Crop Insurance Subsidies and Optimal Policy Design: Evidence from the U.S. Midwest

Job Market Paper of Homa Taheri*

Private copy. Please do not circulate.

Abstract

I study how subsidized crop insurance enrollment affects crop production and what subsidy rate optimizes these responses for maximizing social welfare. Using a county-crop-year panel of corn and soybeans in the U.S. Midwest over 1996-2025, I find that higher enrollment raises mean yield, with an elasticity of about 0.08. At the average insured share, a 0.1 increase in enrollment raises yield by about 1%. Enrollment also increases Commodity Credit Corporation borrowing, consistent with insurance relaxing credit and liquidity constraints. I find no evidence that enrollment raises sensitivity to extreme heat, yield variance, or downside risk. The intensive margin estimates are small and unstable and do not become more adverse under heat. The results therefore provide little support for moral hazard in production. I then use a sufficient-statistics framework to calculate the optimal subsidy. I find that risk protection incentives for farmers alone do not justify any subsidies. However, including disaster relief savings for the government, raises the optimum subsidy to 39%, compared with the observed rate of about 57% for buy up insurance. Including the credit-related production benefits can raise the optimum above the observed rate.

Keywords: Crop Insurance, Agricultural Subsidies, Optimal Subsidy Design, Credit Constraints, Moral Hazard, Disaster Relief

JEL Codes: Q18, G22, H21, Q14, H84

*Homa Taheri is a Ph.D. candidate in Economics at Indiana University Bloomington. Contact: htaheri@iu.edu. I am especially grateful to Ahmad Lashkaripour for his guidance and support throughout this project. I also thank Volodymyr Lugovskyy, Shellye Suttles, Ruli Xiao, Scott Robeson, Marc Bellemare, and Christian Matthes for their valuable advice and feedback. I thank participants at the 89th and 90th Annual Meetings of the Midwest Economics Association, and the 3rd Hoosier Economics Conference for helpful comments. I am also grateful to staff of the U.S. Department of Agriculture, especially the Risk Management Agency, for essential assistance with data and institutional details. My doctoral studies and this research were supported by the Indiana University Economics department's Distinguished Alumni Fellowship, the Fernandus & Elizabeth J. Payne Fellowship, the Frank T. Bachmura Award for an Outstanding Student in Development Economics or Comparative Systems, the McKinney Climate Fellowship, and the Ostrom Workshop Fellowship. All errors are my own.

"*Crop Insurance and a system of storage reserves would help to protect the income of individual farmers against the hazards of crop failure; it would help to protect consumers against shortages of food supply and against extremes of prices; and finally, it would assist in providing a more nearly even flow of farm supplies, thus stabilizing farm buying power and contributing to the security of business and employment.*" - Franklin D. Roosevelt, February 18, 1937, Message to Congress on Crop Insurance (Roosevelt 1937).

1 Introduction

In conversations with members of the Chicago Farmers,[1] I repeatedly hear that federal crop insurance changes how people farm: it discourages growing any produce but the covered commodities, it finances otherwise unaffordable risky practices, and it decreases the incentives to adapt to a changing climate. These are textbook moral hazard concerns, told by the people closest to production. The empirical literature, however, presents a mixed picture. Insurance affects input use and acreage decisions (Mieno, Walters, and Fulginiti 2018; Jisang Yu, A. Smith, and Sumner 2018), but the evidence on realized yields and yield risk is less settled. Roberts, Key, and O'Donoghue (2006) find evidence of moral hazard in Texas but not in Iowa or North Dakota, while a series of studies report that insurance raises the sensitivity of yields to extreme heat or deepens downside risk (Annan and Schlenker 2015; Connor and Katchova 2020; Wang, Rejesus, and Aglasan 2021; Chemeris, Y. Liu, and Ker 2022; Regmi, Tack, and Featherstone 2023).

