THREE ESSAYS IN DEVELOPMENT ECONOMICS

By

Ming Fang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food, and Resource Economics – Doctor of Philosophy
Economics – Dual Major

2023

ABSTRACT

Environmental and technological changes such as climate change and technology development greatly affect rural farmers' decisions in agricultural production, consumption, and employment. As one of the most important issues in agricultural and development economics, understanding how farmers respond to these changes in different institutional contexts has important policy implications. This dissertation addresses three different types of changes—climate change, income shocks, and road construction in three developing countries—Ukraine, India, and China accordingly and investigates rural farmers' behavioral and welfare responses to these changes in agricultural production, consumption, and employment.

The first essay, titled "*Climate and Agricultural Production in Ukraine*", examines the impacts of climate change on agricultural production in Ukraine. Using the combination of large-scale survey data from 46,799 farms and daily temperature and precipitation data from CORDEX, the essay examines the short-term yield response of five main crops in Ukraine—winter wheat, spring barley, sunflower, soybean, and corn to yearly temperature and precipitation changes, and the impacts of climate change on agricultural adaptation. The empirical results first suggest that the overall rising temperature is associated with short-run rises in the yields of all five main crops—winter wheat, spring barley, sunflower, soybean, and corn in Ukraine. While for cold season crops—winter wheat and spring barley, crop yields respond negatively to exposure to heat accumulation above a crop-specific stressful temperature bound (29/30°C). In terms of adaptation, the study unveils the adaptations of winter wheat growers to long-run temperature changes and adaptations of sunflower, soybean, and corn growers to long-run precipitation changes. Moreover, farms are also found with more diversified crop structures in response to negative climate shocks.

The second essay, titled "*Social Network and Consumption Smoothing in Rural India*", explores the impact of the caste-based social network on consumption smoothing behavior of rural households in India. The results first suggest that full insurance or efficient risk-sharing, i.e., consumption is not associated with idiosyncratic income changes when common shocks are

controlled, is achieved within the three backward caste groups but not within the forward caste groups or at the village level. The relatively high expenditure on necessary food items of the backward caste groups is likely to be a potential mechanism that explains the heterogeneity in consumption smoothing across caste groups. Furthermore, the study examines the impact of social interaction norms across castes on consumption smoothing. The result suggests that the consumption of those households who only interact with other households within the same caste or caste groups comoves less with idiosyncratic income changes than those who also interact with other caste or caste groups, which implies the potential positive impact of intra-caste networks on consumption smoothing in India. Moreover, the essay also explores the relationship between migration work and consumption smoothing and finds that households with migrant working experience do better in consumption smoothing.

The third essay, titled "*Road Infrastructure and Rural Employment in China*", examines the impacts of various road infrastructures on farmers' employment choices in rural China, based on farm-level panel data from three rounds of household surveys spanning from 2000 to 2013 and the road infrastructure data from GIS data. The empirical results first indicate that the road infrastructure at the county, provincial/national levels are positively associated with farmers' tendency toward and intensity of local non-agricultural employment while is negatively associated with farmers' tendency toward and intensity of migration work (mainly driven by the impacts on part-time employment). Moreover, denser local roads—village and township roads are found associated with more local part-time non-agricultural employment but less non-local part-time employment. Furthermore, observable heterogeneities in the impacts of road infrastructure on rural employment across gender and region are also discovered in the study.

I would like to dedicate this dissertation to the people I love.

TABLE OF CONTENTS

CHAPTER 1: CLIMATE AND AGRICULTURAL PRODUCTION IN UKRAINE

1.1 Introduction

The impacts of climate change, which refers to the relatively long-term pattern of weather, on land productivity have been felt throughout the world in recent decades. Such changes are reflected not only in the long-term trend of temperature or precipitation but also in their short-term volatilities and the correspondingly more frequent and severe extreme climate events such as heatwaves, droughts, floods, tsunamis, etc. that can affect crop yield dramatically. According to the most recent climate assessment report issued by IPCC, climate change has slowed the growth of the overall increasing agricultural productivity over the past 50 years globally. While some crops in high-latitude regions such as Europe (Raza et al. 2019) and North America (IPCC 2022) might benefit from the global warming trend, climate change in past decades is found associated with crop yields reduction in most mid- and low-latitude regions such as Africa (Kogo et al. 2021, Sultan et al. 2019, Nhamo et al. 2019) and Australasia (Molotoks et al. 2021).

The two components of weather or climate—temperature and precipitation can affect land productivity through various mechanisms such as aerial and soil humidity, nutrient, and growth of other life forms such as weeds, insects, bacteria, etc. Consequently, the exact impacts of climate change on crop yields differ among crop species and can vary with environmental endowments, such as latitude, elevation, soil condition, water efficiency, etc. For example, the effects of rising carbon dioxide concentration, a corresponding phenomenon of global warming, on yield are stronger for plants with the C_3 photosynthetic pathways such as wheat, rice, and soybean than for C_4 plants such as maize, millet, sorghum, and many types of grass due to the difference in the efficiency of carbon dioxide and water use[1]. The other example is the non-identical impacts of

[1] Plants are classified as C_3, C_4, or CAM according to the products formed in the initial phases of photosynthesis. C_3 species respond more to increased CO_2; C_4 species respond better than C_3 plants to higher temperature, and their water-use efficiency increases more than for C_3 plants (Kurukulasuriya and Rosenthal 2013).

floods on crop yields. While floods would normally cause substantial damage to most crops (Mirza, Warrick, and Ericksen 2003*)*, normal floods could support the growth of nitrogen-supplying blue-green algae that fertilizes rice crops, and large floods to higher land may also provide extra moisture and benefit rabi crops such as vegetables and onion, etc. (Brammer 1990). Using data from Taiwan, China, Chang (2002) also finds that warmer temperature is associated with decreasing yields of crops such as rice, soybeans, adzuki beans, sugarcane, banana, pineapple, apple, etc., while is positively related to the yield of tomato. Moreover, the difference in environmental endowments, the initial temperature for instance, also contributes to the variation in the yield effects of climate change. In general, higher temperatures in middle- and high-latitude areas will expand crop-producing areas poleward and lengthen growing seasons, thus benefiting crops in these regions (Rosenzweig and Hillel 1998). While in hot low-latitude tropical areas, the rising temperature would harm crop yields in many lowland areas, particularly in those semiarid regions, but benefit crop yields in highland areas (Fischer and Velthuizen 1996; Downing 1992).

Beyond the crop attributes and environmental endowments, another non-negligible factor that contributes to the variation in crops' yield response to climate change is the adaptive capacity of the agricultural sector. Farmers can first adjust the planting structures, crop cultivars, and the timing of planting to mitigate the impact of climate change. For example, diversifications in seed genetic structure and composition (Mortimore et al. 2000, Fosu-Mensah et al. 2012), the use of intercropping (Hassen et al. 2017), adoptions of drought-tolerance varieties (Tambo and Abdoulaye 2012; Lunduka et al. 2019; Dejene et al. 2011), changes in cropping calendar (Rhodes et al. 2014), etc. are found as strategies against climate change in practice. Adaptations could also happen in other agriculture operations such as weed management (Scott et al. 2014), pest control (Chen and McCarl 2001), nutrient management (Chiotti and Johnson 1995), irrigation (de Loë et al. 1999), etc. Moreover, adaptations at the government or society level such as weather prediction, infrastructure investment, R&D investment in technologies, etc. can also affect the impacts of climate change on crops. While in practice, adaptations to climate change are sometimes constrained by factors such as information availability, knowledge level, economic and social considerations, technology and capital limitation, institutional capacity, etc.

As one of the largest crop exporters in the world, Ukraine plays an important role in global food supply chain, which has recently been highlighted by the effects of the ongoing Russia-Ukraine war on global food supply and price, and the food security of countries that import most of their grain supply from Ukraine (Behnassi and Haiba 2022). With 41.5 million hectares of agricultural land covering 70% of the country and about 25% of black soil (a highly fertile soil with high organic matter contents) reserves on the planet, Ukraine's agriculture sector generated approximately 9.3% of the nation's GDP and 14.11% of employment in 2020. Ukraine is also one of the main exporters of agricultural products in the world, accounting for 47% of sunflower oil, 10% of wheat, 14% of corn, and 17% of barley in global export markets in 2021[2].

Given the importance of the agriculture sector in Ukraine and its vulnerability to climate change, understanding the exact impacts of climate change on agriculture production in Ukraine has important policy implications. However, little is known about the impacts of changing climate on crop yields and agricultural production in Ukraine. And the validities of the small number of studies that examined climate impacts on crop yields in Ukraine are constrained by either the aggregate crop yields data they use (Tarariko et al. 2017) or the relatively short time interval and crop coverage of the data (Müller et al. 2016).

In the past century, extreme weather and climatic shocks have brought challenges to agricultural production in Ukraine. The annual air temperature in Ukraine in the 20th century has increased by 0.6±0.2°C (Boychenko et al. 2016), which corresponds to more frequent arid years and the changing rhythm of seasonal phenomena such as floods and snowfall (Rudych and Batazhok 2018). Since crop growth responses to changes in weather may vary considerably across crops and regions, it is important to develop location-specific estimates of the weather-yield

[2] Data source: USDA https://www.fas.usda.gov/data/grain-world-markets-and-trade

response. The negative relationship between warming and crop yields found in Brazil (Assunção and Chein 2016), France (Gammans et al. 2017), and the United States (Schlenker and Roberts, 2009) may not extend to the cooler climate of Ukraine. Under certain conditions, warming may result in more arable land and increasing crop yields (IPCC 2022). In the context of Ukraine, Taraiko et al. (2017) analyzed the effect of historical and current climate change on grain yields based on historical satellite data and climatic modeling approach and predicted that the grain yields in Ukraine may increase by 25% and 29-30% in 2025 and 2050 accordingly compared with yields in 2015. Müller et al. (2016) also predicted that while the wheat yields in Ukraine may substantially decrease under a high emission scenario (particularly in the southern Steppe zone), rising temperatures and increasing precipitation may increase wheat yields in less fertile areas in northern Ukraine.

To investigate the contextualized impacts of changing climate conditions (especially the temperature) on different crops, we combine large-sample survey data in Ukraine (46,799 farms each year) and climate data from 2004 to 2020 to examine how the yields of five main crops in Ukraine respond to short-run variation in temperature and precipitation in the growing season. Based on these results, we examine whether and how the agricultural sector in Ukraine adapts to long-term climate change. Our results first suggest that the overall rising temperature is associated with short-run yield rises of all five main crops, which contradicts findings in other related studies which conclude that the overall rising temperature harms the crop yields in other European contexts such as France (Gammans et al. 2017). However, the yields of two cold-season crops—winter wheat and spring barley, respond negatively to heat accumulation in temperature intervals above a crop-specific stressful temperature bound (29/30°C). For warm-season crops on the other hand—sunflower, soybean, and corn, we do not find evidence of a stressful temperature threshold. In terms of adaptation, our empirical result unveils the adaptations of winter wheat growers to long-run temperature change and adaptations of sunflower, soybean, and corn growers to long-run precipitation change. Furthermore, we also find that farmers in Ukraine diversify their crops in response to perceived negative climate shocks such as rising temperatures, reducing precipitation, and increasing volatilities of temperature and precipitation in growing seasons.

By examining the heterogeneous relationships between changing climate conditions and land productivity across crops and the responses of the agricultural sector in Ukraine to long-term climate change, our study contributes to the literature by providing a comprehensive analysis that encompasses multiple crop types in a high-latitude context. Our finding that cold-season crops and warm-season crops respond differently to the rising temperature in Ukraine contributes to the broader understanding of the climate impacts on agricultural productivity. With the pressing challenges in agriculture posed by the changing climate, our study has important implications for food security and policymakers and farmers involved in agricultural production.

The rest of the paper is structured as follows. The following section introduces the data and some important descriptive results on climate change and agriculture production in Ukraine. Section 3 introduces the empirical methods we apply for our analysis and the empirical results are shown in section 4. Section 5 concludes the paper.

1.2 Data and Descriptive Results

A. Data

We combine the farm-level survey data collected by Ukraine's State Statistics Department and rayon-level climate data based on data from the WCRP Coordinated Regional Downscaling Experiment program (CORDEX, Giorgi et al. 2009) for our analysis.

The farm-level data collects information on planted areas of crops and outputs from farms in Ukraine, which results in an average sample size of 46,799 farms each year from 2004 to 2020 (except 2015). Compared to macro-level data from other sources, the aggregated cereal and wheat production in our dataset accounts for approximately 84% of the national cereal production (59.79 out of 70.77 million tons, FAO[3]) and 76% of the national wheat production (21.58 out of 28.38

[3] Data source: https://www.fao.org/faostat/en/#country/230

million tons, USDA[4]) accordingly in 2020. We limit our analysis to five major crops in Ukraine—winter wheat, spring barley, sunflower, soybean, and corn. Together, these crops account for approximately 90% of the total cereal and legume production (winter wheat, spring barley, corn, and soybean) and 65% of the total technical crop production (sunflower).

While the survey data we use is advantaged in its large sample size, the survey conducted before and after 2015 could only be matched at the aggregated level (village or rayon) but not at the farm level due to the inconsistency in farm ID in two periods. As a result, our empirical analysis is mainly based on aggregated data at the village level to utilize recent data after 2014. The crop yields at the village level are calculated based on the aggregated output and planting area. Averagely, there are about 13,000 villages each year in our sample.

The temperature and precipitation data we use came from EURO-CORDEX (Jacob et al. 2014), a sub-program of the Coordinated Regional Downscaling Experiment Program (CORDEX) sponsored by the World Climate Research Program (WCRP). EURO-CORDEX provides a set of simulations on climate data in Europe with a horizontal resolution of 12.5 km. Based on the EURO-CORDEX data, we calculate regional climate data including the daily precipitation and daily maximum, average, and minimum temperature for each rayon (sub-district) in Ukraine. The temperature and precipitation data are then processed and merged with the survey data based on rayon information.

B. Climate Change in Ukraine

Ukraine is the second-largest European country located in Eastern Europe. Most of the country, except the southern coast of Crimea with a subtropical Mediterranean climate, has a temperate climate with sufficient sunshine and year-round rainfall. As shown in Figure 1.1, August is the hottest month in Ukraine with an average temperature of 20.01°C, and February is the coldest month with an average temperature of 3.76°C.

[4] Data source: https://www.fas.usda.gov/data/ukraine-grain-and-feed-annual-5

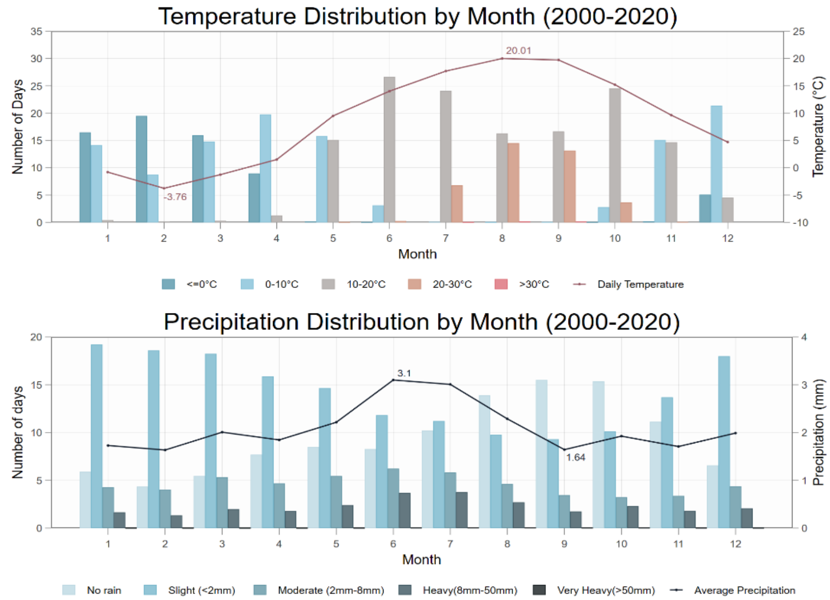

Figure 1.1 Monthly distribution of temperature and precipitation in Ukraine

Precipitation is quite sufficient in Ukraine and June has the most sufficient rainfall of the year (3.1 mm each day). To describe the seasonal precipitation distribution, we categorize the daily rainfall into 5 groups based on the rainfall intensity and count the average number of rainy days in each month from 2000 to 2020. The rainfall intensity of each day is categorized into 5 groups based on the rate of precipitation per hour: No rain (0 mm), Slight (0-2.5 mm), Moderate (2.5-8 mm), Heavy (8-50 mm), and Very Heavy (>50 mm). Figure 1.1 suggests that even in the month with the least rainfall—September, half of the days has some amount of rainfall. Given the sufficiency of rainfall in the context of Ukraine, our empirical analysis focuses more on investigating the impacts of temperature change on agricultural production.

Figure 1.2 visualizes the regional distribution of the 10-year mean daily maximum, average, and minimum temperature in the hottest month (August) and coldest month (February) in 2000, 2010, and 2020. It first presents the geographical distribution of temperature in Ukraine, indicating that the temperature increases from northwest to southeast. Moreover, the temperature map also suggests an observable warming trend in the past two decades, especially in winter—February. On average, the 10-year mean daily temperature in August and February increased by about 0.68°C and 3.18°C accordingly from 2000 to 2020.

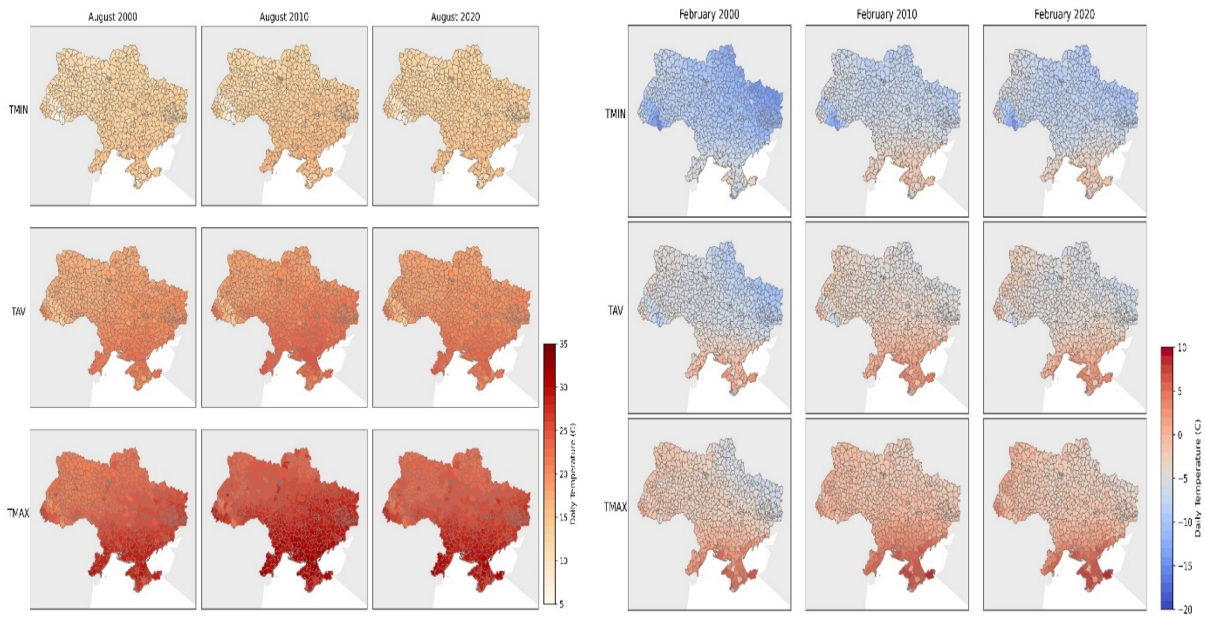

Figure 1.2 Average temperature in August and February of 2000, 2010, and 2020 in Ukraine

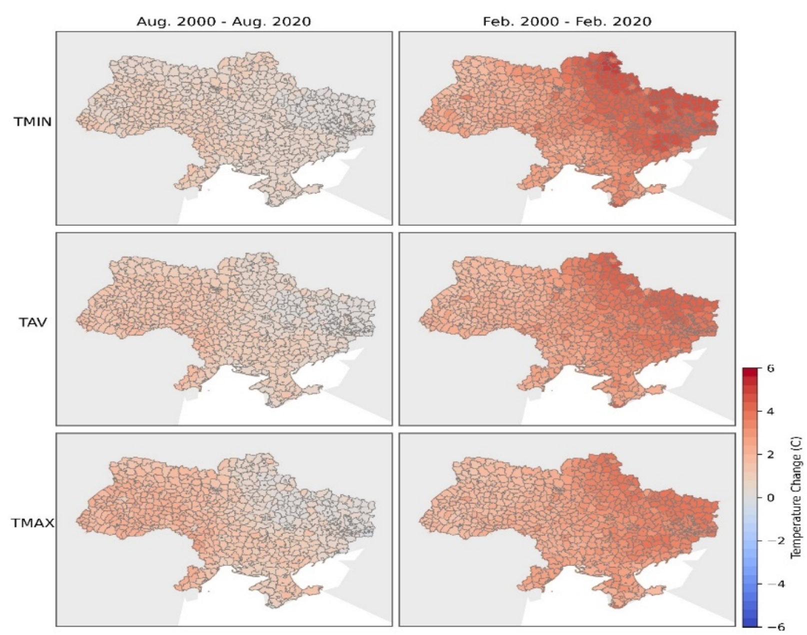

Figure 1.3 Geographical distribution of temperature changes from 2000 to 2020 in August and February

To illustrate the warming trend more explicitly, we calculate the changes in 10-year mean temperature in August and February between 2000 and 2020 of each rayon. As shown in Figure 1.3, the 10-year mean temperature in both August and February has increased in almost all rayons from 2000 to 2020. Moreover, the temperature in northeastern Ukraine increases more in winter (February) than in summer (August) while temperature change in other regions is more synchronous, which implies the geographical difference in climate change patterns.

C. Crop production in Ukraine

As an important industry in Ukraine, agriculture contributes to approximately 9.3% of the nation's GDP and 14.11% of employment in 2020[5]. Ukraine has suitable climate conditions for crop growth and about half of the country is covered with black soil (the greatest reserve on the planet), which is highly fertile with high organic matter contents. Ukraine is also one of the main exporters of agricultural products in the world, accounting for 47% of sunflower oil, 10% of wheat, 14% of corn, and 17% of barley in global export markets in 2021[6].

The total sown area in Ukraine, based on our data, has been falling since 2005 from 19 million hectares to 15 million hectares in 2014 and then starts rising back to 20 million hectares in 2020 (Figure 1.4). However, it needs to be noted that the 2014 Ukraine crisis might have partly contributed to the extremely low level of sown area in 2014 as the agriculture production and statistical operations in some areas are affected by the conflict. Among all crops, cereals and legumes contribute the largest share of the total cropping area while the share of technical crops has been expanding rapidly over time. The cereals and legumes, while slightly decreasing over time, account for more than half of the total cropping area in the country. In the meanwhile, the share of technical crops has doubled from approximately 20% of the total cropping area to 40%

5 Data source: World Bank https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS?locations=UA

6 Data source: USDA https://www.fas.usda.gov/data/grain-world-markets-and-trade

from 2004 to 2020. Correspondingly, the share of feed crops and other crops declined dramatically in the same period.

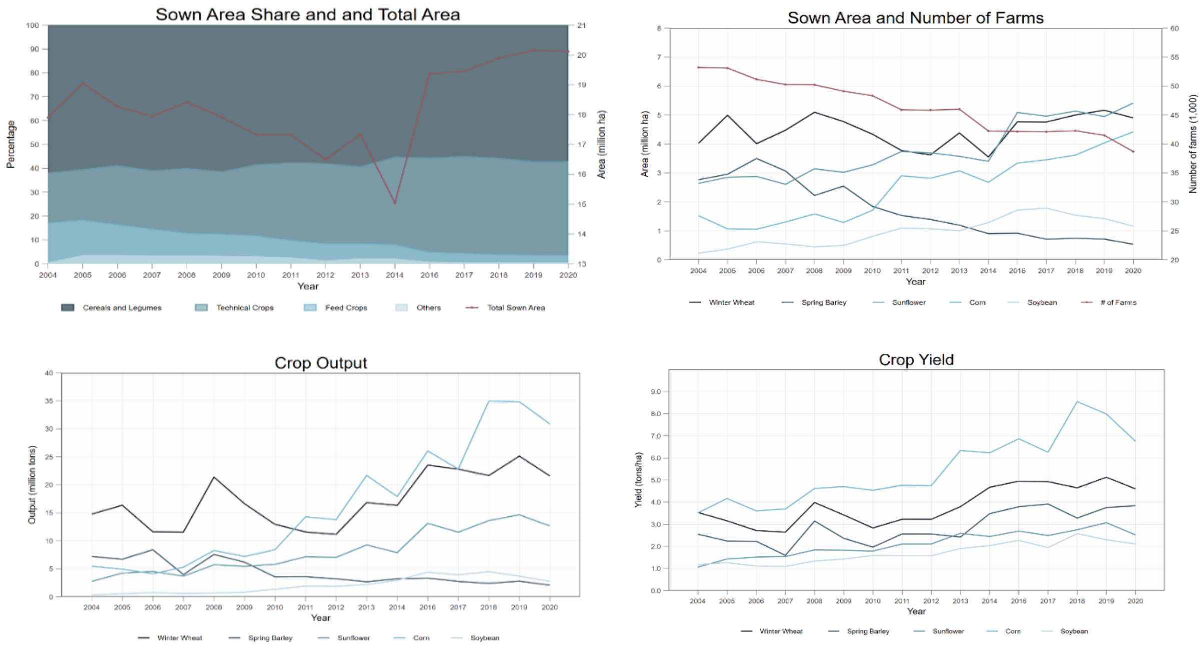

Figure 1.4 Changes in production of five main crops in Ukraine

The five main crops we focus on in our study—winter wheat, spring barley, corn, soybean, and sunflower account for approximately 90% of the total cereal and legumes production (winter wheat, spring barley, corn, and soybean) and 65% of the total technical crops production (sunflower). The production of corn, soybean, and sunflower has been continuously expanding since 2004, while the sown area of soybean starts declining in recent years after 2017. In contrast, the cropping area of spring barley shrank from 2.76 million hectares in 2004 to 0.54 million hectares in 2020. The cropping area of winter wheat, in the same period, has increased overall while varying dramatically over time.

The total outputs of all five crops, except the spring barley as the result of the rapidly shrinking cropping area, have been increasing over time (Figure 1.4). In addition to the expanding sown area of these crops, the adoption of new technologies or new varieties, fertilizer use, and better farm operation have also contributed to the increasing yields of these crops. In 2020, the average yields of winter wheat, spring barley, sunflower, soybean, and corn have reached 4.6, 3.8, 2.5, 2.1, and 6.7 tons/ha accordingly.

1.3 Empirical Methods

A. Yield response to temperature and precipitation

Our analysis focuses on 5 main crops in Ukraine—winter wheat, spring barley, sunflower, soybean, and corn. The growing seasons of the five crops are defined as March-June, April-June, June-August, June-July, and June-September accordingly based on the agronomic calendar in Ukraine (*USDA*[7]), and all climate-related variables are constructed based on crop-specific growing seasons.

