Assessing the Impact of National Air Quality Standards on Agricultural Land Values: Insights from Corn and Soybean Regions

Cécile Couharde
Rémi Generoso

2024-9 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU

Assessing the Impact of National Air Quality Standards on Agricultural Land Values: Insights from Corn and Soybean Regions

Cécile Couharde[1] and Rémi Generoso[2,3]

[1] *EconomiX-CNRS, University Paris Nanterre*
[2] *LEM-CNRS, University of Lille*
[3] *Chaire en macroéconomie et prévisions, Université du Québec Montréal (UQAM)*

Abstract

We examine the impact of the National Ambient Air Quality Standards, as defined by the Clean Air Act, on agricultural land values within the corn and soybean regions of the United States. To achieve this objective, agricultural census data on farmland values are combined with pollution exposure metrics as defined by the Environmental Protection Agency. Using a difference-in-differences approach and conducting various robustness checks, we find that compliance with air quality standards has a statistically significant negative effect on agricultural land values at the county level. Moreover, evidence from quantile regression analysis suggests that counties in the lower quantiles fail to translate the economic and environmental benefits of pollution reduction into increased farmland values, unlike their counterparts with the highest-valued lands.

Keywords: Air Quality Standards; Agricultural Land Values; Difference-in-Differences; Quantile Regression
JEL Classification: C21; Q15; Q53

1. Introduction

In the United States (US), one of the most significant regulations to reduce air pollutant emissions is the Clean Air Act originally passed in 1963 and amended in 1970. The amended Act authorizes the Environmental Protection Agency (EPA) to enforce the National Ambient Air Quality Standards (NAAQS) for total suspended particulates (TSPs). The NAAQS represent a cornerstone of air quality management in the US, aimed at protecting public health and the environment from the adverse effects of air pollution.

Beyond urban and natural ecosystems, this regulatory framework is also expected to benefit agricultural areas. Indeed, conventional wisdom holds that pollution's continuous exposure significantly undermines agricultural outcomes. This is in part because poor air quality tends to reduce crop yields by damaging plants, altering the nutrient balance in soil, and making water sources unsuitable for irrigation (Lobell and Burney, 2021). Indirectly, pollution can also affect agricultural productivity by harming the health of farm workers, leading to decreased labor efficiency (Zivin and Neidell, 2012). On the other hand, there is also some evidence that, for certain farmers, especially those operating on a small scale or with limited resources, the costs associated with complying with pollution regulations can present significant economic hurdles, potentially affecting their livelihoods (Jouzi et al., 2017). This opposite interplay between air quality benefits and regulatory compliance costs raises questions about the overall impact of these regulations on economic returns from agricultural production and ultimately on farmland values. However, despite the significant role of farm real estate in the US as a component of farm sector assets and in the structure of agriculture (Burns et al., 2018), there has been no study that has empirically examined whether the benefits of air pollution reduction are capitalized into farmland values.

This study addresses the gap by examining the effects of the NAAQS, enforced under the Clean Air Act, on agricultural land values within major corn and soybean-producing counties in the United States. Corn and soybeans are among the most valuable and widely cultivated crops in the US, playing a critical role in the agricultural economy. Moreover, counties that are major producers of these crops are likely to be more sensitive to environmental regulations due to the intensive nature of agricultural practices required for high-yield production (Behrer and Lobell, 2022). We focus on significant updates to the NAAQS for particulate matter and ozone in the mid and late 2000s and use the EPAs categorization of counties into “attainment” or “nonattainment” statuses, based on their alignment with the NAAQS for these key air pollutants. We combine these pollution exposure metrics with agricultural land value data at the county level, sourced from the Census of Agriculture conducted by the US Department of Agriculture’s National Agricultural Statistics Service (USDA-NASS) for the years 1997, 2002, 2007, 2012, 2017, and 2022.

To infer a causal relationship between air quality regulations and their economic impacts, we use a difference-in-differences (DID) approach with panel data fixed-effects regressions. A potential concern, when using this approach, is the non-random adoption of pollution standards, which could skew our findings. To account for this potential bias, we estimate DID models using an inverse probability weighted difference-in-differences (IPW-DID) estimator (Abadie, 2005; SantAnna and Zhao, 2020). Additionally, given the variability in achieving NAAQS compliance across the counties over time, we apply a staggered DID analysis, which accounts for varying implementation timelines across different groups (Callaway and SantAnna, 2021). Given the inherent distributional impact of any regulatory policy, we finally examine the effect of the Clean Air Act on the distribution of farmland values across counties, by relying on unconditional quantile regressions (Callaway and Li, 2019).

Our study makes several contributions to the existing literature. First, while a literature examines the effects of the Clean Air Act on property values (Chay and Greenstone, 2005; Grainger, 2012; Bento et al., 2015) and the impact of agricultural policy (Shaik et al., 2005) and farm practices (Chen et al., 2023) on farmland values, to the best of our knowledge, there has been no study that examined the impact of environmental policies and the Clean Air Act more specifically on county-level agricultural land values. Second, the use of recently developed DID models allows us to address bias not accounted for by traditional models. Specifically, the inclusion of accurate covariates and variation in treatment timing gives us the opportunity to provide a relevant counterfactual and in turn more accurately estimate the impact of NAAQS on farmland values. values. Third, by utilizing a staggered framework, we account for the regulatory effects on multiple pollutants. Considering multiple pollutants enables us to broaden the existing literature, which typically focuses on a single pollutant, and to consider the inherent interactions among pollutants and air quality regulations. Finally, this paper explores the distributional effects of environmental policies and the Clean Air Act more specifically, moving beyond the existing body of research that primarily assesses its average effects. We are therefore able to investigate whether compliance with NAAQS has shaped the distribution of farmland values, and identify which quantiles have shown the largest effects.

Our findings indicate that, on average, agricultural farmland values in counties newly subjected to regulation have experienced a decline of approximately 10%, in contrast to counties that meet the standards. Furthermore, the implementation of the NAAQS has led to a widening gap in the value of farmland among counties specializing in corn and soybean production. This trend is largely attributed to a marked decrease in farmland values at the lower end of the distribution range.

The remainder of the paper proceeds as follows. Section 2 delineates the study area. Section 3 details the empirical methodology. In section 4, we discuss our main results. Section 5 concludes the paper.

2. Assessing air quality regulations and county designations

This section discusses the sources and relevant features of the air quality regulations and the characteristics of counties used in our analysis.

2.1. Selection criteria for air quality regulations

In the US, monitoring of air pollution has historically relied on the EPA network of sampling sites. The EPA collects and distributes data from state, local, and tribal agencies through a combination of regulatory and non-regulatory measurements at over 1000 stations, providing near-real-time hourly pollution observations. These monitoring locations are used to assess local and regional "attainment" of the NAAQS, to analyze air pollution impacts on public health, and to validate satellite measurements and air quality models

Stationary source regulations focus primarily on areas that are out of compliance with the NAAQS. Each year, the EPA determines the set of counties that are in violation or "nonattainment" of a particular NAAQS standard based on air pollution monitor measurements.

