offs have fewer members.

Data on Informal Transfers

The borrowing and lending information is taken from the IFLS and is gathered at the individual level about the amount transfers in cash to various relatives living outside the household in the previous 12 months. Transfers to parents, children, and spouses are asked about individually, while siblings are asked about as a group. Lastly there is a category for transfers to any other family members. To arrive at household level transfers we sum transfers across all household members to arrive total for each household. Though we do refer to transfers out of the household as lending and transfers in as borrowing, there are other ways to understand the transfers. Money out could be lending, or it could be the repayment for previous borrowing, or a gift with no expectation of return. There is also the possibility of informal trade between households with one working for the other for money or goods.

Transfers are quantified as any transfer, whether in cash or kind, from a non co-resident family member. This includes all split-off members as well as individuals not surveyed. As such our dependent and independent variable do not perfectly match-up. While not ideal, a household set of linked households is a subset of all non co-resident family members, so effects should exist but with additional noise from the other non co-resident families not surveyed.

Theoretically transfers between two households for the purposes of risk-sharing will be most effective when the correlation of own to network shock is low. For health shocks the correlation is near zero both for all illnesses and those preventing a member from working. For weather shocks there is a strong correlation between own shock and network shock, only if they live nearby. When strong correlation is present between shocks, risk sharing becomes difficult as others do not have free funds to help their links.

When considering transfer flows, several adaptive instruments may be adopted to smooth consumption in response to a shock. A few common strategies are to increase borrowing, obtain more remittances, reduce lending, or lower gift giving and repayment of debt. Studies have shown that some or all of these may be affected by shocks- specifically funds may be diverted away from expenses which may be considered less than a necessity, such as tuition for non-coresident children (Dalton and LaFave, 2017). Inasmuch as the reallocation of household budget serves as a much-studied risk sharing mechanism, in this survey we are unable to disaggregate specific dedication of transfers flowing in or out. Instead we have transfers related to money, tuition, and healthcare categorized as cash in (out) flows and flows related to food and miscellaneous characterized as flows in kind. While it restricts us from making statements about which categories in the household budget get unprioritized, an increase in transfers is indicative of the potency of the network, which is the crucial concern for this study.

3.3 Empirical Strategy

The purpose of this paper and therefore the empirical strategy is to not test for risk sharing but to identify the aspects of the family network that facilitate this risk sharing. We do this in two stages. If risk sharing is efficient, a change in household own income should not affect consumption (Townsend 1994, Morduch 1991). As noted in Fafchamps and Lund (2003) this strategy however, does not shed light on how households actually risk share under inefficient risk sharing. Our approach thus closely resembles the risk sharing model of Udry (1994) where net transfers are modeled as a function of household's own and network shocks. Using net cash transfers T_{idt} to household i alleviates the need to individually look at borrowings and repayments and allows the households to engage in both sides of the informal credit market. Formally,

$$T_{idt} = \beta_0 + \beta_1 \cdot S_{idt} + \beta_2 \cdot C_{idt} + \beta_3 \cdot S_{idt} \cdot C_{idt} + \beta_4 \cdot NS_{idt} + \gamma \cdot X_{idt} + \alpha_d + \Gamma_t + \epsilon_{idt} \quad (3.1)$$

The variables of interest are the household i's own shock (S_{idt}) in district d, in period t, the number of connections C_{idt} a household has, and the corresponding network shock (NS_{idt}).[7] We are also interested in exploring how the interaciton between a shock and the number of connections affect mitigation strategies. A shock to the household or the network is defined as either an aggregate shock or an idiosyncratic shock following the methods discussed above. If household are transferring money for the purposes of risk sharing, then $\beta_1 > 0$ and $\beta_4 < 0$ should both hold true.

We control for observed heterogeniety in household characteristics X_{it} like the age, marital status, and gender of the household head and household size. Also included are aggregated information about their connections such as the number of splitoffs, number of married household connections. Lastly district α_d and time Γ_t fixed effects are included.

Next, we turn to changes in household structure over time. A particularly interesting

[7] To properly identify the effect of a shock on borrowing, it is necessary to include the network shock to avoid creating omitted variable bias. This is especially true if there is a correlation between own shock and network shock. As discussed before, correlation is not a concern between health shocks, but is a concern for weather shocks. If the network shock is not excluded the effect of a shock on net borrowing will be incorrectly estimated as:

$$\tilde{\beta_1} = \beta_1 + \beta_2 \cdot \frac{cov(shock_{it}, netshock_{it}|X_{it}, N_{it})}{var(shock_{it}|X_{it}, N_{it})}$$

Since β_1 and β_2 are expected to have opposite signs this will constituent an underestimate, when a conditional correlation exists between the shocks. With high levels of correlation between the weather shocks the potential underestimation is large.

application of this set up is to explore if split off households are formed as a response to cope against certain shocks. For instance, a member may choose to leave their parent household that is or has been experiencing a shock to search out job opportunities and compensate for lost income. Conversely, family members may move back to parent households to pool resources as a coping mechanism.