To estimate the weather-yield relationship for each of five crops of our interest, we first adopt the fixed effects panel regression and estimate the baseline model:

$$y_{it} = \beta Temp_{it} + \gamma_1 Prec_{it} + \gamma_2 {Prec_{it}}^2 + c_i + \lambda_t + \epsilon_{it} \tag{1}$$

where y_{it} is the logarithm of crop yield at village i in year t. Weather variables include $Temp_{it}$, average daily temperature, along with average daily precipitation, $Prec_{it}$ and its square, which addresses the non-linear relationship between precipitation and yields as suggested in the literature. Location-fixed effects, c_i, absorb time-invariant factors at the village level that might affect crop yields such as geographic conditions or soil type, λ_t denotes the year-fixed effects that control the common factors that could contribute to variation in crop yields over time such as technology availability, and ϵ_{it} is the error term. All climate variables in our analysis are calculated based on crop-specific growing seasons.

To address the nonlinearity in yield response to temperature, we then follow the method that has been widely used in literature by assuming the crop yield is proportional to the total exposure

[7] Data source: https://ipad.fas.usda.gov/countrysummary/default.aspx?id=UP

of heat, and temperature effects on yields are cumulative over time (Schlenker and Roberts 2009; Gong and Chen 2020). The model can be written as:

$$y_{it} = \int_{\underline{h}}^{\overline{h}} g(h)\phi_{it}(h)\ dh + \gamma_1 Prec_{it} + \gamma_2 {Prec_{it}}^2 + c_i + \lambda_t + \epsilon_{it} \tag{2}$$

where y_{it} is the logarithm of crop yield at village i in year t, $g(h)$ is the temperature response function and $\phi_{it}(h)$ is the time distribution of heat over the growing season at village i in year t. $\underline{h}$ and $\overline{h}$ indicate the lower and upper bounds of temperature in the growing season. Control variables $Prec_{it}$, c_i, λ_t, and ϵ_{it} are the same as in equation (1).

We estimate the temperature response function $g(h)$ in equation (2) under two specifications—the temperature exposure step-function approach and the growing degree days (GDDs) approach. For the step-function approach, we define a function that implies a distinct marginal effect within each 3 °C temperature exposure interval. More formally, we can replace the term $\int_{\underline{h}}^{\overline{h}} g(h)\phi_{it}(h)\ dh$ in equation (2) with this step-function to obtain the following regression equation:

$$y_{it} = \sum_{j=\underline{h},\ \underline{h}+3,\ \underline{h}+6,\dots}^{\overline{h}} \rho_j [\Phi_{it}(j+3) - \Phi_{it}(j)] + \gamma_1 Prec_{it} + \gamma_2 {Prec_{it}}^2 + c_i + \lambda_t + \epsilon_{it} \tag{3}$$

where $\Phi_{it}(j+3) - \Phi_{it}(j)$ is the exposure time in the growing season when the temperature is between j °C and $j+3$ °C. Other variables are represented the same as in equation (1). Moreover, it needs to be noted that the 3-degree bins are chosen for simplicity concerns. We also conduct the analysis using 1-degree bins for robustness check and the results are similar to 3-degree specifications.

To calculate the exposure to different 3-degree temperature intervals, we apply a method that has been widely adopted in the literature—the sine curve approach (Allen 1976). More specifically, the within-day temperature distribution of each day is simulated by a sine curve based on daily maximum and minimum temperature, after which the daily exposure to each 3-degree temperature interval is calculated and then aggregated across the growing season.

An alternative specification we apply to capture the nonlinear effect of temperature is based on the growing degree days (GDDs hereafter) approach. Rooted in agronomic research, GDDs are widely used to measure the accumulated heat unit in the growing season, which assumes that crop development will only occur when the temperature exceeds a minimum or lower temperature threshold l_0 while the heat accumulation after the temperature exceeds a maximum or upper-temperature threshold l_1 would either have no effects on crop growth or even harm the crop. Empirically, the accumulated heat unit within the temperature interval $[l_0, l_1]$ is defined as GDDs and the accumulated heat unit above l_1 measures the stress degree days (SDDs hereafter). We apply the GDDs specification by estimating the following equation:

$$y_{it} = \beta_1 GDDs_{it}^{l_0,l_1} + \beta_2 SDDs_{it}^{l_1} + \gamma_1 Prec_{it} + \gamma_2 {Prec_{it}}^2 + c_i + \lambda_t + \epsilon_{it} \tag{4}$$

where $GDDs_{it}^{l_0,l_1}$ is the GDDs in the growing season when the lower and upper-temperature thresholds are set as l_0 and l_1 accordingly, $SDDs_{it}^{l_1}$ denotes SDDs measured by the heat accumulation when the temperature is above the upper-temperature threshold l_1, $Prec_{it}$, c_i, λ_t, and ϵ_{it} are the same as in previous models.

Similar to our temperature exposure variables, the GDDs and SDDs of each day are calculated based on the simulated daily temperature distribution using the sine curve approach, which is the integral area between the given temperature threshold $[l_0, l_1]$ on the simulated daily temperature distribution curve, after which the daily data are aggregated to calculate the GDDs and SDDs for the growing season.

To determine the temperature thresholds for GDDs, we select the lower temperature threshold l_0 based on the agronomic knowledge for each crop[8] and loop for possible upper-temperature threshold l_1 to determine the appropriate thresholds based on changes in estimated coefficients of GDDs/SDDs and R^2 in the regression. The estimated results from the former temperature bins approach are also used to determine the range of upper-temperature thresholds we loop across. More specifically, we choose temperatures around the temperature bins of which the estimated coefficient turns from positive to negative as it unveils the approximate "stressful" temperature for crop growth, above which the rising temperature would negatively affect crop yields.

B. Adaptation to climate change

In literature, the agricultural adaptation to climate change is normally measured by farmers' behavioral responses such as the adoption of new varieties or technologies, the investment in the irrigation system, etc. However, our survey data doesn't contain information on farmers' behavioral changes except for the crop structure. Consequently, we first investigate the adaptation to climate change implicitly based on the yield response of crops to long-term and short-term climate conditions and then examine the explicit adaptation in crop diversification.

To measure the adaptation implicitly, we apply the approach inspired by the model specified by Mérel and Gammans (2021), which provides a new insight into the estimation of short- and long-run yield responses to unveil the long-run adaptation to climate change. The long-run climatic adaptation model can be specified as:

$$y_{it} = \beta_1 x_{it} + \beta_2 x_{it}^2 + \beta_3 (x_{it} - \overline{x}_i)^2 + c_i + \lambda_t + \mu_{it} \tag{5}$$

where y_{it} is the log of crop yield at location i in year t, x_{it} is the weather realization (temperature or precipitation) in year t, the climate variable $\overline{x}_i = \frac{1}{T}\sum_t x_{it}$ is the sample average of x_{it}.

[8] North Dakota Agricultural Weather Network (NDAWN): https://ndawn.ndsu.nodak.edu/help-wheat-growing-degree-days.html

The term $(x_{it} - \overline{x}_i)^2$ in equation (5) reflects the (squared) distance between contemporaneous weather and a location's normal climate, of which the coefficient β_3 is expected to be negative if long-run adaptations to climate change exist as it reflects that, conditional on the contemporaneous weather realization, crop yields in locations with an underlying climate closer to that realization are higher (due to the implicit adaptation to climate) than in locations for which that realization happens to be unusual.

To measure the explicit adaptation to climate change, we apply the two-limit random effect model to estimate farmers' response in crop diversification, measured by the Herfindal index (Bradshaw et al. 2014), to lagged climate-related variables such as the long-run means of daily temperature/precipitation and the lagged standard deviation of daily temperature/precipitation in the growing season. The model is defined as:

$$y_{it}^* = \beta^c c_{it-1} + \beta^x x_{it} + \lambda_t + c_i + \mu_{it} \tag{6}$$
$$y_{it} = 0 \quad if \quad y_{it}^* \leq 0$$
$$y_{it} = y_{it}^* \quad if \quad 0 \leq y_{it}^* \geq 1$$
$$y_{it} = 1 \quad if \quad y_{it}^* \geq 1$$

where y_{it} is the Herfindal Index of location i in period t, y_{it}^* is the underlying latent variable of y_{it}. c_{it-1} denotes the vector of lagged climate variables, x_{it} is the total sown area at location i in period t. c_i is the time-invariant location-fixed effect to control for time-invariant heterogeneity and λ_t is the year fixed effects.

The Herfindal Index (Bradshaw et al. 2004), which ranges from zero (i.e., an infinite number of crops in equal proportions) to one (i.e., single crop), is widely used in literature to measure crop diversity. It is calculated as

$$H = \sum_{n=1}^{N} P_n^2 \tag{7}$$

where P_n is the land share of crop n. The Herfindal index in our analysis is calculated as the sum of the squared land share on 9 categories of crops—winter wheat, winter barley and spring barley, winter rye, sunflower, soybean, corn, sugar beets, and other crops.

1.4 Empirical Results

This section presents our empirical results of the estimated responses in crop yields and adaptations to climate. We start with estimating the short-run response of crop yields to some simple measures of climate conditions, such as average daily temperature and precipitation in growing seasons. To address the potential nonlinearity in the temperature-yield relationship, we then estimate more precise nonlinear yield responses of the crops to exposure to different temperature intervals and heat accumulation (GDDs and SDDs). Furthermore, we investigate the implicit agricultural adaptations from crops' yield response to the changing climate conditions and the farmers' explicit adaptation in crop diversification.

Our analysis focuses on 5 main crops in Ukraine—winter wheat, spring barley, sunflower, soybean, and corn. Since the growing seasons of these crops are not the same, variables related to the growing season such as temperature bins and GDDs are crop specific.

A. Yield responses to short-run weather variation

We first present the results of our baseline model described in equation (1). The climate-related independent variables include the average daily temperature, average daily precipitation, and its square term to address the quadratic relationship between crop yields and precipitation.

Table 1.1 reports the estimated yield responses of crops to weather variations based on equation (1), which suggests that the yields of the five main crops in Ukraine respond positively to the overall rising temperature. Our results indicate that a 1 °C rise in average daily temperature in the growing season is associated with 2.8%, 5%, 3.5%, 8.8%, and 3.5% increases in the yields of winter wheat, sunflower, soybean, and corn respectively. We conclude that, unlike most low-latitude areas, the overall rising temperature (or global warming) is likely to benefit crop yields or land productivity in Ukraine. This finding is consistent with conclusions in other related literature based in colder high-latitude areas—the warming trend may have positive effects on land productivity in some areas such as northern Europe (Alcamo et al. 2007, Olesen et al. 2002) and North America (Motha and Baier 2005).

Table 1.1 Estimated yield response of five crops in Ukraine to weather variation in FE

	(1)	(2)	(3)	(4)	(5)
Log(yield)	Winter wheat	Spring Barley	Sunflower	Soybean	Corn
Daily temperature	0.028***	0.050***	0.035***	0.088***	0.035***
	(0.005)	(0.007)	(0.007)	(0.009)	(0.012)
Prec (mm)	-0.059*	-0.138***	-0.107***	-0.001	0.033
	(0.033)	(0.028)	(0.021)	(0.023)	(0.034)
$Prec^2$ (mm*mm)	-0.003	0.020***	0.015***	-0.004	-0.008
	(0.006)	(0.005)	(0.003)	(0.003)	(0.005)
Observations	155,215	123,716	112,644	57,578	91,383
R-squared	0.579	0.569	0.612	0.537	0.610

Note: All climate-related variables are calculated based on crop-specific growing season; standard errors are clustered at the rayon level; *** p<0.01, ** p<0.05, * p<0.1

The positive outcomes of rising temperatures on crop yields in Ukraine could be explained by various mechanisms such as the extension of growing seasons, lower risk of frost or icing, etc. that are associated with the higher temperature. For example, the prolongation of the vegetation period associated with the warming climate may contribute to the rising crop yields (especially the winter crops) in Ukraine (Tarariko et al. 2017). Moreover, in a regional study based in western Ukraine, Fischer et al. (2014) conclude that the risk of frosting or icing reduces with the rising temperature, which could potentially benefit crop growth. Additionally, the temperature might also have affected crop yields indirectly through soil fertility (Leirós et al., 1999), pests (Skendžiće et al. 2021, Cannon, 1998), diseases (Coakley et al., 1999), etc., that are associated with temperature. However, the exact effects of rising temperature on the interaction of crops and land, pests, and diseases might vary across contexts and farmers' adaptation capacity.

In terms of precipitation[9], our results suggest that the precipitation in the growing season in Ukraine is significantly associated with the yields of spring barley and sunflower only. However,

[9] The average daily precipitation in growing season for the five crops are 2.23 mm (winter wheat), 2.36 mm (spring barley), 2.62 mm (sunflower), 2.95 mm(soybean), and 2.39 mm (corn) accordingly.

the typical inverse-U relationship between precipitation and crop yield is not found on five crops. While the relatively sufficient level of rainfall in Ukraine might contribute to such a result, the fact that farmers could adapt to rainfall change through irrigation could be the other factor that can explain the weak association between precipitation and yields of crops such as winter wheat, soybean, and corn.

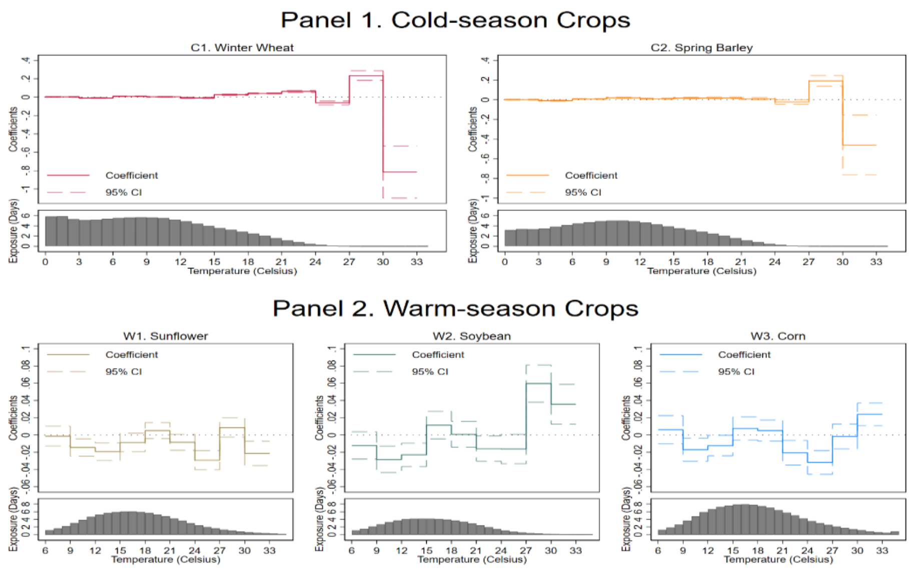

Figure 1.5 Non-linear yield response of five crops to exposure to temperature intervals

Notes: This graph displays changes in log yield if a county is exposed for one extra day to each 3 °C temperature interval relative to a day spent below 0 °C (winter wheat and spring barley) or 6 °C (sunflower, soybean, and corn)

The estimation results of Equation (3) are displayed in Figure 1.5[10]. For winter wheat, we find that exposure to temperatures between 9 °C and 30 °C is positively associated with crop yields, while an additional day of exposure to temperatures above 30 °C decreases yield by 81.7%. Spring barley exhibits a similar temperature response with yield responding positively to exposure temperatures between 6°C and 30°C, while an additional day above 30 °C is associated with a 46% decrease in yield. The relatively small number of observations with daily temperatures over 30°C in the growing season could explain the magnitudes of the coefficients. This is because through March to June, there is only a small number of days when the daily maximum temperature exceeds

[10] The estimated coefficients based on which the graphs are drawn are shown in Table A.1 in the appendix.

30 °C. Therefore, one extra day of exposure to temperatures above 30 °C implies a great increase in the number of days when the daily maximum temperature exceeds 30 °C, indicating a much hotter growing season.

While different in magnitude, several studies have shown the overall negative relationship between the rising temperature and yields of spring-season crops such as winter wheat and spring barley in high-latitude regions such as Denmark (Kristensen et al. 2011), France (Licker et al. 2013), North America (Klink et al. 2014). However, our empirical results suggest that what hurts the yields of these spring-season crops is more likely the exposure to stressful high temperatures, not the temperate warmer environment.

We find strikingly different results on the three warm-season crops— sunflower, soybean, and corn. In contrast with wheat and barley, we do not find any statistical evidence of an upper-temperature threshold above which the crop yields respond negatively to temperature exposure. This contrasts with findings in other contexts such as the U.S. (Burke and Emerick 2016) and China (Chen et al. 2016) which find negative effects for temperatures above 28 or 29 °C. The relatively colder environment in Ukraine could be a possible explanation for such difference, as increased exposure to temperatures above 28 or 29 °C in colder areas with a relatively small scale of exposure might indicate a favorable rising temperature for crop growth rather than an increasing duration of exposure to stressful hot days, such as in hotter regions.

Furthermore, we investigate the heterogeneity in yield-temperature response across regions in Ukraine. For each crop, we divide the sample villages into two categories—warm region and cold region, depending on whether the average daily temperature of the village is above the sample mean. Our results suggest that, for cold-season crops, the negative association between yields and exposure to temperatures above the stressful temperature threshold is more significant in warm regions (Figure A.2 in the appendix) than in cold regions (Figure A.3 in the appendix). While for hot-season crops, it is consistent in both regions that no clear upper-temperature thresholds (above which the crop yields respond negatively to temperature exposure) are detected. The heterogeneity across regions implies that the hotter eastern and southern Ukraine with a higher temperature and less precipitation might be more stressful in mitigating the negative impacts of the warming climate[11].

Based on the results from the temperature intervals approach, we apply the growing degree days (GDDs) specification and estimate equation (4) to examine the relationship between heat accumulation in the growing season and crop yields. To address the previous finding that the temperature in the growing season rarely reaches a potential stressful temperature threshold for the three warm-season crops in Ukraine, the stress degree days (SDDs) term is removed from the equation for sunflower, soybean, and corn. More specifically, we define the lower temperature threshold as 6 °C and loop for all possible upper-temperature thresholds from 25 °C to 34°C and estimate equation (4) for winter wheat and spring barley. On the other hand, for sunflower, soybean, and corn, the lower temperature thresholds are set as 8 °C, 10 °C, and 10 °C accordingly for estimating the impacts of GDDs on the yields of these crops[12].

To determine the appropriate upper-temperature thresholds for winter wheat and spring barley, we utilize both the results from previous temperature bins specification and the results from the regression loop. More specifically, the significant negative coefficients of the temperature intervals above 30 °C from previous results first indicate that the "stressful" temperature thresholds for winter wheat and spring barley are both at around 30 °C. Moreover, the changes in coefficients of GDDs (decreasing) and SDDs (increasing in absolute value) in the regression loop, when a higher temperature threshold is selected, indicate that the turning points for winter wheat and spring barley are at somewhere around 29 °C (Table A.2) and 30 °C (Table A.3) respectively.

---

[11] The time trends of temperature and precipitation in selected growing season of four administrative regions are shown in Figure A.4 in the appendix.

[12] The selection of lower temperature thresholds is mainly based on recommendation from North Dakota Agricultural Weather Network (NDAWN) Center (https://ndawn.ndsu.nodak.edu/help-corn-growing-degree-days.html). We also test with 6 °C, 8 °C, and 10 °C for all five crops as a robustness check. The results are consistent with what we find using our preferred specification.

Table 1.2 Nonlinear impacts of temperature on crop yields (GDDs)

Log(yield)	(1) Winter wheat	(2) Spring barley	(3) Sunflower	(4) Soybean	(5) Corn
$GDDs^{6,29}$ (Mar-Jun, *100)	0.160***				
	(0.015)				
$SDDs^{29}$ (Mar-Jun, *100)	-16.474***				
	(4.304)				
$GDDs^{6,30}$ (Apr-Jun, *100)		0.153***			
		(0.020)			
$SDDs^{30}$ (Apr-Jun, *100)		-31.280**			
		(15.921)			
$GDDs^{8}$ (Jun-Aug, *100)			0.024***		
			(0.007)		
$GDDs^{10}$ (Jun-Jul, *100)				0.169***	
				(0.016)	
$GDDs^{10}$ (Jun-Sep, *100)					0.034***
					(0.010)
Prec (mm)	-0.036	-0.150***	-0.126***	0.000	0.034
	(0.032)	(0.027)	(0.022)	(0.023)	(0.033)
$Prec^2$ (mm*mm)	-0.003	0.024***	0.017***	-0.004	-0.008
	(0.006)	(0.005)	(0.004)	(0.003)	(0.005)
Observations	155,215	123,716	112,644	57,578	91,383
R-squared	0.581	0.569	0.612	0.538	0.610

Note: All climate-related variables are calculated based on crop-specific growing season; standard errors are clustered at the rayon level; *** p<0.01, ** p<0.05, * p<0.1

The results from the GDDs specification after we select 29 °C and 30 °C as upper-temperature thresholds for winter wheat and spring barley respectively, as shown in Table 1.2, are consistent with previous results in the temperature bins specification. For the two cold season crops—winter wheat and spring barley, crop yields respond positively to GDDs while negatively to SDDs, i.e., the heat accumulation when the temperature is below a stressful temperature threshold is positively associated with crop yields, while the heat accumulation after the temperature exceeds the threshold would likely to harm crop yields. However, for the other three hot season crops—

sunflower, corn, and soybean, heat accumulation is positively associated with crop yields. More specifically, one extra growing degree days (GDDs) is associated with a 0.16%, 0.153%, 0.024%, 0.169%, and 0.034% yield increase of winter wheat, spring barley, sunflower, soybean, and corn accordingly. While for winter wheat and spring barley, one extra stress degree days (SDDs) is associated with 16.47% and 31.28% yield loss respectively[13].

B. Implicit adaptations to climate

The approach we apply to unveil the long-run adaptations to climate change is borrowed from Mérel and Gammans (2021), they justified that one can investigate the implicit agricultural adaptations by estimating the yield response to weather and climate conditions in one equation. More specifically, we test whether the crop yields, conditional on the contemporaneous weather realization, are relatively higher in villages where the underlying climate conditions are closer to the realization than in villages where the realization happens to be unusual. More specifically, we estimate equation (5) for each crop and check the negativity of the estimated coefficient (β_3) of the "climate penalty" term $(x_{it} - \overline{x}_i)^2$. The estimated results are shown in Table 1.3.

The results provide evidence of long-run adaptation in winter wheat production to temperature changes with significant negative coefficients of the "climate penalty" term (columns 1), indicating that the yield of winter wheat, conditional on the temperature realization, is relatively higher in locations where farmers are more familiar with the temperature (temperature realization is closer to long-run temperature). Furthermore, the significant negative coefficient of the (squared) precipitation deviation from the long-term mean also provides evidence of adaptations in winter wheat production to precipitation change in the long run (i.e., crop yields in locations where the long-term average precipitation level is closer to the contemporaneous realization of precipitation, conditional on the realization, are relatively greater). However, the result fails to provide evidence of long-run adaptations in spring barley production to temperature or precipitation change (i.e.,

[13] Empirical results are summarized in Table A.4 in the appendix.

negative β_3). For the rest of the three hot season crops—sunflower, soybean, and corn, we find evidence of long-run adaptations to precipitation change but not to temperature change.

Table 1.3 Estimated weather and climate impacts on crop yields

Dependent variable:	Log (yield)				
	(1)	(2)	(3)	(4)	(5)
	Wheat	Barley	Sunflower	Soybean	Corn
Temp in GS	-0.015*	0.016	0.177***	0.229***	-0.185**
	(0.008)	(0.020)	(0.053)	(0.063)	(0.091)
$(\text{Temp in GS})^2$	0.004***	0.002	-0.004***	-0.005***	0.005**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
$(\text{Temp-5-yr Avg.})^2$	-0.005***	0.010***	0.026***	0.073***	0.084***
	(0.001)	(0.002)	(0.005)	(0.008)	(0.010)
Prec in GS	-0.092***	-0.161***	-0.232***	-0.066**	-0.106***
	(0.035)	(0.035)	(0.025)	(0.027)	(0.040)
$(\text{Perc in GS})^2$	0.007	0.026***	0.043***	0.009**	0.025***
	(0.007)	(0.007)	(0.005)	(0.004)	(0.007)
$(\text{Prec-5-yr Avg.})^2$	-0.034**	-0.016	-0.070***	-0.031***	-0.129***
	(0.014)	(0.011)	(0.007)	(0.005)	(0.012)
Constant	1.176***	0.683***	-0.978**	-1.761***	3.096***
	(0.065)	(0.115)	(0.494)	(0.537)	(0.905)
Observations	155,215	123,716	112,644	57,578	91,383
R-squared	0.579	0.569	0.614	0.542	0.615

Note: Standard errors are clustered at the rayon level; *** p<0.01, ** p<0.05, * p<0.1

The surprisingly positive coefficients of the "temperature penalty" term on the three hot season crops implicitly indicate their yields, conditional on the contemporaneous realization of temperature, are even smaller in locations where farmers are already familiar with the temperature. Such results bring up our concern on the applicability of the approach we apply in unveiling the long-run adaptation to climate in contexts where the dome-shaped relationship does not apply so well in relating the crop yields to contemporaneous temperature, as the model is derived based on the commonly used quadratic-in-weather specification.

C. Explicit adaptations in crop diversification

We apply a simple two-limit random-effects Tobit model and estimate equation (6) at the village level to investigate the impacts of (lagged) climate variables (5-year mean of daily temperature and precipitation in the growing season) on crop diversification (measured by Herfindal Index) with the control on (lagged) weather volatilities (standard deviation of daily temperature and precipitation in the growing season) and the total cropping area. The results are shown in Table 1.4.

Table 1.4 Estimated climate effects on crop diversification

Dependent variable: Herfindal Index	(1)
Total cropping area	-0.000***
	(0.000)
Lagged 5-year mean temperature	-0.016***
	(0.002)
Lagged SD. of temperature	-0.002*
	(0.001)
Lagged 5-year mean precipitation	0.061***
	(0.003)
Lagged SD. of precipitation	-0.002***
	(0.001)
Constant	0.489***
	(0.023)
Observations	208,119

Note: *** p<0.01, ** p<0.05, * p<0.1

The results first indicate that the total cropping area is positively associated with crop diversification (negative coefficient), i.e., crops are more diversified as the total land area increases. Moreover, the crop structure is more diversified when the long-term temperature rises or long-term precipitation declines, indicating that farmers in Ukraine are likely to perceive the rising temperature and reducing precipitation as negative shocks for crops, which results in their increasing demand for risk-sharing by crop diversification. Not surprisingly, the negative coefficients of the lagged standard error of temperature and precipitation indicate that crops are more diversified in response to the increasing volatilities of temperature and precipitation in the growing season.

1.5 Conclusion

Given the complexities of crop growth and its interaction with climate conditions, as well as the potential adaptations that can be applied in agricultural practices by farmers, assessing the exact impacts of climate change on agricultural production is never an easy task. In most cases, the results vary across crops and depend largely on the environment of the studied area (geographic location, climate condition, soil type, etc.), as well as the empirical approaches researchers apply.

Our empirical results on yield responses of crops to short-run changes in climate conditions indicate that increasing average temperature is positively associated with the yields of all five main crops in Ukraine. For the two cold season crops—winter wheat and spring barley, crop yield increases with exposure to moderate temperatures. Exposure to temperature intervals above a "stressful" temperature bound (29°C for winter wheat, 30°C for spring barley) harms crop growth. Furthermore, our heterogenous analysis indicates that the negative yield-temperature relationship after the temperature exceeds the stressful temperature thresholds is more significant in warmer regions, suggesting that hotter southern and eastern Ukraine might face more stress in mitigating warming climate in the future.

For the three warm-season crops—sunflower, corn, and soybean, we do not find similar stressful temperature thresholds above which the crop yields respond negatively to temperature increases. Instead, the yields of the three warm-season crops respond positively to overall rises in average temperature and growing degree days (GDDs) in the growing season.