In 1997, the EPA tightened the NAAQS pertaining to ozone and particles further, regulating fine particles less than 2.5 micrometers in diameter ($PM_{2.5}$) for the first time. The new standards were extremely controversial and were challenged in the courts for years, but ultimately the EPA prevailed, and the new standards were implemented in April 2005. The EPA revised the $PM_{2.5}$ (24 hour) standard again in 2006, and the revision went into effect in 2009. However, since all counties that were in "nonattainment" for the annual $PM_{2.5}$ standard in 2009 also failed to meet the 24-hour standard, the 2009 designations did not result in any new areas becoming subject to NAAQS "nonattainment" regulations.

The effects of the 2006 revision on reducing $PM_{2.5}$ concentrations were significant. Only 20 counties failed to comply with the 2012 annual standard revision, a substantial decrease from the 208 counties designated as "nonattainment" under the 1997 $PM_{2.5}$ NAAQS by late 2005, which were already categorized as such in 2006.

Regarding the other pollutant, the NAAQS introduced a standard of 0.08 ppm 1h ozone in 1997 that was revoked in 2008 to 0.075 ppm 8 h ozone. As for $PM_{2.5}$, law enforcement for the 1997 revision started lately, in end of 2004. By 2008, counties in the eastern part of the US that did not meet the ozone standard set by the 2008 NAAQS revision were already as moderate to serious "nonattainment" areas in 2004.

2.2. County designations and anticipation of regulatory changes

This particular timeline highlights a crucial aspect of the regulatory process for NAAQS designations. The timeline from the issuance of a new or revised NAAQS to the official announcement of "nonattainment" designations can indeed span several years. This delay is due to the need for states to monitor air quality, compile data, and for the EPA to review this information before making formal "nonattainment" designations. This situation highlights a common challenge in policy analysis where the timing of regulatory announcements, formal designations, and the onset of enforcement can lead to ambiguity in defining the treatment period. Ambiguity arises particularly with anticipatory effects, where counties or entities might begin compliance efforts prior to formal enforcement, due to expectations of forthcoming regulatory actions. Some counties may have started to implement air quality improvements in anticipation of being designated as "nonattainment" based on the 2006 revision, even before final designations in 2009.

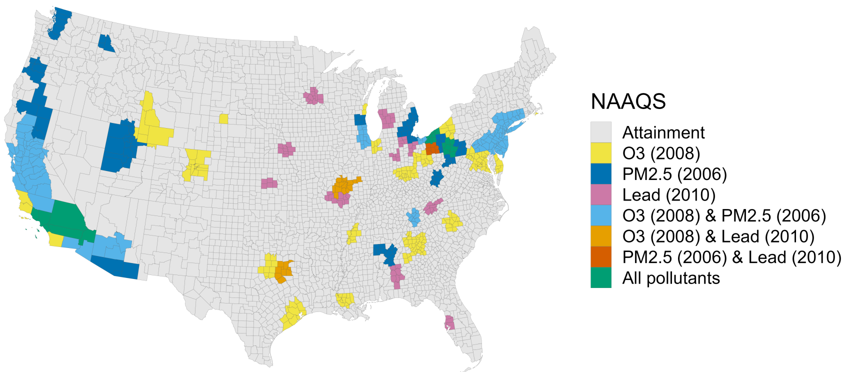
In our study, we consider this anticipatory effect as part of the treatment's impact, as it is a direct consequence of the regulatory change. We classify counties as "nonattainment" based on their non-compliance at the time of the regulatory change, rather than at the time of their final designation. "Nonattainment" counties under NAAQS for each criterion air pollutant are retrieved by year from the NAYRO data set.[1]

Another important aspect of the regulatory process is the fact that the "nonattainment" designations typically apply to air regions or groups of counties in the same local market, rather than to a single county (Cropper et al., 2023). Indeed, whenever a county exceeds the air quality standard based on a local monitoring station, the regulator decides whether nearby or adjacent counties could also have contributed to this violation. Categorizing "nonattainment" status at the county level for a single county might therefore not mimic the regulator's decision and increase the risk of misclassification calling into question any causal interpretation of the results. We then approximate these "nonattainment" designations using county-aggregates in the form of commuting zones (Currie et al., 2023).

Figure 1 provides the geographical localization of counties designated as "nonattainment" for NAAQS standards, grouped by commuting zones for each pollutant. The timeline for NAAQS by pollutant types are detailed in Tables A.1 and A.2 in Appendix A.

[1] https://www.epa.gov/green-book/green-bookdata-download

Fig. 1 *Counties designated "nonattainment" for Clean Air Act's National Ambient Air Quality Standards (NAAQS)*



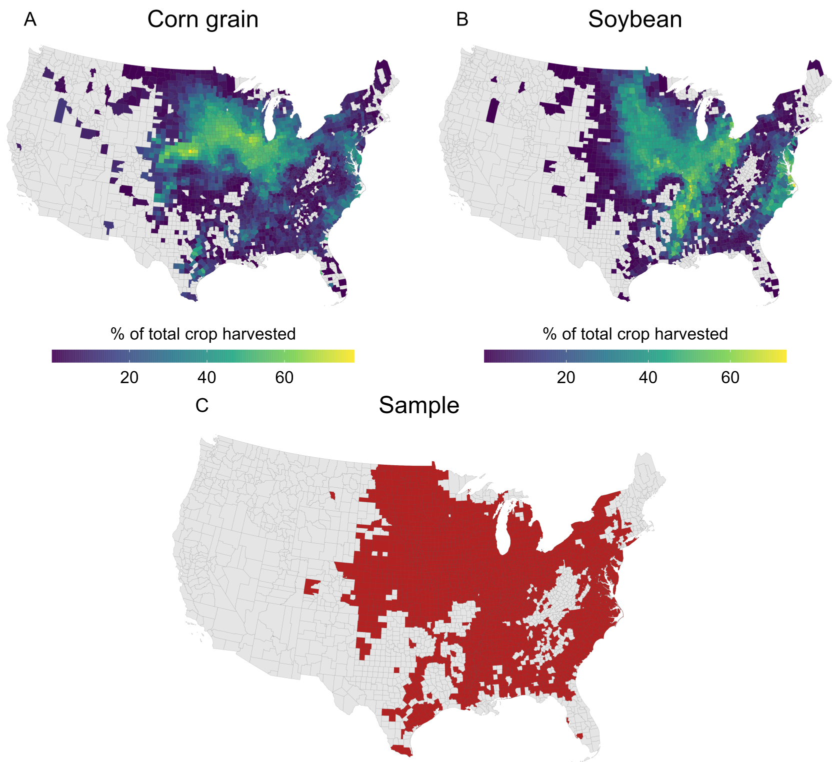
Source: NAYRO data set, EPA Green Book.

2.3. Study area

In our analysis, we focus on counties that are major producers of corn for grain and soybeans. To identify these counties, we utilize data from the USDA National Agricultural Statistics Service (NASS) Census of Agriculture, specifically examining the extent of agricultural reliance on corn and soybean cultivation. We calculate the share of harvested acres dedicated to corn for grain and soybeans as a percentage of total cropland harvested acres for each Census year from 1997 to 2022. Our inclusion criteria involve selecting counties where the proportion of either corn or soybean cultivation exceeds 5% of the total cropland over the 1997-2022 period average. This threshold is chosen to ensure that the sample is sufficiently diverse in terms of geography, given that air pollution emissions are spatially concentrated. Applying this criterion, our sample comprises 1,927 counties. This delineation is illustrated in Figure 2.

