Economic Tipping Points of Permanent Grassland Conversion: Price Shocks, Carbon Costs, and the Extensive Margin

Luoye Chen* Madhu Khanna*

June 2025

Abstract

Permanent grassland conversions pose pressing concerns for carbon storage and ecosystem services due to substantial implications for climate and food security. Despite their importance, there has been extremely limited understanding of economic tipping points driving permanent grassland conversion. Using high-resolution data on 500-meter land parcels in the rainfed United States from 2001 to 2017, we provide the first large-scale investigation of how commodity price fluctuations trigger permanent grassland conversion and quantify the resulting carbon-food security trade-offs. We characterize heterogeneity in conversion responses across land parcels with different biophysical attributes using spatial econometrics with machine learning inference. We find that the contemporary rainfall shocks, long-term climate change, and lagged received commodity prices affect the likelihood of crop-grass conversions. The effects of rainfall shocks and price fluctuations on crop-grass conversions are mainly confined to land parcels at the margin of agricultural production, with relatively lower water storage capacity, less productivity, and steeper slopes. These findings suggest that permanent grassland conversion responds to identifiable economic tipping points and highlight the importance of adaptive conservation policies that adjust for market volatility.

*Chen: luoyechen@hkust-gz.edu.cn, Carbon Neutrality and Climate Change Thrust at the Hong Kong University of Science and Technology (Guangzhou); Khanna: khanna1@illinois.edu, Department of Agricultural and Consumer Economics at the University of Illinois at Urbana-Champaign. We thank seminar participants at the ASSA Annual Meeting, AERE Summer Conference, AAEA & WAEA Joint Annual Meeting, University of Illinois at Urbana-Champaign, University of Wisconsin-Madison, University of Tennessee at Knoxville, Virginia Tech, Hong Kong University of Science and Technology (Guangzhou), and Jinan University for useful comments and suggestions.

1 Introduction

Permanent grasslands store nearly one-third of the world's soil carbon—a vast reservoir that remains stable for decades unless disturbed by agricultural conversion (Scurlock and Hall, 1998; Yang et al., 2019; Bardgett et al., 2021; Bai and Cotrufo, 2022). Yet these irreplaceable carbon stores face accelerating pressure from commodity markets: when crop prices spike, grasslands are often converted to farmland that yields less than existing agricultural lands while releasing centuries of accumulated carbon within years (Lark et al., 2020, 2022)[1]. This pattern reflects a fundamental conflict between short-term market incentives and long-term ecosystem services, with conversion unevenly distributed as climate change reshapes agricultural boundaries and commodity policies create spatially heterogeneous economic incentives (Song et al., 2018; Winkler et al., 2021; Potapov et al., 2022). However, current conservation programs fail to incorporate spatial considerations of conversion probability and do not adequately target the most vulnerable permanent grasslands that remain under natural vegetation. Understanding how grassland conversions respond to crop price fluctuations is therefore essential for designing conservation policies that can prevent irreversible carbon losses while securing global food systems.

Despite its critical importance, quantifying price-induced grassland conversion presents significant empirical challenges that have limited rigorous economic analysis. First, distinguishing permanent from temporary grassland conversions remains difficult without reliable, long-term, high-resolution datasets (Lark et al., 2015; Potapov et al., 2022). This measurement problem has profound implications: managed pastures in crop rotations represent reversible land-use changes, while permanent grassland conversion constitutes effectively irreversible ecosystem destruction on policy-relevant timescales (Dass et al., 2018; Bai and Cotrufo, 2022). Second, establishing causal relationships between economic drivers and land-use decisions requires addressing the endogeneity of commodity prices, as unobserved factors

[1]Lark et al. (2020)find that the representative crop yields of new croplands in in 2008-2016 (which were converted from grassland recently) were lower than the corresponding national average yields on existing cropland (10.9% lower for corn, 8.4% lower for soybean).

such as regional policy changes or supply shocks could simultaneously affect both conversion decisions and market prices (Roberts and Schlenker, 2013; Hendricks et al., 2014; Costinot et al., 2016; Aragón et al., 2021). Third, land-use responses to price signals are highly heterogeneous, varying systematically with underlying biophysical and economic conditions (Costinot et al., 2016; Aragón et al., 2021). A focus on average effects alone masks this critical variation, preventing identification of the specific lands most vulnerable to conversion and undermining the design of efficient, targeted conservation policies.

To answer these questions, we construct a comprehensive dataset combining MODIS satellite data at 500-meter resolution covering the entire U.S. rainfed region from 2001 to 2017, tracking grassland conversions across over 120 million parcel-year observations spanning multiple commodity price cycles. We complement this dataset with historical vegetation records from the Rangeland Analysis Platform (RAP), a comprehensive product that integrates field plot data from monitoring sites with high-resolution Landsat satellite imagery and ecological modeling to characterize vegetation composition at 30-meter resolution from 1984 to 2000. We then classify permanent grasslands as those maintaining consistent perennial grass dominance for over fifteen years, distinguishing them from managed grasslands that experience frequent cultivation, reseeding, or rotation cycles. Our empirical strategy employs a panel framework that exploits year-by-year fluctuations in local weather shocks and state-level commodity prices while controlling for unobserved local heterogeneity and temporal effects. We characterize heterogeneity in conversion responses across parcels with different biophysical attributes using the generic machine learning inference approach of Chernozhukov et al. (2018), which allows for robust causal inference on heterogeneous treatment effects in high-dimensional settings—a significant methodological advance for environmental policy evaluation.

A key econometric challenge is the potential endogeneity of commodity prices, as unobserved factors such as weather shocks, policy changes, or supply disruptions could simultaneously affect both market prices and landowners' conversion decisions. We address this

through an instrumental variable approach grounded in competitive storage theory (Roberts and Schlenker, 2013; Li et al., 2019). Specifically, we instrument state-level crop price indices with lagged state-level crop stocks, which affect current prices through inventory effects but are plausibly unrelated to individual parcel-level conversion decisions in subsequent periods. This identification strategy, combined with our high-resolution panel structure, enables us to isolate causal price effects while controlling for unobserved heterogeneity that could bias standard approaches.

Our findings reveal economically and environmentally significant effects of commodity prices on grassland conversion, with substantial differences between permanent and managed systems. A 1% increase in lagged commodity prices leads to a 0.06% increase in grassland-to-cropland conversions, indicating that crop-grass conversions represent a primary source of extensive margin supply responses to price fluctuations documented in the agricultural economics literature (Roberts and Schlenker, 2013; Li et al., 2019). More critically, permanent grasslands exhibit disproportionately higher conversion rates in response to price increases compared to managed grasslands, despite representing only 35% of total grassland area in our sample. This differential response reflects the lower private opportunity costs of converting less-intensively used lands, which generate minimal direct economic returns compared to managed pastures actively used for livestock production. The environmental consequences are substantial: price-driven conversion systematically targets ecosystems with the highest carbon storage and biodiversity value while producing crops that consistently underperform established agricultural lands by significant margins. This pattern represents a profound market failure where private incentives drive economically inefficient land-use changes that impose large, unpriced costs on society through ecosystem service losses.

Our machine learning estimates further reveal signficant heterogeneity in conversion responses that has critical implications for policy targeting. Weather shocks and price fluctuations have negligible effects on approximately 60% of rainfed lands, while substantially increasing conversion probability for roughly 20% of parcels. Most importantly, we identify

a highly vulnerable subset—the bottom 5% of our distribution—where economic shocks trigger conversion responses nearly ten times larger than the average effect. These vulnerable lands share a distinct biophysical profile: lower soil water storage capacity, reduced inherent productivity, and steeper slopes. This concentration of risk enables precision targeting of conservation resources toward parcels where market failures are most severe and where intervention can generate the highest environmental returns per dollar invested.

This paper contributes to several strands of economic literature while addressing critical policy gaps. First, we contribute to the growing literature on agricultural land-use economics using field-level data (Scott, 2013; Hendricks et al., 2014; Ramsey et al., 2020; Cui, 2020). Although there is a large body of work studying the land-use change in general, most existing work only focuses on the aggregated cultivated cropland acreage (Aragón et al., 2021; Costinot et al., 2016; Li et al., 2019; Miao et al., 2016; Roberts and Schlenker, 2013), crop choices (Ramsey et al., 2020; Hendricks et al., 2014; Scott, 2013), and deforestation (Souza-Rodrigues, 2019; Burgess et al., 2012). The crop-grass conversions are largely ignored due to the difficulty of measuring these conversions and the less awareness of their economic and ecosystem consequences historically (see Bardgett et al. (2021) for detailed review). This article provides the first large-scale investigation of how U.S. farmers and landowners adjust crop-grass conversions in response to weather shocks and commodity price fluctuations and how these adjustments vary across land parcels with different biophysical attributes. To the best of our knowledge, we are the first paper distinguish between permanent and temporary grassland conversion—a distinction with first-order implications for carbon storage and biodiversity.

We also contribute to the broader literature on spatial targeting of conservation policies (Ando et al., 1998; Babcock et al., 1997; Wu et al., 2001). Our machine learning framework demonstrates how advanced econometric methods can identify vulnerable ecosystems at policy-relevant scales, moving beyond broad geographic targeting toward precision approaches that maximize environmental benefits per dollar spent. We identify highly vulner-

able permanent grasslands—comprising the bottom 5% of our distribution—that experience conversion responses ten times larger than average during price shocks. These vulnerable lands share common characteristics: lower soil organic matter, reduced water storage capacity, and steeper slopes. Crucially, they also store above-average carbon while producing below-average crops when converted, representing economically inefficient trades that generate the worst possible outcomes from both climate and food security perspectives.

These findings have immediate implications for the design of conservation programs and carbon markets. Current programs like the Conservation Reserve Program offer uniform payment rates that fail to adjust for market volatility, creating systematic underprotection during high-price periods when conversion pressure peaks. Our analysis provides the empirical foundation for adaptive payment mechanisms that increase conservation incentives when crop prices exceed identified thresholds, potentially preventing economically inefficient grassland loss while reducing program costs during low-price periods. The framework also supports the development of avoided-conversion protocols in voluntary carbon markets, providing rigorous additionality assessments based on economic conversion risk.

The remainder of this paper proceeds as follows. Section 2 develops our conceptual framework for analyzing permanent grassland conversion decisions under economic uncertainty. Section 3 describes our comprehensive dataset integrating satellite observations, weather data, soil characteristics, and economic variables. Section 4 presents our spatial econometric methodology with machine learning enhancement. Section 5 reports results on economic tipping points, carbon storage implications, and spatial targeting opportunities. Section 6 discusses policy implications for conservation program design and carbon market development. Section 7 concludes with broader implications for climate-agriculture policy integration.

2 Conceptual framework

In this section, we develop a stylized framework to underlie the empirical framework developed in the next section. We consider a parcel of land i that belongs to a farm located in m. Assume there is a continuum of such parcels, and for each parcel, a farmer decides whether or not to clear it for agriculture. Let P_s be a vector of input and output prices at the state level and X be other determinants of land use (e.g., productivity factors). We assume farmers are price takers, that all production is sold in nearby markets within each state directly. To simplify the model, we assume there are no additional transportation costs to deliver the agricultural products from farms to nearby markets. The expected profits for using the land parcel for agriculture can then be defined as $\Pi^c(P_s, X_{im})$, and the corresponding value obtained from leaving the plot as grassland or other natural/semi-natural vegetation. The agricultural value Π^c incorporates profits from annual crop production (cereal and broadleaf crops) and upfront conversion costs, if the land parcel is converted from grassland. The grassland value Π^n includes profits from grassland products such as forage and possible non-pecuniary benefits such as wildlife and recreation resources[2]. Then parcel i is converted to agriculture uses when its agricultural value exceeds its grassland value. Let Y_i, a binary indicator, equals one if the parcel i is cleared and cultivated under annual crops and zero otherwise.

$$Y_{ims} = \mathbf{1}\{\Pi^c(P_s, X_{im}) > \Pi^n(P_s, X_{im})\}$$

The vector of determinants of land use can be decomposed as $X_{im} = \{X_i, X_m, U_m, \epsilon^x_{ims}\}$, where X_i is a parcel-level vector of observed biophysical attributes, such as soil quality, slope, and other geographical conditions; X_m is a location-level observed productivity shifter, such as the weather shocks and changes in other agro-climatic conditions, which could af-

[2]For grasslands, conversion decisions involve significant economic and environmental trade-offs. First, conversion costs must account for substantial carbon losses from stored soil organic carbon. Second, permanent grasslands provide ecosystem services that managed grasslands cannot fully replicate. Third, conversion represents an essentially irreversible land use change with long-term implications for both productivity and ecological function.

fect multiple land parcels in this location; and ϵ^x_{ims} captures any unobserved parcel-level or location-level factors, including but not limited to, farmers' unobserved idiosyncratic abilities, efforts (e.g., management practice and input application, insurance policy purchase), common temporal shocks (e.g., changes in government programs or subsidies), and other possible productivity shocks.

