COLUMBIA | SIPA
Center for Environmental Economics and Policy

CEEP Working Paper Series
Working Paper Number 19

October 2021
This Version: October 2022

Environmental Drivers of Agricultural Productivity Growth: CO_2 Fertilization of US Field Crops

Charles A. Taylor and Wolfram Schlenker

https://ceep.columbia.edu/sites/default/files/content/papers/n19_v2.pdf

Environmental Drivers of Agricultural Productivity Growth: CO_2 Fertilization of US Field Crops

Charles A. Taylor[1] and Wolfram Schlenker[2,3]

October 2022

Abstract

Post-war growth in agricultural productivity outpaced the US non-farm economy, spurred by steadily increasing crop yields. We argue that rising atmospheric CO_2 is responsible for a significant share of these yield gains. We present a novel methodology to estimate the CO_2 fertilization effect using data from NASA's Orbiting Carbon Observatory-2 (OCO-2) satellite. Our study complements the many field experiments by regressing county yields on local CO_2 levels across the majority of US cropland under actual growing conditions. For identification, we utilize year-to-year anomalies from county-specific trends, an instrument for those CO_2 anomalies using wind patterns, and a spatial first-differences approach. We consistently find a large CO_2 fertilization effect: a 1 ppm increase in CO_2 equates to a 0.4%, 0.6%, 1% yield increase for corn, soybeans, and wheat, respectively. In a thought exercise, we apply the CO_2 fertilization effect we estimated in our sample from 2015-2021 backwards to 1940, and, assuming no other limiting factors, find that CO_2 was the dominant driver of yield growth—with implications for estimates of future climate change damages.

Acknowledgments: We would like to thank seminar participants at Auburn University, University of British Columbia, University of California (UC-wide) environmental economics seminar, and Yale University for useful comments. This work was supported by the US Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modeling, MultiSector Dynamics, Contract No. DE-SC0016162 as well as by the USDA National Institute of Food and Agriculture, Grant No. 2022-67023-36400.

[1] Department of Agricultural and Resource Economics, UC Berkeley; c.taylor@berkeley.edu

[2] School of International and Public Affairs, Columbia University; ws2162@columbia.edu

[3] National Bureau of Economic Research (NBER)

The Green Revolution brought about a massive increase in crop yields across the world. In the US, for example, corn yields increased six-fold since 1940, while soybeans and wheat increased by a factor of three. Prior to this point, yields had been relatively flat, as shown in Figure 1. By extension, the agricultural sector experienced disproportionately high productivity growth (Jorgenson and Gollop 1992). Before 1950, US farm sector productivity growth was half that of the non-farm sector, but afterwards the relationship reversed with farm productivity growth exceeding the non-farm sector by 62% (Pardey and Alston 2021). Factors such as increased input usage, mechanization, irrigation, and improved crop genetics all contributed to yield growth (Wang, Heisey, Schimmelpfennig, and Ball 2015). However, because aggregate US farm output increased several-fold while the aggregate quantity of inputs (land, capital, labor, and materials) stayed flat, technology is generally understood to have driven agricultural productivity growth.

Understanding the drivers of agricultural productivity is important for economic growth both within the farm sector as well in as other sectors. The adoption of high-yield varieties, for example, had significant positive economic spillovers for the larger economy in India (Gollin, Hansen, and Wingender 2021) and other countries more generally (McArthur and McCord 2017).

This paper argues that carbon dioxide (CO_2) fertilization[1] may help illuminate the puzzling conclusion of Jorgenson and Gollop (1992): why did productivity growth explain over 80% of agriculture's postwar growth but less than 15% in the non-farm economy? During this time, atmospheric CO_2 was steadily increasing alongside steadily increasing crop yields, as shown in Figure 1. The physiological response of plants to CO_2 is well-known: CO_2 drives photosynthesis and has long been used as a greenhouse input to boost yields. Increasing CO_2 has driven global greening: over the last 40 years, half of the world's vegetated area has undergone greening[2], of which 70% is attributed to elevated CO_2 (Zhu et al. 2016).

Our paper investigates the extent to which elevated CO_2 contributed to the observed increase in crop yields during this time. Establishing a causal link between two trending variables is statistically challenging. CO_2 has risen smoothly in tandem with crop yields as well as other factors such as mechanization and input use. Industrialization, both in agriculture and other sectors, might have independently increased CO_2 levels as well as yields—making it all the more difficult to disentangle CO_2 fertilization from other productivity drivers.

[1]The CO_2 fertilization effect is defined in the scientific literature as the increase in photosynthetic activity in response to elevated CO_2. In this paper, we use the term more specifically to refer to an outcome of increased crop yields.

[2]This paper defines 'greening' as an increase in the growing season integrated leaf area index.

To date, field experiments and process-based models are the most common approaches to attribute yield trends to CO_2 levels. These approaches, however, face challenges of their own. The conditions in a well-controlled experiment might not be indicative of real-world farming conditions. Large regional differences in crop responses to CO_2 reflect geographic variation in crop distribution and environmental conditions (McGrath and Lobell 2013). CO_2 fertilization may be negligible in the presence of limiting factors such as nutrient deficiency (Kimball et al. 2001, Hungate, Dukes, Shaw, Luo, and Field 2003, Reich et al. 2006, Ziska and Bunce 2007). The effect is generally stronger under water deficit conditions (Ottman et al. 2001, Leakey, Uribelarrea, Ainsworth, Naidu, Rogers, Ort, and Long 2006, Keenan, Hollinger, Bohrer, Dragoni, Munger, Schmid, and Richardson 2013, Morgan et al. 2011), with the exception of soybeans (Gray et al. 2016) and possibly rice (Zheng, He, Guo, Hao, Cheng, Li, Peng, and Xu 2020). Elevated CO_2 may increase high temperature stress due to stomatal closure (Batts, Morison, Ellis, Hadley, and Wheeler 1997).

Field experiments have several limitations. There is only one major agriculture-focused CO_2 enrichment experiment in the breadbasket Midwest, SOYFACE, which is located at the University of Illinois and focused mainly on soybeans. Such field experiments can suffer from significant measurement error due to the difficulty of controlling elevated CO_2 concentrations in turbulent air (Allen, Kimball, Bunce, Yoshimoto, Harazono, Baker, Boote, and White 2020). Complicating matters further, a decline in the global carbon fertilization effect over time has been documented, likely attributable to changes in nutrient and water availability (Wang et al. 2020). While CO_2 enrichment experiments have generated important insights into the physiological channels of the fertilization effect and its environmental interactions, they are limited in the extent to which they reflect real-world growing conditions in commercial farms at a large geographic scale.

To this end, we employ a new approach to estimating the effect of CO_2 on crop yields that relies neither on process-based models nor on localized field experiments—while also allowing us to analyze the majority of the US cropland. We use observed ambient CO_2 data from NASA's Orbiting Carbon Observatory-2 (OCO-2) satellite and county-level crop yield data. The OCO-2 satellite detects changing ambient CO_2 levels that occur within and across locations and growing seasons (Crisp 2015). While CO_2 mixes in the atmosphere, there are temporal and spatial deviations that can be leveraged (see Appendix Figure A3). We focus on the US, which is the biggest producer of corn and soybeans, accounting for 33% of global production (FAOSTAT) over OCO-2's sample timeframe from 2015-2021, and 7% of global wheat production.

We use several empirical approaches that isolate both time-series and cross-sectional variation in our CO_2 measurement. The major identification concern involves CO_2 anomalies being correlated with other factors that influence crop yields. One could imagine several such confounders related to agricultural practices, fossil fuel production, urbanization, and large-scale weather systems. While we control for such factors when possible, we employ two other empirical strategies to aid in identification: a wind instrument for CO_2 exposure to address endogeneity concerns and a spatial first-difference approach that isolates differences between neighboring counties to reduce the effect of any regional confounders by differencing them out. Our results are robust to myriad sensitivity checks, i.e., the functional form (logarithmic versus levels), whether the temporal trend is by state or county, sample selection, and the choice of controls for co-occurring air pollutants.

We find large fertilization effects in the US: a 1 part per million (ppm) increase in CO_2 equates to a 0.4%, 0.6%, 1% yield increase for corn, soybeans, and winter wheat, respectively, in our baseline panel model. The estimates are high compared to those found in the agronomic literature, a fact we discuss and attempt to rationalize in section 5. More generally, these results shed light on a driver of yield growth that is usually taken as exogenous. The recent literature has used panel variation to estimate climate change damages by relating outcomes of interest to random exogenous year-to-year weather fluctuations (Dell, Jones, and Olken 2014). This approach, which relies on annual variation in weather, does not take into account longer-term dynamics which are correlated with climate change. Consequently, part of the estimated damages may be offset by yield gains from rising CO_2.

Our findings are relevant in several contexts. First, in providing an example of how satellite-based measures of CO_2 can complement field experiments to ensure external validity of the effect of CO_2 on agriculture and ecosystem functioning at a large scale.

Second, our finding that CO_2 fertilization has driven a large portion of the historical increase in crop yields has implications for how we think about the drivers of agriculture productivity growth, which has very large economic spillovers (Gollin et al. 2021).

Third, our results have relevance for estimating climate change impacts on agriculture. There is a gap between process-based studies of climate change which incorporate CO_2 fertilization and statistical ones which tend to omit it (Lobell and Asseng 2017), and the resulting estimates of climate impacts can vary greatly. For example, one study finds the net welfare effect on agriculture to be negative in the absence of CO_2 fertilization but negligible with fertilization (Moore, Baldos, and Hertel 2017).[3] And because the welfare ef-

[3]Some argue that CO_2 fertilization is understood well enough to be directly included in global climate

fects of climate change *vis-a-vis* agriculture vary across regions (Costinot, Donaldson, and Smith 2014, Nath 2020, Hultgren et al. 2022), such inequalities may be exacerbated if the CO_2 fertilization effect varies across crop types and environmental conditions. Finally, we emphasize that the strong fertilization effect we find occurs under current CO_2 levels and current environmental conditions. Given the experimental evidence showing a tapering of the CO_2 fertilization effect at higher levels, a linear extrapolation of our estimates into the future has to be considered with caution.

The paper proceeds as follows: Section 1 provides some background on the CO_2 fertilization effect and current estimates. Section 2 data section describes how we construct our CO_2 anomaly measure from the OCO-2 satellite data product, as well as the other datasets used in this analysis. Section 3 describes our identification strategies and empirical approaches before Section 4 presents our regression results along with robustness tests. Implications of these results are discussed in Section 5 by exploring scientific and policy implications of our study before Section 6 concludes and summarizes our main findings.

1 Background on CO_2 fertilization

Plants respond directly to rising CO_2 through photosynthesis and stomatal conductance, which is the basis for the fertilization effect (Long, Ainsworth, Rogers, and Ort 2004, Ainsworth and Rogers 2007). This response has been known for over 200 years. The role of CO_2 in plant growth was first demonstrated in 1796 by Swiss botanist Jean Senebier, and CO_2 gas has long been pumped into greenhouses to spur photosynthesis and increase the yield of horticultural crops.

The fertilization process varies by crop type. For C3 crops like soybeans, wheat, and rice, mesophyll cells containing RuBisCO[4] are in direct contact with the air. RuBisCO is an enzyme that fixes atmospheric CO_2 during photosynthesis and in oxygenation of the resulting compound during photorespiration. Thus, higher ambient CO_2 increases photosynthetic CO_2 uptake because RuBisCO is not CO_2-saturated at today's atmospheric levels (Long et al. 2004). For C4 crops like corn, on the other hand, RuBisCO is located in bundle sheath cells, where CO_2 levels are several times higher than atmospheric levels. At this concentration, RuBisCO is CO_2-saturated, and there may not be a direct photosynthetic response to changing atmospheric CO_2 levels. However, C4 yields are still indirectly af-

models and impact projections (Toreti et al. 2020).

[4]RuBisCO is the name of an enzyme short for ribulose-1,5-bisphosphate carboxylase-oxygenase

fected through increased water use efficiency via reduction in stomatal conductance (Long, Ainsworth, Leakey, Nösberger, and Ort 2006). All things being equal, one would expect a larger CO_2 fertilization effect for wheat and soybeans than for corn.

Historical estimates of yield responses to CO_2 came from controlled experiments in laboratories and greenhouses where CO_2 levels can easily be controlled. An early survey concluded that doubling ambient CO_2 increased yields by 24 to 43% for C3 crops in the context of full water and nutrient availability (Kimball 1983), which aligned with USDA reporting a 33% increase in yields for most crops under similar settings (Allen Jr., Baker, and Boote 1996). Another study estimated that CO_2 could have accounted for 15% US soybean yield growth from 1972 to 1997 (Specht, Hume, and Kumudini 1999).

In recent decades, free-air concentration enrichment (FACE), a process involving a series of pipes in fields emitting CO_2, has allowed for larger-scale trials in more realistic crop growing conditions. A survey of over 25 years of FACE experiments concludes that increasing CO_2 from 353 to 550 ppm results in 19% higher C3 yields, on average, while C4 crops were only affected under conditions of water scarcity (Kimball 2016). FACE experiments tend to show a lower fertilization effect than either laboratory or greenhouse enclosure studies (Long et al. 2006). However, recent work has pointed out potential measurement error, arguing that FACE estimates should be adjusted upward by 50% to account for the effect of air turbulence and CO_2 fluctuations (Allen et al. 2020). Geographic extent is also limited: there are only two long-standing FACE experiments in the US that focus on agriculture: Arizona FACE in Maricopa, AZ, and SOYFACE in Champaign, IL—with only the latter located in the traditional Midwestern breadbasket. We also note recent work that utilized OCO-2 satellite data to estimate the impact of the 2019 Midwestern floods on CO_2 uptake and crop productivity (Yin et al. 2020).

2 Data

Our primary measure of atmospheric CO_2 comes from the Orbiting Carbon Observatory-2 (OCO-2). Launched in 2014, OCO-2 is NASA's first satellite designed specifically to measure atmospheric CO_2 with the goal of better understanding the geographic distribution of CO_2 sources and sinks and their changes over time. We downloaded the bias-corrected OCO-2 LITE Level 2 v10 product, specifically the 'XCO2' value of averaged dry air CO_2 mole fraction (i.e., part per million, or ppm) over the atmospheric column.[5] The satellite has a

[5]Available: `https://disc.gsfc.nasa.gov/datasets/`

sun-synchronous orbit with an equatorial crossing time at 13:30 hours and a repeat cycle of 16 days. There are 8 distinct soundings with spatial footprints less than 1.29 km by 2.25 km each, encompassing an overall swath less than 10 km. A typical daily output contains over 150,000 XCO2 global readings, including the latitude-longitude point. Readings have quality flags (about 50% of readings), which we exclude from our analysis.

We then calculate the CO_2 anomaly that represents how much more or less CO_2 is observed in a given place and time relative to what would be expected in light of naturally occurring seasonal patterns and global time trends. To this end, we adjust CO_2 values from OCO-2 to account for annual patterns in which ambient concentrations decrease in the spring and summer when plants are actively photosynthesizing and increase in the fall and winter when plants are respiring on net.[6] We identify CO_2 anomalies relative to this pattern by estimating the average seasonality over the contiguous US with a 4^{th}-order Chebyshev Polynomial over the year, which we normalize to [-1,1] by transforming January 1st to equal -1 and December 31st to equal 1 with leap years having an additional day. We restrict the seasonality so the value on January 1 (time -1) equals the value on December 31 (time 1). Figure A1 displays the seasonality in CO_2 in the OCO-2 data. We further include a time trend to account for the annual increase in CO_2 at the global level. We then re-normalize our seasonality-adjusted values to July 1^{st} of each year, i.e., by adjusting a daily value by the differential between the average value for that day of the year and July 1st.

Next we assign each seasonally-adjusted OCO-2 reading to the PRISM grid (1/24° grid in latitude and longitude) in which it falls. Readings are averaged if there are more than one for a grid during the growing season from April to September. The PRISM grids within a county are then averaged using the amount of corn, soybean, or winter wheat area in the PRISM grid, where we aggregate the 30m-resolution from USDA's Cropland Data Layer to the PRISM grid.[7] The resulting seasonality-adjusted CO_2 level is a single value in ppm for each county-crop-year.

Figure A2 displays the resulting number of observations per county and by crop that have both CO_2 readings and annual yield data over the seven years from 2015 to 2021 with OCO-2 data. Given the 16-day revisit time, the high resolution, and the size of the median US county (1,610 km^2), a reading is not obtained for each US county in each year. Since

[6]Local CO_2 concentrations also exhibit strong diurnal within-day variation (Idso, Idso, and Balling Jr 2002, Xueref-Remy et al. 2018), but because the OCO-2 satellite is sun-synchronous and revisits points at the same time each day, this is not a concern for our study.

[7]Note that in the construction of the 4^{th}-order Chebyshev polynomial we include all non-flagged readings over the US, not just those over cropland.

we include both county fixed effects and county-specific time trends, we need at least three degrees of freedom per county, i.e., we can only include counties with at least 3 observations in our regressions.

For weather, we use a recalculated version of the fine-scale PRISM data at 2.5 minute resolution, or 4.5 km by 4.5 km that keeps the set of weather stations constant over time. We follow the approach from Schlenker and Roberts (2009), which found that four weather variables (two temperature, two precipitation) predict yields well. The two temperature variables are degree days 10-29°C (moderate degree days) and degree days above 29°C (extreme degree days) for corn. The upper bound is slightly higher for soybeans, resulting in degree days 10-30°C and degree days above 30°C. We use the same degree day variables for winter wheat as for soybeans. We experimented with using separate temperature measures by trimester (Tack, Barkley, and Nalley 2015), but did not find an improved fit. In each regression, we also include a quadratic of season-total precipitation. Precipitation and degree days are summed across the six-month growing season from April to September and spatially averaged using the same PRISM grid weights that are aggregates of USDA's Cropland Data Layer for each county.

Air pollution data come from the EPA's national network of pollution monitors. We use hourly data from the EPA's Pre-Generated Data Files[8] for five major pollutants: Ozone O_3 (44201), sulphur dioxide SO_2 (42401), carbon monoxide CO (42101), nitrogen dioxide NO_2 (42602), and particulate matter PM_{10} Mass (81102). We use the spatial interpolation approach of Boone, Schlenker, and Siikamäki (2019) to get the pollution variables at the PRISM grid, and then take the area-weighted average (again using the Cropland Data Layer) value of all grids in a county across the six-month growing season from April to September.

For the analysis of long-term trends in vegetation density, we use NOAA's Advanced Very High Resolution Radiometer (AVHRR) satellite measure of Normalized Difference Vegetation Index (NDVI) at 0.05° resolution, or 5.6km at the equator (Vermote et al. 2014) . Accessed through Google Earth Engine, the advantage of AVHRR relative to MODIS and other recent remotely-sensed products is its three decade timespan encompassing growing seasons from 1982 to 2013.

Crop area footprints for each county are derived from the USDA's Cropland Data Layer at 30m resolution. County-level crop yields for corn, soybeans, and winter wheat were obtained from USDA's National Agricultural Statistics Service.

We note that there are other potential CO_2 data sources: OCO-2's GEOS Level 3 daily

[8]Available `https://aqs.epa.gov/aqsweb/airdata/download_files.html`

product, which gap-fills observations in time and space using short transport simulations from the GEOS atmospheric model (Weir and Ott 2022), and which is utilized for the monthly visualization of CO_2 anomalies in A3. Another is NOAA's CarbonTracker, which is based on air sample measurements across 460 global sites and an inverse model of atmospheric CO_2 that adjusts surface-level CO_2 uptake and releases to align with observational constraints (Jacobson et al. 2020). The OCO-2 Level 2 product is our preferred measure of CO_2 because it is directly measured from space, and thus avoids the endogeneity risks around the modelling assumptions behind OCO-2 Level 3 and CarbonTracker in regards to weather and vegetation dynamics or ground-level confounders like local pollution and economic activity. Further, reanalysis products may suffer from promulgation of interpolation errors (Parker 2016).

Nevertheless, CarbonTracker has several advantages including its longer timespan (2000-2018), greater spatial resolution, and the fact that it models CO_2 near ground level, as opposed to OCO-2 which measures the entire air column from the ground up to the satellite. For the sake of robustness, we replicate our main analyses in Appendix B using CarbonTracker rather than OCO-2. We use the CarbonTracker product release CT2019B (Jacobson et al. 2020) and the level 1 estimates which correspond to 25m above the Earth's surface. To construct CO_2 anomalies, we perform a procedure analogous to the one used with the OCO-2 satellite: we take the distance-weighted average of the surrounding four CarbonTracker grids for each PRISM grid to derive PRISM-grid level CO_2 exposure, which is then aggregated to the county level using cropland weights from the Cropland Data Layer. Figure B2 displays the cross-plot of CO_2 anomalies from the OCO-2 satellite and CarbonTracker during the four years (2015-2018) in which the datasets overlap.