Farm real estate values were retrieved for the years 1997 to 2002 from the census data. The NASS Census of Agriculture defines farm real estate value as the value at which all land and building used for agricultural production, including homes, could be sold under current market conditions.

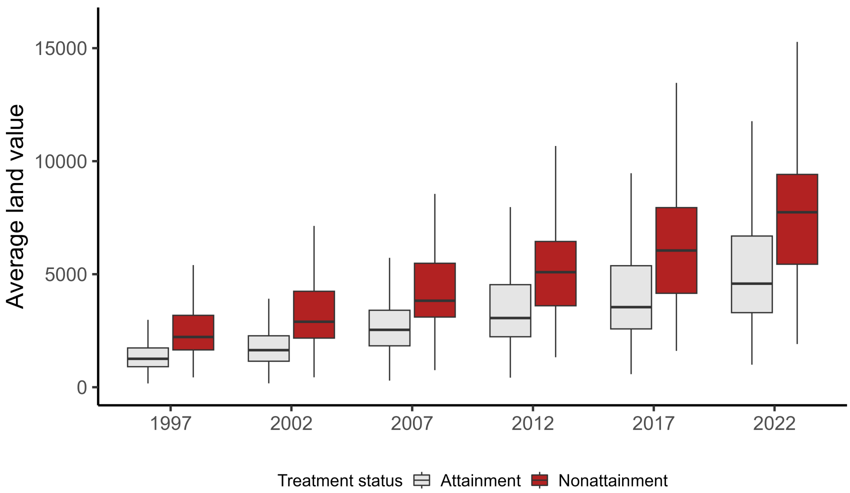
Fig. 2 *Counties included in the sample*



Note: Maps A and B depict the average harvested acres of corn and soybeans, respectively, from 1997 to 2022, expressed as a proportion of total cropland harvested acres in the Contiguous United States (CONUS). The map below depicts our delineation of counties within the CONUS that are major producers of corn and soybeans. It encompasses every county where the proportion of land dedicated to corn and soybean cultivation exceeds a threshold of 5%. This delineation includes a total of 1,927 counties.

Figure 3 illustrates the well-known rise in farmland values across the eastern US, especially in the heartland region, over the last few decades. The median values in the boxplots suggest that "nonattainment" areas have witnessed a relatively steeper growth trajectory compared to their "attainment" counterparts. However, the distribution tails reveal higher end of the land value spectrum with a substantial elongation over time in "attainment" areas. This elongation suggests that while the median land value growth in "attainment" areas may be more modest, the upper segment of the land value distribution has expanded significantly, potentially reflecting a diversification in land value appreciation.

Fig. 3 *Trends in average land value: A comparative analysis of "attainment" vs. "nonattainment" counties (1997-2022)*



Note: the boxplots compare the average land value (expressed in US$ per acre) from 1997 to 2022 between the two groups of counties: those in "attainment" (gray) and those in "nonattainment" (red). They show the median land value (the horizontal line within each box), the interquartile range (the height of each box), and the range excluding outliers (the lines or "whiskers" extending from the boxes).

3. Empirical methodology

To what extend compliance with NAAQS affect farmland values? To address this question, we draw on recent work in Difference-in-Differences (DID) analysis to accurately account for the implementation mechanisms of air pollution regulation and to provide consistent estimates of the Average Treatment effect on the Treated (ATT).

3.1. Identification strategy

Our framework involves two groups of counties ("nonattainment" and "attainment") observed over multiple time periods ($T > 2$) where all treated counties start receiving treatment at time s,

$$Y_{it} = \alpha_i + \phi_t + \sum_{g=g_0}^{G} \sum_{s=g}^{T} \theta(g,s) \times \mathbb{1}(G = g, t = s) + \varepsilon_{it} \tag{1}$$

where $\sum_{g=g_0}^{G} \sum_{s=g}^{T} \theta(g,s) \times \mathbb{1}(G = g, t = s)$ represents the sum over groups g from a starting group g_0 to the last group G, and the sum over time s from the treatment start year for group g to the end of the study period T. $\theta(g,s)$ captures the treatment effect for group g in year s, and $\mathbb{1}(G = g, t = s)$ is an indicator function that equals 1 if unit i belongs to group g and is observed at time s, and 0 otherwise. Equation (1) also controls for county fixed effects, α_i and year fixed effects ϕ_t. The coefficients of interest, $\theta(g,s)$, compare the counties that become newly regulated under the NAAQS to those that are in compliance with these standards according to the specific NAAQS revision.

Given the nature of our economic data (quinquennial data from the USDA Census of agriculture) and the timeline of the regulatory process for NAAQS designations, we first consider all counties in the treatment group receive treatment at time $s = 2007$ and remain treated thereafter as all NAAQS revisions considered in this study occurred before 2012. Counties in the control group, however, do not receive any treatment throughout the entire sample period. Consequently, before the treatment period ($t < 2007$), no counties are treated. This scenario assumes strong anticipation from the counties in the treatment group by setting the treatment period. In this first scenario, the model expressed by Equation (1) falls back to the standard event study design with simultaneous treatment.

We further extend the model in Equation (1)considering the staggered implementation of the NAAQS across the counties in our sample. This extension allows us to account for variations in timing for the designation of "attainment" areas across counties. We consider two treatment groups ($g = 2$).

One comprising counties that were in “nonattainment” to the 2006 NAAQS revision for $PM_{2.5}$ in 2007 and another consisting of counties that failed to comply with the ozone (2008) and lead (2010) NAAQS designation in 2012. This staggered framework is mainly motivated by the timelines of the NAAQS revisions for ozone in 2008 and for lead (Pb) in 2010 that were implemented in 2012.

Our primary challenge in identification arises from the foundational assumption of DID, the parallel-trend assumption (PTA) which is crucial for estimating an unbiased Average Treatment for the Treated (ATT). This assumption posits that, in the absence of treatment, the average outcome for the treated group would have followed a similar trend to the average outcome for the untreated group. It implies that the effects of any confounding factors on the average outcome would have remained constant over time. However, the presence of time-varying confounding factors can cause the violation of the parallel trends assumption. The effect of complying with NAAQS identified by the standard DID could be biased in such a case.

One way to increase the credibility of the parallel trends assumption is to require that it holds only conditional on covariates (SantAnna and Zhao, 2020). We consider the generalization of the ATT proposed by Callaway and SantAnna (2021) that is suitable to setups with multiple treatment groups and multiple time periods:

$$\mathrm{ATT}(g;t) = \mathbb{E}\left[Y_t(g) - Y_t(0) \mid G_g = 1\right] \tag{2}$$

and must satisfy the Limited Treatment Anticipation (LTA) assumption, which states that the treatment effect for a given group in a given year is not affected by the anticipation of future treatments and the Conditional Parallel Trends (CPTA) based on the ”Never-Treated” counties ($C = 1$),

$$\mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid X; G_g = 1\right] = \mathbb{E}\left[Y_t(0) - Y_{t-1}(0) \mid X; C = 1\right] \tag{3}$$

Conditioning on a set of covariates X seeks to account for and mitigate the influence of observable factors that could potentially skew the estimates. The underlying assumption is that, once the observable covariates are accounted for, any remaining unobservable factors affecting the outcomes would have a consistent effect over time. Therefore, these time-invariant unobservable factors would not introduce bias in our estimates of the treatment effect.