A key distinction between our conceptual framework and those in previous studies is that they assume farmers/households may adjust the planting activity based on the realization of weather throughout the year (Aragón et al., 2021), while we assume the planting activities is a one-off activity based on farmers' expectation of productivity and crop price before the planting season. Although farmers in regions with the multi-cropping system, such as Peru, may adjust their planting acreage throughout the year, most of the cultivated land in the rainfed U.S. is under monoculture farming in a year[3]. As documented by the USDA usual planting dates, the active planting dates across the rainfed U.S. are concentrated between March and May. This assumption is also consistent with the literature on climate and land use. For example, recent studies such as Miao et al. (2016) and Huang and Moore (2019) assume farmers would make decisions based on expected profitability and not change the amount of acreage planted over the growing season[4].

Here, we note that the unexceptional shocks such as extreme weather events should be interpreted as farmers' beliefs or expectations of weather shocks in the post-planting period, as we described above. This belief/expectation would be driven by farmers' prior information and observed weather conditions in the pre-planting period of each year. In addition, we note that X_m could include multiple types of weather shock events in the pre-planting season, such as dry anomaly, wet anomaly, and high-temperature events. For example, exceptionally low precipitation in the pre-planting season may affect the root growth of crops by decreasing

[3]Take double cropping as an example, reports from the USDA/ERS shows that, from 1999 to 2012 double cropping occurred on only about 2 percent of total cropland in the US in most years. `https://www.ers.usda.gov/webdocs/publications/43862/46871_eib125.pdf?v=36.6`

[4]Farmers can, however, adjust the amount of land being harvested, as Cui (2020) documented. The discussion on farmers' potential adjustment for harvested acreage is beyond the scope of this paper.

the available water storage in the soil that can be used for crop growth in the growing season; therefore, this might reduce farmers' expectation of profitability[5]. The potential effect of pre-planting precipitation could be further complicated by other weather variables such as pre-planting temperature. Based on the agronomic observations, high temperature contributes to high atmospheric evaporation demand, which can either reduce the potential delay of planting dates in the case of high pre-planting precipitation or put greater stress on available moisture in the case of low pre-planting precipitation (Seager and Hoerling, 2014).

Additionally, the impact of weather conditions on land productivity might vary across fields with different biophysical attributes. For example, the productivity of fields with high and steep slopes may experience serve loss due to the soil erosion caused by rainfalls, while that of flat fields nearby might merely have a negligible loss (Ma et al. 2018). This could lead to potential strategical responses of farmers' planting choices to weather shocks across fields with various biophysical attributes. We utilize these variations in our heterogeneity analysis.

With this framework in mind, we examine the effect of local weather shocks and crop prices on farmers' crop-grass decisions and the heterogeneity in effects across lands with different biophysical attributes in the following sections.

3 Data

3.1 Land Cover Data

We compile annual measures of land-use type at a spatial resolution of 500 meters across the eastern U.S. (rainfed region) based on the latest MODIS/Terra + Aqua Combined Land

[5]The effect of an exceptionally high precipitation in the pre-planting season is conceptually ambiguous in the literature. While the exceptionally high precipitation in the growing season may cause flood events that could damage field crops through root disease, nitrogen loss, soil erosion, and oxygen depletion, as well as delays in the planting dates (Sacks et al., 2010; Weinraub, 2013), the harvested yields might not be substantially affected if the weather conditions (temperature and precipitation) in the post-planting season are proper for crop growth (Nafziger, 2019). `https://farmdoc.illinois.edu/field-crop-production/uncategorized/dealing-with-very-late-planting.html)`

Cover product (collection 6) provided by NOAA[6]. The raw MODIS land cover data contain several classification systems of land types[7]. Here we extract land type for each 500 m pixel using the Plant Functional Types (PFT) classification system, which is widely used in related scientific studies (Jiang et al., 2021). The PFT system classifies each land pixel into one of the 11 cover types (e.g., cereal/broad-leaf crops, grass, trees, etc).

To limit our analysis to agriculture activities only, we remove any land pixels that are (1) classified as water and urban/developed land in any year and (2) always classified as forests across 2001-2017. We limit our sample to the eastern part of U.S. where most of the land is rain-fed and dominated by dryland agriculture practices. All remaining land pixels are then classified as cropland or non-cropland, leaving a balanced panel set of over 6 million pixels and a total of over 115 million pixel-year observations across the entire eastern U.S.. For our main analysis, we also compile a 4 km grid-level sample by matching up each 500 m land pixel with the 4 km PRISM grids based on their geographic information. The share of cropland within each 4 km PRISM grid is calculated by counting the number of 500 m land pixels classified as cropland within each 4 km grid. This leaves a balanced panel set of over 177 thousand grids and a total of over 3 million grid-year observations across the entire eastern U.S..

The major advantage of this state-of-art remote sensing product is the implementation of a multi-temporal post-processing algorithm (Hidden Markov Models) that uses multi-year data to reduce spurious land-use change introduced by classification uncertainty in individual years, which has not been achieved by most existing satellite land cover products such as Cropland Data Layer (CDL) and National Land Cover Data (NLCD) (Sulla-Menashe and Friedl, 2018; Abercrombie and Friedl, 2016). To confirm the quality of MODIS products, we

[6]The specific version of data we are using is Land Cover Type Yearly L3 Global 500m (MCD12Q1, collection 6) `https://lpdaac.usgs.gov/products/mcd12q1v006/`

[7]The MODIS (collection 6) land classifications are performed using the random forests model. Over 67 features (including spectral information from MODIS satellite instruments, surface reflectance and temperature, vegetation continuous fields, crop intensity and many others) are used as input data in the classification procedure. The classification results are then validated against ground-truth observations of land covers from over 3000 sites across the world (see Sulla-Menashe et al. (2019) for more details).

extract the quality assurance (QA, scaled from 0 to 100) information for each parcel-by-year observation used in our sample. We find that the overall confidence of land classification is very high, with a mean of 95.6 and a standard deviation of 2.2. Notably, we do not find any significant difference of QA scores between croplands and grasslands, or any evidence that our main variable of interests – weather shocks and price fluctuations – would be correlated with the measurement errors in land classifications provided by MODIS, which leverages the concerns that our estimates might be biased due to potential measurement errors that correlate with explanatory variables. A detailed discussion on the QA information, potential sources of measurement errors, and their implications for our estimates can be found in the Online Appendix A "Uncertainty and Quality of MODIS data".

Insert Figure 1 about here.

Figure 1 shows the percentage of land converted between crop and non-crop use at least once in our study period within non-overlapping 4 km grids. Unsurprisingly, most of the land converted is concentrated at the boundary of the rain-fed region, where the weather conditions, especially precipitation, are more likely to fluctuate across the year and soil quality is relatively lower compared to other regions such as the Midwest.

3.2 US Crop Calendar

To uncover the underlying mechanisms that influence farmers' planting decisions, we split the agricultural production cycle into three distinct stages: pre-planting, planting, and post-planting. We utilize the crop calendar specific to each state in US, obtained from "Field Crop Usual Planting and Harvesting Dates" reported by USDA/NASS (2010). This calendar provides the start and end dates of the planting and harvesting periods for each crop type in each state commonly observed. We use the beginning and ending dates of the major crops planted in each state to identify the most active planting period (typically one month in

the spring, such as April or May). Based on agronomy knowledge, this is the recommended period for seeding and fertilizer application. We then define the "pre-planting" period as the two months right before the most active planting month[8]. Land preparation such as clearing weeds and tillage practices typically occurs during this period. The months right after the most active planting period are therefore defined as the "post-planting" period, including the main growing seasons and harvest time[9]. We exclude any weather information in the "post-planting" period for this study since planting on a specific land is a one-off activity and will not be affected by the weather information in the "post-planting" period.

As noted in the conceptual framework, the weather shocks in the pre-planting period are likely to affect farmers' expectations for weather conditions of main growing seasons or impede farmers' fieldwork progress. For example, a wet April or May will make it difficult for corn planting on time, and hence corn acreage might be switched to other crops or even abandoned (Sacks et al. 2010, Weinraub 2013, Miao et al. 2016). The latter will lead to a reduction in cultivation acreage. The timing of the planting period varies across states in the rainfed region. For example, the most active planting period in Illinois is in April, while that in Texas is in March. We note that this heterogeneity does not affect our estimates after including fine-scale fixed effects at the 4 km grid level. Any remaining weather variations pertain only to the deviations of annual weather conditions from the long-run average within these 4 km grids. In addition, we show that our point estimates are largely unaffected when we assign a uniform crop calendar to all states, or drop the southern states such as Texas and Oklahoma that have earlier most active planting period compared to the Midwestern states and northern states.

[8]For example, if the active planting month is in May, the pre-planting period is therefore defined as March and April. This approach is similar to that proposed in Jagnani et al. (2021).

[9]As we noted above, the planting activities is a one-off activity based on farmers' expectation of productivity and crop price before the planting season. Therefore, the realized weather conditions in the "post-planting" period would not affect planting acreage.

3.3 Crop Prices and Stocks

Price expectation is one of the major factors that drive farmers' planting decisions, as we noted in the conceptual framework. We consider the state-level crop prices received by farmers in the previous year as a proxy for expected prices, following the convention in the literature[10]. Due to the heterogeneity in the cropping structure across states, we construct a Laspeyres price index to reflect the general price level of major commodity crops in each state. Specifically, we obtain the state-level received crop prices and production levels for ten major crops[11] from NASS. The Laspeyres price index is then constructed based upon deflated state-level received prices[12] and production levels for each of ten major crops as follows:

$$P_{s,t}^{\alpha} = \frac{\sum_{l=1}^{10} p_{sl,t} \times q_{sl,t}}{\sum_{l=1}^{10} p_{sl,2001} \times q_{sl,2001}}$$

where $p_{sl,t}$ and $q_{sl,t}$ is the received price and production level of crop l in state s in year t, respectively; and $q_{sl,2001}$ is the production of crop l in state s in the base year 2001. In all our specifications, we use the one-year lagged Laspeyres price index as a proxy for the (aggregated) price expectation for farmers in the current year.

The temporal and spatial variations of Laspeyres mainly come from three sources: (1) common demand or supply shocks for the entire crop market over the years (e.g., trade war, national corn ethanol mandate etc.); (2) heterogeneity in delivery costs of commodity crops across states (Goodwin, Grennes, and Wohlgenat 1990); (3) heterogeneity in cropping structures across states. Notably, the delivery costs and cropping structures may vary sub-

[10]The lagged received prices are widely used in the literature as the proxy for expected prices (e.g., Miao et al. (2016), Li et al. (2019), Cui et al. (2020) and many others). Some studies focusing on corn and soybean acreage also use futures prices for corn and soybean in the current year as the proxy for expected prices. However, futures prices in the current do not exist for all field crops planted on the rainfed US. In addition, previous studies noted that futures prices are highly correlated with lagged received prices, and that futures prices and lagged received prices can be used interchangeably without strong evidence to suggest that one outperforms the other when modelling farmers' expectations for prices (see Miao et al. (2016) for detailed review on this issue).

[11]Ten major crops include barley, corn, soybean, wheat, sorghum, cotton, rice, oats, peanuts, and rye.

[12]The received crop prices are converted to the year 2000 prices by using the GDP implicit price deflator reported by U.S. Bureau of Economic Analysis.

stantially across states within the rainfed U.S., which allows us to identify the effect of price on land conversions, even adding year fixed effects.

We also obtain data for crop stocks for the ten major crops in each state from USDA/NASS. The crop stocks reflect the storage level on December 1st each year. The state-level aggregated stock level of ten major crops are calculated by simply summing up crop-specific stocks.

3.4 Weather Data

We construct the weather variables using grid-level daily weather data provided by Roberts and Schlenker (2009)[13]. The raw daily data are spatially interpolated based on records from weather stations located throughout the United States and provided by the PRISM weather group. We extract the historical daily weather information, including minimum and maximum temperature and cumulative precipitation, over 1970-2017 at the 4 km spatial resolution for the entire eastern U.S. from this dataset.

To capture the rainfall anomaly events, we construct three indices for the pre-planting and planting periods following the climate literature. We first construct dummy variables to indicate whether rainfall anomalies occur in a year, based on the percentiles of cumulative precipitation for the historical rainfall distribution of each PRISM grid cell in each period. Specifically, we define dry (wet) anomalies equal to 1 as years when the cumulative precipitation is at least below (above) the 25% (75%) percentiles for the long-term distribution of each PRISM grid cell. Additionally, we also construct two alternative indices to measure the concentrations of rainfall in the pre-planting period by calculating the number of consecutive wet rainy days (with daily cumulative precipitation larger than 1mm) and the Herfindal-Hirschman index (HHI) of rainfall for each PRISM grid cell, respectively.