In this paper, I revisit this topic and ask two questions: first, does subsidized enrollment distort production, in the specific sense of making yields more sensitive to extreme heat, more variable, or more exposed to downside losses? Or does it instead raise crop yields, as a credit channel would predict? Second, given the estimated behavioral responses, what is the optimal buy up subsidy rate? I connect the two questions by using the estimated enrollment and production responses as input factors in the optimal policy analysis.

The setting is the Federal Crop Insurance Program (FCIP), the largest farm safety net program in the United States and the dominant risk management instrument for row crops since the 1990s. Farmers choose whether to buy coverage above the catastrophic coverage free of charge, and how much coverage to purchase. The farmers received a

[1] https://chicagofarmers.org/

subsidy to premiums of about 57% for the buy up insurance, and they receive indemnities when realized yields or revenue fall below guaranteed levels. Premium rates are set by the Risk Management Agency's actuarial system and the subsidy schedule by statute, so private insurers service policies but do not set the prices farmers face. Because the subsidy changes both expected returns and the distribution of risk, it can in principle reshape production: crop choice, land allocation, input use, and investment in adaptation. Whether it does, and in which direction, determines whether the subsidy corrects a market failure or causes a distortion.

Although my empirical setting is the United States, the policy question is broader. Premium subsidies are a common feature of agricultural insurance programs around the world, as documented in Appendix C. The U.S. program is institutionally distinctive, however, because crop insurance is organized and subsidized at the federal level, whereas most other insurance markets are regulated primarily by states under the McCarran-Ferguson Act of 1945 (Koijen and Yogo 2023, p. 19). This federal structure creates a nationally standardized subsidy schedule and actuarial framework, making the FCIP a useful setting for studying how public insurance design affects production and welfare.

In this paper, I assemble a county-crop-year panel of corn and soybeans in eight Midwestern states over the 1996-2025 period, and separate the two margins of enrollment in insurance: the *extensive* margin, the share of planted acreage under buy-up coverage, and the *intensive* margin, insurance liability per insured acre. Both are endogenous, since riskier places face higher rates and coverage responds to expected returns. I instrument them with the county's statutory actuarial base rate interacted with predetermined shifters of the cost of participating: driving distance from population-weighted county centroids to the nearest USDA office, and the historical count and density of local weather observation stations in 1996. I benchmark these access-based instruments against a shift share Bartik instrument built from pre-period crop insurance coverage shares and the national subsidy schedule, and I let covariate balance, over-identification, and per-instrument orthogonality tests determine the best instruments. It turns out that the Bartik instrument is mechanically the strongest in the first stage yet fails the validity tests, while the geographic instruments survive them. I proceed with the geographic instruments throughout the paper.

I then test the moral hazard hypothesis on the moments where theory says it must appear. If insurance discourages adaptation, insured yields should become more sensitive to extreme heat, so I interact enrollment with growing-season climate and instrument each interaction. If insurance encourages risk taking, insured yields should

display greater variance and larger downside losses, so I estimate the effect of coverage on the squared yield surprise and its downside deviation, after removing the surprise of insurance's mean effect so that a mean shift cannot appear as a risk shift. Finally, I carry every estimate into a sufficient-statistics calculation of the welfare maximizing buy-up subsidy, so the positive and normative results form one framework:

$$s^* = F\Big(\underbrace{\varepsilon_D}_{\substack{\text{insurance}\\ \text{demand}}} , \underbrace{\text{moral hazard}}_{\approx 0}, \underbrace{\beta_Y}_{\substack{\text{credit / yield}\\ \text{externality}}} , \underbrace{\rho}_{\substack{\text{disaster-relief}\\ \text{coverage rate}}} , \underbrace{m}_{\substack{\text{peer}\\ \text{multiplier}}} \Big).$$

The demand slope ε_D is estimated from base rate variation in farmer-paid premiums. The moral hazard term is set to zero because the direct tests of Section 7 find no effect of enrollment on heat sensitivity, yield variance, or downside risk on either margin. The yield externality β_Y is the estimated mean effect of enrollment on the extensive margin. The disaster relief coverage rate ρ is estimated from federal disaster-payment data; and the peer multiplier m, which county-level variation cannot identify, is bounded by a calibrated sensitivity.