3 Model and empirical strategy

We estimate the CO_2 fertilization effect by linking OCO-2 satellite data on CO_2 levels with county-level yield data in the US. There are several identification challenges to address. While gaseous CO_2 ultimately diffuses across space and becomes uniformly distributed in the atmosphere,[9] this process occurs over the course of weeks to months and is affected by specific emission events, local CO_2 sources and sinks, as well as wind and weather dynamics (Hakkarainen, Ialongo, and Tamminen 2016, Massen and Beck 2011). Spatial variation in CO_2 exposure at any given time is driven by such disturbances. Figure A3 visualizes this

[9]The spatial diffusion of CO_2 is what makes climate change a global public goods problem. It also allows scientists to rely on singular sources of long-term CO_2 measurements, like the Mauna Loa Observatory, to estimate global CO_2 levels, which are then incorporated into global process-based models.

variation across the US during each month of the growing season in an example year, 2019. Taking Nebraska as an example, we see that in April CO_2 exposure is low compared to the US average, high in May, lower in June, neutral in July, high in August, then lower in September.

Our empirical approach links the resulting local variation in CO_2 (i.e., anomalies) to fluctuations in yields. To do this, we match the yield data with local CO_2 readings and weather outcomes over the area where corn, soybeans, and winter wheat are grown within each county, respectively. All models use seasonality-adjusted CO_2 anomalies in ppm, as described in the Data section, and log of county-level yields as the outcome variable unless otherwise noted. We focus on the US, a top global agricultural producer. Our primary analysis encompasses counties east of the 100° meridian for corn and soybeans, the same set of counties as used in Schlenker and Roberts (2009). Because winter wheat is grown further west, we use all states east of the Rocky Mountains as the baseline for wheat. These areas account for the vast majority of US row crop production. As a sensitivity check, we also perform the analyses on the entire continental US and other sub-samples, as visualized in Figure A2.

Panel model

The set up of our panel model is similar to what is commonly used in the literature to estimate climate change damages. We regress yields on CO_2 after controlling for the four weather variables that were found to best predict corn and soybean yields (Schlenker and Roberts 2009) and criteria air pollutants (CO, NO_2, O_3, PM_{10}, SO_2). The panel model includes county fixed effects to account for differences in average yields across counties driven by factors such as soil quality and average climate, as well as county-specific time trends to account for local trends in both CO_2 and yields to rule out a spurious correlation of trending variables. Figure A4 illustrates the variation used in the panel model, highlighting the correlation in Macoupin county, Illinois, which is downwind from an urban area (St. Louis).

The panel model specification is:

$$y_{it} = \alpha_{i0} + \alpha_{i1}t + \beta\ c_{it} + \boldsymbol{\gamma}\mathbf{W}_{it} + \boldsymbol{\delta}\mathbf{P}_{it} + \epsilon_{it} \tag{1}$$

where y_{it} is log crop yield in county i in year t; α_{i0} is a county fixed effect; α_{i1} is a county-specific time trend; β measures the observed CO_2 fertilization effect from the seasonally-adjusted CO_2 reading (c_{it}) in the county i in year t; $\boldsymbol{\gamma}$ is a control vector for weather (two

temperature degree day variables, precipitation and precipitation-squared, all summed over the six-month growing season), while $\boldsymbol{\delta}$ is a control for five criteria air pollutants P_{it} (CO, NO_2, O_3, PM_{10}, SO_2). We use the daily mean for CO, NO_2, PM_{10}, and SO_2, and O_3 averaged over the growing season, which we fix to April-September for all crops. Finally, ϵ_{it} are the errors, which are clustered at the state level to account for spatial correlation and state-level policy. The panel specification has been used to link random year-to-year weather fluctuations to annual yield outcomes (Schlenker and Roberts 2009). The effects are clearly visible in Figure 1 where we see a significant reduction in national corn yields in 2012 when the Corn Belt experienced drought and extreme heat events.

Identification remains challenged by several factors: 1) the possibility that local drivers of CO_2 could also affect yields, 2) possible reverse causality where yield anomalies impact CO_2 levels, although this should downward bias our estimates as higher yields imply less CO_2, and 3) the spatial correlation in ambient CO_2 concentration and possible spatially-correlated confounders that may include agricultural practices, fossil fuel production, urbanization, and large-scale weather systems. While we control for confounders when possible, we exploit differing sources of variation to identify the CO_2 fertilization effect: first, an instrumental variables approach using upwind counties to ensure that local CO_2 levels are not driven by local conditions, and second, a spatial first differences approach to isolate differences between neighboring counties to reduce the effect of any regional confounders that get differenced out.

Wind instrument

Wind direction is often employed in health economics to obtain exogenous variation in pollution exposure (Schlenker and Walker 2016, Deryugina, Heutel, Miller, Molitor, and Reif 2019). In the context of our study, mean CO_2 concentrations and peak fluxes have been shown to correspond with the wind direction of localized CO_2 emission sources (Coutts, Beringer, and Tapper 2007, Massen and Beck 2011, García, Sánchez, and Pérez 2012, Xueref-Remy et al. 2018). And while there are non-wind drivers of CO_2 anomalies, like power plants and vegetation, these features vary far less over time and space than relatively random atmospheric phenomena that drive wind patterns. Such within-year variation can be visualized in Figure A3.

One limitation of this instrument is that wind-driven CO_2 exposure could also be correlated with other co-occurring pollutants. We try to address this be explicitly controlling for the five criteria air pollutants. In any case, given the negative affect of pollutants like ozone on crop yield (Boone et al. 2019), this would likely bias our estimates downward.

Figure 2 shows our approach to deriving a wind instrument that follows Braun and

Schlenker (2022). First we calculate the centroid of each county as the cropland-weighted average of all grids in a county, weighting by the combined corn, soybeans and winter wheat acreage from the Cropland Data Layer. We then pair each county's centroid with the centroids of all its neighbors, which are *not* cropland-area weighted as we are looking for all CO_2 readings as instrument, not just over the agricultural area.

We next determine the upwind county based on the direction in which neighbor centroids are located, which is time invariant, and the county's hourly wind direction over the growing season constructed from hourly North American Land Data Assimilation System (NLDAS) data, again using the cropland-area weighted average of all NLDAS grid cells in a county. We derive the cosine between the direction of neighboring counties and the wind direction and choose the county whose cosine is closest to -1 as the "upwind" that hour.

For each neighboring county, we then sum the number of hours it is "upwind" from our target county over the 4392 hours comprising the April to September growing season in a given year. The neighboring county that is the most hours upwind is thus categorized as the "Upwind" county, which can change year-to-year. We then instrument each county's CO_2 anomaly on that of its most upwind neighbor in that given year. We purposefully pick the CO_2 anomaly in the upwind county that is over the entire county area, not just the agricultural area, as we are interested in all upwind sources.

The strength of the upwind relationship is an important consideration. Figure 3 displays the number of hours that the "Upwind" county is actually upwind in a given year for each crop. A low number of hours implies a weak relationship in which the wind is variable over the growing season, i.e., a value of 1000 hours means that the county most frequently upwind is in fact only upwind 23% of the time (1000 divided by the 4392 hours in the growing season). In our IV regression, we vary the minimum hour cutoff and see that the CO_2 effect persists. Summary statistics related to the IV set up are included in Table 1.

The wind instrument is modelled as:

$$y_{it} = \alpha_{i0} + \alpha_{i1} t + \beta\ c_{it} + \boldsymbol{\gamma}\mathbf{W}_{it} + \boldsymbol{\delta}\mathbf{P}_{it} + \epsilon_{it} \quad (2)$$

$$c_{it} = a_{i0} + a_{i1} t + b c_{i[upwind]t} + \boldsymbol{\gamma}\mathbf{W}_{it} + \boldsymbol{\delta}\mathbf{P}_{it} + e_{it} \quad (3)$$

where items are defined as in equation (1) for the panel model, except that β measures the observed CO_2 fertilization effect from the instrumented CO_2 value (c_{it}), and $c_{i[upwind]t}$ is the instrument using the CO_2 value from the county that is most frequently upwind based on the hourly wind data.

Spatial first differences

We use a spatial first difference (SFD) model that is a generalization of Druckenmiller and Hsiang (2019) in order to leverage another source of variation. It compares the change in the CO_2 and yield anomalies across all county neighbor pairs after removing county fixed effects and county-specific annual time trends, while again controlling for spatial differences in the other control variables (weather and air pollutants). To do this, we first derive anomalies by factoring out county fixed effects and county-specific time trends for all variables:

$$v_{it} = \alpha_{i0} + \alpha_{i1} t + \epsilon_{it} \tag{4}$$

where $v_{it} \in \{y_{it}, c_{it}, \mathbf{W}_{it}, \mathbf{P}_{it}\}$ to obtain the anomalies $\epsilon_{it}^{(v)}$ for each variable. By the Frisch-Waugh-Lovell theorem, regressing $\epsilon_{it}^{(y)}$ on $\epsilon_{it}^{(c)}$ while controlling for $\epsilon_{it}^{(\mathbf{W})}$ and $\epsilon_{it}^{(\mathbf{P})}$ would give the same estimate for β as in equation (1). Instead, we look at the spatial first difference by pairing each county i with all of its neighbors j, defined as having a common coordinate in the county shape file. We take the difference in anomalies in a given year between neighbors: $\Delta_{ijt}^{(v)} = \epsilon_{it}^{(v)} - \epsilon_{jt}^{(v)}$, so any common shock would be differenced out. In a second step we then link these differences in annual anomalies (one observation for each county-pair per year):

$$\Delta_{ijt}^{(y)} = \beta \Delta_{ijt}^{(CO_2)} + \boldsymbol{\gamma} \Delta_{ijt}^{(\mathbf{W})} + \boldsymbol{\delta} \Delta_{ijt}^{(\mathbf{P})} + \epsilon_{ijt} \tag{5}$$

In the appendix we use an alternate SFD approach using a cross-sectional model to examine persistent average gradients in CO_2 and yields in space while again controlling for weather and co-pollutants. Ignoring annual anomalies (i.e., shocks), for each variable we derive the average outcome over all years $\overline{v_i} = \frac{1}{T} \sum_{t=1}^{T} v_{it}$, and again pair county i to all its neighbors j, defined as having a common coordinate in the county shape file. We take the difference in average outcomes between neighbors: $\Delta_{ij}^{(v)} = \overline{v_i} - \overline{v_j}$ and link these differences in space in a cross-sectional regression (one observation for each county-pair):

$$\Delta_{ij}^{(y)} = \beta \Delta_{ij}^{(CO_2)} + \boldsymbol{\gamma} \Delta_{ij}^{(\mathbf{W})} + \boldsymbol{\delta} \Delta_{ij}^{(\mathbf{P})} + \epsilon_{ij} \tag{6}$$

The SFD methods address concerns about regional variation in confounders that might be correlated with regional variation in CO_2 levels as shown in Figure A3. On the other hand, given the small remaining variation in CO_2 levels between neighboring counties, the SFD approach might amplify measurement error by differencing out the common shock in a year and lead to attenuation bias. Panel models and SFD models require different

assumptions, i.e., for panels, that annual CO_2 anomalies are uncorrelated with other omitted explanatory variables, and for the SFD that the average gradient in CO_2 is uncorrelated with omitted variables. The fact that we obtain robust and consistently-positive CO_2 fertilization estimates makes it less likely that results are driven by the particular assumptions of each individual approach.

4 Results

The panel model results showing the aggregate effect of CO_2 on county-level crop yields in the US are included in Table 2. The point estimates for CO_2 are positive in all cases. Using the full models in columns (1c), (2c), and (3c) that control for weather and pollution, we find that a 1 ppm increase in CO_2 equates to yield increases for corn, soybeans, and winter wheat of 0.4%, 0.6%, 1%, respectively. The fertilization effect is less for corn (C4 crop) and greater for soybeans and winter wheat (C3 crops), as observed in controlled experiments.

These results do not appear to be driven by outliers: Figure A5 plots the anomalies for OCO-2 in the preferred panel model after the covariates are factored out with the regression line and a 90% confidence band. Many factors influence yields beyond CO_2, and to that end we see that much variation remains after accounting for CO_2 as well as our controls for weather and other environmental factors. However, as long as CO_2 fluctuations are uncorrelated with the other remaining unaccounted factors, our approach provides an unbiased estimate of the CO_2 fertilization effect. For example, Table 2 starts in columns (a) by not controlling for either weather or pollution in a county, while columns (b) account for the four weather variables that have been shown to be good predictors of corn and soybean yields, yet the CO_2 coefficient is relatively stable across specifications. Finally, columns (c) account for other pollutants that might co-vary with CO_2, e.g., because they are co-generated when fossil fuels are burned, but again we do not find statistically different results. The inclusion and exclusion of these controls known to influence crop yields do not significantly alter our findings, so any omitted variable would have to be correlated with both CO_2 and yields, but not the other controls.

Table 3 includes the results from the IV model. While our baseline model included all counties with at least three CO_2 and yield observation, the wind IV furthermore requires that the upwind county has at least three CO_2 and yield observations, which further limits our dataset. The table therefore first replicates the panel results for the same set of counties in the top row – columns (b) and (c) – before the IV results are shown in the bottom

row to ensure that any possible differences are driven by the IV setup and not sample composition. As described in the modeling section, any possible feedback would downward bias our estimates, i.e., higher yields through increased photosynthesis removes CO_2 from the atmosphere and thereby implies a negative correlation. We would therefore expect the coefficient to increase when we instrument by upwind CO_2 levels, which is indeed the case. Compared to the OLS panel estimates, the coefficients are larger for corn, soybeans, and winter wheat by 36%, 65%, 19%, respectively when averaging across columns (b) and (c). The effect is robust to dropping observations with a less strong upwind relationship (under 1000 hours upwind per growing season).

Table 4 shows the results of the spatial first differences model, which directly isolates variation between neighboring counties. We see consistently positive coefficients—both for annual shocks as well as cross-sectional average levels, the latter being shown in Table A1. The magnitudes are smaller than the panel and IV estimates, but in each case the fertilization effect is largest with winter wheat. The spatial first difference approach clearly shows that the results are not driven by regional anomalies in a year, which would difference out and cause the results to vanish. At the same time, we have infrequent satellite readings over random points in a county in each year, and the spatial first difference approach will therefore amplify measurement error as it differences out the common signal for two neighboring counties and hence suffers from attenuation bias.

We perform a number of sensitivity checks that produce largely similar results. First, we vary the model specification: our baseline model links log yields to CO_2 levels, assuming that a 1 ppm change in CO_2 has the same *relative* (percent) effect on yields. Figure 5 compares the effect of the main specification (Log-Linear) to other functional form combinations: Linear-Linear (constant absolute effect), Log-Log (constant elasticity), and Linear-Log. To make the results comparable, we do not show the coefficients, but instead show the effect of a 1ppm increase on corn yields in each case, as well as the 90% confidence interval. Results are very similar, which is not surprising, as we only have seven years of data and they all provide local linear approximations for this limited duration. It should be noted, though, that in our thought exercise where we extend the coefficient backward to 1940 to simulate what the CO_2 effect was, the functional form makes a very large difference, which would also be the case if we were to project the effect several decades into the future.

Second, we vary the time trends to allow for the possibility that temporal patterns in CO_2 levels and crop yields may be occurring at a geographic level different than the county level—e.g., state-level policies may drive energy or agricultural production. Figure 6 plots

the CO_2 coefficient for all models (Panel, IV, and SFD) alternately using no time trend, a common trend, state-level trends, and county-level trends. All point estimates are positive. The chosen time trend has no effect in the spatial first difference approach as neighboring counties tend to trend in a similar fashion, and hence exclusively focusing on comparisons between neighboring counties absorbs a common time trend. While there is some variation in the panel or IV setup when using a common national time trend or omitting trends altogether, the granularity of the time trend (state-specific versus county-specific) does not matter much.

Third, we run our analyses comprising different US geographies as visualized by the colored regions in Figure A2. Our primary analysis encompasses counties east of the 100° meridian (excluding Florida) for corn and soybeans, an area accounting for the vast majority of US corn and soybean production, as well as counties east of the Rocky Mountains for wheat. Figure 7 includes results for the sample comprising the entire contiguous US, east of the Rockies, or east of the 100° meridian of primarily-rainfed agricultural counties. The results are again fairly stable, mitigating concerns that this relationship is driven by regional dynamics like irrigation. Note that the color coding of subsets in this Figure matches the map in Figure A2.

Finally, one concern about the OCO-2 satellite data we use in our analysis is that it measures CO_2 across the entire atmospheric column, i.e., the area between the satellite and the ground. What matters for plants is CO_2 at ground level, not higher altitudes. While CO_2 concentrations across the air column are related through diffusion processes, if CO_2 disturbances at ground level phase out in altitude, then the variation we observe in the satellite data would be smaller than the ground-level variation, thereby leading to an upward-biased CO_2 fertilization coefficient.

We therefore replicate our entire analysis using a modelled CO_2 product from NOAA's CarbonTracker in Appendix Section B. CarbonTracker provides CO_2 levels at various altitudes and we choose the one closest to the ground. This data is for 2000-2018, and hence has four overlapping years (2015-2018) with our satellite data (2015-2021). The cross-plot of CO_2 anomalies in Figure B2 shows that the variation in OCO-2 and CarbonTracker are comparable. Moreover, we consistently find significant CO_2 fertilization effects in the CarbonTracker data, that in case of corn are even larger than our baseline estimates, alleviating concerns that our coefficients are artificially inflated. Since interpolated re-analysis data products like CarbonTracker can be a 'black box' to users, we prefer the raw OCO-2 satellite measurements—especially in relation to our instrumental variable approach where spatially-

interpolated data may mechanically produce a significant first stage.

5 Discussion

Global ambient CO_2 levels have increased by 2 to 2.5 ppm per year on average since 2000. Our panel models estimate a yield responses between 0.4% to 1% per 1 ppm CO_2. These estimates, which are at the very top of the range found in the literature, imply that CO_2 fertilization was a major contributor to recent crop productivity in the US. Put another way, yields may have increased 1% to 2.5% per year due to CO_2 in recent years, fully accounting for observed yield increases.

Looking further back in time, Figure 1 shows that since 1940 corn yields have increased by 500% and soybeans and winter wheat yields by 200%, while ambient CO_2 levels have increased by about 100 ppm. We can conduct a back-of-the-envelope counterfactual in which we hold CO_2 constant at 1940 levels and assume the CO_2 fertilization effect that we estimated using 2015-2021 data can be applied throughout 1940-2021. Admittedly, this is a strong assumption, as previous studies mentioned in Section 1 have shown that the CO_2 fertilization effect might diminish under stressors, e.g., nutrient or water deficiencies. If crops suffered from those other limiting factors, the CO_2 fertilization effect might have been weaker. And the climate in recent decades would not be the same if CO_2 had remained at 1940 levels. Nevertheless, we find it useful to run this thought experiment to highlight the possible magnitude of the CO_2 fertilization effect. Figure 4 shows the results of this thought experiment, implying that CO_2 fertilization may be responsible for the vast majority of past productivity growth, and that in the absence of CO_2 fertilization, yields may have otherwise started to plateau or even decline in recent decades.

How could this have occurred? One place to draw insights is the period before 1940, when crop yields were largely stagnant during a backdrop of rapid industrialization and economic growth. Olmstead and Rhode (2002) argue that from 1800 to 1940, "wheat production witnessed wholesale changes in varieties and cultural practices...without these changes, vast expanses of the wheat belt could not have sustained commercial production and yields everywhere would have plummeted due to the increasing severity of insects, diseases, and weeds." What if this same dynamic persisted after 1940 in which agricultural innovation served largely to protect crops against loss rather than increase yields? One third of all crop seed patents are related to crop pests or pathogens (Moscona and Sastry 2022), and many agricultural technologies are focused on crop resilience to extreme weather (e.g., flood and

drought tolerance). In addition, only a small share of yield gains since 2005 can be attributed to genetic improvements (Rizzo, Monzon, Tenorio, Howard, Cassman, and Grassini 2022). Taken together, if CO_2 had stayed static, yields could have conceivably stayed flat or only grown modestly over time—especially given that extreme weather and pest pressures have increased with globalization and climate change (Bebber, Holmes, and Gurr 2014, Deutsch, Tewksbury, Tigchelaar, Battisti, Merrill, Huey, and Naylor 2018).