For our control variable, temperature, we use the linear and quadratic terms of daily average temperature over the pre-planting and planting periods in our main specification.

[13]We thank Michael Roberts and Wolfram Schlender for making the data and relevant program codes available at `http://www.columbia.edu/ws2162/`

We also construct the cumulative growing degree days (GDD) and use it in our alternative specifications to measure the intensity of daily exposure to a temperature above a lower bound, beneath which cold stress might impede plant growth, and below an upper bound at which heat stress might begin to impede plant growth as well. Following Schlenker and Roberts (2009), we use daily average temperatures to calculate the number of days each pixel is exposed to a temperature above a lower bound (8 °C), and below an upper bound (29 °C)[14], and then sum the daily exposures for each of the three phases during the main growing season for those bounds.

3.5 Biophysical Attributes

To further control for differences in productivity across fields, we extracted a comprehensive array of biophysical attributes for each land pixel within our study region based on the gridded Soil Survey Geographic Database (gSSURGO), the most detailed soil geographic spatial data in the U.S. developed by the USDA National Cooperative Soil Survey (NCSS). The raw gSSURGO data is at 30m spatial resolution. We then resample and match up the raw gSSURGO data with the 500m MODIS land pixels by averaging the biophysical attributes within each pixel.

We follow the convention in the literature and use National Commodity Crop Productivity (NCCPI) and slope as controls in our main specifications. The NCCPI is a rating index representing inherent capacity to produce dryland (non-irrigated) commodity crops[15]. Studies have found that NCCPI value, ranging from 0 to 1, is highly correlated with multi-year mean crop yields. The slope is a key determinant of erosion and is a common control in the literature as well (Wu et al., 2004)

Although these data provide an abstract portrait of productivity, the heterogeneity in

[14]In the robustness check section, we test the sensitivity of our results using alternative upper bounds such as 32°C and 34°C.

[15]We use the weighted NCCPI, which is calculated as the weighted average of NCCPI values for corn, soybean, small grain crops, and cotton. In our robustness check, we also use include NCCPI values for corn, soybean, small grain crops and cotton separately.

biophysical attributes could be high-dimensional and may not be appropriately represented by the NCCPI index and slope. We also extract a number of additional soil quality variables, including soil depth (the average thickness of soil), soil available water storage, soil organic carbon, and pH values in 6 levels of soil: 0-5 cm, 5-20 cm, 20-50 cm, 50-100 cm, 100-150 cm, and $>$ 150 cm, as well as crop root zone depths, available water storage within crop root zone depths, and drought-vulnerable soil-landscape index. In our main econometric specifications, we do not include them since those variables may correlate with NCCPI and slope and hence lead to biased estimates for NCCPI and slope. However, we include them in the machine learning model described in section 4 as high-dimensional controls to explore the potential heterogeneity in effect across biophysical attributes.

3.6 Summary Statistics

Insert Table 1 about here.

Table 1 reports summary statistics from the data sources documented above. The sample comprises a balanced panel of over 177 thousand 4-km grids[16]. Over our study period, the average cropland ratio across grids in the eastern US varied between 0.45 and 0.48. The total amount of land devoted to farming is declined by roughly 1 percent after 2014, which is consistent with the observations from the 2017 USDA Census Report.

Our preferred approach aims to infer the effects of weather fluctuations, especially rainfall shocks across seasons, on agricultural land-use decisions. We focus on regression models that include the 4 km grid and year fixed effects. It would be ideal if, after adjustment for these fixed effects, the variations in the weather variables are still large enough for us to identify this relationship.

[16] As we noted above, the grid-level sample is compiled from a balanced panel of over 10 million 500-mland pixels.

Table S1 reports the magnitude of the deviations between grid-by-year weather realizations and their long-run averages after taking out grid fixed effect (panel A of Table S1) and year fixed effect (Panel B of Table S1). Therefore, it provides an opportunity to assess the magnitude of the variation in pre-planting season precipitation after adjusting for permanent unobserved grid factors (e.g., whether a 4 km grid is typically wet or dry) and common year shocks (e.g., whether a large-scale drought occurs). Table S1 shows a substantial variation even after adjusting this rich array of fixed effects. For example, panel A shows that one standard deviation of dry anomaly and consecutive rainy days, removing long-run average, is over 0.4 and 1.8 (days), respectively.

4 Empirical strategy

4.1 Average Effects

Our objective is to estimate the effect of exposure to weather shocks and price fluctuations on farmers' crop-grass conversion, net of any potentially confounding factors. We model this relationship using the following regression equation:

$$Share_{ist} = \beta_1 Price_{s,t-1} + \beta_2 Rainfall_{ist} + +\beta_3 Temp_{ist} + \lambda_i + \gamma_t + \epsilon_{it} \tag{1}$$

where the dependent variable is the share of cropland for each 4 km grid i in state s at year t. β_1, one of the main coefficients of interest, measures the impact of county-level price fluctuations on the conversions between cropland and grassland. We include the time-varying and spatially-varying aggregated crop price index $Price_{s,t-1}$ as an explanatory variable in the model. Note that prices are measured at the state level, and therefore the estimated coefficient of β_2 represents the average effect of prices on the share of cropland at the state level.

We control for weather conditions during the pre-planting season, as these may affect

crop yield expectations and influence land use decisions. Our weather controls include both linear and quadratic terms for average temperature during the pre-planting period. To capture localized precipitation variability, we construct two dummy variables—dry shock and wet shock—based on whether cumulative precipitation during the pre-planting season falls in the bottom 25% (dry shock) or top 25% (wet shock) of the historical precipitation distribution for that location. We also test the robustness of our results using alternative measures of precipitation concentration, including the number of consecutive rainy days and the Herfindahl-Hirschman (HH) index of precipitation distribution.

Our estimates also include PRISM (4 km) grid λ_i and year γ_t fixed effects. The grid fixed effects control for unobserved time-invariant confounding factors at the grid level, such as location-specification determinants of productivity, such as climate and soil quality, and other time-invariant determinants of input use like the proximity to markets[17]. Year fixed effects control for any common productivity shocks throughout the rainfed region in each year. It also controls for prices fluctuations in national markets such as livestock and physical inputs (e.g., fertilizer and chemical uses).

There are several reasons to believe that within-region spatial correlation and serial correlation could influence the share of cropland in equation (1). First, observations in each grid across different years might have some serial correlations due to the unobserved time-varying factors, which cannot be controlled by the grid fixed effect in our specification. Second, spatial correlation might affect the share of cropland in grids nearby due to similar geographical, climatic, and socioeconomic factors shared by grids nearby. Therefore, we implement a two-way clustering method to simultaneously account for spatial correlation and serial correlation (Cameron and Miller, 2015). All the regressions are estimated by clustering the standard errors at the PRISM cell and year-by-Crop Reporting District (CRD) level[18].

[17]A potential concern is that the inclusion of fixed effects at this fine scale level could absorb a significant amount of weather variance and amplify measurement error (Fisher et al. 2012, Auffhammer and Schlenker, 2014). However, as we note in the data section, we find that there is still relatively large weather variation, especially for the precipitation, even after including a rich set of fixed effects.

[18]Crop Reporting Districts (CRD), or Agricultural Statistics Districts (ASD), are defined as groups of counties in each States by geographic (i.e., soil type, terrain, and elevation), climate (mean temperature,

A key econometric issue for estimating β_2 in equation (1) is the endogeneity of the price variable. Due to the biological lags in agricultural production, unobserved time-varying factors such as persist supply shock and policy shocks at the subregion level might simultaneously affect farmers' land conversion decisions in the current year and lagged received prices in the last year. For example, an expansion of a land conservation program (such as the Conservation Reserve Program) in a state may reduce available planting acreage while raising the received commodity price, leading to a biased estimate of the price effect in equation (1). To address the potential endogeneity issue of crop prices, we adopt an instrumental variable approach in the spirit of the competitive storage theory described in Roberts and Schlenker (2013). Specifically, a higher level of crop storage in year $t-1$ may increase the crop supply (lagged storage plus new production) in year t, and hence, reduce the received crop prices in year t, which might affect farmers' crop-grass conversion decisions in year $t+1$. We, therefore, use state-level crop stocks as an instrument for the state-level price. The specification for our first stage is:

$$Price_{s,t-1} = \gamma_1 Stock_{s,t-2} + \gamma_2 Rainfall_{ist} + \gamma_3 Temp_{ist} + \lambda_i + \gamma_t + \epsilon_{it} \tag{2}$$

where $Stock_{s,t-2}$ is a measure of aggregated stock levels for ten major crops in each state discussed above. Other variables are defined as the same in equation (1).

The identifying assumption of our instrumental variable is that, after controlling for common supply shocks across the rainfed regions, the crop stocks will cause changes in current year land conversions only through the channel of lagged received crop price. This is likely to be held in our case for several reasons: (1) local crop stocks are largely affected by the local random weather shocks in each year, which will cancel out each other across regions; therefore, the aggregated crop stock at the state level is less prone to be autocorrelated; (2) there does not appear to be any evidence to suggest that the planting decisions of individual price-taking farmers will depend on stock level at the state-level.

annual precipitation, and length of growing seasons), and cropping practices of counties in each State.

4.2 Treatment Effect Heterogeneity

The fine-scale land use data and the rich set of characteristics in the biophysical data provide an excellent opportunity to thoroughly measure and explain the heterogeneity in response to weather shocks across the land with different soil quality and productivity. We pursue a two-pronged approach. First, we extend equation (6) and (7) to a regression specification at the 500 m pixel. This analysis allows us to gauge the potential heterogeneity in effects of weather shocks and price fluctuations on crop-grass land conversions across land parcels with different productivity, as we noted in section 2. We use the land-quality class created by soil quality index (NCCPI) to proxy for high-dimensional biophysical attributes and split our full sample into 10 equal subsamples. We then estimate a similar specification as the one used in section 4.1 for each subsample:

$$\mathbf{1}(\textbf{Cultivated})_{ijst} = \theta_1 Price_{s,t-1} + \theta_2 Rainfall_{ist} + \theta_3 Temp_{ist} + \lambda_i + \gamma_t + \epsilon_{it} \qquad (3)$$

Here the dependent variable is a dummy variable indicating whether the landowner on 500 m land pixel j in 4 km grid i and state s choose to cultivate this pixel (=1) or keep it under grassland (=0) at year t. Other variables are defined as the same in equations (1) and (2). This subsample allows us to explore whether farmers will adjust parcels with relatively low productivity more frequently, especially when the adjustment could be nonlinear across groups[19]. Similarly, the major parameters of interest in equation (4) are the coefficients ϕ_1 and ϕ_2, which indicates the heterogeneous effects in farmers' crop-grass conversion responses to weather shocks and price index, respectively.

Notably, the effectiveness of this subgroup analysis strategy, for example, how much heterogeneity can be captured by this specification, rely heavily on the baseline characteristics such as NCCPI and slope identified by the literature. This method, however, may overlook

[19] As we noted in section 2, the land use on parcels with highest and lowest quality might be stable while the parcels with medium quality (at the margin of agricultural production) are more likely to switch their usage in response to productivity and price shocks.

relevant heterogeneity if there exist additional biophysical factors that matter for the planting decisions but has not been incorporated by our extended specifications (equation 9), which is very likely to be the case since biophysical attributes are typically high-dimensional and affect the agricultural activities through complicated channels (Baylis et al. 2021).

Our second strategy employs a method recently developed by Chernozhukov et al. (2018) (henceforth, CDDF) for estimating heterogeneous effects using machine learning techniques. Compared to the first strategy, CDDF can extend the analysis in heterogeneity to a high-dimensional set of biophysical attributes instead of a few baseline characteristics. This method is valid in high-dimensional settings, allowing us to include a rich set of characteristics in unspecified functional forms. It also allows us to statistically test the existence of heterogeneity effects, construct a reliable causal inference for the heterogeneous effects, and identify the source of heterogeneity, namely, which of the covariates are relatively important in heterogeneity analysis[20].