In terms of the impact of insurance subsidy on crop production, I find that higher enrollment raises mean yield, with a log-log elasticity of about 0.08: at the average insured share in the data (about 0.68), a 0.1 increase in that share raises yield by about 1%. The sign is the opposite of moral hazard, and the evidence points to a credit and liquidity channel in which insurance stabilizes income and serves as collateral. I found the interaction coefficient for enrollment-by-heat to be essentially zero, so insurance does not make yields more sensitive to extreme heat. Once I subtract the mean effect of insurance from the yield surprise, neither margin shifts yield variance or downside risk. The intensive margin is negative but small and its sign is unstable, and it too fails to intensify under heat.

In terms of social welfare consequences of these results, I find that when considering risk protection alone, the optimal buy-up subsidy is zero: marginal enrollees value coverage at what they pay, so subsidizing enrollment spends on the large inframarginal base while creating only second-order surplus (it is because by the definition of being the marginal decision maker, their marginal benefit equals to their marginal cost). But the program was built to replace ad-hoc disaster relief, and relief is not free. Combining USDA disaster outlays with insured loss data for these states, I estimate that ad-hoc relief reimburses about $\hat{\rho} = 0.42$ per dollar of insured disaster loss, and at that crowd out rate, the optimal subsidy is 39%, with a bootstrap confidence interval of $[0.37, 0.40]$, less than the present 57%. Counting the credit-channel yield gain pushes the optimum

higher, and how much higher depends on how much of that gain is a genuine social benefit once input costs are subtracted. Very little is needed: if only a tenth of the estimated gain survives, the observed 57% is exactly optimal, and if all of it does, the optimal subsidy runs to the top of the range I simulate.These mechanisms point to opposite directions. Risk protection and disaster-relief savings justify a subsidy below the current rate, while the credit benefit can justify a higher one. Whether the current subsidy is too high or too low therefore depends on how much of the yield gain is a true net social benefit.

Contribution and related literature. The paper makes two contributions. The first is evidentiary. A long line of work studies behavioral responses to crop insurance through input use and land allocation (Horowitz and Lichtenberg 1993; Bruce A. Babcock and Hennessy 1996; Vincent H. Smith and Barry K. Goodwin 1996; Barry K. Goodwin, Vandeveer, and Deal 2004), and the effect of premium subsidies on acreage planted of subsidized crops (Jisang Yu, A. Smith, and Sumner 2018). For yields, Roberts, Key, and O'Donoghue (2006) find some evidence of moral hazard in Texas but little evidence in Iowa or North Dakota. Annan and Schlenker (2015) find that insured corn and soybeans are 67% and 43% more sensitive to extreme heat, respectively. Connor and Katchova (2020) find that greater participation increases downside yield risk and drought susceptibility, while Wang, Rejesus, and Aglasan (2021) find that participation magnifies the effect of warming on yield risk. Using farm-level data for Kansas, Regmi, Tack, and Featherstone (2023) find that insured dryland corn is about 38% more sensitive to extreme heat. Relative to this literature, I test for moral hazard in the yield outcomes where production distortions would be expected to appear. To do so, I instrument for both margins of enrollment and their climate interactions. I remove the estimated mean effect of insurance from the yield surprise so that a mean shift cannot appear as a change in risk, and I examine heat sensitivity, variance, and downside risk separately. None of these tests finds an adverse production-risk effect. Together, the results show that insurance raises mean yield on the extensive margin without any detectable increase in risk on either margin.

