UNIVERSITY OF CALIFORNIA, SAN DIEGO

Three essays on environment and development economics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Jarrod Ross Welch

Committee in charge:

Professor Richard Carson, Co-Chair
Professor Joshua Graff Zivin, Co-Chair
Professor Prashant Bharadwaj
Professor Gordon Dahl
Professor Craig McIntosh

2011

The dissertation of Jarrod Ross Welch is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co-Chair

Co-Chair

University of California, San Diego

2011

DEDICATION

To my parents, whose love and support made the journey possible;
Danya, for always being there to make fun of me when I got too serious;
Malyk, for keeping life in a healthy perspective;
and Jake, for watching over and guiding me from beyond the threshold.

EPIGRAPH

The most difficult subjects can be explained
to the most slow-witted man if he has not formed any idea of them already;
but the simplest thing cannot be made clear to the most intelligent man
if he is firmly persuaded that he knows already, without a shadow of doubt,
what is laid before him.
—Leo Tolstoy

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge my doctoral committee for their support: Professors Prashant Bharadwaj and Gordon Dahl for helpful comments and suggestions; special thanks to Professor Craig McIntosh for guidance through the years, and to my committee co-chairs, Professor Richard Carson, forever dedicated to his students' success, and Professor Joshua Graff Zivin whose feet pounded the pavement 526,000 times as he logged 263 miles by my side while discussing Chapter 2, seemingly *ad infinitum*.

Thanks to my co-authors, Jen Alix-Garcia and Kate Sims, all the researchers at the International Rice Research Institute and the Food and Agriculture Organization whose efforts made much of this research possible, and especially Jeff Vincent and Max Auffhammer for ongoing and invaluable advice.

Chapter 1, in full, is a reprint of the material as it appears the Proceedings of the National Academy of Sciences, vol. 107, no. 33, 2010. Welch, Jarrod R.; Vincent, Jeffrey R.; Auffhammer, Maximilian; Moya, Piedad F.; Dobermann, Achim; and Dawe, David. The dissertation author shared equally in the research design and analysis, and is lead author of the paper.

Chapter 3, in full, has been submitted for publication of the material as it may appear in the Review of Economics and Statistics. Alix-Garcia, Jennifer; McIntosh, Craig; Sims, Katharine R.E.; Welch, Jarrod R. The dissertation author shared equally in the research design, analysis and authorship.

VITA

2002	B. A. with High Honors in Economics, University of California, Santa Barbara
2006-2011	Graduate Student Researcher, Graduate School of International Relations and Pacific Studies, University of California, San Diego
2007-2011	Teaching Assistant, Graduate School of International Relations and Pacific Studies, University of California, San Diego
2008	M. S. in Economics, University of California, San Diego
2011	Ph. D. in Economics, University of California, San Diego

PUBLICATIONS

Welch, Jarrod R.; Vincent, Jeffrey R.; Auffhammer, Maximilian; Moya, Piedad F.; Dobermann, Achim; and Dawe, David; "Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures", Proceedings of the National Academy of Sciences, 2010, 107(33), p. 14562-14567.

ABSTRACT OF THE DISSERTATION

Three essays on environment and development economics

by

Jarrod Ross Welch

Doctor of Philosophy in Economics

University of California, San Diego, 2011

Professor Richard Carson, Co-Chair
Professor Joshua Graff Zivin, Co-Chair

This dissertation is composed of three original, self-contained essays on environment and development economics. In the first essay we examine the relationship between weather and rice production at the farm level in Asia. Higher minimum temperature reduced yield while higher maximum temperature raised it; radiationŠs impact varied by growth phase. Combined, these effects imply that yield at most sites would have grown more rapidly during the high-yielding season but less rapidly during the low-yielding season if observed temperature trends at the end of the 20th Century had not occurred. Diurnal temperature variation must be considered when investigating the impacts of climate change on irrigated rice in Asia.

In the second essay, I expand on the models used in the first, and incorporate the fact that agricultural yield functions are non-linear, with sharp negative impacts when crops are exposed to temperature in excess of certain thresholds. Exploiting exogenous variation in planting date, I

demonstrate exogenous variation in above-threshold exposure-time comparable to the projected increase due to 100 years of climate change, and analyze the ability of farmers to make adjustments to compensate. I show that farmers do make small adjustments in the quantity of seed that they plant, as well as the amount of nitrogen fertilizer that they apply to the crop. This has important implications for the validity of the typical approach of using observed weather shocks to measure the impact of climate change on agriculture: farmers make adjustments according to expectations. As climate change will be slow and predictable, measuring agricultural output as a function of unpredictable shocks may overstate the true impact of climate change.

The third essay is unrelated to the first two, and in it we study the consequences of poverty alleviation programs for environmental degradation. We exploit the community-level eligibility discontinuity for a conditional cash transfer program in Mexico to identify the impacts of income increases on deforestation, and use the program's initial randomized rollout to explore household responses. We find that additional income raises consumption of land-intensive goods and increases deforestation. The observed production response and deforestation increase are larger in communities with poor road infrastructure. This suggests that better access to markets disperses environmental harm and that the full effects of poverty alleviation can be observed only where poor infrastructure localizes them.

Chapter 1

Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures

1.1 Introduction

The impacts of temperature and solar radiation on rice yield remain imperfectly known despite decades of agronomic research. Current knowledge is based primarily on field trials and greenhouse experiments. These experimental studies indicate that increased temperature (Yoshida and Parao, 1976; Yoshida, Satake and Mackill, 1981; Seshu and Cady, 1984; Wassmann, Jagadish, Heuer, Ismail, Redona, Serraj, Singh, Howell, Pathak and Sumfleth, 2009b) and decreased radiation (Yoshida and Parao, 1976)(Seshu and Cady, 1984; Evans and Datta, 1979) can reduce yield, with the impacts varying across the plant's three growth phases (vegetative = establishment to panicle initiation, reproductive = panicle initiation to flowering, ripening = flowering to mature grain). Unresolved issues remain with respect to the relative impacts of temperature during daytime (T_{max}) vs. nighttime (T_{min}), potentially confounding impacts of temperature and radiation, and the magnitude of impacts in non-experimental settings. Here, we investigate these issues by analyzing data from the largest farm-level rice study conducted in Asia since the mid-1980s. To our knowledge, this paper is the first to use disaggregated data from farmer-managed fields to disentangle the impacts of T_{min}, T_{max}, and solar radiation on rice yield.

With few exceptions (Lobell and Field, 2007; Lobell, 2007), most statistical studies on temperature and rice yield have focused on the impact of daily mean temperature (T_{ave}), despite evidence that that the effects of T_{min} and T_{max} on crop phenological development and physiological processes differ (Wassmann et al., 2009b). It is well-established that extremely high levels of T_{max} during flowering can drastically reduce rice yield due to spikelet sterility, but recent studies have provided evidence that yield might be more sensitive to T_{min} than to T_{max} in locations where spikelet sterility is rarely observed (Peng, Huang, Sheehy, Laza, Visperas, Zhong, Centeno, Khush and Cassman, 2004). Rice simulation models began to include T_{min} and T_{max} as separate variables only recently (Wassmann et al., 2009b). Better understanding of the impacts of temperature at different points in the diurnal cycle is needed, as T_{min} has been rising faster than T_{max} in some important Asian rice-growing countries, including the two largest, China (Zhou, Dickinson, Tian, Fang, Li, Kaufmann, Tucker and Myneni, 2004) and India (B. Padma Kumari, 2007), and is projected to continue doing so in the future (IPCC, 2007).

Potentially confounding impacts of T_{min} and radiation on rice yield in field experiments have attracted recent attention (Peng et al., 2004)(Sheehy, Mitchell and Ferrer, 2006), although this was recognized as a challenge for yield studies decades ago (Evans and Datta, 1979). The difficulty stems from the complex meteorological effects of clouds, which reduce not only insolation but also back radiation, thus possibly increasing T_{min} by enhancing long-wave surface warming at night (Sheehy et al., 2006)(Huang, Dickinson and Chameides, 2006). Understanding the relative impacts of T_{min} and radiation is important in view of evidence of a declining trend in surface radiation ("global dimming") (Stanhill and Cohen, 2001), which is likely due to increased cloudiness caused by a combination of global warming and regional "brown clouds" of aerosol pollution (Huang, Dickinson and Chameides, 2006; Stanhill and Cohen, 2001; Dai, Trenberth and Karl, 1999; Ramanathan, Chung, Kim, Bettge, Buja, Kiehl, Washington, Fu, Sikka and Wild, 2005). A study based on a small number of annual observations (twelve) from a research station in the Philippines reported that the yield of irrigated rice decreased by 10% for each 1°C increase in T_{min} averaged over the growing season (Peng, Huang, Sheehy, Laza, Visperas, Zhong, Centeno, Khush and Cassman, 2004). A reanalysis of the data from that study concluded that the actual impact of T_{min} was much smaller, because T_{min} was negatively correlated with radiation, thus confounding the observed impact of T_{min} with the omitted impact of radiation (Sheehy, Mitchell and Ferrer, 2006). A recent review of the impacts of climate change on rice concluded that "the effect of high night temperature is not understood well" (Wassmann, Jagadish, Heuer, Ismail, Redona, Serraj, Singh, Howell, Pathak and Sumfleth, 2009b).

While experimental studies are essential for understanding physiological relationships and constructing crop simulation models, they do not necessarily replicate real agricultural settings. Researchers typically apply agronomically optimal levels of inputs that are not being investigated, which can accentuate the impact of weather by making it the factor that limits yield. Data from farmer-managed fields allow one to study how weather affects yield in a setting in which farmers make decisions based on the weather they observe every day and the prices they pay for inputs and receive for harvested crops. Although other studies have used non-experimental data to study the relationship between weather (or climate) and agriculture (Mendelsohn, Nordhaus and Shaw, 1994; Schlenker, Hanemann and Fisher, 2005; Deschenes and Greenstone, 2007; Schlenker and Roberts, 2009; Guiteras, 2009), including for rice (Lobell and Field, 2007; Lobell, 2007; Auffhammer, Ramanathan and Vincent, 2006), with one exception they have analyzed aggregate data (e.g., national, state, or county), which precludes careful matching of weather variables with farm-specific planting and harvesting dates and crop growth phases. The exception was a farm-level study on rice yield in a single country (Thailand), which presented no detail on statistical results and evidently did not examine diurnal temperature variation or solar radiation (Felkner, Tazhibayeva and Townsend, 2009).

The data analyzed here are from a multi-year (1994-99) study on productivity of intensively managed irrigated rice farms in Asia by the International Rice Research Institute (IRRI) and its partners in six countries (Dobermann and Witt, 2004). The farms were located in seven important rice-growing regions ("sites"; Figure 1.1) in six of the most important countries in terms of contributions to global rice supply. Each site represented an irrigated rice-growing area of more than 200,000 ha, ranging up to several million ha. All were located in inland plains or large river deltas with humid tropical or subtropical climate, with at least two rice crops grown each year. Such double- and triple-crop rice systems in similar climatic conditions occupy a land area of about 24 million ha in Asia, feed about 1.8 billion Asians, and account for 40% of global rice supply (Dobermann and Cassman, 2004). Most of the sites are in areas where monthly average T_{max} is considered to be high (> 33°C) during the reproductive or ripening phase of one of the annual crops (Wassmann, Jagadish, Sumfleth, Pathak, Howell, Ismail, Serraj, Redona, Singh and Heuer, 2009a).

Farms at each site were selected to represent a range of the most common soil types, cropping systems, farm management practices, and farm sizes. They were early adopters of Green Revolution technologies (modern high-yielding varieties adapted to local conditions, irrigation, fertilizers, pesticides, mechanization) and had been under intensive management for decades. They were generally representative of intensively managed irrigated rice farms in their countries in terms of demographics, access to capital, and capital intensity of production (Moya, 2004). We analyzed all the farms with complete data (227 farms; Table 1.1).

Our objective was to determine the relative sensitivity of rice yield to changes in T_{min}, T_{max}, and radiation in a real-world setting, net of any responses (e.g., input adjustments) by farmers to these changes. Our general approach was to regress yield on weather variables and, in some specifications, exogenously determined economic variables, whose inclusion improved the precision of the estimated weather impacts. IRRI and its partners collected data on crop establishment and harvest dates, production inputs, and yields for each farm in each season of each year. They also collected daily weather data from a single monitoring station at each site, which was within 15-20 km of nearly all farms at a site. This detail enabled us to construct farm-specific measures of weather variables defined according to the rice plantŠs three growth phases (for each phase, means for T_{min}, T_{max}, and radiation and sums for rainfall). The fact that the dataset included observations over multiple growing seasons enabled us to use fixed effects to control for unobserved factors that varied across space (i.e., were unique to each farm, such as soil) or time (were common to all farms at a given site in a given season and year, such as ambient CO_2 concentration). The inclusion of these fixed effects increased the likelihood that the impacts we identified were indeed due to temperature and radiation and not to variables omitted from the regression models.

1.2 Results

1.2.1 Data Variability and Correlations

Yield varied substantially in the sample (5182 ś 1468 kg ha-1; range = 288-10838 kg ha-1), as did weather (Table 1.3). An understanding of correlations among the weather variables is important for interpreting the regression results. Three features of the correlation matrix are most notable (Table 1.4): T_{min} and radiation were not highly correlated, unlike in the Philippines study (Peng et al., 2004; Sheehy et al., 2006); both variables were moderately (and positively) correlated with T_{max}; and their correlations with rainfall were smaller in absolute value than their correlations with T_{max}. These features suggest that the dataset affords the possibility to disentangle the impacts of T_{min} and radiation, and that a failure to control for T_{max} could bias estimates of the impacts more than a failure to control for rainfall.

1.2.2 Regression Results

Multiple regression results support these contentions (Table 1.5). Figure 1.2 shows, for different specifications of the regression model, parameter estimates for the temperature and radiation variables. Given the linear specification of the model, the estimates are interpretable as marginal effects: the impact of a one-unit change in a weather variable on yield, holding other variables constant. For example, in the model that included economic variables (Model 5), a 1°C increase in T_{min} during the ripening phase reduced yield by 322.4 kg ha^{-1}.

Impacts varied by growth phase, with the most significant impacts occurring during the ripening and vegetative phases. T_{min}, T_{max}, and radiation had significant impacts ($P < 0.05$) during both of these phases in the more complete model specifications, with the exception of T_{max} during the ripening phase ($P = 0.087$ in Model 5). The lower significance of the latter variable resulted from its correlation with radiation during the same phase. The two variables were jointly significant ($P < 0.001$ in Model 5), however, and excluding either one sharply improved the significance of the other without affecting the signs or general magnitudes of the parameter estimates on the other weather variables (Table 1.6). These additional results imply that yield was significantly affected by both T_{max} and radiation during the ripening phase.

T_{min} and T_{max} had opposite impacts (negative and positive, respectively), while the impact of radiation differed between phases (positive for ripening, negative for vegetative). Differences in parameter estimates between the vegetative and ripening phases were more significant for radiation and T_{min} than for T_{max} (Table 1.7).

Parameter estimates on T_{min} changed only moderately when radiation was added (Figure 1.2), which is expected given the small correlation between the two variables. The addition of T_{max} had a larger influence, causing the T_{min} parameter estimates to increase 1.5-2 times during both the vegetative and ripening phases. This resulted from the combination of the positive correlation of the two temperature variables and their opposing impacts on yield. Including T_{max} was thus necessary to accurately identify the impact of T_{min}. Otherwise, the T_{min} parameter estimate reflected the net impact of both temperature measures and was biased toward zero. In models that included both temperature measures, the absolute value of the (negative) impact of T_{min} differed significantly from the positive impact of T_{max} during the ripening phase (Table 1.7), which indicates that including both measures is more appropriate than including their mean, T_{ave}. This can also be demonstrated by estimating the same models with T_{ave} included instead of T_{min} and T_{max}. Consistent with previous studies (e.g., 28), T_{ave} tended to have negative impacts on yield during the reproductive and ripening phases, but the impacts were highly insignificant in nearly all specifications (Table 1.8).

Rainfall had a significant impact only during the ripening phase. The addition of rainfall affected the parameter estimates less than the addition of T_{max}, which is not surprising in view of the lower correlations of T_{min} and radiation with rainfall than with T_{max}. Correlations of the economic variables with the weather variables were very small (Table 1.3), and so their addition had a negligible impact on the parameter estimates. It mainly increased the precision of the estimates (lower standard errors and P-values).

The different units of the weather variables in Figure 1.2 impede comparison of the variablesŠ impacts on yield. This can be overcome by expressing the marginal effects per standard deviation of the weather variables (Table 1.9). The standard deviations were calculated after removing any variation explained by the fixed effects for farms and site/season-years, as only this residual variation was used to identify the variablesŠ impacts in the regression models. The largest marginal effect thus expressed was for T_{min} during ripening (-174.4 kg ha^{-1}), followed by T_{max} and radiation during the vegetative phase (122.9 kg ha^{-1} , -124.1 kg ha^{-1}). The smallest was for rainfall (68.4 kg ha^{-1}), which is expected given that the farms were irrigated.

The marginal effects in Figure 1.2 came from regression models that did not allow nonlinear responses of yield to weather. To examine the implications of this restriction, we also estimated a quadratic specification. The estimated parameters on the quadratic terms were mostly insignificant, and the marginal effects evaluated at mean values had the same signs as, and were similar in magnitude to, those in Figure 1.2 (Table 1.10).

1.2.3 Joint Impacts

The opposing effects of T_{min} and T_{max} indicate that warming has an ambiguous impact on rice yield. Which effect dominates depends on the magnitudes of not only the effects but also the trends in the two variables. Even if the absolute values of the variablesŠ opposing effects are re not significantly different, as they are not during the vegetative phase (Table 1.7), differences in the variablesŠ trends could still result in a nonzero net impact of warming.

For each site and season, we investigated the joint impact of recent warming trends by summing the products of the marginal effects and corresponding trends in the two temperature variables during the vegetative and ripening phases. Analyzing observed trends instead of hypothetical future ones that might occur under accelerated warming is appropriate because recent trends have been relatively small; combining marginal effects to calculate the joint impact of multiple temperature changes is valid only if the changes themselves are small (i.e., marginal).

We also included radiation during the vegetative and ripening phases in the analysis, in view of concerns about its potentially confounding effect with T_{min}.

The analysis answered the question, "How would yield growth have been affected if observed weather trends had not occurred?" Although our analysis is not the first to examine the impact of recent climate changes on agricultural yields, including for rice (6, 22), to our knowledge it is the first to be based on farm-level data from multiple countries. Data series from the weather stations at the sites were too short to determine trends. Instead, trends in T_{min} and T_{max} were based on a global analysis of ground-station data for 1979-2004, while trends in surface radiation were based on satellite data for 1983-2004 (see "Data and Methods"). Combining temperature data from ground stations with satellite data for radiation provides reliable estimates of weather impacts on crop yields. Trends were determined separately for each quarter of the year (December-February, etc.) and were assigned to seasons and growth phases using site-specific crop calendars (Dobermann and Witt, 2004).

As expected, evidence of warming was stronger at night (Table 1.11). Sixteen of the 28 site-quarters had significant trends ($P < 0.05$) in the case of T_{min}, with 13 being positive, while only 8 site-quarters had significant trends in the case of T_{max}, with all being positive. Significant trends occurred for radiation in 9 cases, with 8 being negative. Significant warming and dimming thus occurred at some sites, but not all.

Table 1.2 shows the joint impacts of these trends. The most obvious result is that the impacts varied substantially between seasons and sites. At most sites, yield would have grown more rapidly during the high-yielding season but less rapidly during the low-yielding season. The absolute values of the joint impacts were relatively large for one or both seasons at most sites, being equivalent to a fifth or more of the actual annual yield trends for the countries where the sites were located. The direction of the joint impact was influenced more by temperature than by radiation: the absolute value of the joint impact of T_{min} and T_{max} exceeded the absolute value of the impact of radiation for eleven of the fourteen season-sites and matched it for one.

1.3 Discussion

The estimated impacts of weather variables reported here are unique in being based on repeated observations from a large number of farmer-managed fields in multiple countries. This data structure enabled us to investigate the simultaneous impacts of multiple weather variables, broken out by growth phase of the rice plant, and to control for unobserved factors that varied across farms and, at the site level, over time. Despite these methodological differences compared to previous studies, our findings corroborate recent ones that T_{min} has a large, negative impact on yield (Peng, Huang, Sheehy, Laza, Visperas, Zhong, Centeno, Khush and Cassman, 2004). Although the mechanisms responsible for the negative impact have yet to be conclusively identified (Wassmann, Jagadish, Heuer, Ismail, Redona, Serraj, Singh, Howell, Pathak and Sumfleth, 2009b), our results could be explained by increased respiration losses during the vegetative phase (Peng et al., 2004) and reduced grain filling duration and endosperm cell size during the ripening phase.

Our finding of a positive impact of T_{max} during the vegetative and ripening phases is perhaps more surprising, as the literature emphasizes the negative impact of elevated T_{max} during all growth phases, due to reduced photosynthesis caused by chloroplast damage (vegetative phase), spikelet sterility caused by reduced pollen production (reproductive), and increased energy consumption caused by higher respiration demand (ripening) (Wassmann, Jagadish, Heuer, Ismail, Redona, Serraj, Singh, Howell, Pathak and Sumfleth, 2009b). It can be explained by the fact that T_{max} within our sample seldom reached the extremes that cause these negative impacts. For example, fewer than 4% of the observations of T_{max} during the reproductive phase in our sample exceeded the frequently cited threshold of 35°C above which spikelet sterility becomes common under humid conditions (Yoshida et al., 1981; Wassmann et al., 2009b). Field trials for rice grown under ambient temperatures have reported a positive impact of T_{max} (Evans and Datta, 1979), and most controlled-environment studies use $29-30$°C as the optimal daytime growing temperature. Mean T_{max} was within or not much above the latter range at most of our sites (Table 1.3).

Although our finding of a negative correlation between yield and radiation during the vegetative phase contrasts with the literatureŠs emphasis on a positive correlation during the ripening phase (which we found, too), there is experimental evidence that yields of some crops can rise if small reductions in total radiation, which is what we measured, coincide with increases in diffuse radiation (14). Other possible explanations include photoinhibition and excessive production of tillers, which could cause mutual shading and reduced panicle size.

Our most important methodological finding is that it is necessary to analyze the impacts of T_{min} and T_{max} jointly. Because these two variables were moderately correlated in our data and had opposing impacts on yield, excluding T_{max} biased parameter estimates for T_{min} in a position direction. Moreover, the absolute values of the impacts of the two variables were significantly different during the ripening phase. Although the absolute values were not significantly different during the vegetative phase, their opposing effects would cancel only if trends in the two variables were identical, but this has not been the case in recent decades. Recent efforts to develop rice simulation models that include both T_{min} and T_{max} are clearly justified. Our results for these two variables differ from those in two recent studies of national rice yield data (Lobell and Field, 2007; Lobell, 2007), which reported that T_{min} and T_{max} (Lobell and Field, 2007), or T_{ave} and diurnal temperature range $(= T_{max}T_{min})$ (Lobell, 2007), had insignificant impacts in most countries during 1961-2002. Our finding of significant impacts is likely due to a combination of reasons, including the larger number of observations in our sample, our ability to define weather variables specific to farms and rice growth phases, and our inclusion of controls for solar radiation and economic variables, which increased the precision of the estimates.

We emphasize that the impact estimates in Figure 1.2 refer to marginal effects of climate changes. They should not be extrapolated to the non-marginal warming that is projected to occur in Asia by the end of the century (Table 1.12), which lies well outside the residual variation in the weather data that was used to identify warming impacts in our regression models (Table 1.9). For moderate warming in coming decades, however, our results imply a net negative impact on yield, because T_{min} is projected to rise more rapidly than T_{max} and the combined negative marginal effects of T_{min} during the vegetative and ripening phases exceed the combined positive effects of T_{max} (Figure 1.2, Table 1.7). Beyond that, the impact would likely become even more negative, as increases in T_{max} would push it out of the optimal growth range and closer to, or beyond, the extremely high levels where it can damage chloroplasts and cause spikelet sterility.

Another caveat is that our estimates refer to just irrigated rice, not all rice, in tropical and subtropical regions of Asia, not all rice-growing areas of the world. The lack of a substantial rainfall impact in our study does not mean that irrigated rice is ultimately unaffected by rainfall. Rainfall is the source of irrigation water at all seven sites, but it had a negligible impact in our sample simply because we analyzed a period when it did not limit irrigation. This could change in the future, as climate models predict that the area of Asia affected by drought will likely increase (IPCC, 2007).

1.4 Data and Methods

1.4.1 Data

The farms were not selected randomly (Dobermann and Witt, 2004), which is one reason we preferred fixed-effects estimates to random-effects estimates. A consequence of the use of fixed effects is that our results do not necessarily generalize to farms outside the sample, but the regression results changed little if we used random effects instead (see below). Each farm had a parcel dedicated to a nutrient management study, but the parcel was small compared to total farm size. We used only data from the remaining area of each farm, which was controlled by the farmer.