We employ semi-/non-parametric methods that enable consistent estimation of the ATT under the assumption of conditional parallel trends, with less restrictive homogeneity assumptions. The basic idea is to estimate what the entire distribution of counties in "nonattainment" would look like if these counties had the same observable characteristics as counties in "attainment" Abadie (2005). In practice, this amounts to implementing a reweighed strategy to estimate the expected change in the outcome for treated units and subsequently estimate the ATT. This is achieved by adjusting the distribution of characteristics among control units to better resemble those of treated units. We then calculate a weighted average of the change in outcome (ΔY_{it}) among control units to approximate the expected change in outcome for treated units (Abadie, 2005).

This weighting function boils down to estimating a conditional probability of being a county in "attainment" based on observable characteristics (covariates) via a probit regression. This conditional probability is then used, along with the unconditional probabilities of being in "attainment"/"nonattaintment" in our sample to reweight the treated group in a given year.

We use the Inverse Probability Weighting (IPW) estimator for estimating the ATT, following the methodology introduced by Abadie (2005). Nonetheless, the estimator's performance can be compromised by instability when propensity scores are near the limits of the unit interval. This is particularly problematic when scores approach one, leading to exaggerated weights that can disproportionately affect the estimator, amplifying variance and potentially diminishing the reliability of causal inferences. To mitigate such issues, we use stabilized weights, where the typical IPW weights are adjusted to dampen the influence of extreme propensity score values (Imbens, 2000). We also use the Doubly Robust (DR) estimator proposed by SantAnna and Zhao (2020) which combines the IPW with outcome regression. This DR estimator provides a safeguard against misspecification. It remains consistent if either the propensity score model is correctly specified, offering protection against extreme weights, or if the outcome model is accurate, but not necessarily both. Hence, this approach not only addresses the instability associated with IPW but also enhances the reliability of the estimated ATT by allowing for two chances at correct model specification.

3.2. Covariates

We include two types of covariates: (i) covariates correlated with the treatment assignment to address the issue of non-random treatment allocation, and (ii) covariates correlated with the outcome to ensure that the model accounts for potential confounding variables that could affect the outcome independently of the treatment. This point is essential because omitting such variables could lead to biased estimates of the treatment effect, as changes in the outcome might be attributed incorrectly to the treatment rather than these confounders.

We first use covariates reflecting weather and atmospheric conditions as they can simultaneously influence both the outcome of interest (farmland value) and the treatment variable (pollution levels). Specifically, we use the following weather variables: number of growing degree days (8–29°C) and the Palmer Drought Severity Index (PDSI) (Palmer, 1965) to account for drought conditions. These weather variables are key variables because they can be both a driver of expected economic returns from agricultural production and a factor that exacerbates or mitigates pollution levels by affecting atmospheric stability and dispersion. Additionally, we include measurements of thermal inversions. Thermal inversions, which trap pollutants close to the ground, directly influence $PM_{2.5}$ concentrations and can also impact farmland values by affecting local environmental conditions and agricultural productivity. The methodologies used to calculate these various weather and atmospheric covariates are detailed in Appendix B.

We include additional covariates that reflect the various types of returns farmland can generate, which may impact both the outcome and the treatment. In certain areas, the value of farmland may be primarily based on its ability to support agricultural production. Factors specific to each parcel, such as soil quality, proximity to market terminals, and government subsidies, play a significant role in determining land values in these cases. However, farmland located near urban areas may also yield returns from non-agricultural sources, such as the potential for residential, commercial, or industrial development. Factors influencing development potential include proximity to population centers and accessibility to transportation networks (Kuethe et al., 2011). Both agricultural and non-agricultural factors contribute to variations in farmland prices and pollutant exposure. To account for these spatially specific factors, we include the poverty level among county populations, the concentration of farms, off-farm revenue, and agricultural practices such as fertilizer and chemical usage, as well as irrigation. A comprehensive description of all selected covariates is provided in Table A.3 in Appendix A.

3.3. Unconditional quantile regressions

We finally extend our analysis by examining how compliance with the NAAQS affects the observed disparities in farmland values. To explore this distributional effect, we utilize recent advancements in quantile regression, focusing on unconditional quantile regression (UQR) models. Specifically, we employ recentered influence functions (RIF) developed by Firpo et al. (2009).

While conditional quantile regression methods offer insights into treatment effects for individuals at various points in the outcome distribution, UQR models are more aligned with our research question as they analyze the consequences of the treatment on the entire outcome distribution. Specifically, in the context of our study, an UQR model allows us to examine how shifts in the proportion of counties in "attainment" affect the quantile values of the overall distribution of farmland values across counties. Using such estimates, we can infer whether changes in farmland values resulting from compliance with NAAQS contribute to disparities in farmland values among counties and identify the quantiles that are most affected.

We rely on Firpo et al. (2009) that have introduced the UQR model as a solution to estimate influences on unconditional distributions. They propose a two-step approach consisting in (1) re-centering the influence function (RIF) and (2) regressing the RIF on the independent variables in an OLS model. The RIF is defined as:

$$RIF\left(y;Q_{\tau}\right)=Q_{\tau}+\frac{\tau-\mathbb{1}\left(y\leq Q_{\tau}\right)}{f_{Y}Q_{\tau}} \tag{4}$$

were $\mathbb{1}$ describes the indicator function, $f_Y Q_\tau$ refers to the density of the outcome at quantile τ, and Q_τ is the sample quantile value of τ. The RIF has only two values, one for units with an outcome value below or equal to the quantile value Q_τ and one for units with an outcome value larger than Q_τ.

From a technical point of view, the UQR approach compares two counterfactual distributions: one where all units are treated and one where none are treated. Thus, the UQR coefficient refers to changes from 0% of counties being classified in "nonattainment" to 100% being classified in that category. However, the main weakness of RIF regressions is that coefficients provide only local approximations of the effect of changes in the distribution of the independent variables and should not be used to make inferences about the coefficients of categorical variables as changes from 0 to 1 (Essama-Nssah and Lambert, 2012).

Recent studies, including Rothe (2010), Donald and Hsu (2014), Firpo and Pinto (2016), and Firpo et al. (2018), have proposed methodologies relying on reweighting schemes to address this weakness. These methodologies rely on parametric or nonparametric strategies to obtain inverse probability weights that can be used to identify counterfactual distributions and identify the treatment effects on distributional statistics. As in the DID framework, we use the IPW estimator on the same covariates to estimate the UQR model. We also employ the Doubly Robust estimator to check the robustness of our findings.

4. Results

4.1. Average effects

We begin with the analysis of the average effects of compliance with the NAAQS for the treated counties, relative to the control counties. The results are presented in Table 1.

The treatment coefficient is negative and statistically significant at the 1 percent level. This finding indicates that air quality regulations adversely affect farmland values in newly regulated counties, relative to those already in "attainment". The economic impact of this effect is considerable. With the dependent variable measured as a logarithm, farmland values, expressed in dollars per acre, decrease by 10% in counties meeting NAAQS requirements compared to counties already in "attainment". Importantly, this effect remains consistent across the two estimators used in the analysis and is robust to the implementation of the treatment. Indeed, our results suggest that the staggered treatment has a statistically significant negative effect on farmland values in the post-treatment period, particularly for the Group (2007).