The second contribution is normative. It builds on the sufficient-statistics literature on social insurance (Baily 1978; Einav, Finkelstein, and Cullen 2010; Chetty and Finkelstein 2013; Hendren and Sprung-Keyser 2020). Related applications study crop insurance (Jialing Yu 2017; Klosin and Solomon 2024, see ch. 4) and flood insurance (Gruber and Solomon 2026). The disaster-relief argument follows the Samaritan's dilemma (Coate 1995) and the role of crop insurance in reducing reliance on ad hoc

disaster payments (Glauber 2013). Here every input of the formula is estimated inside the paper: the demand slope from base-rate variation in farmer-paid premiums, the moral-hazard term from the direct tests, which set it to zero, the credit externality from the yield equation, and the crowd-out rate from federal disaster-payment data. Enrollment also clusters sharply in space, with a correlation of 0.91 between a county's buy-up share and the average share of its nine nearest neighbors, so a social multiplier belongs in the formula. Section 8 estimates it with the network design of Bramoullé, Djebbari, and Fortin (2009), but subsidy variation is nearly common across neighboring counties, and the estimate fails the paper's own instrument-validity tests. I calibrate the multiplier instead, over the range the experimental adoption literature supports (Glaeser, Sacerdote, and Scheinkman 2003; Cai, Janvry, and Sadoulet 2015), and the optimum barely moves across it. The observed 57–60% subsidy then sits between the program's two quantified rationales: disaster-relief crowd-out justifies 39% on its own, and a modest net credit benefit carries the optimum past the rate in force.

2 Background

2.1 Institutional Background

Federal Crop Insurance Program (FCIP) is the largest protective program of the government of United States for the farmers. Initiated in 1937, it is also the oldest of the world's public crop insurance schemes. Since 1990s, the Federal crop insurance program has raised to a prominent Government support for U.S. crop farmers. The Federal Crop Insurance was first developed under the Agricultural Adjustment Act of 1938, as a part of the New Deal. Its introduction was a response to the Great Depression and large declines in crop yields resulting from the Dust Bowl. The FCIP initially started off without premium subsidies. But farmers were found to be unwilling to purchase unsubsidized multiperil crop insurance. The Department of Agriculture initiated the premium subsidies to encourage the participation of farmers (Glauber 2004). Before the passage of Federal Crop Insurance Act of 1980, only a small percentage of farmers enrolled in the program. But once the 1980 act authorized premium subsidies and included private insurers into the program for marketing purposes, the participation rate increased significantly. The goal of the 1980 Act was to create an insurance program that would replace disaster relief measures while operating on an actuarially sound basis with limited government interference (Barry K Goodwin 1993). Since then, the Federal Crop Insurance has operated through a shared public private arrangement funded by tax payers and producers and offers three types of insurance coverage: yield-based, revenue-based and area-based. In its current design, it is a heavily subsidized

program - from 2007 to 2016, in aggregate, Government paid 61% of total premium, and producers paid 39% of total premium (Rosa 2018)- and is one of the largest public insurance programs in the United States. With nearly $15 billion in annual subsidies. Like health, unemployment, and disability insurance, crop insurance embodies the fundamental trade-off in social insurance design: public subsidies provide risk protection, but they also shape behavior in ways that can amplify or dampen aggregate risks. This paper examines how subsidy design in crop insurance interacts with extreme weather, offering evidence on both the risk-sharing benefits and the incentive distortions of one of the most consequential U.S. social insurance programs.

A major structural shift occurred with the Federal Crop Insurance Act of 1980, which introduced Multiple-Peril Crop Insurance (MPCI) as the core federal product. MPCI expanded coverage to nearly all unavoidable natural causes of yield loss and substantially broadened the set of eligible crops and counties (Vincent H Smith and Glauber 2012). Subsequent reforms—most notably the Crop Insurance Reform Act of 1994—further consolidated MPCI as the dominant risk management instrument by introducing heavily subsidized catastrophic (CAT) coverage and linking access to ad hoc disaster assistance to insurance participation (Glauber 2013). By the mid-1990s, the FCIP had effectively transitioned from a narrowly targeted, named-peril system to a nationwide, multi-peril income stabilization program.