In particular, we model dynamic link formation as a function of past and present shocks. We further disaggregate this effect based on variations in geographic distance of the links from parent households a la Fafchamps and Gubert 2007. This allows us to (a) generate some additional insights as to why certain links may form further away and (b) isolate the nature of link formation by shock type.

$$C_{idt} = \gamma_0 + \gamma_1 \cdot S_{idt} + \gamma_2 \cdot S_{idt-1} + \gamma_3 \cdot C_{idt-1} + \pi \cdot X_{idt-1} + \delta_d + \Sigma_t + \upsilon_{idt} \tag{3.2}$$

Apart from the general list of controls described above, we include past connections C_{idt-1} in this model as well.

3.4 Results

3.4.1 Differences in coping strategies

If households risk share, net transfers flowing into the household will rise as a function of their own shock and decrease as a function of network shock. We begin with the effect of an aggregate weather shock.

Columns 1, 3 and 5 in Table 3 report the regression coefficients of the effect of dry season temperature on net borrowing. We distinguish between the the original and split off households in order to parse out the differences in borrowing patterns. The result that emerges here is that while having more connections is deterrent to net transfers fro the newly formed split off households, it is greatly beneficial to original households. This is intuitive given that split off households may be formed in response to shocks and are thus more active in lending. [8]

Aggregate shocks do not happen in isolation. This means that a negative weather shock is likely to be experienced by the network links of a given household at the same time. As such, it is quite plausible that the effect on net borrowing is dependent on the ability to borrow, under varying levels of temperature. We estimate regressions of the form of equation (1) in columns 2, 4, and 6 of Table 3 by household type. For above average temperatures, every additional network link leads to more net borrowing, an outcome which is entirely driven by original households.

Against this backdrop, we turn to idiosyncratic health shocks which are more likely to not affect network members at the same time. Columns 1, 3, and 5 replicate the results of the previous table wherein more number of links lead to less and more borrowing for split-off and original households respectively. In columns 2, 4, and 6, we further show that an additional connection is not efficient in risk sharing for every other sick family member.

3.4.2 Evolution of networks

Split-off households are younger, have less likelihood of being married, are smaller in size and mostly live in urban areas (see Table 1). While this may simply be a result of maturation of the household structure, it also motivates the hypothesis that changes in household structure is a coping strategy adopted to spread risk. We thus turn to evolution of networks in this section.

Our data and empirical set up is ideal to study these changes as they addresses several

[8] While we do not report the coefficients on other determinants of net borrowing, it is interesting to note some of the associations that are indicative of underlying patterns in family networks . For example, bigger and married households or households with an older and male head borrow significantly less. A newly split-off household borrows significantly more, indicative of fixed costs associated with setting up a new household. Results are available upon request.

sources of known biases (Fafchamps and Lund 2003) in network variables. First, we rule out endogenous link formation consisting only of members who are most likely to help. Since, in our link formation strategy, we only consider members who are related by common ancestry, households have limited ability to select network links based on whether they experienced a shock. Households could have had sons and daughters move out prior to the survey who we do not observe, but cannot create links where there were none. Second, besides the actual formation of new links, we are able to focus on the distance between the new links and the original household. Geographic distance especially under aggregate shocks, may be an efficient way to diversify risk.

The results in the previous section point towards two important implications that may play a role in how links form- First, networks are more active under aggregate shocks than under idiosyncratic shocks. Second, we show that there may be upfront fixed costs in the creation of new links as these are predominantly newer and younger households.

Estimates of equation (2) are shown in Tables 5 and 6 where we attempt to explain link formation as a response to shocks in the past. We observe that households do not engage in link formation under aggregate shocks, indicating the motivation to pool resources together when the shock is more pervasive and new link formation is costly. Second, we observe that link formation is stronger when a household member is sick last period, suggesting a time lag in the adaptation to a shock. The propensity to form these new links is higher within the district than outside a district.

3.5 Conclusion

In this paper we explore how transfers between extended family networks are affected by shocks in Indonesia. While we see networks behave in standard risk-sharing transfers under aggregate shocks, we find that they more likely to undertake costly transformations to household structures under idiosyncratic health shocks. The key distinguishing factor in both types of strategies is the immediate versus lagged requirement of risk pooling.