Moreover, our results unveil significant adaptation in winter wheat production to long-run temperature change but no clear evidence of adaptation to temperature is shown for spring barley. We also find evidence of adaptation to long-term precipitation changes for sunflower, soybean, and corn. However, as we explained in previous sections, the applicability of the approach we apply in unveiling the long-run adaptation to climate change might be undermined if the dome-shaped relationship between crop yields and temperature is not suggested. Furthermore, our empirical results on explicit agricultural adaptations to climate change indicate that the long-run temperature is positively associated with crop diversification while the precipitation is negatively associated with crop diversification. We also find that the crop structure is more diversified when the volatilities of temperature and precipitation increase.

While earlier studies have pointed out the importance of threshold temperature effects, our study adds to the literature by highlighting that such threshold effects would be crop specific. To the extent that for a subset of crops, relevant thresholds do not yet seem to have been crossed in Ukraine, changes in cropping patterns seem to offer an opportunity for Ukrainian farmers to adjust to climate change. Follow-up research to assess the extent to which such changes already happened as well as the likelihood that the threshold to stressful temperatures for the more heat-resistant crops will be crossed in the future, will be important to assess how climate change may impact agricultural production in Ukraine and also global food security.

REFERENCES

Adams, R. M., Rosenzweig, C., Peart, R. M., Ritchie, J. T., McCarl, B. A., Glyer, J. D., ... & Allen, L. H. (1990). Global climate change and US agriculture. Nature, 345(6272), 219-224.

Alcamo, J., Moreno, J. M., & Shvidenko, A. (2007). Europe. Climate change 2007: impacts, adaptation and vulnerability.

Allen, J. C. (1976). A modified sine wave method for calculating degree days. Environmental Entomology, 5(3), 388-396.

Annan, F., & Schlenker, W. (2015). Federal crop insurance and the disincentive to adapt to extreme heat. American Economic Review, 105(5), 262-66.

Arellano-Gonzalez, J., & Moore, F. C. (2020). Intertemporal Arbitrage of Water and Long-Term Agricultural Investments: Drought, Groundwater Banking, and Perennial Cropping Decisions in California. American Journal of Agricultural Economics, 102(5), 1368-1382.

Assunção, J., & Chein, F. (2016). Climate change and agricultural productivity in Brazil: future perspectives. Environment and Development Economics, 21(5), 581-602.

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

Behnassi, M., & El Haiba, M. (2022). Implications of the Russia–Ukraine war for global food security. Nature Human Behaviour, 1-2.

Berry, S. L., & Roderick, M. L. (2002). CO2 and land-use effects on Australian vegetation over the last two centuries. Australian Journal of Botany, 50(4), 511-531.

Boychenko, S., Voloshchuk, V., Movchan, Y., Serdjuchenko, N., Tkachenko, V., Tyshchenko, O., & Savchenko, S. (2016). Features of Climate Change on Ukraine: Scenarios, Consequences for Nature and Agroecosystems.

Bradshaw, B., Dolan, H., & Smit, B. (2004). Farm-level adaptation to climatic variability and change: crop diversification in the Canadian prairies. Climatic change, 67(1), 119-141.

Brammer, H. (1990). Floods in Bangladesh: geographical background to the 1987 and 1988 floods. Geographical journal, 12-22.

Branco, D., & Féres, J. (2021). Weather shocks and labor allocation: Evidence from rural Brazil. American Journal of Agricultural Economics, 103(4), 1359-1377.

Brklacich, M., Bryant, C., Veenhof, B., & Beauchesne, A. (2000). Agricultural adaptation to climatic change: A comparative assessment of two types of farming in central Canada. Agricultural and environmental sustainability in the new countryside. Hignell Printing Limited, Winnipeg, 40, 51.

Burke, M., & Emerick, K. (2016). Adaptation to climate change: Evidence from US agriculture. American Economic Journal: Economic Policy, 8(3), 106-40.

Cannon, R. J. (1998). The implications of predicted climate change for insect pests in the UK, with emphasis on non‐indigenous species. Global change biology, 4(7), 785-796.

Chang, C. C. (2002). The potential impact of climate change on Taiwan's agriculture. Agricultural Economics, 27(1), 51-64.

Chen, S., & Gong, B. (2021). Response and adaptation of agriculture to climate change: Evidence from China. Journal of Development Economics, 148, 102557.

Chen, C. C., & McCarl, B. A. (2001). An investigation of the relationship between pesticide usage and climate change. Climatic Change, 50(4), 475-487.

Chen, S., Chen, X., & Xu, J. (2016). Impacts of climate change on agriculture: Evidence from China. Journal of Environmental Economics and Management, 76, 105-124.

Chiotti, Q., & Johnston, T. (1995). Extending the boundaries of climate change research: modelling the farm-level decision-making complex. J Rural Studies, 11(3), 335-350.

Coakley, S. M., Scherm, H., & Chakraborty, S. (1999). Climate change and plant disease management. Annual review of phytopathology, 37(1), 399-426.

Darwin, R. (2001). Climate change and food security (No. 1474-2016-120874).

Darwin, R., Tsigas, M., Lewandrowski, J., & Raneses, A. (1996). Land use and cover in ecological economics. Ecological Economics, 17(3), 157-181.

Dejene, A., Midgley, S., Marake, M. V., & Ramasamy, S. (2011). Strengthening capacity for climate change adaptation in agriculture: experiences and lessons from Lesotho. FAO.

De Loe, R. C., & Kreutzwiser, R. D. (2000). Climate variability, climate change and water resource management in the Great Lakes. Climatic Change, 45(1), 163-179.

de Loë, R., Kreutzwiser, R., & Mararu, L. (1999). Climate change and the canadian water sector: Impacts and adaptation. Guelph: Natural Resources Canada. Retrieved March, 1, 2016.

Dietz, K. J., Zörb, C., & Geilfus, C. M. (2021). Drought and crop yield. Plant Biology.

Downing, T. E. (1992). Climatic change and vulnerable places: global food security and country studies in Zimbabwe, Kenya, Senegal and Chile.

Fang, J. Q., & Liu, G. (1992). Relationship between climatic change and the nomadic southward migrations in eastern Asia during historical times. Climatic Change, 22(2), 151-168.

Fischer, G., & van Velthuizen, H. T. (1996). Climate change and global agricultural potential project: A case study of Kenya. International Institute for Applied Systems Analysis, Laxenburg, Austria, 96.

Fischer, S., Pluntke, T., Pavlik, D., & Bernhofer, C. (2014). Hydrologic effects of climate change in a sub-basin of the Western Bug River, Western Ukraine. Environmental earth sciences, 72, 4727-4744.

Fisher, M., Abate, T., Lunduka, R. W., Asnake, W., Alemayehu, Y., & Madulu, R. B. (2015). Drought tolerant maize for farmer adaptation to drought in sub-Saharan Africa: Determinants of adoption in eastern and southern Africa. Climatic Change, 133(2), 283-299.

Fosu-Mensah, B. Y., Vlek, P. L., & MacCarthy, D. S. (2012). Farmers' perception and adaptation to climate change: a case study of Sekyedumase district in Ghana. Environment, Development and Sustainability, 14(4), 495-505.

Gammans, M., Mérel, P., & Ortiz-Bobea, A. (2017). Negative impacts of climate change on cereal yields: statistical evidence from France. Environmental Research Letters, 12(5), 054007.

Giorgi, F., Bates, G. T., & Nieman, S. J. J. Colin, and A. Ghassem, 2009: Addressing climate information needs at the regional level: The CORDEX framework. WMO Bull, 58, 175-183.

Hassen, A., Talore, D. G., Tesfamariam, E. H., Friend, M. A., & Mpanza, T. D. E. (2017). Potential use of forage-legume intercropping technologies to adapt to climate-change impacts on mixed crop-livestock systems in Africa: a review. Regional Environmental Change, 17(6), 1713-1724.

Hillel, D. (1988). The greenhouse effect and its implications regarding global agriculture. Research bulletin/Massachusetts Agricultural Experiment Station (USA).

Holdridge, L. R. (1947). Determination of world plant formations from simple climatic data. Science, 105(2727), 367-368.

Howden, S. M., Soussana, J. F., Tubiello, F. N., Chhetri, N., Dunlop, M., & Meinke, H. (2007). Adapting agriculture to climate change. Proceedings of the national academy of sciences, 104(50), 19691-19696.

Hulme, M. (Ed.). (1996). Climate Change and Southern Africa: An Exploration of Some Potential Impacts and Implications in the SADC Region: a Report Commissioned by WWF International and Co-ordinated by the Climatic Research Unit, UEA, Norwich, UK. Climatic Research Unit, University of East Anglia.

Hussain, S., Hussain, S., Qadir, T., Khaliq, A., Ashraf, U., Parveen, A., ... & Rafiq, M. (2019). Drought stress in plants: An overview on implications, tolerance mechanisms and agronomic mitigation strategies. Plant Science Today, 6(4), 389-402.

Iglesias, A., Erda, L., & Rosenzweig, C. (1996). Climate change in Asia: a review of the vulnerability and adaptation of crop production. Climate Change Vulnerability and Adaptation in Asia and the Pacific, 13-27.

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. In Press.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., ... & Yiou, P. (2014). EURO-CORDEX: new high-resolution climate change projections for European impact research. Regional environmental change, 14, 563-578.

Kaiser, H. M., Riha, S. J., Wilks, D. S., Rossiter, D. G., & Sampath, R. (1993). A farm-level analysis of economic and agronomic impacts of gradual climate warming. American journal of agricultural economics, 75(2), 387-398.

Kogo, B. K., Kumar, L., & Koech, R. (2021). Climate change and variability in Kenya: a review of impacts on agriculture and food security. Environment, Development and Sustainability, 23, 23-43.

Kurukulasuriya, P., & Rosenthal, S. (2013). Climate change and agriculture: A review of impacts and adaptations.

Lansigan, F. P., De los Santos, W. L., & Coladilla, J. O. (2000). Agronomic impacts of climate variability on rice production in the Philippines. Agriculture, ecosystems & environment, 82(1-3), 129-137.

Lobell, D. B., & Asner, G. P. (2003). Climate and management contributions to recent trends in U. S. agricultural yields. Science, 299(5609), 1032-1032.

Lunduka, R. W., Mateva, K. I., Magorokosho, C., & Manjeru, P. (2019). Impact of adoption of drought-tolerant maize varieties on total maize production in south Eastern Zimbabwe. Climate and development, 11(1), 35-46.

Lutze, J. L., Roden, J. S., Holly, C. J., Wolfe, J., Egerton, J. J. G., & Ball, M. C. (1998). Elevated atmospheric [CO2] promotes frost damage in evergreen tree seedlings. Plant, Cell & Environment, 21(6), 631-635.

Matarira, C. H., Kamukondiwa, W., Mwamuka, F. C., Makadho, J. M., & Unganai, L. S. (1996). Vulnerability and adaptation assessments for Zimbabwe. In Vulnerability and Adaptation to Climate Change (pp. 129-140). Springer, Dordrecht.

Mendelsohn, R., Nordhaus, W. D., & Shaw, D. (1994). The impact of global warming on agriculture: a Ricardian analysis. The American economic review, 753-771.

Mérel, P., & Gammans, M. (2021). Climate Econometrics: Can the Panel Approach Account for Long‐Run Adaptation?. American Journal of Agricultural Economics, 103(4), 1207-1238.

Mirza, M. M. Q., Warrick, R. A., & Ericksen, N. J. (2003). The implications of climate change on floods of the Ganges, Brahmaputra and Meghna rivers in Bangladesh. Climatic change, 57(3), 287-318.

Mohamed, A. B., Van Duivenbooden, N., & Abdoussallam, S. (2002). Impact of climate change on agricultural production in the Sahel–Part 1. Methodological approach and case study for millet in Niger. Climatic Change, 54(3), 327-348.

Molotoks, A., Smith, P., & Dawson, T. P. (2021). Impacts of land use, population, and climate change on global food security. Food and Energy Security, 10(1), e261.

Mortimore, M. J., & Adams, W. M. (2001). Farmer adaptation, change and 'crisis' in the Sahel. Global environmental change, 11(1), 49-57.

Motha, R. P., & Baier, W. (2005). Impacts of present and future climate change and climate variability on agriculture in the temperate regions: North America. Climatic Change, 70(1-2), 137-164.

Mu, J. E., McCarl, B. A., & Wein, A. M. (2013). Adaptation to climate change: changes in farmland use and stocking rate in the US. Mitigation and Adaptation Strategies for Global Change, 18(6), 713-730.

Müller, D., Jungandreas, A., Koch, F., & Schierhorn, F. (2016). Impact of climate change on wheat production in Ukraine. Kyiv: Institute for Economic Research and Policy Consulting, 41.

Newman, J. E. (1980). Climate change impacts on the growing season of the North American" corn belt".

Nhamo, L., Matchaya, G., Mabhaudhi, T., Nhlengethwa, S., Nhemachena, C., & Mpandeli, S. (2019). Cereal production trends under climate change: Impacts and adaptation strategies in southern Africa. Agriculture, 9(2), 30.

Olesen, J. E., & Bindi, M. (2002). Consequences of climate change for European agricultural productivity, land use and policy. European journal of agronomy, 16(4), 239-262.

Patterson, D. T., Westbrook, J. K., Joyce, R. J. V., Lingren, P. D., & Rogasik, J. (1999). Weeds, insects, and diseases. Climatic change, 43(4), 711-727.

Qaim, M., & Zilberman, D. (2003). Yield effects of genetically modified crops in developing countries. Science, 299(5608), 900-902.

Raza, A., Razzaq, A., Mehmood, S. S., Zou, X., Zhang, X., Lv, Y., & Xu, J. (2019). Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. Plants, 8(2), 34.

Rhodes, E. R., Jalloh, A., & Diouf, A. (2014). Review of research and policies for climate change adaptation in the agriculture sector in West Africa. Future agricultures working paper, 90.

Rojas-Downing, M. M., Nejadhashemi, A. P., Harrigan, T., & Woznicki, S. A. (2017). Climate change and livestock: Impacts, adaptation, and mitigation. Climate Risk Management, 16, 145-163.

Rosenzweig, C. (1985). Potential CO 2-induced climate effects on North American wheat-producing regions. Climatic Change, 7(4), 367-389.

Rosenzweig, C., & Hillel, D. (1998). Climate change and the global harvest: potential impacts of the greenhouse effect on agriculture. Oxford University Press.

Rosenzweig, C., & Hillel, D. (2000). Soils and global climate change: Challenges and opportunities. Soil science, 165(1), 47-56.

Rosenzweig, C., & Parry, M. L. (1994). Potential impact of climate change on world food supply. Nature, 367(6459), 133-138.

Rudych, O., & Batazhok, S. (2018). Natural and climatic conditions as a risk factor for agricultural production in Ukraine.

Sakurai, T., & Reardon, T. (1997). Potential demand for drought insurance in Burkina Faso and its determinants. American Journal of Agricultural Economics, 79(4), 1193-1207.

Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to US crop yields under climate change. Proceedings of the National Academy of sciences, 106(37), 15594-15598.

Scott, J. K., Webber, B. L., Murphy, H., Ota, N., Kriticos, D. J., & Loechel, B. (2014). AdaptNRM Weeds and climate change: supporting weed management adaptation.

Sivakumar, M. V. K. (1992). Climate change and implications for agriculture in Niger. Climatic change, 20(4), 297-312.

Skendžić, S., Zovko, M., Živković, I. P., Lešić, V., & Lemić, D. (2021). The impact of climate change on agricultural insect pests. Insects, 12(5), 440.

Smit, B., McNabb, D., & Smithers, J. (1996). Agricultural adaptation to climatic variation. Climatic change, 33(1), 7-29.

Smithers, J., & Blay-Palmer, A. (2001). Technology innovation as a strategy for climate adaptation in agriculture. Applied Geography, 21(2), 175-197.

Sultan, B., Defrance, D., & Iizumi, T. (2019). Evidence of crop production losses in West Africa due to historical global warming in two crop models. Scientific reports, 9(1), 12834.

Tambo, J. A., & Abdoulaye, T. (2012). Climate change and agricultural technology adoption: the case of drought tolerant maize in rural Nigeria. Mitigation and Adaptation Strategies for Global Change, 17(3), 277-292.

Tarariko, O., Ilienko, T., Kuchma, T., & Velychko, V. (2017). Long-term prediction of climate change impact on the productivity of grain crops in Ukraine using satellite data. Agricultural science and practice, 4(2), 3-13.

Tubiello, F. N., Soussana, J. F., & Howden, S. M. (2007). Crop and pasture response to climate change. Proceedings of the National Academy of Sciences, 104(50), 19686-19690.

World Bank. Asia Region. Country Dept. I. (1989). Bangladesh: action plan for flood control. World Bank.

APPENDIX

Table A.1 Nonlinear impacts of temperature on crop yields (3-degree intervals)

Log(yield)	(1) Wheat	(2) Barley	(3) Sunflower	(4) Soybean	(5) Corn
Exposure					
0-3 ºC	0.002	0.001			
	(0.002)	(0.003)			
3-6 ºC	-0.009***	-0.008***			
	(0.002)	(0.003)			
6-9 ºC	0.008***	0.008***	-0.001	-0.012	0.006
	(0.002)	(0.002)	(0.006)	(0.008)	(0.008)
9-12 ºC	0.002	0.020***	-0.014***	-0.028***	-0.017**
	(0.003)	(0.003)	(0.005)	(0.008)	(0.007)
12-15 ºC	-0.008**	0.011***	-0.019***	-0.023***	-0.012**
	(0.003)	(0.003)	(0.005)	(0.007)	(0.006)
15-18 ºC	0.025***	0.018***	-0.009	0.012	0.008
	(0.003)	(0.004)	(0.005)	(0.008)	(0.007)
18-21 ºC	0.040***	0.020***	0.005	0.001	0.005
	(0.004)	(0.005)	(0.005)	(0.008)	(0.006)
21-24 ºC	0.061***	0.010*	-0.008*	-0.016**	-0.021***
	(0.005)	(0.006)	(0.005)	(0.007)	(0.007)
24-27 ºC	-0.063***	-0.022**	-0.029***	-0.016*	-0.032***
	(0.010)	(0.011)	(0.006)	(0.009)	(0.007)
27-30 ºC	0.234***	0.193***	0.009	0.060***	-0.002
	(0.027)	(0.028)	(0.006)	(0.011)	(0.007)
30-33 ºC	-0.817***	-0.460***	-0.021***	0.036***	0.024***
	(0.144)	(0.154)	(0.007)	(0.012)	(0.007)
Observations	155,215	123,716	112,644	57,578	91,383
R-squared	0.589	0.573	0.614	0.542	0.615

Note: Daily precipitation and its square term in growing seasons are controlled; Standard errors are clustered at the rayon level; *** p<0.01, ** p<0.05, * p<0.1

Table A.2 GDDs estimates on winter wheat for various temperature thresholds

Log(yield)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Threshold (°C)	25	26	27	28	29	30	31	32	33	34
GDDs: 6 - Threshold (*100)	0.159***	0.156***	0.157***	0.159***	0.160***	0.160***	0.157***	0.154***	0.152***	0.152***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)	(0.014)
SDDs: > Threshold (*100)	-0.750**	-0.708	-1.638	-5.525***	-16.474***	-54.787***	-175.969***	-533.727***	-19,556.538***	-
	(0.375)	(0.641)	(1.127)	(2.094)	(4.304)	(11.608)	(25.617)	(131.380)	(781.268)	
Observations	155,215	155,215	155,215	155,215	155,215	155,215	155,215	155,215	155,215	155,215
R-squared	0.580	0.580	0.580	0.581	0.581	0.581	0.581	0.581	0.580	0.580

Note: Daily precipitation and its square term in growing seasons are controlled; standard errors are clustered at the rayon level; *** p<0.01, ** p<0.05, * p<0.1

Table A.3 GDDs estimates on spring barley for various temperature thresholds

Log(yield)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Threshold (°C)	25	26	27	28	29	30	31	32	33	34
GDDs: 6 - Threshold (*100)	0.148***	0.146***	0.147***	0.151***	0.153***	0.153***	0.153***	0.149***	0.147***	0.147***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
SDDs: > Threshold (*100)	0.028	0.284	-0.032	-2.293	-8.521	-31.280**	-135.351***	-452.297***	-14,972.049***	-
	(0.451)	(0.792)	(1.476)	(2.874)	(6.045)	(15.921)	(36.216)	(145.184)	(778.662)	
Observations	123,716	123,716	123,716	123,716	123,716	123,716	123,716	123,716	123,716	123,716
R-squared	0.569	0.569	0.569	0.569	0.569	0.569	0.569	0.569	0.569	0.569

Note: Daily precipitation and its square term in growing seasons are controlled; standard errors are clustered at the rayon level; *** p<0.01, ** p<0.05, * p<0.1

Table A.4 Summary of empirical methods and main results

Empirical Methods	Fixed effects (FE)	Temperature bins	Growing degree days (GDDs)
Model specification[1]	Linear	Non-linear	Non-linear
Cold-season crops			
Winter wheat	1 °C rise in average daily temperature is associated with a 2.8% yield increase	Yields respond positively to exposure temperatures between 9°C and 29°C, while an additional day above 30 °C is associated with an 81.7% decrease in yields	1 extra growing degree days ($GDDs^{6,29}$) is associated with a 0.16% increase in yields, 1 extra stressful degree days ($SDDs^{29}$) is associated with a 16.47% yields reduction
Spring barley	1 °C rise in average daily temperature is associated with a 5% yield increase	Yields respond positively to exposure temperatures between 6°C and 30°C, while an additional day above 30 °C is associated with a 46% decrease in yields	1 extra growing degree days ($GDDs^{6,30}$) is associated with a 0.15% increase in yields, 1 extra stressful degree days ($SDDs^{30}$) is associated with a 31.28% yields reduction
Warm-season crops			
Sunflower	1 °C rise in average daily temperature is associated with a 3.5% yield increase	No stressful upper temperature is detected	1 extra growing degree days ($GDDs^{8}$) is associated with a 0.02% increase in yields
Soybean	1 °C rise in average daily temperature is associated with an 8.8% yield increase	No stressful upper temperature is detected	1 extra growing degree days ($GDDs^{10}$) is associated with a 0.18% increase in yields
Corn	1 °C rise in average daily temperature is associated with a 3.5% yield increase	No stressful upper temperature is detected	1 extra growing degree days ($GDDs^{10}$) is associated with a 0.03% increase in yields

Note: [1] specification on the temperature-yield relationship.

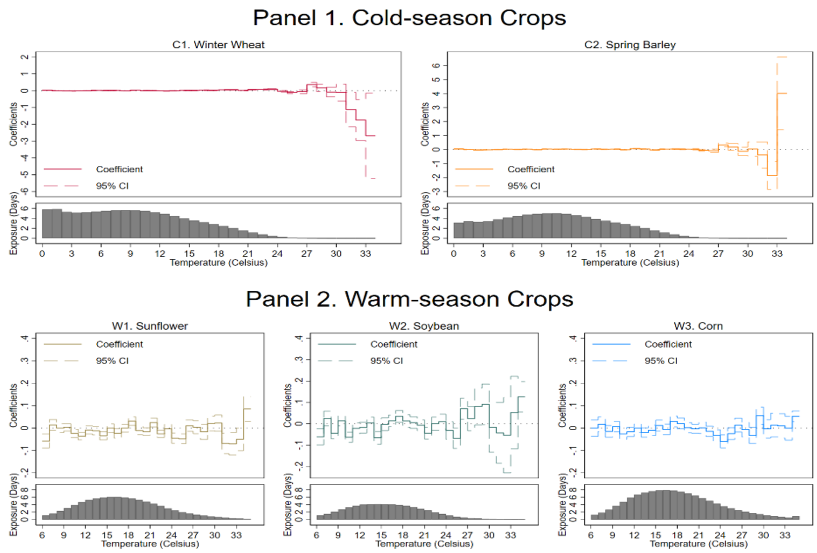

Figure A.1 Non-linear yield response of crops to exposure to 1-degree temperature intervals

Notes: This graph displays changes in log yield if a county is exposed for one extra day to each 1 °C temperature interval relative to a day spent below 0 °C (winter wheat and spring barley) or 6 °C (sunflower, soybean, and corn).

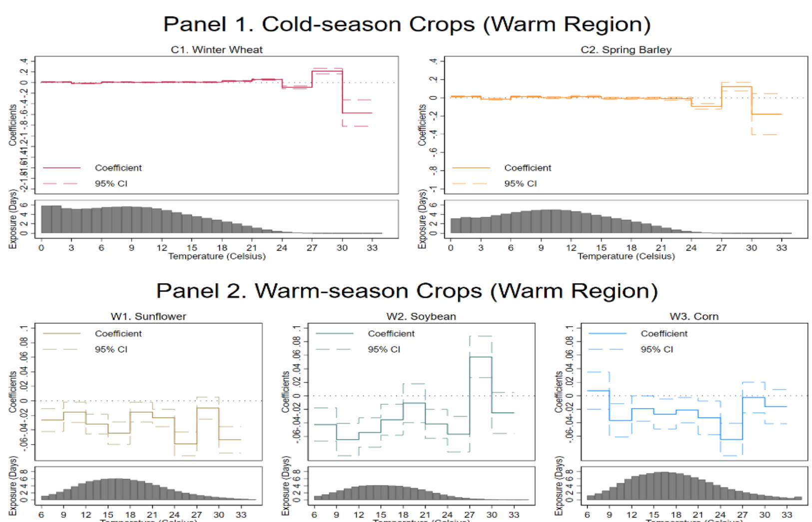

Figure A.2 Non-linear yield response of five crops to exposure to 3-degree temperature intervals in warm regions

Notes: This graph displays changes in log yield if a county is exposed for one extra day to each 3 °C temperature interval relative to a day spent below 0 °C (winter wheat and spring barley) or 6 °C (sunflower, soybean, and corn). For each crop, the sub-sample of warm regions includes all villages where the average daily temperature in the growing season is above the national mean.

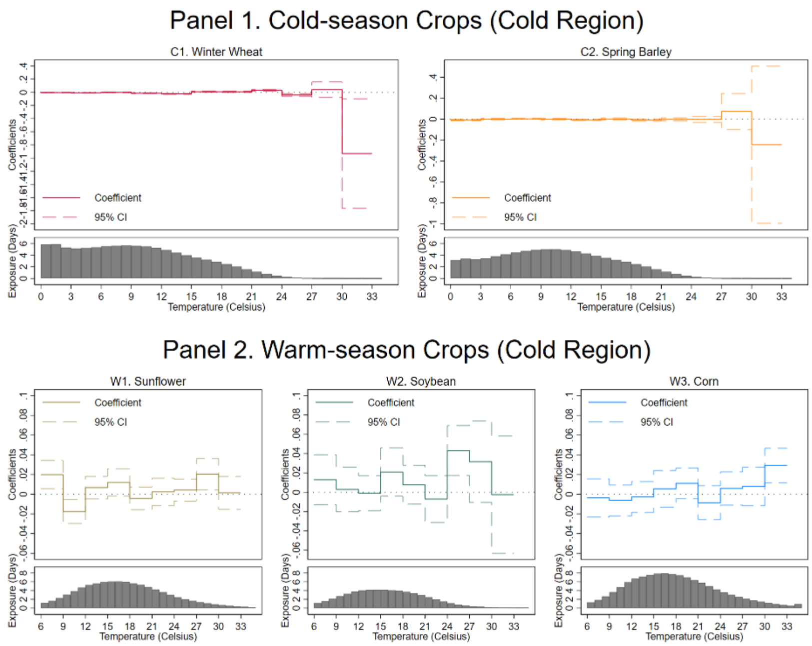

Figure A.3 Non-linear yield response of five crops to exposure to 3-degree temperature intervals in cold regions

Notes: This graph displays changes in log yield if a county is exposed for one extra day to each 3 °C temperature interval relative to a day spent below 0 °C (winter wheat and spring barley) or 6 °C (sunflower, soybean, and corn). For each crop, the sub-sample of cold regions includes all villages where the average daily temperature in the growing season is below the national mean.