Although IRRI collected weather data from just a single station at each site, the staggering of crop establishment and harvest dates across farms created variation in the weather variables within each site even for a given season-year. We included only farms with no more than two days of weather data missing in a particular season. We used standard definitions of the three growth phases of rice in constructing the weather variables: vegetative = crop establishment through 66 days before harvest; reproductive = 31-65 days prior to harvest; and ripening = last 30 days of the growing season, excluding the harvest date. We constructed the temperature and radiation variables as farm-specific means, and the rainfall variables as farm-specific sums, of the daily observations for each phase.

Some recent studies on future agricultural impacts of climate change, most (Deschenes and Greenstone, 2007; Schlenker and Roberts, 2009) but not all (Guiteras, 2009) of which focus on the U.S., have defined temperature variables in other ways, including growing-degree days (GDD) and the number of days in 1-degree temperature bins, with the latter fitted either linearly or with flexible polynomials. We did not use GDD because low temperature rarely constrains rice growth at tropical and subtropical sites. Temperature bins are useful for identifying significant nonlinearities in the relationship between yield and temperature, which is especially important when simulating the impacts of projected large future increases in temperature. We were unable to implement this approach for two reasons: we had too few observations to estimate precisely the large number of parameters involved, and the tails of our temperature distributions were too thin to detect the nonlinearities.ă It is for such reasons that we restricted the simulation analysis to the relatively small climate changes that have already occurred at the sites.ă

Rice price was farm-specific. It reflected variation in the varieties grown, which changed little over time on a given farm, and the quality of the harvested crop. Some farms sold parts of their harvest at different prices; in those cases, the rice price variable was the average of the various prices reported. The wage rate was calculated at the site level by dividing aggregate expenditure on hired labor across farms by the aggregate number of person-days hired. This was done separately for each season in a given year. The price of nitrogen fertilizer was also calculated at the site/season-year level. Nutrient-specific fertilizer prices were generally not available due to the prevalence of compound fertilizers. The price of nitrogen fertilizer was approximated by the corresponding parameter estimate from a regression of total fertilizer expenditure on the total quantities of nutrients applied (i.e., an implicit price). The use of uniform wages and fertilizer prices across farms at a given site in a given season is reasonable because the farms at each site were located in villages adjacent to each other and were well served by transportation infrastructure.

1.4.2 Regression analysis

We used multiple regression to estimate the following statistical model,

$$y_{it} = c_i + \theta_{jt} + w_{it}\beta + u_{it} \tag{1.1}$$

where y_{it} is the yield of farm i in season-year t; c_i is a farm-level fixed effect, which equals 1 for observations from farm i and 0 otherwise; θ_{jt} is a site-specific season-year fixed effect, which equals 1 for observations from site j in season-year t and 0 otherwise; w_{it} is an NK matrix of weather variables, where N is the number of observations across farms and season-years and K is the number of variables; β is a $K1$ vector of parameters that give the impact of weather; and u_{it} is a random error term that represents the impacts of factors other than weather on yield. (We discuss the inclusion of economic variables below.) Because the model included farm-level fixed effects, the impacts of climate change were identified from the random variation in weather over time as opposed to the mean differences between farms. This identification strategy has been used in other recent studies (Deschenes and Greenstone, 2007; Schlenker and Roberts, 2009; Guiteras, 2009; Auffhammer, Ramanathan and Vincent, 2006; Felkner, Tazhibayeva and Townsend, 2009).

We used a Box-Cox transform to guide model specification. The estimate of the Box-Cox theta parameter for a model with the same variables as Model 5 in Table 1.5 was 0.886, which implied that a linear specification was more appropriate than log-log, semi-log, or inverse specifications.

We intentionally excluded from the right-hand side of Equation 1.1 any variable over which farmers had control, which could have caused endogeneity bias. As a result, the parameters in β are more inclusive than the marginal effects of weather that would be obtained from a regression model that controlled for farm inputs such as labor and fertilizer. To see this, suppose that instead of Equation 1.1 the model were

$$y_{it} = a_i + \theta_{jt} + w_{it}\alpha + \gamma z_{it} + \varepsilon_{it} \tag{1.2}$$

The key change is the addition of z_{it}, which is a farmer-controlled input (an $N1$ vector) whose impact on yield is given by the parameter γ. The farm-level fixed effects are now given by a_i, the parameters on w_{it} by α, and the error term by ε_{it}. Suppose further that farmersŠ decisions about how much of the input to use are affected by weather in the following way:

$$z_{it} = \delta_{0i} + w_{it}\delta_1 + \xi_{it} \tag{1.3}$$

δ_{0i} and δ_1 are parameters, and ξ_{it} is a random error term. Inserting Equation 1.3 into Equation 1.2 yields

$$y_{it} = c_i + \theta_{jt} + w_{it}\beta + u_{it} \tag{1.4}$$

where $c_i = a_i + \gamma\delta_{0i}$, $\beta = \alpha + \gamma\delta_1$, and $u_{it} = \varepsilon_{it} + \gamma\xi_{it}$. This is the same as Equation 1.1. Hence, the expected value of estimates of β obtained by regressing yield on just the weather variables is the total marginal effect of weather on yield: the sum of the direct impact on yield (α) and the indirect impact through weatherŠs influence on input use ($\gamma\delta_1$).

The addition of exogenous economic variables does not fundamentally change the preceding explanation of the parameter estimates on the weather variables. According to standard producer theory, input demand by farmers is determined by not only weather but also crop price (rice in our model), prices of inputs (labor, nitrogen), and stocks of fixed inputs (area planted with rice). The exclusion of these variables from Equation 1.3, and thus from Equation 1.4, can bias estimates of β when the variables are significantly correlated with the weather variables. When the correlations are small, however, as they are in our dataset, then the bias is small and exclusion of these variables mainly makes estimates of β less precise. (Hence, P-values for the temperature and radiation variables in Table 1.4 are larger for Model 4 than Model 5.) Consistent estimates also require that the economic variables are not simultaneously determined with yield. This condition was met in our data: farmers were price takers in rice, labor, and fertilizer markets, and area planted was determined months before each seasonŠs crop was harvested.

The panel structure of our data (i.e., both cross-sectional and time-series variation) allowed the estimation of models that included either fixed effects or random effects to control for unobserved farm characteristics. We used the generalized form of the Hausman test to test the validity of the random effects model. We rejected the null that the regressors were uncorrelated with the farm-level random effects ($P < 0.0001$ in all cases). The random-effects estimates did not differ greatly from the fixed effects estimates (Table 1.13), however, which implies that the correlation of the random effects with the regressors did not bias the parameter estimates greatly. The results remained similar if we did not include either fixed or random effects to control for unobserved farm characteristics, but they changed substantially if we excluded fixed effects for site/season-years (Table 1.13). Evidently, the most influential unobserved effects in our sample were ones that varied over time at the sites. This suggests that parameter estimates from future studies that use cross-sectional, farm-level data instead of panel data might not be very biased if the data are from multiple sites and the regression analysis includes site-level fixed effects.

Residuals in the models could be spatially correlated across farms within a site and serially correlated over time despite the inclusion of the site-specific season-year fixed effects. We addressed this issue by clustering the standard errors at the village/district level. The number of clusters was relatively small (just 32), which could cause the standard errors to be inconsistent. To check this, we implemented a recently developed bootstrapping method for estimating consistent t-statistics when the number of clusters is small. Parameter estimates on T_{min}, radiation, and T_{max} during the vegetative phase and T_{min} during the ripening phase remained significant ($P < 0.05$) according to the bootstrapped t-statistics, but the parameter estimates for T_{max} and radiation during the ripening phase did not ($P = 0.118$ and 0.148, respectively). Each of the latter two variables became significant ($P = 0.002$ and 0.02, respectively), however, if the other was excluded.

1.4.3 Analysis of joint impacts

Estimated quarterly trends in T_{min} and T_{max} ($^{\circ}C\ yr^{-1}$) during 1979-2004 were provided by the U.S. National Climatic Data Center and were generated using the methods described in Vose et al. (2005). They referred to 5°5° grid cells containing the sites. Trends in surface radiation were based on analysis of the series, "Insolation on Horizontal Surface" (MJ $m^{-2}\ day^{-1}$), from the NASA Climatology Resource for Agroclimatology website (http://earth-www.larc.nasa.gov/cgi-bin/cgiwrap/solar/agro.cgi?email=agroclim@larc.nasa.gov). Daily data for this series were downloaded by entering the latitude, longitude, and elevation of each site. Data were averaged within each quarter of the year during 1983-2004 (1983 was the first year in the dataset), and then the natural logarithm of each quarterly series for each site was regressed on an annual time trend. Hence, the radiation trends were expressed in % change yr^{-1}. Significance was tested using Newey-West standard errors, which were robust to heteroskedasticity and first-order serial correlation.

Impacts in Table 1.1 were calculated by multiplying: (i) temperature trends (Table 1.11) by the corresponding regression coefficients (Model 5 in Figure 1.2 and Table 1.5), and (ii) radiation trends by not only the corresponding regression coefficients but also the means of the radiation variables used in the regression analysis. National yield trends were estimated by regressing the natural logarithm of national yield data (FAOStat, http://faostat.fao.org/default.aspx), on an annual time trend. Impacts changed little if they were based on weather trends that were significant at $P < 0.1$ instead of $P < 0.05$.

1.5 Acknowledgments

We thank all the researchers who participated in the project, “Reversing Trends of Declining Productivity in Intensive Irrigated Rice Systems,” which generated the data we analyzed, and the funders of that project (International Rice Research Institute, IRRI; Swiss Agency for Development and Cooperation; International Fertilizer Industry Association; Potash and Phosphate Institute; International Potash Institute). We also thank A. Rala (IRRI) for preparing Figure 1; R. Vose (U.S. National Climatic Data Center) for providing estimates of historical temperature trends; W. Li (Duke University) for providing temperature and radiation projections; seminar participants at the Food and Agriculture Organization of the United Nations (FAO), the University of California at San Diego, and IRRI (especially S.V.K. Jagadish, S. Peng, R. Wassmann) for their suggestions; and FAO and the University of CaliforniaŠs Institute on Global Conflict and Cooperation for partial financial support of the preparation of this paper.

Chapter 1, in full, is a reprint of the material as it appears the Proceedings of the National Academy of Sciences, vol. 107, no. 33, 2010. Welch, Jarrod R.; Vincent, Jeffrey R.; Auffhammer, Maximilian; Moya, Piedad F.; Dobermann, Achim; and Dawe, David. The dissertation author shared equally in the research design and analysis, and is lead author of the paper.

1.6 Figures

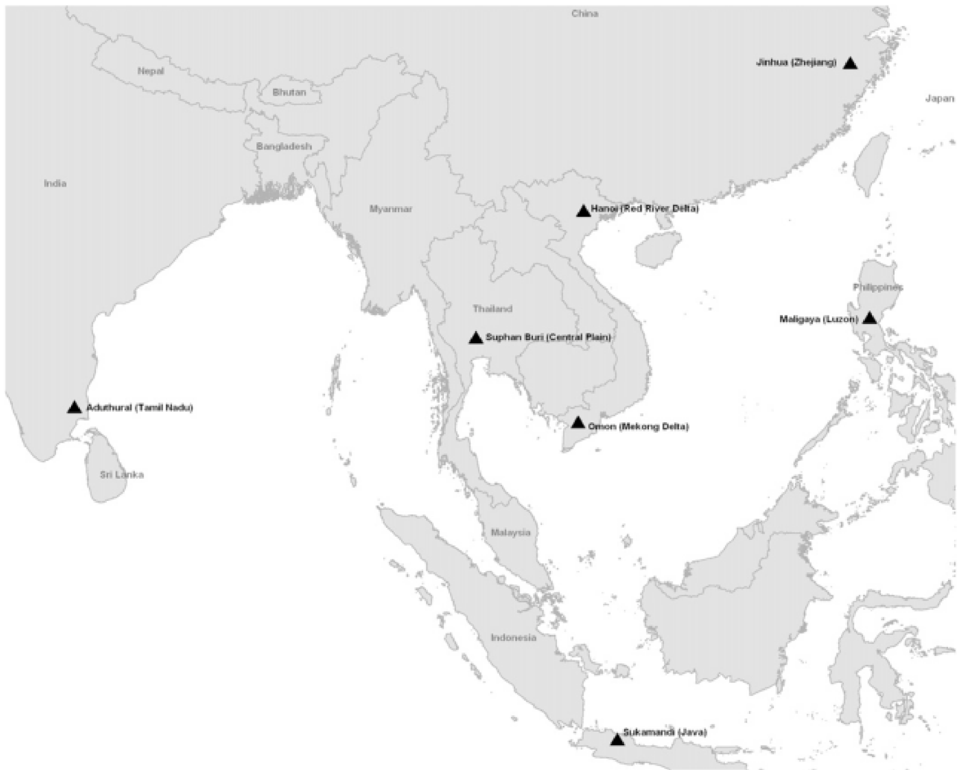

Figure 1.1: Location of study sites

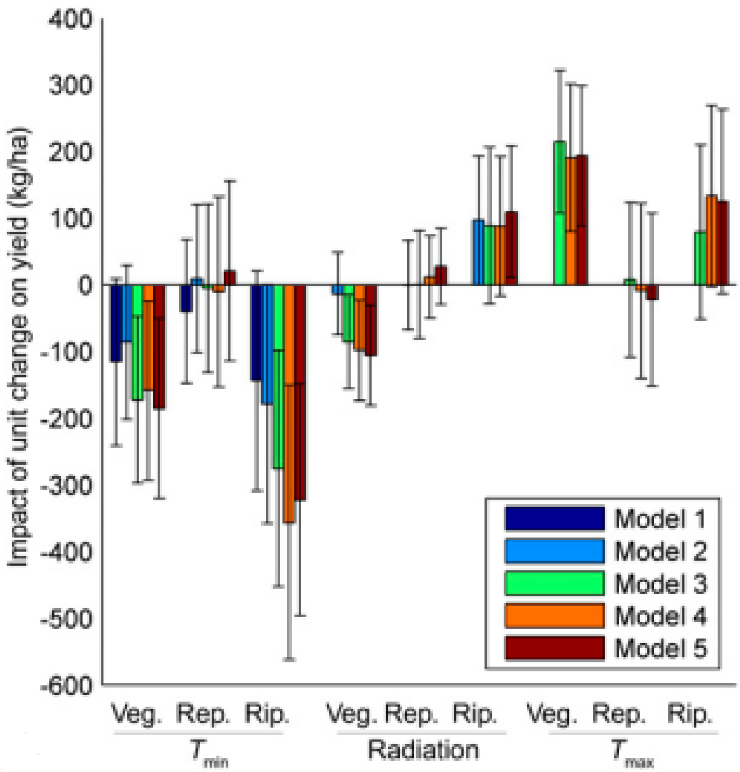

Figure 1.2: Impacts of temperature and radiation on rice yield, expressed per °C for temperature and per MJ m^{-2} d^{-1} for radiation. Each cluster shows estimates for a given variable from different regression-model specifications, distinguished by rice-growth phase (vegetative, reproductive, and ripening). Model 1 included only T_{min}. Model 2 added radiation. Model 3 added T_{max}. Model 4 added rainfall. Model 5 added economic variables. Bars show 95% confidence intervals.

1.7 Tables

Table 1.1: Characteristics of study sites: locations of weather stations, and observations included in sample

Site	Lat/Long	Climatic zone	Years	Villages	Farms	Obs.
China (Jinhua)	29°5’ N 119°47’ E	Subtropics	1998-99	7	25	91
India (Aduthurai)	11°1’ N 79°29’ E	Tropics (subhumid)	1994-99	4	37	214
Indonesia (Sukamandi)	6°21’ S 107°40’ E	Tropics (humid)	1995-99	3	30	159
Philippines (Maligaya)	15°23’ N 120°54’ E	Tropics (humid)	1994-99	4	48	361
Thailand (Suphan Buri)	14°28’ N 100°10’ E	Tropics (subhumid)	1994-98	3	31	169
Vietnam (Hanoi)	21°1’ N 105°53’ E	Subtropics	1997-99	3	24	144
Vietnam (Omon)	10°08’ N 105°32’ E	Tropics (subhumid)	1994-99	8	32	234
Totals	-	-	-	32	227	1372

Sample: observations in regression models that included only weather variables.
Lat and long refer to weather stations at the sites.

Table 1.2: Predicted changes in annual growth rate of rice yield if observed weather trends at end of 20th Century had not occurred at each site

	Predicted change in yield growth (kg ha^{-1} yr^{-1}) resulting from elimination of trend in:				Net impact relative to mean observed yield	Observed national yield growth rate
Season and site	T_{min}	T_{max}	Radiation	Net impact	for site/season (% yr^{-1})	(annual, not seasonal) (% yr^{-1})
High-yielding						
China (Jinhua)	0	0	0	0	0	1.48
India (Aduthurai)	6.3	-12.6	4.1	-2.2	-0.04	2.05
Indonesia (Sukamandi)	12	0	0.3	12.3	0.22	1.13
Philippines (Maligaya)	21.8	0	-2.3	19.5	0.31	1.51
Thailand (Suphan Buri)	11.7	0	0	11.7	0.23	1.55
Vietnam (Hanoi)	16.6	0	6.2	22.8	0.38	3.18
Vietnam (Omon)	-9	0	-10.7	-19.8	-0.33	3.18
Low-yielding						
China (Jinhua)	9.8	-14	0	-4.1	-0.07	1.48
India (Aduthurai)	7.2	-10.7	0	-3.5	-0.07	2.05
Indonesia (Sukamandi)	7.1	0	0	7.1	0.17	1.13
Philippines (Maligaya)	12.5	0	0	12.5	0.32	1.51
Thailand (Suphan Buri)	0	-10.9	-6.8	-17.7	-0.35	1.55
Vietnam (Hanoi)	0	0	-11.8	-11.8	-0.22	3.18

Second through fourth columns show annual changes in yield growth due to elimination of trends in individual weather variables, summed across rice growth phases. Fifth column ("Net impact") shows sum of these changes. Sixth column shows net impact expressed as change in annual growth rate. Seventh column shows observed growth rate in rice yield at national level (not site level) for both seasons combined (seasonal data were not available). Time periods for estimating trends: temperature and observed growth rate, 1979-2004; radiation, 1983-2004.

Table 1.3: Characteristics of study sites: means (SDs) of yield, temperature, and solar radiation

	Yield	T_{min} (°C)			T_{max} (°C)			Radiation (MJ m^{-2} day^{-1})		
Site	(kg ha^{-1})	Veg	Rep	Rip	Veg	Rep	Rip	Veg	Rep	Rip
China	6287.7	22.1	21.1	20.9	30.2	28.7	28.2	13.7	12.5	12.0
	(1257.3)	(3.6)	(1.0)	(3.9)	(3.7)	(1.2)	(3.6)	(8.5)	(6.8)	(8.6)
India	5450.5	24.6	23.4	22.7	33.1	31.7	31.6	17.4	17.3	18.3
	(872.8)	(1.3)	(2.2)	(2.1)	(2.9)	(3.0)	(2.2)	(1.8)	(1.6)	(1.5)
Indonesia	4852.4	24.1	23.5	23.0	31.6	31.3	31.4	17.7	18.0	19.0
	(1614.2)	(0.7)	(0.8)	(1.0)	(1.1)	(0.9)	(1.0)	(2.7)	(2.3)	(2.9)
Philippines	4890.5	23.0	22.7	22.7	30.8	31.0	31.8	21.0	22.4	22.3
	(1858.6)	(1.4)	(1.4)	(1.1)	(1.2)	(1.0)	(1.0)	(3.9)	(3.5)	(3.8)
Thailand	5056.1	23.4	23.8	23.8	33.3	33.8	34.1	16.7	18.1	18.6
	(923.6)	(2.1)	(1.4)	(1.7)	(1.3)	(1.5)	(2.3)	(1.7)	(1.5)	(1.9)
Vietnam (Hanoi)	5699.3	21.7	23.7	24.7	27.6	29.5	31.3	13.6	16.7	17.1
	(1153.8)	(4.7)	(2.7)	(1.1)	(5.7)	(3.4)	(1.1)	(3.9)	(1.3)	(1.3)
Vietnam (Omon)	4951.1	23.3	23.0	22.8	32.1	31.2	31.1	21.7	21.4	22.2
	(1362.1)	(1.3)	(1.3)	(1.5)	(1.4)	(1.6)	(1.1)	(3.7)	(3.5)	(3.6)

See Table 1.1 for numbers of observations. Standard deviations are in parentheses under the means.

Table 1.4: Correlations among weather variables, by rice growth phase

Phase	Variable	T_{min}	T_{max}	Radiation
Vegetative	T_{max}	0.399 (0.000)	-	-
	Radiation	0.117 (0.000)	0.597 (0.000)	-
	Rainfall	0.034 (0.215)	-0.284 (0.000)	-0.324 (0.000)
Reproductive	T_{max}	0.567 (0.000)	-	-
	Radiation	-0.118 (0.000)	0.396 (0.000)	-
	Rainfall	0.383 (0.000)	-0.053 (0.052)	-0.337 (0.000)
Ripening	T_{max}	0.505 (0.000)	-	-
	Radiation	0.055 (0.043)	0.412 (0.000)	-
	Rainfall	0.358 (0.000)	-0.158 (0.000)	-0.250 (0.000)

To match the regression models, correlations were calculated using residual variation in the variables, after demeaning them by farm and site/season-years to remove the fixed effects of unobserved factors unique to each farm or common to all farms at a given site in a given season-year. Number of observations: 1372. P-values are in parentheses.

Table 1.5: Regression results: Impacts of weather and economic variables on rice yield (kg ha^{-1}), for model specifications in Figure 1.2

Variables	Model 1	Model 2	Model 3	Model 4	Model 5
T_{min}: veg	-116.0*	(86.1)	-172.4**	-158.5**	-185.2**
	(0.1)	(0.2)	(0.0)	(0.0)	(0.0)
T_{min}: rep	(40.1)	8.9	(5.5)	(11.0)	20.5
	(0.5)	(0.9)	(0.9)	(0.9)	(0.8)
T_{min}: rip	-143.8*	-178.3*	-275.4***	-356.3***	-322.4***
	(0.1)	(0.1)	(0.0)	(0.0)	(0.0)
Radiation: veg		(13.7)	-85.75**	-98.17**	-106.1***
		(0.7)	(0.0)	(0.0)	(0.0)
Radiation: rep		(0.5)	0.5	11.4	27.4
		(1.0)	(1.0)	(0.7)	(0.4)
Radiation: rip		96.67*	88.7	88.0	109.4**
		(0.1)	(0.2)	(0.1)	(0.0)
T_{max}: veg			214.9***	190.4***	193.9***
			0.0	(0.0)	(0.0)
T_{max}: rep			7.4	(9.4)	(22.3)
			(0.9)	(0.9)	(0.7)
T_{max}: rip			79.1	133.2*	124.9*
			(0.2)	(0.1)	(0.1)
Rainfall: veg				(0.6)	(0.5)
				(0.1)	(0.1)
Rainfall: rep				0.4	0.3
				(0.6)	(0.7)
Rainfall: rip				1.325*	1.284**
				(0.1)	(0.0)
ln(Farm size)					-779.5***
					0.0
Rice price/Wage					7073**
					(0.0)
Rice price/N price					348.6
					(0.6)
R^2	0.5	0.6	0.6	0.6	0.6
Observations	1372.0	1372.0	1372.0	1372.0	1248.0
Number of farms	227.0	227.0	227.0	227.0	219.0

All models included fixed effects for farms and site/season-years, in addition to variables shown in table. Units for explanatory variables: °C for T_{min} and T_{max}, MJ m^{-2} day^{-1} for radiation, mm for rainfall, and ha for farm size. Robust P-values are in parentheses, for standard errors clustered by village;
***P < 0.01, **P < 0.05, *P < 0.1.

Table 1.6: Regression results: Impacts of weather and economic variables on rice yield (kg ha^{-1}) in Model 5 when excluding either solar radiation or T_{max} during the ripening phase

Variables	Full model (Model 5)	Exclude Radiation: rip	Exclude T_{max}: rip
T_{min}: veg	-185.2**	-200.2***	-185.9**
	(0.0)	(0.0)	(0.0)
T_{min}: rep	20.5	21.9	22.3
	(0.8)	(0.7)	(0.8)
T_{min}: rip	-322.4***	-328.8***	-248.2***
	(0.0)	0.0	(0.0)
Radiation: veg	-106.1***	-126.8***	-91.83**
	(0.0)	(0.0)	(0.0)
Radiation: rep	27.4	15.8	37.5
	(0.4)	(0.6)	(0.2)
Radiation: rip	109.4**		138.5***
	(0.0)		(0.0)
T_{max}: veg	193.9***	190.3***	180.6***
	(0.0)	(0.0)	(0.0)
T_{max}: rep	(22.3)	(60.7)	(15.5)
	(0.7)	(0.4)	(0.8)
T_{max}: rip	124.9*	209.5***	
	(0.1)	(0.0)	
Rainfall: veg	(0.5)	-0.743**	(0.5)
	(0.1)	(0.0)	(0.2)
Rainfall: rep	0.3	0.0	0.3
	(0.7)	(1.0)	(0.7)
Rainfall: rip	1.284**	1.117*	0.9
	(0.0)	(0.1)	(0.1)
ln(Farm size)	-779.5***	-741.8***	-799.7***
	0.0	0.0	0.0
Rice price/Wage	7073**	6860*	6823*
	(0.0)	(0.1)	(0.1)
Rice price/N price	348.6	380.1	441.1
	(0.6)	(0.6)	(0.5)
R^2	0.6	0.6	0.6
Observations	1248	1248	1248
Number of farms	219	219	219

All models included fixed effects for farms and site/season-years, in addition to variables shown in table. Units for explanatory variables: °C for T_{min} and T_{max}, MJ m^{-2} day^{-1} for radiation, mm for rainfall, and ha for farm size. Robust P-values are in parentheses, for standard errors clustered by village;
***P < 0.01, **P < 0.05, *P < 0.1.