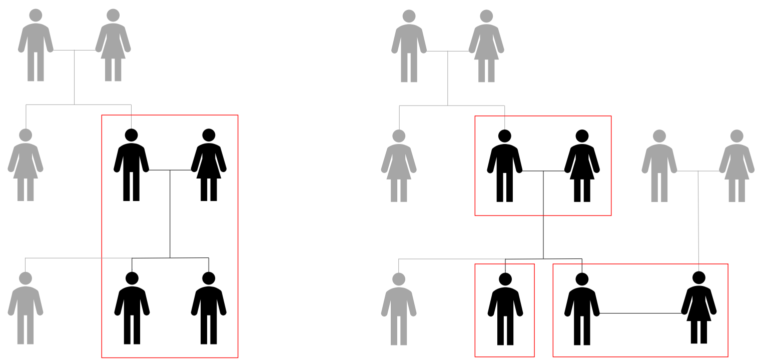
One of our key contributions is detailing the importance of including network shocks when analyzing transfers between households. Under aggregate shocks where shocks will be correlated, not including these constitutes omitting a key variable and will lead to underestimation of the overall effect of the shock. Moreover, we identify portions of the family network who are more likely to be nodal in risk sharing behaviors and may therefore be instrumental in transmitting information and resources.

Exploring the effects of shocks in developing countries is important as understanding how existing institutional structures mitigate some risks while collapse under others. Some shocks cannot be properly insured through networks of families, and thus create a constant threat of poverty over those living in less developed countries. This research helps illuminate where government and non-government interventions are most useful in insuring against shocks.

Figures

Figure 3.1: Family Networks

(a) 1st Period **(b)** 2nd Period



Note: (a) Figures in black denote split-off links from original household who are surveyed. Figures in grey are links who are not followed after. (b) Split-offs of up to three generations can be observed in IFLS.

Tables

Table 3.1: Descriptive Statistics by Household Types

	Parent Households		Split-off Households	
	Mean	Std Dev	Mean	Std Dev
Borrowing (Millions Rupiah)$^{\gamma}$	42.7	378.0	26.2	242.3
Lending (Millions Rupiah)	43.5	345.2	33.3	298.1
Dry Season Temperature (C)	29.8	3.04	30.2	3.19
Sick Household Members$^{\delta}$	0.950	1.40	0.326	0.731
Number of Links	0.923	0.979	1.00	0.323
Age of Household Head	48.8	20.0	34.0	35.66
Marital Status	0.819	0.385	0.794	0.405
Male Head of Household	0.822	0.383	0.855	0.352
Household Size	4.14	1.86	3.20	1.67
Split-off Connections	0.329	0.643	0	0
Married Connections	0.503	0.487	0.736	0.436
Observations	24,896		8,194	

Notes: A split-off household is defined as a household that is first observed in the current wave of the IFLS, while original households have been observed for multiple periods. $^{\gamma}$In December of 2014, when IFLS 5 took place, 1 US Dollar was equal to 12,650 Indonesian Rupiah. $^{\delta}$Detailed information about sick household members is only recorded in IFLS 4 and 5, thus there are only 18,773 original observations and 6,635 splitoff observations for that variable.

Table 3.2: Descriptive Statistics on total Network Links

	Total # of links	% of Links in same district as HH	% Links in same province as HH	Transfers as proportion of consumption
IFLS 3	1941	73.16	90.06	0.109
IFLS 4	4629	70.12	87.00	0.171
IFLS 5	5882	64.23	84.03	0.238
Total	12,452	67.81	86.07	0.185

Notes: Total consumption for a household is calculated by adding up value of food, rent or rental value of their home, monthly, and yearly purchases. For Column 3, the top 2% of household are dropped as some had very low consumption values leading to high levels of normalized borrowing.

Table 3.3: Results: Aggregate Shock and Networks

	All Households		Splitoffs		Originals	
	(1)	(2)	(3)	(4)	(5)	(6)
Dry Season Temperature (C)	0.022*	0.011	0.044*	0.046	0.014	0.003
	(0.013)	(0.014)	(0.026)	(0.029)	(0.014)	(0.015)
Number of Links	0.073***	-0.247**	-0.148**	-0.092	0.084***	-0.239**
	(0.017)	(0.110)	(0.066)	(0.456)	(0.018)	(0.121)
Temp X Links		0.011***		-0.002		0.011***
		(0.004)		(0.015)		(0.004)
Network Temperature (C)	-0.008**	-0.013***	-0.013	-0.013	-0.001	-0.008
	(0.004)	(0.004)	(0.008)	(0.009)	(0.005)	(0.005)
Other Controls	X	X	X	X	X	X
Time FE	X	X	X	X	X	X
District FE	X	X	X	X	X	X
Observations	33,090	33,090	8,194	8,194	24,896	24,896
R^2	0.045	0.045	0.098	0.098	0.048	0.048

Notes: Standard Errors are clustered at the District-IFLS level. * indicates signfigance at the 10% level, ** indicates signifigance at the 5% level, and *** indicates signifigance at the 1% level. Dependent variable is the log of net cash transfers. A splitoff household is defined as a household that is first observed in the current wave of the IFLS, while original households have been observed for multiple periods. Other varriables controled for included the number of household members, the age and age squared of the household head, an indicator for if household head is male, an indicator for it the household head is married, the number of splitoffs a household is connected to and the number of married households a household is connected to. In columns 1, 3, and 5 a dummy for if a household is a new splitoff is also included.