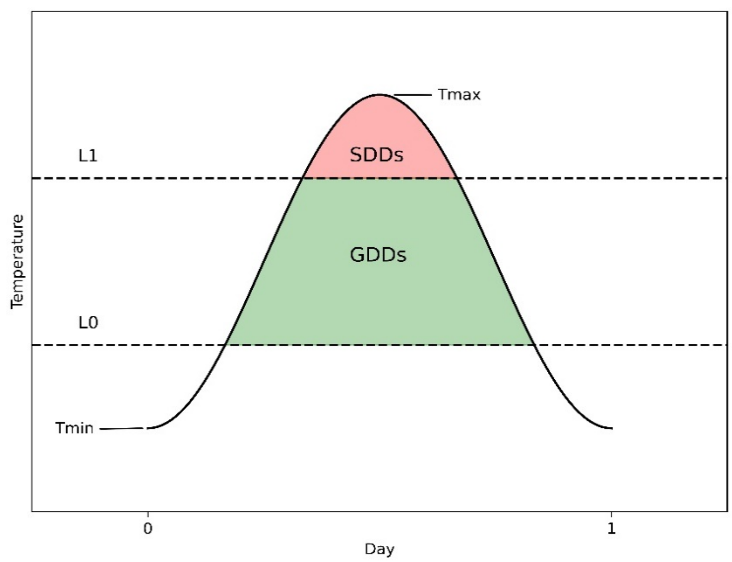

Figure A.4 GDDs and SDDs in sine curve approach

Notes: The sine curve is a simulation of temperature change on a given day based on daily maximum (Tmax) and minimum temperature (Tmin). The green area and red area denote the daily $GDDs^{L0, L1}$ and $SDDs^{L1}$ accordingly, in which the L0 and L1 are the lower and higher temperature thresholds used in the GDDs approach, which are then aggregated across the growing season to calculate the GDDs and SDDs for specific crops.

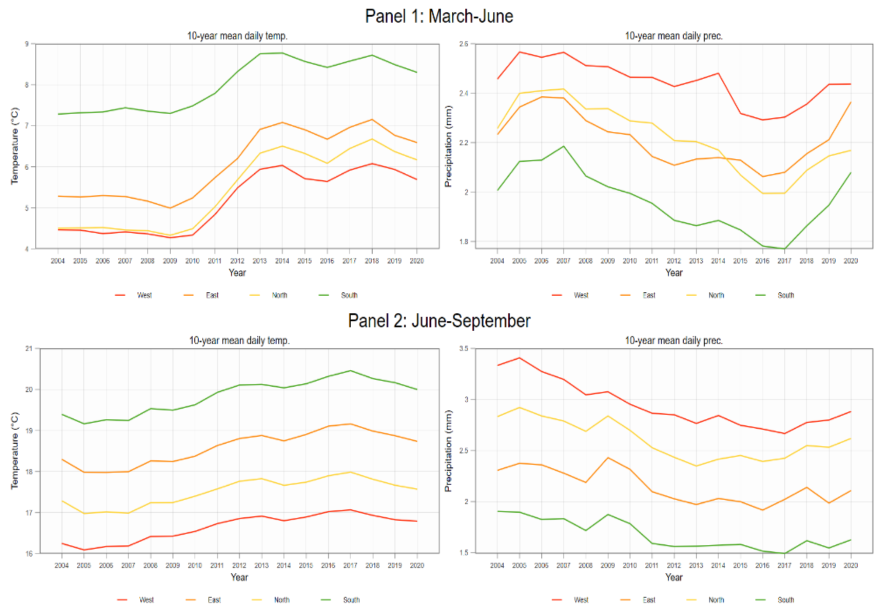

Figure A.5 Time trend of 10-year mean daily temperature and precipitation in selected growing seasons by region

CHAPTER 2: SOCIAL NETWORK AND CONSUMPTION SMOOTHING IN RURAL INDIA

2.1 Introduction

Rural households in developing countries face considerable covariate and idiosyncratic risks or shocks that could result in substantial income variability. Death or illness, weather vagaries, and price risks are just a few of the important examples. However, the shocks or temporary variation in idiosyncratic income does not necessarily cause variation in households' consumption if households can apply buffer strategies. Understanding whether consumption smoothing is achieved in practice, through what specific mechanisms, and how they are shaped by institutional factors has important policy implications for the welfare improvement of rural households in developing countries.

In response to shocks or potential shocks, rural households adopt various *ex-ante* and *ex-post* coping mechanisms to stabilize or smooth consumption. In terms of *ex-ante* strategies, people can store food (Kazianga and Udry 2006), work in the off-farm sector (Rosenzweig and Stark 1989, Grimard 1997, Morten 2019), accumulate assets (Rosenzweig and Wolpin 1993, Fafchamps 1998) and savings or securities (Deaton 1989, Carroll et al. 1992). *Ex-post* strategies in practice include labor pooling (Fafchamps 1992), child fostering (Ksoll 2007), gifts or transfers from others (Stark and Lucas 1988, Rosenzweig and Stark 1989, Fafchamps and Lund 2003), and formal and informal insurance (Udry 1994, Ligon et al. 2000, Fafchamps and Gubert 2007).

Beyond the identification of specific consumption smoothing mechanisms, a large body of related literature focuses on investigating whether one's consumption is smoothed in practice against various shocks with all available buffering strategies. Full consumption smoothing could be achieved through self-insurance or risk-sharing with others. A large literature in this area tested whether efficient risk sharing is achieved within specific groups (Cochrane 1991, Townsend 1994, Jalan and Ravallion 1999, Ravallion and Chaudhuri 1997). For instance, Cochrane (1991) shows that full insurance within the village is rejected for long illness and involuntary job loss. Townsend (1994) tests the full insurance model in the context of India and finds that while household consumption comoves with village average consumption, the full insurance model is rejected significantly. Using data from China, Jalan and Ravallion (1999) also provide evidence of partial insurance rather than full insurance within the village.

In this paper, we first construct a benchmark full insurance model and derive the testing strategy of full insurance or efficient risk-sharing in rural India. Using a large sample dataset from India, we find evidence that rural households' consumption in India is significantly associated with their idiosyncratic income when the co-movement shocks are controlled, which rejects the assumption of full insurance within the village. While this finding is consistent with what is found in similar literature such as Townsend (1994), Jalan and Ravallion (1999), Ravallion and Chaudhuri (1997), etc., our method contributes to the literature by instrumenting the idiosyncratic income in the analysis with a more exogenous instrument variable—rainfall deviation to address the endogeneity issue related to income. Comparably, income variables in similar studies are either not instrumented (e.g., Townsend 1994) or instrumented with relatively more endogenous variables such as lagged income variables (e.g., Jalan and Ravallion 1999).

Beyond specific mechanisms, social network is another important determinant of consumption smoothing that has been investigated in the literature, as it can not only determines the potential risk-sharing groups but also affect various consumption smoothing mechanisms. For instance, different forms of risk-sharing groups in practice such as health insurance groups, funeral societies, or other mutual fire insurance groups, are more or less formed within the social network. Social networks also serve in relaying information about jobs or business opportunities, which helps households buffer shocks by earning alternative income. For example, Granovetter (2000) and Munshi (2003), and Beaman (2012) provide evidence of how family and ethnic networks provide information on business opportunities and help individuals deal with unemployment and business risks. Other consumption smoothing mechanisms such as transfer or gifts, or even formal credits depend largely on social networks as well.

In the context of India, the caste system embedded deeply in the society is perhaps the most important form of social networks or social interactions. As a social identity or stratification system, caste plays an important role at almost every stage of most Indian's life (Munshi 2019). A large literature has studied the role of caste in determining factors that affect consumption smoothing or risk-sharing. For example, Munshi (2019) suggests that in the context of rural India, the caste is the social group around which mutual insurance arrangements are organized. The caste institution also determines the decisions on schooling and employment (Dr`eze and Kingdon 2001; Munshi and Rosenzweig 2016), education performance (Hanna and Linden 2012), and even marriage (Munshi and Rosenzweig 2006; Luke and Munshi 2011; Banerjee et al. 2013). Under the caste system, people in the same caste intend to reside close, do similar jobs, share similar ideas, and marry each other.

While the literature has provided rich evidence on how the caste institution affects the formation of social networks and therefore consumption smoothing, the heterogeneity in consumption smoothing abilities across caste groups, to our knowledge, has not been investigated in literature so far. In this paper, we compare the consumption smoothing ability within the village or caste group, measured by the degree to which household consumption depends on idiosyncratic income change conditional on covariate income shocks. Our results soundly reject the full risk-sharing assumption within the village. However, our analysis at the caste group level indicates that except for the highest caste group, the consumption of the household is fully insured among each of the three backward caste groups.

Compared to most related literature suggesting that the poorer households are disadvantaged in consumption smoothing due to the lack of enough buffering mechanisms (e.g., Jalan and Ravallion 1996), our finding that consumption is smoother in backward caste groups is quite unexpected. To explore the underlying explanations for this unexpected finding, we first investigate how the degree of consumption smoothing varies across households with different expenditure shares on food items. We find that consumption is smoother in households with a larger share of expenditure on food items, suggesting that the relatively higher food expenditure of households in backward caste groups might contribute to their smoother consumption against the idiosyncratic income shocks, as food consumption is less adjustable and less expensive. The other related possible explanation is that while the backward caste groups are generally more constrained by self-insurance ability, they are likely to rely more on their communities to help them buffer income shocks.

Furthermore, our paper explores how caste-based social interaction norms correspond to consumption smoothing—the other contribution we try to make with this study. Comparably, most related studies focus more on how the social network affects specific consumption smoothing mechanisms such as gifts and informal loans (Fafchamps and Lund 2003), migration work or business opportunities (Granovetter 2000, Munshi 2003, and Beaman 2012), etc., and how the shock passes through the social networks (Kinnan et al. 2020). None of the literature, to our knowledge, examines the correspondence between the caste-based social interaction norms, defined as whether households interact with other castes or caste groups, and consumption smoothing.

Whether people interact with other castes or caste groups, in the context of rural India, reflects to what extent is the households' social interaction constrained by the barriers between castes that embedded in the caste institution in India. We posit that while those households who only interact with people within their own caste or caste group are disadvantaged in sharing risk across castes or caste groups, they are more likely to receive stronger support from members of their own caste or caste group to help buffer income shocks. This is supported by our empirical result, as we find that the consumption of households with no inter-caste social interactions (i.e., interactions with other castes or *jatis*) or inter-group social interactions (i.e., interactions with other caste groups) is relatively smoother. This finding provides indirect evidence of the mechanism of risk-sharing within the caste and contributes to the literature by providing extra evidence of how the caste institution corresponds to consumption smoothing in rural India.

The last social network-related issue we investigate is the association between migration working experience and consumption smoothing. While it is argued in the literature that households with migrant members are likely to have reduced access to rural caste networks as they cannot be as easily punished by the network and the superior outside options they can have when they are excluded from the network (Munshi and Rosenzweig 2016), we find that the consumption of households with migration working experience is relatively smoother against the idiosyncratic income shock. This indicates that the positive impacts of migration working on consumption smoothing, such as the alternative income source from the non-agricultural sector and the possible stronger social network outside the village, dominate potential disadvantages of migration in terms of weakened local networks.

The rest of the paper is organized as follows. In section 2, a benchmark theoretical model is constructed and the testing strategy of efficient risk-sharing or full insurance is derived. Section 3 introduces the dataset we use for empirical analysis and some important descriptive statistics. Empirical results are reported in section 4 and section 5 concludes the paper.

2.2 Theoretical Framework

To test the hypothesis of full insurance (or efficient risk sharing), we adopt the benchmark model that has been widely used in literature (Townsend, 1994; Ravallion and Chaudhuri, 1992; Jalan and Ravallion, 1999). One of our main objectives is to test whether full insurance or efficient risk sharing is achieved within the risk-sharing group, i.e., households' consumption does not respond to idiosyncratic income shocks when the co-movement in income at the group level is controlled. Furthermore, we extend our analysis by incorporating caste and caste-based social networks. We start by constructing a general theoretical framework of risk-sharing, from which the test of efficient risk-sharing that guides our empirical analysis is derived.

Assume a simple small open economy of resource-sharing or coinsurance group with N households. The uncertain income of each household $i = 1,2 \dots N$ at period t, $y_t^i(s_{\tau t})$ depends on the state of nature, $s_{\tau t}$. The nature state $\tau = 1,2,\dots,S$ achieves in period t with probability $\pi(s_{\tau t})$. The common utility function of household i is $u\left(c_t^i(s_{\tau t})\right)$, in which $c_t^i(s_{\tau t})$ denotes household consumption. The preference is assumed to be unchanged over time and states. Savings or any form of assets are not allowed in the economy.

The intertemporal expected utility of HH i is represented as

$$EU = \sum_{t=0}^{\infty} \gamma^t \sum_{\tau=1}^{S} \pi(s_{\tau t}) u(c_t^i(s_{\tau t})) \tag{1}$$

where γ is the time preference.

The Pareto-optimal allocation of consumption within the group, therefore, optimizes the social planning problem:

$$\max_{\{c_t^i(s_{\tau t})\}} \sum_{i=1}^{N} w^i \sum_{t=0}^{\infty} \gamma^t \sum_{\tau=1}^{S} \pi(s_{\tau t}) u\left(c_t^i(s_{\tau t})\right) \tag{2}$$

$$s.t. \sum_{i=1}^{N} c_t^i(s_{\tau t}) = \sum_{i=1}^{N} y_t^i(s_{\tau t})$$

where w^i is the welfare weight of each household i. F.O.Cs of the optimization problem could be derived as:

$$w^i \gamma^t \pi(s_{\tau t}) u'\left(c_t^i(s_{\tau t})\right) = \lambda(s_{\tau t}) \sum_{i=1}^{N} y_t^i(s_{\tau t}) \tag{3}$$

or,

$$w^i u'(c_t^i) = \lambda_t \tag{4}$$

for each period t, where w^i is the household-specific welfare weight that remains unchanged over time and λ_t measures the group-specific resource constraint in period t.

The reduced form of equation (4) could be written as

$$c_{ivt} = \alpha\lambda_{vt} + \beta y_{ivt} + u_i + \varepsilon_{ivt} \tag{5}$$

where the notation v denotes the risk-sharing group (e.g., village, or one of the four caste groups). Note that the coefficient of household income y_{ivt}, β is expected to be 0 if full insurance is achieved within the group according to equation (4).

The first-difference version of equation (4) is

$$c_{ivt} - c_{ivt-1} = \alpha(\lambda_{vt} - \lambda_{vt-1}) + \beta(y_{ivt} - y_{ivt-1}) + (\varepsilon_{ivt} - \varepsilon_{ivt-1})$$

$$i.e.\ \ \Delta c_{ivt} = \sum_{j,k} \delta_{jk} D_{ivt}^{jk} + \beta \Delta y_{ivt} + \Delta\varepsilon_{ivt} \tag{6}$$

where D_{ivt}^{jk} is a group-time dummy variable that equals 1 if group $j = v$ and period $k = t$, and zero otherwise, which controls for the aggregate shock at the group level. This suggests the empirical strategy of testing full insurance we use in this paper, i.e., testing the null hypothesis H0: $\beta = 0$.

The other widely used test of full insurance developed by Townsend (1994) could also be derived by imposing a specific exponential utility function of the following form in the framework:

$$u(c_t^i) = \left(-\tfrac{1}{\sigma}\right) \exp\left(-\sigma c_t^i\right) \tag{7}$$

Substituting into equation (4), we have

$$w^i \exp(-\sigma c_t^i) = \lambda_t \tag{8}$$

or,

$$c_t^i = (\log w^i - \log \lambda_t)/\,\sigma \tag{9}$$

Add up for all households:

$$\sum_{i=1}^{N} c_t^i = (\sum_{i=1}^{N} \log w^i - N \cdot \log \lambda_t)/\,\sigma$$

$$N \cdot \overline{c_t} = (\sum_{i=1}^{N} \log w^i - N \cdot \log \lambda_t)/\,\sigma$$

$$\log \lambda_t = \frac{\sum_{i=1}^{N} \log w^i}{N} - \sigma \cdot \overline{c_t} \tag{10}$$

where $\overline{c_t}$ denotes the average consumption within the group in period t.

Substituting into equation (9), we have

$$c_t^i = \overline{c_t} + (\log w^i - \frac{\sum_{i=1}^{N} \log w^i}{N})/\,\sigma \tag{11}$$

which predicts that household consumption is independent of the idiosyncratic income shock when the aggregate consumption is controlled. The resulting test of full insurance parallel to equation (6) can be derived using the first difference method:

$$\Delta c_{ivt} = \alpha \Delta c_{vt} + \beta \Delta y_{ivt} + \Delta \varepsilon_{ivt} \tag{12}$$

where Δc_{vt} indicates the change in average consumption within the group. The test of full insurance will be conducted by testing the null hypothesis H0: $\alpha = 1$ and $\beta = 0$. However, as suggested by Ravallion and Chaudhuri (1997), the estimates from equation (12) will be biased if there exists a co-movement in household-level income changes within the group. As a result, our empirical test will be mainly based on equation (6).

While the above framework is a benchmark model that is widely used in the literature, the recent development in the literature has extended the framework by adding savings or other buffer assets, endogenous market prices, heterogeneous preferences, etc. into the model. However, since this study aims more on exploring the role of social networks on consumption smoothing, we didn't extend the benchmark model further by including these variables or factors.

2.3 Data and Descriptive Statistics

This section introduces the dataset used in our analysis and presents some key descriptive statistics. The key data we use come from the most recent 3 rounds of the Rural Economic and Demographic Survey (REDS) conducted in 1999, 2006, and 2016. As a micro survey dataset administered by the National Council for Applied Economic Research of India, REDS provides comprehensive information on thousands of rural households in 242 villages within 17 major states of India. Each round of REDS contains three components: 1) the village questionnaire that records detailed information on village governance, economy, and history, etc., 2) the census data that collects basic information of all households in surveyed villages, and 3) the household questionnaire that records detailed information on household demographics, income, assets, consumption, employment, agricultural inputs and outputs, etc. Since the census data in the first two rounds do not include information on household consumption, data from the household questionnaire is mainly used in our empirical analysis. The census data is only used to describe the rural employment distribution and construct variables of caste-based social interaction norms.

In terms of sample size, 7,474 households in REDS were surveyed in 1999, in which 4,562 households (61%) did not split and were traced in 2006, 537 households (7%) are split into 1,223 new households. In the second round in 2006, 2,745 new households were added to maintain the representativeness level of the survey. As a result, the total sample size increased to 8,569 in 2006, in which 4,569 households (53%) remains without split and were traced in 2016. After cleaning, we have a total number of 15,851 households across three panels (4,562, 6,720, and 4,569 in 1999, 2006, and 2016 accordingly). Note that 2,151 households in the sample are in the first two panels only, and 2,158 households are in the latest two panels only, 2,411 households exist in both 3 panels.

Detailed rainfall data from National Oceanic and Atmospheric Administration (NOAA) is also collected and merged into our data based on the geographic location of the village. For each village, we calculate the geographic distances between the village and all observation sites and select the closest site to match each village.

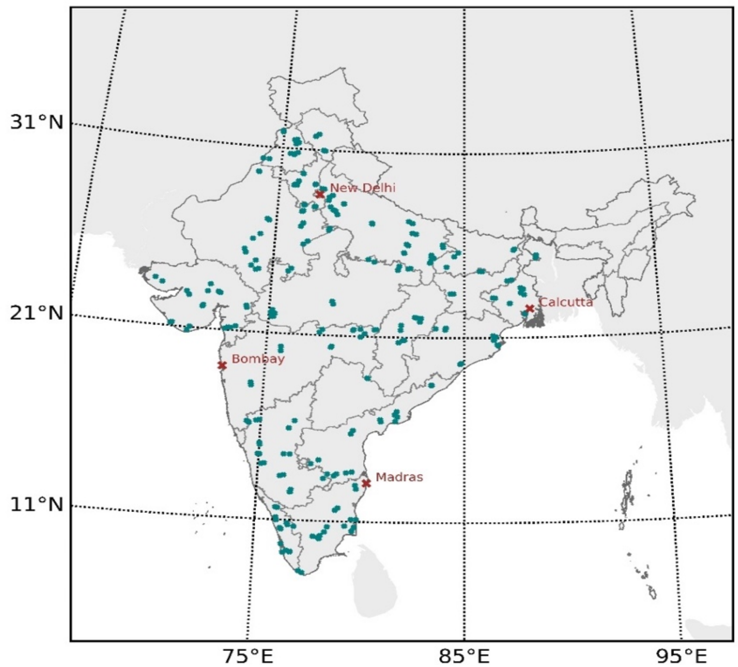

Figure 2.1 Geographical distribution of surveyed villages in REDS

The geographic locations of all surveyed villages are shown in Figure 2.1, in which the blue dots marked the location of each village. The distribution map suggests that, as a representative dataset in India, REDS covers most Indian states except for several northeastern states.

2.3.1 Descriptive Statistics

The descriptive statistics of all samples are shown in Table 2.1. The total sample size of our data is 15,849, of which 90% of households have a male household head. The average age of household heads is 51, and about 86% of them are married. The average household size is 5.42, and the mean number of male and female laborers (aged between 15 and 65) and children are 1.88, 1.80, and 1.49 accordingly. Household heads' mean length of education is about 6 years, suggesting a relatively low education level.

Table 2.1 Household descriptive statistics of all samples by castes

	Total		Caste Groups			
			SC	ST	OBC	OC
	Mean	SD	Mean	Mean	Mean	Mean
Male head	0.90	0.30	0.90	0.91	0.90	0.90
Age of head	50.86	13.87	49.15	48.19	50.80	52.64
Religion: Hindu	0.90	0.30	0.92	0.99	0.90	0.87
Head: Married	0.86	0.35	0.86	0.87	0.85	0.85
Household size	5.42	2.88	5.27	5.26	5.48	5.45
No. of male labor force	1.88	1.21	1.77	1.81	1.89	1.94
No. of female labor force	1.80	1.07	1.72	1.75	1.80	1.84
No. of child	1.49	1.61	1.55	1.52	1.54	1.34
Mean number of years of education (HH head)	5.92	3.86	5.81	6.36	5.93	5.82
Size of own land (acres)	3.09	5.92	1.28	2.81	3.01	4.26
Irrigated land (acres)	1.87	4.09	0.62	1.31	1.84	2.73
Unirrigated land (acres)	1.22	4.30	0.65	1.50	1.17	1.53
Annual HH income (1000 Rs.)	70.89	191.90	47.99	46.63	68.73	93.69
Income (Agriculture)	33.07	127.36	15.53	26.08	31.88	46.29
Income (Livestock)	3.40	14.37	1.98	1.88	3.24	4.86
Income (Ag wage)	3.12	8.56	5.68	3.83	3.05	1.69
Income (Non-Ag wage)	7.46	16.57	10.80	7.37	7.69	5.34
Income (Salary)	9.15	30.30	7.55	2.62	8.06	13.79
Income (Self-employment)	7.05	119.39	2.10	1.86	8.58	8.57
Income (Others)	7.65	68.65	4.35	2.98	6.23	13.15
Log of HH income per capita	9.07	0.98	8.78	8.84	9.03	9.35
Annual HH expenditure (1000 Rs.)	32.51	130.99	26.26	22.68	32.81	38.17
Expenditure (Cereal)	5.27	4.12	4.74	4.29	5.37	5.68
Expenditure (Pulse)	1.29	1.28	1.11	1.09	1.27	1.48
Expenditure (Food)	17.07	10.53	14.87	13.21	16.79	19.84
Expenditure (Non-food)	15.44	129.85	11.40	9.46	16.02	18.33
Log of HH expenditure per capita	8.61	0.55	8.48	8.35	8.59	8.78
No. of observations	15,849	15,849	2,347	1,336	7,661	4,505

Note: All income and expenditure-related variables are measured in 1999 Indian Rupees (1000 Rs); SC-Scheduled Castes, ST-Scheduled Tribes, OBC-Other Backward Castes, OC-Other Castes.

Households in our sample own 3.09 acres of land on average, and the mean areas of owned irrigated land and unirrigated land are 1.87 and 1.22 acres accordingly. Households earn 70,890 Rs annually on average, measured in 1999 Indian Rupees[14]. Not surprisingly, income from agriculture or farming is the major source of household income. The mean annual household expenditure is 32,510 Rs, which measures the household-level consumption in our empirical analysis. Averagely, households spend about 17,070 Rs on food items and 15,440 Rs on non-food items each year.

Before describing the difference in summary statistics across caste groups, as implied in statistics in the last four columns of Table 2.1, it is necessary to briefly introduce the origin of the Indian caste system. Our empirical analysis on consumption smoothing is also largely based on the caste-based social network.

As summarized in a comprehensive review paper of caste and Indian society by Munshi (2019), the caste system in India is a system of social stratification that comprises four hierarchical classes or *varnas*—Brahmins, Kshatriyas, Vaishya, and Shudras, which could be dated back as far as 1500-500 BCE. The other population groups in India that were historically excluded from the *varna* system and were regarded as untouchables are known as *Dalits* today. The five hierarchical classes—the four *varnas* and *Dalits* (or untouchables), stratified the majority of Indian society in almost every aspect of people's life. Within each class, there are hundreds of castes (or *jatis)*.

Historically, the caste plays an important role in determining the economic, social, and political activities of most Indians. For example, the caste in India could determine education resources people have access to (Dr`eze and Kingdon 2001) and performance in school (Hanna et al. 2012), the marriage partners (Munshi and Rosenzweig 2006, Luke and Munshi 2011, Banerjee et al. 2013), and jobs and employment (Hnatkovska et al. 2012), etc. More importantly, the power of castes is reinforced through social networks and social interactions. These networks are found in the job market (Munshi 2011, Munshi and Rosenzweig 2016), industry and business (Banerjee

[14] Official exchange rate in 1999 is 43.06 Rs per US$ (Data source: The Word Bank)

and Munshi 2004, Iyer et al. 2013), and informal financial activities such as loans or insurance (Munshi and Rosenzweig 2016), etc.

The disparities or discriminations induced by caste identity evoked public concern for the rights or welfare of those historically disadvantaged caste groups. As a result, affirmative action administered by the government is introduced in India, which reserves positions in central government and educational institutions for historically disadvantaged castes, aiming at relieving inequalities in education, employment, and public governance across castes. In the political area, the reservation policy is applied in the election of local politics. A large body of literature has evaluated the effect of affirmative action on politics (Munshi and Rosenzweig 2006, Luke and Munshi 2011, Chattopadhyay and Duflo 2004), education (Bertrand et al. 2010, Bagde et al. 2016, Cassan 2019), and access to public resources (Pande 2003, Besley et al. 2004).