We focus on the heterogeneous effect of rainfall shocks and price fluctuations on the crop-grass conversion decisions of each individual land parcel. As we denoted in our main specification, the main variables of interest are price index and two dummy variables that indicate whether the cumulative precipitation in the pre-planting period is at least 25% (75%) lower (higher) than those in the historical distributions. Since the CDDF method requires us to identify the "treatment effect" and only allows one binary treatment variable in the analysis, we then partition our parcel-year observations into three "treatment" groups and one "control" group based on (1) whether a dry shock occurs ($\mathbf{1}_{\mathbf{DryShock}}$) (2) whether the percentile of the predicted price index is above the medium value ($\mathbf{1}_{\mathbf{HighPrice}}$)[21]. The

[20]This is the key distinction between the CDDF and naive/generic machine learning algorithms applied in the literature (Chernozhukov et al. 2018). While naive machine learning algorithms are successful in prediction empirically, it is typically difficult to obtain uniformly, and valid causal inferences in high dimensional settings; generic machine learning algorithms may not even produce consistent estimates of the conditional average treatment effects (CATE) if absent strong assumptions. The CDDF method overcomes this issue by focusing only on the presence of heterogeneity, the group average treatment effects (or average treatment effects by heterogeneity groups), and the average characteristics of the most and least affected units.

[21]We estimated the predicted price index using the first-stage specification presented in equation (8). Three treatment groups include (1) observations experiencing dry shocks and above-medium predicted crop prices; (2) observations experiencing dry shocks and below-medium predicted crop prices; (3) observations

group of central interest is two "treatment" groups with observations that are experiencing dry shocks ($\mathbf{1}_{\mathbf{DryShock}} = 1$) and with high/low crop price levels ($\mathbf{1}_{\mathbf{HighPrice}}$=1 or 0). By comparing these two "treatment" groups to the control group ($\mathbf{1}_{\mathbf{DryShock}} = 0$ and $\mathbf{1}_{\mathbf{HighPrice}} = 0$) respectively, we can then isolate the effect of dry shocks and crop prices on the crop-grass conversion probability on each land parcel. This leaves us two samples for the machine learning exercises; each of them contains one treatment group of central interest and the control group. In the result section, we present two sets of results based on these two treatment variables. We also include all control variables used in equation (1) and a high-dimensional array of soil quality attributes (see more detail in the data section). To reduce the computational burden, we are unable to include a fine-scale grid fixed effect in our machine learning model. Instead, we replace it with the county fixed effect[22].

Next, we split each of the two samples randomly into an "auxiliary sample" (used to estimate the prediction model) and a "main sample" (used to estimate the regression model). We tune and train machine learning models by using the auxiliary sample. The training process is separate for observations in the treatment group (places where rainfall anomalies occur) and control group (places with normal rainfall), resulting in two "best-tuned" prediction models. We then apply these predictions models in the main sample to predict the planting probability of individual land pixels in the treatment group (denoted as $\mathbf{1}(\mathbf{Cultivated})^T$) and the control group (denoted as $\mathbf{1}(\mathbf{Cultivated})^C$). The difference between these two predicted probabilities $S = \mathbf{1}(\mathbf{Cultivated})^T - \mathbf{1}(\mathbf{Cultivated})^C$, is referred as the proxy predictor of (true) conditional average treatment effect. Valid causal inference about the key features of this conditional average treatment effect, such as the best linear predictor of this effect, the presence of heterogeneity in effect, and characteristics associated with the most and least affected observations can be identified. The detailed explanations

do not experience dry shocks but have above-medium predicted crop prices. The control group includes observations that do not experience dry shocks and have below-medium predicted crop prices.

[22]Even with a small subsample with county and year fixed effects and 1 million observations, the CDDF requires around 100 CPU hours (20 CPU cores×5 hours) and over 200 GB memory on a decent cluster. The estimation time and memory requirement will be ten times more if we use 4 km grid fixed effects instead of county fixed effects.

for the CDDF and the algorithm we used in this analysis can be found in Online Appendix B). To estimate the machine learning model on a sample this large (over 115 million pixel-by-year observations), we split the sample evenly into 100 subsets, each containing 1% (1.15 million) of the observations. The machine learning estimation process is applied to each of these 100 subsets and then takes the appropriate average of the coefficients and standard errors.

5 Results

5.1 Average Effects

Insert Table 2 about here.

Table 2 reports our baseline estimates of the average effect of rainfall shocks and crop price index, using equation (1). Columns (1) and (2) report the estimated effect of rainfall anomalies in the pre-planting and planting period and crop price index on the share of cropland, respectively, in separate regressions. In column (3), we include both rainfall shocks and crop price index. In columns (1)-(3), we do not address the endogeneity issue of the crop price index. In column (4), our preferred model, we utilize the lagged stock of ten major crops in each state as the instrumental variable to address the endogeneity issue of the crop price index. The associated first-stage results are shown in column (5). All regressions in Table 2 include a set of controls including linear and quadratic terms of average temperature and 4 km grid fixed effects as well as year fixed effects. For all models that involve instrumental variables (column 4), the p-values of weak identification test (Kleibergen-Paap wald rk F statistics) are much larger than 10 in all cases, indicating the lagged stocks are valid instrumental variables, which is consistent with the literature (Robert and Schlenker, 2013; Miao et al. 2016; Li et al., 2019).

Column (1) - (3) of Table 2 allow us to examine the presence of omitted variable bias when the crop price or weather variables are excluded as determinants of crop-grass conversions over our study period. Comparing column (2) with column (1), we find that the inclusion of price effect (column 2) will lead to a slightly negative estimated coefficient for the dry shock (increased by 6%) and a more positive estimated coefficient for the wet shock (increased by 12.5%), indicating a potential substitution effect between weather conditions and price when farmers form the expectations of profitability. Additionally, by comparing results under columns (3) and (4), we find that ignoring the endogeneity of price variables will substantially attenuate the estimated coefficient of price towards zero, thus underestimating the true underlying effect at least by 46% = ((0.0183-0.00988)/0.00988).

Holding all else being equal, we find that crop price index and dry shock significantly affect farmers' crop-grass conversion decisions. A one-unit increase in the price index, which represents about one standard deviation of the price index, as is shown in Table 2, will increase the share of cropland in the 4 km grid by about 0.0183 (Column 4 of Table 2). The sample mean of share of cropland in 4 km grids is 0.54 over the entire eastern US. These results indicate a price effect of 3.3% (=0.0183/0.54) or a price-acreage elasticity of 0.068. This is slightly smaller but close to the price-acreage elasticities (0.07-0.08) reported in previous studies that do not distinguish between grassland and other non-cropland (Schlenker and Roberts, 2013; Miao et al. 2016; Li et al. 2019). This finding indicates that most of the supply response from the extensive margin, i.e., an expansion of cropland area, comes from the grass-to-crop conversion rather than other non-croplands such as idled land.

The estimated effect of dry shock (cumulative precipitation is at least 25% lower than the historical precipitation distribution) significantly decrease the share of cropland by 0.8% (=-0.00451/0.54) while that of wet shock (cumulative precipitation is at least 75% higher than the historical precipitation distribution) are generally statistically insignificant[23]. Notably, in

[23]Regarding the magnitudes, the wet shock slightly increases the share of cropland in 4 km grids by 0.2%. However, these results should be interpreted with cautions due to the statistical insignificance. In the following section, we show that the estimated effect of high precipitation may be sensitive to the model specification and potential spatial heterogeneity.

our data, the change in the state-level crop price index in consecutive two years is typically below or around a half unit, as shown in Figure S1. By this metric, the price effect we reported above appears a factor of two times larger than the effect of dry shock and seven times larger than the effect of wet shock.

Insert Table 3 about here.

Timing. – Do the effects and responses to rainfall shocks vary according to the time at which extreme precipitation events are experienced within the pre-planting period? Answering this question is relevant to understanding of the average effects we obtained above. As noted in related studies (i.e., Aragon et al. 2021), farmers may be more sensitive to rainfall shocks at some stages of agricultural production cycle, or face time-varying constraints to determine crop/grass uses on their lands.

We then examine the sensitivity of our main results to the timing of rainfall shocks within the pre-planting period. In column 3 of table3, we break down the pre-planting period into two stages – preparation stage (two months right before the planting activities) and planting stage (the month when planting activities are undertaken) – to examine whether the response to rainfall shocks varies according to the time at which rainfall shocks occurs. We find that the precipitation effect at the planting stage is slightly larger than that at the preparing stage. Interestingly, we also find that farmers' crop-grass conversion decisions would be significantly affected even if rainfall shocks only occur at one of these two stages (i.e., a dry shock event happens at the preparation stage while the precipitation condition at the planting stage is normal). For instance, a dry shock that occurs only at the preparation stage leads to a reduction of 0.3 percentage points in the share of cropland (column 3, table3), which is about one-third of the estimated effect of a prolonged dry shock that lasts for the entire pre-planting period (column 3, table3).

Insert Figure 2 about here.

Short-term Effects versus Long-term Effects. –Due to the biological lags in agricultural production, farmers may determine the crop-grass conversions based on weather conditions in previous seasons instead of the contemporaneous shocks; this potential dynamic link between weather shocks and medium-run/long-run outcomes may not be captured by our main specification which focus on the contemporaneous weather shocks only. In figure 2, we present the cumulative effects of rainfall shocks from the distributed lag model (equation 3), calculated by summing the respective rainfall variables and their lags. For the purpose of comparison, we begin by repeating the estimates with no lags, which is our main specification, and then add lags of rainfall shock variables up to ten years.

We find that generally , long-term effects for dry shocks are not significant; the cumulative effect of dry shock is only statistically significant with no lag or one lag and then becomes insignificant as more lags are included. A contemporary dry shock reduced the share of cropland by 0.8 percentage points with no lags, as we reported above. With one lag included, the cumulative effect is a reduction of 0.2 percentage points. The cumulative effect is no longer statistically significant when we include three lags or more. The change in the signs of individual coefficients of the first two individual lags indicates that the effect of the dry shock on crop-grass conversion appears to be reversed once the shock disappears, rather than being persistent. The cumulative effects of wet shock, however, remain positive with fairly stable point estimates around 0.4 percentage points with first two lags, and then substantially increase in the magnitudes as more lags are included. For instance, a persist wet shocks in the past 5 years[24] will increase the share of cropland by 1.7%, indicating the effects of wet shocks appear to be persist in the medium/long term, rather than being reversed. We interpret these findings as suggestive evidence that the grass-to-crop conversion is more likely

[24]The cumulative precipitation in the pre-planting period remain in the top 25 percent of the historical distribution of precipitation at this region.

to be affected by the medium/long-term climate conditions instead of contemporary weather shocks, while the crop-to-grass conversion is likely to be driven by the contemporary weather shocks[25].

Economic Implications of the Estimates Effects. – We utilize the estimates from our main specification (column 4 in table 2, our preferred main specification) to compute changes in planting probability due to the changes in state-specific crop price index and local weather shocks across different agricultural production periods. Specifically, to calculate the changes in the share of cropland in each 4 km cells from 2002 to 2016, we use the coefficient of dry anomaly (-0.00431) and wet anomaly (0.00138) and multiply it with the change in local precipitation (whether it is 25% (75%) lower (or higher) than the average precipitation in the historical distribution) to obtain cell-specific changes in acreage over this period. The same procedure is applied when calculating cropland changes due to the fluctuations of aggregated price index. The sum of changes in each cell provides an estimate of the change due to the local weather shocks and price changes across the entire rainfed region, holding all else constant.

Our results show that, on average, the fluctuations in aggregated crop prices, *ceteris paribus*, leads to an 0.6-million-ha net increase in cropland converted from grassland in the period of 2002-2016, whereas the dry shock leads to a 0.3-million-ha net decrease in cropland acreage in the same period. This finding indicates that the impact of aggregated price is partly surpassed by the effect of weather shocks such as the dry shocks in that period.

[25]There are, however, several important caveats. First, we are unable to disentangle the effects of rainfall shocks on different types of crops (corn, soybean, sorghum etc) and grass (native/natural vegetation, managed pastureland for grazing etc). One might concern that the grass-to-crop conversion, especially those driven by the medium/long-term precipitation conditions, happen more frequently on permanent grasslands (native/natural vegetation) while the crop-to-grass conversions, occur more frequently on managed grasslands such as pasture/grazing lands. This would not bias our estimates but could lead to slightly different interpretations for crop-to-grass conversions and grass-to-crop conversions. In the robustness check section, we provide suggestive evidence to show that both conversions occurs more frequently on the permanent grasslands rather than managed grasslands. That being said, we can rule out that our results are only reflecting the less sensitivity of managed grasslands to weather shocks.

5.2 Permanent Grasslands versus Managed Grasslands.

Due to the limited land classifications provided by MODIS products, we cannot distinguish between permanent grasslands and managed grasslands. The permanent grasslands are covered by natural and semi-natural perennial vegetation while the managed grasslands might be replanted frequently. The conversions between these two grassland types and cropland could lead to substantial difference in economic and environmental consequences[26], and therefore different policy implications.