Notwithstanding these explanatory factors, how do these results square with existing CO_2 fertilization estimates? Most FACE experiments raise CO_2 levels by 190 to 200 ppm over a 350 ppm baseline, on average, to which yield responses averaged 18-19% (Kimball 2016, Ainsworth and Long 2021), or 0.1% per ppm. Our estimates of 0.4 to 1% per ppm are thus 4 to 10 times larger. However, the average effect conceals significant variation across crops, location, and growing conditions. A FACE study of dryland wheat in Australia showed that a 180 ppm increase in CO_2 was associated with yield increases of 24% and 53% in two sites, with some yield responses reaching 79% (Fitzgerald et al. 2016). The latter estimate, equivalent to 0.44% per ppm, is closer to what we find. Similarly, under varying environmental conditions yield responses above 35% have been observed for corn, rice, cotton, as well as various leguminous and root crops (Kimball 2016, Ainsworth and Long 2021). Given the range in FACE results and the complexities of environmental interactions, it is difficult to benchmark our results.

It is also likely that FACE experiments underestimate CO_2 responses due to measurement error related to the difficulty of maintaining an elevated gas concentration in an open space. FACE experiments regulate CO_2 through a series of pipes in the field that inject the gas at high velocity based on sensor feedback. CO_2 concentrations in FACE experiments fluctuate widely due to air turbulence, varying 10 times more than what plants experience under natural conditions (Kimball 2016, Allen et al. 2020). When elevated CO_2 is supplied in cycles or pulses, crop responses are lower than if the CO_2 is supplied more steadily (Bunce 2012).[10] Just as CO_2 levels can be better controlled in chamber studies than FACE experiments, our study's smaller absolute variation in ambient CO_2 would imply less fluctuation as well. A recent review of FACE experiments by USDA researchers found that they underestimate yield responses by a factor of 1.5 (Allen et al. 2020) due to CO_2 fluctuations. With this adjustment, our estimates become more even more reasonable.

It is also worth noting that there are only two long-standing FACE experiments in the US

[10]Short-term fluctuations in CO_2 can affect photosynthetic activity in part because leaves have little storage capacity for gaseous CO_2 and the half-life of CO_2 in the gas space is short, e.g., 0.20 seconds for wheat (Hendrey, Long, McKee, and Baker 1997).

that focus on agriculture: Arizona FACE in Maricopa, AZ, and SOYFACE at the University of Illinois in Champaign, IL (Ainsworth and Long 2021). Other FACE experiments study non-cropland ecosystems like forests, grasslands, and tundra, as well as crops in other countries. Only SOYFACE in Illinois has the potential to approximate agricultural conditions in the Midwest, where most crop production occurs in the US—though SOYFACE's primary focus on soybeans limits what can be said about other crops. Moreover, SOYFACE consists of 16 octagonal experimental sites that are each 20m wide ($283m^2$), covering about $4{,}500m^2$ in total, or slightly more than one acre. For comparison, the average farm in the US is 445 acres (USDA ERS), which raises questions about how generalizable the results are for the Midwest—especially considering the large variation in crop yields across counties and even within fields (Lobell and Azzari 2017).

Therefore, it is possible that FACE experiments do not reflect the growing conditions and farming practices of the major growing regions. Given the well-documented interactions between CO_2 and environmental conditions[11], CO_2 fertilization effects could vary between FACE experiments and commercial agricultural operations in response to differing fertilization and input regimes, soil and water management practices, and local air pollution and climate anomalies across regions—as well as conditions that vary over time. Our experimental design utilizing OCO-2 satellite measures of ambient CO_2 allows us to account for this variation at a larger scale and across multiple years of observations.

Nevertheless, we offer another potential explanation for why our CO_2 fertilization estimates are higher than what's generally found in the literature. First, our study looks only at small increases in CO_2, and it may be inappropriate to extrapolate out fertilization effects which may diminish at higher CO_2 levels. As noted, most studies (including FACE, open top chamber, and greenhouse experiments) involve a large increase in CO_2 levels by 200 ppm or more over ambient levels. In contrast, our paper relies on variation in the range of 15 ppm during the OCO-2 timeline from 2015 to 2021. Such marginal increases could produce relatively higher fertilization effects given the diminishing photosynthetic response curve of plants to elevated CO_2. The rate of CO_2 assimilation in C4 plants, for example, is nearing saturation at current global CO_2 concentrations (Lambers and Oliveira 2019). Our results may reflect higher yield responses around current ambient CO_2 levels that occur at a steeper part of the photosynthetic response curve. This same dynamic could explain part of the

[11]Including nutrient availability (Kimball et al. 2001, Hungate et al. 2003, Reich et al. 2006, Ziska and Bunce 2007), water availability (Ottman et al. 2001, Leakey et al. 2006, Keenan et al. 2013, Morgan et al. 2011, Zheng et al. 2020, Gray et al. 2016), and combined nutrient-water-CO_2 interactions (Markelz, Strellner, and Leakey 2011)

observed decline in the global carbon fertilization effect (Wang et al. 2020).

But more generally, a strong positive relationship between CO_2 and yields should not be inherently surprising. CO_2 is a purchased input in many agricultural settings. As mentioned earlier, the gas has long been pumped into greenhouses to spur photosynthesis and increase the yield of horticultural crops. Optimal CO_2 concentrations of 900 ppm have been suggested, which is over twice current ambient levels (Mortensen 1987).

An alternative way of contextualizing our results involves looking at trends in non-cropland vegetation near to the US breadbasket where our analysis is focused. As mentioned earlier, studies have documented a global greening trend associated with CO_2 fertilization (Zhu et al. 2016). In a similar vein, Figure C1 analyzes trends in NDVI, a measure of vegetative density, over 32 years from 1982 to 2013 using AVHRR satellite data. We find that NVDI increases 0.48% per year across the entire US, on average. Focusing just on forested land, which is still subject to CO_2 fertilization but far less actively managed than cropland, NVDI growth is 0.64% per year.[12]. We can look to isolated and/or protected forests like the Adirondacks or the Ozarks to further limit ourselves to locations untouched by agricultural innovation[13]. The bottom panel shows that several of these locations experienced an even higher greening trend, closer to 1% per year. While vegetation indices like NDVI are not directly comparable to crop yields, this analysis implies that CO_2 fertilization likely played a material role in greening the forestland that is proximate to US croplands—in such a way that cannot be attributed technology-driven productivity drivers—by an order of magnitude similar to what we find in managed croplands.

6 Conclusion

We find a significant and robust CO_2 fertilization effect by linking OCO-2 satellite-measured CO_2 fluctuations to yield fluctuations of corn, soybeans, and winter wheat from 2015 to 2021. Our study spans more than half of the commercially-farmed area of these crops in the US and offers a test of whether the fertilization effects found in controlled experiments can be verified under real-world growing conditions. While panel models linking weather and yield anomalies have shown the possible detrimental effect of extreme heat on yield, the same setup can be used to show that localized CO_2 anomalies drive significant yield changes—outcomes

[12]The higher forestland average aligns with FACE experiments which find that trees are more responsive than herbaceous species like row crops to elevated CO_2 (Ainsworth and Long 2005)

[13]The selected forests span a range of biomes and age classes to address concerns that forest growth and succession dynamics are driving these trends

also reflected when utilizing alternate empirical approaches including a wind instrument and spatial first differences across neighboring counties. Our results suggest that a significant proportion of observed yield gains for corn, soybeans, and winter wheat since 1940 may be attributable to increases in CO_2, an important driver of agricultural productivity growth.

Our paper shows how satellite-based measures of CO_2 can be useful in complementing FACE field experiments, especially in the context of ensuring the external validity of estimates of the effect of CO_2 on agriculture and ecosystem functioning at a large scale. The approach can be extended to study real-world crop responses globally. Our results also merit consideration in the context of climate models used to estimate climate change impacts, but we caution against extrapolating the fertilization effect far into the future, which requires further assumptions about the functional form and the extent there are decreasing returns to further CO_2 increases, as well as uncertainty about future environmental interactions.

Relatedly, our analysis is focused on the US, and it is possible that fertilization effects will differ greatly across countries based on the prevailing crops and environmental conditions (McGrath and Lobell 2013), especially given that climate change alters the coupling of temperature, soil moisture and precipitation which determine crop yields (Proctor, Rigden, Chan, and Huybers 2022). Under future climate change, such heterogeneity could exacerbate spatial inequalities (Cruz Álvarez and Rossi-Hansberg 2021) and alter the comparative advantage of different regions (Costinot et al. 2014, Nath 2020) with large potential welfare effects that are worth investigating. While recent research has shown that mechanization significantly increased productivity and welfare (Caunedo and Kala 2022), as do property rights (Wüpper, Schlenker, Jain, Wang, and Finger 2022), we argue that environmental factors like CO_2 also play a crucial role.

We reiterate that climate change will likely have a negative impact on agriculture in aggregate, especially in regions exposed to extreme heat, and that CO_2-driven yield increases may be offset by effects on food nutrition and quality (Loladze 2002, Taub and Allen 2008, Myers et al. 2014). Nevertheless, this paper demonstrates that marginal increases in CO_2 can also have a strong countervailing fertilization effect—and that such effects may account for a material proportion of historical productivity improvements in US agriculture with implications for climate modelling and the literature on agricultural productivity and structural transformation.

References

Ainsworth, Elizabeth A, and Stephen P Long. 2005. "What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2." *New phytologist* 165 (2): 351–372.

Ainsworth, Elizabeth A., and Stephen P. Long. 2021. "30 years of free-air carbon dioxide enrichment (FACE): What have we learned about future crop productivity and its potential for adaptation?" *Global Change Biology* 27 (1): 27–49.

Ainsworth, Elizabeth A., and Alistair Rogers. 2007. "The response of photosynthesis and stomatal conductance to rising CO_2: mechanisms and environmental interactions." *Plant, Cell & Environment* 30 (3): 258–270.

Allen Jr., L. Hartwell, Jeff T. Baker, and Ken J. Boote. 1996. "The CO_2 fertilization effect: higher carbohydrate production and retention as biomass and seed yield."

Allen, LH, BA Kimball, JA Bunce, M Yoshimoto, Y Harazono, JT Baker, KJ Boote, and JW White. 2020. "Fluctuations of CO2 in Free-Air CO2 Enrichment (FACE) depress plant photosynthesis, growth, and yield." *Agricultural and Forest Meteorology* 284 107899.

Batts, G.R., J.I.L. Morison, R.H. Ellis, P. Hadley, and T.R. Wheeler. 1997. "Effects of CO_2 and temperature on growth and yield of crops of winter wheat over four seasons." *European Journal of Agronomy* 7 (1-3): 43–52.

Bebber, Daniel P, Timothy Holmes, and Sarah J Gurr. 2014. "The global spread of crop pests and pathogens." *Global Ecology and Biogeography* 23 (12): 1398–1407.

Boone, Christopher, Wolfram Schlenker, and Juha Siikamäki. 2019. "Ground-Level Ozone and Corn Yields in the United States." *CEEP Working Paper 7*.

Braun, Thomas, and Wolfram Schlenker. 2022. "Cooling Externality of Large-Scale Irrigation." *Working Paper*.

Bunce, J.A. 2012. "Responses of cotton and wheat photosynthesis and growth to cyclic variation in carbon dioxide concentration." *Photosynthetica* 50 (3): 395–400.

Caunedo, Julieta, and Namrata Kala. 2022. "Mechanizing Agriculture." *Working Paper.*

Costinot, Arnaud, Dave Donaldson, and Cory B. Smith. 2014. "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields around the World." *NBER Working Paper 20079.*

Coutts, Andrew M, Jason Beringer, and Nigel J Tapper. 2007. "Characteristics influencing the variability of urban CO2 fluxes in Melbourne, Australia." *Atmospheric Environment* 41 (1): 51–62.

Crisp, David. 2015. "Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2)." In *Earth Observing Systems XX*, Volume 9607. 960702. https://doi.org/10.1117/12.2187291.

Cruz Álvarez, José Luis, and Esteban Rossi-Hansberg. 2021. "The economic geography of global warming." *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2021-130): .

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What Do We Learn from the Weather? The New Climate-Economy Literature." *Journal of Economic Literature* 53 (3): 740–798.

Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif. 2019. "The mortality and medical costs of air pollution: Evidence from changes in wind direction." *American Economic Review* 109 (12): 4178–4219.

Deutsch, Curtis A, Joshua J Tewksbury, Michelle Tigchelaar, David S Battisti, Scott C Merrill, Raymond B Huey, and Rosamond L Naylor. 2018. "Increase in crop losses to insect pests in a warming climate." *Science* 361 (6405): 916–919.

Druckenmiller, Hannah, and Solomon Hsiang. 2019. "Accounting for Unobservable Heterogeneity in Cross Section Using Spatial First Differences." *NBER Working Paper 25177.*

Fitzgerald, Glenn J., Michael Tausz, Garry O'Leary et al. 2016. "Elevated atmospheric CO_2 can dramatically increase wheat yields in semi-arid environments and buffer against heat waves." *Global Change Biology* 22 (6): 2269–2284.

García, Ma Ángeles, Ma Luisa Sánchez, and Isidro A Pérez. 2012. "Differences between carbon dioxide levels over suburban and rural sites in Northern Spain." *Environmental Science and Pollution Research* 19 (2): 432–439.

Gollin, Douglas, Casper Worm Hansen, and Asger Mose Wingender. 2021. "Two Blades of Grass: The Impact of the Green Revolution." *Journal of Political Economy* 129 (8): 2344–2384. https://doi.org/10.1086/714444.

Gray, Sharon B, Orla Dermody, Stephanie P Klein et al. 2016. "Intensifying drought eliminates the expected benefits of elevated carbon dioxide for soybean." *Nature Plants* 2 (9): 1–8.

Hakkarainen, Janne, Iolanda Ialongo, and Johanna Tamminen. 2016. "Direct space-based observations of anthropogenic CO2 emission areas from OCO-2." *Geophysical Research Letters* 43 (21): 11–400.

Hendrey, GR, SP Long, IF McKee, and NR Baker. 1997. "Can photosynthesis respond to short-term fluctuations in atmospheric carbon dioxide?" *Photosynthesis Research* 51 (3): 179–184.

Hultgren, Andrew, Tamma Carleton, Michael Delgado et al. 2022. "Estimating Global Impacts to Agriculture from Climate Change Accounting for Adaptation." *Available at SSRN*.

Hungate, Bruce A., Jeffrey S. Dukes, M. Rebecca Shaw, Yiqi Luo, and Christopher B. Field. 2003. "Nitrogen and climate change." *Science* 302 (5650): 1512–1513.

Idso, Sherwood B, Craig D Idso, and Robert C Balling Jr. 2002. "Seasonal and diurnal variations of near-surface atmospheric CO2 concentration within a residential sector of the urban CO2 dome of Phoenix, AZ, USA." *Atmospheric environment* 36 (10): 1655–1660.

Jacobson, Andrew R., Kenneth N. Schuldt, John B. Miller et al. 2020. "CarbonTracker CT2019B." https://doi.org/10.25925/20201008.

Jorgenson, Dale W., and Frank M. Gollop. 1992. "Productivity Growth in U.S. Agriculture: A Postwar Perspective." *American Journal of Agricultural Economics* 74 (3): 745–750, `http://www.jstor.org/stable/1242588`.

Keenan, Trevor F., David Y. Hollinger, Gil Bohrer, Danilo Dragoni, J. William Munger, Hans Peter Schmid, and Andrew D. Richardson. 2013. "Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise." *Nature* 499 (7458): 324–327.

Kimball, B. A., C. F. Morris, P. J. Pinter Jr et al. 2001. "Elevated CO_2, drought and soil nitrogen effects on wheat grain quality." *New Phytologist* 150 (2): 295–303.

Kimball, Bruce A. 1983. "Carbon Dioxide and Agricultural Yield: An Assemblage and Analysis of 430 Prior Observations." *Agronomy Journal* 75 (5): 779–788.

Kimball, Bruce A. 2016. "Crop responses to elevated CO_2 and interactions with H_2O, N, and temperature." *Current Opinion in Plant Biology* 31 36–43.

Lambers, Hans, and Rafael S. Oliveira. 2019. "Photosynthesis, Respiration, and Long-Distance Transport: Photosynthesis." In *Plant Physiological Ecology*, 11–114, Springer.

Leakey, Andrew D.B., Martin Uribelarrea, Elizabeth A. Ainsworth, Shawna L. Naidu, Alistair Rogers, Donald R. Ort, and Stephen P. Long. 2006. "Photosynthesis, Productivity, and Yield of Maize Are Not Affected by Open-Air Elevation of CO_2 Concentration in the Absence of Drought." *Plant Physiology* 140 (2): 779–790.

Lobell, David B., and Senthold Asseng. 2017. "Comparing estimates of climate change impacts from process-based and statistical crop models." *Environmental Research Letters* 12 (1): 015001.

Lobell, David B, and George Azzari. 2017. "Satellite detection of rising maize yield heterogeneity in the US Midwest." *Environmental Research Letters* 12 (1): 014014.

Loladze, Irakli. 2002. "Rising atmospheric CO_2 and human nutrition: toward globally imbalanced plant stoichiometry?" *Trends in Ecology & Evolution* 17 (10): 457–461.

Long, Stephen P., Elizabeth A. Ainsworth, Andrew D. B. Leakey, Josef Nösberger, and Donald R. Ort. 2006. "Food for Thought: Lower-Than-Expected Crop Yield Stimulation with Rising CO_2 Concentrations." *Science* 312 (5782): 1918–1921.

Long, Stephen P., Elizabeth A. Ainsworth, Alistair Rogers, and Donald R. Ort. 2004. "Rising atmospheric carbon dioxide: plants FACE the future." *Annual Review of Plant Biology* 55 591–628.

Markelz, R. J. Cody, Reid S. Strellner, and Andrew D. B. Leakey. 2011. "Impairment of C_4 photosynthesis by drought is exacerbated by limiting nitrogen and ameliorated by elevated CO_2 in maize." *Journal of Experimental Botany* 62 (9): 3235–3246.

Massen, Francis, and Ernst-Georg Beck. 2011. "Accurate estimation of CO 2 background level from near ground measurements at non-mixed environments." In *The Economic, Social and Political Elements of Climate Change*, 509–522, Springer.

McArthur, John W, and Gordon C McCord. 2017. "Fertilizing growth: Agricultural inputs and their effects in economic development." *Journal of development economics* 127 133–152.

McGrath, Justin M., and David B. Lobell. 2013. "Regional disparities in the CO_2 fertilization effect and implications for crop yields." *Environmental Research Letters* 8 (1): 014054.

Moore, Frances C., Uris Lantz C. Baldos, and Thomas Hertel. 2017. "Economic impacts of climate change on agriculture: a comparison of process-based and statistical yield models." *Environmental Research Letters* 12 (6): 065008.

Morgan, Jack A., Daniel R. LeCain, Elise Pendall et al. 2011. "C4 grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland." *Nature* 476 (7359): 202–205.

Mortensen, Leiv M. 1987. "CO_2 Enrichment in Greenhouses. Crop Responses." *Scientia Horticulturae* 33 (1-2): 1–25.

Moscona, Jacob, and Karthik Sastry. 2022. "Inappropriate Technology: Evidence from Global Agriculture." *Available at SSRN 3886019*.

Myers, Samuel S., Antonella Zanobetti, Itai Kloog et al. 2014. "Increasing CO_2 threatens human nutrition." *Nature* 510 (7503): 139–142.

Nath, Ishan B. 2020. "The food problem and the aggregate productivity consequences of climate change."Technical report, National Bureau of Economic Research.

Olmstead, Alan L, and Paul W Rhode. 2002. "The red queen and the hard reds: Productivity growth in American wheat, 1800–1940." *The Journal of Economic History* 62 (4): 929–966.

Ottman, M. J., B. A. Kimball, P.J. Pinter et al. 2001. "Elevated CO_2 increases sorghum biomass under drought conditions." *New Phytologist* 150 (2): 261–273.

Pardey, Philip G, and Julian M Alston. 2021. "Unpacking the Agricultural Black Box: The Rise and Fall of American Farm Productivity Growth." *The Journal of Economic History* 81 (1): 114–155.