Table 3.4: Results: Idiosyncratic Shock and Networks

	All Households		Splitoffs		Originals	
	(1)	(2)	(3)	(4)	(5)	(6)
Sick Household Members	0.030***	0.021	-0.030	0.012	0.027***	0.021
	(0.009)	(0.013)	(0.025)	(0.096)	(0.009)	(0.014)
Number of Links	0.078***	0.067***	-0.155**	-0.142**	0.088***	0.081***
	(0.019)	(0.022)	(0.068)	(0.070)	(0.019)	(0.023)
Sick X Links		0.007		-0.040		0.005
		(0.008)		(0.084)		(0.008)
Network Sick Household Members	-0.015**	-0.016**	-0.003	-0.003	-0.016*	-0.016*
	(0.007)	(0.007)	(0.013)	(0.013)	(0.009)	(0.009)
Other Controls	X	X	X	X	X	X
Time FE	X	X	X	X	X	X
District FE	X	X	X	X	X	X
Observations	25,408	25,408	6,635	6,635	18,773	18,773
R^2	0.048	0.048	0.109	0.109	0.051	0.051

Notes: Standard Errors are clustered at the District-IFLS level. * indicates signfigance at the 10% level, ** indicates signifigance at the 5% level, and *** indicates signifigance at the 1% level. Dependent variable is the log of net cash transfers. A splitoff household is defined as a household that is first observed in the current wave of the IFLS, while original households have been observed for multiple periods. Other variables controled for included the number of household members, the age and age squared of the household head, an indicator for if household head is male, an indicator for it the household head is married, the number of splitoffs a household is connected to and the number of married households a household is connected to. In columns 1, 3, and 5 a dummy for if a household is a new splitoff is also included.

Table 3.5: Results: Aggregate Shocks and Network Formation

	Total Links (1)	Outside Links (2)	Inside Links (3)	New Links (4)
Dry Season Temperature (C)	-0.009**	0.008	-0.017***	-0.006*
	(0.004)	(0.006)	(0.006)	(0.003)
Dry Season Temperature (C) T-1	-0.043***	-0.022***	-0.021**	-0.016**
	(0.012)	(0.008)	(0.010)	(0.008)
Number of Links T-1	1.005***	0.041***	0.170***	-0.007
	(0.009)	(0.008)	(0.014)	(0.006)
Dependent Varriable T-1		0.801***	0.792***	0.138***
		(0.016)	(0.015)	(0.013)
Other Controls	X	X	X	X
Time FE	X	X	X	X
District FE	X	X	X	X
Observations	18,063	18,063	18,063	18,063
R^2	0.602	0.422	0.583	0.191

Notes: Standard Errors are clustered at the District-IFLS level. * indicates signfigance at the 10% level, ** indicates signifigance at the 5% level, and *** indicates signifigance at the 1% level.

Table 3.6: Results: Idiosyncratic Shocks and Network Formation

	Total Links (1)	Outside Links (2)	Inside Links (3)	New Links (4)
Sick Household Members	0.010*	0.006*	0.004	0.010***
	(0.005)	(0.003)	(0.004)	(0.003)
Sick Household Members T-1	0.035***	0.008	0.026***	0.019***
	(0.010)	(0.008)	(0.009)	(0.007)
Number of Links T-1	1.007***	0.038***	0.165***	-0.005
	(0.010)	(0.010)	(0.016)	(0.007)
Dependent Varriable T-1		0.826***	0.795***	0.140***
		(0.020)	(0.018)	(0.017)
Other Controls	X	X	X	X
District FE	X	X	X	X
Observations	10,479	10479	10,479	10,479
R^2	0.631	0.469	0.615	0.216

Notes: Standard Errors are clustered at the District-IFLS level. * indicates signfigance at the 10% level, ** indicates signifigance at the 5% level, and *** indicates signifigance at the 1% level.

Bibliography

Akerman, A., Gaarder, I., and Mogstad, M., 2015. The skill complementarity of broadband internet, *The Quarterly Journal of Economics*, 130 (4), 1781–1824.

Ambrus, A., Mobius, M., and Szeidl, A., 2010. Consumption risk-sharing in social networks (vol. no. w15719).

Angelucci, M., De Giorgi, G., Rangel, M.A., and Rasul, I., 2010. Family networks and school enrolment: Evidence from a randomized social experiment, *Journal of public Economics*, 94 (3-4), 197–221.

Angrist, J.D. and Keueger, A.B., 1991. Does compulsory school attendance affect schooling and earnings?, *The Quarterly Journal of Economics*, 106 (4), 979–1014.