Here we augment our main analysis by utilizing the Rangeland Analysis Product (RAP), the only existing fine-scale dataset that provides additional information on grassland types, such as annual grasses, perennial grasses, and shrubs, over the period of 1984 to 2019[27]. We extract the grass type information from 1984 to 2000 from the RAP dataset and match up with the 4 km grids in our sample. Since the RAP data product only covers the Great Plain states[28] in our main sample, this will restrict our sample and drop almost half of the observations. However, since these Great Plain states are the places where crop-grass conversion occurs most frequently, a majority of variations of crop-grass conversion is still kept in the remaining sample. We then construct a "historical grassland status" index by calculating the average percentages of permanent grasses over the period of 1984-2000 for each 4 km grid. The higher the value of "grassland status", the more likely the grid is covered by permanent grasses instead of managed grasses.

[26]For example, the permanent grassland generally has relatively lower productivity and higher ecosystem services.

[27]The RAP is developed by USDA's National Resources Conservation Services (NRCS) and the university of Montana. The estimates were produced by combing approximating 60,000 field plots from the Bureau of Land Management (BLM) Assessment, Inventory, and Monitoring (AIM) datasets and NRCS National Resources Inventory (NRI) along with the historical high-frequency Landsat satellite records. These input data are incorporated with a biophysical model to determine the vegetation cover types across the western half of the U.S. at the 30 meter resolution. See https://doi.org/10.1016/j.rama.2021.04.003 for more technical details.

[28]North Dakota, South Dakota, Nebraska, Kansas, Oklahoma, Texas

We replicate the main specification using the observations with "grassland index" information. Unsurprisingly, the estimates of price fluctuations and rainfall shocks obtained from the new sample are slightly larger than those from the full sample (Columns (1) and (2) of Table S4) due to the higher frequency of crop-grass conversions in the Great Plain states. We then split the new sample into two portions: (1) lands dominated by permanent grasses (percentage of permanent grasses larger than the medium value of "historical grassland status"–56%); (2) lands dominated by annual managed grasses (percentage of permanent grasses less than the medium). We also relax the main specification and include interaction relationship between "historical grassland status" and our main variables of interests, which allows us to statistically test the difference in price and weather effects across different types of grasslands. The results shows that both price fluctuations and dry shocks appear to have larger impacts on crop-grass conversions on lands that was dominated by permanent grasses historically (Column 3 and 4 of Table S4). While the difference in weather effects are not statistically significant, we do find supporting evidence that the change in price index are more likely to induce grass-to-crop conversions on land that was dominated by permanent grasslands (Column 5 of Table S4).

A possible interpretation of these findings is that the crop-grass conversion activities are restricted by the opportunity costs, which are determined by their previous land use. For managed grasslands, many of them are well-maintained for livestock production and may provide competitive revenue compared to those from annual crop production. Oppositely, the permanent grasslands can barely provide much economic revenues; the positive externality of these lands, such as recreation functions and ecosystem services, cannot be easily internalized due to the lack of corresponding markets. Therefore, farmers or landowners have stronger incentive to convert these land when crop prices are high[29].

[29]We would like note that there might be, however, other important factors that drive the different responses to price effects on permanent grassland and managed grassland, including but not limited to, conversion costs, owners' preferences etc. Therefore, we interpret these results with caution as only suggestive evidence.

5.3 Heterogeneity Effects across Biophysical Attributes and Regions

5.3.1 Regression Method

The analysis presented in the previous section relies on data aggregation at the 4 km grid level. In this section, we extend the main specification and utilize the variations of biophysical attributes across 500 m land cover parcels within each 4 km grid to further study the potential heterogeneity in effects of weather shocks and price fluctuations across fields.

As described in section 4.2, we divide the full sample into nine subsamples based on NCCPI index and estimate equation (8) separately for each subsample. This allows us to explore the marginal responsiveness of crop-grass conversion to weather shocks and price fluctuations across land with different quality. Figure 3 reveals that crop-grass conversions in medium quality groups (with NCCPI between 0.3-0.6) are more frequent than those in lands at the bottom or top of the quality distribution. Holding all else equal, for each unit increase in price index (around one standard deviation), the grass-to-crop conversions in the medium quality groups increase by 5.6%-7.4% percentage points, implying a price-acreage elasticity lies in the range of 0.12 to 0.16, while the corresponding effects on the bottom or top quality groups are statistically indistinguishable from zero and have small magnitudes. Similarly, dry shocks have the largest effects on the medium quality groups as well; an occurrence of dry shock would lead to an increase in crop-to-grass conversions by 0.9%-1.3%. These results echo the agronomy discussion on land at the margin of agricultural production (see the review in Khanna et al. 2021) and could be interpreted as evidence that farmers' crop-grass conversion decisions depend on soil quality. However, we also would like to note that these confidence intervals of most of quality groups are overlapping with each other and statistically indistinguishable; this noisy results indicate that there might be other important factors that drive farmers' crop-grass conversion decisions.

Insert Figure 3 about here.

5.3.2 Machine Learning Method

As we discussed above in section 4, if there are factors that affect the farmers' planting responses to rainfall shocks and price fluctuations even after conditioning on biophysical indices such as NCCPI, then the above analysis may miss meaningful heterogeneous effects, which are relevant and important for policy implications. In this section, we investigate this possibility by utilizing the method developed by CDDF to exploit the high-dimensional set of biophysical attributes[30]

Insert Table 4 about here.

Table4 presents the best linear predictor of the average treatment effect of being assigned to different treatment groups (columns 1 and 3 of table4) and statistical testing on heterogeneity in effects. The average effects reported in column 1 and 3 indicates that being assigned to the treatment group with dry shock and high price index does not significantly affect the cultivation probability of an individual land parcels, while being assigned to the treatment group with dry shock and low price index significantly reduce the cultivation probability of a land parcel by 1.1% (=-0.006/0.54), which are consistent with reduced-form estimates reported from our main specification. Columns 2 and 4 report estimates of heterogeneity loading parameters in the BLP. The significance of coefficients indicates that we can strongly reject the null hypothesis that effects of rainfall shocks are homogeneous and statistically indistinguishable from the average effect across land with different biophysical attributes.

[30]Following the recommendation in literature, we mainly rely on the extreme gradient decision trees algorithm (XGboost) to improve the efficiency of the estimation procedure.

Insert Figure 4 about here.

Figure 4 visualize the heterogeneity in effects across groups of observations with rainfall shocks. The groups of observations are divided based on percentiles of their predicted conditional average treatment effects . We find a negative and statistically significant effect for land pixels in the bottom 20 percent of the distribution under the dry shock regardless of the level of crop prices. For instance, the cultivation probability of land parcels in bottom 5 percent reduces by around 10% when experiencing a dry shock in the pre-planting period, which is almost ten times larger than the average effect from our main specification. Similarly, the corresponding effect for the group with treatment effects falling in the 5 to 20 percentiles is also profound and about four times larger than the average effect reported. This indicates a substantial vulnerability of agricultural production in these regions under the dry shock[31]. The estimated treatment effects for those falling in the 20 to 80 percentiles, however, are either insignificant or marginally small and close to zero. The estimated effect becomes larger and positive for observations in the top 20 percent of the distribution, experiencing around a 1% increase in their cultivation probability.

Insert Table 5 about here.

We future investigate potential factors that may drive the pattern of heterogeneous treatment effects that we find in the above analysis. We compare the 22 different biophysical characteristics of the bottom 5 percent and middle 60 percent groups and find significant differences for each one (Table 5). For example, the bottom 5 percent group has at least 20 percent lower organic matter in all layers of soil, 10 percent less water capacity, 5 percent less thickness of soil, and is more likely to have steeper slope. While the heterogeneity associated with these characteristics cannot be causally interpreted, it still shed light on the potential

[31]Notably, although the points estimates are slightly smaller under the high price regime, the confidence interval of these estimated effects are overlapping with each other and statistically indistinguishable.

important biophysical attributes that might be considered in farmers's land use decisions and implied the importance of incorporating these indicators into agricultural policies such as crop insurance to encourage farming methods to build soil health and resilience against extreme weather events (Woodard et al. 2017). Interestingly, despite their importance on crop growth and ecosystem services although widely being used in the crop modelling in agronomic literature, some features of soil, such as soil PH values, thickness of soil, maximum depth of root, do not statistically different across groups. These features are closely related to the ecosystem services such as soil erosion and nitrogen leaching but seems to be seldomly considered in farmers land use decisions[32]. This finding suggest that farmers would only consider a small set of biophysical attributes and may not internalize the potential ecosystem consequence in their land use decisions[33].

5.4 Robustness

Quality Assessment. – As discussed in the data section, one might concern the quality, especially the classification accuracy, of the satellite data product such as MODIS. We take the advantage of the QA scores provided by the MODIS product and use it as an indicator to identify the quality of each parcel-by-year observations. We replicate our main specifications in a "high quality" sample that only include land parcels with QA scores higher than or equal to 95 (the mean of QA score in our sample) in any of the years within our study period. This drop around xx percent of the baseline sample, but reduce concerns of measurement errors due to the potential misclassification between cropland, grassland and other land cover types. We obtain similar estimates as those from the main specifications using the restricted sample (Table S3).

[32]There are two potential explanations: (1) these attributes are difficult to be measured and not easily to interpret by individual farmers without proper instruments or consultations with agronomists; (2) individual farmers do not internalize ecosystem services in their cultivation decisions and therefore may undervalue these attributes. Due to the lack of farmer-level information, we are unable to test and rule out any of these explanations here. We leave it for future research to provide more generalized evidence on this issue.

[33]This finding echoes the recent discussion in the literature on the absence of ecosystem consideration in individual farmers' land use decisions. Khanna et al. (2021) provides a detailed review of the findings in related studies.

Alternative Rainfall Measures. - Different precipitation measures may reflect different aspects of rainfall effects on agricultural land use decisions. To evaluate the sensitivity of our main results, we examine alternative measures of rainfall shocks constructed using (1) different percentile thresholds, (2) the number of consecutive rainy days, and (3) the Herfindahl-Hirschman Index (HHI) of rainfall distribution in the pre-planting period. Overall, we find that the magnitude and signs of estimates from these alternative specifications remain consistent with those reported in our main results. For comparison purposes, we repeatedly report the estimates from our main specifications in column (1) of Table S5. Column (2) of Table S5 decomposes the dry and wet shocks into more refined subcategories to examine potential heterogeneity in precipitation effects across different percentile thresholds. We find slightly larger changes in cropland share for fields experiencing extreme rainfall events. Specifically, fields with pre-planting precipitation in the bottom 5% of the historical distribution experience a 1.1 percentage point reduction in cropland share, while fields with precipitation in the top 5% experience a 0.7 percentage point increase. However, the signs of these estimates remain consistent with our main specification.

Table S6 presents the effects of rainfall duration and concentration on farmers' planting decisions. In columns (2) and (3), we replace the rainfall shock indices based on cumulative precipitation with linear and quadratic terms of consecutive rainy days and the HHI index. We calculate marginal effects of consecutive rainy days and the HHI index at median values (reported in Table S6) and across the entire distribution of consecutive rainy days (shown in Figure S2). Although the magnitude of these estimates is not directly comparable to our main specification, the general patterns—particularly for consecutive rainy days—remain consistent. Holding all else equal, at median values across the eastern United States, a higher number of consecutive rainy days leads to a slight increase in cropland share, reinforcing our primary findings about the relationship between precipitation patterns and land use decisions.

Additionally, the effect of a consecutive rainy day can be negative if there are prolonged

consecutive rainy days (i.e., more than seven days), especially in the planting period when it is close to farmers' fieldwork activities. In column (4) and (5) of Table S6, we also interacted the number of consecutive rainy days and HHI with the wet shock index to test whether the characteristic of high precipitation events, such as the duration or concentration of rainfall, will affect the estimated coefficients we obtain from the main specification. We find the interaction terms are generally insignificant, indicating the estimated effects of dry or wet shocks in these augmented specifications are qualitatively consistent with those from the main specification.

Temperature Controls. – We also investigate the alternative formulations of the temperature variables as the controls in our main specification. First, we construct a set of degree days (number of days with average temperature within 3° C range) for the pre-planting period. Second, we construct the growing degree days (GDD) using a sinusoidal interpolation between the daily minimum and maximum temperatures for days in the pre-planting period[34]. These two specifications produce similarly sized impacts (column (2) of Table S7). We also obtain similar results when allowing for different GDD thresholds (columns 3-5 of Table S7).