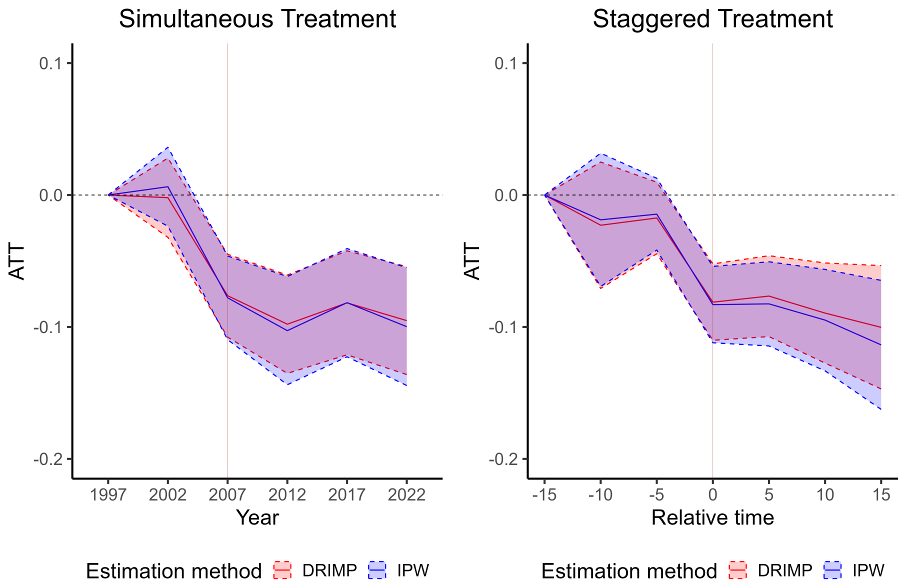
Tab. 1 *DID results*

	Inverse Probability Weight				Doubly Robust Estimator			
	Coeff.	se	95% CI lower	95% CI upper	Coeff.	se	95% CI lower	95% CI upper
Simultaneous treatment (2007)								
ATT (pre)	0.0054	0.0156	-0.0253	0.0359	0.0024	0.0154	-0.0278	0.0326
ATT (post)	-0.0915***	0.0169	-0.1246	-0.0584	-0.0953***	0.0151	-0.1248	-0.0657
pre-trend	0.1175				0.0243			
p-value	0.7318				0.8761			
N	9019				9019			
Staggered treatment								
ATT (pre)	-0.0092	0.0121	-0.0329	0.0145	-0.0102	0.0121	-0.0338	0.0135
ATT (post)	-0.0957***	0.0153	-0.1257	-0.0657	-0.1021***	0.0149	-0.1314	-0.0729
Group mean	-0.0893***	0.0147	-0.1182	-0.0605	-0.0942***	0.0139	-0.1216	-0.0669
Group (2007)	-0.1148***	0.018	-0.1502	-0.0795	-0.1263***	0.0183	-0.1621	-0.0905
Group (2012)	-0.0452	0.0246	-0.0935	0.0031	-0.0387	0.0209	-0.0796	0.0022
pre-trend	1.4520				2.0495			
p-value	0.6934				0.5622			
N	9390				9390			

Note: The dependent variable is the logarithm of farmland values. All regressions include county fixed effects and year fixed effects and control for parallel trends for the treated and control groups in the pretreatment period. The t-statistics are based on bootstrapped standard errors clustered at the county level with 10,000 repetitions. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Figure 4 presents the estimated event-study coefficients from Equation (1).

Fig. 4 *Dynamic treatment effects*



Note: The left plot displays the event-time coefficient estimates derived from Equation (1), where the dependent variable is farmland values. The regression analysis adjusts for fixed effects at both the county and year level. The right plot is based on relative time rather than calendar years, centered around the implementation of the treatment (time 0). The dashed lines in blue and red represent the estimates obtained using the Inverse Probability Weighting (IPW) estimator and the Doubly Robust estimator, respectively. The shaded areas around these lines indicate the bootstrapped (10,000 reps) 95% confidence intervals associated with each set of point estimates.

There are two main findings. First, in the years leading up to the implementation of the regulation, the trends in farmland values between the newly regulated counties and the unregulated counties are not statistically different from zero. Five years following the policy's implementation, farmland values per acre in counties subject to new regulations experience a decline of approximately 10%, compared to those in unregulated counties. This reduction persists until the end of our observation period. The patterns observed in the staggered treatment appear even more pronounced, compared to the simultaneous treatment. Consistent with the estimate results in Table 1, the findings suggest that the compliance with NAAQS, whether applied simultaneously or in a staggered fashion, has exerted a negative effect on newly regulated counties over the considered period.

4.2. Balance and positivity diagnostics

The validity of our previous results relies on the assumption that there are no differential trends for the treated and control groups in the pretreatment period. To address any initial imbalances between the two groups, we have employed reweighting procedures, as previously stated. The primary objective of these reweighting methods is to enhance balance between the treatment and control groups in observational studies, where randomization is not possible, by aligning the covariate distributions of the two groups more closely, based on pretreatment periods. When executed correctly, this process significantly reduces model dependence, diminishes the potential for bias, decreases variance, and consequently, lowers the mean squared error. Thus, validating these procedures is crucial to ensuring they effectively yield unbiased and accurate estimates of treatment effects (Wei et al., 2023).

In practice, two important diagnostic evaluations are conducted when using Inverse Probability Weighting (IPW). The first is evaluating the covariate balance (before and after weighting) and ensuring that weighting leads to comparable treatment groups, with respect to the measured covariates. The second involves assessing the positivity assumption by looking at the overlap of the propensity score distributions between the treatment groups and their common support. While the common support assumption is not required for estimators based on IPW, parametric models are only approximations of the true model, and one may suspect that their out-of-support predictions might be particularly unreliable. A lack of sufficient overlap may be indicative of a violation of the positivity assumption, which can result in extremely large IPW weights.

Table 2 displays the unadjusted and adjusted standardized mean differences (SMD) and variance ratios (VR) for the selected covariates. These metrics are used to evaluate whether the covariates are balanced between the treatment and control groups both before and after reweighed adjustments. We observe that adjusting for the selected covariates leads to a significant reduction in the standardized mean difference and variance ratio between the treatment and control groups. This adjustment has, therefore, improved the precision of our treatment effect estimations.