Auffhammer, M., Hsiang, S.M., Schlenker, W., and Sobel, A., 2013. Using weather data and climate model output in economic analyses of climate change, *Review of Environmental Economics and Policy*, 7 (2), 181–198.

Barreca, A.I., 2012. Climate change, humidity, and mortality in the united states, *Journal of Environmental Economics and Management*, 63 (1), 19–34.

Besley, T. and Burgess, R., 2000. Land reform, poverty reduction, and growth: Evidence from india, *The Quarterly Journal of Economics*, 115 (2), 389–430.

Bessen, J., 2015. *Learning by doing: the real connection between innovation, wages, and wealth*, Yale University Press.

Bharadwaj, P., Fenske, J., Kala, N., and Mirza, R.A., 2020. The green revolution and infant mortality in india, *Journal of health economics*, 71, 102314.

Brainerd, E. and Menon, N., 2014. Seasonal effects of water quality: The hidden costs of the green revolution to infant and child health in india, *Journal of Development Economics*, 107, 49–64.

Corno, L., Hildebrandt, N., and Voena, A., 2020. Age of marriage, weather shocks, and the direction of marriage payments, *Econometrica*, 88 (3), 879–915.

Dalton, M. and LaFave, D., 2017. Mitigating the consequences of a health condition: The role of intra-and interhousehold assistance, *Journal of Health Economics*, 53, 38–52.

Dasgupta, A., 2018. Technological change and political turnover: The democratizing effects of the green revolution in india, *American Political Science Review*, 112 (4), 918–938.

Dasgupta, B., 1977. India's green revolution, *Economic and political weekly*, 241–260.

Dell, M., Jones, B.F., and Olken, B.A., 2014. What do we learn from the weather? the new climate-economy literature, *Journal of Economic Literature*, 52 (3), 740–98.

Dercon, S., 2002. Income risk, coping strategies, and safety nets, *The World Bank Research Observer*, 17 (2), 141–166.

Deschenes, O., 2014. Temperature, human health, and adaptation: A review of the empirical literature, *Energy Economics*, 46, 606–619.

Deschênes, O. and Greenstone, M., 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us, *American Economic Journal: Applied Economics*, 3 (4), 152–85.

Donaldson, D. and Storeygard, A., 2016. The view from above: Applications of satellite data in economics, *Journal of Economic Perspectives*, 30 (4), 171–98.

Duflo, E., 2001. Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment, *American economic review*, 91 (4), 795–813.

Duflo, E., Dupas, P., and Kremer, M., 2010. Education and fertility: Experimental evidence from kenya, *Unpublished Manuscript, J-PAL/MIT. Cambridge, MA*.

Ebenstein, A., Lavy, V., and Roth, S., 2016. The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution, *American Economic Journal: Applied Economics*, 8 (4), 36–65.

Eisner, E., 2019. Have robots taken all of our jobs before? the case of mechanization in us agriculture.

Enander, A.E. and Hygge, S., 1990. Thermal stress and human performance, *Scandinavian journal of work, environment & health*, 44–50.

Fafchamps, M. and Gubert, F., 2007. The formation of risk sharing networks, *Journal of development Economics*, 83 (2), 326–350.

Fafchamps, M. and Lund, S., 2003. Risk-sharing networks in rural philippines, *Journal of development Economics*, 71 (2), 261–287.

Foster, A.D. and Rosenzweig, M.R., 2002. Household division and rural economic growth, *The Review of Economic Studies*, 69 (4), 839–869.

Frankel, F.R., 2015. *India's green revolution: Economic gains and political costs*, Princeton University Press.

Garg, T., Jagnani, M., and Taraz, V., 2018. Temperature and human capital in india.

Geruso, M. and Spears, D., 2018. Heat, humidity, and infant mortality in the developing world, Tech. rep., National Bureau of Economic Research.

Gollin, D., Hansen, C.W., and Wingender, A., 2019. Two blades of grass: Agricultural innovation, productivity, and economic growth, Tech. rep., working paper.

Graetz, G. and Michaels, G., 2018. Robots at work, *Review of Economics and Statistics*, 100 (5), 753–768.

Graff Zivin, J., Hsiang, S.M., and Neidell, M., 2018. Temperature and human capital in the short and long run, *Journal of the Association of Environmental and Resource Economists*, 5 (1), 77–105.

Grimard, F., 1997. Household consumption smoothing through ethnic ties: evidence from cote d'ivoire, *journal of Development Economics*, 53 (2), 391–422.

Hancock, P.A., 1986. Sustained attention under thermal stress., *Psychological bulletin*, 99 (2), 263.

Heft-Neal, S., Lobell, D.B., and Burke, M., 2017. Using remotely sensed temperature to estimate climate response functions, *Environmental Research Letters*, 12 (1), 014013.