Panel Specifications and Clustering Choices. – We find that the results are broadly consistent across a range of alternative panel specifications. In all of the specifications presented in Table S8, we keep control variables and grid and year fixed effect as we reported in our main results. Columns (2) and (3) of Table S8 reports the results when we progressively add state-specific linear time trends and region-by-year fixed effects in our regression specification, respectively. This rich set of interactive fixed effects and linear trends can further control for any potential omitted variables such as the technological changes and/or policy shocks at the state or region level. These estimates are similar to those from our

[34]See more technical details on GDD construction in D'Agostino and Schlenker (2016) and Roberts and Schlenker (2009)

main specifications in terms of the magnitudes but less precise .Our results are also robust to alternative clustering choices (Table S9)

6 Carbon Sequestration, Food Security, and Welfare Implications

7 Conclusion

We estimate the effect of local weather shocks and changes in price on land conversion between crop and grass uses with a fine spatial resolution land use history data for the rainfed US from 2001 to 2017. We find that a 1% increase in state-level aggregated price index leads to an increase in grass-to-crop conversions by 0.06%. This indicates the crop-grass conversion, which is largely ignored in national and regional agricultural policies, is the main channel of extensive margin supply response to price fluctuations. The estimated effect of weather shocks is slightly lower but still comparable to the effect of price. An occurrence of a contemporary dry shock in the pre-planting season could lead to a 0.8% increase in crop-to-grass conversion. While the contemporary wet shocks does not significantly affect crop-grass conversions, a climate-induced wet shock that is persist for five years can lead to an increase of grass-to-crop conversion by 1.1%. The estimates of treatment effect heterogeneity further reveal a substantial heterogenous response of crop-grass conversion to rainfall shocks across the rainfed region. Of the most vulnerable group (bottom 5 percent), the effect of a contemporary dry shock on their crop-grass conversion probability is almost ten times larger than the average effect estimated, regardless of the price level, on the most vulnerable group of land (5 percent) that is associated with lower storage capacity, less productivity, and/or steeper slope. This implies the importance of incorporating these indicators into agricultural policies such as crop insurance to encourage farming methods to build soil health and resilience against extreme weather events.

Our findings indicate that the change in the use of land for crop production may not be inelastic and static even in the short term. Previous studies using aggregated-level data may fail to reveal the substantial fine-scale changes and underestimate the effects of extreme weather events on individual farmers' land-use decisions, especially for land at the margin of crop production that is vulnerable to extreme weather and price fluctuations, which has been overlooked in the literature. We acknowledge that the fine-scale land use decisions might also be affected by other important factors such as demographic characteristics, culture, farmers' own risk, time, and aesthetic preference (Eaton et al., 2018, 2019; Khanna et al. 2021). Incorporating the satellite-based land use data and micro-data on farmers' decisions (i.e., farm-level survey, large-scale administration data) could provide a more comprehensive understanding of the heterogeneous effects of extreme weather and price fluctuations beyond the biophysical attributes of land discussed in this study. We leave it to future research to quantify the heterogeneity across these factors.

References

Abercrombie, S. P. and M. A. Friedl (2016). Improving the Consistency of Multitemporal Land Cover Maps Using a Hidden Markov Model. *IEEE Transactions on Geoscience and Remote Sensing 54*(2), 703–713.

Aragón, F. M., F. Oteiza, and J. P. Rud (2021, feb). Climate Change and Agriculture: Subsistence Farmers' Response to Extreme Heat. *American Economic Journal: Economic Policy 13*(1), 1–35.

Bai, Y. and M. F. Cotrufo (2022). Grassland soil carbon sequestration: Current understanding, challenges, and solutions. *Science 377*(6606), 603–608.

Bardgett, R. D., J. M. Bullock, S. Lavorel, P. Manning, U. Schaffner, N. Ostle, M. Chomel, G. Durigan, E. L. Fry, D. Johnson, J. M. Lavallee, G. Le Provost, S. Luo, K. Png, M. Sankaran, X. Hou, H. Zhou, L. Ma, W. Ren, X. Li, Y. Ding, Y. Li, and H. Shi (2021, sep). Combatting global grassland degradation. *Nature Reviews Earth & Environment 2021 2:10 2*(10), 720–735.

Burgess, R., M. Hansen, B. A. Olken, P. Potapov, and S. Sieber (2012, nov). The Political Economy of Deforestation in the Tropics*. *The Quarterly Journal of Economics 127*(4), 1707–1754.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernández-Val (2018, jun). Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India. Technical report, National Bureau of Economic Research, Cambridge, MA.

Costinot, A., D. Donaldson, and C. Smith (2016). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. *Journal of Political Economy 124*(1), 205–248.

Cui, X. (2020, may). Climate change and adaptation in agriculture: Evidence from US cropping patterns. *Journal of Environmental Economics and Management 101*, 102306.

Dass, P., B. Z. Houlton, Y. Wang, and D. Warlind (2018, jul). Grasslands may be more reliable carbon sinks than forests in California. *Environmental Research Letters 13*(7), 074027.

Hendricks, N. P., A. Smith, and D. A. Sumner (2014, oct). Crop Supply Dynamics and the Illusion of Partial Adjustment. *American Journal of Agricultural Economics 96*(5), 1469–1491.

Huang, H.-H. and M. R. Moore (2019). Farming Under Weather Risk: Adaptation, Moral Hazard, and Selection on Moral Hazard. In *Agricultural Productivity and Producer Behavior*, pp. 77–124. University of Chicago Press.

Jiang, C., K. Guan, M. Khanna, L. Chen, and J. Peng (2021). Assessing marginal land availability based on high-resolution land use change information in the Contiguous United States.

Lark, T. J., N. P. Hendricks, A. Smith, N. Pates, S. A. Spawn-Lee, M. Bougie, E. G. Booth, C. J. Kucharik, and H. K. Gibbs (2022). Environmental outcomes of the us renewable fuel standard. *Proceedings of the National Academy of Sciences 119*(9), e2101084119.

Lark, T. J., J. Meghan Salmon, and H. K. Gibbs (2015, apr). Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environmental Research Letters 10*(4).

Lark, T. J., S. A. Spawn, M. Bougie, and H. K. Gibbs (2020, dec). Cropland expansion in the United States produces marginal yields at high costs to wildlife. *Nature Communications 11*(1), 4295.

Li, Y., R. Miao, and M. Khanna (2019). Effects of ethanol plant proximity and crop prices on land-use change in the United States. *American Journal of Agricultural Economics 101*(2), 467–491.

Miao, R., M. Khanna, and H. Huang (2016). Responsiveness of Crop Yield and Acreage to Prices and Climate. *American Journal of Agricultural Economics 98*(1), 191–211.

Nafziger, E. (2019). Dealing with very late planting • farmdoc.

Potapov, P., S. Turubanova, M. C. Hansen, A. Tyukavina, V. Zalles, A. Khan, X.-P. Song, A. Pickens, Q. Shen, and J. Cortez (2022). Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century. *Nature Food 3*(1), 19–28.

Ramsey, S. M., J. S. Bergtold, and J. L. Heier Stamm (2020, nov). Field-Level Land-Use Adaptation to Local Weather Trends. *American Journal of Agricultural Economics*, ajae.12157.

Roberts, M. J. and W. Schlenker (2013, oct). Identifying supply and demand elasticities of agricultural commodities: Implications for the US ethanol mandate. *American Economic Review 103*(6), 2265–2295.

Sacks, W. J., D. Deryng, J. A. Foley, and N. Ramankutty (2010, jun). Crop planting dates: an analysis of global patterns. *Global Ecology and Biogeography*, no–no.

Scott, P. T. (2013). Dynamic Discrete Choice Estimation of Agricultural Land Use. *Working Paper*, 1–55.

Scurlock, J. and D. Hall (1998). The global carbon sink: a grassland perspective. *Global Change Biology 4*(2), 229–233.

Song, X.-P., M. C. Hansen, S. V. Stehman, P. V. Potapov, A. Tyukavina, E. F. Vermote, and J. R. Townshend (2018). Global land change from 1982 to 2016. *Nature 560*(7720), 639–643.

Souza-Rodrigues, E. (2019, nov). Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis. *The Review of Economic Studies 86*(6), 2713–2744.

Sulla-Menashe, D. and M. A. Friedl (2018). User Guide to Collection 6 MODIS Land Cover (MCD12Q1 and MCD12C1) Product.

Weinraub, M. (2013). U.S. soybean acreage to rise due to planting delays — Reuters.

Winkler, K., R. Fuchs, M. Rounsevell, and M. Herold (2021). Global land use changes are four times greater than previously estimated. *Nature communications 12*(1), 2501.

Yang, Y., D. Tilman, G. Furey, and C. Lehman (2019). Soil carbon sequestration accelerated by restoration of grassland biodiversity. *Nature communications 10*(1), 718.

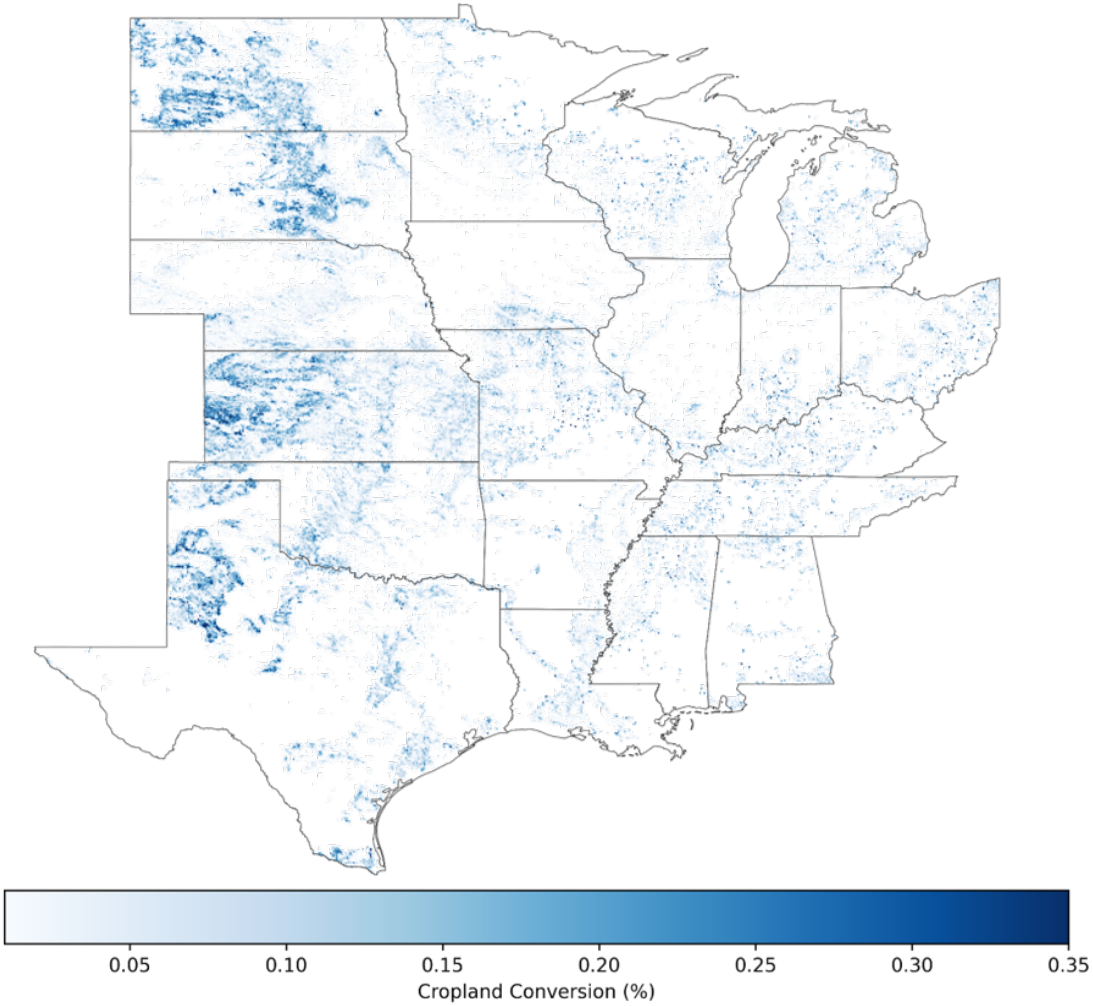


Figure 1: Net Land Conversion 2001-2017

Notes: Rates of conversion calculated as a percentage of total land switching at least one time between crop and non-crop usages within non-overlapping 4 km grids. Net land conversion was most concentrated in the boundary of the rain-fed region, including the eastern halves of North and South Dakota, western portions of Kansas, Kentucky.