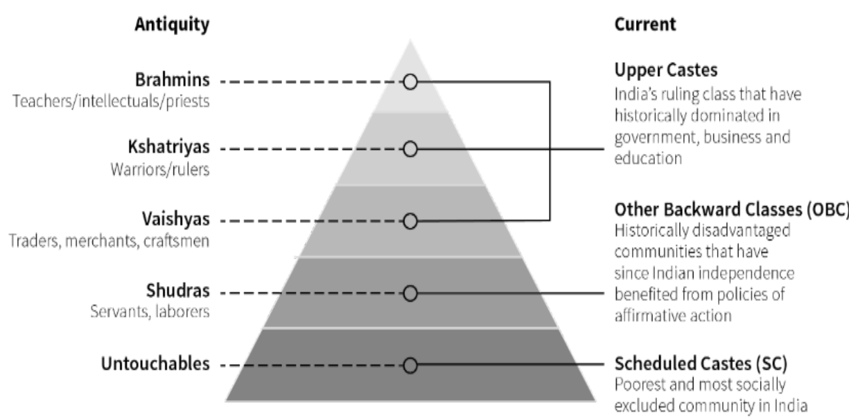

Figure 2.2 Antique and current caste classes in India[15]

The historically disadvantaged groups that benefit from affirmative action include Scheduled Castes (SC), indigenous ethnic groups that are classified as Scheduled Tribes (ST), and Other Backward Castes (OBC) that are included more recently. The rest castes that are not on the list of

---

[15] Figure source : http://fingfx.thomsonreuters.com/gfx/rngs/INDIA-ELECTION/010031Y54EE/index.html

disadvantaged castes are classified as Other (forward) Castes (OC). Figure 2.2 describes the difference between the antique and current caste classes in India, in which the Upper Castes in the current stratification are equivalent to the Other Castes (OC) group in our classification. The other category of class we used in our analysis—Scheduled Tribes (ST) is not included in the caste system as they are the indigenous ethnic groups.

The inequality across the four caste groups, SC, ST, OBC, and OC as documented in other literature is suggested in the REDS survey as well. As indicated in the last four columns of Table 2.1, while the demographics of household heads are almost indifferent across four caste groups, an observable larger size of own land, higher household income and expenditure of forwarding caste groups unveils significant economic disparity across caste groups.

Table 2.2 Employment of household heads by castes in 2016

	Total	**SC**	**ST**	**OBC**	**OC**
Self-employment Agriculture	11.90%	4.15%	9.41%	13.90%	16.00%
Self-employment Non-Ag	7.86%	3.69%	2.05%	8.24%	13.80%
Casual Labor	48.90%	58.40%	55.80%	46.80%	41.40%
NREGA Worker	9.41%	13.00%	17.20%	9.30%	2.56%
Salary Worker	7.51%	7.47%	5.83%	7.08%	9.55%
Others	14.40%	13.30%	9.70%	14.70%	16.70%
No. of observations	53,225	10,904	3,845	28,869	9,607

Note: The data is from the census survey of REDS (2016), which has a relatively larger sample size than the survey data; The sample only includes the employed household heads in 2016; NREGA (Mahatma Gandhi National Rural Employment Guarantee Act), an act passed in 2005 that aims to guarantee the "right to work" by providing at least 100 days of wage employment in a financial year to volunteered adult labors.

The other aspect of inequality that corresponds to the caste identity is reflected in the difference in employment or labor allocation across caste groups. Table 2.2 describes the employment type of household heads from the census data of REDS in 2016, suggesting that the proportions of self-employment agriculture and non-agriculture, salary worker that normally require more capital, technical ability, or knowledge are significantly higher in forward caste groups. While in comparison, the proportion of casual labor is higher in backward caste groups.

Correspondingly, the participation in the NREGA (The Mahatma Gandhi National Rural Employment Guarantee Act) program, a labor program that provides wage employment opportunities to poor adult applicants, is also significantly higher in backward caste groups while the participation rate of other (forward) castes (OC) in NREGA is only 2.56%.

2.3.2 Social Networks or Interaction Within and Across Castes

This section introduces the method we use to measure the social network and interaction norms within and across caste or caste groups, which is also the core of our empirical analysis on caste-based social networks and consumption smoothing.

Answers to the following two questions from REDS (2006) are used to measure the social network:

Q1: *If you wish to borrow Rs.10,000 or any amount to meet a family need, whom would you approach to in the village?*

Q2: *If you wish to borrow simple food items such as chilies, spices, vegetables, etc., whom do you generally go to in the village?*

For each question, each household in the survey is asked to list at most 3 residents within the village, of which the name, member ID, and household ID are recorded and could be matched with the census data to collect their caste information. The second question is answered by members in households who make decisions on cooking.

In Table 2.3, each number in the first 4 columns indicates the average number of households from each of the four caste groups with whom an average household of a given caste group interacts. Panel 1 and 2 report the results from two network-related questions accordingly, and panel 0 shows the merged results from two questions. As an example, the number 0.13 in row 2, column 1 of panel 1 indicates that each household in the ST group would approach 0.13 households from the SC group to borrow money on average. Note that for the merged result in panel 0, the same household that has been listed in response to both question 1 and question 2 is counted only once.[16]

Table 2.3 Social interactions within and across caste and caste groups

	SC	ST	OBC	OC	**All**	Av. share of interactions within the same group	Av. share of interactions within the same caste	Share of HHs with intra-group interactions only	Share of HHs with intra-caste interactions only
Panel 0: Total									
SC	**1.33**	0.05	0.80	0.63	**2.85**	46.94%	38.32%	32.83%	26.35%
ST	0.21	**1.85**	0.67	0.23	**2.99**	60.25%	54.83%	46.81%	40.77%
OBC	0.35	0.09	**2.07**	0.42	**2.98**	68.54%	39.49%	50.22%	24.60%
OC	0.36	0.05	0.66	**1.98**	**3.08**	61.94%	48.53%	41.82%	30.99%
Panel 1: Q1									
SC	**0.71**	0.03	0.45	0.32	**1.53**	45.81%	37.14%	37.26%	29.84%
ST	0.13	**1.04**	0.39	0.15	**1.72**	58.48%	53.09%	50.36%	44.70%
OBC	0.20	0.05	**1.14**	0.23	**1.65**	68.52%	38.36%	57.18%	29.48%
OC	0.18	0.03	0.36	**1.06**	**1.66**	61.41%	47.92%	48.70%	37.08%
Panel 2: Q2									
SC	**0.90**	0.03	0.51	0.43	**1.89**	47.99%	39.53%	38.81%	31.58%
ST	0.14	**1.25**	0.41	0.13	**1.95**	62.14%	56.71%	53.52%	47.48%
OBC	0.23	0.05	**1.30**	0.27	**1.88**	68.74%	40.31%	56.88%	30.78%
OC	0.24	0.04	0.43	**1.26**	**1.99**	62.18%	48.90%	48.16%	36.19%

Note: The data is from the census survey of REDS (2006).

The descriptive result first indicates that there is no significant difference in the total number of people one would approach for financial support across caste groups, which is likely because each household is asked to list 3 residents for each question at most. However, more information on differences in interaction norms across caste groups could be drawn from the composition of these interactions.

The last four columns of Table 2.3 list several summarized results of the response to the network-related questions, from which the key features of the social interaction norm within and

[16] Therefore, the number in panel 0 is not a simple arithmetic sum of those in panel 1 and panel 2.

across castes or caste groups are indicated. As shown in column 6, the average share of social interactions within the same caste group exceeds 60% except for the SC group (one possible explanation is that they have to approach economically more advanced households, who are likely coming from other caste groups, for help when they face financial difficulties). The average share of social interactions within the same caste in column 7 follows a similar pattern with smaller numbers, as normally there is more than one caste within each caste group. The last two columns present the proportions of households with intra-group interactions only and intra-caste interactions only accordingly, suggesting that a large portion of households would only find households within the same caste or the same caste group for financial help if needed.

2.4 Empirical Results

This section reports results from our empirical analysis guided by the theoretical frameworks developed in section 2. We start with the empirical test of efficient risk-sharing or full insurance within the village based on the tests we derived in Section 2. Based on this, we investigate the heterogeneity in consumption smoothing ability across different caste groups and more importantly, examine how is the social network associated with consumption smoothing and explore the potential mechanisms behind the results.

2.4.1 Rejection of Full Insurance Assumption Within the Village

As discussed in our theoretical framework, we could test the full insurance assumption based on the estimation results of equation (6). More specifically, the null hypothesis to be tested is H0: $\beta = 0$.

As an analogy to the specifications in Cochrane (1991) and Jalan and Ravallion (1999), we use our panel data to estimate an empirical analog to equation (6):

$$\Delta\log(c_{ivt}) = \beta\Delta\log(y_{ivt}) + \gamma x_{ivt} + \delta_{vt} + \Delta\varepsilon_{ivt} \tag{13}$$

where c_{ivt} is the consumption per capita of household i in village v at period t, y_{ivt} is the income per capita of the household, x_{ivt} is the vector of control variables, δ_{vt} is the village-year fixed effect that controls for the covariate shock at the village level and ε_{ivt} is an unobserved identically distributed random variable. $\Delta\log(c_{ivt}) = \log(c_{ivt}) - \log(c_{ivt-1})$ is the change in the log of consumption per capita of household i from period $t-1$ to period t, and $\Delta\log(y_{ivt}) = \log(y_{ivt}) - \log(y_{ivt-1})$ is the change in the log of income per capita of household i between two periods, or panels in our case.

Similar specifications have been used in a large body of literature that tests the full insurance assumption (Cochrane 1991, Ravallion and Chaudhuri, 1992; Jalan and Ravallion, 1999). While unlike most literature, we address the potential endogeneity issue corresponding to household income by instrumenting the income change with the rainfall deviation and applying the IV estimation method. More specifically, we instrument household income with the interaction of rainfall deviation at the village level and the land size at the household level to capture the idiosyncratic income change induced by rainfall deviation when the village-level aggregate shock is controlled. Similar strategies for measuring the idiosyncratic income shock are also applied in other income-related literature (Kazianga and Udry 2006, Fafchamps et al. 1998).

To estimate equation (13), the rainfall deviation could serve as a good instrument of household income as the rainfall deviation is generally unpredictable and therefore could be considered exogenous. Moreover, it is also safe to assume that the rainfall deviation would only affect household consumption through income by assuming that the consumption preference doesn't vary with rainfall and therefore the IV satisfies the exclusion restriction. Since the village-time fixed effect is controlled in equation (13), the interaction of rainfall deviation at the village level and the land size at the household level captures the relative size of the impact of rainfall deviation on households' agricultural income within the village.

In the context of rural India, agricultural income is the major income source for most households, which is indicated in our data as well (see Table 2.1). Additionally, more than 70% of households in our sample own land and participate in farming activities. From this perspective, the instrument variable we adopt is suitable for capturing the exogenous income variation for the majority of households in our sample.

In the first stage of our 2SLS pooled IV estimation, we estimate the following equation:

$$\Delta \log (y_{ivt}) = \rho_1 R_{vt} L_{ivt} + \rho_2 \Delta R_{vt} L_{ivt} + \mu x_{ivt} + \sigma_{vt} + \varphi_{ivt} \tag{14}$$

where the dependent variable is the change in the log of income per capita of household *i* in village *v* from period $t-1$ to period t. $R_{vt}L_{ivt}$, the interaction of the standardized rainfall deviation of village v—R_{vt} and the land size of household i—L_{ivt} at period *t*, serves as our first instrument variable. Since we are instrumenting the change in the log of income per capita of each household, the first-difference of $R_{vt}L_{ivt}$—$\Delta R_{vt}L_{ivt}$, or $R_{vt}L_{ivt} - R_{vt-1}L_{ivt-1}$, is used as our second instrument variable. x_{ivt} is the vector of control variables as in equation (13), σ_{vt} is the village-year fixed effect and φ_{ivt} is an unobserved identically distributed random variable.

The estimated parameters in equation (14) are reported in Table 2.4. The *F* statistic for the joint significance test of two instrument variables is reported in the table, which is significant at the 1% significance level. The significant coefficients of both IV variables indicate that the rainfall interactions capture idiosyncratic income shocks quite well when the village-time fixed effect is controlled.

Table 2.5 reports the results for the full insurance test, in which the first 3 columns of panel 1 report the estimates of equation (13), and the rest columns in panel 2 report the estimates of an analog of equation (13)—equation (12) by simply replacing the village-time fix effects with the change in the log of average consumption per capita within the village. The latter is another empirical testing strategy of full insurance applied in literature (Townsend 1994), while as suggested by Ravallion and Chaudhuri (1997), the estimates from (12) will be biased if there exists co-movement in household-level income changes within the group. Since the overall testing results are quite similar, our discussion will be mainly based on results in panel 1 based on the first specification.

Table 2.4 First-stage IV regression results on income change

	Change in the log of income per capita
Rainfall deviation * land area	-0.075***
	(0.014)
Change in rainfall deviation * land area	0.057***
	(0.010)
Land area	0.009***
	(0.003)
Change in land area	0.051***
	(0.003)
Household size	0.001
	(0.006)
Change in household size	-0.080***
	(0.005)
Male household head	0.040
	(0.039)
Age of household head	-0.003***
	(0.001)
Length of education of household head	-0.002
	(0.003)
Religion: Hindu	-0.039
	(0.057)
Caste group: ST	-0.061
	(0.075)
Caste group: OBC	-0.056
	(0.040)
Caste group: OC	-0.116**
	(0.045)
Constant	0.089
	(0.229)
Observations	7,686
R-squared	0.257
F-test: Rainfall interaction and change in interaction	17.59

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 2.5 The test of full insurance assumption

	(1)	(2)	(3)	(4)	(5)	(6)
		Panel 1			Panel 2	
	OLS	OLS	IV	OLS	OLS	IV
Change in income[1]	0.124***	0.115***	0.307***	0.114***	0.106***	0.495*
	(0.005)	(0.005)	(0.091)	(0.005)	(0.005)	(0.255)
Change in average consumption[2]				0.784***	0.780***	0.632***
				(0.012)	(0.012)	(0.099)
Land area		-0.003**	-0.006***		-0.004***	-0.012**
		(0.001)	(0.002)		(0.001)	(0.006)
Change in land area		0.009***	-0.000		0.009***	-0.009
		(0.001)	(0.004)		(0.001)	(0.012)
Household size	-0.007***	-0.003	-0.003	-0.004*	-0.000	-0.004
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)
Change in household size	-0.058***	-0.062***	-0.047***	-0.055***	-0.059***	-0.028
	(0.002)	(0.002)	(0.008)	(0.002)	(0.002)	(0.021)
Male household head		-0.054***	-0.061***		-0.054***	-0.064***
		(0.018)	(0.019)		(0.018)	(0.024)
Age of household head		-0.001	0.000		-0.001	0.001
		(0.000)	(0.001)		(0.000)	(0.001)
Education of household head		-0.003*	-0.002		-0.001	-0.003
		(0.002)	(0.002)		(0.001)	(0.002)
Religion: Hindu		-0.046*	-0.036		-0.026	-0.014
		(0.026)	(0.028)		(0.019)	(0.026)
Caste group: ST		0.023	0.034		-0.041*	-0.066*
		(0.034)	(0.036)		(0.024)	(0.036)
Caste group: OBC		-0.025	-0.012		-0.021	-0.014
		(0.018)	(0.021)		(0.017)	(0.023)
Caste group: OC		-0.072***	-0.045*		-0.037**	-0.001
		(0.021)	(0.025)		(0.018)	(0.034)
Constant	-0.123	0.015	0.361***	-0.053***	0.079**	-0.043
	(0.099)	(0.104)	(0.116)	(0.016)	(0.037)	(0.117)
Observations	7,655	7,605	7,605	7,655	7,605	7,605
R-squared	0.549	0.555	0.473	0.499	0.504	0.105

Note: Standard errors in parentheses; 1- Change in the log of income per capita;2-Change in the log of average consumption per capita within the village; *** p<0.01, ** p<0.05, * p<0.1

The first two columns of Table 2.5 report the OLS estimates, while column (1) controls for the household size-related variable only and column (2) also controls other listing variables in the table. The income coefficients are similar to related literature on the full insurance test (ranging from 10% to 40%), indicating that a 1% income increase is associated with about 0.115% increase in consumption on average. The estimated coefficient of income change in IV regression (column 3) is larger than in OLS regression (column 2), suggesting that 1% idiosyncratic income change is associated with 0.307% of consumption change. As for other control variables, the result also suggests that households' consumption per capita increases more in households with smaller household sizes and households with female heads. The age and education level of household heads are not associated with the consumption change significantly (at 5% significant level) conditional on other control variables.

The null hypothesis H0: $\beta = 0$ is rejected at the 1% confidence level in both OLS and IV estimation, which rejects the full insurance assumption empirically. The result is not surprising as a large literature has shown that full insurance or efficient risk sharing is not achieved within the villages (Cochrane 1991, Townsend 1994, Jalan and Ravallion 1999, Fafchamps and Kebede 2007). While comparing to these studies, our result contributes to the literature by instrumenting the idiosyncratic income change with rainfall deviation using a dataset with a comparably larger sample size for the test, which provides a relatively exogenous instrument of households' income to address the endogeneity issue.

2.4.2 Heterogeneity in Consumption Smoothing Across Caste Groups

In this section, we explore the heterogeneity in consumption smoothing across caste groups based on IV estimation results on equation (13). The $\beta,$ coefficient of household income change, captures to what extent the household consumption comoves with idiosyncratic income change when the group-level covariate shock is controlled. A larger estimate of β is an indication of less smoothed consumption against the idiosyncratic income change.

The first column of Table 2.6 reports the IV estimates of all samples as the reference of the sub-sample estimates in the following 4 columns, which report the result of IV estimates among 4 caste groups. The reported results indicate that the idiosyncratic income change is associated with households' consumption change significantly only in caste group OC—the forward caste group. In other backward caste groups (SC, ST, and OBC), households' consumption is smoothed, i.e., not significantly associated with the idiosyncratic income change induced by weather shock.

Table 2.6 Heterogeneity in consumption smoothing across caste groups from IV estimation

	(1)	(2)	(3)	(4)	(5)
	Total	Caste groups			
		SC	ST	OBC	OC
Change in income[1]	0.307***	0.064	-0.054	0.182	0.574***
	(0.091)	(0.114)	(0.470)	(0.115)	(0.160)
Land area	-0.006***	-0.009	-0.024***	-0.005*	-0.007**
	(0.002)	(0.008)	(0.008)	(0.003)	(0.003)
Change in land area	-0.000	0.014**	0.032	0.005	-0.012
	(0.004)	(0.005)	(0.047)	(0.006)	(0.008)
Household size	-0.003	0.006	-0.023**	-0.001	-0.001
	(0.003)	(0.008)	(0.010)	(0.004)	(0.007)
Change in household size	-0.047***	-0.065***	-0.094*	-0.054***	-0.031**
	(0.008)	(0.013)	(0.056)	(0.009)	(0.014)
Constant	0.361***	0.176	2.240**	0.451***	-0.263
	(0.116)	(0.255)	(0.934)	(0.119)	(0.444)
Household demographics	Yes	Yes	Yes	Yes	Yes
Cate groups	Yes	No	No	No	No
Village-time	Yes	Yes	Yes	Yes	Yes
Observations	7,605	1,111	657	3,668	2,168
R-squared	0.473	0.708	0.683	0.519	0.221

Note: Standard errors in parentheses; 1- Change in log of income per capita; *** p<0.01, ** p<0.05, * p<0.1

Several mechanisms or channels might contribute to such a surprising finding that while the backward caste groups are generally disadvantaged economically, they seem to have a relatively smoother consumption. First, as indicated in the descriptive statistics, the backward castes own less land on average and would therefore be exposed relatively less to weather shocks. Second, the lower castes are more flexible in finding casual jobs in non-agriculture sectors as a buffer strategy against yield shock. As shown in Table 2.2, the proportions of casual labor and NREGA worker are significantly higher in backward caste groups.

Table 2.7 Expenditure structure and consumption smoothing (IV estimates)

	(1)
	IV
Change in income[1]	1.340***
	(0.410)
Food expenditure interaction[2]	-1.933***
	(0.718)
Land area	-0.008***
	(0.002)
Change in land area	0.004
	(0.005)
Household size	-0.000
	(0.003)
Change in household size	-0.055***
	(0.008)
Constant	0.284**
	(0.132)
Household demographics	Yes
Cate groups	Yes
Village-time	Yes
Observations	7,605
R-squared	0.364

Note: Standard errors in parentheses; 1- Change in the log of income per capita; 2- Change in the log of income per capita*Proportion of food expenditure; *** p<0.01, ** p<0.05, * p<0.1

Moreover, the difference in expenditure structure might also contribute to the difference in consumption smoothing capability between backward caste groups and the forward caste group. The demographic statistics in Table 2.1 suggest that the household expenditure and income of the forward caste groups are significantly higher in the forward caste group, hence the proportion of necessary consumption such as food within the backward caste groups is expected to be higher (Engel 1857), which is indicated in our data as well. As a result, the consumption of households in backward caste groups can be less adjustable compared to the forward castes and therefore comoves less with idiosyncratic income changes.

The last possible mechanism corresponds to informal credits or finance. Unlike the forward castes who are normally economically advantaged in terms of asset owning and formal credits to buffer against shocks, informal credits or finance can be a more important risk-coping mechanism for backward caste groups. As a result, the backward caste groups might have a stronger incentive in forming informal insurance groups or networks against potential risks, which could help in explaining their smoother consumption. The comparably smaller amount of expenditure on backward castes might also make it easier for them to get support from others since their needs are relatively easier to be satisfied.

As a test of the third mechanism, i.e., the expenditure structure, we estimate equation (13) after adding the interaction term of households' income change and the proportion of food expenditure into the equation. Table 2.7 reports the results for the IV estimation, in which the negative and significant coefficient of the interaction term indicates that consumption of households with a higher proportion of expenditure on food items comoves less with idiosyncratic income change. It suggests that the consumption of households that spend relatively more on necessary goods such as food is less affected by the idiosyncratic income shock, which could be a possible explanation of the relatively smoother consumption of backward caste groups we observed.

2.4.3 Caste-Based Social Interaction Norms and Consumption Smoothing

In this section, we examine the association between consumption smoothing and the caste-based social interaction norms, defined as whether households interact with those from other castes or other caste groups. Whether people are interacting with other castes or caste groups, in the context of rural India, reflects to what extent is the households' social interaction constrained by the interaction barriers between castes that are embedded in India's caste system. While households who only interact with people within the same caste or caste groups are disadvantaged in sharing risk across castes or caste groups, they might have stronger social interactions within the caste or caste group they belong to, which can strengthen their capability in consumption smoothing.

Table 2.8 Caste-based social network and consumption smoothing (IV estimates)

	(1) All	(2) Caste group	(3) Caste
Change in income[1]	0.307***	0.506***	0.320***
	(0.091)	(0.160)	(0.093)
Income change*Within caste-group only[2]		-0.504**	
		(0.244)	
Income change*Within caste only[3]			-0.386*
			(0.219)
Constant	0.361***	0.327**	0.494***
	(0.116)	(0.133)	(0.133)
Household demographics	Yes	Yes	Yes
Cate groups	Yes	Yes	Yes
Village-time	Yes	Yes	Yes
Observations	7,605	7,605	7,605
R-squared	0.473	0.323	0.447

Note: Standard errors in parentheses; 1- Change in the log of income per capita; 2- Change in the log of income per capita*The indicator of within-caste group interaction only; 3- Change in the log of income per capita*The indicator of within-caste interaction only; *** p<0.01, ** p<0.05, * p<0.1.

Based on the social interaction variables we constructed, we generate two dummy variables on whether the household is only interacting with households within the same caste groups or the same caste. The interaction terms of two dummy variables and income change are added to our benchmark equation separately, and results from IV estimation are reported in Table 2.8. The significant negative coefficients of the interaction terms in the last two columns of the table indicate that the consumption of households with intra-caste or intra-group interaction only comoves less with idiosyncratic income change, suggesting that households with stronger intra-caste interactions have a smoother consumption against the idiosyncratic income change.

Moreover, the T-test on households with and without inter-caste or inter-caste group interactions suggests that the difference in caste-based social interaction norms didn't generate any significant difference in most demographic and socioeconomic features across groups. The relatively smoother consumption of households with no inter-caste or inter-group interaction is therefore likely to be a result of their potentially stronger intra-caste social network, which improves their ability in consumption smoothing through informal credits or risk-sharing within the same caste or caste group. While for households who have interactions with other caste groups, the inter-caste network might not be as efficient as the intra-caste networks that could help them buffer shocks.

2.4.4 Migration Experience and Consumption Smoothing

The other important social network-related factor we investigate is migration experience. It is argued in some literature that households with migrant members are likely to have reduced access to rural caste networks, as they cannot be easily punished by local networks and the superior outside options they can have when they are excluded from the network (Munshi and Rosenzweig 2016). While on the other hand, migration work may have positive impacts on consumption smoothing against agricultural shocks as it provides alternative income sources from the non-agricultural sector and generates a potentially stronger social network outside the village.

A dummy variable that measures whether there are household members who worked outside the village between 1996 and 2006, a period of ten years before the second panel of our survey is conducted, is generated. It turns out that 47% of households in our data have members with working experience as migrants. Similar to what was done in the previous sections, the interaction term of the dummy variable and the income change is generated and added to our benchmark model.

Table 2.9 reports the result of the IV estimation. The significant negative coefficient of the interaction term in column (2) indicates that the association between income change and consumption change is weaker in households with migration experience, i.e., their consumption comoves less with idiosyncratic income change.

Table 2.9 Migration working experience and consumption smoothing (IV estimates)

	(1)	(2)
Change in income[1]	0.307***	0.456***
	(0.091)	(0.115)
Income change*Migration experience[2]		-0.630***
		(0.168)
Constant	0.361***	0.327**
	(0.116)	(0.133)
Household demographics	Yes	Yes
Cate groups	Yes	Yes
Village-time	Yes	Yes
Observations	7,605	7,605
R-squared	0.473	0.323

Note: Standard errors in parentheses; 1- Change in log of income per capita; 2- Change in log of income per capita*The indicator of migration working experience; *** p<0.01, ** p<0.05, * p<0.1.

The result suggests that the positive impact of migration experience on consumption smoothing exceeds its potential disadvantage induced by the reduced access to the local caste network. Several factors might contribute to the result. First, the income from the non-agricultural sector can help buffer income shocks in agriculture. The fact that most households with members working as migrants in the past still have members working as migrant workers is consistent with this possible explanation. Moreover, the stronger out-village social network of households with migrant members in the past might help them find casual jobs in cases of local income shocks. Unfortunately, the survey data we have doesn't provide information on specific out-village social networks or social interactions, such mechanism is hard to verify empirically using the dataset we have.

2.5 Conclusion

There has been considerable growth in the literature on consumption smoothing and risk-sharing in the last decades. Much of the literature, using the benchmark model, rejects the hypothesis of efficient risk-sharing in Indian villages (Townsand 1994, Jallan and Ravallion 1999). As a result, emerging literature starts exploring risk-sharing within the closer socially interacted groups such as caste groups or ethnic groups.

Consumption smoothing is a broader topic than risk-sharing as the informal insurance arrangement is just one source of buffering strategy and households could also be self-insured even without group insurance. The social network of a household could very well affect consumption smoothing through many channels. And this challenges the literature that tries to identify the exact role of social networks in consumption smoothing. The other issue related to social network is that the impact of social networks might vary across cultures or institutions, which makes it hard to compare related literature conducted in different countries or under different contexts.

In this paper, we test the hypothesis of full insurance in rural India. We find very limited evidence of full insurance within villages in India, which is consistent with conclusions of other studies in the literature. Additionally, we test the role of caste in consumption smoothing and find that surprisingly, consumption of lower caste groups comoves less with the idiosyncratic income change when the group-level shock is controlled. This suggests that the consumption of lower caste groups is relatively smoother than advanced caste groups. Among several mechanisms we proposed in explaining such a result, we empirically test the impact of expenditure structure and discovered that the relatively higher expenditure on food items could explain why lower castes' consumption seems to be "smoother", as the necessary consumptions are less adjustable. Moreover, we find that households that interact mainly within the same caste or caste groups are better off against income shocks. This implies that the caste-based social network might be an important mechanism of consumption smoothing in India, which is consistent with conclusions in the literature (Munshi, 2019), suggesting that such informal institutions will continue to exist and play a significant economic role.