Table 1.7: Equality tests for regression parameters in Models 4 and 5 in Table 1.5

Null hypothesis	P-values	
	Model 4	Model 5
Equality across rice growth phases		
T_{min}: veg = T_{min}: rep	0.12	0.02
T_{min}: veg = T_{min}: rip	0.04	0.11
T_{min}: rep = T_{min}: rip	0.01	0.01
T_{min} equal for all 3 phases	0.04	0.02
Radiation: veg = Radiation: rep	0.01	0
Radiation: veg = Radiation: rip	0.01	0
|Radiation: veg| = Radiation: rip	0.87	0.96
Radiation: rep = Radiation: rip	0.14	0.12
Radiation equal for all 3 phases	0.03	0
T_{max}: veg = T_{max}: rep	0.04	0.02
T_{max}: veg = T_{max}: rip	0.47	0.35
T_{max}: rep = T_{max}: rip	0.2	0.19
T_{max} equal for all 3 phases	0.11	0.05
Rainfall: veg = Rainfall: rep	0.19	0.25
Rainfall: veg = Rainfall: rip	0.04	0.02
Rainfall: rep = Rainfall: rip	0.24	0.21
Rainfall equal for all 3 phases	0.13	0.06
Equality within growth phases		
$\|T_{min}$: veg$\|$ = T_{max}: veg	0.58	0.9
$\|T_{min}$: rip$\|$ = T_{max}: rip	0.03	0.04

Table 1.8: Regression results: Impacts of weather and economic variables on rice yield (kg ha^{-1}), for models that included T_{ave} instead of T_{min} and T_{max}

Variables	T_{ave} only	Add radiation	Add rainfall	Add economic variables
T_{ave}: veg	-22.64	43.29	31.74	10.05
	(0.8)	(0.4)	(0.6)	(0.9)
T_{ave}: rep	-115.0**	-38.22	-54.31	-34.26
	(0.0)	(0.6)	(0.4)	(0.6)
T_{ave}: rip	-45.77	-153.9	-163.4*	-142.9
	(0.6)	(0.1)	(0.1)	(0.1)
Radiation: veg		-12.43	-27.95	-30.32
		(0.7)	(0.3)	(0.3)
Radiation: rep		18.7	19.03	28.22
		(0.6)	(0.5)	(0.3)
Radiation: rip		121.1**	118.3**	142.5***
		(0.0)	(0.0)	(0.0)
Rainfall: veg			-0.820**	-0.689**
			(0.0)	(0.0)
Rainfall: rep			-0.13	-0.12
			(0.9)	(0.9)
Rainfall: rip			0.44	0.47
			(0.5)	(0.4)
ln(Farm size)				-816.1***
				0.0
Rice price/Wage				7199**
				(0.1)
Rice price/N price				559.2
				(0.4)
R^2	0.55	0.55	0.55	0.59
Observations	1372	1372	1372	1248
Number of farms	227	227	227	219

All models included fixed effects for farms and site/season-years, in addition to variables shown in table. Units for explanatory variables: °C for T_{min} and T_{max}, MJ m^{-2} day^{-1} for radiation, mm for rainfall, and ha for farm size. Robust P-values are in parentheses, for standard errors clustered by village;
***$P < 0.01$, **$P < 0.05$, *$P < 0.1$.

Table 1.9: Marginal effects of weather variables, expressed per standard deviation: regression model that included economic variables

Variable	Growth phase	Standard deviation, based on residual variation	Marginal effect ($kg\ ha^{-1}$ per standard deviation)
T_{min}	Vegetative	0.42 °C	-78.6
	Ripening	0.54 °C	-174.4
Radiation	Vegetative	1.17 $MJ\ m^{-2}\ day^{-1}$	-124.1
	Ripening	0.90 $MJ\ m^{-2}\ day^{-1}$	98.2
T_{max}	Vegetative	0.63°C	122.9
	Ripening	0.55°C	69
Rainfall	Ripening	53.3 mm	68.4

Marginal effects are shown only for weather variables whose estimated regression parameters were significant at $P < 0.1$ (see Model 5 in Table 1.5) and were calculated by multiplying regression parameters by the standard deviations of the corresponding weather variables. Standard deviations refer to residual variation after removing variation explained by fixed effects for farms and site/season-years.

Table 1.10: Quadratic model: significance of linear and quadratic terms, marginal effects, and turning points

Variable	Growth phase	P-values of variables:			Marginal	Turning	Impact when variable
		Linear	Quadratic	Joint	effect	point	exceeds turning point
T_{min}	Vegetative	0	0	0	-225.4	21.5°C	Yield falls
	Reproductive	0.26	0.26	0.53	25.3	22.7°C	Yield rises
	Ripening	0.99	0.79	0	-343.4	*	Yield falls
Radiation	Vegetative	0.26	0.11	0.08	-77.5	12.8 MJ/m2/day	Yield falls
	Reproductive	0.06	0.02	0.02	29.7	17.5 MJ/m2/day	Yield rises
	Ripening	0.82	0.82	0.06	107.4	*	Yield rises
T_{max}	Vegetative	0	0	0	139.6	29.9°C	Yield rises
	Reproductive	0.98	0.99	0.73	-56.5	7.7°C	Yield falls
	Ripening	0.23	0.3	0.14	122.9	37.3°C	Yield falls
Rainfall	Vegetative	0.39	0.59	0.34	-0.66	768 mm	Yield rises
	Reproductive	0.12	0.21	0.29	0.84	341 mm	Yield falls
	Ripening	0.11	0.52	0.05	1.8	548 mm	Yield falls

Model also included economic variables and fixed effects for farms and site/season-years. Significance of individual variables was based on robust standard errors clustered by villages/districts; joint significance of linear and quadratic terms was determined using F tests. Marginal effects were calculated at means of the weather variables and can be compared to parameter estimates for the corresponding weather variables in Model 5 in Table 1.5. An asterisk denotes a monotonic function over the range of positive values.

Table 1.11: Observed weather trends at study sites at end of 20th Century, by quarter of the year

	T_{min} (°C decade^{-1})				T_{max} (°C decade^{-1})				Surface radiation (% yr^{-1})			
Site	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
China	0.62	0.53	0.15	0.28	0.58	0.72	0.04	0.32	-0.09%	0.21%	-0.59%	0.25%
India	0.19	0.1	0.26	0.39	0.39	0.09	0.42	0.3	0.31%	0.16%	0.43%	-0.18%
Indonesia	0.39	0.19	0.22	0.23	0.13	0.05	0.11	0.05	-0.24%	-0.15%	-0.17%	-0.10%
Philippines	0.55	0.36	0.24	0.25	0.05	-0.17	0.02	0.12	-0.51%	-0.36%	-0.15%	-0.23%
Thailand	0.63	0	0.1	0.12	0.16	-0.06	0.38	0.28	-0.35%	-0.30%	-0.40%	-0.18%
Vietnamă (north)	0.59	0.44	0.1	0.21	0.13	-0.38	0	0.23	-0.21%	-0.25%	-0.65%	0.18%
Vietnam (south)	0.03	-0.28	-0.32	-0.34	0.27	0.11	0.46	0.38	-0.50%	-0.22%	-0.29%	-0.10%

DJF = December-February, MAM = March-May, JJA = June-August, SON = September-November. Numbers in bold indicate trends that were significant at $P < 0.05$. Time periods for estimating trends: temperature, 1979-2004; radiation, 1983-2004.

Table 1.12: Projected changes in decadal means of weather variables at study sites between 1991-2000 and 2091-2100, by quarter of the year

	T_{min} (°C decade^{-1})				T_{max} (°C decade^{-1})				Surface radiation (% yr^{-1})			
Site	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
China	3	1.29	1.96	3.44	3.77	1.58	2.08	3.01	-0.07	0.11	-0.5	-2.85
India	2.16	3.56	3.58	3.09	2.08	4.16	4.39	3.36	1.4	-6.99	-7.64	-7.84
Indonesia	2.27	2.43	2.46	2.48	2	2.03	2.11	2.4	-11.88	-3.82	-8.33	-6.28
Philippines	2.97	3.31	3.28	3.13	0.32	3.03	3.24	2.79	-28.62	-16.94	-15.3	-16.38
Thailand	2.16	3.46	3.2	2.96	1.37	3.25	3.08	2.62	-8.86	-14.97	-15.43	-17.76
Vietnamă (north)	1.83	2.25	2.84	2.98	1.28	2.4	2.72	2.46	-7.93	-13.5	-9.4	-4
Vietnam (south)	2.06	3.02	2.92	2.8	1.17	2.95	2.75	2.6	-8	-16.61	-10.43	-11.91

Projections are from climate simulations conducted for the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. They refer to scenario A1B, which was the scenario used in the Report's regional climate projections (11), and were generated by the Geophysical Fluid Dynamics Laboratory (GFDL) model.
DJF = December-February, MAM = March-May, JJA = June-August, SON = September-November. Numbers in bold indicate trends that were significant at $P < 0.05$. Time periods for estimating trends: temperature, 1979-2004; radiation, 1983-2004.

Table 1.13: Regression results: alternative specifications of fixed effects (FE)

Variables	FE: farm, site/season-year	RE: farm FE: site/season-year	FE: site/season-year	FE: farm	FE: none
T_{min}: veg	-185.2**	-170.3**	-155.4*	-89.92	-92.68
	(0.0)	(0.0)	(0.1)	(0.3)	(0.3)
T_{min}: rep	20.49	35.71	83.42	-9.13	4.89
	(0.8)	(0.7)	(0.4)	(0.9)	(1.0)
T_{min}: rip	-322.4***	-373.0***	-381.6***	-160.7	-92.82
	(0.0)	0.0	(0.0)	(0.1)	(0.4)
Radiation: veg	-106.1***	-132.0***	-147.6***	-53.84*	-95.75***
	(0.0)	0.0	0.0	(0.1)	(0.0)
Radiation: rep	27.38	11.27	-1.2	-3.64	-11.84
	(0.4)	(0.7)	(1.0)	(0.9)	(0.8)
Radiation: rip	109.4**	133.9**	188.5**	46.77	71.30*
	(0.0)	(0.0)	(0.0)	(0.2)	(0.1)
T_{max}: veg	193.9***	215.5***	231.2***	147.4**	150.2**
	(0.0)	0.0	0.0	(0.0)	(0.0)
T_{max}: rep	-22.27	6.85	29.86	-77.87	-99.67
	(0.7)	(0.9)	(0.7)	(0.3)	(0.3)
T_{max}: rip	124.9*	98.73	-2.3	75.98	4.1
	(0.1)	(0.3)	(1.0)	(0.4)	(1.0)
Rainfall: veg	-0.53	-0.842**	-1.010**	-0.836*	-1.549***
	(0.1)	(0.0)	(0.0)	(0.1)	0.0
Rainfall: rep	0.31	0.26	0.26	-1.662**	-1.176*
	(0.7)	(0.7)	(0.7)	(0.0)	(0.1)
Rainfall: rip	1.284**	1.09	0.58	-0.97	-1.37
	(0.0)	(0.1)	(0.5)	(0.3)	(0.2)
ln(Farm size)	-779.5***	-228.7**	-167.4*	-742.4***	-201.1*
	0.0	(0.0)	(0.1)	(0.0)	(0.1)
Rice price/Wage	7073**	7813**	7794**	359.6	861.3
	(0.0)	(0.0)	(0.0)	(0.8)	(0.6)
Rice price/N price	348.6	857.2	1442*	645.5**	366.2
	(0.6)	(0.2)	(0.1)	(0.0)	(0.3)
R^2	0.6	-	0.51	0.29	0.22
Observations	1248	1248	1248	1248	1248
Number of farms	219	219	219	219	219

Robust P-values are in parentheses, for standard errors clustered by village;
***P < 0.01, **P < 0.05, *P < 0.1.

Chapter 2

Climate change, agriculture, and adaptation: A natural experiment

2.1 Introduction

The relationship between climate, agriculture, and socioeconomic outcomes is subject to significant non-linearities. Agronomists have long understood that crop yield functions are sensitive to high temperature, exhibiting strong negative responses to temperatures that exceed certain crop-specific thresholds. With the exception of Schlenker and Roberts (2009), most economic research on the topic of climate change and agriculture has focused on mean temperature. With increasing evidence that not only are mean temperatures increasing, but we are likely to see increased variability also, we should be particularly concerned about these threshold effects, and how farmers respond to them.

The agronomic literature has significantly advanced the understanding of the physical crop responses to weather variables, typically using simulation and experimental methods that do not necessarily reflect farmer decision processes. Using observational data from the U.S., Schlenker and Roberts (2009) estimate non-linear yield functions for corn and soybeans, two important U.S. crops, but to my knowledge this relationship has not been estimated for rice. Disentangling the relationship between climate change, rice production, and human response is an important issue; rice is an important global food crop contributing up to a third of the daily caloric intake of over three billion people, among them some of the world's poorest. Although the detection of temperature thresholds is relatively straightforward, a complete understanding of the relationship involves understanding farmer responses. Among the difficulties in conducting empirical analysis of this in particular, is that existing farmers are fully adapted to their environments, and the probability of exposure to temperature exceeding the threshold is changing very slowly. One could attempt to identify adaptation strategies using cross-sectional variation, however, there are often many confounding factors. In this paper, I propose and use a natural experiment that effectively assigns farmers to different amounts of exposure to high temperature, at random. Further, individual farmers experience variation in the amount of time their crops are exposed to high temperature from year to year, and in such a way that they are able to anticipate. This allows me to not only estimate the temperature threshold for rice, but also analyze how individual farmers change their behavior when faced with a less favorable climate regime.

Several studies have been conducted analyzing the relationship between temperature and rice, many pointing to the importance of considering exposure to high temperature at night. Seshu and Cady (1984) and Dey and Hossain (1995) both analyze data from International Rice Research Institute (IRRI) test sites, Seshu and Cady using only climate variables as explanatory variables, Dey and Hossain using climate and fertilizer. Both studies find that higher minimum temperature is damaging to yield. Peng, Huang, Sheehy, Laza, Visperas, Zhong, Centeno, Khush and Cassman (2004) find similar results in their analysis of field trials, though they do not conduct a multivariate analysis, they include only one weather variable at a time, leading to a subsequent critique by Sheehy, Mitchell and Ferrer (2006). Peng et al. find that minimum temperatures rose significantly during the 23 years leading up to their study, and that this increase had a significant, negative impact on rice yields. It should be noted that the crops studied were on an IRRI test site in the Philippines, and they were managed to achieve highest possible yield, not following practices of farmers responding to fluctuations in weather and price signals. This makes controlled crop experiments and agronomic simulation methods suboptimal when considering the relevance of behavioral responses. Auffhammer, Ramanathan and Vincent (2006) and Welch, Vincent, Auffhammer, Moya, Dobermann and Dawe (2010) both use observational data, Auffhammer et al. (2006) aggregate state-level data from India, Welch et al. (2010) farm-level data from six Asian countries, and also demonstrate the importance of considering nighttime temperature.

One way the economic literature has tried to account for behavioral responses is with hedonic studies. These studies use rents or farmland values as the dependent variable, regressed on weather variables. Theoretically, in the presence of perfect land markets, farmland rents should reflect the optimized use of the land. Any effect of weather on farmland rents (or value) then, should be interpretable as the total effect, allowing for full behavioral response. In one of the original hedonic studies, Mendelsohn, Nordhaus and Shaw (1994) find that higher temperatures in all seasons except autumn reduce farmland values, but that coupling their findings with global warming scenarios shows a significantly lower estimated impact on agriculture than previous studies that use a production function approach. Dinar, Mendelsohn, Evenson, Parikh, Sanghi, Kumar, McKinsey and Lonergen (1998) find that in India, higher temperatures were harmful to net farm revenues. This type of study is relevant for understanding economic effects of climate change after account for adaptation, but all choice variables are swept into the error term so these studies fail to shed light on any of the adaptation mechanisms.

In a more recent study, Schlenker and Roberts (2009) provide the first example in the economics literature of consideration of the entire temperature distribution, enabling them to isolate the threshold, and estimate the non-linearity of the yield function. On the other hand, means may be appropriate when attempting to model behavioral responses. Kelly, Kolstad and Mitchell (2005) argue that in the context of adaptation, means are the appropriate measure because it is possible for farmers to adjust to long term changes in climate (means), but given the time horizon of many of their decisions (often the entire growing season) it is quite difficult for then to adjust to daily weather fluctuations. In this paper, I show that farmers who plant later in the season experience a significant increase in exposure to high temperature, and use exogenous variation in crop establishment date driven by the availability of irrigation water, to demonstrate that even though farmers can in fact anticipate changes in exposure to high temperature, they do little about it.

Several researchers have studied agricultural adaptation to climate change. Smit and Skinner (2002) discuss a variety of adaptation strategies from technological developments to government programs and farm financial management, using specific examples from Canada. Nath and Behera (2011) discuss adaptation in a more broad, social sense, and argue for the necessity of government intervention and a strong role for social institutions. Seo, Mendelsohn and others conduct a series of cross-sectional, hedonic analyses of adaptation in South America and Africa, focusing primarily on crop choice and diversification (e.g. Seo (2010), Seo and Mendelsohn (2008)). A similar study, specific to rice, is Wang, Mendelsohn, Dinar and Huang (2008). Wu, Shibasaki, Yang, Tang and Sugimoto (2010) use a crop decision model and simulation methods to analyze future changes in rice sown areas in Asia. Wassmann, Jagadish, Sumfleth, Pathak, Howell, Ismail, Serraj, Redona, Singh and Heuer (2009a) analyze the specific physiological effects of climate on rice, and identify adaptation roles for irrigation infrastructure and development of heat tolerant germplasm. To my knowledge, this is the first study that analyzes adaptation using farm-level data that are structured such that we are able to observe individual farmers who knowingly face different climate scenarios from year to year.

In the next section, I provide a theoretical framework within which to discuss the analysis, in Section 3 I present the data and analysis, ending that section with a discussion of the results in the context of projected climate change, and in Section 4 I discuss a few caveats, and provide concluding remarks.

2.2 Theoretical Framework

When estimating the relationship between climate and agricultural and economic outcomes, it is important to consider thresholds. It is well understood from a physiological perspective that plant growth, and therefore grain yield, is initially increasing in temperature until temperature reaches a certain threshold at which point yield falls dramatically with further increases in temperature. Denote by τ this temperature threshold, and let

$$Y_t = F(X_t, g(\tau_t), W_t) \tag{2.1}$$

represent the agricultural production function, where Y_t represents yield for year t, $g(\tau_t)$ is some function of crop exposure to temperature above the threshold, W_t is a vector of average weather, and X_t is a vector of inputs composed of inputs X_{1t} which are fixed at the beginning of each season (e.g. area planted, seed quantity, planting method, some fertilizer), and X_{2t} which are applied throughout the season (e.g. pesticides, herbicides, labor for certain activities, additional fertilizer). Assume that the marginal effect of weather on yield is the only stochastic component of this function.

There are two conceptually distinct ways to quantify the relationship between temperature and yield. The first is simply the direct, marginal effect:

$$\alpha_W = \tfrac{\partial Y}{\partial W}$$

This is only the partial effect of temperature on yield, holding all other inputs constant. I assume that $\alpha_\tau \geq 0$ for $T \leq \tau$ and $\alpha_\tau < 0$ for $T > \tau$.

It is important to remember, however, that agriculture is a coupled human-ecological system, and that γ_W might not capture all factors that comprise this relationship. Alternatively, we can talk about the total effect, allowing for the fact that farmers observe weather, formulate expectations, develop beliefs about climate distributions, and that input choices may in fact be functions of weather expectations and observations. Let

$$\beta_W = \frac{dY}{dW} = \frac{\partial Y}{\partial W} + \frac{\partial Y}{\partial X}\frac{\partial X}{\partial W}$$

In earlier work, we estimate β, and argue that this is the relevant parameter when discussing how agricultural output will change as the climate changes. Little is known about the second term, $\beta - \alpha$ however, and that is the topic of this paper.

The farmer's problem is:

$$\max_X E[U(P_Y F(X_t, g(\tau_t), W_t)] - P_X X_t)] \tag{2.2}$$

With the assumption that weather is the only source of stochasticity[1], the solution can be written as

$$X_1^*(E[g(\tau_t)], E[W_t]) \tag{2.3}$$

and

$$X_2^*(g(\tau_t), W_t) \tag{2.4}$$

Theoretical predictions regarding these relationships are largely ambiguous. Particularly in the context of climate change, we do not know how farmers will react; whether they will continue to farm the same crops in the same locations while taking measures to mitigate the harmful effects of climate change, or adapt by changing crops or moving to cooler areas is largely an empirical question. In the long run, as the probability of temperature exceeding the threshold increases dramatically, farmers will presumably move to cooler areas or plant less temperature sensitive crops. What remains unclear is how large changes will have to be for this to take place. Given the projected changes over the course of this century, how soon should we expect to see changes of this sort?

Without making any further assumptions about functional form, I can take a first-order Taylor approximation to get the following estimation equation:

$$X_t = \delta_1 g(\tau_t) + \delta_2 W_t + u_t \tag{2.5}$$

where $\delta_1 = \frac{\partial X}{\partial g(\tau_t)}$, $\delta_2 = \frac{\partial X}{\partial W}$ and u_t is an error term. In the case of X_{1t}, this is a function of expected weather which is inherently unobservable. In this paper, I exploit a natural experiment that generates exogenous variation in the amount of time a farmer's crop is exposed to temperature over the threshold, and provides information on predicted exposure as of the beginning of the season.

[1] This assumption amounts to ignoring the fact that the spot price at the time of harvest is unknown; the futures price is, however, and the assumption is essentially that the farmer has entered into a futures contract for delivery at the time of harvest of an unknown quantity at a specified price.

2.3 Empirical analysis

2.3.1 Data

The data come from a multi-year study of rice farming conducted by researchers at the International Rice Research Institute (IRRI) and various local agricultural research centers in Asia. Starting in 1994, data were collected from farms in Tamil Nadu, India; Mekong Delta, Vietnam; Central Luzon, Philippines; and the Central Plain of Thailand. The study expanded in subsequent years, and additional farms in the following sites were added: in West Java, Indonesia (1995); Red River Delta, Vietnam (1997); and Zhejiang, China (1998). Data were collected through 1999. Altogether, the data set I use contains data for 227 intensively irrigated farms in 32 different villages, from seven important rice growing regions in six of the most important countries in terms of their contribution to global rice supply. Farms were not chosen at random, but were chosen to be representative of the main cropping systems and soil types of the region, as well as a range of small and large farms. Each farm had a parcel dedicated to the experiment which IRRI was conducting, but this parcel was small with respect to total farm size and the remainder of the farm was controlled by the farmer. I use only data from the farmer controlled plots. Some farms were dropped during the course of the study, but the attrition was generally unrelated to farmer characteristics. Reasons include breakdown of irrigation systems, problems with agronomic data collection (on the part of the experimenters), or project financial constraints.

All farms are irrigated, and most harvest two crops per year, a high-yield season (season HY—typically the dry season) and a low-yield season (season LY—typically the wet season). Detailed household surveys were conducted, collecting information on yield, rice price, and cost and quantity of various inputs including land use, labor (highly disaggregated by sourceŮfamily, hired, etc.Ůand task), pesticides, and fertilizers (Table 2.1 provides summary statistics of a few key variables). Perhaps most importantly, the surveys document farm-specific crop establishment and harvest dates, allowing me to match each farm with specific weather realizations actually observed over the growing season of the crop. In each location, the irrigation authority controls the allocation and timing of water availability. This imposes a constraint on the farmers' choice of crop-establishment date as they cannot plant before their water is available, and storage is not an option. Irrigation water is turned on, and farmers immediately plant and flood their field. This effectively generates exogenous variation is the climate distribution to which each farmer is exposed. To my knowledge, this is the first paper to exploit such year-to-year variation at the farm-level, in the climate distribution (not just weather realizations), to explore how individual farmers adjust their behavior in response to different climate scenarios.

Wages were calculated by dividing total labor cost by quantity of hired labor. Some farms did not hire labor in some seasons, so I calculate the wage at the site level, separately for each growing season (e.g. wet season 1997, dry season 1998, etc.) in order to not lose observations due to a missing wage variable. Often compound fertilizers are used and cost is not nutrient specific, so prices of nitrogen, phosphorus, and potassium fertilizers were approximated by their respective coefficients from a regression of total fertilizer cost on the quantity of each of the nutrients. This was also done at the site-season level. The assumption of constant wages and fertilizer prices for a given site in a given season is reasonable in intensively irrigated areas in Asia because they are well served by infrastructure. Farm profit per hectare is defined as the difference between production value per hectare and total non-land cost per hectare. All economic variables are in constant 1999 US dollar terms.