Tab. 2 *Balance table*

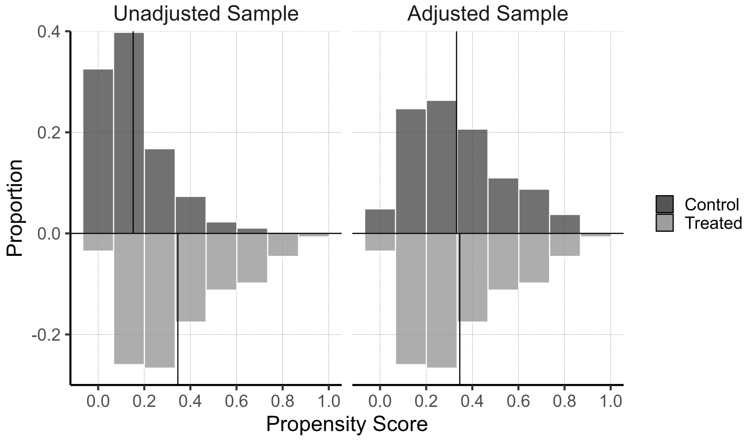
	Unadjusted		Adjusted	
	SMD	VR	SMD	VR
Drought	0.223	1.188	0.019	1.242
Inversion	0.409	3.296	0.068	1.863
GDD	0.276	1.371	0.089	1.106
Poverty	0.987	1.858	0.081	1.307
Offfarm	0.064	1.135	0.012	1.272
Chemical	0.433	1.263	0.012	1.092
Fertilizer	0.547	1.285	0.008	1.042
Irrigation	0.549	1.917	0.044	1.160
Concentration	0.610	1.480	0.021	1.245

Note: Unadjusted (Ajusted) Standardized Mean Difference (SMD) measures the difference in means between the treatment and control groups for each covariate, standardized by the pooled standard deviation. It provides an indication of the covariate imbalance (balance) between the groups before any adjustments (after adjusting for potential confounders). Unadjusted (Ajusted) Variance Ratio (VR) compares the variances of the covariates between the treatment and control groups. It helps assess whether the variances are similar between the groups before any adjustments (after adjusting for potential confounders).

We conduct an additional, in-depth diagnostic assessment to evaluate the quality of the matching between the treatment and control groups. Precisely, we compare the propensity scores calculated for each county based on pre-treatment conditions, before and after applying weighting, to assess the effectiveness of the IPW method in reducing imbalances between groups. Figure 5 illustrates the propensity score distribution before and after the adjustment method.

In the unadjusted sample, the distribution of propensity scores for the treated and control groups are not well-aligned. The control group (in darker shade) shows a different distribution pattern compared to the treated group (in lighter shade). There is some overlap in the middle range of propensity scores, but the peaks of the distributions are offset from each other. Figure 5 suggests that, before adjustment, there is a selection bias in the treatment allocation, with the treated group having overall higher or lower propensity scores. In

Fig. 5 *Propensity score distribution before and after the adjustment method*



Note: The two histograms compare the distribution of propensity scores for the treated group (in lighter shade) and the control group (in darker shade) in both the unadjusted sample and the adjusted sample.

the adjusted sample, the overlap between the treated and control groups is much more substantial. The distributions are more similar, indicating that the weighting scheme has improved the balance between the two groups.

Therefore, Figure 5 indicates that our adjustment method has effectively balanced the propensity score distributions between the treated and control groups. This balance is crucial for making causal inferences about the treatment's effect. Specifically, the absence of significant pre-trends further strengthens our previous results by confirming that the effects of compliance with NAAQS can be attributed to the treatment rather than to pre-existing conditions or trends. This adjustment is also crucial for inferring unbiased estimates of how compliance with the NAAQS affects various points in the distribution of farmland values.

4.3. Distributional effects

While the estimates from Table 1 and Figure 4 tell us about the average effects in the treated counties, relative to the controls, they tell us little about other parts of the farmland value distribution that might otherwise be affected by the air quality regulations. To examine this distributional impact, we now turn to unconditional quantile regression estimates.

Table 3 presents the results from the estimation of the Recentered Influence Function (RIF) over several time intervals.

Tab. 3 *RIF-OLS results*

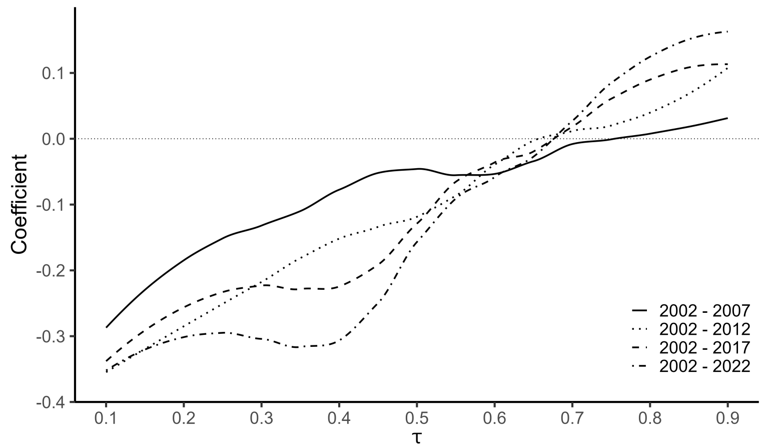
	$\tau = 0.1$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$
ATT (2002 - 2007)	-0.321***	-0.156**	-0.0447	0.0157	0.0869
	(0.0703)	(0.0558)	(0.0463)	(0.0533)	(0.0699)
ATT (2002 - 2012)	-0.341***	-0.264***	-0.116*	-0.0374	0.140
	(0.0753)	(0.0589)	(0.0590)	(0.0616)	(0.0769)
ATT (2002 - 2017)	-0.343***	-0.205**	-0.134	0.0865	0.132
	(0.0792)	(0.0678)	(0.0696)	(0.0731)	(0.0890)
ATT (2002 - 2022)	-0.370***	-0.298***	-0.203**	0.0967	0.180**
	(0.0853)	(0.0803)	(0.0775)	(0.0714)	(0.0684)

Note: The Table presents results from the RIF regression at different quantiles of the farmland value distribution. The ATT is measured over several time intervals: 2002 to 2007, 2002 to 2012, 2002 to 2017, and 2002 to 2022. The quantiles (τ) range from the 10th percentile (0.1) to the 90th percentile (0.90) of the farmland value distribution. *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively.

The estimates indicate that the ATT is consistently negative and statistically significant at the lower quantiles (0.1 and 0.25) across all time periods. This suggests that the implementation of NAAQS has had a negative effect on the value of farmland, particularly at the lower end of the value distribution. This effect diminishes and loses statistical significance as we move towards the median (0.50) and is not significant at higher quantiles (0.75 and 0.90) except for the 90th percentile in the 2002 to 2022 interval where it becomes positive and significant. For the earliest period (2002 - 2007), the impact of NAAQS on farmland values is highly significant at the lower quantile (0.1) but becomes insignificant as we reach the median and upper quantiles. Over time, the treatment effect remains significantly negative at the lower quantiles (e.g., 2002 - 2022), indicating a persistent adverse impact of NAAQS on farmland values at these levels. Comparing ATT across years for each quantile, it appears that the negative impact of NAAQS on lower-value farmland (0.1 and 0.25 quantiles) has either remained stable or slightly increased over time, indicating that the regulatory standards imposed by NAAQS have been more burdensome on lower-valued farmland.

Figure 6 illustrates the evolution over time of the estimated coefficients from the RIF estimation. The results show that the estimated impact of NAAQS on farmland values has generally been negative across all quantiles over the years, with the magnitude of the effect varying by quantile and time period. For the earliest period (2002-2007), the impact is negative across all quantiles and is less negative as the quantile increases, suggesting that higher-valued farmland experienced a lower negative effect from NAAQS during this time. As we move to later periods, the magnitude of the negative effect seems to increase at the lower quantiles (indicating a greater negative impact on lower-valued farmland), while the effect at the higher quantiles appears to be less negative and even positive by 2022. The increasing negative impact at lower quantiles in later years suggests that NAAQS have progressively placed a heavier burden on less valuable farmland or that the standards have become stricter over time. On the contrary, the lesser negative impact and the shift towards a positive effect at higher quantiles in later years indicate that more valuable farmland has been able to adapt or benefit from the regulations implied by NAAQS over time, possibly due to having more resources to comply with or benefit from air quality improvements.