REFERENCES

Bagde, S., Epple, D., & Taylor, L. (2016). Does affirmative action work? Caste, gender, college quality, and academic success in India. American Economic Review, 106(6), 1495-1521.

Banerjee, A., Duflo, E., Ghatak, M., & Lafortune, J. (2013). Marry for what? Caste and mate selection in modern India. American Economic Journal: Microeconomics, 5(2), 33-72.

Banerjee, A., & Munshi, K. (2004). How efficiently is capital allocated? Evidence from the knitted garment industry in Tirupur. The Review of Economic Studies, 71(1), 19-42.

Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US. The Review of Economic Studies, 79(1), 128-161.

Bertrand, M., Hanna, R., & Mullainathan, S. (2010). Affirmative action in education: Evidence from engineering college admissions in India. Journal of Public Economics, 94(1-2), 16-29.

Besley, T., Pande, R., Rahman, L., & Rao, V. (2004). The politics of public good provision: Evidence from Indian local governments. Journal of the European Economic Association, 2(2-3), 416-426.

Carroll, C. D., Hall, R. E., & Zeldes, S. P. (1992). The buffer-stock theory of saving: Some macroeconomic evidence. Brookings papers on economic activity, 1992(2), 61-156.

Cassan G. Affirmative action, education and gender: Evidence from India[J]. Journal of Development Economics, 2019, 136: 51-70.

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in India. Econometrica, 72(5), 1409-1443.

Cochrane, J. H. (1991). A simple test of consumption insurance. Journal of political economy, 99(5), 957-976.

Deaton, A. (1989). Saving and liquidity constraints.

Drèze, J., & Kingdon, G. G. (2001). School participation in rural India. Review of Development Economics, 5(1), 1-24.

Engel, E. (2021). Die vorherrschenden Gewerbszweige in den Gerichtsämtern mit Beziehung auf die Productions-und Consumtionsverhältnisse des Königreichs Sachsen. WISTA–Wirtschaft und Statistik, 73(2), 126-136.

Fafchamps, M. (1992). Solidarity networks in preindustrial societies: Rational peasants with a moral economy. Economic development and cultural change, 41(1), 147-174.

Fafchamps, M., & Gubert, F. (2007). Contingent loan repayment in the Philippines. Economic Development and Cultural Change, 55(4), 633-667.

Fafchamps, M., & Lund, S. (2003). Risk-sharing networks in rural Philippines. Journal of development Economics, 71(2), 261-287.

Fafchamps, M., Udry, C., & Czukas, K. (1998). Drought and saving in West Africa: are livestock a buffer stock?. Journal of Development economics, 55(2), 273-305.

Granovetter, M. (2000). The economic sociology of firms and entrepreneurs. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Grimard, F. (1997). Household consumption smoothing through ethnic ties: evidence from Cote d'Ivoire. Journal of development Economics, 53(2), 391-422.

Hanna, R. N., & Linden, L. L. (2012). Discrimination in grading. American Economic Journal: Economic Policy, 4(4), 146-168.

Hnatkovska, V., Lahiri, A., & Paul, S. (2012). Castes and labor mobility. American Economic Journal: Applied Economics, 4(2), 274-307.

Jalan, J., & Ravallion, M. (1999). Are the poor less well insured? Evidence on vulnerability to income risk in rural China. Journal of development economics, 58(1), 61-81.

Kazianga, H., & Udry, C. (2006). Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso. Journal of Development economics, 79(2), 413-446.

Kinnan, C., Samphantharak, K., Townsend, R. M., & Vera Cossío, D. A. (2020). Insurance and propagation in village networks (No. IDB-WP-1155). IDB Working Paper Series.

Ksoll, C. (2007). Family networks and orphan caretaking in Tanzania.

Iyer, L., Khanna, T., & Varshney, A. (2013). Caste and entrepreneurship in India. Economic and Political Weekly, 52-60.

Ligon, E., Thomas, J. P., & Worrall, T. (2000). Mutual insurance, individual savings, and limited commitment. Review of Economic Dynamics, 3(2), 216-246.

Luke, N., & Munshi, K. (2011). Women as agents of change: Female income and mobility in India. Journal of development economics, 94(1), 1-17.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village india. Journal of Political Economy, 127(1), 1-46.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. The Quarterly Journal of Economics, 118(2), 549-599.

Munshi, K. (2019). Caste and the Indian economy. Journal of Economic Literature, 57(4), 781-834.

Munshi, K., & Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. American Economic Review, 96(4), 1225-1252.

Munshi, K., & Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. American Economic Review, 106(01), 46-98.

Pande, R. (2003). Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India. American economic review, 93(4), 1132-1151.

Ravallion, M., & Chaudhuri, S. (1992). Tests of risk-sharing in three Indian villages. Unpublished manuscript, Columbia University.

Ravallion, M., & Chaudhuri, S. (1997). Risk and insurance in village India: Comment. Econometrica: Journal of the Econometric Society, 171-184.

Rosenzweig, M. R., & Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural India. Journal of political Economy, 97(4), 905-926.

Rosenzweig, M. R., & Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India. Journal of political economy, 101(2), 223-244.

Stark, O., & Lucas, R. E. (1988). Migration, remittances, and the family. Economic development and cultural change, 36(3), 465-481.

Townsend, R. M. (1994). Risk and insurance in village India. Econometrica: journal of the Econometric Society, 539-591.

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

CHAPTER 3: ROAD INFRASTRUCTURE AND RURAL EMPLOYMENT IN CHINA

3.1 Introduction

Road infrastructure, along with other fundamental infrastructures such as electricity and water supply, has been found as the key to rural development in developing countries (World Bank 1994). It is documented in the literature that road infrastructure could promote rural economy through mechanisms such as improving market efficiency (Mu and Van de Walle 2011), reducing transportation costs and price distortion (Minten and Kyle 1999), promoting regional trade (Buys et al. 2010; Donaldson 2018) and enhancing rural mobility (Bryceson et al. 2008), catalyzing the adoption of agricultural technologies (Aggarwal 2018), etc. Consequently, road infrastructure in developing countries has been found to contribute to improvements in various rural socioeconomic indicators such as income improvement and poverty reduction (Warr 2005; Khandker et al. 2009; Dave 2018), consumption (Dercon et al., 2009), school participation and performance (Adukia et al. 2020), social-service accessibility (Bryceson et al. 2008), farmland value (Jacoby 2000; Shrestha 2020), etc.

While the positive role of road infrastructure in promoting rural development has been confirmed in most developing countries, the answer to how would road infrastructure affects rural farmers' employment choice can be highly contextualized, mainly because road infrastructure would affect different employment sectors simultaneously. Moreover, rural farmers' employment choice is also largely affected by factors that can vary across contexts such as working opportunities for migrants, access to health and educational services, culture and social norms, etc. For instance, studies have confirmed that road infrastructure is positively associated with rural off-farm job opportunities in different developing countries such as Nicaragua (Corral and Reardon 2001), Mexico (De Janvry and Sadoulet 2001), Peru (Escobal 2001), and China (Qiao et al. 2014). In a more recent study, Bai et al. (2021) find that road accessibility in China only promotes more local off-farm employment in high-income villages but has no effects on rural nonfarm employment on average. While based on data from India, Asher and Novosad (2018) find that rural road construction is associated with more off-farm work outside the village but not within the village.

Given that road investments are considered as an effective way of promoting rural development, accurately assessing the impacts of road infrastructure on rural farmers' employment choices and welfare has important policy implications for developing countries. However, a main factor that undermines the robustness and reliability of most related works based on cross-sectional data is the endogeneity of the road placement, i.e., road construction is non-arbitrary (Gachassin et al. 2010). Until recent years, the application of more available paneled datasets (Khandker et al. 2009), matching methods (Mu and Van de Walle 2011), and instrumental variable approach (Gibson and Rozelle 2003) makes it possible to address the endogeneity issue and more accurately estimate the road impacts on rural farmers.

In this study, we take advantage of the paneled data at the household level and explore the road impacts on farmers' employment choices in rural China. As the heart of China's national development strategy, the intensive construction of transportation infrastructures since the country initiated the economic reform and opening-up policy in 1978 has been viewed as one of the main driving forces of its rapid economic growth in past decades. In 2020, the total mileage of railways, expressways, and highways in China reached 0.15, 0.16, and 5.2 million kilometers accordingly from almost nothing in 1949[17]. While urban road mileage also experienced unprecedented growth, most highway mileage growth in China are contributed by rural roads in past decades. Rural road mileage in China reached 4.2 million kilometers in 2020, accounting for 80% of the total highway mileage of the country. Many studies have also demonstrated the key role of road infrastructure in promoting China's rural development through improvement in agricultural productivity and household income (Wang and Sun 2016; Fan and Chan-Kang 2008). Given the great success

[17] Data source: National Bureau of Statistics of China (https://data.stats.gov.cn/easyquery.htm?cn=C01)

China has achieved in rural development in past decades, understanding how Chinese farmers' employment choice is affected by road infrastructure has important policy implications for other developing countries to make investment decisions on rural road construction.

Among the large body of literature that examined the impact of rural road infrastructure in China, only a small number of studies (Qiao et al. 2014; Bai et al. 2021; Huang et al. 2022) have addressed the specific role of road construction in determining rural farmers' employment choice using farm-level or individual-level data, and the rest of studies are mainly conducted at the macro level. Also, most of the related studies conducted based on micro-level data are disadvantaged due to the constraints in road or employment data used in their analysis. For example, the road infrastructure in these studies is normally measured by the shortest travel distance between the village center to local road network, which fails in measuring the road improvement accurately as the shortest distance can remain unchanged even when more roads are constructed around the village. Such bias can even be greater if the subjective recall data is used to measure the distance. Also, road data used in these studies are not classified based on road type, which constrains the study in addressing the potential heterogeneous impacts of different types of roads. Moreover, the employment data in some related studies fail in providing information on the location of non-farm employment, which makes it hard to examine the impact of road infrastructure on migration patterns in rural China.

To address these disadvantages in the literature, our study combines farm-level household survey data and road information from Gord Maps and Google Earth to investigate the heterogeneous impacts of different types of roads on rural employment in China. Based on various non-linear models, we first find that the densities of county roads and provincial/national highways are positively associated with rural farmers' tendency towards and working intensity of local off-farm employment, while negatively associated with migration work. However, the association between local roads and rural employment is more ambiguous, which is likely related to the complexity of factors that can determine the construction of local roads. For agricultural employment, while the densities of county roads and provincial/national highways are positively associated with the participation in local part-time employment, full-time agricultural employment and part-time migration work are found negatively associated with road expansion. Consequently, farmers' working intensity on agriculture is negatively associated with the densities of county roads and provincial/national highways. Surprisingly, the association between expressway intensity and rural employment is found subtle when the intensities of other types of roads are controlled. Furthermore, we explore and find certain heterogeneities in the impacts of road infrastructure on rural employment across gender, education level, and region.

Our study contributes to the literature in several ways. First, this is the first study, to our knowledge, that examines the distinct impacts of various types of roads on farmers' employment choices in rural China using micro-level data. Second, we explore the road impacts on rural employment in both extensive and extensive margins by addressing farmers' employment choices in three dimensions—local vs. non-local, farming vs. off-farm, full-time vs. part-time, which contributes to broadening the understanding of the exact impacts of road infrastructure on rural employment and hasn't been done in any other related studies. Our conclusion that road infrastructure is associated with more local off-farm employment (especially local part-time employment) and less non-local employment, contradicts the long-standing view on the role of road infrastructure in promoting rural farmers' migration working in China.

The rest of the paper is organized as follows. Section 2 introduces the data we use for empirical analysis and some key descriptive results for a better understanding of road infrastructure and rural employment in China. Empirical methods and results are introduced in section 3, and section 4 concludes the paper.

3.2 Data and Descriptive Statistics

This section introduces the data we use and some key descriptive statistics on rural road infrastructure and employment in China.

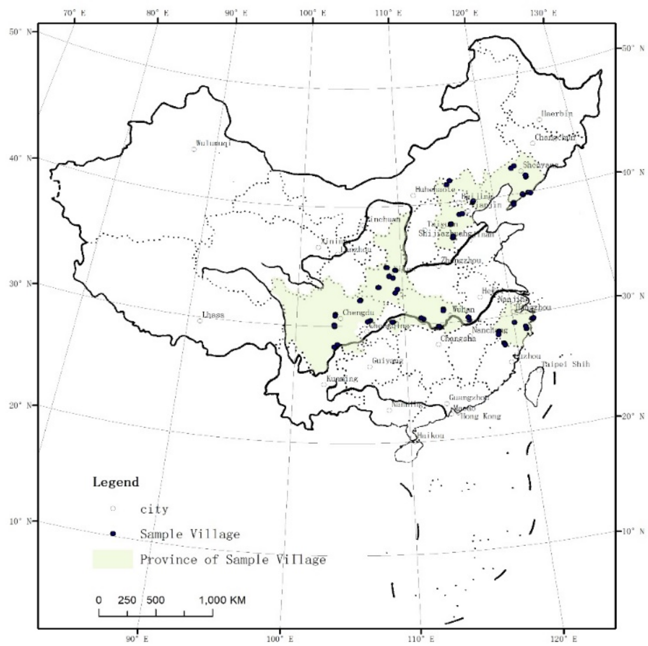

Figure 3.1 Sample area of the China National Rural Survey (CNRS)

Datasets from various sources are collected and used in our analysis. The first is the farm-level dataset from the China National Rural Survey (CNRS hereafter), which collects detailed information on demographics, employment, income, etc. on members of more than 1000 rural households each year (2000, 2008, and 2013) in China. Samples in CNRS are randomly selected from 60 villages in six provinces—Hebei, Liaoning, Shannxi, Zhejiang, Hubei, and Sichuan. The geographic locations of all sampled villages are presented on the map in Figure 3.1.

To measure the road infrastructure, we process and extract historical information on road mileage of each surveyed village in CNRS based on GIS data from Gord Maps and Google Earth. More specifically, we analyze historical images of roads on Google Earth, combined with the survey method, to determine when each road was built and opened to the public. We then calculate the length and density of each type of road within different village boundaries— 2, 5, 10, and 20 kilometers. Such information is assembled into six types of roads—village road, township road, provincial highway, national highway, and expressway (see example images of different types of roads in Figure B.1 in the appendix).

Moreover, some macro data from the National Bureau of Statistics of China (NBSC hereafter) are also used for descriptive analysis on the development of road infrastructure and changes in rural employment over time in China.

3.2.1 Rural Road Infrastructure in China

Road infrastructure in China has experienced rapid growth in past decades, as shown in Figure 3.2. The figure clearly showed that the transportation infrastructure until the early reform period was rather poor. For instance, there has been almost zero mileage of expressway, and less than a million mileage of highway in 1990. Since the mid-1990s, the Chinese government started to invest heavily to improve its transportation infrastructure, resulting in massive road and railway mileage growth since then. In 2020, the total mileage of railways, expressways, and highways (or public roads, most are paved) in China reached 0.15, 0.16, and 5.2 million kilometers accordingly from almost nothing in 1949[18].

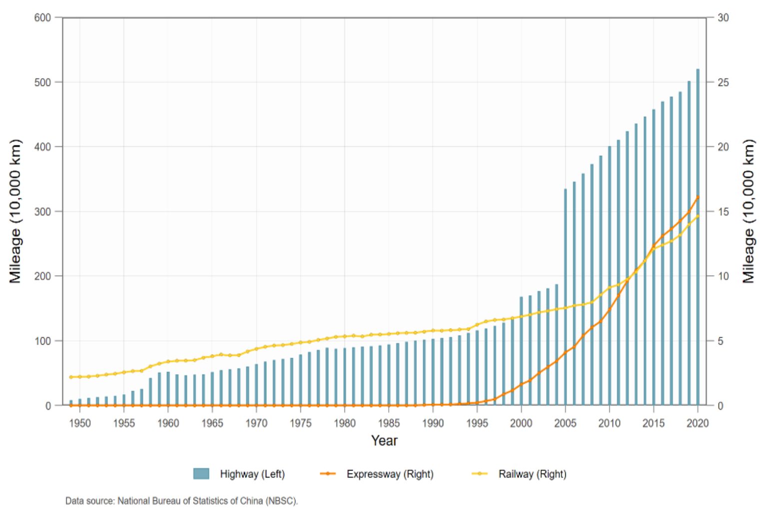

Figure 3.2 Highway, railway, and expressway mileages in China from 1949 to 2020

While the urban road mileage in China also experienced unprecedented growth, most highway mileage growth in past decades are contributed by rural roads, of which the mileage reaches 4.2 million kilometers in 2020 and accounts for 80% of total highway mileage in the country.

[18] Data source: National Bureau of Statistics of China (https://data.stats.gov.cn/easyquery.htm?cn=C01)

Table 3.1 Descriptive statistics of road infrastructure in CNRS villages

	Average road mileage of sampled villages (km)				Share of villages with mileage change	
	2000	*2008*	*2013*	***Total***	*2000-08*	*2008-13*
1. Village Road						
Radius: 2 km	4.23	5.49	6.17	5.29	45%	22%
Radius: 5 km	22.16	31.2	36.96	30.01	66%	53%
Radius: 10 km	92.17	125.11	149.53	121.93	88%	79%
Radius: 20 km	350.46	461.99	535.06	448.05	100%	93%
2. Township Road						
Radius: 2 km	3.48	5.11	6.49	5.01	43%	24%
Radius: 5 km	19.57	31.9	40.29	30.46	67%	48%
Radius: 10 km	69.03	115.83	143.9	109.13	81%	74%
Radius: 20 km	252.69	408.23	484.63	380.38	98%	90%
3. County Road						
Radius: 2 km	1.29	1.26	1.26	1.27	7%	0%
Radius: 5 km	7.54	7.77	7.89	7.73	24%	5%
Radius: 10 km	30.08	34.78	36.04	33.59	45%	21%
Radius: 20 km	119.72	140.34	148.11	135.87	74%	47%
4. Provincial Highway						
Radius: 2 km	1.2	1.57	1.57	1.45	9%	0%
Radius: 5 km	5.52	7.38	7.38	6.74	14%	0%
Radius: 10 km	23.39	29.46	29.58	27.43	19%	2%
Radius: 20 km	78.61	96.92	97.91	91	22%	3%
5. National Highway						
Radius: 2 km	0.79	0.74	0.74	0.76	0%	0%
Radius: 5 km	4.1	4.49	4.49	4.36	5%	0%
Radius: 10 km	13.39	14.47	14.47	14.1	7%	0%
Radius: 20 km	45.51	48.74	48.74	47.64	9%	0%
6. Expressway						
Radius: 2 km	0.28	0.82	1.17	0.75	7%	5%
Radius: 5 km	1.06	2.91	5.51	3.13	9%	16%
Radius: 10 km	2.83	13.75	23.18	13.13	34%	26%
Radius: 20 km	9.62	52.71	85.08	48.69	52%	40%

The descriptive statistics on road infrastructure in our studied area (based on data from map services) are consistent with the overall trend of road construction in China at the national level in past decades. However, the construction progress in our studied period (from 2000 to 2013), as indicated in our data, varies across various types of roads.

Table 3.1 summarizes changes in mileages of various types of roads in our studied area from 2000 to 2013, in which columns 1-4 report the average mileages of all types of roads within different radiuses of the village center by years and in total sample, and columns 5-6 present the statistic results on the share of villages with changes in road mileage between two panels. The result unveils several features of the rural road construction progress in our studied areas, which provides important information for identifying the appropriate radius for each type of road in describing the local road infrastructure.

First, compared to the county road and provincial/national highway, the village/township roads and the expressway have experienced more significant mileage growth over the period. Moreover, the construction of provincial and national highways in our studied area almost stagnated between 2000 to 2013. This is not a surprising result as the provincial or national roads were mainly constructed for connecting cities before the extensive constriction of expressways in China. However, most newly built roads in China in the recent two decades are more advanced expressways with better quality, greater speed limits and width. A main function of these expressways is to connect cities. This is consistent with the rapid mileage growth of the expressway mileages in China after 2000, as shown in Figure 3.2.

Second, the variation in road mileage, along with the function of the road, provides useful information for selecting appropriate radius for different types of roads to construct road-related variables for our empirical analysis. For example, the village and township roads are similar in quality and density and have almost the same function in practice. As a result, these two types of roads are merged into one category—local roads in our empirical analysis. Moreover, 2 km is chosen as the radius for the local road as it corresponds to a relatively narrower local area compared with a 5 km or greater radius, which might result in covered areas that are too broad and fail to measure the local road density within the village accurately.

Moreover, 10 km is selected as the radius of the county road for similar considerations, mainly based on the average area of all counties in our sample. For provincial and national highways, we merge them into one category considering their similarity in function, density, and quality. We use the largest 20 km as the appropriate radius for these two types of roads, as the main role of these two types of roads is to connect the main cities and this makes it sparser than the county and local roads. The consideration of road length variation is another concern we have in choosing 20 km for provincial and national roads. For the expressway, we choose 20 km as the appropriate radius considering the sparsity of roads with smaller radiuses.

3.2.2 Rural Employment in China

A. Rural employment at the national level

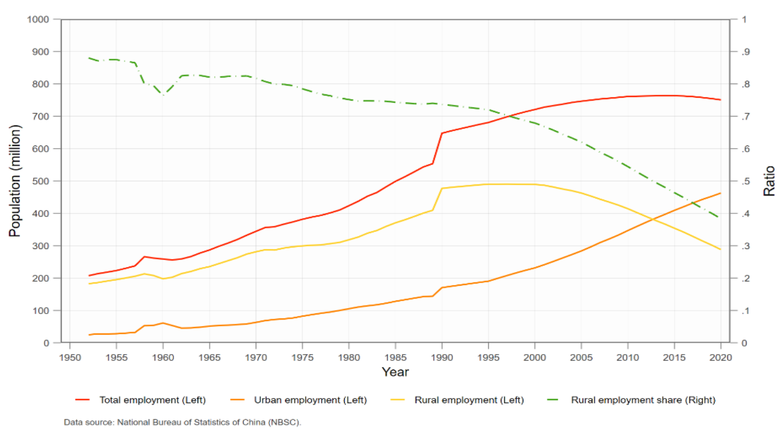

Figure 3.3 Total, urban, and rural employment in China from 1952 to 2020

The populations of total, rural, and urban employment in China have experienced rapid growth from 1952 to 2020. As shown in Figure 3.3, the populations of rural and urban workers have increased from 182.43 million and 24.86 million in 1952 to 287.93 million and 462.71 million accordingly in 2020, mainly due to population and economic growth in the period. However, the growth over the entire period masks significant differences in growth patterns between rural and urban employment. While rural and urban employment shared a similar up-trend growth pattern before 1990, they started to behave very differently after the mid-1990s. While urban employment continued to grow at an accelerated pace until today, rural employment has been declining since 2000.

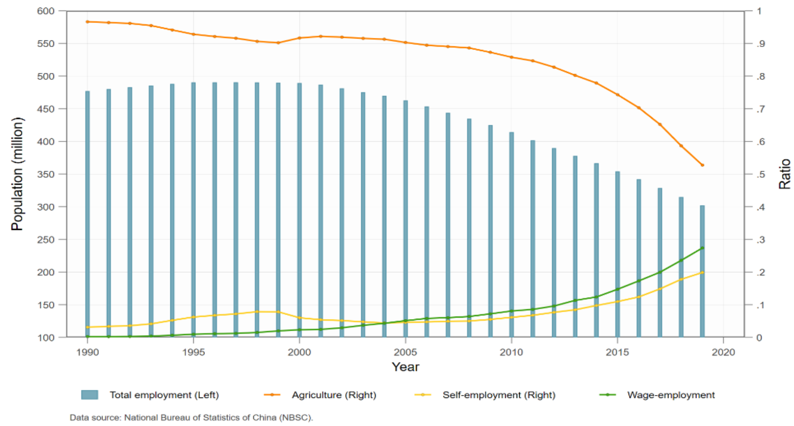

Figure 3.4 Rural employment by type in China from 1990 to 2019

Table 3.2 Share of rural employment by year (CNRS)

	2000	2008	2013	**All**
Employment Type				
Unemployment	0.08	0.1	0.13	0.1
Agriculture (Full-time)	0.45	0.37	0.3	0.37
Local off-farm (Part-time)	0.24	0.19	0.15	0.19
Migration working (Part-time)	0.04	0.04	0.03	0.04
Local off-farm (Full-time)	0.09	0.13	0.21	0.15
Migration working (Full-time)	0.09	0.16	0.18	0.15
Number of observations	3,304	2,875	3,163	9,342

While the total rural employment has been declining in the past decades, as shown in Figure 3.3, the share of self-employed labor or wage employment in rural areas kept increasing during the same period. As shown in Figure 3.4, the shares of wage-employment and self-employment in rural employment have increased from 0.23% and 3.12% in 1990 to 27.38% and 19.87% in 2019 accordingly. In contrast, the share of agricultural workers has declined from 96.64% to 52.76% over the same period. This trend suggests that while fewer rural residents are involved in agricultural activities, there are more participation in the rural non-agricultural economy in the past 2 or 3 decades.

B. Rural employment of CNRS samples

The CNRS data presents a similar trend of rural employment as in the national-level data, i.e., less agricultural employment and more off-farm employment over time. Table 3.2 presents the shares of different employments in three panels, indicating a decline in agricultural employment (including both full-time and part-time), and an increase in full-time non-agricultural employment. The other interesting trend in rural employment, as shown in the table, is the observable reduction of participation in part-time employment over time.

Table 3.3 Share of rural employment by gender, education, and province (CNRS)

	Unemployment	Agriculture (Full-time)	Local off-farm (Part-time)	Migration (Part-time)	Local off-farm (Full-time)	Migration (Full-time)	Obs
Gender							
Female	0.15	0.48	0.12	0.02	0.13	0.11	4,504
Male	0.05	0.27	0.27	0.06	0.16	0.18	4,838
Education							
Illiteracy	0.13	0.54	0.17	0.03	0.07	0.06	2,957
Primary	0.08	0.41	0.21	0.05	0.12	0.13	1,898
Junior High	0.09	0.29	0.22	0.05	0.18	0.17	3,207
Senior High	0.1	0.16	0.16	0.03	0.27	0.29	1,280
Province							
Sichuan	0.12	0.43	0.23	0.03	0.12	0.07	1,594
Hebei	0.09	0.48	0.19	0.04	0.1	0.1	1,408
Zhejiang	0.13	0.21	0.22	0.01	0.32	0.11	1,645
Hubei	0.09	0.31	0.2	0.02	0.16	0.22	1,692
Liaoning	0.07	0.45	0.15	0.05	0.08	0.18	1,419
Shaanxi	0.09	0.4	0.18	0.08	0.07	0.18	1,584

Table 3.3 presents the average shares of different employments by gender, education, and region in our data. It first shows that almost half of the female samples are doing full-time agricultural work while males participate more in part-time work and non-agricultural work.

Moreover, better-educated rural residents, not surprisingly, participate less in agriculture and more in non-agricultural employment.

Rural employment in CNRS also shows a significant regional difference. The province-level spatial distribution and differences in terms of industrial structure, endowments, and economic growth are crucial in understanding the Chinese economy. The division of eastern, central, western, and northeastern provinces is a wide-accepted form in most analysis frameworks in understanding the regional economy difference in China. Ever since 1978, the eastern, especially the coastal provinces of China contribute to most of the national GDP while the other three regions are poorer but provide a huge amount of labor for industries in the eastern provinces. Figure 3.5 presents the GDP contribution of all provinces in China in 2008, which explicitly indicates the benchmark spatial distribution of the Chinese economy.