Weather data come from seven weather stations (one at each site) located such that no farm in a site is more than 25km from the weather station. I use daily observations on minimum and maximum temperature (measured in degrees Celcius). Since all farms in this study are intensively irrigated, contemporaneous rainfall is expected to have little effect on farm-level outcomes and is therefore excluded. Despite the fact that weather observations are recorded by a single weather station for all farms within a particular site, I observe farm-level variation in weather due to the staggered crop-establishment dates.

2.3.2 Results

First, I estimate the threshold. I start with a model motivated by Welch et al. (2010), separately controlling for nighttime (minimum) and daytime (maximum) temperature averages over the three growth phases–vegetative (crop establishment to panicle initiation–66 days before harvest), reproductive (panicle initiation to flowering–30 days before harvest), and ripening (flowering to harvest). Rice is known to be particularly sensitive to heat during a brief period of time around flowering (Wassmann et al., 2009a), so I focus on the two-week window centered at the flowering date (30 days before harvest). Define threshold-degree days around the flowering date as $TDDays\tau_{it} = \sum_{d=1}^{14} T_{idt} - \tau$, where T_{idt} is the maximum temperature observed by farmer i on day d of the window around flowering in season t, and τ is the threshold. I estimate the following equation:

$$Y_{it} = \beta TDDays\tau_{it} + \lambda Tmean_{it} + \theta_i + \nu_{it} \tag{2.6}$$

where Y_{it} is the yield of farm i in season t, $Tmean_{it}$ is a vector of minimum and maximum temperature, averaged over the three growth phases, θ_i is the individual farm fixed effect, and ν_{it} is the error term. I run a series of these regressions, increasing the cutoff for threshold-degree days into the negatively sloped region in 1°C increments, selecting the threshold (32°) that best fits the data; this is my estimate of τ. Given the importance discussed above of considering nighttime temperature effects, I repeat this process for minimum temperature, maximum temperature, and both combined. Included alone, minimum temperature does have a significant threshold above which yield decreases, but when I control for both minimum and maximum temperature, only the coefficient on maximum temperature remains significant. I also test a second specification, one which counts the number of days above the threshold. This treats all days with temperature above the threshold equally, while threshold-degree days assigns greater importance to hotter days. Table 2.3 displays the yield effects of threshold-degree days. Column 1 is the primary model, showing the effect of degree days above 32°C, where the coefficient on $TDDays32$ is the effect in the high-yield season. Focusing on within variation to interpret these results, we see that a one standard deviation change in threshold-degree days ($\sim$ 11), holding mean temperature constant, leads to a reduction in yield of 363 kg/ha in the high-yield season (about 7% at the mean, or 0.3 within SDs). An F-test that the sum of the coefficients on $TDDays32$ and the interaction term is equal to zero results in an p-value of 0.15, so the effect in the low-yield season is statistically insignificant. This could be due to the fact that there is not only a differential threshold effect, but possibly a differential threshold.

Columns 3-5 are included to demonstrate robustness (I discuss the relevance of column 2 below): column 3 shows that controlling for daytime thresholds, nighttime temperature thresholds do not have any additional effect, and controlling for nighttime temperature has little effect on the significance of daytime thresholds; column 4 shows that as I increase the cutoff for the threshold, the fit does not decline dramatically, but the magnitude of the effect increases sharply; column 5 shows that under the alternative specification, counting all days above the threshold equally, we still detect a significant effect.

With respect to farmer responses in preseason input decisions, the identification strategy relies on the assumption that farmers have no control over their planting day, and the extent to which planting day predicts threshold-degree days within the brief window around flowering. First, I create a new variable Şplanting dayŤ. For each site and season, I take the first crop-establishment date to be the date on which irrigation is made available to the first farmer; Şplanting dayŤ is defined as $Pdate_{ist} - Sdate_{is}$, where $Pdate_{ist}$ is the date on which farmer i

in site s plants his crop in season t, and $Sdate_s$ is the first date on which a crop is planted in site s. To demonstrate the relevance of the farm-specific date of availability of irrigation water as an instrument, Figure 2.1 displays a scatter-plot of farm-demeaned threshold-degree days against planting day, along with a lowess plot and a linear fit, for the high-yielding season[2]. We see that in the high-yielding season, the probability of observing temperatures that exceed the threshold increases as farmers plant their crops later in the season. As further demonstration that the probability of a given day exceeding the temperature threshold increases the later is a particular farmer's planting day, Figures 2.3 and 2.4 display the distribution (using between and within variation) of daily maximum temperature for the first and last days of the window around flowering date, separated by farms in the first and last quintile of planting day. Though all regressions are identified using only within variation, I use this display to help make more concrete the concept of climate as a distribution of potential outcomes on a given day.

Table 2.4 demonstrates the same relationship as in Figure 2.1, in regression form for both high- and low-yielding seasons combined, including an interaction term to account for the fact that planting day is likely to have a differential effect on threshold-degree days depending on the season. (These are the first-stage regressions for the IV estimates in Tables 2.5-2.7). The coefficient on $TDDays32$ indicates the effect for the high-yielding season; we can see that planting day effectively predicts threshold-degree days in the high-yielding season, but has negligible predictive power in the low-yielding season. Focusing on column 3, we see that a one- (within) SD change in planting day results in an increase of 3.6 threshold-degree days, resulting in a subsequent decrease in yield of 119 kg/ha (0.1 SDs; not huge, but meaningful). In this case, I barely reject the hypothesis of zero effect in the low-yield season (p-value of 0.05), but the coefficient is an order of magnitude smaller than that for the high-yield season, and translates to a meaningless (and statistically insignificant effect on yield).

An additional concern is that farmers may be able to influence the irrigation authority and receive their water at an optimal time (or at least a time of their choosing), and that this influence might be correlated with ability or socioeconomic status. There is no test for the validity of planting day as an instrument, but Figure 2.2 displays a scatter-plot of farm size (as estimated by maximum over all seasons, of area planted to rice on a particular farm) against planting day, along with a lowess plot; farm size is the best measure of socioeconomic status, and this figure shows that there is no relationship between farm size and planting day. Additionally, column 2 of Table 2.3 shows that, controlling for the effect of threshold-degree days, planting day

[2]I focus here on the high-yield season due to the lack of a significant yield effect in the low-yield season

has no additional marginal effect on yield. This evidence helps support the idea that farmers do not have influence over the timing of their water allocation, and that the timing is not correlated with ability or socioeconomic status or ability.

Tables 2.5-2.7 focus on the effect of predicted threshold-degree days on preseason input choices. Columns 1 and 4 show the results of naïve OLS regressions of input choice on observed whether. We should expect no effect here as weather observations are not realized until after these decisions have been made; for the most part, these coefficients are indeed insignificant. Columns 2 and 5 show the reduced-form estimates from regressing input choice on planting day, and finally, columns 3 and 6 display the IV results, focusing on the high-yield season only. Table 2.5 shows the effects on area planted and planting method[3], Table 2.6 the effects on seed and fertilizer quantities, and Table 2.7 labor use for land preparation and crop establishment. I detect no effect of planting day or predicted threshold-degree days on land use, crop-establishment method, or preseason labor use. (Note that the standard error on the IV estimate of the effect of threshold-degree days on area planted indicates a minimum detectable effect of about 0.006 standard deviations; we can be fairly confident that the effect here is truly zero).

Table 2.6 is perhaps the most interesting of the three: we actually do see some response with respect to quantity of seed and nitrogen fertilizer[4]. Looking at columns 2 and 3, we see that farmers respond to a later planting day by increasing the quantity of seed that they use, and increase of almost 2 kg/ha per threshold-degree day that is predicted by the planting day (or a 0.16 standard deviation change in seed quantity per one standard deviation change in planting day). Comparing this to column 1, we see that relying on the naïve OLS estimate gives the wrong sign. Nitrogen fertilizer also displays an interesting pattern of results. We see that both OLS and IV give us a negative coefficient, though very imprecisely estimated, while the reduced-form regression on planting day gives us a positive and significant coefficient. It would seem that farmers respond to a bad draw in planting day initially by increasing seed and fertilizer use, but when we look at the part of this that is due to predicted increase in threshold-degree days, on the the seed effect remainsŮthe effect on fertilizer is ambiguous[5]. Note that the precision of the estimates of the coefficients on nitrogen fertilizer and the fact that they are sensitive to specification is likely a result of the heterogeneity of the sample of farms that I use; the probability of fertilizer as a

[3]Flowering date is a function of the number of growing-season degree-days from germination, so it is possible that transplanting could be a way to mitigate the harmful effects of a late planting date. However, transplanting is also associated with a reduction in yield and an increase in planting costs; this could help explain the ambiguous effect.

[4]Nitrogen fertilizer is by far the largest component of cost (25% on average, and up to 80%).

[5]Interestingly, if I add in a term for threshold-degree days in the vegetative phase, the result is a strong, negative coefficient on fertilizer use. It would seem that farmers initially try to compensate for a suboptimal planting day by farming more intensively, but that as they see confirmation of a bad weather draw, they decide to cut their losses.

mitigation strategy is likely to depend on soil quality and seed variety, and within that there may even be heterogeneity of farmer beliefs.

Finally, Table 2.8 shows the effects of threshold-degree days on harvest labor and the application of herbicides and pesticides. These quantities are all chosen during or at the end of the season, so we can expect them to respond to weather observations; I do not consider IV estimates here. We see in the high-yield season, there is an increase in herbicide and pesticide use, and that the coefficient on harvest labor is negative (as we would expect given the reduction in yield) but it is statistically insignificant.

2.3.3 Relevance for climate change

In the previous section, I demonstrated that exposure of a farmer's crop to above-threshold temperature depends significantly and predictably on the exogenously determined day on which the crop is planted, but the question remains: how does this relate to projected changes in the global climate? The average projected warming from IPCC (2007), which takes into account several different emissions scenarios and multiple climate models is about 0.2°C per decade over the next few decades. Warming after that depends heavily on the emissions scenario used, but in general the rate is projected to increase. For example, projected warming for the decade 2090-2100 is around 0.6°C, with a range of $0.3^\circ - 0.9^\circ$C. I make the conservative assumption that the rate of warming will remain constant at 0.2°C over the entire century, and that this increase will be uniform across the temperature distribution. Figures 2.5 and 2.6 display the baseline distribution (average over the sample from 1994-1999) of temperature on a typical day close to flowering, and the projected distribution for 2100, based on warming of 2°C over the century. Based on this projected distribution, I calculate projected values of $TDDays32$, and compare the difference between those values and the ones observed in the sample, to the observed variation generated by variation in planting day. Table 2.2 tests the difference between farms which plant in the first and last quintiles for a particular site in a particular season, and the difference between the average actual and projected $TDDays32$. The assumed warming of 2°C, results in an average increase of 11.1 threshold-degree days, very close to the difference of 10.7 between farms in the first and last quintiles of planting day, and even closer to the observed within standard deviation of 11.0. The variation in threshold-degree days observed in these data is substantial, and comparable to 100 years of projected climate change, yet I find that the farmers do little differently in light of this exposure.

2.4 Conclusion

In this paper, I exploit a natural experiment which provides unique, exogenous variation in crop establishment date for rice farmers in several different important rice producing regions. This exogenous variation in planting day in turn generates variation in the amount of time a particular crop is exposed to temperature above 32°C, a threshold that I estimated to be harmful for rice. I demonstrated that this variation is significant and large, comparable to 100 years of global warming, but that even when faced with the less-favorable climate scenario, farmers seem to do little differently. I interpret the pattern of results presented here as modest evidence that adaptation to climate change is going to depend largely on technological change. In the long run, farmers will have to move, grow different crops, grow their crops differently, or a combination of all three.

These results should be interpreted with caution. In analyzing farmer decisions, I estimate many equations; many more than I report here. We should expect a few coefficients to be statistically significant by chance, and should not assign too much weight to the interpretation of any one single coefficient. The impact of threshold-degree days on yield are robust to a variety of definitions and specifications, and despite the very strong assumptions I make about the nature of climate change, the exogenous variation in these threshold-degree days is meaningful. For the inputs where the lack of response is precisely estimated (e.g. area planted), we can conclude that those inputs will not be used as part of mitigation and adaptation strategies. The imprecision of other estimates is probably due in part to the relatively small sample size and large degree of heterogeneity across season and location. Further research into the source and nature of this heterogeneity is needed, but unfortunately, despite the opportunity for clean identification that these data provide, they fail in terms of identifying this heterogeneity; broken down into the relevant temporal and geographic categories, the cell sizes are simply too small.

2.5 Figures

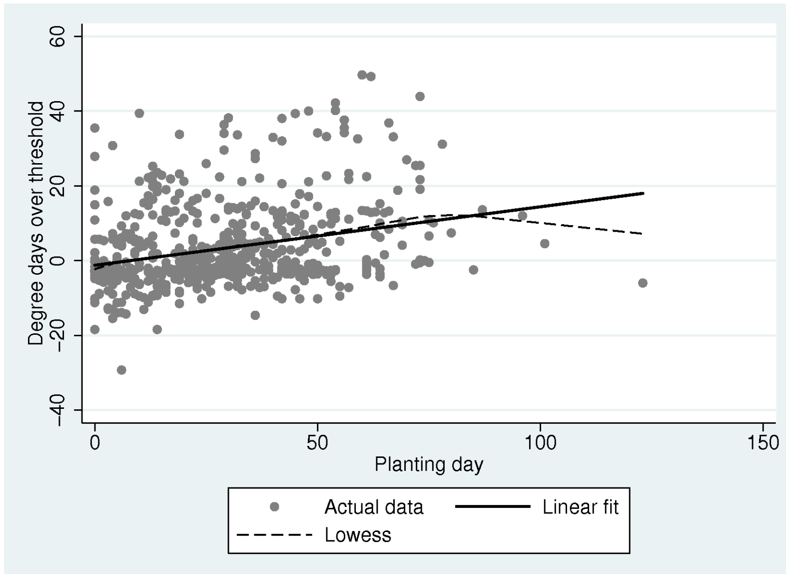

Figure 2.1: Planting day and threshold-degree days

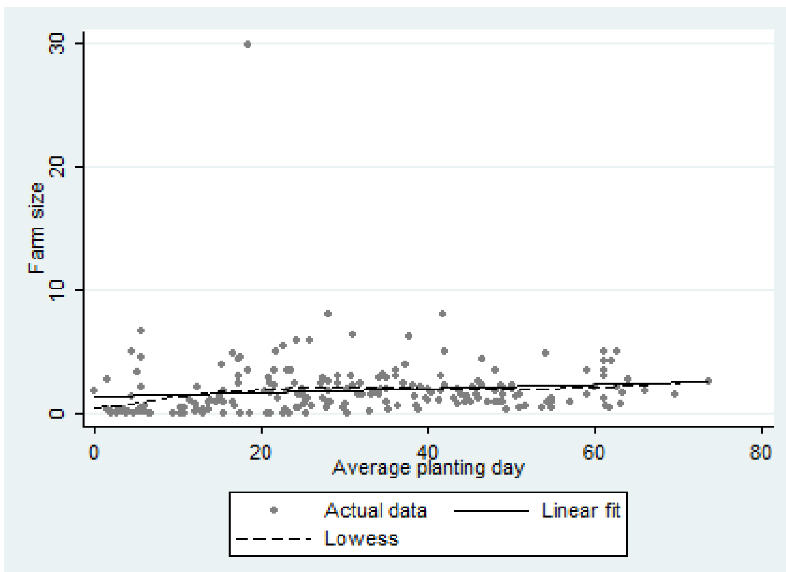

Figure 2.2: Planting day and farm size

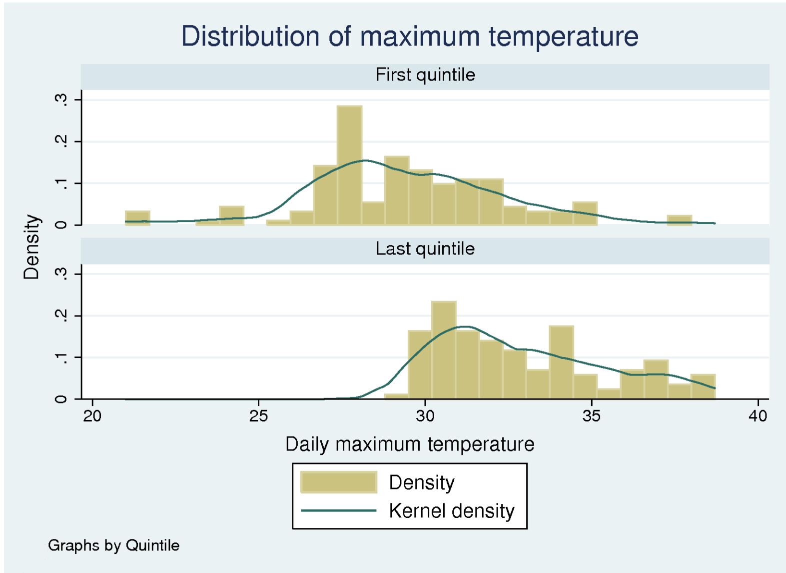

Figure 2.3: Distribution of threshold-degree days by planting day - 37 days before harvest

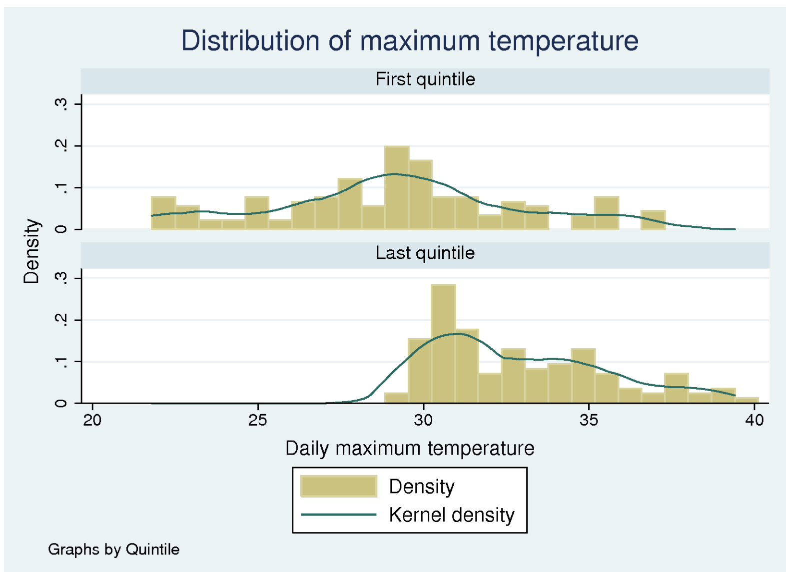

Figure 2.4: Distribution of threshold-degree days by planting day - 23 days before harvest

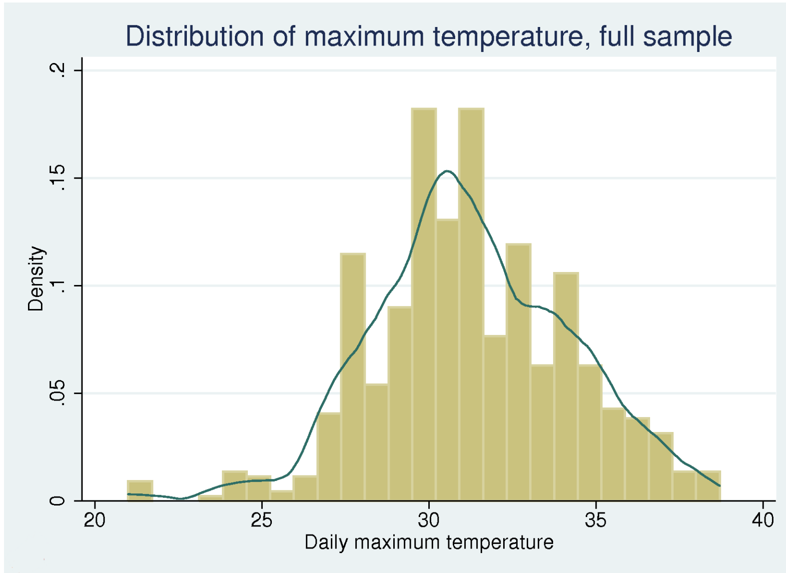

Figure 2.5: Distribution of threshold-degree days by planting day - day 1 of flowering, baseline

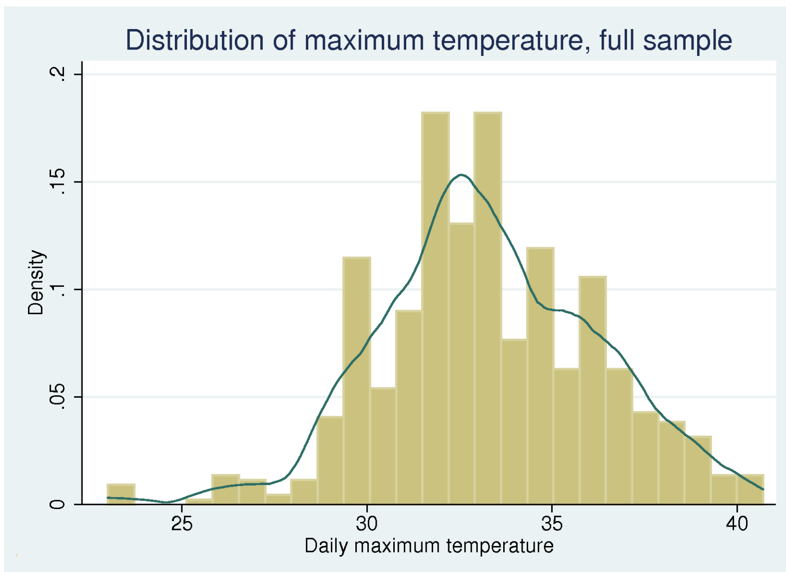

Figure 2.6: Distribution of threshold-degree days by planting day - day 1 of flowering, 2100 projected

2.6 Tables

Table 2.1: Summary statistics

Variable		Mean	Std. Dev.	Min.	Max.
Yield (kg/ha)	overall	5179.317	107.644	288.018	10837.5
	between		938.121	1308.756	7608.345
	within		1198.072	-489.208	10300.48
Planting day	overall	30.342	23.765	0	136
	between		16.897	.444	79
	within		18.157	-15.229	118.914
TDDays32	overall	9.984	15.045	0	88.4
	between		10.580	0	53.95
	within		10.984	-36.766	59.684
Area planted (ha)	overall	1.587	1.946	0.023	29.73
	between		1.670	0.029	15.599
	within		0.869	-8.682	15.718
Seed (kg/ha)	overall	151.169	84.035	4.8	674.286
	between		77.715	19.025	413.083
	within		41.756	-64.880	632.667
Nitrogen (kg/ha)	overall	107.644	39.981	10.5	328.6
	between		31.488	48.809	249.323
	within		28.574	-46.577	256.463
Obs.					
	Season LY	753			
	Season HY	628			

Table 2.2: Test of differences

Variable		Obs.	Mean	Std. Err.
TDDays32 (within)				
	First quintile	129	-1.669	.735
	Last quintile	121	9.031	1.271
Diff			10.7	1.446
TDDays32 (overall)				
	Baseline	1381	10.089	.408
	2100, projected	1381	21.215	.645
Diff			11.125	.274

Table 2.3: Effects of thresholds on yield

	Dependent variable: kg/ha				
	(1)	(2)	(3)	(4)	(5)
TDDays32	-33.346	-33.071	-32.950		
	(7.396)***	(7.052)***	(7.685)***		
TDDays32*Season LY	45.237	43.662	45.009		
	(7.077)***	(7.581)***	(7.151)***		
TDDays37				-146.938	
				(23.876)***	
TDDays37*Season LY				327.505	
				(183.882)*	
NDays32					-115.973
					(28.353)***
NDays32*Season LY					135.137
					(34.778)***
TDNights27			-585.770		
			(432.106)		
TDNights27*Season LY			583.492		
			(429.408)		
Planting day		-3.762			
		(3.183)			
Planting day*Season LY		-.728			
		(3.355)			
Farm FE	yes	yes	yes	yes	yes
Mean temp (max&min)	yes	yes	yes	yes	yes
Obs.	1372	1372	1372	1372	1372
Farms	227	227	227	227	227
F statistic	25.017	29.944	29.877	17.853	26.465
R^2 (within)	0.447	0.452	0.447	0.411	0.426

Standard errors clustered at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.4: Effects of planting day on temp

	TDDays32			TDDays37	NDays32
	(1)	(2)	(3)	(4)	(5)
Planting day	0.232	0.192	0.208	0.025	0.041
	(0.104)**	(0.069)***	(0.027)***	(0.003)***	(0.014)***
Planting day*Season LY	-.235	-.229	-.243	-.028	-.041
	(0.096)**	(0.087)***	(0.029)***	(0.003)***	(0.017)**
1[Season = LY]	1.017	0.298	0.948	0.423	-.477
	(3.741)	(3.731)	(1.066)	(0.108)***	(0.86)
FE	none	site	farm	farm	farm
Obs.	1381	1381	1381	1381	1381
Farms	227	227	227	227	227
F statistic	6.073	3.346	57.274	45.583	4.629
Adjusted R^2	0.078	0.486			
R^2 (within)			0.13	0.106	0.131

Standard errors clustered at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.5: Pre-season inputs: Area and method