Hicks, D.L. and Hicks, J.H., 2013. Lucky late bloomers: The impact of early marriage on adult outcomes in western kenya, *Unpublished report.*

Janzen, J.P., Smith, A., and Carter, C.A., 2018. Commodity price comovement and financial speculation: the case of cotton, *American Journal of Agricultural Economics*, 100 (1), 264–285.

Kazianga, H. and Udry, C., 2004. Consumption smoothing and livestock in rural burkina faso, *Economic Growth Center, Working Paper*, 898, 368.

Kinsey, B., Burger, K., and Gunning, J.W., 1998. Coping with drought in zimbabwe: Survey evidence on responses of rural households to risk, *World development*, 26 (1), 89–110.

Kleemans, M., 2015. Migration choice under risk and liquidity constraints, Tech. rep.

LoPalo, M., 2019. Temperature,worker productivity, and adaptation:evidence from survey data production.

Manuelli, R.E. and Seshadri, A., 2014. Frictionless technology diffusion: The case of tractors, *American Economic Review*, 104 (4), 1368–91.

McKenzie, D.J., 2003. How do households cope with aggregate shocks? evidence from the mexican peso crisis, *World Development*, 31 (7), 1179–1199.

Mohan, R. and Evenson, R.E., 1975. The intensive agricultural districts programme in india: A new evaluation, *The Journal of Development Studies*, 11 (3), 135–154.

Morduch, J., 1991. Consumption smoothing across space: Tests for village-level responses to risk, Tech. rep., Mimeo, Harvard University.

Moscona, J., 2017. The impact of india's green revolution: An empirical investigation of modern agricultural development, *Unpublished Manuscript*.

Munshi, K., 2004. Social learning in a heterogeneous population: technology diffusion in the indian green revolution, *Journal of development Economics*, 73 (1), 185–213.

Murgai, R., 1999. *The Green Revolution and the productivity paradox: evidence from the Indian Punjab*, The World Bank.

Murgai, R., Winters, P., Sadoulet, E., and De Janvry, A., 2002. Localized and incomplete mutual insurance, *Journal of development Economics*, 67 (2), 245–274.

Nakamura, S., Bundervoet, T., and Nuru, M., 2020. Rural roads, poverty, and resilience: Evidence from ethiopia, *The Journal of Development Studies*, 56 (10), 1838–1855.

Nunn, N. and Qian, N., 2011. The potato's contribution to population and urbanization: evidence from a historical experiment, *The quarterly journal of economics*, 126 (2), 593–650.

Olmstead, A.L. and Rhode, P.W., 2001. Reshaping the landscape: the impact and diffusion of the tractor in american agriculture, 1910-1960, *Journal of Economic history*, 663–698.

Ozier, O., 2015. *The impact of secondary schooling in Kenya: A regression discontinuity analysis*, The World Bank.

Parayil, G., 1992. The green revolution in india: A case study of technological change, *Technology and culture*, 33 (4), 737–756.

Park, J., 2017. *Temperature, test scores, and human capital production*, Ph.D. thesis, Doctoral dissertation. Harvard University Economics Department.

Pingali, P., 2007. Will the gene revolution reach the poor?-lessons from the green revolution, *Mansholt Lecture, Wageningen University, the Netherlands, January*, 26.

Prahladachar, M., 1983. Income distribution effects of the green revolution in india: A review of empirical evidence, *World Development*, 11 (11), 927–944.

Ramsey, J.D., 1995. Task performance in heat: a review, *Ergonomics*, 38 (1), 154–165.

Riniolo, T.C. and Schmidt, L.A., 2006. Chronic heat stress and cognitive development: An example of thermal conditions influencing human development, *Developmental review*, 26 (3), 277–290.

Rosenzweig, M.R., 1988. Risk, implicit contracts and the family in rural areas of low-income countries, *The Economic Journal*, 98 (393), 1148–1170.

Sanyal, B., 1983. How revolutionary was india's green revolution?, *South Asia Bulletin*, 3 (2), 31–44.

Sheffield, J., Goteti, G., and Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *Journal of climate*, 19 (13), 3088–3111.

Smith, J.P., Thomas, D., Frankenberg, E., Beegle, K., and Teruel, G., 2002. Wages, employment and economic shocks: Evidence from indonesia, *Journal of Population Economics*, 15 (1), 161–193.

Townsend, R.M., 1994. Risk and Insurance in Village India, *Econometrica: Journal of the Econometric Society*, 539–591.

Udry, C., 1994. Risk and insurance in a rural credit market: An empirical investigation in northern nigeria, *The Review of Economic Studies*, 61 (3), 495–526.

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

Westerterp-Plantenga, M.S., van Marken Lichtenbelt, W.D., Strobbe, H., and Schrauwen, P., 2002. Energy metabolism in humans at a lowered ambient temperature, *European journal of clinical nutrition*, 56 (4), 288–296.