Parker, Wendy S. 2016. "Reanalyses and observations: What's the difference?" *Bulletin of the American Meteorological Society* 97 (9): 1565–1572.

Proctor, Jonathan, Angela Rigden, Duo Chan, and Peter Huybers. 2022. "More accurate specification of water supply shows its importance for global crop production." *Nature Food* 3 (9): 753–763.

Reich, Peter B., Sarah E. Hobbie, Tali Lee et al. 2006. "Nitrogen limitation constrains sustainability of ecosystem response to CO_2." *Nature* 440 (7086): 922–925.

Rizzo, Gonzalo, Juan Pablo Monzon, Fatima A Tenorio, Réka Howard, Kenneth G Cassman, and Patricio Grassini. 2022. "Climate and agronomy, not genetics, underpin recent maize yield gains in favorable environments." *Proceedings of the National Academy of Sciences* 119 (4): e2113629119.

Schlenker, Wolfram, and Michael J. Roberts. 2009. "Nonlinear Temperature Effects Indicate Severe Damages to U.S. Crop Yields under Climate Change." *Proceedings of the National Academy of Sciences* 106 (37): 15594–15598.

Schlenker, Wolfram, and W Reed Walker. 2016. "Airports, Air Pollution, and Contemporaneous Health." *Review of Economic Studies* 2 (83): 768–809.

Specht, JE, DJ Hume, and SV Kumudini. 1999. "Soybean yield potential—a genetic and physiological perspective." *Crop science* 39 (6): 1560–1570.

Tack, Jesse, Andrew Barkley, and Lawton Lanier Nalley. 2015. "Effect of Warming temperatures on US Wheat Yields." *Proceedings of the National Academy of Sciences* 112 (22): 6931–6936.

Taub, Daniel R., and Brian Millerand Holly Allen. 2008. "Effects of elevated CO_2 on the protein concentration of food crops: a meta-analysis." *Global Change Biology* 14 (3): 565–575.

Toreti, Andrea, Delphine Deryng, Francesco N. Tubiello et al. 2020. "Narrowing uncertainties in the effects of elevated CO_2 on crops." *Nature Food* 1 (12): 775–782.

Vermote, Eric, Chris Justice, Ivan Csiszar et al. 2014. "NOAA Climate Data Record (CDR) of Normalized Difference Vegetation Index (NDVI), Version 4." https://doi: 10.7289/V5PZ56R6.

Wang, Songhan, Yongguang Zhang, Weimin Ju et al. 2020. "Recent global decline of CO_2 fertilization effects on vegetation photosynthesis." *Science* 370 (6522): 1295–1300.

Wang, Sun Ling, Paul Heisey, David Schimmelpfennig, and V. Eldon Ball. 2015. "Agricultural Productivity Growth in the United States: Measurement, Trends, and Drivers." *Economic Research Service, Paper No. 189.*

Weir, Brad, and Lesley Ott. 2022. "OCO-2 GEOS Level 3 daily, 0.5x0.625 assimilated CO2 V10r." https://doi.org/10.5067/Y9M4NM9MPCGH.

Wüpper, David, Wolfram Schlenker, Meha Jain, Haoyu Wang, and Robert Finger. 2022. "Property Rights and Global Crop Yields." *Working Paper.*

Xueref-Remy, Irène, Elsa Dieudonné, Cyrille Vuillemin et al. 2018. "Diurnal, synoptic and seasonal variability of atmospheric CO 2 in the Paris megacity area." *Atmospheric Chemistry and Physics* 18 (5): 3335–3362.

Yin, Yi, Brendan Byrne, Junjie Liu et al. 2020. "Cropland carbon uptake delayed and reduced by 2019 Midwest floods." *AGU Advances* 1 (1): e2019AV000140. https://doi.org/10.1029/2019AV000140.

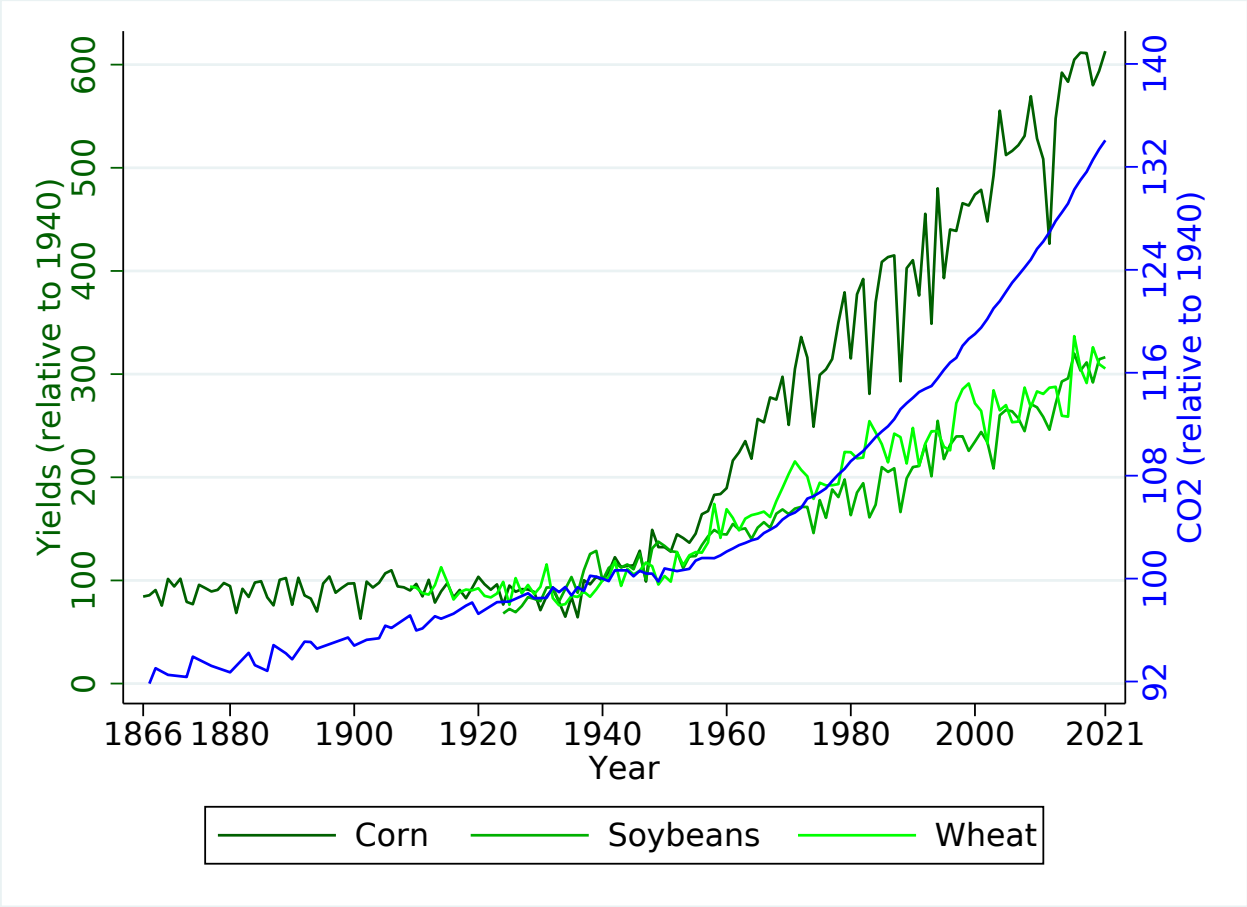
Zheng, Yunpu, Chunlin He, Lili Guo, Lihua Hao, Dongjuan Cheng, Fei Li, Zhengping Peng, and Ming Xu. 2020. "Soil water status triggers CO_2 fertilization effect on the growth of winter wheat (Triticum aestivum)." *Agricultural and Forest Meteorology* 291 108097.

Zhu, Zaichun, Shilong Piao, Ranga B. Myneni et al. 2016. "Greening of the Earth and its Drivers." *Nature climate change* 6 (8): 791–795.

Ziska, Lewis H., and James A. Bunce. 2007. "Predicting the impact of changing CO_2 on crop yields: some thoughts on food." *New Phytologist* 175 (4): 607–618.

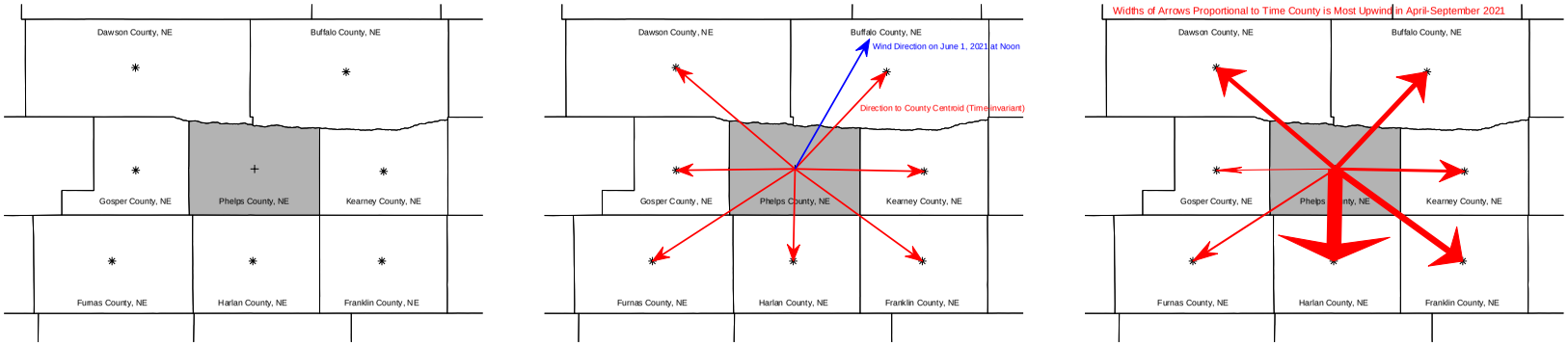
Figures and Tables

Figure 1: Annual Yields and CO_2



Notes: Figure displays the evolution of yearly aggregate US yields (left axis) and annual CO_2 averages (right axis). Each time series is normalized relative to 1940 (value = 100). Aggregate US corn yields are shown in dark green from 1866-2021, aggregate US soybeans yields in green from 1924-2021, and aggregate US winter wheat yields in light green from 1909-2021, the years for which the data is reported by the National Agricultural Statistics Service (NASS) by USDA. The annual average CO_2 level is added in blue from NOAA.

Figure 2: IV Setup



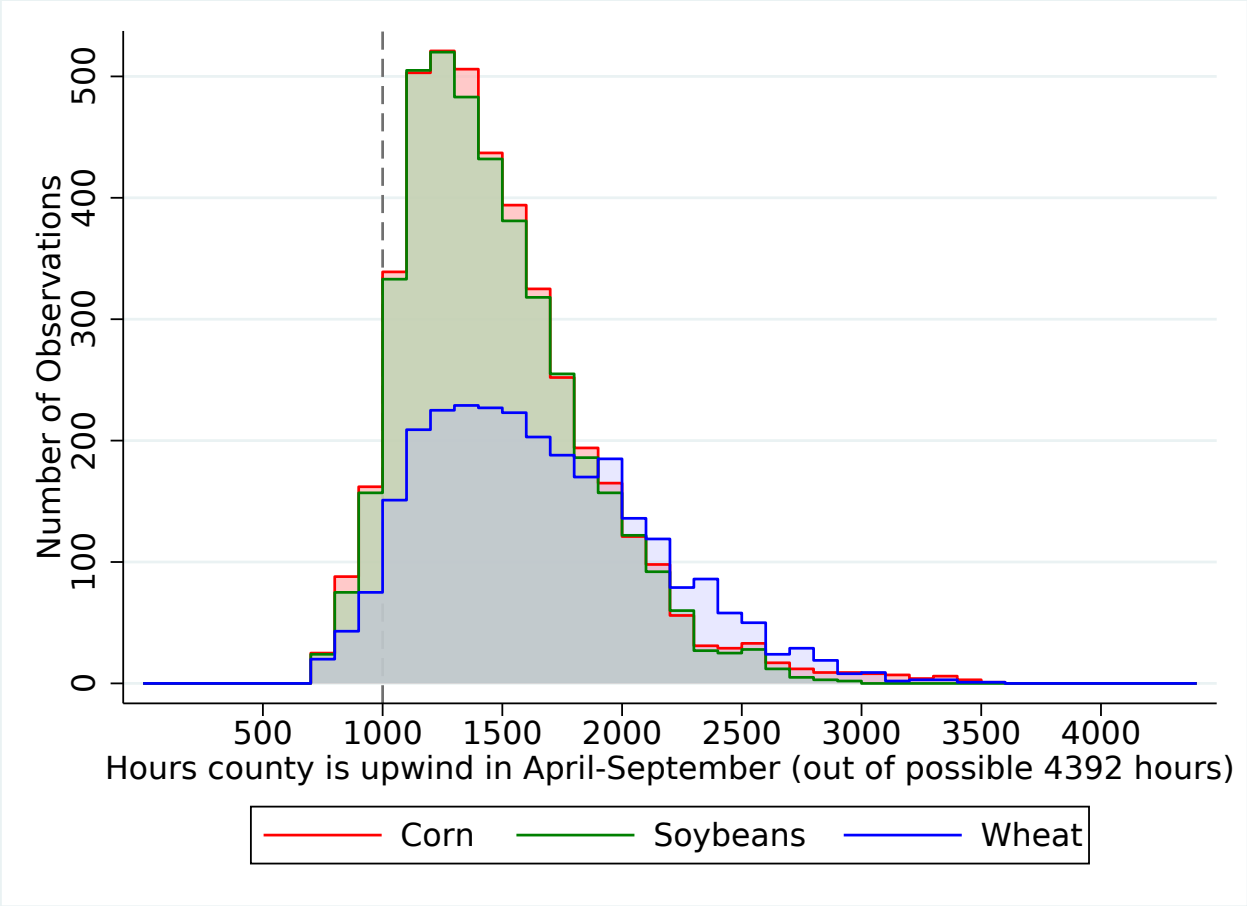
Notes: Figure outlines the construction of the instrument using Phelps County, Nebraska (fips code 31137) as an example.

The left panel demonstrates the first step: we derive the centroid (shown as +) for Phelps county (shown in grey) as the cropland-weighted average of all grids in a county. The weight is the combined corn, soybeans and winter wheat acreage from the Cropland Data Layer. We pair Phelps' centroid with the centroids of all its seven neighbors (shown as $*$), which are *not* cropland-area weighted as we are looking for all CO_2 readings as instrument, not just over the agricultural area.

The middle panel demonstrates how we derive an upwind county: the red arrows show the direction in which neighbor centroids are located, which are time invariant. The blue arrow shows the wind direction at noon on June 1st, 2021. It is constructed from the hourly NLDAS data, where we again take the cropland-area weighted average of all NLDAS grid cells in Phelps county. We derive the cosine between the direction of neighboring counties (red arrows) and the wind direction (blue arrow) and choose the county whose cosine is closest to -1 as "upwind".

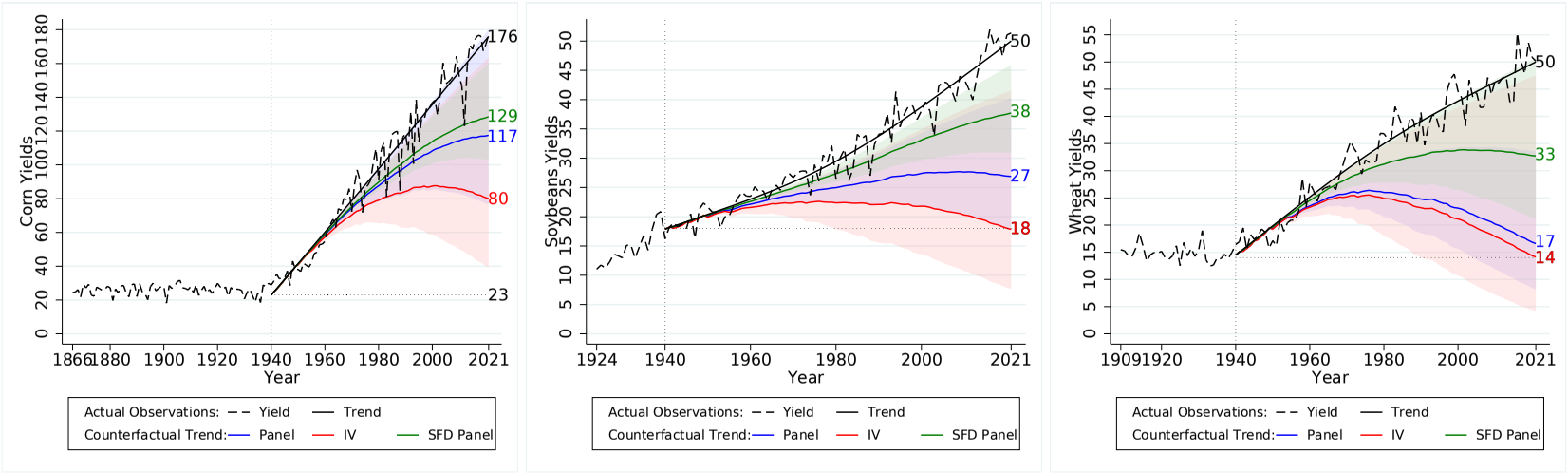
The right panel shows how often a neighboring county is "upwind" during the 4392 hours in April-September of a year, where the widths of the arrows are proportional to the frequency. he CO_2 readings over Harlan county (entire county, not just agricultural area) would hence be used as instrument for the CO_2 over the agricultural area of Phelps county in 2021 as it is 1789 hours (41% of the time) upwind. "Upwind" counties can change year-to-year.

Figure 3: Histogram of Hours a County is Upwind



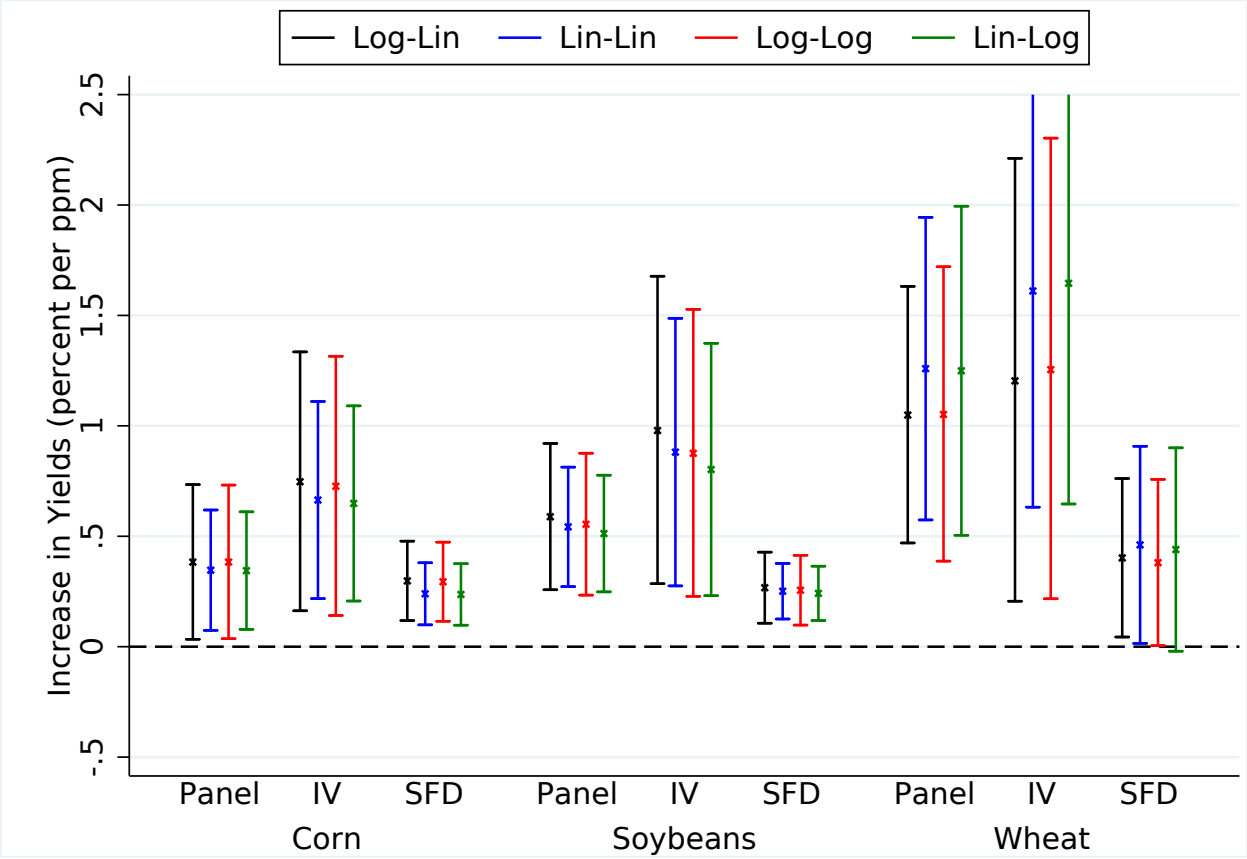
Notes: Figure displays histograms of the number of hours the county used in the IV is upwind (out of the possible 4392 hours in April-September of a year). Data are shown for counties east of the 100° meridian for corn and soybeans (shown in green in Figure A2), and east of the Rocky Mountains for winter wheat (shown in green and blue in Figure A2).