	Area planted to rice			1[Method = Transplant]		
	OLS		IV	OLS		IV
	(1)	(2)	(3)	(4)	(5)	(6)
TDDays32	0.008		0.007	0.002		-.002
	(0.004)**		(0.006)	(0.001)		(0.002)
TDDays32*Season LY	-.002			-.004		
	(0.003)			(0.001)**		
Planting day		0.003			0.0007	
		(0.003)			(0.0007)	
Planting day*Season LY		-.005			0.0007	
		(0.005)			(0.0007)	
1[Season = LY]	0.316	0.357		0.076	0.058	
	(0.179)*	(0.281)		(0.033)**	(0.043)	
Farm FE	yes	yes	yes	yes	yes	yes
Mean temp (max&min)	yes	no	no	yes	no	no
Obs.	1372	1381	598	1372	1381	598
Farms	227	227	183	227	227	183
F statistic	1.478	0.839	1.477	3.306	2.067	1.022
R^2 (within)	0.071	0.018		0.079	0.055	

Standard errors clustered at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.6: Pre-season inputs: Seed and fertilizer

	Seed quantity			Nitrogen fertilizer		
	OLS		IV	OLS		IV
	(1)	(2)	(3)	(4)	(5)	(6)
TDDays32	-.443 (0.155)***		1.932 (0.915)**	-.052 (0.177)		-.201 (0.514)
TDDays32*Season LY	0.963 (0.16)***			0.094 (0.171)		
Planting day		0.427 (0.136)***			0.198 (0.095)**	
Planting day*Season LY		-.289 (0.164)*			-.291 (0.135)**	
1[Season = LY]	1.017 (3.906)	14.467 (7.182)**		-7.212 (4.016)*		
Price	no	no	no	yes	yes	yes
Farm FE	yes	yes	yes	yes	yes	yes
Mean temp (max&min)	yes	no	no	yes	no	no
Obs.	1372	1381	598	1372	1381	598
Farms	227	227	183	227	227	183
F statistic	29.157	3.981	4.318	5.264	1.961	0.085
R^2 (within)	0.077	0.02		0.119	0.033	

Standard errors clustered at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.7: Pre-season inputs: Land prep and planting labor

	Land prep labor			Planting labor		
	OLS		IV	OLS		IV
	(1)	(2)	(3)	(4)	(5)	(6)
TDDays32	-.018		-.016	0.025		0.042
	(0.015)		(0.067)	(0.025)		(0.083)
TDDays32*Season LY	0.026			-.074		
	(0.013)*			(0.028)***		
Planting day		-.018			0.043	
		(0.013)			(0.035)	
Planting day*Season LY		0.03			-.016	
		(0.013)**			(0.035)	
1[Season = LY]	-.260	-1.102		1.840	2.385	
	(0.505)	(0.482)**		(0.924)**	(1.311)*	
Price	no	no	no	yes	yes	yes
Farm FE	yes	yes	yes	yes	yes	yes
Mean temp (max&min)	yes	no	no	yes	no	no
Obs.	1372	1381	598	1372	1381	598
Farms	227	227	183	227	227	183
F statistic	1.81	2.552	0.56	3.478	1.901	2.857
R^2 (within)	0.014	0.009		0.054	0.026	

Standard errors clustered at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2.8: Mid/late season inputs: Pesticides and harvest labor

	Harvest labor	Herbicides	Pesticides
	(1)	(2)	(3)
TDDays32	-.076	0.004	0.024
	(0.089)	(0.002)	(0.005)***
TDDays32*Season LY	0.089	0.003	-.007
	(0.094)	(0.003)	(0.004)*
1[Season = LY]	-3.456	0.03	-.331
	(3.141)	(0.042)	(0.096)***
Price	yes	yes	yes
Farm FE	yes	yes	yes
Mean temp (max&min)	yes	yes	yes
Obs.	1369	1349	1372
Farms	227	227	227
F statistic	1.479	2.573	25.513
R^2 (within)	0.023	0.029	0.419

Standard errors clustered at village level.
* significant at 10%; ** significant at 5%; *** significant at 1%

Chapter 3

The ecological footprint of poverty alleviation: Evidence from Mexico's Oportunidades program

3.1 Introduction

Environmental quality and natural resource stocks are key components of welfare for the world's poor but are being degraded at an alarming rate (MEA, 2005). Are efforts to alleviate poverty likely to mitigate or exacerbate this degradation? This is a crucial question for policy-makers pursuing sustainable development goals and has been a perennial debate in the economics literature (e.g. Grossman and Krueger (1995); Dasgupta, Laplante, Wang and Wheeler (2002); Harbaugh, Levinson and Wilson (2002); Foster and Rosenzweig (2003)). Poverty alleviation may raise demand for goods which are resource-intensive in production, increasing degradation. However, increased wealth may also augment demand for environmental resources, inducing households to invest in those resources, or may raise the opportunity cost of extractive activities, reducing degradation. As noted in a recent review (World Bank, 2008), empirical work on the environmental effects of poverty alleviation has been significantly limited by the possible endogeneity of household income changes. In this paper, we exploit the discontinuity in the community-level eligibility rule for a conditional cash transfer program in Mexico, as well as random variation in the pilot phase of the program, to study the consequences of poverty alleviation programs for environmental degradation.

Previous work has also not adequately considered problems in estimating the response to income changes when impacts are market-mediated and therefore can be spatially dispersed. Recent work on the effects of local rainfall shocks (Keller and Shiue, 2008; Donaldson, 2009) shows that as infrastructure improves, price changes become less correlated with local shocks. Similarly, we show that even if the true impact of a wealth increase on production is constant, we will detect apparently heterogeneous impacts. Stronger effects will be found where infrastructure is poor and thus the source of environmental resources for production is more geographically constrained. The market-mediation of impacts is a fundamental causal inference issue but is often difficult to disentangle because markets are relatively homogenous. Here we take advantage of large variation in transportation infrastructure to investigate whether observed heterogeneity in impacts is consistent with these theoretical predictions.

Our analysis focuses on deforestation as a measure of environmental quality. Deforestation is locally and globally important and in our dataset can be consistently measured for the more than 105,000 localities in Mexico. Locally, forests contribute to welfare through fuel wood, fodder, timber, watershed protection and wildlife habitat. Globally, forest loss is a major environmental concern. Net forest cover is estimated to have fallen by 9.4 million hectares (just under one percent) per year during the 1990s (FAO, 2005). Carbon emissions from deforestation are estimated at approximately 20% of the global total (IPCC, 2007) and have been an important focus of recent international climate negotiations. We link spatial data on deforestation in Mexico from the period 2000-2003 to the location and eligibility of every locality in Mexico, and exploit this data structure to examine whether deforestation rates are affected by the program.

Oportunidades represents an ambitious attempt to increase consumption among the poor in Mexico by building human capital. The program funnels large cash payments to households conditional upon their children's school attendance and receipt of regular health checkups. The program has an annual budget of $2.6 billion, or half a percent of GDP, and treats 40% of rural households, increasing per-capita income among recipients by an average of one-third. The program's rollout featured centralized eligibility thresholds at both the locality and the household level, with eligibility defined according to a marginality index. It therefore introduced a large income shock which is discontinuous where localities are defined as just "poor enough" to participate in the program. While a relatively large literature exists using the household-level discontinuity in Oportunidades (Angelucci and de Giorgi, 2009), few previous analyses use the community-level discontinuity (exceptions are Barham (2009)'s paper on the impact of Oportunidades on child health and Green (2005)'s study of political impact). This structure provides us with an unusual ability to study economy-wide effects from the nation-wide introduction of a conditional cash transfer program in a large and diverse country.

We find that exposure to Oportunidades increases deforestation. The results imply roughly a doubling in the probability that any deforestation occurs in a locality. The probability that any deforestation occurs in a locality not eligible for the program is 4.9%, so this represents an increase in an already high likelihood of deforestation. Among communities who do deforest, the results indicate an increase in the rate of deforestation ranging from 15 to 33 percent. To understand the micro-behavior that might explain this increase in deforestation, we turn to household data from the randomized pilot phase of the program. These experimental data show that additional household income significantly increases consumption, and recipient households shift strongly into land-intensive goods such as beef and milk. Consumption increases appear to be constant across localities, but the corresponding production increases and deforestation patterns are not. We observe significant household-level production responses only in treated localities which are more isolated. We also find larger deforestation effects in treated localities that have poor road infrastructure and thus are more isolated from outside markets. Finally, we investigate spatial spillovers of treatment using a new method for calculating spatial lag functions in a regression discontinuity context. This analysis shows the spatial contour of impacts to be flat where roads are good, and to be concentrated around the location of treatment where roads are bad. These results are consistent with the hypothesis that transportation infrastructure is a significant determinant of the spatial profile of market-mediated production impacts.

Our results suggest that there are significant environmental impacts of poverty alleviation. There is an increase in deforestation as households shift demand from less land-intensive goods to more land-intensive goods, increasing their "ecological footprint" (Wackernagel and Rees, 1996). This contrasts with Foster and Rosenzweig (2003)'s finding that as incomes rise, household demand for forest products increases, strengthening incentives to conserve forests. It implies that in cases where the demand for agricultural products is likely to rise faster than the demand for forest products in response to higher incomes, poverty alleviation programs should be accompanied by environmental regulations that correctly price externalities or clearly establish property rights to environmental goods (i.e. carbon markets). The results also indicate that policymakers should be cautious in interpreting the magnitude of apparent impact estimates without taking into account how these are mediated through markets. Given a set of localized demand shocks, better-integrated local markets will allow demand to be sourced from a broader set of producers. To the extent that new demand is satisfied by national or global markets, we will not observe a clear link between local consumption increases and local environmental degradation. Therefore where local infrastructure is good, impact studies are unlikely to capture the full magnitude of the “ecological footprint" effect[1].

The paper is organized as follows: we begin in the next section by discussing the literature on poverty and deforestation and the empirical problem introduced by the study of micro-interventions when agents may participate in market transactions on a broader spatial scale. Section 3.3 describes the Oportunidades program in more detail, and presents the estimation strategy and results of the discontinuity analysis. Section 3.4 seeks to disentangle the mechanisms through which this impact occurs by using household data from the randomized evaluation phase of the program. Section 3.5 presents results on the heterogeneity and spatial distribution of observed impacts, and the final section concludes with a discussion of the policy implications of our findings.

3.2 Poverty, Deforestation, and Spatial Impact Analysis

Conditional cash transfer programs that seek to alleviate household poverty and improve access to education or health are increasingly popular in developing countries, but may have unintended secondary effects. One possibility that has not received adequate attention is the potential for environmental consequences. It is not clear, *ex ante*, whether we should expect income increases to exacerbate or reduce environmental degradation: a large previous literature on the Environmental Kuznets Curve suggests the relationship is complex and non-linear (Stern, 2004; Dasgupta, Laplante, Wang and Wheeler, 2002; Panayotou, 1997). Disentangling this relationship requires examination of three distinct yet interrelated issues: the existence of a correlation or causal link; the micro-foundations of the relevant household production and consumption decisions; and the role of local markets in mediating the relationship.

3.2.1 Does alleviating poverty increase or decrease forest cover?

We focus on forests as an environmental outcome of interest. Forests are a key local resource and global public good. Understanding how to prevent further deforestation would significantly contribute to efforts to limit greenhouse-gas emissions (Kaimowitz, 2008; Stern,

[1] It is possible that by sourcing production more broadly, goods will be produced more efficiently and thus the true impacts might actually be smaller in better-integrated markets rather than constant. Caution is still warranted because environmental goods may not be efficiently priced and therefore not efficiently sourced.

2008). However, even if we limit the scope to the relationship between income and deforestation, previous empirical results and theory are ambiguous (Pfaff, Kerr, Cavatassi, Davis, Lipper, Sanchez and Timmins, 2008; Chomitz, 2006).

Whether higher household incomes increase or decrease pressure on forest resources depends on multiple factors (Barbier and Burgess, 1996; Wunder, 2001; Pfaff, Kerr, Cavatassi, Davis, Lipper, Sanchez and Timmins, 2008) including prices of agricultural and pastoral goods (Pfaff, 1999), demand for forest products (Baland, Bardhan, Das, Mookherjee and Sarkar, 2007; Fisher, Shively and Buccola, 2005; Foster and Rosenzweig, 2003), credit constraints (Zwane, 2007), returns to alternative household activities (Deininger and Minten, 1999, 2002), agricultural intensification and extensification (Shortle and Abler, 1999; World Bank, 1992), and demand for environmental amenities (Cropper and Griffiths, 1994). The complexity of the relationship between household incomes and deforestation means that research has generated few unambiguous theoretical predictions, and the search for sufficiently large, plausibly exogenous sources of income variation for empirical analysis has been a challenging one.

Initial work on the development-deforestation link focused primarily on the presence and shape of an Environmental Kuznets Curve (Cropper and Griffiths, 1994; Pfaff, 2000), positing that forest cover initially decreases as income rises but then recovers as income increases beyond some turning point. Subsequent work has shown both increases and decreases in forest cover as income increases. Foster and Rosenzweig (2003) use a general equilibrium framework to show that devotion of land to the production of forest products should rise as demand rises. They confirm this relationship using long-term changes in income and forest cover across Indian states. Deininger and Minten (1999, 2002) suggest that as countries grow richer, relative returns to off-farm labor would increase and reduce pressure on forests. They illustrate such a relationship in data from Mexico. Zwane (2007) finds that the relationship between income and deforestation in Peru is positive at low levels of income but may be negative at higher levels. Baland, Bardhan, Das, Mookherjee and Sarkar (2007) assesses the impacts of income growth on firewood collection in Nepal and find a net negative but very small effect.

The empirical literature on the relationship between income and deforestation has been hampered by concerns about the endogeneity of income growth. Rates of deforestation are clearly influenced by multiple factors which could be correlated with income shocks. These include population growth, agricultural returns, forest product prices, capital availability, technology, accessibility and institutional variables (see reviews by Angelsen and Kaimowitz (1999); Barbier and Burgess (2001); Chomitz (2006)). The endogeneity problem may be particularly severe for studies using cross-sectional variation to identify impacts. Conversely, in studies using panel variation in income (Zwane, 2007; Baland, Bardhan, Das, Mookherjee and Sarkar, 2007), the relatively small income changes observed in a short-term panel may not reflect true economic development. Also, these short-term fluctuations are different in nature than permanent income changes. Households are likely to respond differently to income changes that are perceived to be substantial and permanent versus small and temporary.

Exploiting Mexico's rollout of Oportunidades allows us to make two contributions to the existing empirical literature. First, the implementation of the Oportunidades program creates an exogenous source of variation in income, allowing for clean identification of causal effects. Second, the magnitude and duration of the program represents a substantial and durable increase in income for a large share of the households in poor communities. We are thus able to estimate impacts using a positive shock to income that is as large as is likely to be achievable by any actual poverty alleviation program.

3.2.2 The household response to income shocks

In the set of empirical studies discussed above, several potential mechanisms are proposed to explain how changes in household income might affect deforestation. On the production side, Deininger and Minten (1999, 2002) suggest that income increases which occur through increased returns to off-farm labor would reduce agricultural land use and ease pressure on land, also reducing deforestation. Although a conditional cash transfer program might not directly raise off-farm wages, it could raise the opportunity cost of leisure, and therefore discourage on-farm production through a similar mechanism. Other researchers have suggested that income increases could spur capital improvements or technological adoption, which would facilitate agricultural intensification and reduce pressure on forests (Shortle and Abler, 1999; World Bank, 1992). Zwane (2007), in contrast, suggests that the expected effect of relaxing a credit constraint depends on initial income. At low incomes, relaxing the credit constraint increases deforestation while at higher incomes there is an offsetting increase in the marginal utility of leisure which may result in less deforestation.

On the consumption side, Foster and Rosenzweig (2003) propose that higher incomes will decrease deforestation through increased demand for forest products and a corresponding supply response by households where there is clear ownership of forest resources. However, their results depend on the demand for forest products rising faster than the demand for agricultural products in response to an income increase. If instead households rapidly increase demand for land-intensive agricultural goods, we would expect to see increased deforestation. This pattern might be particularly pronounced if inferior goods are relatively more land-efficient than normal goods. As incomes increase, households would substitute consumption away from these land-efficient inferior goods (e.g. beans) to land-intensive normal goods (e.g. beef), thus expanding their “ecological footprint".

3.2.3 The ecological footprint of market-mediated shocks

If income changes lead to consumption-driven impacts on deforestation, we must address an issue that is fundamental to the estimation of all market-mediated impacts: there is by no means a one-to-one mapping between the location of the consumption change and the location of the corresponding adjustment in production. Particularly when the treatment unit (and therefore the source of variation in demand) is small relative to the geographic coverage of the program, the extent to which production impacts spill over will determine what is measured by comparing treated and untreated units. In trying to understand how these local shocks alter market demand and supply of forest-intensive resources, we can draw an analogy with the literature estimating the effect of localized rainfall shocks on prices. A well-established result from this literature is that as infrastructure improves, prices become less correlated with localized rainfall shocks and more correlated with the rainfall shocks of adjacent areas (Keller and Shiue, 2008; Donaldson, 2009). This effect occurs because demand within a given area is sourced from more distant producers when infrastructure is improved, and hence shocks are spread over a greater area.

When we measure market-mediated treatment effects from localized experiments (even randomized ones), this same phenomenon will generate observed heterogeneity in the measured treatment effect across infrastructure quality. This heterogeneity will be present even if the true, total treatment effect is constant. To see this, we can think of a market as a grouping of a set of units into a single price-setting mechanism, so that shocks to one unit within a market are transmitted to the other units. Let the number of units per market be given by η, which proxies for infrastructure quality. A treatment induces a constant increase in demand equal to τ per unit, and this increase in demand is sourced on average from itself and the $\eta - 1$ other members of the market.

The increase in outcomes within a unit as a function of its own treatment is the part of the effect that does not spill over, namely $\frac{\tau}{\eta}$. In addition to the direct effect of treatment, each unit will receive an expected spillover effect equal to the indirect treatment effect from the number of individuals within the market who were treated. Writing the share treated as σ, then $\sigma\eta$ units per market will be treated and the expected spillover effect will be $\sigma\eta\frac{\tau}{\eta} = \sigma\tau$. The average treatment effect is given by the difference between treated and untreated units, or

$$E(Y \mid T) - E(Y \mid C) = (\frac{\tau}{\eta} + \sigma\tau) - \sigma\tau = \frac{\tau}{\eta}.$$

This says that the experiment measures not the total effect of treatment but only the component of it that does not spill over to other members of the same market. Now if we think of infrastructure (in our case roads) as being an intermediating variable that determines the size of the market, it can be thought of as determining the number of units on to which the treatment effect τ spills. In environments where the road network is excellent, η moves towards infinity and we have a single national market where the measured difference between treatment and control units is zero. With poor road infrastructure, consumption is localized to the spatial unit of treatment, η goes to one and the estimated difference between treatment and control converges on the true total treatment effect, τ. If what we set out to do with our experiment was to measure the total environmental impact of the treatment, then the error, meaning the difference between the true total treatment effect and the result of the micro-experiment is given by $\tau(\frac{\eta-1}{\eta})$, which vanishes as markets become completely autarkic.

In a sample with variability over the quality of local infrastructure, we will observe heterogeneity in impacts even when the actual treatment effect is constant. The reason for this differential is that spatial arbitrage removes the difference between treated and control units when the pixel size of treatment is small and transport costs are low. Under the assumption of homogenous treatment effects, such an argument implies that we only get the correct estimated treatment effect when spatial arbitrage is shut off. This argument is consistent with the results of Foster and Rosenzweig (2003), who observe a positive feedback effect of higher income on forest reserves only in closed economies, but not in open ones. Presumably the reason for this heterogeneity is that closed economies do not arbitrage their increased demand for forest products across global markets, and hence they manifest the full treatment effect on internal markets. In what follows we investigate the heterogeneity in impacts across infrastructural quality and confirm that our largest observed treatment effects occur precisely where they are the most localized.

3.3 Oportunidades and Deforestation: Overall Impact

3.3.1 Program description

The intention of Oportunidades is to increase school attendance and health care among poor families in Mexico. The financial scope of Oportunidades is large. The annual budget is approximately $2.6 billion a year, about half of Mexico's anti-poverty budget. It treats some four million households, providing cash transfers conditional on health care provision and school attendance. On average the transfers are about one-third of total income in these poor households, clearly meaningful income changes.

The program has been widely studied and lauded for its success in achieving these objectives (Schultz, 2004; Fernald, Gertler and Neufeld, 2008; Skoufias and McClafferty, 2001). The transparent nature of its enrollment criteria and benefits has contributed to the attractiveness of the program, and it is currently being replicated in various other countries. The program was implemented in stages. A pilot implementation of the program (beginning in 1997) was randomized, and combined with detailed household-level data collection. The full rural roll-out of the program occurred mainly in 1998-2000, but new communities continued to enroll at a slower rate after this. This phase was not randomized, but was targeted to localities based on a marginality index; this created the discontinuity in treatment which we use. Eligible rural villages were first selected according to their level of marginality, and then surveys were conducted within villages to determine who would receive payments.

3.3.2 Data description

Our analysis of the national rollout focuses on rural localities[2]. We combine information on locality eligibility and program rollout with national deforestation data.

The spatial coordinates of each locality (village) in Mexico, along with the population and marginality index numbers for 1995, are from the National Institute of Geography and Statistics in Mexico (INEGI), and the data describing the roll-out of Oportunidades come from the Oportunidades office. We have information on enrollment by village through 2003[3]. Locality-

[2]We exclude villages with more than 2,500 inhabitants as these are defined as "urban" communities in Mexico and were not eligible for the program until after 2000. Focusing only on rural localities means that we are likely to underestimate the total environmental impacts of Oportunidades because we are not taking into account possible consumption increases resulting from the urban roll-out.

[3]Although the bulk of enrollment in rural areas occurred before 2000, some villages were enrolled after this date. We include these villages although the presence of these villages, which were not enrolled according to the eligibility cutoff, potentially biases the results towards zero and against finding any impact of the program. Leaving them in the dataset therefore generates the most conservative estimates. Our results hold and are in fact stronger if we exclude

level eligibility for the program is based upon marginality indices calculated by CONAPO for 105,749 localities[4].

To measure deforestation at the locality level we rely on data from the Mexican National Forestry Commission (CONAFOR). The data are based on mosaics of Landsat satellite images from 2000 and 2003 (30 m resolution) and were created by CONAFOR under a mandate to accurately measure and monitor deforestation (Monitoreo Nacional Forestal). CONAFOR's data pieces together a large number of Landsat scenes in order to achieve wall-to-wall coverage for the entire country. This is in contrast to the method used by Foster and Rosenzweig (2003) which looks at forest cover for a representative sample of villages. Here we are measuring deforestation for all of the more than 105,000 localities with a marginality index in 1995[5]. We restrict the analysis to localities which had at least 10 hectares of land classified as forest in the 2000 National Forest Inventory, focusing on localities in which deforestation is possible[6]. Figure 3.1 shows the distribution of forest across Mexico in 2000. In order to assign each part of the landscape to a unique locality, we use the method of Thiessen polygons. (INEGI gives point data on the locations of each locality, but data on the detailed boundaries of the localities does not exist.) This method assigns land to localities based on the closest locality point and has the advantage of avoiding the problem of double counting caused by other shapes such as circles around each locality. Figure 3.2 shows a zoomed in picture of land use in 2000 along with the locality boundaries assigned by the Thiessen polygons method. Finally, because CONAFOR was primarily concerned with identifying areas of new deforestation, we do not have data on afforestation. We correct for this potential censoring problem in the data analysis by using Tobit estimations. Practically speaking, we believe our measure picks up the key land use change dynamic of the study period because Mexico was a net deforester across this time. In fact, FAO's 2005 Global Forest Resources Assessment places Mexico in 13th place in the world in

villages enrolled in and after 2000 or before 1998

[4]By 2000, points were available for approximately 200,000 localities; the missing points in 1995 are very small localities: ninety-three percent of the villages for which there is no marginality index in 1995 had fewer than 25 inhabitants in 2000. The index is a continuous measure and was created using a principal components analysis based on seven variables from the 1995 Conteo (short census) and 1990 census, including illiteracy rates, dwelling characteristics, and proportion of the population working in the primary sector (Skoufias et al., 1999).

[5]The correct georeferencing and interpretation of Landsat data is a specialized and labor intensive process. After putting images together from several Landsat "scenes," the classification of deforestation is based on changes in the Normalized Difference Vegetation Index (NDVI) values across time. Comparisons are made using images from the dry season. NDVI is an indicator of vegetation cover and is used worldwide to measure changes in forest cover. Although NDVI change is the best available indicator of changes in forest cover, we note that the measure can have some errors due to weather shocks such as unusually high rainfall or drought conditions. These errors are in the dependent variable but are unlikely to be correlated with variation in treatment.

[6]The NFI data are based on a combination of remote sensing using Landsat images and field sampling to verify the classification system. The results are not sensitive to using lower thresholds.

terms of net forest loss over the period 2000-2005 (FAO, 2005). We present results using the percent of each locality deforested as the dependent variable, but all results in the paper are robust to alternative specifications of the dependent variable, including ln(total deforestation) and percent of baseline forest area deforested.