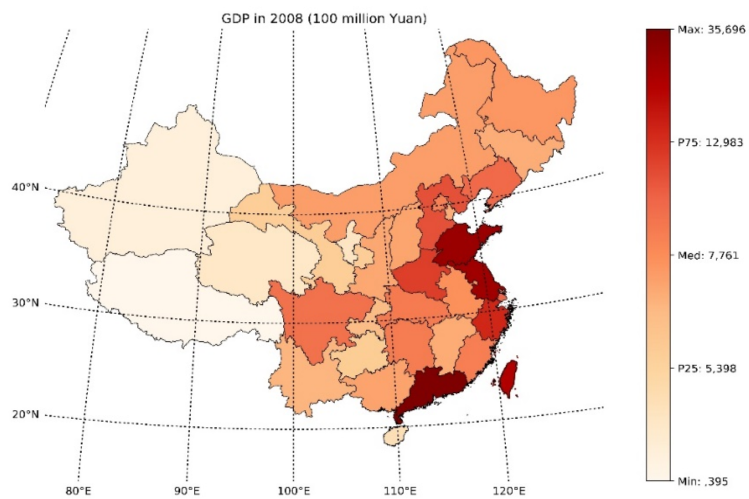

Figure 3.5 Regional GDP distribution of China in 2008

The regional difference in terms of rural employment in our dataset is highly consistent with the regional difference at the macro level. While the eastern province—Zhejiang, has a relatively higher share of local off-farm employment, other non-eastern provinces have a higher ratio of out-county off-farm employment. This pattern of regional differences unveils a relatively less developed local economy, rare local employment opportunities, and therefore higher potential for migration in the non-eastern areas.

3.3 Empirical Model and Results

3.3.1 Empirical Model

To investigate the impact of road infrastructure on rural employment, we estimate a series of non-linear models including pooled logit model, fixed effects (FE) logit model, and correlated random effects (CRE) logit model based on the following equation:

$$y_{i,t}^{*} = \alpha + \beta_1 Road_{i,t-1} + \beta_2 X_{i,t} + \lambda_t + \theta_{p,t} + c_i + \mu_{i,t} \quad (1)$$

$$y_{i,t} = 1\left[y_{i,t}^{*} > 0\right]$$

where $y_{i,t}$ is a binary variable of employment participation, which equals 1 if the latent variable $y_{i,t}^{*} > 0$. $Road_{i,t-1}$ is a vector of densities of roads within the village boundary at year $t-1$, the lagged measures of roads are used to reduce the potential impacts of road construction itself at year t on labor demand and therefore farmers' employment choice. $X_{i,t}$ is a vector of control variables including individual demographics such as gender, age, marital status, education, etc., and household characteristics such as family size, land area, assets, etc. λ_t and c_i denote the year and individual fixed effects accordingly, while the province-year fixed effects—$\theta_{p,t}$ capture the fixed and time-varying unobservable attributes at the province level, and $\mu_{i,t}$ is the error term.

We first apply the pooled logit model, in which the unobserved time-invariant individual heterogeneity c_i is omitted, as the baseline model. To address the unobserved heterogeneity, we first apply the FE logit model, which assumes that the individual heterogeneity is correlated with the independent variables and treats c_i as parameters. However, the FE logit model normally suffers from the incidental parameters problem, especially when T is small (Wooldridge 2019). Therefore in addition to the FE logit model, we apply the CRE logit model that allows the dependence between the individual heterogeneity and independent variables in a restricted way for estimation. Compared to the random effects (RE) logit model, the CRE logit model is advantaged in not assuming the strong RE assumption, i.e., individual heterogeneity is uncorrelated with independent variables. In practice, CRE approaches have been widely used in empirical studies for estimating nonlinear panel data models (Altonji and Matzkin 2005, Wooldridge 2005). More specifically, a general nonparametric assumption we make in the CRE logit model is:

$$D(c_i|X_i) = D(c_i|\overline{X}_i) \tag{2}$$

where X_i is the vector of all independent variables in all periods and $\overline{X}_i$ is the mean of X_{it} over time. Empirically, the sample averages of time-variant independent variables are included in the regression for modeling the c_i term.

3.3.2 Empirical Results

Table 3.4 reports the estimation results on rural farmers' participation in agricultural work, local off-farm (within the county), and migration working (outside the county) accordingly.

For local roads, the results from pooled logit model suggest that in villages with denser local roads, farmers are more likely to do local off-farm work but less likely to participate in agricultural and migration work. Moreover, the fact that the directions and magnitudes of estimated impacts of local road density on employment in pooled logit model significantly differ from results from FE and CRE models might suggest that the local roads are more "endogenous" than other types of roads. This is not surprising as rural farmers or village collectives are more influential in the construction of village or township roads, while the construction of other types of roads is normally determined by higher-level governments. And another factor that could contribute to the endogeneity or the complexity of local roads is the economic conditions of different villages. On the one hand, economically more disadvantaged villages might receive more financial support on local road construction (Zhang et al. 2006). On the other hand, more developed villages are more able to finance the construction of local roads (Zhang et al. 2006). Consequently, the economic development at the village level might affect the construction of local roads in different ways, which makes it harder to identify the impacts of local roads on rural employment empirically.

The impacts of county roads and provincial/national highways on rural employment are similar. For agricultural works, the estimated impacts of densities of county roads and provincial/national highways on agricultural employment from the FE logit model and CRE logit model, while insignificant, are positive. A more detailed investigation of agricultural employment (See Table B.1 in the appendix) indicates that while the densities of these two types of roads are negatively associated with full-time agricultural employment, part-time agricultural employment (mainly contributed by local part-time employment) increases with road expansion. As a result, the estimated impacts of roads on agricultural participation are more ambiguous.

For off-farm employment, the results from the FE logit model and CRE logit model both suggest that the densities of these two types of roads are associated with more local off-farm employment but less migration working. The results provide evidence that road expansion in rural China, at least in our studied context, might have incentivized more migrant workers to go home and do more local off-farm jobs. Such a conclusion is consistent with what is found by Qiao et al. (2014), while road infrastructure is measured in a relatively simple way in their study.

Furthermore, a more detailed investigation (see Table B.2 in the appendix) on local off-farm employment shows that the positive impacts of expansion in county roads and provincial/national highways are more significant on part-time local off-farm employment. This might suggest that the development of rural road infrastructure can not only boost the local off-farm sector and generate more employment opportunities but would also reduce the cost of doing farm and off-farm work at the same time. For migration working, the results (see Table B.3 in the appendix) show that participation in full-time and part-time migration working are both negatively associated with expansion in county roads and provincial/national highways.

Surprisingly, the improvement in expressway density doesn't seem to have significant impact on farmers' employment choices in our studied context. This first implies that it might be the accessibility, rather than the density, of the expressway affecting farmers' employment choice, as farmers might just access the closest expressway regardless of how dense the expressway network is. Also, the fact that a large portion of rural migrant workers in China is transported through the well-developed railway system could also explain the relatively insignificant impacts of expressway density on rural farmers we find. Moreover, an important implication from the result is that the significant impacts of expressway expansion on rural employment in China found in previous studies might come from the corresponding expansion of other roads (county roads or provincial/national highways), rather than the pure impacts of the expressway expansion.

For other control variables, the results show that educated and professionally trained male laborers are more likely to do off-farm work but less likely to do farm work. We also find that older and married laborers are more likely to work locally while are less likely to migrate. Moreover, not surprisingly, the land area owned by the household is positively associated with the tendency to work in agriculture against off-farm employment.

To address the potential collinearity issue on various types of roads, we report the results from pooled logit and CRE logit model using the aggregated road density and the milage share of local roads (village and township roads) in Table 3.5. The results from CRE logit models on aggregate road density first indicate that the overall increasing road mileage is associated with more local off-farm employment but less out-county migration working. Moreover, the results from pooled logit models on the share of local roads seemingly suggest that there is relatively less local employment (including both farming and local non-farm employment) and more migration working in regions with a higher share of local roads. This implies that, unlike other types of roads, the local roads may not be a good indicator of a strong local non-agricultural economy. While instead, there are more migrant workers in villages with relatively more local roads.

Table 3.4 Road infrastructure and rural employment participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:		**Agriculture**			**Local off-farm**			**Migrant working**	
Participation	Pooled Logit	FE Logit	CRE Logit	Pooled Logit	FE Logit	CRE Logit	Pooled Logit	FE Logit	CRE Logit
Road Densities									
Local	-0.43***(0.12)	0.01(0.50)	-0.16(0.47)	0.34***(0.10)	-0.36(0.39)	-0.58(0.36)	-0.60***(0.13)	0.24(0.61)	0.36(0.54)
County	-0.04(0.56)	5.83(3.57)	5.13(3.36)	0.21(0.45)	7.09***(2.63)	6.78***(2.49)	1.47**(0.60)	-14.75***(4.35)	-13.55***(4.02)
Prov./National	1.37(0.87)	1.42(5.00)	0.05(3.93)	3.68***(0.70)	6.36**(3.12)	7.14***(2.72)	-2.75***(0.96)	-6.81*(4.12)	-7.70**(3.65)
Expressway	-1.91*(1.06)	-0.13(2.58)	-1.24(2.22)	-0.34(0.89)	-1.50(1.71)	-1.36(1.62)	1.33(1.16)	-1.75(2.66)	-0.12(2.46)
Year_2008	-1.23***(0.20)	-1.28***(0.46)	-1.01***(0.28)	-0.72***(0.16)	-1.13***(0.37)	-1.22***(0.21)	1.32***(0.18)	1.57***(0.55)	1.38***(0.26)
Year_2013	-2.04***(0.20)	-2.21***(0.59)	-1.80***(0.30)	-0.32**(0.16)	-0.61(0.49)	-0.75***(0.22)	1.57***(0.19)	1.94**(0.77)	1.57***(0.29)
Male	-0.35***(0.06)		-0.50***(0.10)	0.73***(0.05)		0.93***(0.08)	0.76***(0.07)		1.03***(0.11)
Age	0.09***(0.00)		-0.01(0.02)	-0.02***(0.00)		0.04***(0.01)	-0.08***(0.00)		-0.04*(0.02)
Marriage	1.10***(0.09)		0.35(0.24)	0.89***(0.08)		0.76***(0.20)	-0.80***(0.09)		-0.85***(0.24)
Junior high	-0.32***(0.07)		-0.45***(0.10)	0.34***(0.05)		0.38***(0.08)	0.02(0.07)		0.08(0.10)
Tech training	-0.37***(0.08)		-0.39***(0.11)	0.46***(0.06)		0.51***(0.09)	0.22***(0.08)		0.33***(0.11)
HH size	-0.12***(0.02)		0.03(0.06)	-0.05**(0.02)		0.00(0.04)	0.07***(0.02)		-0.14**(0.07)
Land	0.51***(0.05)	0.56***(0.14)	0.49***(0.11)	-0.26***(0.04)	-0.20**(0.10)	-0.18**(0.09)	0.05(0.05)	-0.09(0.15)	-0.08(0.13)
House value	0.02(0.02)	0.02(0.04)	0.04(0.03)	0.02(0.01)	-0.01(0.03)	-0.01(0.02)	-0.03*(0.02)	-0.05(0.04)	-0.07**(0.03)
Durable assets	-0.08***(0.02)	-0.08(0.06)	-0.02(0.05)	0.19***(0.02)	0.11***(0.04)	0.11***(0.04)	-0.17***(0.02)	-0.15***(0.05)	-0.16***(0.05)
Constant	-1.45***(0.26)		-2.69***(0.45)	-2.59***(0.21)		-3.08***(0.34)	2.41***(0.27)		3.40***(0.45)
Observations	8,512	1,549	8,512	8,512	2,436	8,512	8,512	1,238	8,512

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.5 Road infrastructure (aggregated) and rural employment participation

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Agriculture**		**Local off-farm**		**Migrant working**	
Participation	Pooled Logit	CRE Logit	Pooled Logit	CRE Logit	Pooled Logit	CRE Logit
Road Densities						
Road density (20km)	-0.06(0.17)	0.65(0.71)	1.07***(0.14)	1.00*(0.54)	-0.94***(0.19)	-2.40***(0.84)
Share of local road	-0.77**(0.35)	0.10(0.75)	-0.70***(0.26)	-0.38(0.54)	0.42(0.38)	0.72(0.81)
Year_2008	-1.27***(0.20)	-1.17***(0.28)	-0.76***(0.16)	-1.22***(0.21)	1.34***(0.18)	1.48***(0.25)
Year_2013	-2.12***(0.19)	-2.07***(0.29)	-0.43***(0.15)	-0.87***(0.21)	1.69***(0.18)	1.90***(0.28)
Male	-0.35***(0.06)	-0.50***(0.10)	0.73***(0.05)	0.93***(0.08)	0.75***(0.07)	1.01***(0.11)
Age	0.09***(0.00)	-0.00(0.02)	-0.02***(0.00)	0.04***(0.01)	-0.08***(0.00)	-0.05**(0.02)
Marriage	1.10***(0.09)	0.35(0.24)	0.90***(0.08)	0.74***(0.20)	-0.79***(0.09)	-0.80***(0.23)
Junior high	-0.32***(0.07)	-0.45***(0.10)	0.32***(0.05)	0.36***(0.08)	0.05(0.07)	0.10(0.10)
Tech training	-0.37***(0.08)	-0.40***(0.11)	0.46***(0.06)	0.50***(0.09)	0.22***(0.08)	0.34***(0.11)
HH size	-0.13***(0.02)	0.03(0.06)	-0.05**(0.02)	0.00(0.04)	0.07***(0.02)	-0.13**(0.07)
Land	0.54***(0.05)	0.49***(0.11)	-0.27***(0.04)	-0.16*(0.09)	0.06(0.05)	-0.09(0.12)
House value	0.01(0.02)	0.04(0.03)	0.02(0.01)	-0.01(0.02)	-0.04**(0.02)	-0.08**(0.03)
Durable assets	-0.07***(0.02)	-0.02(0.05)	0.18***(0.02)	0.12***(0.04)	-0.15***(0.02)	-0.16***(0.05)
Constant	-0.99***(0.34)	-1.70***(0.59)	-2.20***(0.27)	-2.64***(0.44)	2.20***(0.36)	2.93***(0.61)
Observations	8,512	8,512	8,512	8,512	8,512	8,512

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.6 Road infrastructure and share of working days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:		**Agriculture**			**Local off-farm**			**Migrant working**	
Share of days	Pooled Tobit	RE Tobit	CRE Tobit	Pooled Tobit	RE Tobit	CRE Tobit	Pooled Tobit	RE Tobit	CRE Tobit
Road Densities									
Local	-0.16***(0.06)	-0.15**(0.06)	0.14(0.14)	0.39***(0.08)	0.37***(0.08)	-0.19(0.20)	-1.26***(0.33)	-1.26***(0.35)	-0.34(0.92)
County	-0.67**(0.27)	-0.72**(0.30)	-0.44(0.96)	0.22(0.35)	0.42(0.39)	3.08**(1.36)	3.52**(1.45)	2.96*(1.56)	-19.89***(6.76)
Prov/Nation	-0.71*(0.40)	-1.01**(0.43)	-0.05(1.03)	2.41***(0.55)	2.77***(0.58)	3.99***(1.53)	-8.27***(2.33)	-7.02***(2.42)	-12.64**(6.17)
Expressway	-0.65(0.52)	-0.31(0.49)	0.75(0.64)	0.01(0.69)	-0.07(0.67)	-0.46(0.90)	5.13*(2.81)	3.17(2.77)	1.28(4.16)
Year_2008	-0.33***(0.09)	-0.39***(0.07)	-0.12(0.08)	-0.50***(0.13)	-0.46***(0.11)	-0.73***(0.12)	3.27***(0.47)	3.53***(0.43)	2.56***(0.45)
Year_2013	-0.78***(0.09)	-0.85***(0.08)	-0.47***(0.09)	-0.02(0.12)	0.04(0.11)	-0.32**(0.13)	3.97***(0.49)	4.23***(0.46)	2.92***(0.50)
Male	-0.58***(0.03)	-0.55***(0.03)	-0.55***(0.03)	0.48***(0.04)	0.45***(0.05)	0.46***(0.05)	1.83***(0.19)	1.68***(0.20)	1.71***(0.20)
Age	0.05***(0.00)	0.05***(0.00)	0.00(0.01)	-0.02***(0.00)	-0.02***(0.00)	0.03***(0.01)	-0.21***(0.01)	-0.21***(0.01)	-0.06*(0.04)
Marriage	0.51***(0.05)	0.47***(0.05)	-0.02(0.08)	0.49***(0.06)	0.46***(0.06)	0.39***(0.11)	-2.39***(0.25)	-2.37***(0.25)	-1.48***(0.40)
Junior high	-0.21***(0.03)	-0.21***(0.03)	-0.19***(0.03)	0.30***(0.04)	0.28***(0.04)	0.26***(0.04)	0.13(0.17)	0.22(0.18)	0.24(0.18)
Tech training	-0.34***(0.04)	-0.29***(0.04)	-0.27***(0.04)	0.33***(0.05)	0.26***(0.05)	0.24***(0.05)	0.47**(0.20)	0.55***(0.19)	0.51***(0.19)
HH size	-0.03***(0.01)	-0.03***(0.01)	0.01(0.02)	-0.01(0.01)	0.00(0.01)	0.02(0.02)	0.19***(0.06)	0.15**(0.06)	-0.23**(0.11)
Land	0.24***(0.02)	0.21***(0.02)	0.14***(0.03)	-0.28***(0.03)	-0.25***(0.03)	-0.15***(0.05)	0.00(0.12)	0.01(0.12)	-0.10(0.21)
House value	0.00(0.01)	0.01(0.01)	0.02**(0.01)	0.00(0.01)	0.00(0.01)	-0.00(0.01)	-0.04(0.04)	-0.04(0.04)	-0.10*(0.06)
Durable assets	-0.06***(0.01)	-0.05***(0.01)	-0.01(0.01)	0.18***(0.01)	0.15***(0.01)	0.08***(0.02)	-0.39***(0.06)	-0.36***(0.06)	-0.26***(0.08)
Constant	-0.38***(0.12)	-0.47***(0.12)	-0.59***(0.15)	-1.85***(0.17)	-1.78***(0.17)	-1.53***(0.20)	5.74***(0.69)	5.68***(0.70)	6.10***(0.83)
Observations	8,295	8,295	8,295	8,295	8,295	8,295	8,295	8,295	8,295

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

In addition to the employment tendency, we also investigate the impacts of rural road density on farmers' working intensity, measured by the share of working days[19], by estimating pooled Tobit, random effects (RE) Tobit, and correlated random effects (CRE) Tobit model. The results, as shown in Table 3.6, are consistent with the results on employment tendency. More specifically, the densities of local roads, county roads, and provincial/national highways are found associated with relatively higher intensity of local off-farm employment and lower intensities of agricultural work and migration work. While for expressways, the road density seems to have vaguely positive impacts on migration working. The results on other control variables are consistent with what is shown in Table 3.4.

To sum up, our empirical results suggest that the densities of county roads and provincial/national highways are positively associated with rural farmers' tendency towards and working intensity of local off-farm employment while negatively associated with migration working. However, the association between local roads and rural employment is more ambiguous, possibly due to the endogeneity of the construction of this type of road. For agricultural employment, while the densities of county roads and provincial/national highways are positively associated with the participation in local part-time employment, full-time agricultural employment and part-time migration employment are found negatively associated with road expansion. Consequently, farmers' working intensity on agriculture is negatively associated with the densities of these roads. Moreover, the association between expressway intensity and rural employment is found subtle when the intensities of other types of roads are controlled.

3.3.3 Heterogeneity Across Gender, Education, and Region

This section reports the results on the heterogeneous impacts of road infrastructure on rural employment based on CRE models. More specifically, we consider the heterogeneity across gender, education, and region.

---

[19] For robustness check, we also measure the employment intensity with the absolute working days and the results are consistent (See Table B.4 in the appendix).

The results on gender, as shown in Table 3.7, indicate that the positive association between the densities of county roads and provincial/national highways is more significant among female laborers. While comparably, the density of county roads is associated with less male participation in migration work but more male participation in agriculture. This suggests that improving road infrastructure, as an indicator of a better local economy, affects male and female rural laborers in different ways. Compared to males who are normally physically advantaged and would perhaps benefit more from labor-intensive agricultural production with better roads, females could benefit comparably more from participating in the local non-farm economy when more employment opportunities are available.

Table 3.8 summarizes the results on education levels. The association between road infrastructure and rural employment, while varying in significance levels, is consistent across two groups with different education levels. The result suggests that the overall increasing density of county roads and provincial/national highways is associated with more local non-farm employment but less migration working in both groups, which is consistent with the results based on total samples.

Tables 3.9, 3.10, and 3.11 report the results on agricultural employment from CRE models by province. As shown in the results, there is observable heterogeneity in the impacts of road infrastructure on rural employment across regions. However, given that there are too many factors that could affect rural farmers' employment choices such as economic and employment structure, endowments in land ownership and road infrastructure, culture, language, etc., and vary across provinces in China, the heterogeneous road impacts on rural employment we observed is not that surprising.

In northern provinces with traditionally more agricultural employment such as Hebei and Liaoning, the density of local roads is negatively associated with rural farmers' participation in agricultural employment while positively associated with the tendency to migration work. In the same area, the density of expressway is also negatively associated with local employment participation including agricultural production (e.g., Liaoning) and local non-farm employment (e.g., Hebei). And the density of county roads is negatively associated with rural farmers' participation in migration work.

In the only southeastern province—Zhejiang where the local non-agricultural economy is more advanced, the densities of county roads and provincial/national highways are positively associated with local non-farm employment while negatively associated with migration working. While for expressways, road density is negatively associated with local non-farm employment while positively associated with migration work.

In the only province in our sample in mid-China—Hubei, the density of provincial and national highways is positively associated with migration work while negatively associated with agricultural employment. In southwest—Sichuan, we didn't find any significant association between road density and employment tendency. In northwest—Shaanxi, the densities of local roads and county roads are negatively associated with agricultural employment, while the densities of the other two types of roads are positively associated with agricultural employment.

3.4 Conclusion

Road infrastructure could affect rural farmers' employment choices through numerous mechanisms such as employment opportunities, information, transportation cost, crop price, etc. And due to the relatively rich employment choices farmers have in well-developed non-agricultural sectors, the long-standing household registration system that constrained free migration, cultural factors, etc., farmers' decision-making process on employment can be complex in the context of China. All these factors, including the endogeneity of road construction and the association between different types of roads, make it challenging to explicitly identify the impacts of road infrastructure on rural employment in China.

Table 3.7 Road infrastructure and rural employment participation by gender in CRE model

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Agriculture**		**Local off-farm**		**Migrant working**	
Participation	Female	Male	Female	Male	Female	Male
Road Densities						
Local	-0.75(0.86)	0.13(0.58)	-0.96(0.63)	-0.41(0.46)	-0.24(1.08)	0.49(0.63)
County	1.64(5.98)	7.83*(4.16)	13.47***(4.44)	3.30(3.17)	-6.49(8.33)	-15.91***(4.71)
Prov/Nation	0.77(7.42)	-0.88(4.74)	9.83**(4.97)	6.04*(3.38)	-8.58(7.08)	-6.89(4.37)
Expressway	0.31(4.03)	-1.79(2.69)	-2.51(2.84)	-0.83(2.03)	-0.06(5.20)	-0.97(2.83)
Year_2008	-0.20(0.51)	-1.38***(0.35)	-1.00**(0.39)	-1.33***(0.27)	1.48***(0.47)	1.37***(0.32)
Year_2013	-1.72***(0.50)	-1.90***(0.39)	-0.17(0.37)	-1.06***(0.29)	2.17***(0.49)	1.36***(0.36)
Age	0.00(0.03)	-0.02(0.02)	0.04(0.02)	0.03*(0.02)	-0.06(0.04)	-0.03(0.03)
Marriage	0.43(0.66)	0.26(0.27)	0.50(0.53)	1.03***(0.22)	-1.07(0.68)	-0.84***(0.26)
Junior high	-0.59***(0.18)	-0.31**(0.12)	0.53***(0.13)	0.22**(0.10)	0.06(0.18)	0.03(0.13)
Tech training	-0.44*(0.24)	-0.35***(0.13)	0.61***(0.18)	0.48***(0.10)	0.44*(0.23)	0.26**(0.13)
HH size	-0.06(0.11)	0.06(0.07)	-0.07(0.08)	0.03(0.05)	-0.08(0.14)	-0.14*(0.08)
Land	0.49**(0.20)	0.49***(0.14)	-0.08(0.15)	-0.23**(0.11)	-0.34(0.25)	0.00(0.15)
House value	0.03(0.06)	0.04(0.04)	-0.01(0.04)	-0.02(0.03)	-0.10(0.07)	-0.07*(0.04)
Durable assets	-0.07(0.09)	-0.00(0.06)	0.19***(0.06)	0.09**(0.04)	-0.21**(0.09)	-0.14**(0.06)
Constant	-3.36***(0.78)	-3.01***(0.57)	-3.21***(0.57)	-1.99***(0.42)	3.30***(0.74)	4.58***(0.59)
Observations	3,878	4,634	3,878	4,634	3,878	4,634

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.8 Road infrastructure and rural employment participation by education level in CRE model

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Agriculture**		**Local off-farm**		**Migrant working**	
Participation	< Junior High	Junior High	< Junior High	Junior High	< Junior High	Junior High
Road Densities						
Local	-0.39(0.72)	0.12(0.69)	-0.21(0.53)	-0.91*(0.53)	-0.68(0.85)	1.04(0.75)
County	4.62(5.20)	6.12(4.78)	2.78(3.57)	9.12**(3.75)	-11.69*(6.06)	-14.84**(5.79)
Prov/Nation	1.22(5.84)	0.34(5.59)	10.97***(4.07)	3.53(3.96)	-13.01**(5.38)	-2.27(5.33)
Expressway	0.68(3.32)	-4.25(3.23)	-3.35(2.40)	1.16(2.41)	3.27(3.83)	-3.44(3.40)
Year_2008	-0.73(0.49)	-1.07***(0.37)	-1.65***(0.37)	-0.93***(0.28)	0.85**(0.41)	1.64***(0.35)
Year_2013	-2.22***(0.49)	-1.58***(0.40)	-0.60*(0.35)	-0.78***(0.30)	1.35***(0.44)	1.72***(0.39)
Male	-0.66***(0.15)	-0.38***(0.14)	1.32***(0.12)	0.59***(0.10)	1.29***(0.16)	0.84***(0.14)
Age	-0.02(0.03)	0.01(0.03)	0.04**(0.02)	0.03(0.02)	-0.01(0.03)	-0.07**(0.03)
Marriage	0.57(0.44)	0.30(0.31)	0.50(0.36)	0.88***(0.25)	-0.30(0.43)	-1.10***(0.31)
Tech training	-0.39**(0.20)	-0.46***(0.14)	0.58***(0.15)	0.50***(0.10)	0.44**(0.19)	0.31**(0.14)
HH size	-0.06(0.09)	0.08(0.09)	-0.08(0.06)	0.12*(0.07)	-0.14(0.10)	-0.16*(0.09)
Land	0.62***(0.17)	0.36**(0.17)	-0.29**(0.13)	-0.05(0.12)	-0.10(0.20)	-0.09(0.17)
House value	0.09*(0.05)	-0.02(0.05)	-0.04(0.04)	0.01(0.04)	-0.06(0.05)	-0.07(0.05)
Durable assets	-0.05(0.07)	-0.00(0.07)	0.12**(0.05)	0.10*(0.05)	-0.16**(0.07)	-0.12*(0.07)
Constant	-2.77***(0.68)	-2.99***(0.61)	-2.63***(0.50)	-3.05***(0.46)	3.50***(0.66)	3.47***(0.62)
Observations	4,342	4,170	4,342	4,170	4,342	4,170