Figure 4: Counterfactual if CO_2 Remained at 1940 Levels



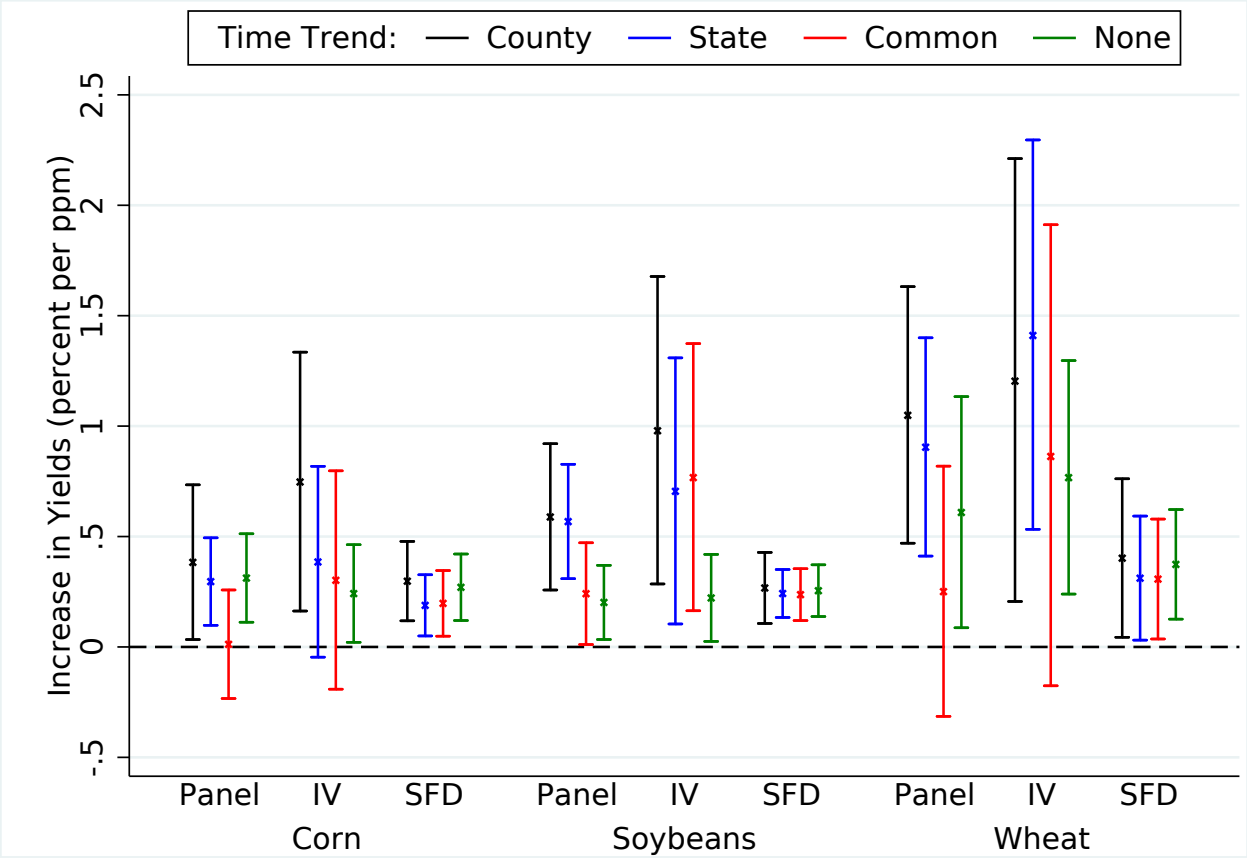
Notes: Figure displays the evolution of yearly aggregate US yields as dashed black line. The trend using restricted cubic splines with 3 knots (two variables) since 1940 is fitted and shown as solid black line. The graph adds three counterfactuals where CO_2 levels are kept constant at 1940 level, assuming the CO_2 fertilization effect that was estimated using 2015-2021 data applied throughout 1940-2021. The blue line uses the coefficient estimates from columns (1c), (2c) and (3c) of Table 2 to simulate the effect on corn, soybean and winter wheat yields, respectively. The red line uses the IV estimates from columns (1c), (2c) and (3c) of Table 3, and the green line uses the Spatial First Difference (SFD) estimates from columns (1c), (2c) and (3c) of Table 4. We only adjust the yield trend by the difference in actual CO_2 compared to CO_2 in 1940 while leaving all other factors unchanged. The 95% confidence confidence bands are added.

Figure 5: Sensitivity to Functional Form



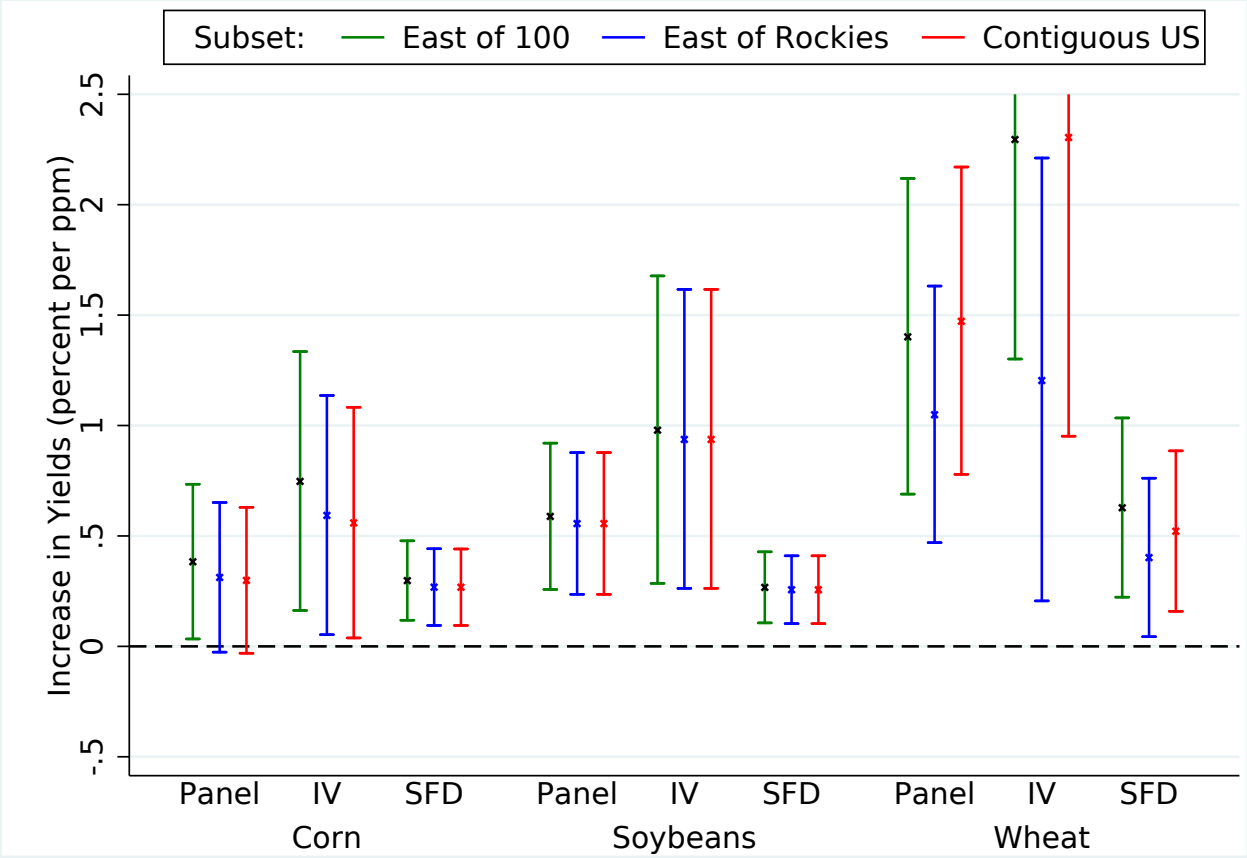
Notes: Figure presents sensitivity check to what functional form is chosen. Graphs shows the effect of a one ppm increase in CO_2 on aggregate yields in percent as well as the 90% confidence bands. Black lines show the baseline results from columns (c) in Table 2 for the panel regression, Table 3 for the wind IV, and Table 4 for the spatial first difference that regress log yields on CO_2 levels (Log-Lin model). Blue lines regress yields on CO_2 (Lin-Lin model), red lines regress log yields on log CO_2 (Log-Log model), while green lines regress yields on log CO_2 (Lin-Log model). All regressions include county fixed effects as well as county-specific time trends and control for four weather and five criteria air pollution variables. Errors are clustered at the state level.

Figure 6: Sensitivity to Included Time Trend



Notes: Figure presents sensitivity check to what time controls are included. Graphs shows the effect of a one ppm increase in CO_2 on aggregate yields in percent as well as the 90% confidence bands. Black lines show the baseline results from columns (c) in Table 2 for the panel regression, Table 3 for the wind IV, and Table 4 for the spatial first difference that all included county-specific time trends. Blue lines instead include state-specific time trends, red lines include a common time-trend, and green lines include no time trend at all. All regressions include county fixed effects and control for four weather and five criteria air pollution variables. Errors are clustered at the state level.

Figure 7: Sensitivity to Geographic Subset



Notes: Figure presents sensitivity check to what what geographic subset is included in the analysis. Graphs shows the effect of a one ppm increase in CO_2 on aggregate yields in percent as well as the 90% confidence bands. Green lines show the results when counties east of the 100 meridian are used in the analysis, while blue lines show the results when counties east of the Rocky Mountains are used, and red lines show the results when all counties of the contiguous US are used. The subsets are shown in Figure A2 and exclude Florida. All regressions include county fixed effects as well as county-specific time trends and control for four weather and five criteria air pollution variables. Errors are clustered at the state level.

Table 1: Summary Statistics of IV Setup

	Corn	Soybeans	Winter wheat
	Panel A: All counties		
Hours Upwind			
Mean	1482	1461	1604
Range	[722,3511]	[722,2970]	[722,3511]
Standard deviation	(1482)	(1461)	(1604)
County Upwind			
Always the same	738	694	397
One of two counties	202	209	78
One of three or more counties	21	17	6
	Panel B: At least 1000 hours		
Hours Upwind			
Mean	1527	1503	1658
Range	[1000,3511]	[1000,2970]	[1001,3511]
Standard deviation	(1527)	(1503)	(1658)
County Upwind			
Always the same	721	683	379
One of two counties	169	172	67
One of three or more counties	6	5	1

Notes: Tables provides summary statistics for the IV setup that is outlined in Figure 2. The first three rows of each panel give the number of hours a county is upwind in the IV setup (The corresponding histogram is given in Figure 3). The last three rows in each panel display how much variation there is year-to-year in which county is upwind. For the majority of counties, the upwind county is the same in every year. Panel A includes all counties using the most frequent upwind neighbor irrespective of how many hours it is upwind, while panel B forces the upwind county to be at least 1000 hours upwind in April-September.

Table 2: Baseline Panel Regression

	Corn			Soybeans			Winter Wheat		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
CO_2 (ppm)	0.299	0.405**	0.382*	0.396*	0.502**	0.587***	1.305***	1.540***	1.044***
	(0.192)	(0.184)	(0.206)	(0.223)	(0.227)	(0.195)	(0.385)	(0.424)	(0.340)
Moderate DDay		0.224**	0.223***		0.322***	0.525***		-0.117	0.107
		(0.086)	(0.080)		(0.101)	(0.090)		(0.250)	(0.251)
Extreme DDay		-0.277***	-0.281***		-0.518***	-0.522***		-0.026	-0.069
		(0.063)	(0.063)		(0.091)	(0.083)		(0.164)	(0.165)
Precipitation		0.468***	0.493**		0.249	0.241		1.455**	1.099*
		(0.169)	(0.203)		(0.172)	(0.196)		(0.562)	(0.578)
Prec. squared		-0.380***	-0.392***		-0.193*	-0.201		-1.203***	-1.037**
		(0.113)	(0.133)		(0.112)	(0.137)		(0.400)	(0.409)
Mean CO			-0.121			-0.741**			0.016
			(0.310)			(0.337)			(0.548)
Mean NO_2			-0.017*			-0.026			0.009
			(0.009)			(0.017)			(0.024)
Mean O_3			0.439			-1.791**			-0.859
			(0.854)			(0.694)			(1.367)
Mean PM_{10}			0.023			0.088			-2.128**
			(0.562)			(0.326)			(0.770)
Mean SO_2			-0.010			-0.034**			-0.019
			(0.014)			(0.013)			(0.017)
R^2	0.8553	0.8671	0.8676	0.8514	0.8694	0.8787	0.8495	0.8584	0.8653
Observations	5798	5798	5798	5512	5512	5512	3545	3545	3545

Notes: Tables regresses log yields on CO_2 readings from OCO-2. All regressions include county fixed effects as well as county-specific time trends. Columns (b) furthermore include four weather: moderate degree days ×1000 (degree days 10-29°C for corn and degree days 10-30°C for soybeans and winter wheat) and extreme degree days ×100 (degree days above 29°C for corn and degree days above 30°C for soybeans and winter wheat) as well as a quadratic in precipitation (in meters). Columns (c) additionally include mean pollution levels of five criteria air pollutants: CO (in ppm), NO_2 (in ppm), O_3 (in 100 ppb), PM_{10} (in 100ppm) and SO_2 (in ppm). All weather and pollution variables are constructed over the growing season April-September. Corn and soybean regression use counties east of the 100° meridian (green in Figure A2), while wheat regressions include all counties east of the Rocky Mountains (green and blue in Figure A2). Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

Table 3: IV Regression using Upwind CO_2

	Corn			Soybeans			Winter Wheat		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
Min hours	-	0	1000	-	0	1000	-	0	1000
OLS Regression									
CO_2 (ppm)	0.382*	0.521**	0.561**	0.587***	0.576**	0.610**	1.044***	1.079**	1.053**
	(0.206)	(0.213)	(0.218)	(0.195)	(0.263)	(0.277)	(0.340)	(0.421)	(0.459)
IV Regression									
CO_2 (ppm)		0.724**	0.744**		0.976***	0.974**		1.274**	1.197**
		(0.336)	(0.344)		(0.375)	(0.408)		(0.533)	(0.586)
F (1st stage)		15.99	15.91		16.08	16.20		13.15	12.83
Observations	5798	4355	4014	5512	4202	3882	3545	2775	2598

Notes: Tables regresses log yields on CO_2 readings from OCO-2. Each coefficient is from a separate regression. Columns (a) replicates Table 2 specification (c). Columns (b) and (c) replicate the same OLS regression for the set of observations that are included in the IV regression in the second row, where CO_2 readings are instrumented with the CO_2 reading (averaged over the entire county, not just agricultural area) from the upwind county. Columns (b) and (c) differ in what is classified as upwind county. Column (b) uses the county that is most frequently upwind during April - September using the hourly wind data for NLDAS (there are 4392 hours in April-September). Columns (c) only uses a county as upwind if it at least 1000 hours upwind. All regressions include county fixed effects as well as county-level time trends and control for four weather and five pollution variables. Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

Table 4: Spatial First Difference Regression

	Corn		Soybeans		Winter Wheat				
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)			
CO_2 (ppm)	0.277**	0.294***	0.297***	0.259**	0.276***	0.267***	0.462**	0.513**	0.401*
	(0.105)	(0.104)	(0.106)	(0.099)	(0.097)	(0.095)	(0.200)	(0.205)	(0.211)
Moderate DDay		0.211*	0.115		0.303**	0.389***		-0.028	0.022
		(0.103)	(0.110)		(0.113)	(0.102)		(0.211)	(0.223)
Extreme DDay		-0.254***	-0.261***		-0.432***	-0.468***		-0.010	-0.052
		(0.079)	(0.083)		(0.075)	(0.072)		(0.124)	(0.133)
Precipitation		0.468***	0.485***		0.404***	0.412***		1.052**	0.889**
		(0.123)	(0.126)		(0.137)	(0.144)		(0.377)	(0.359)
Prec. squared		-0.368***	-0.366***		-0.279***	-0.288***		-0.841***	-0.768***
		(0.082)	(0.085)		(0.089)	(0.099)		(0.243)	(0.228)
Mean CO			0.009			-0.578**			-0.085
			(0.158)			(0.226)			(0.374)
Mean NO_2			-0.010			-0.024**			-0.011
			(0.008)			(0.010)			(0.017)
Mean O_3			1.132			-0.180			0.661
			(0.723)			(0.434)			(0.859)
Mean PM_{10}			0.335			-0.054			-1.711***
			(0.302)			(0.264)			(0.531)
Mean SO_2			-0.020**			-0.030***			-0.018
			(0.008)			(0.011)			(0.021)
Observations	10224	10224	10224	9960	9960	9960	5296	5296	5296

Notes: Table is analogous to Table 2 but uses spatial first difference rather than a standard fixed effect regression. It regresses the change in log yield residuals (after removing county fixed effects as well as county-specific time trends) between neighboring counties on the corresponding difference in CO_2 residuals among neighboring counties. Columns (b) furthermore include the difference in residuals from four weather variables, while columns (c) include the difference in residuals from the five criteria air pollutants. The weather variables are moderate degree days $\times 1000$ (degree days 10-29°C for corn and degree days 10-30°C for soybeans and winter wheat) and extreme degree days $\times 100$ (degree days above 29°C for corn and degree days above 30°C for soybeans and winter wheat) as well as a quadratic in precipitation (in meters). The pollution variables are the average concentration of CO (in ppm), NO_2 (in ppm), O_3 (in 100 ppb), PM_{10} (in 100ppm) and SO_2 (in ppm). All weather and pollution variables are constructed over the growing season April-September. Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

Appendix - For Online Publication

Environmental Drivers of Agricultural Productivity Growth: CO_2 Fertilization of US Field Crops

Charles A. Taylor[1] and Wolfram Schlenker[2,3]

List of Figures

A1 Seasonality in CO_2 ii
A2 Number of Observations per County in 2015-2021 iii
A4 Identifying Variation Used in Analysis - Residuals From Trend v
A5 Scatter Plots of Log Yield Anomalies against CO_2 Anomalies vi
B1 CarbonTracker: Number of Observations per County in 2000-2018 x
B2 Crossplots of CO_2 Anomalies in CarbonTracker and OCO-2 (2015-2018) . . . xi
B3 CarbonTracker: Counterfactual if CO_2 Levels Remained at 1940 Levels . . . xii
B4 CarbonTracker: Sensitivity to Functional Form xiii
B5 CarbonTracker: Sensitivity to Included Time Trend xiv
B6 CarbonTracker: Sensitivity to Geographic Subset xv
C1 Annual Trends in NDVI Vegetation from the AVHRR Satellite, 1982-2013 . . xix

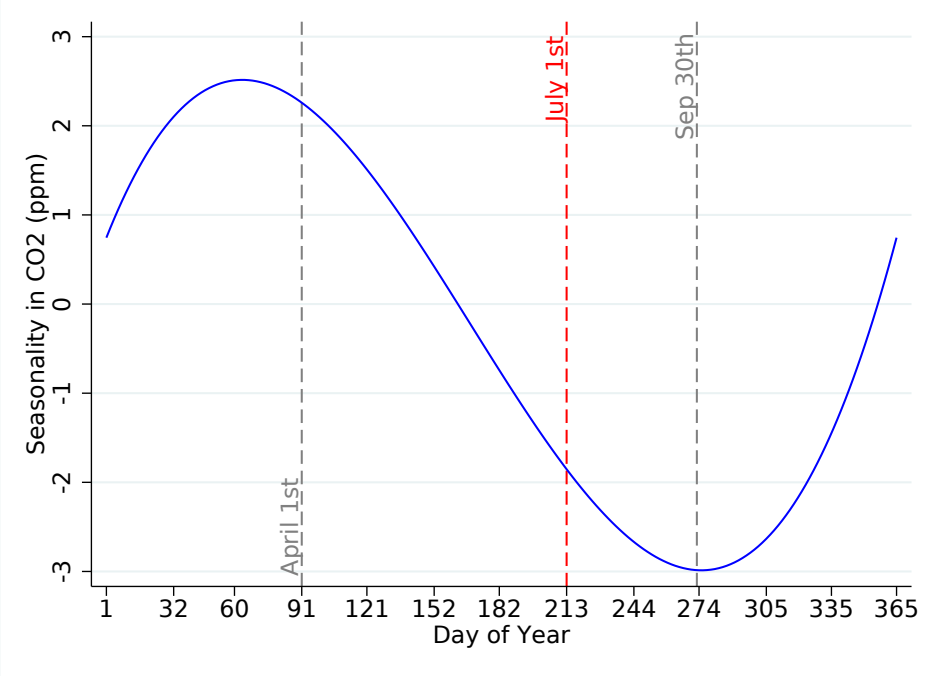
List of Tables

A1 Spatial First Difference Regression - Cross-Section viii
B1 CarbonTracker: Panel Regression xvi
B2 CarbonTracker: IV Regression using Upwind CO_2 xvii
B3 CarbonTracker: Spatial First Difference Regression xviii

[1] Department of Agricultural and Resource Economics, UC Berkeley, 207 Giannini Hall, Berkeley, CA 94720.
[2] School of International and Public Affairs, Columbia University, 420 West 118th St., New York, NY 10027.
[3] National Bureau of Economic Research (NBER), 1050 Massachusetts Ave., Cambridge, MA 02138.