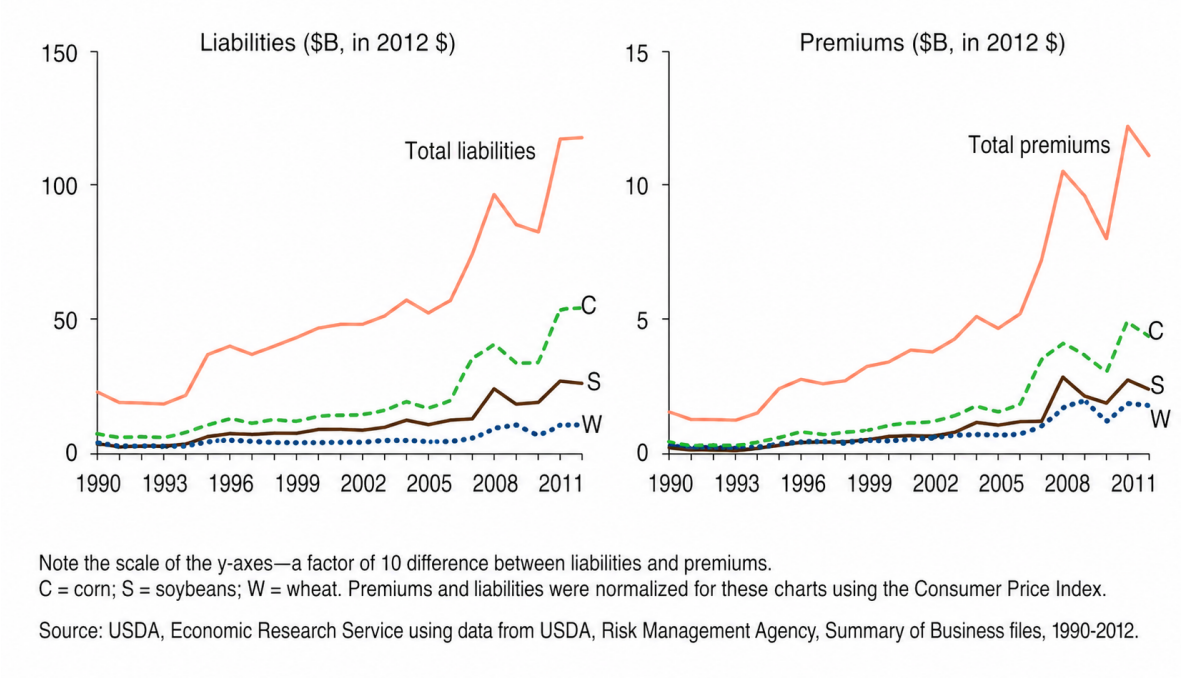


Figure 1: Growth in participation in crop insurance: alternate measures, normalized to 2012 dollars (O'Donoghue 2014, Fig. 2)

2.2 The Premiums, Subsidies and Types of Protection

The US government has played a historically active role in protecting producers against yield and revenue risks by developing, promoting, and subsidizing crop insurance. By reducing the costs of crop yield and crop revenue insurance to farmers, USDA encourages insurance participation (Young, Vandeveer, and Schnepf, 2001). USDA subsidizes insurance premiums based on the riskiness of insured crops at rates that vary by type of insurance, coverage levels, and unit types; although for a given combination of coverage level, insurance product, and unit type, the subsidy rates of all crops are the same Jisang Yu, A. Smith, and Sumner (2018). This makes the premium rates and subsidy rates that participating farms face endogenous to their production strategies. In terms of type of insurance, the two most common products for major field crops are yield and revenue protection. For yield protection, the indemnity is triggered when actual yield is smaller than historical average yield. For revenue protection, the indemnity is triggered when actual yield times harvest price is less than historical average yield times the larger of either projected price or harvest price. Revenue protection generally has a higher premium rate per dollar of liability than yield protection. Update the Table for the recent years.

Table 1: Major Legislative Changes in U.S. Farm Policy and the Federal Crop Insurance Program, 1985-2014. Compiled using Rosch (2021, Table A-1), Edwards (2025), DeSimone (2021), and Library of Congress (2026).