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.9 Road infrastructure and agricultural employment by province in CRE model

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Hebei**	**Liaoning**	**Zhejiang**	**Hubei**	**Sichuan**	**Shaanxi**
Participation	CRE	CRE	CRE	CRE	CRE	CRE
Road Densities						
Local	-1.66*(0.97)	-8.47***(2.36)	2.52(1.54)	2.09(2.39)	-0.66(1.83)	-2.77*(1.58)
County	2.81(5.31)	44.02(79.69)	-7.93(11.33)	10.50(10.15)	-0.77(17.81)	-60.44***(18.21)
Prov/Nation		-21.56(15.95)	92.40***(35.17)	-34.80**(16.27)	-5.69(7.77)	22.65**(9.28)
Expressway	-4.05(5.76)	-29.92***(8.67)	1.59(4.82)	-4.38(5.82)	8.97(8.65)	15.30**(7.74)
Year_2008	-0.40(0.31)	-0.52(0.60)	-2.15***(0.39)	-1.54***(0.44)	-0.48(0.46)	-1.10***(0.39)
Year_2013	-0.54(0.37)	-1.60*(0.88)	-3.57***(0.54)	-2.32***(0.63)	-0.67(0.61)	-1.93***(0.49)
Male	-0.51**(0.22)	-1.01***(0.38)	0.12(0.25)	-0.88***(0.25)	-0.57**(0.29)	-0.46**(0.23)
Age	-0.05(0.04)	0.19**(0.09)	0.02(0.04)	0.06(0.06)	-0.14**(0.06)	-0.00(0.05)
Marriage	0.26(0.56)	0.13(0.84)	-0.54(0.62)	1.04*(0.63)	2.46***(0.76)	0.07(0.54)
Junior high	-0.46**(0.22)	0.05(0.35)	-0.87***(0.28)	0.12(0.24)	-0.87***(0.29)	-0.72***(0.24)
Tech training	-0.12(0.26)	-0.98***(0.38)	0.12(0.25)	-0.18(0.28)	-0.54(0.34)	-0.63**(0.27)
HH size	-0.07(0.14)	0.02(0.26)	0.21(0.15)	-0.06(0.13)	-0.05(0.18)	0.13(0.15)
Land	0.52(0.39)	0.35(0.33)	0.35(0.29)	0.60**(0.26)	0.74**(0.31)	0.16(0.33)
House value	-0.10(0.08)	-0.01(0.12)	0.16**(0.07)	0.05(0.08)	-0.02(0.10)	0.04(0.08)
Durable assets	0.02(0.12)	-0.27(0.17)	-0.28**(0.13)	-0.03(0.11)	0.07(0.12)	0.03(0.11)
Constant	-0.29(0.91)	-7.52***(1.89)	0.33(1.11)	-3.98***(1.21)	-4.43***(1.28)	-4.50***(0.93)
Observations	1,415	1,300	1,449	1,575	1,323	1,450

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.10 Road infrastructure and local non-farm employment by province in CRE model

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Hebei**	**Liaoning**	**Zhejiang**	**Hubei**	**Sichuan**	**Shaanxi**
Participation	CRE	CRE	CRE	CRE	CRE	CRE
Road Densities						
Local	0.13(0.66)	0.50(1.13)	-1.48(1.17)	-5.66***(1.79)	1.46(1.68)	-1.42(1.35)
County	4.77(3.98)	-21.68(43.10)	14.76*(8.61)	-0.35(6.96)	3.38(15.69)	13.11(13.70)
Prov/Nation		8.66(8.29)	6.74(24.38)	-4.80(11.30)	1.34(4.86)	14.44**(6.66)
Expressway	-9.71**(4.15)	3.67(4.59)	-8.42**(3.53)	1.08(4.29)	6.41(6.72)	9.91(6.16)
Year_2008	-0.19(0.24)	-0.64*(0.36)	-0.24(0.26)	0.03(0.32)	-0.91**(0.39)	-1.12***(0.32)
Year_2013	-0.62**(0.31)	-0.69(0.51)	0.42(0.33)	-0.16(0.44)	-0.75(0.50)	-0.42(0.40)
Male	1.18***(0.18)	1.25***(0.21)	0.24(0.17)	1.11***(0.18)	1.25***(0.25)	1.00***(0.20)
Age	0.00(0.03)	0.04(0.05)	0.03(0.03)	0.05(0.04)	0.09*(0.05)	-0.05(0.04)
Marriage	0.87*(0.49)	0.77(0.58)	0.19(0.43)	1.21***(0.45)	0.72(0.61)	1.00**(0.48)
Junior high	0.34*(0.17)	0.50***(0.19)	-0.00(0.20)	0.25(0.18)	0.23(0.23)	0.53***(0.19)
Tech training	0.49**(0.22)	0.23(0.21)	0.40**(0.18)	0.50**(0.20)	0.58**(0.26)	0.89***(0.22)
HH size	0.13(0.10)	0.04(0.14)	-0.08(0.10)	-0.10(0.09)	0.10(0.14)	-0.01(0.11)
Land	-0.10(0.27)	0.10(0.20)	-0.41*(0.22)	-0.12(0.19)	0.10(0.23)	-0.64**(0.29)
House value	-0.02(0.06)	-0.07(0.07)	-0.03(0.05)	0.16**(0.06)	-0.06(0.08)	-0.00(0.06)
Durable assets	-0.01(0.09)	0.11(0.10)	0.18*(0.10)	0.20**(0.08)	-0.02(0.10)	0.21**(0.09)
Constant	-1.40*(0.74)	-2.46***(0.90)	-0.08(0.76)	-5.45***(0.90)	-3.91***(1.01)	-2.92***(0.73)
Observations	1,415	1,300	1,449	1,575	1,323	1,450

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.11 Road infrastructure and migrant working by province in CRE model

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	**Hebei**	**Liaoning**	**Zhejiang**	**Hubei**	**Sichuan**	**Shaanxi**
Participation	CRE	CRE	CRE	CRE	CRE	CRE
Road Densities						
Local	3.22*(1.76)	3.72*(2.19)	-0.19(1.94)	3.30(2.55)	-1.58(1.68)	-1.00(1.49)
County	-16.48**(8.25)	-150.86*(81.50)	-15.39(14.70)	-2.72(9.88)	-1.28(15.68)	-33.03**(16.07)
Prov/Nation		-22.00(15.74)	-82.15**(38.28)	40.84**(17.32)	-4.81(6.58)	-11.14(8.09)
Expressway	5.75(8.28)	0.57(8.39)	9.65*(5.81)	-7.83(6.35)	-3.15(7.77)	-2.00(6.83)
Year_2008	0.86**(0.43)	2.08***(0.65)	1.75***(0.42)	1.42***(0.43)	1.60***(0.43)	1.08***(0.34)
Year_2013	0.47(0.52)	3.09***(0.96)	1.97***(0.52)	2.16***(0.60)	1.34**(0.54)	1.22***(0.42)
Male	1.44***(0.37)	1.18***(0.38)	0.83***(0.27)	0.50**(0.25)	0.99***(0.26)	1.46***(0.22)
Age	-0.04(0.05)	-0.02(0.09)	-0.10*(0.05)	-0.06(0.06)	-0.03(0.05)	0.03(0.04)
Marriage	-0.67(0.75)	-1.68**(0.83)	-0.09(0.56)	-0.63(0.53)	-2.02***(0.67)	-1.21**(0.51)
Junior high	0.01(0.31)	-0.15(0.35)	0.23(0.31)	0.02(0.25)	0.64**(0.26)	0.15(0.20)
Tech training	0.35(0.36)	1.54***(0.40)	0.30(0.25)	-0.02(0.27)	0.51(0.32)	-0.17(0.24)
HH size	-0.50**(0.23)	0.07(0.26)	-0.31*(0.18)	-0.03(0.14)	-0.13(0.17)	0.00(0.13)
Land	-0.58(0.57)	0.76*(0.42)	0.34(0.34)	-0.23(0.29)	-0.59**(0.30)	-0.42(0.31)
House value	0.05(0.13)	-0.12(0.12)	-0.05(0.08)	-0.32***(0.09)	-0.04(0.09)	-0.02(0.07)
Durable assets	-0.14(0.17)	-0.23(0.16)	0.02(0.15)	-0.27**(0.12)	-0.05(0.11)	-0.27***(0.10)
Constant	2.87**(1.32)	3.98**(1.63)	-1.67(1.16)	5.46***(1.23)	4.31***(1.13)	2.15***(0.79)
Observations	1,415	1,300	1,449	1,575	1,323	1,450

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Building upon the advantage of the road information from GIS data, our study combines road data and farm-level survey data and aims at assessing the distinct impacts of various rural road infrastructures on farmers' employment choices in China, attempting to address rural farmers' employment choices in three dimensions—agriculture vs. non-agriculture, full time vs. part-time, and local vs. non-local. Our results suggest that the overall improvements in densities of local roads, county roads, and provincial/national highways are positively associated with farmers' tendency to participate and work longer in local off-farm employment while negatively associated with full-time agricultural production and migration working. The increasing local off-farm employment opportunities and lower transportation costs associated with the increasing road intensity might be two main factors contributing to such results.

Surprisingly, we didn't find a significant association between expressway density and rural farmers' employment choice in our studied context as what is suggested in previous studies. Several factors might contribute to such results. First, for expressways, it might be the accessibility rather than the density that matters more to farmers' employment choices. Second, a large portion of migrant workers in China might rely on the well-developed railway system rather than the expressway for long-distance transportation. And more importantly, it is likely that the impacts of expressway density on rural farmers' employment choice are through indirect impacts on the local economy, which might be controlled by the densities of other types of roads in our empirical analysis.

Our study provides a new perspective on understanding the role of rural road infrastructure in determining farmers' employment choices. Rather than promoting more migration working, the improvement in rural road infrastructure in our studied context boosts more local off-farm employment, especially local part-time employment. Such a conclusion could contribute to a better understanding of the impacts of road construction on rural farmers' employment choices in China. In the future, more studies on identifying specific mechanisms through which road infrastructure affects farmers' employment choices are needed.

REFERENCES

Adukia, A., Asher, S., & Novosad, P. (2020). Educational investment responses to economic opportunity: evidence from Indian road construction. American Economic Journal: Applied Economics, 12(1), 348-376.

Aggarwal, S. (2018). Do rural roads create pathways out of poverty? Evidence from India. Journal of Development Economics, 133, 375-395.

Altonji, J. G., & Matzkin, R. L. (2005). Cross section and panel data estimators for nonseparable models with endogenous regressors. Econometrica, 73(4), 1053-1102.

Asher, S. and Novosad, P. (2018), Rural roads and local economic development, The World Bank.

Bai, Y., Zhou, T., Ma, Z., & Zhang, L. (2021). Does road accessibility benefit rural poor? Evidence on the extent of household off-farm employment from 2004 to 2018. China Agricultural Economic Review, 13(3), 639-672.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. Review of Economics and Statistics, 99(3), 435-448.

Banerjee, A., Duflo, E., & Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. Journal of Development Economics, 145, 102442.

Buys, P., Deichmann, U., & Wheeler, D. (2010). Road network upgrading and overland trade expansion in Sub-Saharan Africa. Journal of African Economies, 19(3), 399-432.

Bryceson, D. F., Bradbury, A., & Bradbury, T. (2008). Roads to poverty reduction? Exploring rural roads' impact on mobility in Africa and Asia. Development Policy Review, 26(4), 459-482.

Corral, L., & Reardon, T. (2001). Rural nonfarm incomes in Nicaragua. World development, 29(3), 427-442.

DeBrauw, A., Jikun, H., Scott, R., Linxiu, Z. and Yigang, Z. (2002), "The evolution of China's rural labor markets during the reforms", Journal of Comparative Economics, Vol. 30 No. 2, pp. 329-353.

De Janvry, A., & Sadoulet, E. (2001). Income strategies among rural households in Mexico: The role of off-farm activities. World development, 29(3), 467-480.

Dercon, S., Gilligan, D. O., Hoddinott, J., & Woldehanna, T. (2009). The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages. American Journal of Agricultural Economics, 91(4), 1007-1021.

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. American Economic Review, 108(4-5), 899-934.

Escobal, J. (2001). The determinants of nonfarm income diversification in rural Peru. World development, 29(3), 497-508.

Escobal, J., & Ponce, C. (2002). The benefits of rural roads: Enhancing income opportunities for the rural poor.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. Review of Economic Studies, 81(3), 1046-1070.

Fan, S., & Chan-Kang, C. (2008). Regional road development, rural and urban poverty: Evidence from China. Transport Policy, 15(5), 305-314.

Gachassin, M., Najman, B., & Raballand, G. (2010). The impact of roads on poverty reduction: a case study of Cameroon. World Bank Policy Research Working Paper, (5209).

Gibson, J., & Olivia, S. (2010). The effect of infrastructure access and quality on non-farm enterprises in rural Indonesia. World Development, 38(5), 717-726.

Gibson, J., & Rozelle, S. (2003). Poverty and access to roads in Papua New Guinea. Economic development and cultural change, 52(1), 159-185.

Huang, Q., Zheng, X., & Wang, R. (2022). The impact of the accessibility of transportation infrastructure on the Non-Farm employment choices of rural laborers: Empirical analysis based on China's micro data. Land, 11(6), 896.

Jacoby, H. G. (2000). Access to markets and the benefits of rural roads. The economic journal, 110(465), 713-737.

Khandker, S. R., Bakht, Z., & Koolwal, G. B. (2009). The poverty impact of rural roads: Evidence from Bangladesh. Economic development and cultural change, 57(4), 685-722.

Minten, B., & Kyle, S. (1999). The effect of distance and road quality on food collection, marketing margins, and traders' wages: evidence from the former Zaire. Journal of Development Economics, 60(2), 467-495.

Mu, R., & Van de Walle, D. (2011). Rural roads and local market development in Vietnam. The Journal of Development Studies, 47(5), 709-734.

Pinstrup-Andersen, P., & Shimokawa, S. (2006). Rural infrastructure and agricultural development. World Bank.

Qiao, F., Rozelle, S., Huang, J., Zhang, L., & Luo, R. (2014). Road expansion and off-farm work in rural China. The China Quarterly, 218, 428-451.

Shamdasani, Y. (2021). Rural road infrastructure & agricultural production: Evidence from India. Journal of Development Economics, 152, 102686.

Shrestha, S. A. (2020). Roads, participation in markets, and benefits to agricultural households: Evidence from the topography-based highway network in Nepal. Economic Development and Cultural Change, 68(3), 839-864.

Wang, Z., & Sun, S. (2016). Transportation infrastructure and rural development in China. China Agricultural Economic Review.

Warr, P. (2005). Road development and poverty reduction: the case of Lao PDR (No. 64). ADBI Research Paper Series.

Wooldridge, J. M. (2005). Unobserved heterogeneity and estimation of average partial effects. Identification and inference for econometric models: Essays in honor of Thomas Rothenberg, 27-55.

Wooldridge, J. M. (2019). Correlated random effects models with unbalanced panels. Journal of Econometrics, 211(1), 137-150.

World Bank. (1994). World development report 1994: Infrastructure for development. The World Bank.

Yamauchi, F., Muto, M., Chowdhury, S., Dewina, R., & Sumaryanto, S. (2011). Are schooling and roads complementary? Evidence from income dynamics in rural Indonesia. World Development, 39(12), 2232-2244.

Yamauchi, F. (2016). The effects of improved roads on wages and employment: Evidence from rural labour markets in Indonesia. The Journal of Development Studies, 52(7), 1046-1061.

Zhang, L., Dong, Y., Liu, C., & Bai, Y. (2018). Off-farm employment over the past four decades in rural China. China Agricultural Economic Review.

Zhang, L., Luo, R., Liu, C., Rozelle, S., (2006). Investing in rural China: tracking China's commitment to modernization. The Chinese Economy 39 (4), 57–84.

APPENDIX

1. Village Road

2. Township Road

3. County Road

4. Provincial/National Highway

5. Expressway

Figure B.1 Example images of different types of roads in China

Table B.1 Road infrastructure and agricultural employment

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:		**Full-time**			**Part-time**	
Participation	Pooled Logit	FE Logit	CRE Logit	Pooled Logit	FE Logit	CRE Logit
Road Densities						
Local	-0.17(0.11)	-0.06(0.44)	-1.57(2.75)	-0.24**(0.11)	-0.39(0.39)	-0.43(0.37)
County	-1.40***(0.52)	-1.04(2.90)	-2.73(2.80)	1.28***(0.49)	0.51(2.70)	3.14(2.49)
Prov/Nation	-2.72***(0.78)	-1.77(3.10)	2.57(1.83)	3.90***(0.74)	4.14(2.89)	3.39(2.63)
Expressway	-1.33(1.01)	3.17(1.96)	0.20(0.22)	-0.80(0.96)	-3.41*(1.78)	-2.30(1.68)
Year_2008	-0.15(0.16)	0.12(0.40)	-0.52**(0.24)	-0.70***(0.16)	-0.61(0.38)	-0.80***(0.21)
Year_2013	-0.77***(0.17)	-0.68(0.55)	-1.72***(0.09)	-0.73***(0.16)	-0.42(0.51)	-0.78***(0.23)
Male	-1.35***(0.06)		0.01(0.02)	1.09***(0.06)		1.33***(0.08)
Age	0.07***(0.00)		-0.04(0.27)	-0.00(0.00)		0.00(0.01)
Marriage	0.61***(0.09)		-0.39***(0.08)	1.07***(0.09)		-0.04(0.21)
Junior high	-0.31***(0.06)		-0.99***(0.11)	0.05(0.06)		0.04(0.08)
Tech training	-0.86***(0.08)		0.05(0.05)	0.32***(0.07)		0.40***(0.09)
HH size	-0.00(0.02)		0.28***(0.10)	-0.12***(0.02)		-0.05(0.04)
Land	0.31***(0.04)	0.30***(0.11)	0.06**(0.03)	0.11***(0.04)	0.09(0.09)	0.10(0.09)
House value	0.01(0.02)	0.07**(0.03)	-0.03(0.04)	0.02(0.01)	-0.02(0.03)	-0.02(0.03)
Durable assets	-0.10***(0.02)	-0.05(0.04)	-2.85***(0.38)	0.04**(0.02)	0.05(0.04)	0.03(0.04)
Constant	-2.03***(0.23)			-2.46***(0.22)		-3.42***(0.34)
Observations	8,512	2,045	8,512	8,512	2,884	8,512

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table B.2 Road infrastructure and local off-farm employment

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:		**Full-time**			**Part-time**	
Participation	Pooled Logit	FE Logit	CRE Logit	Pooled Logit	FE Logit	CRE Logit
Road Densities						
Local	0.68***(0.13)	-0.04(0.54)	0.24(0.50)	-0.23**(0.11)	-0.84**(0.40)	-0.93**(0.38)
County	-0.69(0.59)	-0.52(3.96)	-1.81(3.49)	0.90*(0.52)	4.48(2.78)	6.83***(2.59)
Prov/Nation	1.62*(0.95)	12.26(8.56)	11.17**(5.18)	3.98***(0.78)	5.37*(3.07)	5.69**(2.81)
Expressway	-0.55(1.12)	0.94(2.63)	1.59(2.28)	-0.49(1.03)	-2.14(1.83)	-1.76(1.74)
Year_2008	-0.08(0.34)	-0.37(0.66)	-0.79*(0.41)	-0.80***(0.17)	-0.54(0.41)	-0.87***(0.23)
Year_2013	1.24***(0.28)	1.30(0.80)	0.61*(0.36)	-0.90***(0.18)	-0.36(0.55)	-0.90***(0.25)
Male	0.02(0.07)		0.08(0.10)	0.96***(0.06)		1.17***(0.09)
Age	-0.04***(0.00)		0.08***(0.02)	0.01***(0.00)		0.01(0.02)
Marriage	0.21**(0.10)		0.33(0.25)	1.26***(0.11)		0.32(0.23)
Junior high	0.39***(0.08)		0.44***(0.10)	0.13**(0.06)		0.13*(0.08)
Tech training	0.35***(0.08)		0.34***(0.11)	0.29***(0.07)		0.35***(0.09)
HH size	0.07***(0.02)		0.04(0.06)	-0.13***(0.02)		-0.05(0.05)
Land	-0.53***(0.05)	-0.38***(0.13)	-0.38***(0.12)	0.08*(0.04)	0.10(0.10)	0.07(0.09)
House value	0.00(0.02)	0.00(0.04)	-0.01(0.03)	0.03**(0.02)	0.00(0.03)	0.00(0.03)
Durable assets	0.21***(0.02)	0.12*(0.06)	0.10*(0.05)	0.09***(0.02)	0.08**(0.04)	0.07**(0.04)
Constant	-3.19***(0.34)		-3.15***(0.49)	-3.68***(0.24)		-4.92***(0.37)
Observations	8,512	1,300	8,512	8,512	2,242	8,512

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table B.3 Road infrastructure and migration working

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:		**Full-time**			**Part-time**	
Participation	Pooled Logit	FE Logit	CRE Logit	Pooled Logit	FE Logit	CRE Logit
Road Densities						
Local	-0.66***(0.15)	-0.43(0.72)	-0.63(0.64)	-0.28(0.24)	0.97(1.09)	1.71**(0.78)
County	0.76(0.65)	-10.90**(5.01)	-7.27(4.90)	2.83**(1.11)	-16.79**(6.78)	-12.90**(5.19)
Prov/Nation	-4.04***(1.08)	-10.42*(5.79)	-6.96(4.53)	1.05(1.56)	-1.18(5.50)	-5.32(4.58)
Expressway	1.86(1.27)	-0.86(3.15)	0.84(2.85)	-0.65(2.08)	-0.99(4.05)	-1.47(3.54)
Year_2008	1.84***(0.22)	2.69***(0.69)	2.08***(0.32)	-0.02(0.25)	-0.10(0.77)	-0.17(0.30)
Year_2013	2.08***(0.22)	2.95***(0.95)	2.21***(0.34)	0.14(0.25)	0.05(1.10)	-0.05(0.33)
Male	0.48***(0.08)		0.67***(0.12)	1.25***(0.14)		1.32***(0.15)
Age	-0.09***(0.00)		-0.03(0.02)	-0.04***(0.01)		0.01(0.03)
Marriage	-0.93***(0.10)		-0.80***(0.26)	0.36**(0.16)		-0.68*(0.35)
Junior high	0.11(0.08)		0.20*(0.12)	-0.22*(0.12)		-0.23*(0.13)
Tech training	0.18**(0.09)		0.26**(0.12)	0.23*(0.13)		0.25*(0.15)
HH size	0.09***(0.03)		-0.15**(0.08)	0.00(0.05)		-0.04(0.09)
Land	-0.02(0.05)	-0.20(0.18)	-0.19(0.14)	0.19**(0.09)	0.12(0.22)	0.17(0.18)
House value	-0.02(0.02)	-0.04(0.05)	-0.04(0.04)	-0.05*(0.03)	-0.09*(0.05)	-0.09*(0.05)
Durable assets	-0.15***(0.02)	-0.10(0.06)	-0.10*(0.05)	-0.14***(0.04)	-0.08(0.07)	-0.13**(0.06)
Constant	1.83***(0.31)		2.80***(0.52)	-1.45***(0.41)		-1.65***(0.54)
Observations	8,512	910	8,512	8,512	642	8,512

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table B.4 Road infrastructure and working days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable:		**Agriculture**			**Local off-farm**			**Migrant working**	
Working days	Pooled Tobit	RE Tobit	CRE Tobit	Pooled Tobit	RE Tobit	CRE Tobit	Pooled Tobit	RE Tobit	CRE Tobit
Road Densities									
Local	-8.51(5.33)	-7.33(5.86)	23.41*(13.86)	42.80***(12.94)	39.94***(13.98)	-55.60(34.40)	-79.01***(19.69)	-76.78***(20.98)	-1.12(58.52)
County	-69.57***(24.59)	-75.71***(27.37)	-16.06(93.74)	66.47(61.89)	88.92(68.13)	624.78***(239.80)	142.72*(86.60)	127.60(93.16)	-1,411.01***(426.93)
Prov/Nation	-60.27(36.79)	-80.84**(39.27)	-10.74(94.63)	483.44***(95.64)	557.02***(101.00)	756.01***(276.64)	-374.14***(139.38)	-328.25**(145.71)	-803.62**(397.53)
Expressway	-77.97(47.63)	-49.09(46.81)	3.45(61.76)	-106.07(119.29)	-123.81(115.61)	-186.99(157.98)	186.67(168.36)	120.90(168.09)	127.82(263.62)
Year_2008	-12.68*(7.62)	-17.66***(6.73)	10.65(7.47)	-76.46***(22.59)	-74.76***(20.34)	-112.18***(22.15)	194.44***(26.79)	205.03***(24.68)	144.55***(26.87)
Year_2013	-118.97***(8.00)	-127.05***(7.35)	-86.00***(8.62)	-26.64(21.71)	-21.54(20.01)	-71.00***(23.18)	193.99***(27.32)	204.91***(25.78)	121.13***(29.54)
Male	-15.16***(2.71)	-13.68***(3.09)	-13.42***(3.08)	83.76***(7.17)	80.01***(7.98)	81.73***(7.98)	116.67***(10.28)	108.72***(11.00)	112.37***(11.05)
Age	3.44***(0.13)	3.45***(0.14)	-1.17**(0.55)	-1.95***(0.35)	-1.98***(0.37)	5.22***(1.46)	-12.08***(0.57)	-11.99***(0.60)	-2.08(2.19)
Marriage	76.25***(4.35)	72.89***(4.49)	19.04**(8.34)	104.10***(10.79)	92.82***(10.93)	51.81***(19.41)	-116.55***(13.26)	-118.33***(13.51)	-91.11***(24.77)
Junior high	-14.33***(2.89)	-15.14***(3.08)	-14.44***(3.08)	53.95***(7.53)	50.16***(7.92)	46.85***(7.93)	5.96(10.47)	9.03(10.87)	10.37(10.90)
Tech training	-24.56***(3.70)	-21.43***(3.70)	-19.39***(3.69)	59.23***(8.56)	48.47***(8.41)	46.22***(8.39)	34.12***(11.68)	37.52***(11.68)	34.75***(11.67)
HH size	-5.00***(0.99)	-4.50***(1.02)	-1.10(1.59)	-4.63*(2.53)	-2.49(2.59)	3.03(4.28)	13.75***(3.68)	12.40***(3.77)	-10.61(7.02)
Land	23.49***(2.00)	20.81***(2.07)	12.96***(3.30)	-37.62***(5.05)	-33.81***(5.17)	-20.45**(8.45)	5.00(7.28)	4.12(7.42)	-2.75(13.48)
House value	0.12(0.72)	0.15(0.71)	0.51(0.96)	2.04(1.73)	1.30(1.67)	-0.70(2.31)	-2.18(2.45)	-2.05(2.43)	-5.64(3.69)
Durable assets	-3.35***(0.92)	-2.81***(0.92)	-1.13(1.27)	29.36***(2.45)	25.25***(2.43)	14.55***(3.47)	-21.98***(3.23)	-20.47***(3.22)	-14.10***(5.06)
Constant	-38.19***(10.70)	-38.01***(11.09)	-70.26***(13.31)	-389.64***(28.85)	-371.21***(29.43)	-339.63***(34.95)	280.77***(38.79)	272.18***(39.58)	313.74***(46.72)
Observations	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512	8,512

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1