Fig. 6 *RIF-OLS results*



Note: The plot illustrates the relationship between the treatment effect (coefficients) and the quantiles (τ) of the farmland value distribution for different time periods from 2002 to 2022.

5. Conclusion

In this paper, we use a panel of 1,927 corn and soybean producing counties within the United States to examine the impact of air quality regulations on disparities in farmland values from 1997 to 2022.

Our results firstly suggest that the implementation of NAAQS has adversely affected farmland values, particularly evident after the policy was put in place. We find that the impact of NAAQS has been more pronounced in the immediate aftermath (2007) and appears to have persisted but was not be as strong in later years (2012). The negative effects are consistent across a simultaneous treatment and a staggered one, suggesting that the treatment effects are robust to different difference-in-differences frameworks. Moreover, the use of two different but converging estimation methods strengthens the validity of the results.

Secondly, we provide evidence that the benefits of the NAAQS measured through the capitalization of air quality improvements in farmland values were regressive. Indeed, a consistent pattern of significant negative effects at lower quantiles across all time frames indicates that the NAAQS have had a sustained adverse effect on the value of less valuable farmland. This outcome may be attributed to the compliance costs or changes in land use prompted by the regulations, which seem to have disproportionately affected counties with lower farmland values.

Therefore, our findings offer essential insights into the environmental policy aspects of agricultural production. Given that the NAAQS aim to safeguard public health through air quality regulation, a detrimental effect on farmland values suggests that policymakers should strive to balance environmental goals with the economic repercussions for landowners. Specifically, our results highlight the importance of implementing economic support or compensation for landowners negatively impacted by NAAQS, particularly those at the lower spectrum of the farmland value distribution.

References

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The review of economic studies*, 72(1):1–19.

Avila Uribe, A. (2023). The effect of air pollution on us aggregate production. Technical report, London School of Economics and Political Science, LSE Library.

Behrer, A. P. and Lobell, D. (2022). Higher levels of no-till agriculture associated with lower $PM_{2.5}$ in the Corn Belt. *Environmental Research Letters*, 17(9):094012.

Bento, A., Freedman, M., and Lang, C. (2015). Who Benefits from Environmental Regulation? evidence from the Clean Air Act Amendments. *Review of Economics and Statistics*, 97(3):610–622.

Burns, C., Key, N., Tulman, S., Borchers, A., and Weber, J. (2018). Farmland Values, Land Ownership, and Returns to Farmland, 2000-2016. Report ERR-245, U.S. Department of Agriculture, Economic Research Service.

Butler, E. E., Mueller, N. D., and Huybers, P. (2018). Peculiarly pleasant weather for us maize. *Proceedings of the National Academy of Sciences*, 115(47):11935–11940.

Callaway, B. and Li, T. (2019). Quantile treatment effects in difference in differences models with panel data. *Quantitative Economics*, 10(4):1579–1618.

Callaway, B. and SantAnna, P. H. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Chay, K. Y. and Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, 113(2):376–424.

Chen, L., Rejesus, R. M., Aglasan, S., Hagen, S., and Salas, W. (2023). The impact of no-till on agricultural land values in the United States Midwest. *American Journal of Agricultural Economics*, 105(3):760–783.

Chen, S., Oliva, P., and Zhang, P. (2018). Air Pollution and Mental Health: Evidence from China. NBER Working Papers 24686, National Bureau of Economic Research.

Cropper, M., Muller, N., Park, Y., and Perez-Zetune, V. (2023). The impact of the clean air act on particulate matter in the 1970s. *Journal of Environmental Economics and Management*, 121:102867.

Currie, J., Voorheis, J., and Walker, R. (2023). What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality. *American Economic Review*, 113(1):71–97.

Dechezleprtre, A., Rivers, N., and Stadler, B. (2019). The economic cost of air pollution: Evidence from europe. OECD Economics Department Working Papers 1584, OECD Publishing.

Donald, S. G. and Hsu, Y.-C. (2014). Estimation and inference for distribution functions and quantile functions in treatment effect models. *Journal of Econometrics*, 178:383–397.

Essama-Nssah, B. and Lambert, P. (2012). Influence Functions for Policy Impact Analysis. In Bishop, J. and Salas, R., editors, *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*, volume 20 of *Research on Economic Inequality*, pages 135–159. Emerald Group Publishing Limited, Leeds.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.

Firpo, S. and Pinto, C. (2016). Identification and Estimation of Distributional Impacts of Interventions Using Changes in Inequality Measures. *Journal of Applied Econometrics*, 31(3):457–486.

Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2):28.

Grainger, C. A. (2012). The distributional effects of pollution regulations: Do renters fully pay for cleaner air? *Journal of Public Economics*, 96(9-10):840–852.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, 87(3):706–710.

Jouzi, Z., Azadi, H., Taheri, F., Zarafshani, K., Gebrehiwot, K., Passel, S., and Lebailly, P. (2017). Organic Farming and Small-Scale Farmers: Main Opportunities and Challenges. *Ecological Economics*, 132:144–154.

Kuethe, T. H., Ifft, J., and Morehart, M. (2011). The Influence of Urban Areas on Farmland Values. *Choices*, 26(2).

Lobell, D. B. and Burney, J. A. (2021). Cleaner air has contributed one-fifth of US maize and soybean yield gains since 1999. *Environmental Research Letters*, 16(7):074049.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth system science data*, 13(9):4349–4383.

Palmer, W. C. (1965). Meteorological drought. Research Paper 45, U.S. Weather Bureau. [Available from NOAA Library and Information Services Division, Washington, DC 20852].

Rothe, C. (2010). Nonparametric estimation of distributional policy effects. *Journal of Econometrics*, 155(1):56–70.

SantAnna, P. H. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

Shaik, S., Helmers, G. A., and Atwood, J. A. (2005). The Evolution of Farm Programs and Their Contribution to Agricultural Land Values. *American Journal of Agricultural Economics*, 87(5):1190–1197.

Wei, Y., Epland, M., and Liu, J. H. (2023). Inverse Probability Weighting Difference-in-Differences (IPWDID). *Observational Studies*, 9(3):73–81.

Zivin, J. G. and Neidell, M. (2012). The Impact of Pollution on Worker Productivity. *American Economic Review*, 102(7):3652–3673.

Appendix

Appendix A - NAAQS timeline and covariates description

Tab. A.1 *$PM_{2.5}$ NAAQS timeline*

Publication	Type	Norm	Level	Form
1997 / Jul 18, 1997	Primary & Secondary	24 hour	65 g/m^3	98th percentile, averaged over 3 years
1997 / Jul 18, 1997	Primary & Secondary	Annual	15.0 g/m^3	Annual arithmetic mean, averaged over 3 years
2006 / Oct 17, 2006	Primary & Secondary	24 hour	35 g/m^3	98th percentile, averaged over 3 years
2006 / Oct 17, 2006	Primary & Secondary	Annual	15.0 g/m^3	Annual arithmetic mean, averaged over 3 years
2012 / Jan 15, 2013	Primary	Annual	12.0 g/m^3	Annual arithmetic mean, averaged over 3 years
2012 / Jan 15, 2013	Secondary	Annual	15.0 g/m^3	Annual arithmetic mean, averaged over 3 years
2012 / Jan 15, 2013	Primary & Secondary	24 hour	35 g/m^3	98th percentile, averaged over 3 years
2020 / Dec 18, 2020	No revision.			