Westley, J.R., 2019. *Agriculture and equitable growth: The case of Punjab-Haryana*, Routledge.

World Bank, 2018. Educational attainment, upper secondary indonesia.

Appendix

Table A1.1: IV Estimates: Long Term Education Outcomes

	Years of Education (1)	I(Secondary Enrolment>0) (2)	I(Secondary Completion>0) (3)	Grade Repitition (4)
Math score	0.528**	0.113***	0.061*	-0.061**
	(0.224)	(0.036)	(0.035)	(0.029)
N	6,119	6,134	4,073	2,201

Note: *** p<0.01, ** p<0.05, * p<0.1.Standard errors are robust and in parenthesis. Math score, which is available for the largest sample group, is the endogenous independent variable for the second stage. It is instrumented by exam-month wet bulb temperatures in the first stage. It is calculated on a scale of 0-10. Results include month and cohort fixed effects. For full list of controls, see text.

Table A1.2: IV Estimates : Long Term Labor Market Outcomes

	Labor Force Participation (1)	Primary Activity: Housework (2)	Primary Activity: Paid work (3)
Math score	0.101***	-0.104***	0.084**
	(0.034)	(0.032)	(0.034)
N	6,133	5,965	6,132

Note: *** p<0.01, ** p<0.05, * p<0.1.Standard errors are robust and in parenthesis. Math score, which is available for the largest sample group, is the endogenous independent variable for the second stage. It is instrumented by exam-month wet bulb temperatures in the first stage. It is calculated on a scale of 0-10. Results include month and cohort fixed effects. For full list of controls, see text.

Table A1.3: IV Estimates of the Effect of Test Scores on Labor Market Returns

	Hours Worked Last Week (1)	Hours Worked per Week (2)	Log(Monthly wages) (3)	Log(Wages earned per hour) (4)
Math score	1.479	1.597	0.155	0.137
	(1.747)	(1.614)	(0.363)	(0.327)
N	8,566	8,566	3,021	2,918

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis. Wages is in millions of Indoensian Rupiah. Math score, which is available for the largest sample group, is the endogenous independent variable for the second stage. It is instrumented by exam-month wet bulb temperatures in the first stage. It is calculated on a scale of 0-10. Results include month and cohort fixed effects. For full list of controls, see text.

Table A1.4: IV estimates: Marital Outcomes

	Age of Marriage (1)	Age of Marriage Male (2)	Age of Marriage Female (3)	I(Married by 18) (4)	Log(Dowry Value) (5)	Spousal Age gap (6)
Math score	0.557	0.917	0.532	-0.047	-1.370***	-0.170
	(0.495)	(0.580)	(0.961)	(0.036)	(0.521)	(0.662)
N	3,736	2,014	1,722	3,736	3,555	1,761

Note: *** p<0.01, ** p<0.05, * p<0.1.Standard errors are robust and in parenthesis. Dowry Value is in millions of Indoensian Rupiah. Math score, which is available for the largest sample group, is the endogenous independent variable for the second stage. It is instrumented by exam-month wet bulb temperatures in the first stage. It is calculated on a scale of 0-10. Results include month and cohort fixed effects. For full list of controls, see text.

Table A1.5: Impact of Hot weather during the previous Dry Agricultural Season

	Math (1)	Science (2)	Language (3)	Religion (4)	Social Studies (5)	Total (6)
# days with T_{wb}>27.5	-0.016	-0.023	-0.029	-0.028	-0.020	-0.058
in August of year $t-1$	(0.021)	(0.024)	(0.018)	(0.022)	(0.025)	(0.062)
# days with T_{wb}>27.5	0.020	-0.004	-0.017	-0.033	-0.015	0.001
in September of year $t-1$	(0.021)	(0.023)	(0.016)	(0.022)	(0.024)	(0.059)
# days with T_{wb}>27.5	-0.022	-0.022	-0.010	-0.012	0.002	-0.063
in October of year $t-1$	(0.015)	(0.016)	(0.011)	(0.014)	(0.016)	(0.039)
N	12,283	7,995	12,140	7,882	7,738	7,952
R^2	0.177	0.155	0.205	0.159	0.139	0.169

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis.Controls include temperature during exam month, individual controls and year, province fixed effects.