3.3.3 Illustrating the discontinuity

Figures 3.3 and 3.4 illustrate the variation in program enrollment and deforestation across the marginality index. The marginality index, which is continuous, is divided into bins with width = .1 for these illustrations. In each of these figures the left axis measures the percent of each locality deforested and the right the proportion of localities treated.

Figure 3.3 shows the relationship between enrollment, deforestation, and marginality for the full sample of localities[7]. As expected by program rules, we see a sharp increase in enrollment to the right of values of -1.2 on the marginality index. The discontinuity is not perfect – there is a small jump in enrollment before the eligibility criteria. This jump is due almost entirely to the enrollment of villages post-2000, when the program became more demand-driven[8].

Figure 3.3 also shows that deforestation rates vary with poverty in a roughly inverse-U relationship. This is an interesting confirmation of the empirical environmental Kuznets curve relationship: we see lower rates of deforestation for very poor communities (high marginality index), higher rates of deforestation for poor communities, and lower deforestation rates among less poor communities[9].

Figure 3.4 zooms in on the range of the marginality index around the eligibility cutoff, showing the discontinuity more clearly. The figure uses a kernel regression to estimate the relationship between deforestation and the marginality index (the results are robust to larger and

[7] It is important to note that the number of observations in each bin varies considerably across bins because the marginality index itself has frequencies which are roughly normally distributed. Therefore there are few observations per bin in the extreme bins and many more per bin towards the middle. This means that outliers have more influence on the points at either end of the marginality distribution.

[8] The proportion enrolled remains high for intermediate values of the marginality index and then is lower at high levels of marginality; we suspect that the decreases in enrollment at very high levels of marginality may be related to the fact that the very poorest villages may not have been eligible as a result of their lack of infrastructure.

[9] Note that because income is decreasing as we move to the right, a treatment that increases income is effectively pushing households to the left on this figure. The implication is that while the cross-sectional data are supportive of a Kuznets-style relationship (deforestation highest in the middle part of the distribution) the eligibility discontinuity lies above this value, and so if we took the Kuznets relationship to be causal, we would have expected an income increase in this part of the poverty distribution to decrease deforestation. This would appear to provide another piece in the already substantial body of evidence suggesting that cross-sectional Kuznets relationships do not depict a causal link between income and environmental changes.

smaller windows). The data range in Figure 3.4 includes marginality levels from -2 to -.2, which constitutes 27% of the total sample with baseline forest and populations less than 2,500. This is referred to as the "restricted sample" in the sections that follow. We can see the clear increase in the proportion of localities to the right of -1.2. We also see the increase in deforestation rates around the discontinuity. Deforestation rates average around .03 percent on the richer end of the discontinuity, but once a locality becomes just poor enough to qualify for Oportunidades, average deforestation jumps to nearly .08 percent.

3.3.4 Empirical strategy

We observe a cross-sectional relationship between enrollment in Oportunidades by the year 2003, and deforestation between 2000 and 2003. One way to estimate the effect would be to apply OLS to the equation:

$$\Delta f_i = \alpha + \delta T_i + \beta' X_i + \varepsilon_i \tag{3.1}$$

where Δf_i represents the percent deforestation in polygon i over the period 2000-2003, T_i is equal to one if the locality associated with the polygon was enrolled in the program by 2003, X_i represents a vector of locality-level characteristics which might also affect deforestation, including poverty, and ε_i are unobserved factors affecting deforestation. If the program had been randomly assigned, then this would be an appropriate way to measure its effect on environmental outcomes. However, it is not randomly assigned; it is offered to those who are poor, and who may be likely to have different rates of deforestation even in the absence of the program. In addition, since enrollment in the program is voluntary, it is possible that those communities where enrollment is very high are systematically different than those where enrollment is very low – i.e., that selection problems could bias the estimates of the parameters in equation 3.1.

If the discontinuity is sharp, meaning that the rule for eligibility perfectly predicts treatment, then one can simply include the eligibility cutoff as a proxy for the treatment itself. In our case, this is a dummy variable (E_i) equal to one if the locality's marginality index exceeds -1.22. This corresponds to the boundary between "medium" and "low" levels of poverty, as classified by the marginality index. We use this simple approach in several specifications, noting that it captures the intention to treat effect, rather than the treatment effect on the treated.

Our situation differs from a sharp discontinuity in two ways. First, enrollment is not one hundred percent beyond any threshold. Second, the probability of enrollment increases rapidly over a range of the marginality index between -1.2 to -0.9. The first problem can be dealt with in the standard way by using the eligibility cutoff to instrument for the probability of enrollment[10]. We address the second problem following approaches developed by Hahn et al. (2001), Green (2005) and Jacob and Lefgren (2004). Nonlinear combinations of the eligibility rule and the marginality index (equation (3)) are used to instrument for treatment in the main regression. The two equations are given as:

$$\Delta f_i = \alpha + \delta T_i + \gamma I_i + \beta' X_i + \varepsilon_i \tag{3.2}$$

$$T_i = \omega + \tau_1 E_i + \tau_2 E_i I_i + \tau_3 M_i + \tau_4 M_i I_i + \mu I_i + \Gamma' X_i + \nu_i \tag{3.3}$$

where T_i represents treatment, E_i is the eligibility cutoff dummy, I_i is the marginality index and M_i is a dummy equal to one over the zone where enrollment increases rapidly and zero otherwise. Other variables are as defined above. Note that all specifications include a control for the marginality index, I_i, in order to control for the underlying relationship between deforestation and poverty. We also estimate results both for the full sample and a narrow window around the discontinuity. Within a narrow window around the discontinuity, it is reasonable to assume that the relationship between poverty and deforestation is linear. When we use a wider window, we include higher-order terms of the index (up to a fourth-order polynomial, following the example of (Lee, Moretti and Butler, 2004). We also include additional controls, represented above by the vector X_i and including the size of the polygon in kilometers squared, the population in 1995, the percentage of the polygon that was forested in 2000, kilometers of roads in a 10 kilometer buffer around the locality (“road density"), and regional ecosystem dummy variables. Finally, in order to address issues surrounding the appropriateness of the IV Tobit estimator when the endogenous variable is binary, we also estimate the equation substituting the continuous proportion of households treated in lieu of the binary treatment variable.

Valid estimates based on a regression discontinuity design rely on the assumption that the discontinuity in the outcome can be attributed to the discontinuity in treatment; i.e. there is not another unobservable variable which also changes discontinuously over the relevant marginality range which could be driving the results. To test this assumption, we analyzed all covariates using the kernel regression specification applied in Figure 3.4. No variables showed a significant jump at the discontinuity, with the exception of slope, which is slightly higher among the eligible population. Given that deforestation generally decreases with increases in slope, we feel that this strengthens our results. In addition, we control for slope in all specifications.

[10]For a review of regression discontinuity approaches, see Imbens and Lemieux (2008).

As a falsification test, we check for a discontinuity in forest cover around the eligibility cutoff prior to the start of the program, using data on 1994 land use. We find no difference in 1994 forest levels (measured in percent of polygon in forest) at the point of the discontinuity either visually or statistically[11].

3.3.5 Results

Table 3.1 presents some simple summary statistics from the two samples comparing average deforestation levels and other covariates across the eligibility criteria for the program. In both the full and restricted samples, there are significant differences in both the probability of deforestation and in the level. These simple comparisons of means across the running variable seem to indicate the presence of a jump in deforestation around the discontinuity. They do not, however, control for the underlying relationship between poverty and deforestation, nor do they control for any other covariates which might be correlated with both of these.

Simple approach

We first present results from the simplest approach of regressing deforestation outcomes on the eligibility cutoff as a proxy for treatment (i.e. intention to treat; which replaces T_i in equation 1 with E_i). Table 3.2 shows the results of this approach. The first three columns are estimated using a Tobit. Columns (1) and (2) show results from the full sample and the last column from the restricted sample (marginality index between -2 and .2). Column 1 includes in addition to the eligibility cutoff: the marginality index, the area of each locality, the baseline percentage of the locality in forest, locality population, road density, slope, and ecoregion controls. Column 2 shows results with a fourth order polynomial of the marginality index[12]. The third column shows results from the restricted sample and includes the marginality index linearly.

We see that the coefficients on eligibility are positive and significant (10% level) in all specifications, suggesting that the program increased deforestation. Marginal effects of eligibility on the probability of deforestation and on the rate of deforestation among deforesters calculated at the mean of the covariates are given at the bottom of Table 3.2. As a robustness check, we also consider OLS estimates of the probability of deforestation in a given polygon (column (4)), and on the percent deforested in those polygons with positive deforestation (column (5)).

[11] Unfortunately, the data on 1994 forest areas is missing large tracts of land in northwest Mexico and in parts of the state of Guerrero; but at least 30,000 relevant observations remain. We also note that the classification of this data into land uses is not directly comparable with the 2000 Forest Inventory so we must use forest cover rather than changes in forest cover for this test.

[12] Results are robust to including just second and third order polynomials of the index as well.

Note that the estimates from the linear probability model are nearly identical to the marginal impact of eligibility on the probability of deforestation estimated using the Tobit. The impact on percent deforestation among the deforesters is larger in the linear model that in the marginal effect estimated with the Tobit, but it is also not adjusted for the probability of deforestation in the sample.

Relying on this simple methodology, we also conduct a basic falsification test of the results using pseudo eligibility rules. We chose the eligibility cutoff based on the defined boundary between “medium" and “low" levels of poverty (-1.2). Using other cutoffs should not indicate deforestation effects. We re-run the specification in Column 2 of Table 3.2 on subsamples both to to the left and to the right of the discontinuity, but re-define eligibility at each tenth of the marginality index. We do not find any significant results using these placebo eligibility thresholds[13].

Instrumental variables approach

Results from the instrumental variables discontinuity approach are presented next. We begin by examining the predictive power of the instruments and then show the impact estimation results. Table 3.3 shows the results of the first stage OLS regressions (corresponding to equation (3)) of a dependent variable equal to one if the locality was treated by 2003. The first column tests the significance of the simple instrument of eligibility using the full sample, and columns 2-3 test the power of the set of fuzzy discontinuity instruments on the full sample. Column (4) shows results for the restricted sample. Column (5) shows an estimation of the fuzzy discontinuity variables on the proportion of households receiving Oportunidades in a locality between 1997 and 2003. The variables have the expected signs – being eligible for the program (in the zone above -1.2) increases the probability of enrollment, as does being in the marginal zone. The slope of the increase in probability of enrollment in the marginal zone is given by the interaction of the marginality index with the marginal zone, and is positive and significant as predicted. Estimations 3 and 5 include nonlinear terms of the marginality index. F-tests of the set of excluded instruments show that the instruments have excellent power.

Table 3.4 shows the estimated impact of the program on deforestation using the eligibility as the sole instrument. The results are consistent with those of the simplest approach, showing participation in the program increasing the probability and amount of deforestation. Two robustness checks in Table 3.4 warrant discussion. First, IV OLS is used in columns (5) and (6), and

[13]Results available upon request.

yields a nearly equivalent marginal effect of treatment on the probability of deforestation, and, as in the simplest approach, a slightly larger impact on percent deforestation. Column (3) uses a continuous variable to measure impact – the average proportion of the locality treated – and the marginal impact on the probability of deforestation is substantially larger than using the binary treatment. It is important to note, however, that the binary treatment variable should pick up the treatment for the average locality, which in terms of proportion treated is .42. Multiplying .120 by .42 yields a marginal effect estimate nearly identical to the marginal effect estimated using the binary treatment in column (2).

Table 3.5 shows the estimated impact of the program on deforestation using the fuzzy discontinuity approach. The estimates are similar to the simple approach. The marginal effects for the binary treatment indicate an increase in the probability of deforestation of 1.8 to 3.8 percentage points. Given the baseline probability of deforestation among the non-eligible population of 4.9%, this suggests nearly a doubling of deforestation probability around the discontinuity. The baseline percent deforested among deforesters in the non-eligible population is .6, which means that the marginal effects implied by the estimation amount to a 15-33% increase in the percentagearea deforested among deforesters.

The discontinuity results indicate that Oportunidades is associated with an acceleration of deforestation. Localities that received treatment show greater deforestation than localities with very similar poverty levels that did not receive treatment. In order to try to understand the household-level changes that might underlie these broader impacts, we turn to the evaluation data from the randomized pilot of the program.

3.4 Understanding Household Channels using a Randomized Trial

3.4.1 The Progresa data

The initial, experimental phase of Oportunidades was known as Progresa. The pilot phase featured a three-year period during which the intervention was directly randomized at the locality level. This evaluation design provides a unique opportunity to study the microfoundations of the household production and consumption decisions that underlie the observed deforestation impacts. Of the pool initially identified for participation in the program, 506 localities were randomized into 320 "treatment" (initial intervention) and 186 "control" (delayed intervention) groups. Within each locality, households were assigned eligibility status for the program depending on their degree of poverty; eligible households within the treatment localities received the program. The experiment included several baseline and evaluation surveys that have been used in previous studies (see Skoufias (2005), Section 3 for a description of the evaluation design). For our analysis, we combine the 1997-98 baseline surveys with the 2000 follow-up survey which occurred at the end of the experimental phase.

Since the program was randomized among households in this dataset, we apply a difference in difference specification. We use the sample of eligible (poor) households to estimate direct treatment effects and the sample of non-eligible (non-poor) to estimate spillover effects:

$$y_{it} = \gamma_0 + \gamma_1 T_i + \gamma_2 P_t + \gamma_3 T_i P_t + \upsilon_{it} \tag{3.4}$$

where y_{it} is the household-level outcome variable related to consumption or production, T_i equals 1 if the household is in a treated locality, P_t is equal to one in the post-treatment period, $T_i P_t$ is the interaction of T_i and P_t, and υ_{it} is the household specific error. Because randomization was at the locality level we cluster standard errors at the locality level.

We test first for relevant consumption impacts of the program. Given the previous results by Foster and Rosenzweig (2003), we might suspect that there would be an increase in demand for forest products. Since the survey does not contain direct measures of timber demand, we use measures of new housing construction (number of rooms) as a proxy for timber demand. Previous literature on the consumption impacts of Progresa has indicated that the program increased the intake of meat and animal products (Hoddinott and Skoufias, 2004). Given the well-documented significant increase in the resources required to supply an animal-intensive diet (White (2000), Gerbens-Leenes and Nonhebel (2002), Bouma et al. (1998)) and the intense competition between cattle-rearing and forest resources in Mexico (Barbier and Burgess (1996), Kaimowitz (1995)) this seems a natural place to look for a demand-driven increase in pressure on forest cover. We therefore examine changes in consumption of beef and milk products.

As mentioned in Section 3.2, there is not necessarily a one-to-one relationship between the location of consumption changes and the corresponding production adjustment, but we might expect that some increased production could come directly from the treated households. We therefore assess changes by treated households in the number of cattle owned, number of plots of land that households report using for livestock grazing or agricultural purposes, and total area of all plots. Since these goods are also traded in markets, increased production could come from neighboring non-recipient households. Therefore, we also examine changes in production behavior by neighboring households were in treated localities but were not eligible for the program.

We would expect that the degree to which we should observe local production responses (and therefore local environmental consequences) depends on the extent to which local markets are connected. To this end, we will use road density (as measured by total kilometers of roads within a 10km buffer of each locality) as a proxy for market-connectedness. To test for heterogeneity, we include a second specification for each outcome variable which examines the interaction between treatment effect and the inverse road density in the locality (R_i):

$$y_{it} = \beta_0 + \beta_1 T_i + \beta_2 P_t + \beta_3 T_i P_t + \beta_4 R_i + \beta_5 R_i T_i + \beta_6 R_i P_t + \beta_7 R_i T_i P_i + \varepsilon_{it} \tag{3.5}$$

The coefficient β_7 measures the variation in the intention to treat effect according to infrastructure quantity.

3.4.2 Progresa results

The experimental household data confirm the findings in previous literature that Oportunidades strongly increased consumption of land-intensive resources (Hoddinott and Skoufias, 2004). Table 3.6 shows regression results for demand-side outcome variables. We see no increase in the direct demand for timber products in the context of the home improvements proxy, but we do see increases in beef and milk consumption. The estimated treatment effects represent increases relative to the baseline mean of 29% and 23%, respectively. The interactions with road density however show that these demand-side impacts do not vary significantly with the quality of local road networks–it appears as though the treatment effect on consumption of these resource-intensive goods is homogeneous across infrastructure quality.

Table 3.7 presents production-side results on number of cows, total hectares of land in production and number of plots in production. The baseline distribution of total hectares in production is highly skewed so we use the natural logarithm of this variable in both specifications. We do not see significant increases in the number of cows owned, plots used, or the total area cultivated by recipient households, nor do these effects vary with road density[14]. Progresa does not appear to provoke a substantial increase in agricultural production among beneficiary households, regardless of the level of isolation[15].

[14]The results indicate that we can rule out increases in land use and cow ownership greater than 9% and 18% respectively, with 95% confidence. Given the 29% and 23% increase in beef and milk consumption, it seems unlikely that recipient households are supplying their entire increase in demand. Skoufias (2005) documents a significant decrease in child labor (not surprising given the conditionality of the program). Since this type of labor is disproportionately used on the family farm, this provides a possible reason for why households eligible for Progresa/Oportunidades may produce less on their own household farms and consume more goods produced elsewhere

[15]This result would seem to contradict the findings of Gertler, Martinez and Rubio-Codina (2006). In that study the authors show that recipient households do invest a small portion of Oportunidades transfers in livestock and land.

The discussion in Section 3.2 motivates the analysis of market-mediated spillovers which may vary with the depth of local markets despite the very constant increases in consumption observed so far. In order to address this question using the Progresa data, we examine the extent to which non-recipient households (households that reside in eligible localities but who do not themselves qualify as poor) adjust their production behavior in response to the arrival of program transfers. In Table 3.8 we observe that while the program does not have significant effects on production in this group overall, in road-poor areas there is a significantly stronger increase in the number of hectares under cultivation and in the number of cows owned by non-recipient households. The estimate of the coefficients on the interaction of inverse road density with the spillover effect in Column 4 indicates less than a one-percent increase in hectares in production at the 90th percentile of road density, and a 1.2% and 3.2% increase at the median and 10th percentile, respectively. The estimate of the same interaction effect on the number of cows owned (Column 6) indicates a 3% and 5% increase in the number of cows owned when evaluated at the 90th percentile and the median respectively, and a 12% increase when evaluated at the 10th percentile.

The micro-data from the randomized pilot phase of the program therefore provide evidence that the consumption increases caused by Progresa were similar across localities with different connection to markets, but the corresponding production increases among nearby wealthier households were not. Specifically, in localities with good road infrastructure there is no production-side response among local ineligibles, but where poor infrastructure localizes economic activity the increased consumption engendered by the program is met by an increase in output. This is in accordance with our hypothesis that even homogenous treatment effects will appear heterogeneous when they are mediated by markets of different sizes.

Given these estimated consumption increases by households, are the deforestation impacts previously estimated of a reasonable magnitude? To explore this question we conducted a back-of-the-envelope calculation using the marginal effects on milk and beef consumption combined with estimates of consumption and the resource intensivity of cattle-raising to estimate the additional land required[16]. Our simulation indicates that the average locality would require

However, they aggregate all animals into two categories: "production" animals which include cows, pigs, chickens, turkeys; and "draft" animals (horses, oxen). While they do find a significant increase in ownership of production animals, this appears to be driven by landless and non-agricultural households in their sample, indicating that the increase is unlikely to be due to large animals. Our data confirm this. We concentrate on the demand for animal protein but previous studies also suggested a diversification of fruit and vegetable consumption in response to the program Hoddinott and Skoufias (2004) which could also increase deforestation.

[16]Our simulation assumes each household consumes a quarter gallon of milk and a pound of beef each day they consume it, that a beef cow produces 400 pounds of beef and a dairy cow 1500 gallons of milk per year – these numbers in the US are 500-650 and 2400, respectively. Given the Progresa treatment effects, this gives us a number

maintenance of eight additional cows, more than twice the number that Table 8 shows were being provided by ineligibles in local villages. This would suggest that even in isolated places more than half of demand was being satisfied from production outside of the locality. If we then estimate the land required to support these cows, we come to a figure roughly 20 times the observed deforestation estimated in column 3 of Table 5. This demonstrates that the measured consumption increases are more than large enough to account for the observed deforestation. That the predicted amount of land needed is larger than the observed effects is not surprising, both because much of the marginal land is likely not to be forested and because the market-mediated spillovers cause us to underestimate total treatment effects[17].

3.5 Heterogeneity in the Impact of Oportunidades

3.5.1 Road Density and Treatment Effects

If the most plausible mechanism underlying an increase in deforestation is increasing demand for land-intensive goods, we should expect to observe heterogeneity in estimated treatment impacts across localities consistent with this mechanism. To this end, we test for variation in estimated effects by the quality of local transportation infrastructure. We expect that the estimated impact of the program should be greater where the supply response is more localized by poor infrastructure.

The problem of estimating responses when shocks can be dispersed through market transactions suggests that we will be more likely to detect impacts where road networks are poor. Table 3.9 shows the apparent differential impact of treatment at different categories of road density. The first six columns divide the entire sample into three equal sized groups according to road density. Results are shown for both IV OLS and IV Tobit specifications. Here we observe that the program only has a significant positive local impact on deforestation where road densities are low. We also see much larger point estimates for the marginal effects on the

of beef cattle slaughtered over the 3-year period, and the incremental size of the dairy herd needed. We assume that 9 acres is needed to support a cow, and that the resource intensity of the counterfactual vegetable-based diet is 1/5th of the animal-based diet, and this gives us the additional number of square kilometers needed for the dietary change: just under a quarter of a square kilometer per locality. The simulation of observed average deforestation per locality multiplies locality size times the fraction of localities in the treatment group with any deforestation and the marginal effect where deforestation occurs. The estimated deforestation is roughly a hundredth of a square kilometer per locality.

[17]This market demand mechanism between treated and ineligible households within treatment villages provides an alternative channel for the well-documented spillover effects of Progresa. Rather than working through peer effects or insurance and credit markets (Angelucci and de Giorgi, 2009), ineligible households may have realized benefits by increasing output to satisfy local demand.

probability of deforestation for the low road density class. The results are nearly identical for the restricted sample (not shown). Columns 7 and 8 interact treatment with low road density for the full sample. We find the percent deforested difference to be marginally significant in the Tobit estimation (although the marginal effects in the low-density areas are several times those in the other groups), but isolated localities have a significantly higher probability of seeing some deforestation. The coefficient estimates for the sum of the interaction with treatment in both the tobit and OLS estimates are almost identical to the estimates from the low road density sample.

3.5.2 Spatial ACFs in a RD framework

An alternate test of our hypothesis that production is sourced from surrounding markets is to examine the spatial contours of program effects directly. Since treatment is potentially endogenous, we cannot calculate spatial lag functions in the standard way. Instead, we adapt techniques introduced by Conley and Topa (2002) to the regression discontinuity framework. This mirrors the logic of the discontinuity analysis in that while the distribution of outcomes may be endogenous across the broader distribution of the eligibility score, it is plausibly exogenous within a window around the discontinuity.

The underlying information used here is the same as that used in the discontinuity analysis, but the structure of the data is slightly different. Here we divide the country in a grid of equally-sized cells of 10x10 km. For each cell we calculate deforestation and a “saturation" of treatment, which is composed of a ratio where the numerator is the number of villages out of the “study" localities that receive Oportunidades and the denominator is the number of “study” localities in the cell. We define a study village as one which is in the restricted subsample that we used for the discontinuity analysis, i.e., one which is located between -2 and -0.2 on the poverty index. This provides a conservative way of using “as if random” saturation in the intensity of treatment in the window around the discontinuity to measure spillover effects.

s_{i0} represents this saturation ratio in each cell, which we refer to as “own" saturation. For each cell, we then calculate saturation for all of the neighboring cells, excluding the own cell (saturation at 10 kilometers, s_{i10}). We proceed outwards in a similar fashion, calculating saturation in successive rings around a given cell up to 40 kilometers. We also calculate the density of road networks in the 50 kilometers surrounding each cell. We call this variable c_i and interact it with each of the saturation variables to help us understand how road access might affect the probability of deforestation. For areas which have no “study” localities in them, we include a dummy variable equal to one when there are no localities, and for these observations include zeros in the saturation observations[18]. We then drop all cells with no baseline forest cover and estimate:

$$d_i = \alpha + \sum_{k=0,10,20,30,40} [\beta_k s_{ik} + \theta_k s_{ik} c_i] + \Gamma X_i + \varepsilon_i, \tag{3.6}$$

where $d_i = 1$ if there is deforestation in the cell, s_{ik} is the saturation at each distance, c_i is road density, X_i are control variables including average poverty level, road density within 0-50 kilometers around cell, latitude and longitude fixed effects, and baseline forest. ε_i is the error term. We calculate standard errors using bootstrapping in order to avoid the problem of spatial autocorrelation of error terms (for a discussion of spatial autocorrelation in the probit, tests, and estimation strategies, see Pinkse and Slade (1998)). Our theory tells us that deforestation should be most strongly correlated with nearby treatment intensity where infrastructure is poorest.