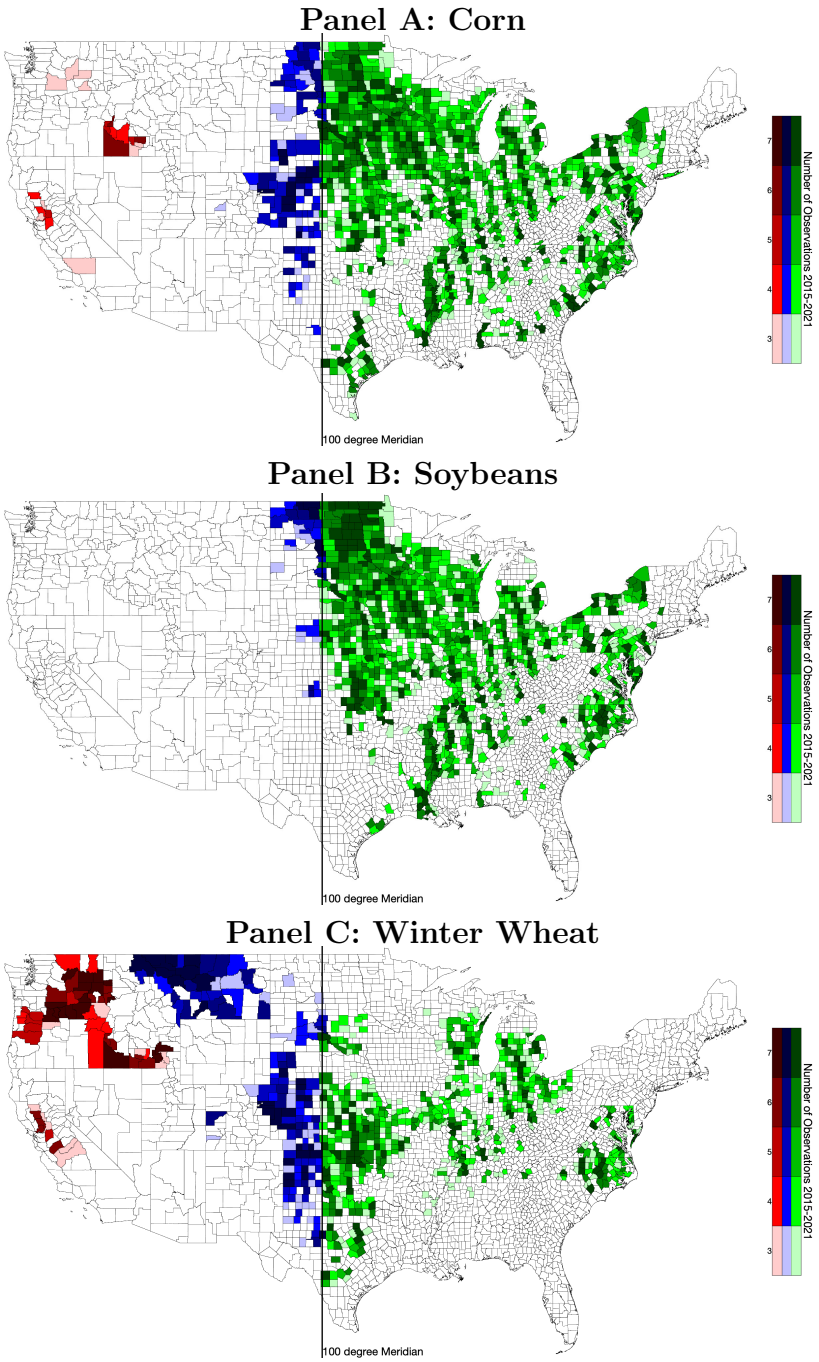
A Additional Analysis Using OCO-2 Data

Figure A1: Seasonality in CO_2



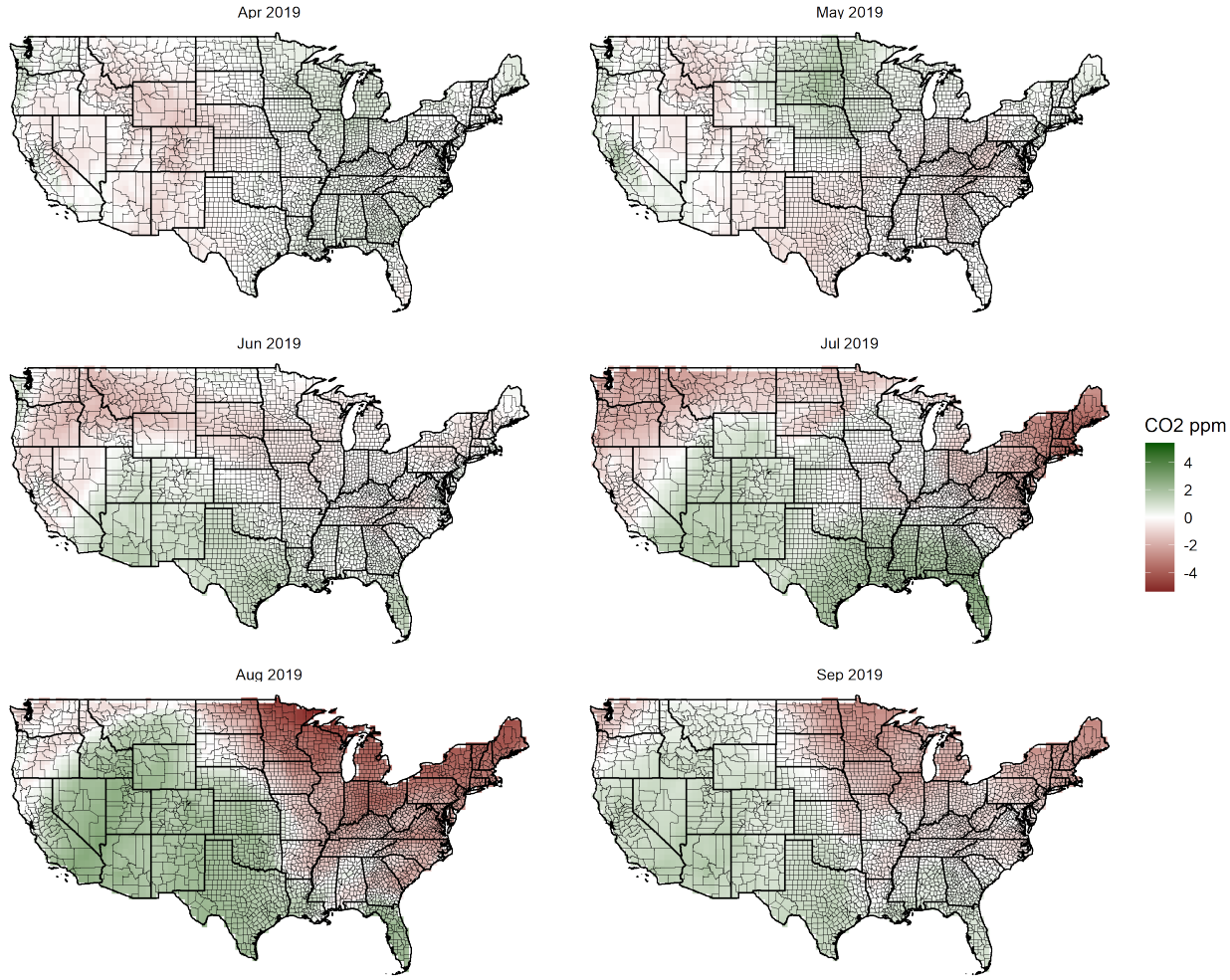
Notes: Chart displays the seasonality in CO_2. To make readings comparable, they are seasonality-adjusted to July 1st (red dashed line) of that year, i.e., a reading on a particular day is corrected by the difference between the July 1st value of the above seasonality curve and the value of the seasonality curve on the day of the measurement. The seasonality curve are estimated using all OCO2 readings without quality flags over the contiguous US using a 4^{th}-order Chebyschev polynomial in the day of year as well as a linear time trend. Since years have different numbers of days, we normalize January 1st to -1 and December 31st to 1. The seasonality regression is constrained so the value at the end of the year (December 31st) equals the value at the beginning of the year (January 1st). The main growing season for corn and soybeans (April-September) is added as grey dashed lines.

Figure A2: Number of Observations per County in 2015-2021



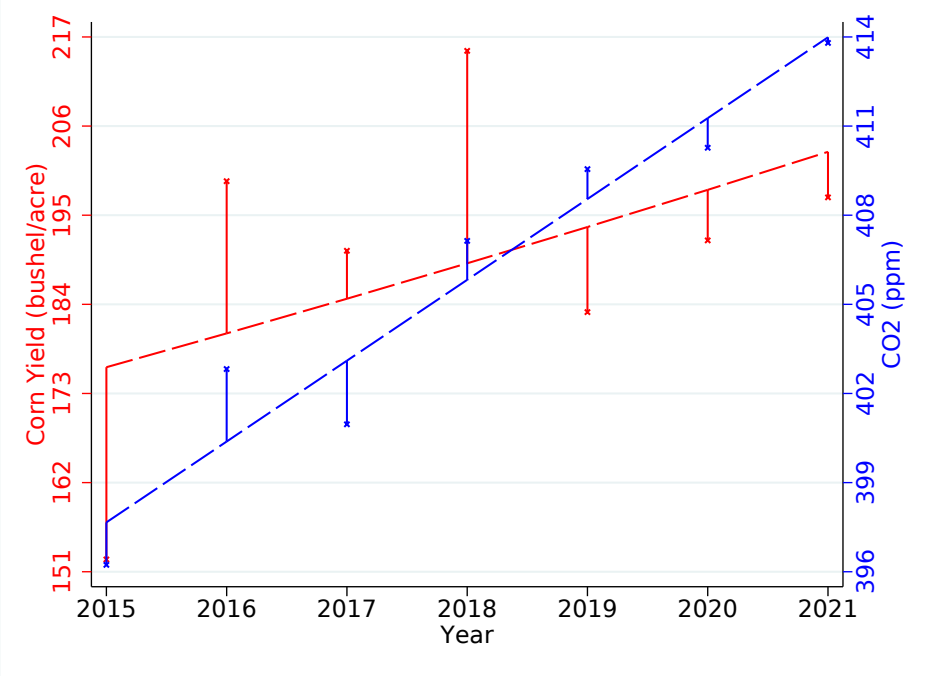
Notes: Figure displays the number of observations per county in the date set, i.e., where yield, weather, criteria air pollution, as well as CO_2 data from OCO-2 are available over our sample period 2015-2021. We split the analysis into three geographic subsets: east of the 100° meridian excluding Florida (Schlenker and Roberts 2009) shown in shades of green, inter-mountain states (Montana, Wyoming, Colorado, and New Mexico) shown in blue, and western states (California, Arizona, Utah, Nevada, Oregon, Idaho, and Washington) shown in red. Since our specification includes county fixed effects and county-specific time trends, we require at least 3 observation to be included in the dataset.

Figure A3: Spatial Variation in CO_2 Anomalies within a Growing Season, 2019



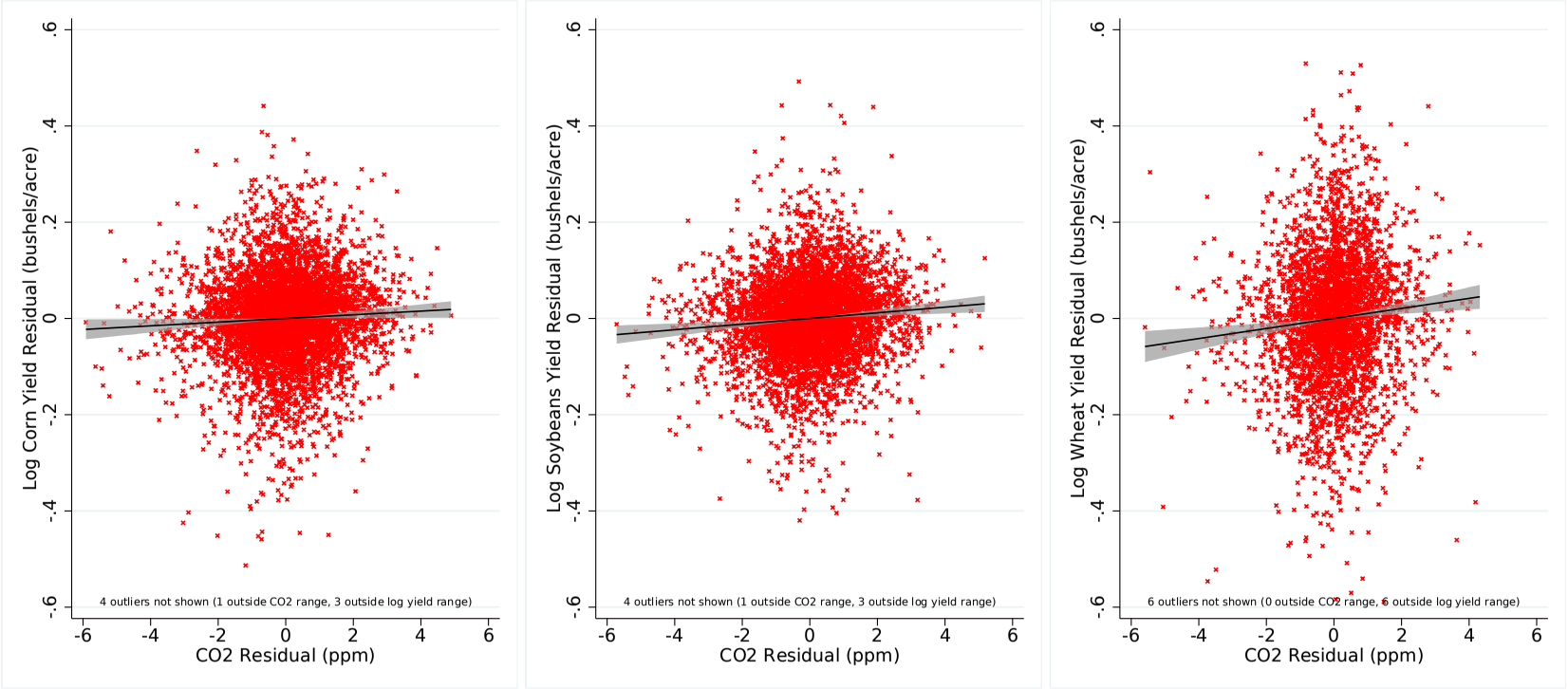
Notes: Figure displays CO_2 anomalies relative to the mean value on the first day of each month during the growing season in 2019 using OCO-2's GEOS Level 3 daily modelled product (Weir and Ott 2022). Our analysis uses the raw satellite measurements, but we are showing here the spatial extend of the anomalies in an interpolated product.

Figure A4: Identifying Variation Used in Analysis - Residuals From Trend



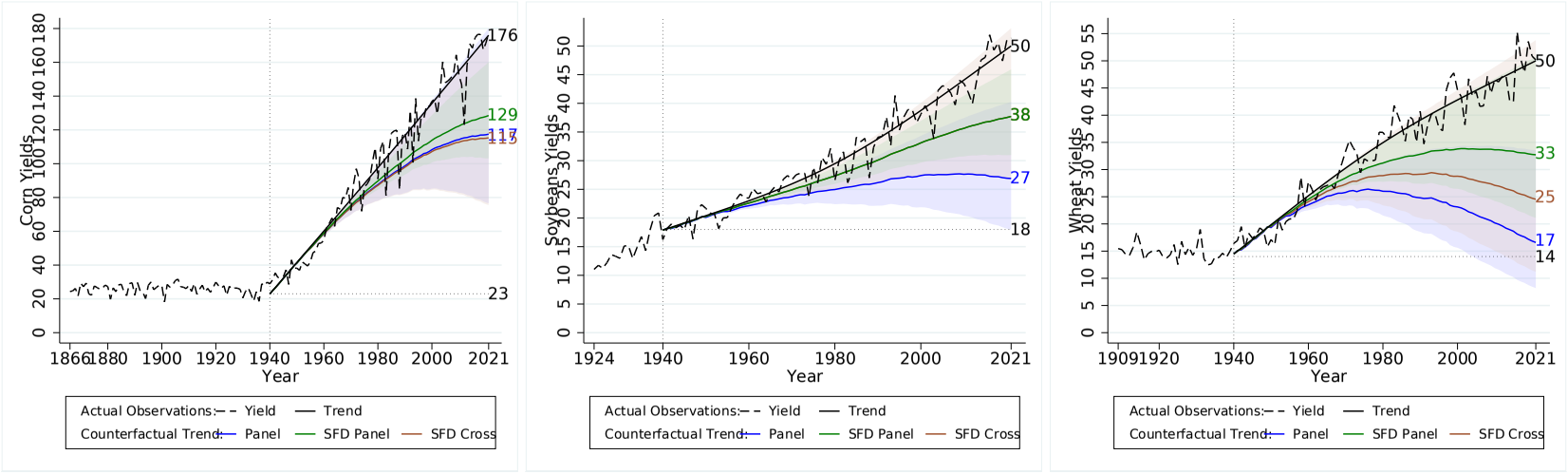
Notes: Figure displays the variation used in the statistical analysis. We include county-fixed effects and county-specific time trends. This is equivalent to fitting a time trend (shown as dashed lines) to both yields and CO_2 readings for each county and then looking at the residuals. The above figure shows this for Macoupin county in Illinois (FIPS code 17117), which has Madison County, IL as most frequent upwind county (eastern-edge of Saint Louis metropolitan area, where CO_2 anomalies should mainly be driven by non-agricultural factors). Corn anomalies are shown as solid red lines, while CO_2 anomalies are shown as blue lines. When CO_2 positively (negatively) deviates from the trend, so do yields. The correlation of the residuals is 0.57. Figure A5 shows the cross-plot for all observations (counties and years) after additionally removing the effect of weather and criteria air pollutants.

Figure A5: Scatter Plots of Log Yield Anomalies against CO_2 Anomalies



Notes: Figure summarizes anomalies after county fixed effects, county-specific time trends, the four weather and five criteria air pollution variables are partialled out. This is the variation we use to identify the effect of CO_2 on yields in the panel regressions in Table 2. The left column displays the results for corn, the middle for soybeans, and the right for winter wheat. There are a few outliers that are not shown for clarity, but including or excluding them has no effect on the estimate coefficient given the large sample size.