Table A1.6: Chronic Heat Stress: Impact of hot months in preceding School years

	Math	Science	Language	Religion	Social Studies	Total
	(1)	(2)	(3)	(4)	(5)	(6)
# days with T_{wb}>27.5	-0.010	-0.020**	-0.027***	-0.013	-0.015	-0.041*
in School year $t-1$	(0.008)	(0.009)	(0.006)	(0.008)	(0.010)	(0.022)
# days with T_{wb}>27.5	-0.008	-0.004	0.001	0.003	0.016*	-0.015
in School year $t-2$	(0.009)	(0.009)	(0.007)	(0.008)	(0.009)	(0.023)
# days with T_{wb}>27.5	0.003	-0.008	0.006	-0.011	-0.019*	-0.004
in School year $t-3$	(0.009)	(0.009)	(0.007)	(0.008)	(0.010)	(0.024)
N	11,627	7,358	11,494	7,234	7,113	7,318
R^2	0.173	0.145	0.207	0.165	0.146	0.165

Note: *** p<0.01, ** p<0.05, * p<0.1. Standard errors are robust and in parenthesis.Controls include temperature during exam month, individual controls and year, province fixed effects.

Table A2.1: Effect of Productivity Increase on Employment and Wages using State Fixed Effects

	Dependent Variables			
	Log Daily Wages	Log Aggregate Employment^	Log # of Agricultural Laborers	Log # of Cultivators
	(1)	(2)	(3)	(4)
Panel A: All Districts				
Log Rice Yield (tons/ha)	0.0021	0.074***	0.150***	0.060***
	(0.011)	(0.015)	(0.019)	(0.018)
Log Wheat Yield (tons/ha)	0.031***	-0.0059	-0.022	0.012
	(0.011)	(0.013)	(0.017)	(0.016)
Mean of Dependent Variable	0.039	286,475	932	334
N	5667	5667	5667	5667
Adjusted R^2	0.571	0.390	0.528	0.337
Panel B: Sample restricted to Rice Districts				
Log Rice Yield (tons/ha)	0.078***	0.180***	0.240***	0.150**
	(0.029)	(0.058)	(0.062)	(0.070)
Mean of Dependent Variable	0.032	364,723	446	1101
N	2607	2607	2607	2607
Adjusted R^2	0.382	0.487	0.556	0.318
Panel B: Sample restricted to Wheat Districts				
Log Wheat Yield(tons/ha)	0.085***	0.039	0.140**	0.022
	(0.021)	(0.025)	(0.058)	(0.024)
Mean of Dependent Variable	0.043	235,422	261	822
N	4096	4096	4096	4096
Adjusted R^2	0.487	0.505	0.552	0.529
District controls	Y	Y	Y	Y
Geographic controls	Y	Y	Y	Y
Time Fixed Effects	Y	Y	Y	Y

Note: Standard errors clustered at district level are shown in parenthesis. *p<.10 **p<.05 ***p<.001. All regressions include district level controls, soil controls, climate controls and time fixed effects from OLS regression specification. Aggregate Employment is the sum of cultivators and agricultural laborers weighted by the number of days worked in the state by farm workers. The Census defines agricultural laborers as those who work for wages on private/Government owned farms. Cultivators are defined as a person engaged as employer/family worker paid in cash or kind or share of crop.

Table A2.2: Effect of productivity in Neighboring Districts that switch

	Dependent Variables			
	Log Daily Wages	Log Aggregate Employment^	Log # of Agricultural Laborers	Log # of Cultivators
	(1)	(2)	(3)	(4)
Wheat 1970, Rice 1985	0.118***	-0.050	-0.103	-0.0367
	(0.024)	(0.060)	(0.065)	(0.064)
Rice 1970, Wheat 1985	0.442***	-0.175**	-0.119**	-0.175**
	(0.062)	(0.049)	(0.037)	(0.050)

Note: Standard errors clustered at district level are shown in parenthesis. *p<.10 **p<.05 ***p<.001. All regressions include district level controls, soil controls, climate controls and time fixed effects from OLS regression specification. Aggregate Employment is the sum of cultivators and agricultural laborers weighted by the number of days worked in the state by farm workers. The Census defines agricultural laborers as those who work for wages on private/Government owned farms. Cultivators are defined as a person engaged as employer/family worker paid in cash or kind or share of crop.

Table A2.3: Effect of Productivity on Employment and Wages w/o Bengal and Gujarat

	Dependent Variables			
	Log Daily Wages	Log Aggregate Employment^	Log # of Agricultural Laborers	Log # of Cultivators
	(1)	(2)	(3)	(4)
Log Rice Yield (tons/ha)	0.004	0.092***	0.157***	0.086***
	(0.011)	(0.017)	(0.020)	(0.020)
Log Wheat Yield(tons/ha)	0.027**	0.001	-0.030	0.008
	(0.012)	(0.014)	(0.022)	(0.016)

Standard errors clustered at district level are shown in parenthesis. *p<.10 **p<.05 ***p<.001. All regressions include district level controls, soil controls, climate controls and time fixed effects from OLS regression specification. Aggregate Employment is the sum of cultivators and agricultural laborers weighted by the number of days worked in the state by farm workers. The Census defines agricultural laborers as those who work for wages on private/Government owned farms. Cultivators are defined as a person engaged as employer/family worker paid in cash or kind or share of crop.