3.5.3 Spatial analysis: Results

The results from the spatial regression are shown in Table 3.10. The table contains only partial results – in all cases, 10 latitude/longitude fixed effects and the mean poverty level in each buffer is included, along with the variables indicating zero observations in a buffer. The fixed effects capture spatial variation in ecosystem, as well as cultural heterogeneity, to the extent that it varies geographically in Mexico. We use two variables capturing infrastructure quality: the natural log of total road density (measured as total length of roads in all the cells around a sample cell), and a dummy variable equal to 1 if the density is less than the median[19]. In the simplest specification, which does not include interactions of saturations with road density, saturations have no significant effect on the probability of deforestation. In the two versions where interactions are included, however, we observe that road density is very important in determining the effect of program concentration on deforestation, but that the key determinant is the interaction of saturation with infrastructure. In both cases, in more remote areas (those with low road density), the probability of deforestation as a result of Oportunidades recipients nearby increases.

Figure 3.5 graphs out the reported coefficients from column (2) by distance, calculating the interaction effects at 90% road density ("high") and at 10% road density ("low")[20]. The horizontal axis indicates the distance to the baseline cell in kilometers, and the plots include dotted lines indicating 95% confidence intervals. At each cell distance, the marginal effect is

[18]This follows Foster and Rosenzweig (2003)'s approach for dealing with missing data.
[19]Results are robust to various cutoff points less than the median as well.
[20]The graph looks nearly identical using coefficients from column (3).

calculated for a one standard deviation change in saturation. This provides a visual image of the effect of the program on deforestation according to distance, and shows that the spatial contour of deforestation is not significantly different from zero with respect to the location of treatment for well-connected cells, whereas in isolated cells the deforestation effect is more localized – increases in saturation increase the probability of deforestation, but at a decreasing rate. The impact of increases in saturation goes to zero at the 20-30 kilometer band. This confirms our hypothesis that good infrastructure may help spread the impacts of the program to the point where they are non-detectable locally.

In summary, the results discussed above are consistent with the framework introduced in Section 2. Oportunidades appears to induce greater consumption of resource-intensive goods everywhere, and hence increases pressure on resources regardless of network quality. However, since treatment does not increase output among recipient households, this additional demand is mediated through market networks. With poor transportation infrastructure, demand must be met locally and so we see greater production responses. Where infrastructure is better, increases in demand will be sourced from a greater variety of locations.

3.6 Conclusions

This paper conducts an analysis of the impact of large income transfers on deforestation, taking advantage of the discontinuity created by the eligibility rule for Oportunidades. We find that the income transfer increases deforestation, at least in the population that is just below the marginality level required to be able to receive payments. We then use household data to test for a plausible mechanism consistent with this increase in forest loss. Here we observe that households increase their consumption of two relatively land-intensive goods – beef and milk. We do not detect a corresponding increase in consumption of a good that might increase forest cover through increasing demand for forest products– housing construction. Nor do we detect consistent changes on the production side triggered by exposure to Progresa, suggesting that the observed deforestation effects of the program arise from consumption changes, in other words through an expansion of each household's "ecological footprint" of land use.

Average household income increases by one-third as a result of the transfers, which leads the probability of deforestation to nearly double and the rate of deforestation among deforesters to increase by 15 to 33 percent. These increases are significant in the entire sample, but are strongest in places with poor infrastructure. These results underline the importance of considering spatial spillovers in the analysis of micro-experiments, and provide no support for the argument that increasing incomes will translate into improved environmental outcomes. Although we demonstrate that there were potential negative secondary environmental effects of the Oportunidades program, we cannot draw firm overall welfare conclusions. Welfare losses due to deforestation may have been outweighed by the health and education benefits of the Oportunidades program. In addition, a full welfare analysis of the program would take into account how long-term changes in income might affect environmental quality. Income growth may improve education or institutional quality, potentially leading to better environmental outcomes in the long term.

In recent years the use of local average treatment effects in the analysis of development program impacts has come under fire for answering small questions using a non-representative sample, and for obfuscating important sources of heterogeneity in outcomes (Deaton, 2009). Although we estimate local average treatment effects in this paper, our use of the national rollout means that we have a very large and heterogeneous sample at the discontinuity. Therefore we are able to exploit the jump in program participation to cleanly identify impacts of poverty reduction but also to investigate a critical source of heterogeneity. Furthermore, the eligibility cutoff that we use for identification in this paper is close to the extensive margin of the actual program, and hence measures plausibly the impact of expanding the current program, as in Karlan and Zinman (2009). Hence we submit that the treatment effect estimated in this paper is both policy relevant and has substantial richness in terms of the analysis of heterogeneity.

In terms of the generalizability of these results, it is important to recognize the dimensions in which impacts of a CCT program may not reproduce the dynamics of a more endogenous long-term increase in income. Most obvious is the conditionality; it explicitly seeks to alter the prices faced by households in the use of one input to production, child labor. The program also features conditionality on regular health checkups for beneficiary children, and this increase in focus on their health may lead to dietary changes that would not be replicated with a simple increase in income. Further, Oportunidades payments are made monthly and hence provide a cash flow that may be more suited to consumption than investment. It is quite possible, for example, that an alternative program delivering the same total amount of cash to beneficiary households in one lump sum would have seen more investment and less consumption, particularly if credit markets are imperfect. Finally, no particular household receives Oportunidades payments for longer than they have children of eligible age, and so the program features a rolling beneficiary pool and is not likely to generate the real wealth effects that would be seen if permanent income had increased. Despite these caveats, CCT programs have emerged as a major policy tool in the fight against global poverty, and so to the extent that they present one of the most obvious policy levers for decreasing poverty our results are relevant even if we interpret impacts as limited to these programs.

Our findings, particularly the spatial contours of estimated treatment effects, motivate the idea that transportation infrastructure plays a critical role in determining the location of environmental impacts–i.e. where the “ecological footprint" lands. This underlines the empirical issues generated by spatial spillover effects when we examine the production response to market-mediated increases in local demand. A well-established result in the literature on rainfall shocks and on famines is the idea that infrastructure decreases the correlation between localized shocks and local market prices (Keller and Shiue, 2008; Donaldson, 2009). Extended to a program evaluation context, this logic suggests that when treatment is administered at small spatial units, market-driven spillovers cause an underestimation of the true harm from treatment. By this logic, the strong deforestation impact seen in isolated parts of Mexico when treated with Oportunidades is troubling, because it is precisely in these environments that we are closest to capturing the full impact of treatment. We therefore see these results not as a criticism of poverty-alleviation programs but rather as a cautionary tale. Should we wish to achieve increases in wealth simultaneously with improvements in environmental quality, our study suggests that carefully designed environmental management schemes should accompany poverty alleviation programs.

Chapter 3, in full, has been submitted for publication of the material as it may appear in the Review of Economics and Statistics. Alix-Garcia, Jennifer; McIntosh, Craig; Sims, Katharine R.E.; Welch, Jarrod R. The dissertation author shared equally in the research design, analysis and authorship.

3.7 Figures

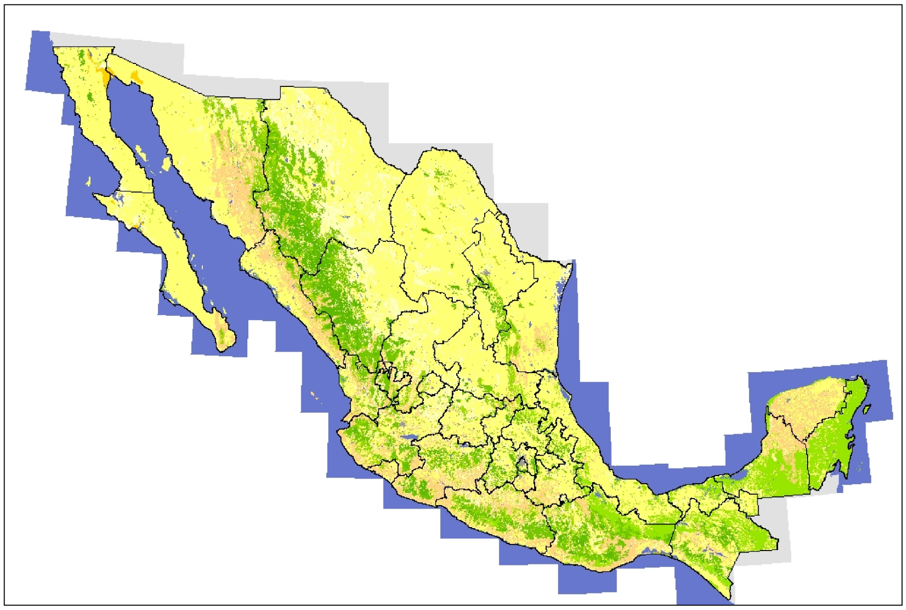

Figure 3.1: Forest Cover in Mexico, 2000

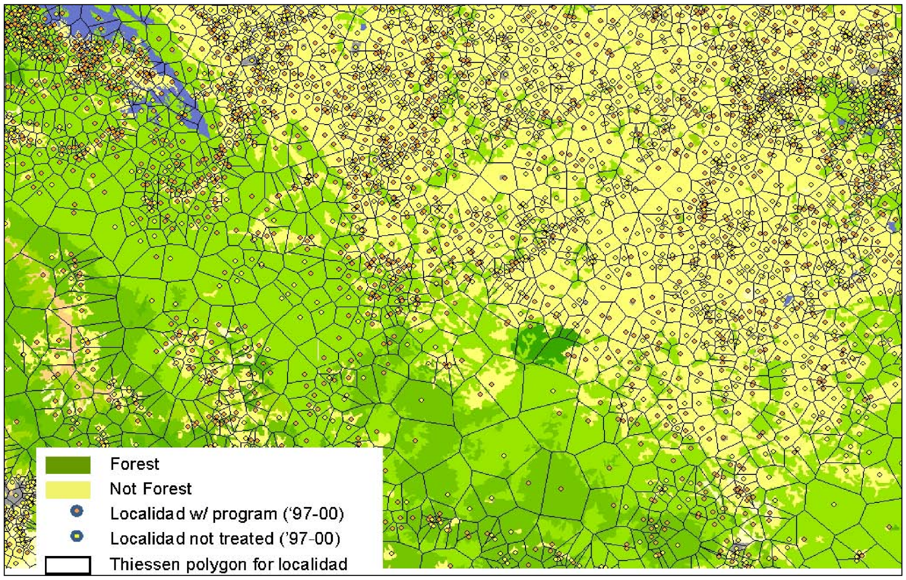

Figure 3.2: Thiessen Polygons

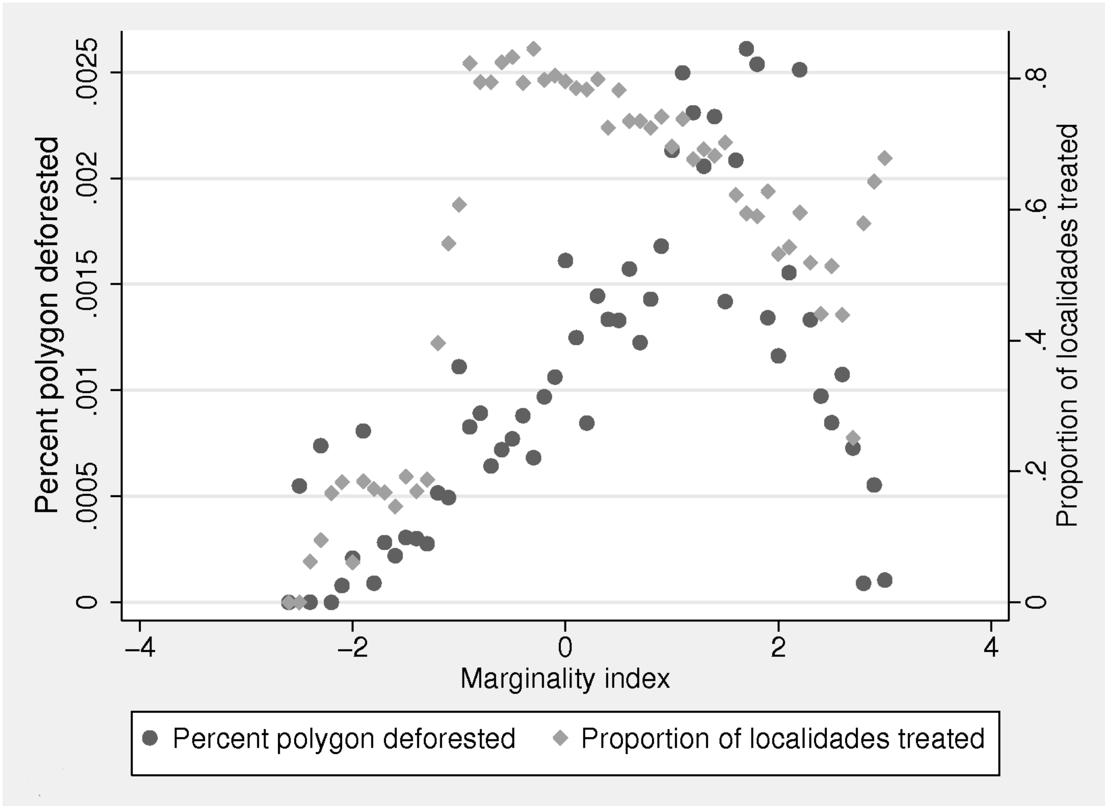

Figure 3.3: Entire sample minus observations with index > 3 (51 observations missing)

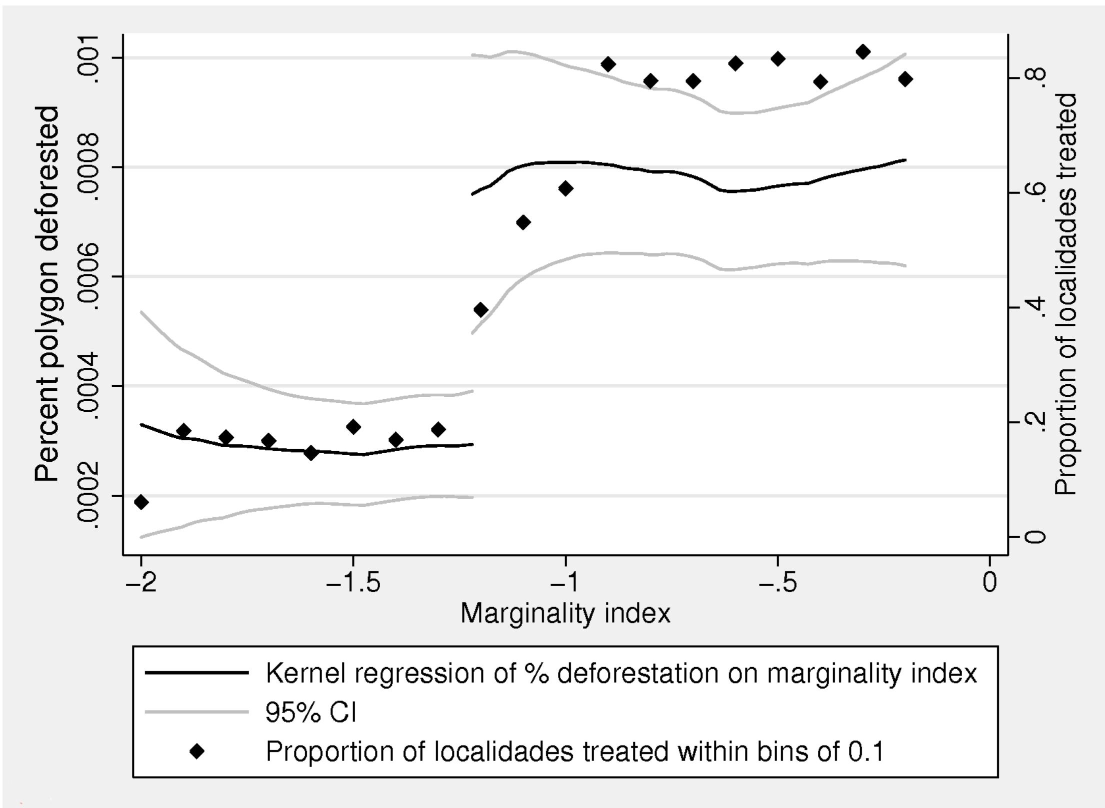

Figure 3.4: Kernel estimation of deforestation on marginality index – restricted sample

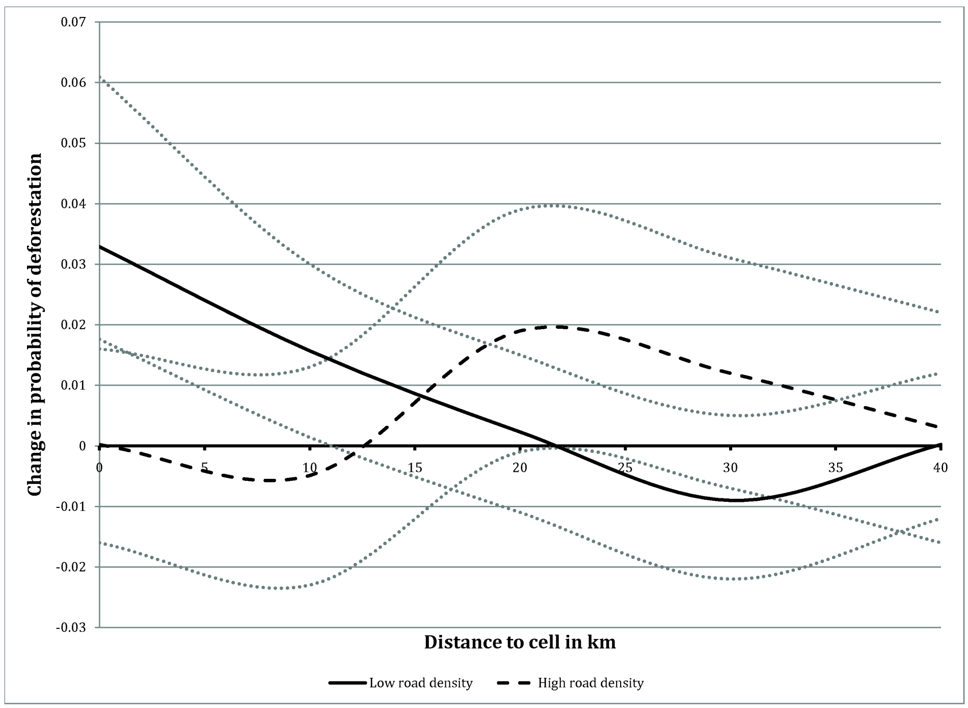

Figure 3.5: Own deforestation probability as a function of treatment within distance bands

3.8 Tables

Table 3.1: Summary statistics across eligibility

	Non-eligible <-1.2	Eligible $>= -1.2$	Test of difference	Normalized difference
Full sample				
Polygon area (km^2)	37.9	18.9	18.17	-.163
Average slope in polygon (degrees)	5.63	9.63	34.4	.482
% polygon forested in 2000	12.1	10.5	3.24	0.035
Km roads in 10 km buffer	47.0	32.7	32.7	-.36
% polygon polygon deforested	.0003	.0014	6.78	.11
Proportion with deforestation	.048	.098	9.64	
Observations	3510	55077		
Restricted sample				
Polygon area (km^2)	37.9	25.6	7.43	-.095
Average slope in polygon (degrees)	5.61	6.95	12.5	.18
Percent forested in 2000	12.2	10.4	3.37	-.042
Km roads in 10 km buffer	46.4	41.2	9.88	-.129
Proportion polygon deforested	.0003	.0008	4.14	.139
Proportion with deforestation	.049	.072	4.89	
Observations	3350	12408		

Table 3.2: Simple approach – eligibility as proxy

	Tobit			OLS	
	% polygon deforested			Deforestation (0/1)	% deforested if 1
	(1)	(2)	(3)	(4)	(5)
Eligible	.004	.005	.004	.013	.004
	(.002)**	(.003)*	(.002)*	(.008)*	(.002)*
Marginality index	.005	.008	.002	.031	.0008
	(.0004)***	(.0008)***	(.002)	(.003)***	(.0008)
Index^2		.0006		.002	.0005
		(.0007)		(.003)	(.0008)
Index^3		-.001		-.004	-.0002
		(.0004)***		(.001)***	(.0003)
Index^4		-.00002		-.0001	-.0001
		(.0002)		(.0005)	(.0001)
Baseline area in forest, 2000	-3.72e-06	-4.78e-06	.00004	.0006	.00005
	(9.77e-06)	(9.78e-06)	(.00002)**	(.0001)***	(1.00e-05)***
Ln(polygon area)	.010	.010	.007	.046	-.010
	(.0004)***	(.0004)***	(.0007)***	(.002)***	(.0006)***
Ln(total population in 1995)	.001	.001	.0004	.010	-.0004
	(.0002)***	(.0002)***	(.0003)	(.001)***	(.0003)*
Ln(slope)	-.0005	-.0005	-.00009	-.003	-.0003
	(.00005)***	(.00005)***	(.0001)	(.0002)***	(.00006)***
Ln(road density)	-.0006	-.0006	.0003	-.004	-.0001
	(.0003)**	(.0003)**	(.0005)	(.001)***	(.0003)
Obs.	58587	58587	15758	58587	5551
Ecoregion controls	yes	yes	yes	yes	yes
Marginal effects of eligibility					
$\Pr(y > 0)$	.011	.015	.011	.013	
	(.005)**	(.021)**	(.007)*	(.008)*	
$y > 0$	.0006	.0008	.0005		.004
	(.0003)**	(.00042)**	(.0003)*		(.002)*

In column (4) the dependent variable is an indicator for any deforestation, and in column (5) is percent polygon deforested, but only for those polygons experiencing positive deforestation. Standard errors in parentheses. * significant at 10%; ** significant at 5%.