Tab. A.2 *O_3 NAAQS timeline*

Publication	Norm	Level	Form
1997, Jul 18, 1997	8 hours	0.080 ppm	Annual fourth-highest daily maximum 8-hr concentration, averaged over 3 years
2008, Mar 27, 2008	8 hours	0.075 ppm	Annual fourth-highest daily maximum 8-hr concentration, averaged over 3 years
2015, Oct 26, 2015	8 hours	0.070 ppm	Annual fourth-highest daily maximum 8 hour average concentration, averaged over 3 years
2020, Dec 31, 2020	No revision.		

Tab. A.3 *Covariates description and data sources*

Variables	Description	Source
Drought	Log of yearly cumulative counts of moderate to extreme dry conditions based on the Palmer Drought Severity Index	Authors calculation based on the TerraClimate
Inversion	Log of yearly sum of days with surface thermal inversion	Authors calculation based on ERA5 - ECMWF
GDD	Log of cumulative sum of Growing Degree Days (GDD)	Authors calculation based on Gridmet
Poverty	All ages in Poverty, Rate Estimate	Census Bureau, Small Area Income and Poverty Estimates (SAIPE)
Off farm	Log of off-farm income per operation	USDA - Census of Agriculture
Chemical	Log of expenses for insecticides, herbicides, fungicides, and other pesticides excluding commercial fertilizer purchased (per operation)	USDA - Census of Agriculture
Fertilizer	Log of expenses for fertilizer, lime, rock phosphate, and gypsum and the costs of custom application (per operation)	USDA - Census of Agriculture
Irrigation	Log of irrigated culivated lands (per operation)	USDA - Census of Agriculture
Concentration	Log of number of operation per crop acres harvested	USDA - Census of Agriculture

Appendix B - Climate and atmospheric covariates

Growing degree days

The calculation of degree days, often referred to as "growing degree days" (GDD), is tailored to the specific thermal requirements of a crop. The method involves using a base temperature (the minimum temperature required for crop growth) and an upper threshold (the temperature beyond which additional warmth does not accelerate growth).

We follow Butler et al. (2018) to calculate GDDs using daily maximum and daily minimum 2 m temperature data from ERA5-Land. ERA5-Land is a reanalysis dataset that offers gridded data at a resolution of 0.1°× 0.1° over the global land surface (Muñoz-Sabater et al., 2021). For any grid point g at any given day d, we compute daily heat unit, $GDD_{g,d}$, as:

$$GDD_{g,d} = \frac{T^*_{min,g,d} + T^*_{max,g,d}}{2} - T_{low} \tag{B.1}$$

where,

$$T^*_{max,g,d} = \begin{cases} T_{max,g,d} & \text{if } T_{low} < T_{max,g,d} < T_{high}, \\ T_{low} & \text{if } T_{max,g,d} \leq T_{low}, \\ T_{high} & \text{if } T_{max,g,d} \geq T_{high} \end{cases}$$

$T^*_{min,g,d}$ is defined by using the same low and high bounds of $T_{low} = 8$°C and $T_{high} = 29$°C.

We then sum $GDD_{g,d}$ over each annual corn and soybeans growing season (from April 1st to September 30th) from 1997 to 2022.

To ensure that we incorporate only regions where corn cultivation is prevalent before computing average EDD values at the county level, we utilize the USDA NASS Crop Frequency Layer (CFL) for corn and soybeans. This layer provides information on the number of years, spanning from 2008 to 2022, during which corn and soybeans have been cultivated at a particular 30-meter grid point. We calculate the average values from the EDD grid points to obtain EDD values at the county level. These values correspond specifically to areas where frequent corn and soybeans cultivation is substantial. Finally, wecalculate the annual total of "growing degree days" for each county..

Drought conditions

We use the Palmer Drought Severity Index (PDSI) developed by Palmer (1965), which is a widely used measure for assessing drought severity based on meteorological data. Since we are interested in the effects of drought on agricultural outcomes, we compute county-level PDSI that reflect drought severity in corn and soybeans areas. To do this, we match each PDSI grid cell value to agriculture production areas using the USDAs Cropland Data Layer (CDL) hosted on CropScape. We then spatially aggregate monthly PDSI grid cell values at the county level. We use a weighted spatial mean that considers the fraction of the cell covered by each county's borders, to obtain agricultural county-specific measures of relevant climate conditions. Finally, we cumulate over each year the number of times each county was affected by moderate to extreme dry conditions drought conditions.

Thermal inversions

Several studies have shown that the use of high-altitude weather conditions, especially thermal inversions, allows for the isolation of the causal effect of pollutants on economic activity (Avila Uribe, 2023; Dechezleprtre et al., 2019). During a thermal inversion, the atmosphere is stable and air circulation is very limited, the inversion layer then acts as a "lid" blocking pollutants along the layer and thus promoting the occurrence of pollution peaks.

Data on thermal inversions comes from the ERA5 reanalysis. We obtain mean air temperature measures over the North American domain at a spatial resolution of 0.1° from January 1, 1995 00:00:00 UTM to December 31, 2020 18:00:00 UTM (every 6 hours). Temperature (*tmp*) are retrieved at multiple pressure levels from 1000 hPa (approximately 111m above the surface) to 1 hPa (top of atmosphere) divided into $j = 1, ..., 37$ categories where $j = 1$ denotes the lowest atmospheric level above surface (higher pressure level). Since, surface pressure may be lower than atmospheric pressure at higher pressure levels due to land surface elevation or low pressure system, j is defined dynamically in each grid cell and for each time unit so that the index $j = 1$ always corresponds to the lowest pressure level above surface.

Following Dechezleprtre et al. (2019) and Chen et al. (2018), we define the presence of thermal inversions τ if temperature inversions occur between the lowest level and the second lowest level above the local surface for any gridcell and any time level unit:

$$Inv_s = \frac{temp_{j=2} - temp_{j=1}}{z_{j=2} - z_{j=1}} \times \mathbb{1}\{I_J > \gamma\} \tag{B.2}$$

where $\mathbb{1}\{I_J > \gamma\} = 1$ if $I_J > \gamma$ and 0 otherwise. $temp$ is the air temperature and z is the altitude. $temp_{j=2} - temp_{j=1}$ is the strength of the temperature inversion in Kelvin, and $z_{j=2} - z_{j=1}$ is the depth of the inversion measured in hPa. The parameter γ ($0 \leq \gamma \leq 0.5$) measures the critical adiabatic thermal gradient in $K \cdot hPa^{-1}$, characterizing the magnitude of the inversion.

Our final index is calculated as the sum of days with at least one thermal inversion during a year. Our benchmark index is calculated using $\gamma = 0$. We conducted tests on various threshold levels for the parameter γ to identify days with the most significant thermal inversions, which yielded comparable results in our econometric analyses.