Figure A6: Counterfactual if CO_2 Levels Remained at 1940 Levels - SFD Cross-section



Notes: Figure replicates Figure 4 and also includes the spatial first difference approach (SFD) using cross-sectional differences in average outcomes. Figure displays the evolution of yearly aggregate US yields as dashed black line. The trend using restricted cubic splines with 3 knots (two variables) since 1940 is fitted and shown as solid black line. The graph adds three counterfactuals where CO_2 levels are kept constant at 1940 level, assuming the CO_2 fertilization effect that was estimated using 2015-2021 data applied throughout 1940-2021. The blue line uses the coefficient estimates from columns (c) of Table 2 to simulate the effect on yields, while the the green line uses the Spatial First Difference (SFD) estimates from Table 4 using residuals from each year (panel variation), and the brown lines uses the Spatial First Difference (SFD) estimates using averages across years (cross-sectional variation) from columns (c) in Table A1. We only adjust the yield trend by the difference in actual CO_2 compared to CO_2 in 1940 while leaving all other factors unchanged. The 95% confidence confidence bands are added.

Table A1: Spatial First Difference Regression - Cross-Section

	Corn			Soybeans			Winter Wheat		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
CO_2 (ppm)	0.338*	0.376*	0.400*	0.174	0.282	0.267	0.635	0.683*	0.673*
	(0.198)	(0.186)	(0.204)	(0.199)	(0.178)	(0.165)	(0.451)	(0.362)	(0.379)
Moderate DDay		0.182**	0.172**		0.164**	0.184**		0.047	0.047
		(0.081)	(0.081)		(0.079)	(0.082)		(0.108)	(0.117)
Extreme DDay		-0.200**	-0.192**		-0.341***	-0.335***		-0.086*	-0.087*
		(0.077)	(0.074)		(0.108)	(0.109)		(0.047)	(0.044)
Precipitation		0.018	-0.006		0.652**	0.616*		0.291	0.237
		(0.309)	(0.298)		(0.315)	(0.307)		(0.605)	(0.506)
Prec. squared		0.040	0.061		-0.451*	-0.429*		-0.271	-0.243
		(0.229)	(0.221)		(0.249)	(0.245)		(0.380)	(0.330)
Mean CO			-0.392			-0.572**			-0.375
			(0.234)			(0.233)			(0.383)
Mean NO_2			0.011**			0.003			0.013
			(0.005)			(0.005)			(0.009)
Mean O_3			-4.025			-11.116			-15.192
			(7.883)			(7.300)			(11.072)
Mean PM_{10}			0.003			0.000			-0.001
			(0.002)			(0.002)			(0.004)
Mean SO_2			0.007			0.013			0.023
			(0.008)			(0.010)			(0.029)
Observations	2700	2700	2700	2559	2559	2559	1496	1496	1496

Notes: Table is analogous to Table 4 but uses spatial first difference in cross-sectional averages rather than in each year. It regresses changes in average log yields across neighboring counties on the corresponding difference in average CO_2 residuals. Columns (b) furthermore include the difference in residuals from four weather variables, while columns (c) include the difference in residuals from the five criteria air pollutants. The weather variables are moderate degree days ×1000 (degree days 10-29°C for corn and degree days 10-30°C for soybeans and winter wheat) and extreme degree days ×100 (degree days above 29°C for corn and degree days above 30°C for soybeans and winter wheat) as well as a quadratic in precipitation (in meters). The pollution variables are the average concentration of CO (in ppm), NO_2 (in ppm), O_3 (in 100 ppb), PM_{10} (in 100ppm) and SO_2 (in ppm). All weather and pollution variables are constructed over the growing season April-September. Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

B Replication using CarbonTracker Data on CO_2

Our main paper relies on direct satellite readings from OCO-2 Level 2 product, which are available for the years 2015-2021. In this section, we replicate our analyses using modelled data from NOAA's CarbonTracker, which provides spatially-resolved estimates of CO_2 from 2000 to 2018 derived from measurements of air samples collected at 460 sites around the world by 55 laboratories. Unconnected to OCO-2, CarbonTracker involves an inverse model of atmospheric CO_2 that adjusts surface-level CO_2 uptake and releases to align with observational constraints. We use use product release CT2019B (Jacobson et al. 2020) and the level 1 estimates which correspond to 25m above the Earth's surface.

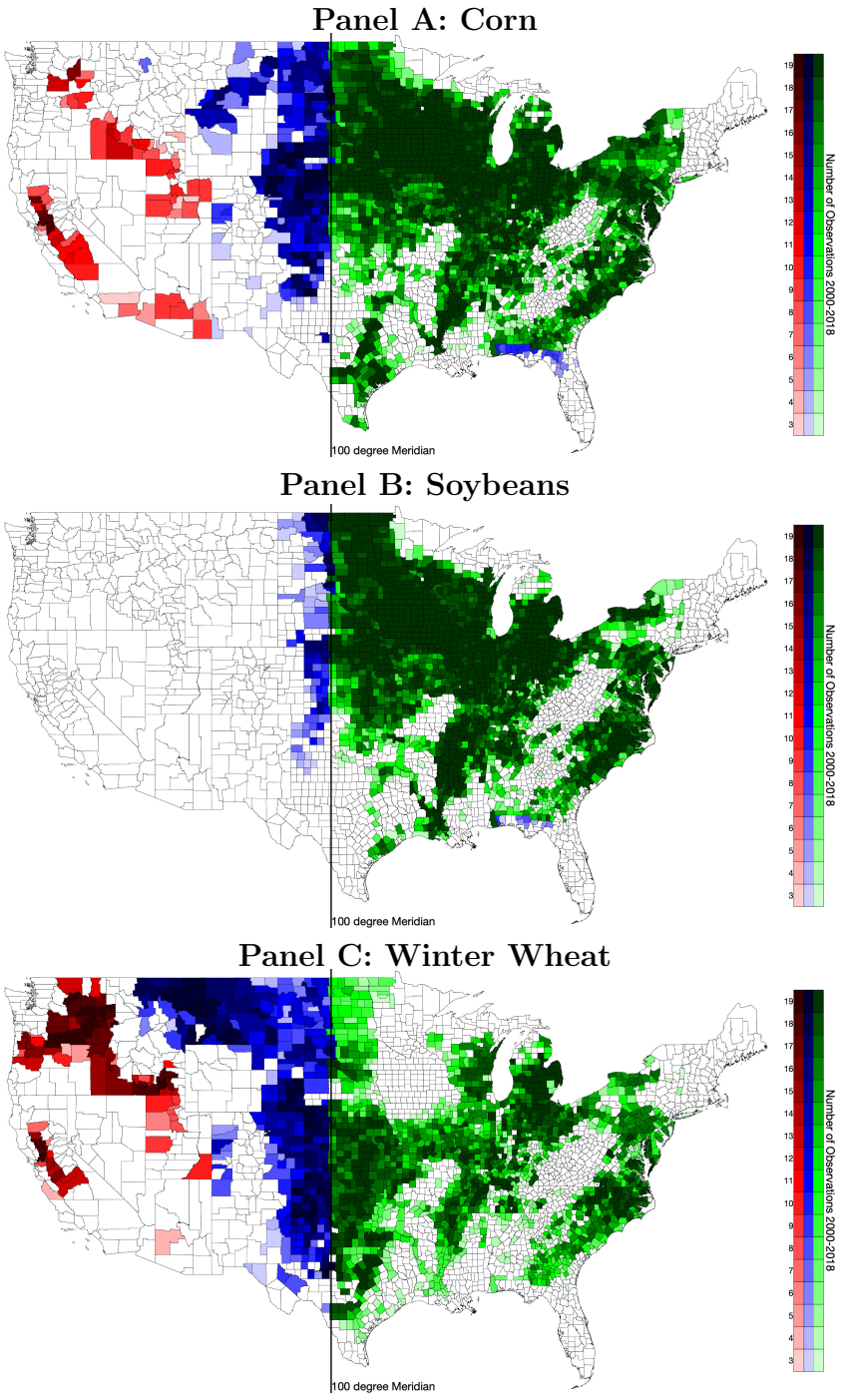
CarbonTracker has an advantage of explicitly modelling surface-level CO_2, while OCO-2 readings are air column-averaged. Figure B2 provides a comparison of CO_2 anomalies for the counties and years (2015-2018) where the two dataset overlap where we find that they have comparable standard deviations. If the satellite-based product of the entire column had a lower variance, it would inflate our estimate of the CO_2 fertilization effect as ground-level varies by more than what is measured in the column. However, this is not the case.

OCO-2 are raw measurements from a satellite, while CarbonTracker is a reanalysis product that might suffer from promulgation of interpolation errors and whose modelling assumptions may cause endogeneity concerns, especially for our IV regression as the underlying spatial interpolation in CarbonTracker is including neighboring counties, invalidating the concept of using upwind neighbors (which are themselves a smoothed estimates of surrounding stations. For that reason we prefer the raw satellite measurements.

Like with OCO-2, we seasonally adjust CO_2 levels from CarbonTracker to account for annual patterns in which ambient concentrations decrease in the spring and summer when plants are actively photosynthesizing and increase in the fall and winter when plants are respiring on net. To identify CO_2 anomalies relative to this seasonality pattern, we estimate the average seasonality over the contiguous US with a 4^{th}-order Chebychev Polynomial over the year which we normalize to [-1,1] by transforming January 1st to equal -1 and December 31st to equal 1 with leap years having an additional day as well. We restrict the seasonality so the value on January 1 (time -1) equals the value on December 31 (time 1).

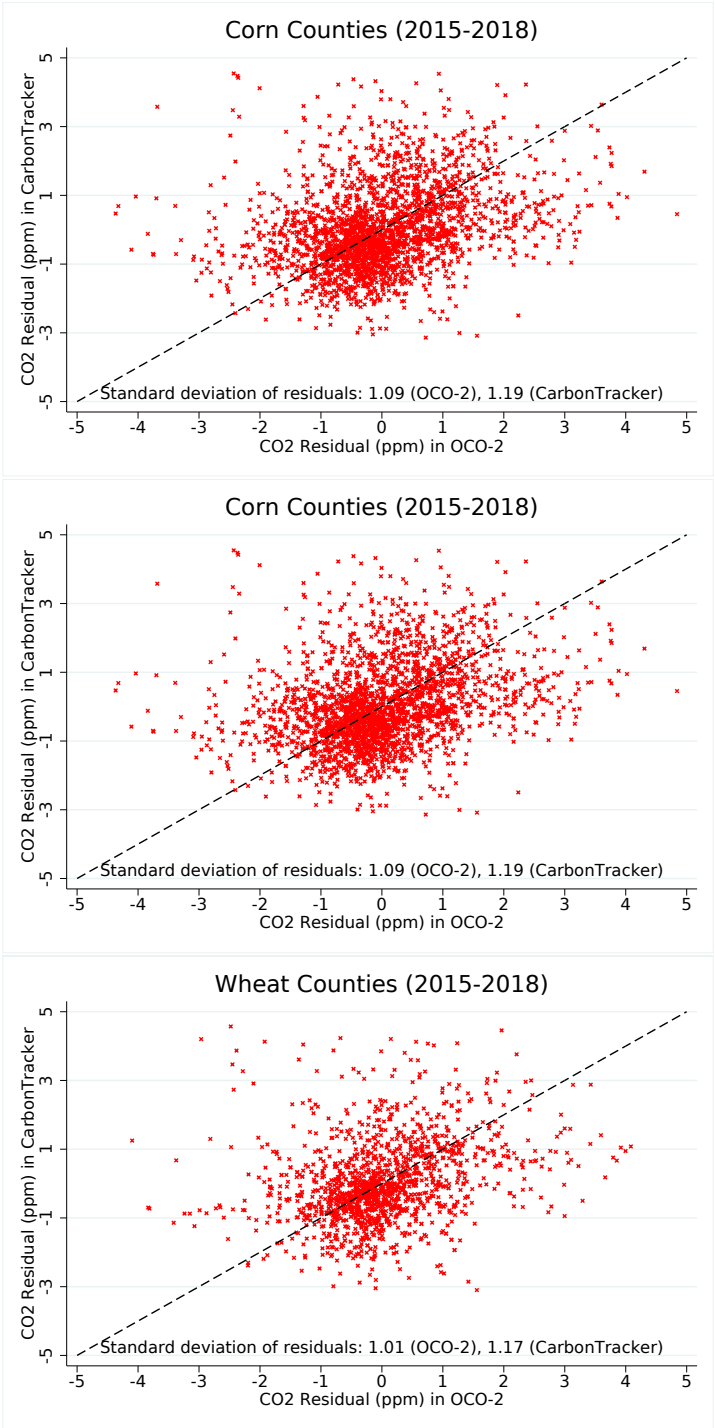
The seasonally-adjusted CO_2 values are averaged during the growing season from April to September. We take the distance-weighted average of the surrounding four CarbonTracker grids for each PRISM grid to derive the PRISM-grid level CO_2 exposure, which is then aggregated to the county level using cropland weights from the Cropland Data Layer, where we aggregate the 30m-resolution from USDA's Cropland Data Layer to the PRISM grid. This gives county-level CO_2 estimates.

Figure B1: CarbonTracker: Number of Observations per County in 2000-2018



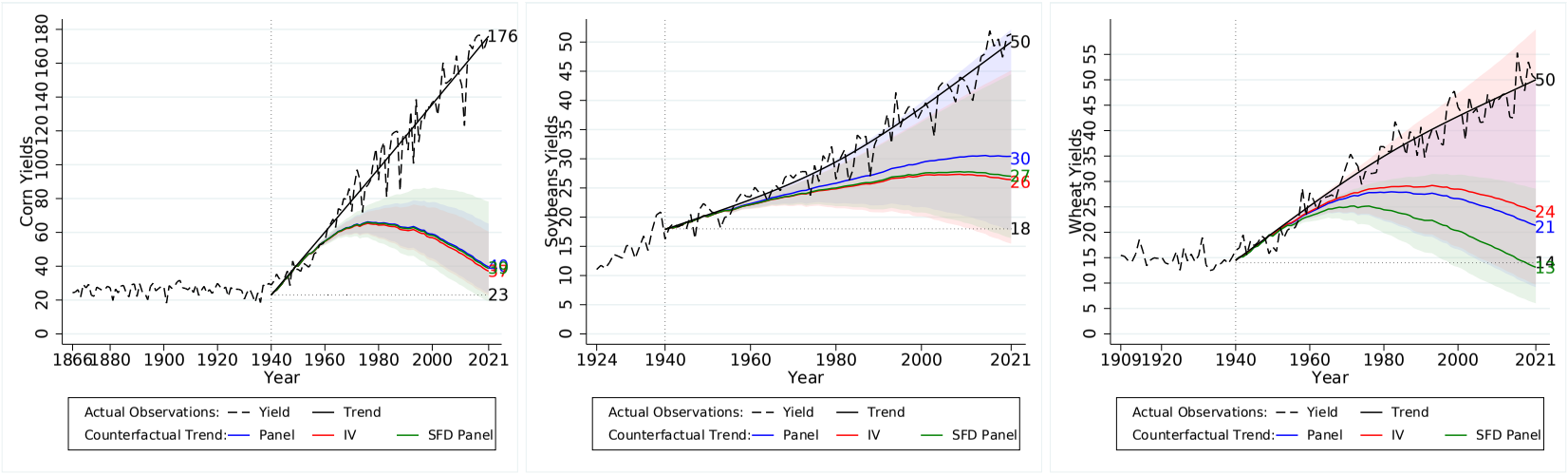
Notes: Figure displays the number of observations per county in the date set, i.e., where yield, weather, criteria air pollution, as well as CO_2 data from CarbonTracker are available over our sample period 2000-2018. We split the analysis into three geographic subsets: east of the 100° meridian excluding Florida (Schlenker and Roberts 2009) shown in shades of green, Western United States (California, Arizona, Utah, Nevada, Oregon, Idaho, and Washington) shown in red, and the remaining in-between counties shown in blue. Since our specification includes county fixed effects and county-specific time trends, we require at least 3 observation to be included in the dataset.

Figure B2: Crossplots of CO_2 Anomalies in CarbonTracker and OCO-2 (2015-2018)



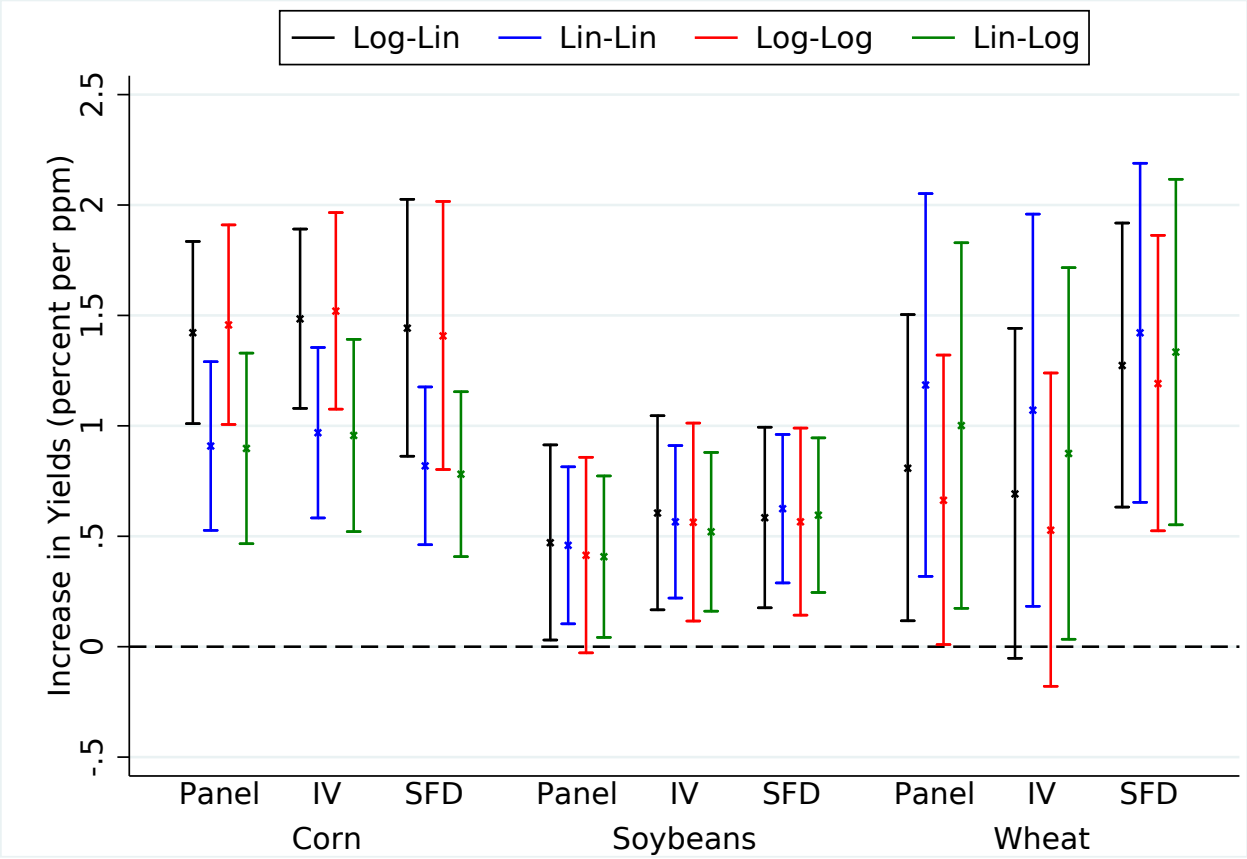
Notes: Figure displays the cross-plot of CO_2 anomalies in CarbonTracker and OCO-2 satellite readings. The two data sources only overlap for four years (2015-2018). We include county-year observations east of the 100° meridian for corn and soybeans and east of the Rocky Mountains for wheat that have at least three observations for both CO_2 measures so we can fit county fixed effects and county-specific time trends. The dashed line is the 45-degree line. The standard deviation of the anomalies is given in each panel.