Table 3.3: First stage regressions (OLS)

	Full sample			Restricted sample	Proportion treated
	(1)	(2)	(3)	(4)	(5)
Eligible	.645 (.008)***	.621 (.040)***	.843 (.097)***	.676 (.046)***	.751 (.060)***
Marginal		1.117 (.087)***	1.041 (.089)***	1.077 (.087)***	.351 (.052)***
Marginal x index		1.331 (.084)***	1.156 (.088)***	1.194 (.085)***	.416 (.051)***
Eligible x index		-.063 (.025)**	.222 (.081)***	.078 (.034)**	.427 (.051)***
Marginality index	.007 (.002)***	.041 (.025)*	-.189 (.083)**	.021 (.029)	-.329 (.052)***
Index^2			-.046 (.005)***		-.073 (.004)***
Index^3			.006 (.005)		.010 (.003)***
Index^4			-.001 (.001)		.0003 (.0009)
Baseline area in forest, 2000	.0003 (.00009)***	.0003 (.00009)***	.0003 (.00008)***	.0006 (.0001)***	.0002 (.00007)***
Ln(polygon area)	-.030 (.002)***	-.029 (.002)***	-.029 (.002)***	-.027 (.004)***	-.034 (.002)***
Ln(total population in 1995)	.158 (.001)***	.158 (.001)***	.159 (.001)***	.129 (.002)***	.119 (.0008)***
Ln(slope)	.004 (.0003)***	.004 (.0003)***	.003 (.0003)***	.003 (.0005)***	.003 (.0002)***
Ln(road density)	.028 (.001)***	.028 (.001)***	.028 (.001)***	.009 (.003)***	.016 (.001)***
Obs.	58587	58587	58587	15758	58587
Adjusted R-squared	.314	.330	.334	.462	.336
Ecosystem controls	yes	yes	yes	yes	yes
F-test of instruments		2017	425	349	239

The dependent variable in columns (1)-(4) is equal to 1 if the locality received Oportunidades before 2004, and 0 otherwise.
In column (5), the dependent variable is the average proportion of households receiving Oportunidades from 1997 to 2003, inclusive.
Robust standard errors in parentheses. * significant at 10%;** significant at 5%; *** significant at 1%

Table 3.4: Simple discontinuity approach – instrumentation with eligibility

	IV Tobit				IV OLS	
	Full estimation sample			Restricted sample	Deforestation (0/1)	% deforested if 1
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	.006 (.003)**	.013 (.007)*		.010 (.006)*	.031 (.019)*	.013 (.007)*
Proportion treated			.038 (.021)*			
Marginality index	.005 (.0004)***	.006 (.001)***	.002 (.003)	-.0007 (.003)	.028 (.003)***	-.00005 (.001)
Index^2		.002 (.001)	.004 (.002)*		.004 (.004)	.002 (.001)
Index^3		-.0009 (.0003)***	-.0005 (.0003)		-.003 (.001)***	-.0004 (.0003)
Index^4		-.0001 (.0002)	-.0004 (.0002)*		-.0003 (.0005)	-.0002 (.0001)
Baseline area in forest, 2000	-5.11e-06 (9.75e-06)	-7.95e-06 (9.96e-06)	-1.00e-05 (1.00e-05)	.00003 (.00002)**	.0006 (.0001)***	.00004 (1.00e-05)***
Ln(polygon area)	.010 (.0004)***	.010 (.0005)***	.011 (.0008)***	.008 (.0007)***	.047 (.002)***	-.009 (.0007)***
Ln(total population in 1995)	.0005 (.0005)	-.0006 (.001)	-.003 (.002)	-.001 (.0009)	.005 (.003)	-.003 (.001)**
Ln(slope)	-.0005 (.00005)***	-.0006 (.00006)***	-.0006 (.00008)***	-.0001 (.0001)	-.003 (.0002)***	-.0003 (.00007)***
Ln(road density)	-.0008 (.0003)***	-.0009 (.0003)***	-.001 (.0004)***	.0002 (.0005)	-.005 (.001)***	-.0005 (.0004)
Obs.	58587	58587	58587	15758	58587	5545
Ecoregion controls	yes	yes	yes	yes	yes	yes
Marginal effects of treatment						
$\Pr(y > 0)$	.018 (.008)**	.038 (.019)**	.12 (.067)*	.030 (.017)*	.031 (.019)*	
$y > 0$	.0009 (.0004)**	.002 (.001)*	.005 (.003)*	.001 (.0008)*		.013 (.007)*

Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3.5: Fuzzy discontinuity estimates

	IV Tobit				IV OLS	
	Full estimation sample			Restricted sample	Deforestation (0/1)	% deforested if 1
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	.006 (.003)**	.012 (.006)**		.010 (.005)**	.035 (.015)**	.009 (.005)*
Proportion treated			.022 (.012)*			
Marginality index	.005 (.0004)***	.007 (.0009)***	.005 (.002)**	-.0007 (.003)	.028 (.003)***	.0002 (.0009)
Index^2		.002 (.001)	.002 (.001)		.005 (.003)	.001 (.001)
Index^3		-.0009 (.0003)***	-.0006 (.0003)**		-.003 (.001)***	-.0002 (.0003)
Index^4		-.00008 (.0002)	-.0002 (.0002)		-.0003 (.0005)	-.0002 (.0001)
Baseline area in forest, 2000	-5.06e-06 (9.75e-06)	-7.74e-06 (9.89e-06)	-9.81e-06 (1.00e-05)	.00004 (.00002)**	.0006 (.0001)***	.00005 (1.00e-05)***
Ln(polygon area)	.010 (.0004)***	.010 (.0005)***	.010 (.0006)***	.008 (.0007)***	.047 (.002)***	-.010 (.0006)***
Ln(total population in 1995)	.0006 (.0005)	-.0005 (.0009)	-.001 (.001)	-.001 (.0007)	.004 (.003)	-.002 (.0008)**
Ln(average slope)	-.0005 (.00005)***	-.0006 (.00005)***	-.0006 (.00007)***	-.0001 (.0001)	-.003 (.0002)***	-.0003 (.00006)***
Ln(road density)	-.0007 (.0003)***	-.0009 (.0003)***	-.0009 (.0003)***	.0002 (.0005)	-.005 (.001)***	-.0004 (.0003)
Obs.	58587	58587	58587	15758	58587	5545
Ecoregion controls	yes	yes	yes	yes	yes	yes
Marginal effects of treatment						
$Pr(y > 0)$	.018 (.007)***	.037 (.016)**	.070 (.039)*	.030 (.013)**	.035 (.015)**	
$y > 0$	.0009 (.0004)**	.002 (.0008)**	.003 (.0019)*	.001 (.0006)**		.009 (.005)*

Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3.6: Household-level Consumption Impacts, Progresa

	Rooms in home		Days Ate Beef		Days Drank Milk	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	.014	.017	.114	.118	.337	.331
	(.033)	(.035)	(.030)***	(.031)***	(.081)***	(.087)***
Treatment x inverse road density		-.034		-.070		.183
		(.148)		(.097)		(.669)
Village chosen to receive Progresa	.0001	.002	-.025	-.031	-.133	-.143
	(.037)	(.038)	(.029)	(.030)	(.111)	(.118)
Post treatment year	.053	.049	-.137	-.138	-.655	-.664
	(.028)*	(.029)*	(.024)***	(.025)***	(.061)***	(.065)***
Inverse of road density		.266		-.156		.051
		(.169)		(.069)**		(.499)
Village x inverse road density		.043		.102		.232
		(.236)		(.140)		(.682)
Post treatment x inverse road density		.067		.016		.155
		(.140)		(.068)		(.252)
Obs.	23318	23318	33128	33128	33128	33128
Mean dependent variable in baseline	1.557		0.388		1.440	
	(0.930)		(0.661)		(2.367)	

* significant at 10% ** significant at 5%; *** significant at 1%

Table 3.7: Household-level Production Impacts, Progresa

	No. of Plots		Log 1+ Total Hectares		No. of Cows	
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment effect	.030	.031	-.014	-.015	.092	.036
	(.039)	(.040)	(.038)	(.039)	(.057)	(.057)
Treatment x inverse road density		-.107		.142		.936
		(.210)		(.223)		(.522)*
Village chosen to receive Progresa	.014	.037	-.004	.017	-.004	.058
	(.056)	(.057)	(.040)	(.040)	(.087)	(.085)
Post treatment year	-.094	-.077	.312	.317	-.239	-.180
	(.032)***	(.033)**	(.033)***	(.033)***	(.046)***	(.046)***
Inverse of road density		.833		.820		2.122
		(.161)***		(.227)***		(.799)***
Village x inverse road density		-.263		-.217		-.760
		(.317)		(.258)		(.872)
Post treatment x inverse road density		-.275		-.235		-.982
		(.149)*		(.128)*		(.402)**
Obs.	45087	45087	32631	32631	34248	34248
Mean dependent variable in baseline	0.824		1.724		0.604	
	(0.955)		(3.535)		(2.304)	

* significant at 10% ** significant at 5%; *** significant at 1%

Table 3.8: Local Spillover Impacts of Progresa
Impacts on Ineligible Households in Treatment Villages

	No. of Plots		Log 1+ Total Hectares		No. of Cows	
	(1)	(2)	(3)	(4)	(5)	(6)
Spillover effect	-.001	-.017	-.037	-.052	.153	-.021
	(.038)	(.040)	(.041)	(.042)	(.125)	(.125)
Spillover x inverse road density		.372		.535		3.605
		(.240)		(.167)***		(1.243)***
Village chosen to receive Progresa	.042	.094	-.015	.022	-.121	.052
	(.055)	(.056)*	(.047)	(.047)	(.219)	(.215)
Post treatment year	-.208	-.202	.254	.256	-.702	-.551
	(.028)***	(.028)***	(.034)***	(.034)***	(.108)***	(.106)***
Inverse of road density		1.009		1.207		6.036
		(.196)***		(.387)***		(2.021)***
Village x inverse road density		-1.051		-.620		-2.846
		(.249)***		(.430)		(2.395)
Post treatment x inverse road density		-.119		-.257		-3.060
		(.159)		(.122)**		(1.180)***
Obs.	40569	40569	30068	30068	31184	31184
Mean dependent variable in baseline	1.031		2.844		1.577	
	(1.667)		(5.322)		(4.675)	

* significant at 10% ** significant at 5%; *** significant at 1%

Table 3.9: Deforestation and infrastructure

	Low density		Medium density		High density		Interactions	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	%	0/1	%	0/1	%	0/1	%	0/1
Treated	.016 (.009)*	.075 (.037)**	.006 (.008)	.019 (.030)	.018 (.015)	.023 (.021)	.008 (.006)	.008 (.015)
Treated x low road density							.010 (.005)*	.059 (.017)***
Low road density							-.004 (.004)	-.028 (.012)**
Obs.	19529	19529	19529	19529	19529	19529	58587	58587
Ecoregion controls	yes	yes	yes	yes	yes	yes	yes	yes
Marginal effects of treatment								
Pr($y > 0$)	.080 (.040)**		.020 (.030)		.026 (.019)			
$y > 0$	.002 (.001)**		.001 (.001)		.002 (.002)			
							.018^{+} (.007)***	.067^{+} (.020)***

Partial results. All estimations contain full set of covariates.
Standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.
+ indicates the sum of the treatment coefficient with the treatment and low road density interaction.

Table 3.10: Spatial regressions

Dependent variable = 1 if deforestation

	(1)	(2)	(3)
Own saturation	.017	.192	-.006
	(.018)	(.044)***	(.019)
Within 10-20 km	.025	.120	-.027
	(.019)	(.047)**	(.033)
Within 20-30 km	.038	-.070	.067
	(.023)*	(.067)	(.056)
Within 30-40 km	-.006	-.154	.066
	(.028)	(.079)*	(.076)
Within 40-50 km	.012	-.014	-.068
	(.031)	(.080)	(.075)
Ln(road density, 0-50km)	-.027	-.081	
	(.016)*	(.039)**	
Baseline forest	.001	.001	.001
	(.0002)***	(.0002)***	(.0002)***
Density x own saturation		-.090	
		(.020)***	
Density x 10-20 km		-.063	
		(.029)**	
Density x 20-30 km		.068	
		(.045)	
Density x 30-40 km		.099	
		(.051)*	
Density x 40-50 km		.013	
		(.054)	
Density $<$ median			-.061
			(.044)
Density $<$ median x own saturation			.058
			(.020)***
Density $<$ median x 10-20 km			.069
			(.032)**
Density $<$ median x 20-30 km			-.043
			(.058)
Density $<$ median x 30-40 km			-.093
			(.079)
Density $<$ median x 40-50 km			.095
			(.078)
Obs.	11007	11007	11007
R^2	.195	.198	.196
Lat-long fixed effects	yes	yes	yes

OLS with bootstrapped standard errors. ** significant at 5%; *** significant at 1%.

Bibliography

Angelsen, Arild and David Kaimowitz, "Rethinking the Causes of Deforestation: Lessons from Economic Models," *World Bank Res Obs*, February 1 1999, *14* (1), 73–98.

Angelucci, M and G de Giorgi, "Indirect effects of an aid program: how do cash injections affect ineligibles' consumption," *American Economic Review*, 2009, *99*, 486–508.

Auffhammer, Maximilian, V. Ramanathan, and Jeffrey R. Vincent, "Integrated model shows that atmospheric brown clouds and greenhouse gases have reduced rice harvests in India," *Proceedings of the National Academy of Sciences*, 2006, *103* (52).

Baland, Jean-Marie, Pranab Bardhan, Sanghamitra Das, Dilip Mookherjee, and Rinki Sarkar, "The Environmental Impact of Poverty: Evidence from Firewood Collection in Rural Nepal," 2007.

Barbier, E. B. and J. C. Burgess, "The Economics of Tropical Deforestation," *Journal of Economic Surveys*, July 2001, *15* (3), 413–433.

Barbier, Edward B. and Joanne C. Burgess, "Economic analysis of deforestation in Mexico," *Environment and Development Economics*, 1996, *1* (02), 203–239.

Barham, Tania, "A Healthier Start: the Effect of Conditional Cash Transfers on Neonatal and Infant Mortality in Rural Mexico," 2009. Working Paper, University of Colorado, Boulder.

Bouma, J., N. H. Batjes, and J. J. R. Groot, "Exploring land quality effects on world food supply," *Geoderma*, 1998, *86* (1-2), 43 – 59.

Chomitz, Ken, *At Loggerheads?: Agricultural Expansion, Poverty Reduction, And Environment in the Tropical Forests*, World Bank, Washington, D.C., 2006.

Conley, Timothy and Giorgio Topa, "Socio-Economic Distance and Spatial Patterns in Unemployment," *Journal of Applied Econometrics*, 7 2002, *17* (4), 303–327.

Cropper, Maureen and Charles Griffiths, "The Interaction of Population Growth and Environmental Quality," *The American Economic Review*, May 1994, *84* (2, Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association), 250–254.

Dai, Aiguo, Kevin E. Trenberth, and Thomas R. Karl, "Effects of Clouds, Soil Moisture, Precipitation, and Water Vapor on Diurnal Temperature Range," *Journal of Climate*, 1999, *12* (8), 2451–2473.

Dasgupta, Susmita, Benoit Laplante, Hua Wang, and David Wheeler, "Confronting the Environmental Kuznets Curve," *The Journal of Economic Perspectives*, Winter 2002, *16* (1), 147–168.

Deaton, Angus, "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," 2009. NBER Working Paper No. 214690.

Deininger, Klaus and Bart Minten, "Determinants of Deforestation and the Economics of Protection: An Application to Mexico," *American Journal of Agricultural Economics*, 2002, *84* (4), 943–60.

Deininger, Klaus W. and Bart Minten, "Poverty, Policies, and Deforestation: The Case of Mexico," *Economic Development and Cultural Change*, 01/01 1999, *47* (2), 313–344.

Deschenes, Olivier and Michael Greenstone, "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *The American Economic Review*, 2007, *97.*

Dey, M. and M. Hossain, "Yield potentials of modern rice varieties: an assessment of technological constraints to increase rice production," in "Proceedings of the Final Workshop of the Projections and Policy Implications of Medium- and Long-Term Rice Supply and Demand Project" April 24-26 1995.

Dinar, A., R. Mendelsohn, R. Evenson, J. Parikh, A. Sanghi, K. Kumar, J. McKinsey, and S. Lonergen, "Measuring the Impact of CLimate Change on Indian Agriculture," *World Bank - Technical Papers*, 1998, *402.*

Dobermann, A. and C. Witt, *Increasing Productivity of Intensive Rice Systems Through Site-Specific Nutrient Management*, Science Publishers and IRRI, 2004.

_ **and K. Cassman**, *Encyclopedia of Plant and Crop Science*, Marcel Dekker, 2004.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," 2009. Working paper, MIT Department of Economics.

Evans, L.T. and S.K. De Datta, "The relation between irradiance and grain yield of irrigated rice in the tropics, as influenced by cultivar, nitrogen fertilizer application and month of planting," *Field Crops Research*, 1979, *2*, 1 – 17.

FAO, "FAO Global Forest Resources Assessment, 2005," 2005.

Felkner, John, Kamilya Tazhibayeva, and Robert M. Townsend, "Impact of Climate Change on Rice Production in Thailand," *American Economic Review*, 2009, *99* (2), 205–10.

Fernald, Lia CH, Paul J. Gertler, and Lynnette M. Neufeld, "Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades," *The Lancet*, 2008, *371* (9615), 828–837.

Fisher, Monica, Gerald E. Shively, and Steven Buccola, "Activity Choice, Labor Allocation, and Forest Use in Malawi," *Land Economics*, Nov. 2005, *81* (4), 503–517.

Foster, Andrew D. and Mark R. Rosenzweig, "Economic Growth And The Rise Of Forests," *The Quarterly Journal of Economics*, 2003, *118* (2), 601–637.

Gerbens-Leenes, P. W. and S. Nonhebel, "Consumption patterns and their effects on land required for food," *Ecological Economics*, 2002, *42* (1-2), 185 – 199.

Gertler, Paul J., Sebastian Martinez, and Marta Rubio-Codina, "Investing Cash Transfers to Raise Long Term Living Standards," *World Bank Policy Research*, 2006.

Green, Tina, "Do Social Transfer Programs Affect Voter Behavior? Evidence from PROGRESA in Mexico, 1997 ï£¡- 2000," 2005.

Grossman, Gene and Alan Krueger, "Economic Growth and the Environment," *Quarterly Journal of Economics*, 1995, *110*, 353–377.

Guiteras, Raymond, "The impact of climate change on Indian agriculture," *Working paper*, 2009.

Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 2001, *69*, 201–09.

Harbaugh, William, Arik Levinson, and Dave Wilson, "Reexamining the Empirical Evidence for an Environmental Kuznets Curve," *Review of Economics and Statistics*, 2002, *84*, 541–555.

Hoddinott, John and Emmenuel Skoufias, "The Impact of PROGRESA on Food Consumption," *Economic Development and Cultural Change*, 2004, *53*, 37–61.

Huang, Yan, Robert E. Dickinson, and William L. Chameides, "Impact of aerosol indirect effect on surface temperature over East Asia," *Proceedings of the National Academy of Sciences*, 2006, *103* (12), 4371–4376.

Imbens, Guido W. and Thomas Lemieux, "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 2 2008, *142* (2), 615–635.

IPCC, "Fourth Assessment Report," 2007. Intergovernmental Panel on Climate Change. http://www1.ipcc.ch/.

Jacob, Brian A. and Lars Lefgren, "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis," *Review of Economics and Statistics*, 02/01 2004, *86* (1), 226–244.

Kaimowitz, David, "Livestock and deforestation in Central America in the 1980s and 1990s: a policy perspective," Technical Report 9 1995.

__ , "The prospects for Reduced Emissions from Deforestation and Degradation (REDD) in Mesoamerica," *International Forestry Review*, 2008, *10* (3), 485–495.

Karlan, Dean and Jon Zinman, "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts," *Review of Financial Studies*, 2009.

Keller, Wolfgang and Carol Shiue, "Tariffs, Trains, and Trade: The Role of Institutions versus Technology in the Expansion of Markets," 2008.

Kelly, David L., Charles D. Kolstad, and Glenn T. Mitchell, "Adjustment costs from environmental change," *Journal of Environmental Economics and Management,*, 2005, *50* (3), 468–495.

Lee, David, Enrico Moretti, and Matthew Butler, "Do Voters Affect or Elect Policies? Evidence from the US House," *Quarterly Journal of Economics*, 2004, pp. 807–859.

Lobell, David B., "Changes in diurnal temperature range and national cereal yields," *Agricultural and Forest Meteorology*, 2007, *145* (3-4), 229 – 238.

_ **and Christopher B. Field**, "Global scale climateŰcrop yield relationships and the impacts of recent warming," *Environmental Research Letters*, 2007, *2* (1), 014002.

Londhe, S. Daniel D. B. Jadhav B. Padma Kumari A. L., "Observational evidence of solar dimming: Offsetting surface warming over India," *Geophysical Research Letters*, 2007, *34* (21).

MEA, "Millennium Ecosystem Assessment, Ecosystems and Human Well-Being: Policy Responses, Findings of the Responses Working Group," 2005.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw, "The Impact of Global Warming on Agriculture: A Ricardian Analysis," *The American Economic Review*, 1994, *84* (4), 753–771.

Moya, P., *Increasing Productivity of Intensive Rice Systems Through Site-Specific Nutrient Management*, Science Publishers and IRRI, 2004.

Nath, Pradosh and Bhagirath Behera, "A critical review of impact of and adaptation to climate change in developed and developing economies," *Environment, Development and Sustainability*, 2011, *13.*

Panayotou, Theodore, "Demystifying the environmental Kuznets curve: turning a black box into a policy tool," *Environment and Development Economics*, 1997, *2* (04), 465.

Peng, Shaobing, Jianliang Huang, John E. Sheehy, Rebecca C. Laza, Romeo M. Visperas, Xuhua Zhong, Grace S. Centeno, Gurdev S. Khush, and Kenneth G. Cassman, "Rice yields decline with higher night temperature from global warming," *Proceedings of the National Academy of Sciences*, 2004, *101* (27), 9971–9975.

Pfaff, Alexander, *From Deforestation to Reforestation in New England, USA* Global Prospects of Deforestation and Forest Transition, Helsinki: Kluwer, 2000.

Pfaff, Alexander S. P., "What Drives Deforestation in the Brazilian Amazon?: Evidence from Satellite and Socioeconomic Data," *Journal of Environmental Economics and Management*, 1 1999, *37* (1), 26–43.

Pfaff, Alexander, Suzi Kerr, Romina Cavatassi, Benjamin Davis, Leslie Lipper, Arturo Sanchez, and Jason Timmins, *Effects of Poverty on Deforestation*, Vol. 25 of *Economics of Poverty, Environment and Natural Resource Use*, Netherlands: Springer,

Pinkse, J. and M. Slade, "Contracting in space: an application of spatial statistics to discrete choice models," *Journal of Econometrics*, 1998, *85*, 125–154.

Ramanathan, V., C. Chung, D. Kim, T. Bettge, L. Buja, J. T. Kiehl, W. M. Washington, Q. Fu, D. R. Sikka, and M. Wild, "Atmospheric brown clouds: Impacts on South Asian climate and hydrological cycle," *Proceedings of the National Academy of Sciences*, 2005, *102* (15), 5326–5333.

Schlenker, Wolfram and Michael J. Roberts, "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change," *Proceedings of the National Academy of Sciences*, 2009, *106* (37).

_ , W. Michael Hanemann, and Anthony C. Fisher, "Will U.S. Agriculture Really Benefit from Global Warming? Accounting for Irrigation in the Hedonic Approach," *The American Economic Review*, 2005, *95* (1), pp. 395–406.

Schultz, T. Paul, "School subsidies for the poor: evaluating the Mexican Progresa poverty program," *Journal of Development Economics*, 6 2004, *74* (1), 199–250.

Seo, S. Niggol, "Is an integrated farm more resilient against climate change? A micro-econometric analysis of portfolio diversification in African agriculture," *Food Policy*, 2010, *35* (1).

_ and Robert Mendelsohn, "An analysis of crop choice: Adapting to climate change in South American farms," *Ecological Economics*, 2008, *67* (1).

Seshu, D. V. and F. B. Cady, "Response of Rice to Solar Radiation and Temperature Estimated from International Yield Trials," *Crop Science*, July 1 1984, *24* (4), 649–654.

Sheehy, John E., P. L. Mitchell, and Anaida B. Ferrer, "Decline in rice grain yields with temperature: Models and correlations can give different estimates," *Field Crops Research*, 0 2006, *98* (2-3), 151–156.

Shortle, J.S. and D.G. Abler, *Agriculture and the Environment* Handbook of Environmental and Resource Economics, Cheltenham, UK: Edward Elgar,

Skoufias, Emmanual and Bonnie McClafferty, "Is Progresa Working? A Summary of the Results of an Evaluation by IFPRI," Technical Report FCND Discussion Paper No. 118, IFPRI 2001.

_ , Benjamin Davis, and Jere R. Behrman, "An Evaluation of the Selection of Beneficiary Households in the Education, Health, and Nutrition Program (PROGRESA) of Mexico," Technical Report, International Food Policy Research Institute 1999.

Skoufias, Emmanuel, "PROGRESA and its impacts on the welfare of rural households in Mexico," Research Report 139, International Food Policy Research Institute (IFPRI) 2005.

Smit, Barry and Mark W. Skinner, "Adaptation options in agriculture to climate change: a typology," *Mitigation and Adaptation Strategies for Global Change*, 2002, *7*.

Stanhill, Gerald and Shabtai Cohen, "Global dimming: a review of the evidence for a widespread and significant reduction in global radiation with discussion of its probable causes and possible agricultural consequences," *Agricultural and Forest Meteorology*, 2001, *107* (4), 255 – 278.

Stern, David I., "The Rise and Fall of the Environmental Kuznets Curve," *World Development*, 8 2004, *32* (8), 1419–1439.

Stern, Nicholas, "The Economics of Climate Change," *The American Economic Review*, 2008, *98*, 1–37(37).

Wackernagel, Mathis and William Rees, "Our Ecological Footprint," *Green Teacher*, 00/00 1996, *45* (1996), 5–14.

Wang, Jinxia, Robert O. Mendelsohn, Ariel Dinar, and Jikun Huang, "How China's Farmers Adapt to Climate Change," *SSRN eLibrary*, 2008.

Wassmann, R., S.V.K. Jagadish, K. Sumfleth, H. Pathak, G. Howell, A. Ismail, R. Serraj, E. Redona, R.K. Singh, and S. Heuer, "Chapter 3 Regional Vulnerability of Climate Change Impacts on Asian Rice Production and Scope for Adaptation," in Donald L. Sparks, ed., *Donald L. Sparks, ed.*, Vol. 102 of *Advances in Agronomy* 2009.

— , — , S. Heuer, A. Ismail, E. Redona, R. Serraj, R.K. Singh, G. Howell, H. Pathak, and K. Sumfleth, "Chapter 2 Climate Change Affecting Rice Production: The Physiological and Agronomic Basis for Possible Adaptation Strategies," in Donald L. Sparks, ed., *Donald L. Sparks, ed.*, Vol. 101 of *Advances in Agronomy* 2009.

Welch, Jarrod R., Jeffrey R. Vincent, Maximilian Auffhammer, Piedad F. Moya, Achim Dobermann, and David Dawe, "Rice yields in tropical/subtropical Asia exhibit large but opposing sensitivities to minimum and maximum temperatures," *Proceedings of the National Academy of Sciences*, 2010, *107* (33).

White, Thomas, "Diet and the distribution of environmental impact," *Ecological Economics*, 2000, *34* (1), 145 – 153.

World Bank, "1992 World Development Report," Technical Report, World Bank 1992.

— , "Poverty and the Environment: Understanding Linkages at the Household Level," 2008.

Wu, Wenbin, Ryosuke Shibasaki, Peng Yang, Huajun Tang, and Kenji Sugimoto, "Modeling changes in paddy rice sown areas in Asia," *Sustainability Science*, 2010, *5.*

Wunder, Sven, "Poverty Alleviation and Tropical Forestsï£¡What Scope for Synergies?," *World Development*, 11 2001, *29* (11), 1817–1833.

Yoshida, S. and FT Parao, *Climatic influence on yield and yield components of lowland rice in the tropics*, International Rice Research Institute, 1976.

— , T. Satake, and D. Mackill, "High temperature stress in rice," *IRRI Research Paper*, 1981, (67).

Zhou, Liming, Robert E. Dickinson, Yuhong Tian, Jingyun Fang, Qingxiang Li, Robert K. Kaufmann, Compton J. Tucker, and Ranga B. Myneni, "Evidence for a significant urbanization effect on climate in China," *Proceedings of the National Academy of Sciences*, 2004, *101* (26), 9540–9544.

Zwane, Alix Peterson, "Does poverty constrain deforestation? Econometric evidence from Peru," *Journal of Development Economics*, 9 2007, *84* (1), 330–349.