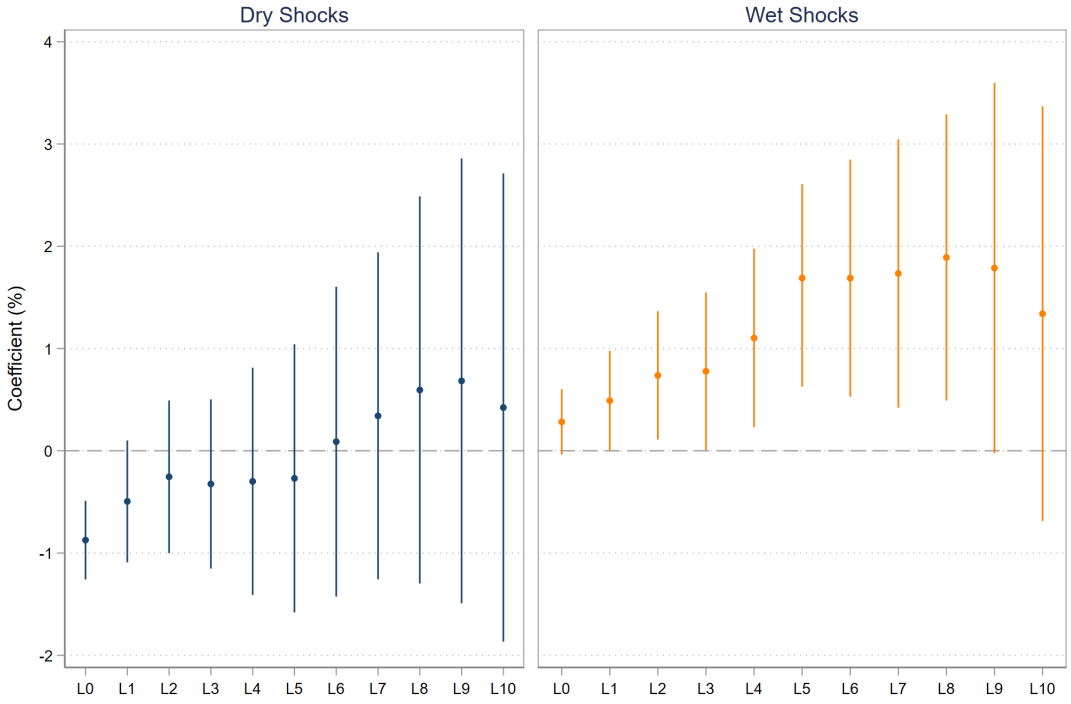


Figure 2: Distributed Lag Models

Notes: Blue (yellow) dots and lines refer to the point estimates and 95% confidence interval for dry (wet) anomaly and its lag terms from specification (7), the model with lags. All specifications use the 4 km grid level data and include grid and year fixed effects. Standard errors are in parentheses, adjusted for two-way clustering at grid and district-year levels.

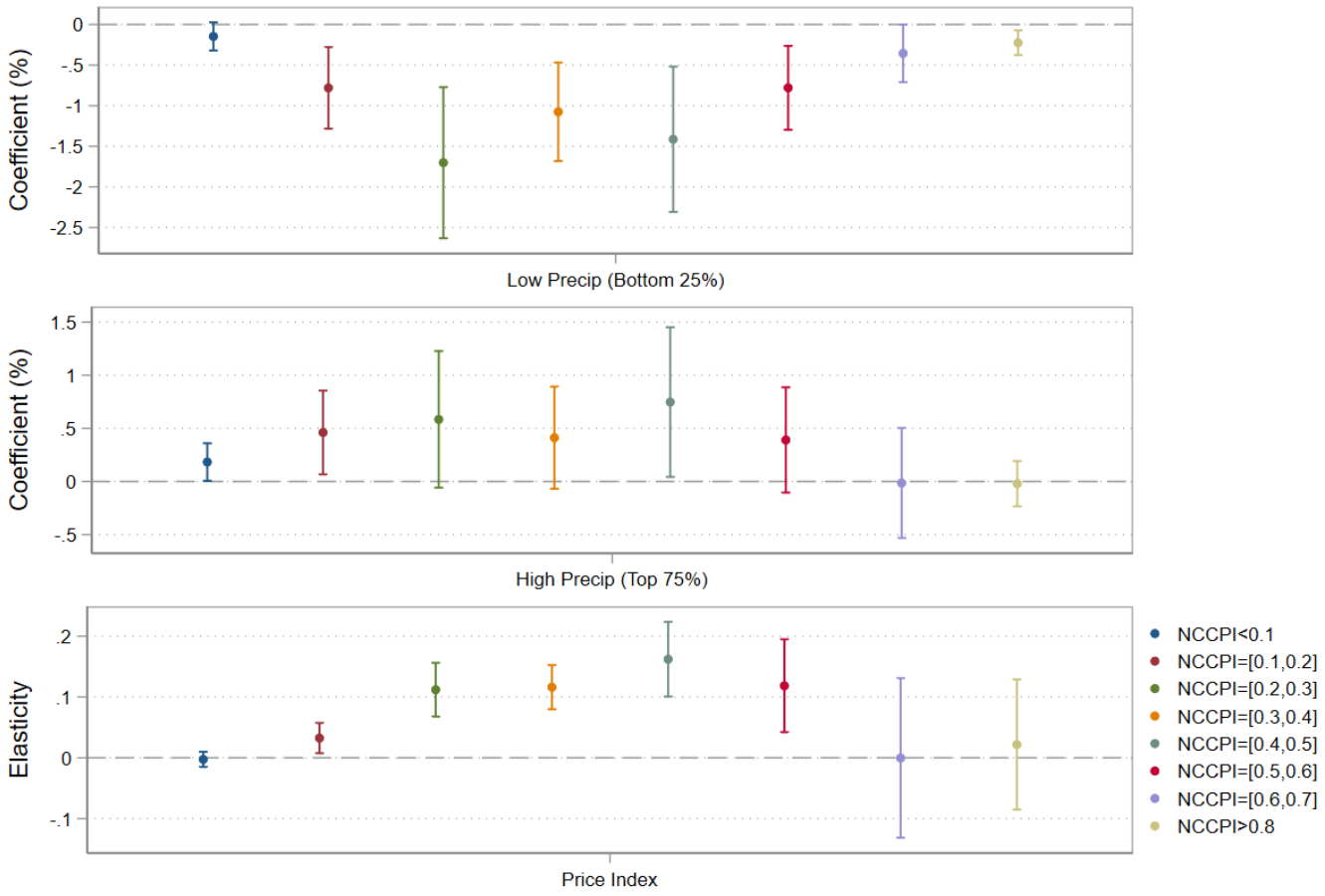


Figure 3: Effects of Weather Shocks and Price Fluctuations across land quality

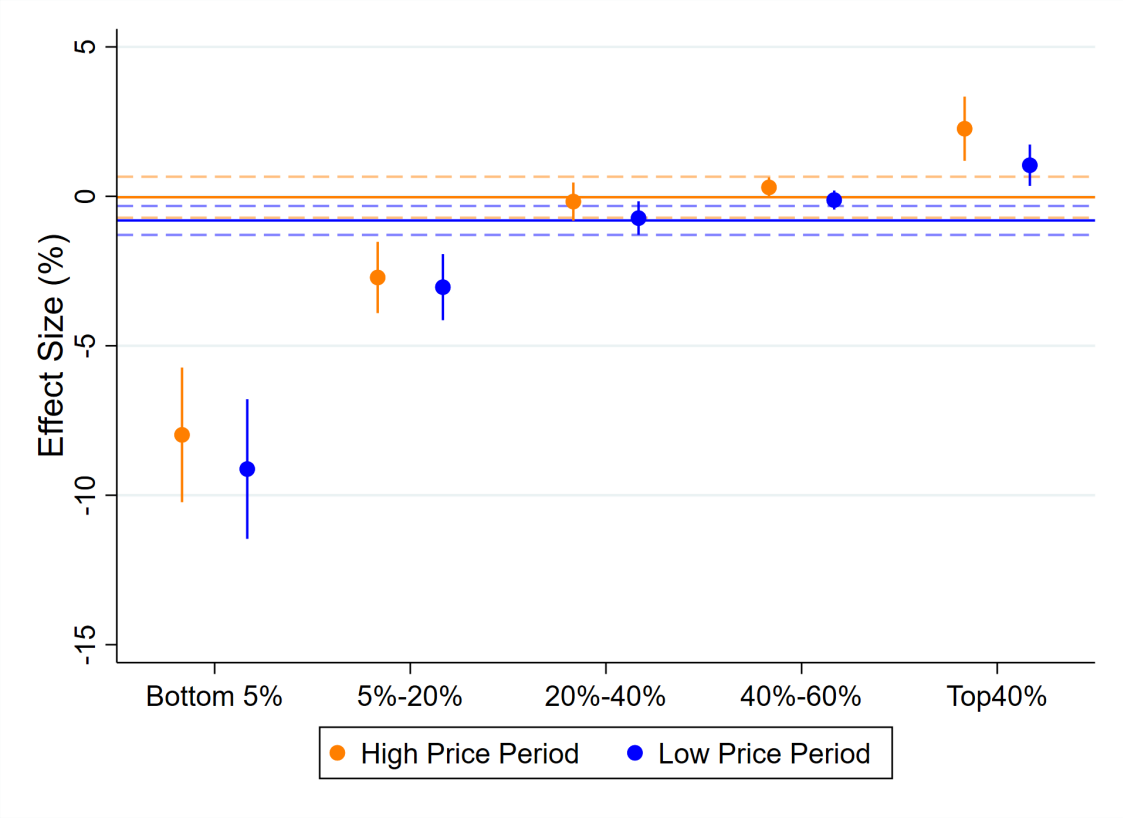


Figure 4: Effects of Dry Shocks and Price Fluctuations by Heterogeneous Groups

Notes: This Figure illustrates the heterogeneous crop-grass conversion probability under dry shocks and different price regimes across land parcels. The horizontal axis shows the percentiles of individual treatment effects in the entire distribution. Blue (orange) dots refer to point estimates of dry shock effects on crop-grass conversion under low (high) price regime. The bars represent 95 percent confidence intervals. Standard errors are clustered by county.

Table 1: Summary Statistics, 2001-2017

	Mean	SD	Min.	Max
Share of Cropland (ratio)	0.54	0.45	0.00	1.00
Received Price Index	2.15	0.98	0.78	5.49
Lagged Stock Level (Thousand Tons)	17000.32	17438.76	0.00	68007.88
Dry Anomaly (Bottom 25%)	0.25	0.43	0.00	1.00
Wet Anomaly (Top 75%)	0.25	0.43	0.00	1.00
Consecutive Rainy Days	5.92	2.17	1.00	84.00
Concentration of Rainfall (mm)	0.09	0.07	0.02	1.00
Average Temperature (Celsius Degree)	10.95	4.10	-2.13	23.82

Notes: Table reports statistics for the sample used for the main specification. Unit of observation is grid(4 km)-year. The temporal framework is 2001-2017. Prices are in year 2000 U.S. dollars.

Table 2: Average Effects

	(1) (OLS)	(2) (OLS)	(3) (OLS)	(4) (IV)	(5) (IV-First)
Dry Shock (Bottom 25%)	-0.00389***	-0.00412***		-0.00431***	0.0318
	(0.000887)	(0.000885)		(0.000937)	(0.0213)
Wet Shock (Top 75%)	0.00112	0.00126*		0.00138*	-0.0157
	(0.000760)	(0.000744)		(0.000768)	(0.0167)
Price Index		0.00988***	0.0246***	0.0183***	
		(0.00114)	(0.00709)	(0.00536)	
Stock					2.71e-08***
					(3.46e-09)
PRISM Grid FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	3008243	3008243	3008243	3008243	3008243
Mean of Dep. Variable	0.540	0.540	0.540	0.540	
R2	0.989	0.989	-0.00955	0.00359	

Notes: The dependent variable is the ratio of cropland of each 4 km PRISM grid. All specifications include the linear and quadratic average temperature in the pre-planting periods, 4 km grid-level fixed effects and year fixed effects as controls. Column 1-3 use the OLS specification and do not address the endogeneity issue of received price index. Column 4 instruments the received price index with lagged stock level for major crops. Column 5 presents the first-stage estimates for the instrument variable approach used in Column 4. Standard errors clustered at the 4 km PRISM grid and CRD-by-year levels (in parentheses). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Alternative Specification Timing of Rainfall Shock

	(1)	(2)	(3)
Price Index	0.0183***	0.0183***	0.0190***
	(0.00536)	(0.00535)	(0.00555)
Dry Shock (Bottom 25%)	-0.00431***		
	(0.000937)		
Wet Shock (Top 75%)	0.00138*		
	(0.000768)		
Dry Shock (Bottom 25%) at Preparing Stage			-0.00162**
			(0.000801)
Wet Shock (Bottom 75%) at Preparing Stage			0.00218***
			(0.000833)
Dry Shock (Bottom 25%) at Planting Stage			-0.00300***
			(0.00100)
Wet Shock (Bottom 75%) at Planting Stage			0.00223***
			(0.000749)
PRISM Grid FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	3008243	3008243	3008243
Mean of Dep. Variable	0.540	0.540	0.540
R2	0.00359	0.00412	0.00308