Figure B3: CarbonTracker: Counterfactual if CO_2 Levels Remained at 1940 Levels



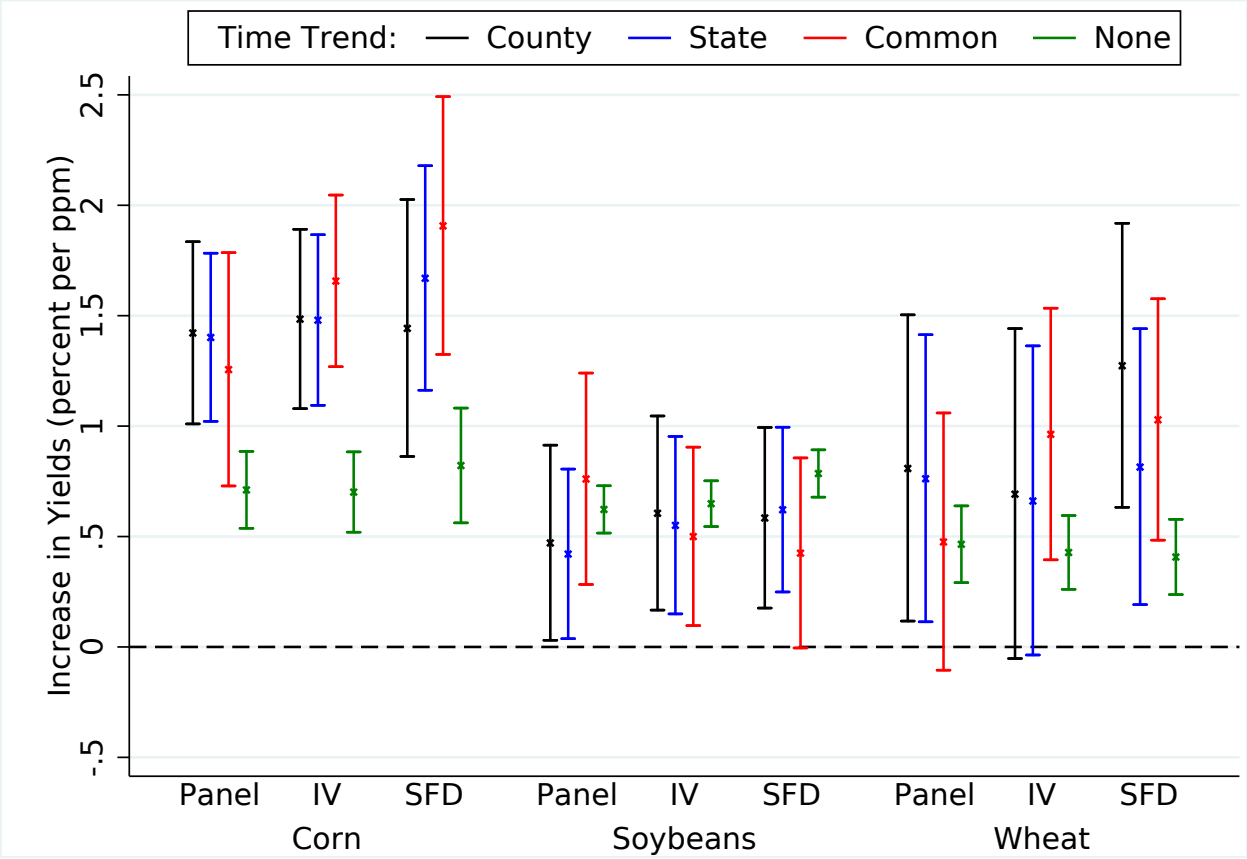
Notes: Figure is the equivalent to Figure 4 except that it uses regression coefficients for CarbonTracker data. Figure displays the evolution of yearly aggregate US yields as dashed black line. The trend using restricted cubic splines with 3 knots (two variables) since 1940 is fitted and shown as solid black line. The graph adds three counterfactuals where CO_2 levels are kept constant at 1940 level, assuming the CO_2 fertilization effect that was estimated using 2015-2021 data applied throughout 1940-2021. The blue line uses the coefficient estimates from columns (c) of Table B1 to simulate the effect on yields, while the red line uses the IV estimates from columns (c) of Table B2, and the green line uses the Spatial First Difference (SFD) estimates. We only adjust the yield trend by the difference in actual CO_2 compared to CO_2 in 1940 while leaving all other factors unchanged. The 95% confidence confidence bands are added.

Figure B4: CarbonTracker: Sensitivity to Functional Form



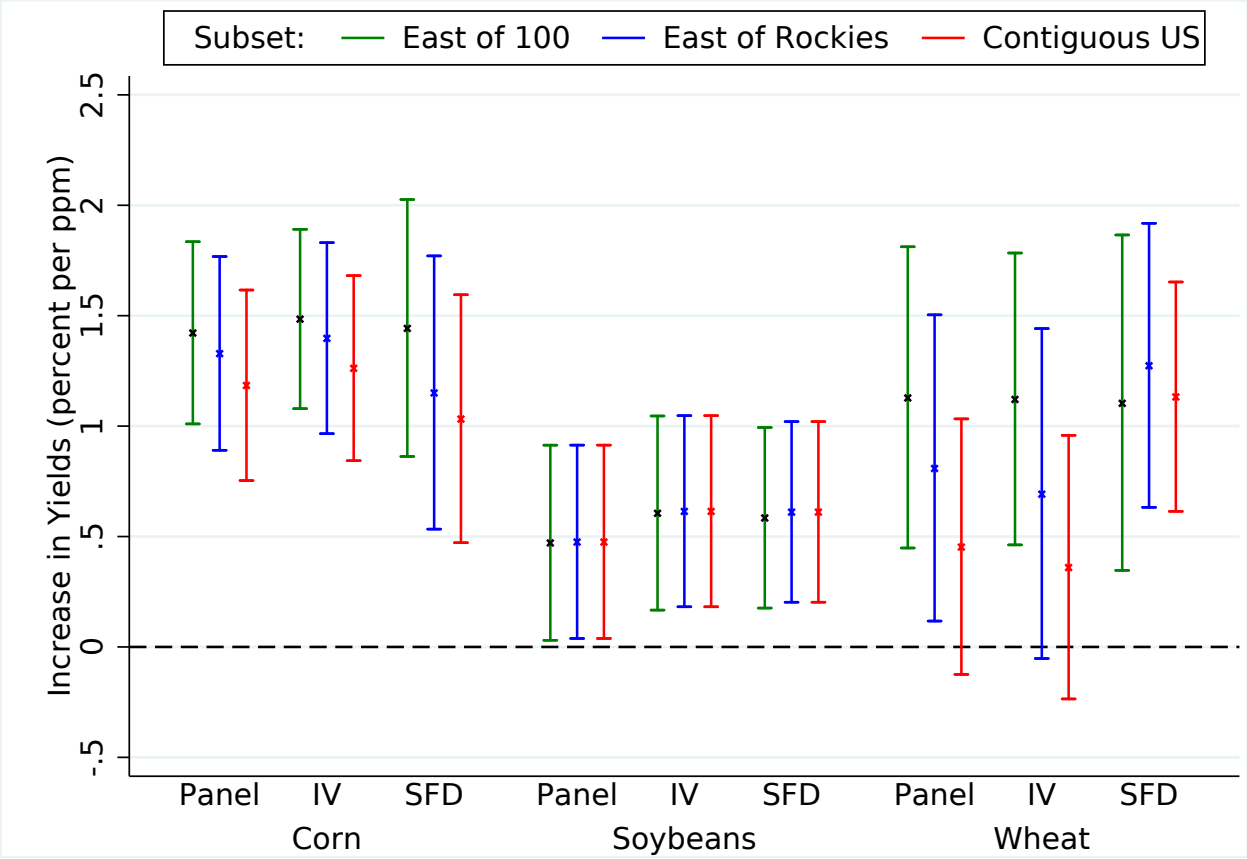
Notes: Figure presents sensitivity check to what functional form is chosen. Graphs shows the effect of a one ppm increase in CO_2 on aggregate yields in percent as well as the 90% confidence bands. Black lines show the baseline results from columns (c) in Table B1 for the panel regression, Table B2 for the wind IV, and Table B3 for the spatial first difference that regress log yields on CO_2 levels (Log-Lin model). Blue lines instead regress yields on CO_2 (Lin-Lin model), red lines regress log yields on log CO_2 (Log-Log model), while green lines regress yields on log CO_2 (Lin-Log model). All regressions include county fixed effects as well as county-specific time trends and control for four weather and five criteria air pollution variables. Errors are clustered at the state level.

Figure B5: CarbonTracker: Sensitivity to Included Time Trend



Notes: Figure presents sensitivity check to what time controls are included. Graphs shows the effect of a one ppm increase in CO_2 on aggregate yields in percent as well as the 90% confidence bands. Black lines show the baseline results from columns (c) in Table B1 for the panel regression, Table B2 for the wind IV, and Table B3 for the spatial first difference that all included county-specific time trends. Blue lines instead include state-specific time trends, red lines include a common time-trend, and green lines include no trend at all. All regressions include county fixed effects and control for four weather and five criteria air pollution variables. Errors are clustered at the state level.

Figure B6: CarbonTracker: Sensitivity to Geographic Subset



Notes: Figure presents sensitivity check to what what geographic subset is included in the analysis. Graphs shows the effect of a one ppm increase in CO_2 on aggregate yields in percent as well as the 90% confidence bands. Green lines show the results when counties east of the 100 meridian are used in the analysis, while blue lines show the results when counties east of the Rocky Mountains are used, and red lines show the results when counties when all counties of the contiguous US are used. The subsets are shown in Figure B1. All regressions include county fixed effects and control for four weather and five criteria air pollution variables. Errors are clustered at the state level.

Table B1: CarbonTracker: Panel Regression

	Corn			Soybeans			Winter Wheat		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
CO_2 (ppm)	1.904***	0.968***	1.412***	1.087***	0.102	0.470*	1.243***	0.944**	0.805*
	(0.363)	(0.215)	(0.241)	(0.300)	(0.178)	(0.260)	(0.341)	(0.362)	(0.407)
Moderate DDay		0.038	0.243**		0.435***	0.521***		-0.304***	-0.336***
		(0.116)	(0.094)		(0.081)	(0.079)		(0.076)	(0.071)
Extreme DDay		-0.425***	-0.335***		-0.567***	-0.527***		0.014	-0.003
		(0.083)	(0.069)		(0.048)	(0.048)		(0.032)	(0.031)
Precipitation		0.775***	0.581***		1.257***	1.155***		1.099***	1.032***
		(0.170)	(0.157)		(0.194)	(0.197)		(0.161)	(0.162)
Prec. squared		-0.496***	-0.427***		-0.788***	-0.736***		-0.957***	-0.904***
		(0.135)	(0.122)		(0.121)	(0.121)		(0.120)	(0.120)
Mean CO			-0.087			-0.038			-0.014
			(0.053)			(0.082)			(0.066)
Mean NO_2			-0.005			-0.008			0.009*
			(0.004)			(0.005)			(0.005)
Mean O_3			-3.143***			-0.969***			0.669**
			(0.506)			(0.338)			(0.259)
Mean PM_{10}			-0.158			-0.209			-0.754***
			(0.227)			(0.162)			(0.223)
Mean SO_2			0.038***			-0.005			0.002
			(0.011)			(0.008)			(0.012)
R^2	0.6036	0.7198	0.7323	0.6321	0.7543	0.7566	0.7454	0.7594	0.7613
Observations	29149	29149	29149	26879	26879	26879	23196	23196	23196

Notes: Tables is the equivalent to Table 2 except that it use CO_2 data from the reanalysis product CarbonTracker for the years 2000-2018 rather than raw satellite readings from OCO-2 in 2015-2021. It regresses log yields on CO_2 values. All regressions include county fixed effects as well as county-specific time trends. Columns (b) furthermore include four weather: moderate degree days $\times 1000$ (degree days 10-29°C for corn and degree days 10-30°C for soybeans and winter wheat) and extreme degree days $\times 100$ (degree days above 29°C for corn and degree days above 30°C for soybeans and winter wheat) as well as a quadratic in precipitation (in meters). Columns (c) additionally include mean pollution levels of five criteria air pollutants: CO (in ppm), NO_2 (in ppm), O_3 (in 100 ppb), PM_{10} (in 100ppm) and SO_2 (in ppm). All weather and pollution variables are constructed over the growing season April-September. Corn and soybean regression use counties east of the 100° meridian (green in Figure B1), while wheat regressions include all counties east of the Rocky Mountains (green and blue in Figure B1). Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

Table B2: CarbonTracker: IV Regression using Upwind CO_2

	Corn			Soybeans			Winter Wheat		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
Min hours	-	0	1000	-	0	1000	-	0	1000
OLS Regression									
CO_2 (ppm)	1.412***	1.412***	1.441***	0.470*	0.470*	0.573**	0.805*	0.805*	0.737*
	(0.241)	(0.241)	(0.247)	(0.260)	(0.260)	(0.257)	(0.407)	(0.407)	(0.407)
IV Regression									
CO_2 (ppm)		1.431***	1.473***		0.498*	0.604**		0.759*	0.689
		(0.230)	(0.237)		(0.262)	(0.258)		(0.434)	(0.439)
F (1st stage)		16.28	16.48		15.60	15.77		15.52	15.39
Observations	29149	29149	26806	26879	26879	24717	23196	23196	21570

Notes: Tables is the equivalent to Table 3 except that it use CO_2 data from the reanalysis product CarbonTracker for the years 2000-2018 rather than raw satellite readings from OCO-2 in 2015-2021. It regresses log yields on CO_2 values. Each coefficient is from a separate regression. Columns (a) replicates Table B1 specification (c). Columns (b) and (c) replicate the same OLS regression for the set of observations that are included in the IV regression in the second row, where CO_2 readings are instrumented with the CO_2 reading (averaged over the entire county, not just agricultural area) from the upwind county. Columns (b) and (c) differ in what is classified as upwind county. Column (b) uses the county that is most frequently upwind during April - September using the hourly wind data for NLDAS (there are 4392 hours in April-September). Columns (c) only uses a county as upwind if it at least 1000 hours upwind. All regressions include county fixed effects as well as county-level time trends and control for four weather and five criertia air pollution variables. Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

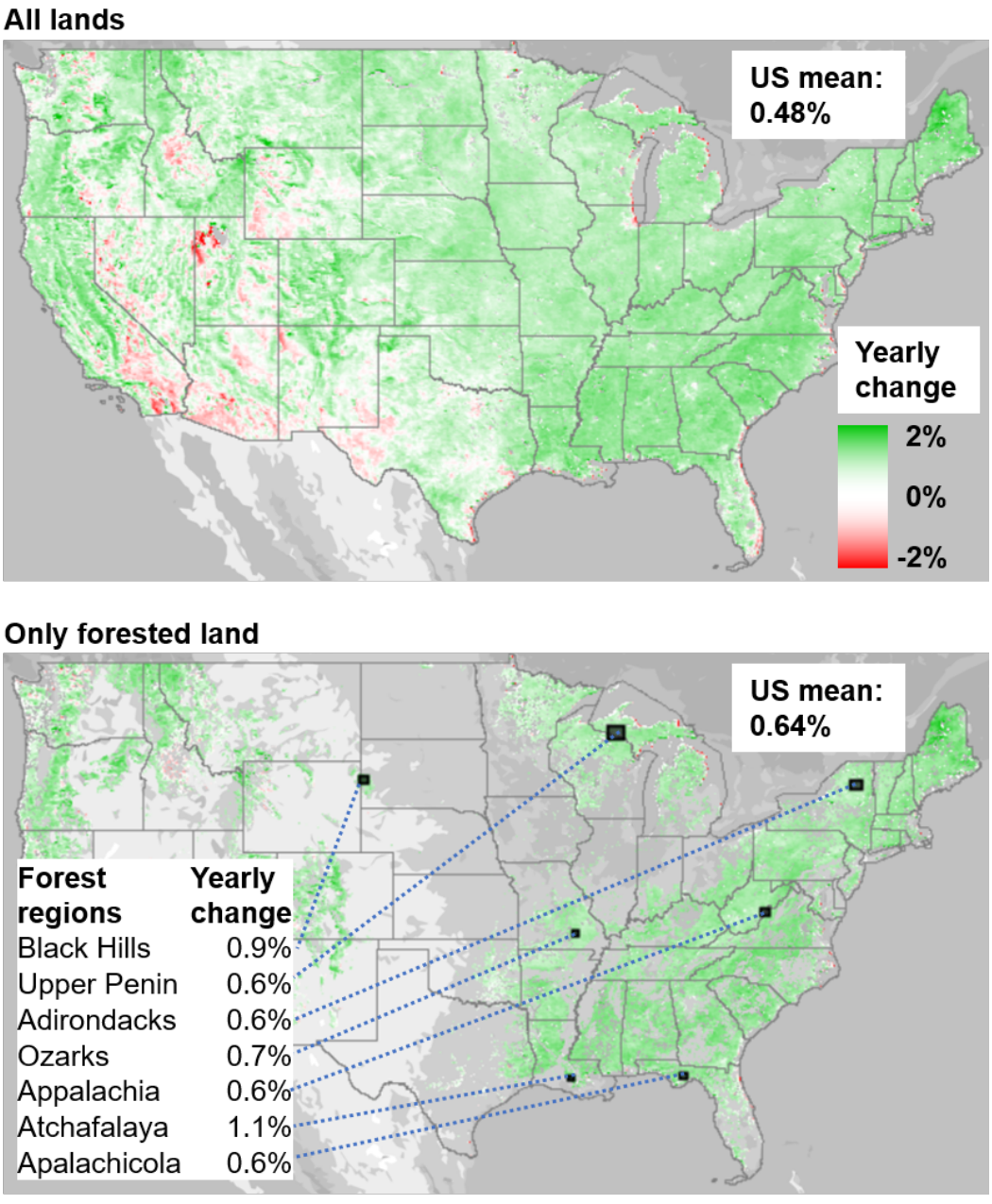
Table B3: CarbonTracker: Spatial First Difference Regression

	Corn		Soybeans		Winter Wheat				
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)			
CO_2 (ppm)	3.064***	1.468***	1.432***	2.125***	0.569**	0.582**	1.306***	1.273***	1.265***
	(0.483)	(0.361)	(0.339)	(0.416)	(0.264)	(0.241)	(0.373)	(0.385)	(0.376)
Moderate DDay		-0.002	0.028		-0.019	0.006		-0.343***	-0.337***
		(0.107)	(0.101)		(0.079)	(0.075)		(0.088)	(0.084)
Extreme DDay		-0.386***	-0.364***		-0.508***	-0.494***		0.001	-0.002
		(0.061)	(0.061)		(0.038)	(0.038)		(0.025)	(0.024)
Precipitation		1.006***	0.988***		1.164***	1.155***		0.460***	0.447***
		(0.096)	(0.096)		(0.133)	(0.132)		(0.161)	(0.157)
Prec. squared		-0.642***	-0.636***		-0.707***	-0.703***		-0.441***	-0.431***
		(0.071)	(0.070)		(0.078)	(0.077)		(0.118)	(0.115)
Mean CO			0.016			0.005			-0.010
			(0.052)			(0.040)			(0.075)
Mean NO_2			-0.000			-0.003			0.002
			(0.002)			(0.003)			(0.003)
Mean O_3			-1.104***			-0.633**			0.142
			(0.358)			(0.234)			(0.206)
Mean PM_{10}			-0.090			0.047			-0.354**
			(0.110)			(0.105)			(0.133)
Mean SO_2			0.001			-0.005			-0.015***
			(0.005)			(0.004)			(0.004)
Observations	74205	74205	74205	69160	69160	69160	52807	52807	52807

Notes: Tables is the equivalent to Table 4 except that it use CO_2 data from the reanalysis product CarbonTracker for the years 2000-2018 rather than raw satellite readings from OCO-2 in 2015-2021. It regresses the change in log yield residuals (after removing county fixed effects as well as county-specific time trends) between neighboring counties on the corresponding difference in CO_2 residuals among neighboring counties. Columns (b) furthermore include the difference in residuals from four weather variables, while columns (c) include the difference in residuals from the five criteria air pollutants. The weather variables are moderate degree days $\times$1000 (degree days 10-29°C for corn and degree days 10-30°C for soybeans and winter wheat) and extreme degree days $\times$100 (degree days above 29°C for corn and degree days above 30°C for soybeans and winter wheat) as well as a quadratic in precipitation (in meters). The pollution variables are the average concentration of CO (in ppm), NO_2 (in ppm), O_3 (in 100 ppb), PM_{10} (in 100ppm) and SO_2 (in ppm). All weather and pollution variables are constructed over the growing season April-September. Errors are clustered at the state level. Stars indicate significance levels: * 10%, ** 5%, *** 1%.

C Robustness Check using NDVI data on Plant Growth

Figure C1: Annual Trends in NDVI Vegetation from the AVHRR Satellite, 1982-2013



Notes: Figure displays 30m pixel-level linear trends of log NDVI values by year for the six months of the growing season from April to September over the 31 years from 1982 to 2013 from the AVHRR satellite (Vermote et al. 2014). Map visualization and calculations produced using Google Earth Engine.