Notes: The dependent variable is the ratio of cropland of each 4 km PRISM grid. All specifications include the linear and quadratic average temperature in the pre-planting periods, 4 km grid-level fixed effects and year fixed effects as controls. Column 1 replicates our baseline specification. Column 2 decomposes the pre-planting period into two stages: preparation stage and planting stage and incorporate the rainfall shocks indicators at these two stages separately into the baseline specification. Standard errors clustered at the 4 km PRISM grid and CRD-by-year levels (in parentheses). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Best Linear Predictor of Cultivation Probability

	(1) Average Treatment Effects	(2) Heterogeneity
$\mathbf{1}_{\mathbf{DryShock}}$=1 & $\mathbf{1}_{\mathbf{HighPrice}}$=1	0.00074	0.29***
	(0.0032)	(0.033)
$\mathbf{1}_{\mathbf{DryShock}}$=1 & $\mathbf{1}_{\mathbf{HighPrice}}$=0	-0.0068***	0.24***
	(0.0025)	(0.026)

Notes: The dependent variable is the cultivation probability of 500 m land parcels in each year. The point estimate in column (1) measures the average treatment effects of being exposed to a dry shock in the pre-planting period on cultivation probability of each land parcel under different price regimes. Column (2) presents a statistical testing of the heterogeneity in effects across observations with different biophysical attributes. Rejecting the null hypothesis (H_0 : treatment effect is homogeneous and statistically indistinguishable from average treatment effect.) implies that heterogeneity is present. Standard errors are clustered at the county level (in parenthesis).* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Summary Statistics for the Biophysical Characteristics of Lands Most Affected By Rainfall Shocks and Price Fluctuations

	(1)	(2)	(3)	(4)	(5)
Biophysical Attributes	Bottom 5%	20%-80%	Diff	95% Conf.(lo.)	95% Conf.(hi.)
Water Capacity (0-5cm)	6.06	6.78	-0.73	-0.77	-0.68
Water Capacity(20-50cm)	32.52	34.41	-1.89	-2.12	-1.67
Water Capacity (5-20cm)	18.16	19.65	-1.49	-1.61	-1.37
Crop root zone water capacity	147.00	150.70	-3.70	-4.81	-2.59
Soil Organic Matter (0-5cm)	688.55	848.59	-160.04	-169.18	-150.89
Soil Organic Matter (20-50cm)	1797.97	2124.85	-326.89	-353.69	-300.08
Soil Organic Matter (50-100cm)	1617.17	1883.42	-266.24	-292.58	-239.90
Soil Organic Matter (5-20cm)	1646.53	2014.92	-368.39	-388.17	-348.61
Thickness of soil (0-5cm)	4.03	4.20	-0.17	-0.18	-0.15
Thickness of soil (20-50cm)	22.71	23.62	-0.91	-1.02	-0.80
Thickness of soil (5-20cm)	11.96	12.48	-0.52	-0.58	-0.47
Soil PH (0-5cm)	6.85	6.82	0.04	0.03	0.04
Soil PH (5-20cm)	6.88	6.84	0.04	0.03	0.05
Soil PH (20-50cm)	7.01	6.94	0.07	0.06	0.08
Percentage of parcel area falling into eight levels of slope classes:					
0-0.5 Degree	7.41	10.33	-2.92	-3.04	-2.79
0.5-2 Degree	47.30	48.20	-0.89	-1.17	-0.62
2-5 Degree	35.24	31.59	3.65	3.41	3.90
5-10 Degree	8.68	8.54	0.13	-0.03	0.30
10-15 Degree	1.12	0.98	0.14	0.10	0.19
15-30 Degree	0.31	0.34	-0.03	-0.06	0.00
30-45 Degree	0.01	0.02	-0.02	-0.02	-0.01
>45 Degree	0.00	0.00	0.00	-0.01	0.00

Notes: Column (1) presents means for parcel-year observations predicted to have a treatment effects at the bottom 5 percent of the distribution. Column (2) presents means for those in the middle 20% - 80% percentiles. Column (3) reports the difference between column (2) and (1). Column (4) and (5) report the 95% confidence intervals for this difference.Standard errors are clustered at the county level (in parenthesis).* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A Uncertainty and Quality of MODIS Data

Although the satellite-imagery data offers the opportunity to detect and track land-use history at the fine-scale spatial resolution, there are some concerns raised in the literature on their validity of land classifications, which might lead to and the current practices for processing and analysing these data, including concerns on the classification errors and time inconsistency.

First, the raw spectral data from sensor loaded on satellites have to be translated into interpretive land categories that can be used in empirical analysis. The validity and reliability of land cover products may vary substantially due to potential incompetence of the classification algorithm or ground-truth observations they remote sensing scientists used to train and validate classification algorithms. For example, it has been found very challenging to distinguish native grassland from cultivated hay and fallow/idle cropland (Kline et al. 2013) in CDL land cover products. Second, satellite land cover products are usually produced with the intent of mapping annual land cover ("snap-shot") rather than for tracking changes in land pixels over time (Lark et al. 2017). As a result, land cover transitions in consecutive years are typically not well accounted for in satellite land cover products if temporal inconsistencies aren't first addressed. Multi-temporal land cover products without further post-processing steps often show unrealistically high level of year-to-year changes in land cover (Abercrombie and Friedl 2016).

These classification errors and time inconsistency, if not corrected, may lead to substantial measurement errors and therefore, lead to potential biased and inconsistent estimates in our analysis. In the following sections, we first use a simple statistical model to describe the impact of measurement errors in a general case. We then provide more details on the validation and general accuracy information of the MODIS products. Although the reported quality assessment scores for land pixels used in our analysis are generally high, we

nevertheless undertake two sets of robustness exercises to ensure that our estimates do not suffer from biasness caused by measurement errors.

Measurement Errors: A Simple Model. Let L_{it} be the true land use type, X_{it} represents for a vector of explanatory variables such as weather shocks and price index. Thus, the "true" relationship between land use type and explanatory variables can be represented by

$$L_{it} = \beta X_{it} + \epsilon_{it} \tag{4}$$

Now assume that we can only obtain $\widehat{L_{it}}$, the observed land use types from satellite products with some measurement/classification error v_{it}. The relationship between observed land use type and explanatory variables can be represented by

$$\hat{L_{it}} = \beta X_{it} + v_{it} + \epsilon_{it} \tag{5}$$

where the $\widehat{\beta}$ is the coefficients estimated from our empirical model, which can be noted by

$$\begin{aligned} \widehat{\beta} &= \frac{cov(\widehat{L_{it}}, X_{it})}{var(X_i t)} \\ &= \frac{cov(\beta X_i t + \epsilon_i + v_i, X_{it})}{var(X_{it})} \\ &= \beta \frac{cov(X_{it}, X_{it})}{var(X_{it})} + \frac{cov(\epsilon_{it}, X_{it})}{var(x_{it})} + \frac{cov(v_{it}, X_{it})}{var(X_{it})} \end{aligned} \tag{6}$$

In the above model, the magnitude of bias (or inconsistency) of estimated coefficient β mainly depends on structures of measurement errors in data. When the error in satellite data is "classical", which is random and uncorrelated with the variable of interest (i.e., price index), we have both of $cov(\epsilon_{it}, X_{it})$ and $covcov(v_{it}, X_{it})$ equal to zero. In that case, the estimated β would be consistent and unbiased. The "classical" measurement error may

only increase the standard errors due to the additional term in the error which reduces the power of our statistical tests. In the worst case, it might lead to an insignificant results even though the underlying "true" effect is there. However, much of the error in the satellite data product, as noted by the remote sensing literature (e.g., Lark et al. 2017; Jain, 2020; Scott et al. 2021) could be "nonclassical". First, the error may not be random: due to the finite set of possible values that land cover types can take, the measurement errors may only move in one direction and hence an zero-mean error (conditional on the true discrete value) is impossible (Jain, 2020). Second, the error may correlate with some variables of interest (i.e., weather conditions). In these cases, the second and third term in equation (7) will not equal to zero and therefore substantially bias the estimates in a standard regression specification. Without understanding or proper correction of the dataset, it is possible that coefficients from otherwise well-identified economic analysis may be incorrectly estimated.

Accuracy of MODIS dataset The recent version of MODIS product (collection 6) incorporates substantial improvements on raw data and algorithms (Sulla-Menashe et al. 2019). Collection 6 data include corrections for sensor degradation in the visible bands of MODIS-Terra satellite and improvements to lower level products (e.g., cloud screening, atmospheric correction, and surface reflectance) that are used as inputs for land cover classification. In addition, MODIS (collection 6) implement a multi-temporal post-processing algorithm (Hidden Markov Models[35]) that uses multi-year data to reduce spurious land-use change introduced by classification uncertainty in individual years, which has not been achieved by most existing satellite land cover products such as Cropland Data Layer (CDL) and National Land Cover Data (NLCD) (Sulla-Menashe and Friedl, 2018; Abercrombie and Friedl, 2015).

Table S10 lists all the potential issues and sources of uncertainty identified by the MODIS group. Most of the known issues and sources of uncertainty identified by the MODIS groups

[35]Hidden Markov Model is widely used in recent econometric (Scott et al. 2021; Hu 2020) and remote-sensing studies that correcting the potential misclassification and spurious land use changes across years.

would not affect our measurement of crop-grass conversions or results. For the issue 6 (some grassland areas are classified as sparse forest), we note that this might lead to potential underestimation of grassland and hence attenuate our estimates. Therefore, our estimates should be cautiously interpreted as lower bounds.

Additional Checks. In addition, we also extracted the quality assurance (QA) information for each land pixel from the MODIS land cover product. The QA scores represent the assessment of the relative classification quality (scaled from 0-100) for each 500 m pixel in each year [36]. Figure S2 shows the distribution of QA scores for the entire sample and two subsamples: croplands and grasslands. We find that the overall quality of land classifications in this version of MODIS product is very high, with a mean of 97 and a standard deviation of 4.7. Although the overall land classification is accurate, our estimates could still be biased if the variables of interest (such as weather shocks) were correlated with QA scores, as we noted above. The direction of potential biases depends on the correlation between these variables and QA scores across different land classifications. For instance, in the main specification, we find that a dry shock would increase the probability of crop-to-grass conversion. However, if a dry shock itself could lead to a higher probability of misclassifying cropland as grassland, then the estimated coefficient may only reflect the "classification error". We then use the following specification to formally test the difference of QA scores across weather shocks and land classifications:

$$QA_{ijt} = \beta_1 Rainfall_{it} + \beta_2 Temp_{it} + \beta_3 Rainfall_{it} * \mathbf{1}(\textbf{Cropland}) + u_j + \gamma_t + \epsilon_{it} \quad (7)$$

where the QA_{ijt} is the quality assessment scores for land parcel j at 4 km grid i in year t, $Rainfall_{it}$ is the dry shock and wet shock variables, defined as those in our main specification, $Temp_{it}$ is the temperature controls. One significant difference between this specification and

[36] More detailed information of QA scores is documented on page 15 of MCD12 User Guide (V6) `https://lpdaac.usgs.gov/documents/101/MCD12_User_Guide_V6.pdf)`

the main specification is the fixed effects. Here we include the 500 m parcel-level fixed effects instead of the 4 km grid level fixed effects. That is because MODIS classify and post-process individual 500 m land pixels instead of 4 km grid. The classification for each land parcel should not be affected by any of the nearby parcels. For the same reason, we also cluster the standard errors at the 500 m parcel level to address any temporal autocorrelations in the error terms. Table S11 reports the effects of weather shocks on QA scores. Column (1) reports the difference in means of QA scores between cropland observations and grassland observations. We do find the QA scores for cropland are generally higher than grassland with a magnitude of 2 units (or 2% compared to the mean of QA for the entire sample). Column (2) reports the effects of rainfall shocks on QA scores for the entire sample. Columns (3) and (4) restrict our sample to land parcels being classified as cropland and grassland, respectively. These specifications show that rainfall shocks have statistically significant impact on the QA scores, but the coefficients are very close to zero. Finally, in column (6) of Table S11, we interact the weather shocks variables with the dummy variable whether a land parcel is classified as cropland in a specific year. Again, the results confirm that weather shocks could downward bias the cropland classification, i.e., the probability of being classified as cropland would significantly decrease under the dry or wet shocks. However, the magnitudes are small and negligible, which leverages the above concerns. In the robustness check, we create a "high QA quality" subsample by excluding all land pixels with QA scores below 95 (mean of QA score in our sample) in any of the years within our study period. We replicate our main specifications using this high-quality sample and obtain similar estimates as those from the main specifications.