Year	Legislation	Major Changes to FCIP and Related Farm Support Policy
1985	Food Security Act of 1985	Passed amid the farm financial crisis of the 1980s. Introduced conservation compliance requirements tied to crop insurance purchases. Also expanded farm support through new subsidy-related programs, including the Export Enhancement Program and the Conservation Reserve Program.
1994	Federal Crop Insurance Reform and Department of Agriculture Reauthorization Act of 1994†	Increased premium subsidies and introduced catastrophic (CAT) coverage. Required crop insurance purchases for eligibility for commodity support and related farm programs. Authorized prevented planting coverage.
1996	Federal Agriculture Improvement and Reform Act of 1996	The "Freedom to Farm" law aimed to increase planting flexibility and better align producer decisions with market conditions. For FCIP, it removed the requirement linking commodity support and conservation compliance to crop insurance purchases, created USDA's Risk Management Agency (RMA), and disallowed sales of crop insurance through USDA county offices unless approved insurance-provider coverage was unavailable. Although intended as a reform, subsequent supplemental subsidy bills beginning in 1998 expanded support beyond original expectations.
2000	Agriculture Risk Protection Act of 2000†	Increased premium subsidies to near-present levels. Expanded and strengthened the crop insurance program by granting authority for revenue coverage and insurance for livestock and authorized the creation of Section 508(h) pilot programs.
2002	Farm Security and Rural Investment Act of 2002	Reversed some of the reform orientation of the 1996 law and returned toward more traditional subsidy expansion and price support. Added a countercyclical price support program, expanded existing programs to additional crops, and substantially increased projected subsidy payments over the following decade.
2008	Food, Conservation, and Energy Act of 2008	Expanded farm subsidies. Within FCIP, it introduced enterprise and whole-farm units and increased subsidies for area coverage plans.
2014	Agricultural Act of 2014	Reorganized farm support without reducing the overall level of benefits. It ended the direct payment, countercyclical payment, and Average Crop Revenue Election programs while expanding crop insurance and adding the Agriculture Risk Coverage (ARC) and Price Loss Coverage (PLC) programs. Within FCIP, it introduced area-based shallow-loss coverage and yield-exclusion options and relinked conservation-compliance requirements.

Note: Entries marked with † were major standalone laws affecting the Federal Crop Insurance Program (FCIP), but they were not omnibus farm bills of the type typically listed in standard farm-bill timelines.

3 Data

1- USDA RMA Base and Reference Rates for Insurance Premium The actuarial rates that the USDA Risk Managemnt Agency (RMA) uses to set the price of the federal crop insurance offers is available for years 1996 onwards on the RMA's Actuarial Data Master (ADM) archive (U.S. Department of Agriculture, Risk Management Agency 2026a)[2]. The stored rate is for yield-risk insurance at the 0.65 coverage. This county-crop panel covers two distinct ADM eras. The modern era covers years 2011-2027 in which the data is reported with a modern data structure of pipe-delimited. The legacy era reports data in fixed-width record-type-1 files. In both data structure, the *A01010* files report the actuarial rates. The data structure in legacy era changes twice. I call the structure in year 1996, Layout A1, in year 1997, Layout A2, and for years 1998-2010, Layout B. Using a parser module, I converted the legacy era data to the modern era data. Appendix I provides more details.

2- Climate data To produce a 1996-2025 panel, I used NOAA nClimGrid-Daily Version 1, a 1951-present daily climate dataset for the contiguous United States produced by the National Centers for Environmental Information (NCEI). NCEI provides nClimGrid-Daily as native 5 km grid files (NetCDF), plus pre-computed area-averaged time series at several aggregation scales, including counties. I used this data set to calculate the number of heating degree days, number of medium degree days, precipitation and its square over the 2013-2025 period, and compared them with the values reported in the replication package of Annan and Schlenker (2015) to minimize any errors . Appendix J provides more details.

3-Area planted and area insured The construction follows the data methodology of Annan and Schlenker (2015), which combines county-level planted acreage from USDA NASS QuickStats (U.S. Department of Agriculture, National Agricultural Statistics Service 2026). Appendix K provides more details.

4- Expected price of farmers The price control I use is the state-crop-year level expected price constructed by Jisang Yu, A. Smith, and Sumner (2018) for their study

[2]Although years 1994 and 1995 are also listed on this webpage, both are repositories are empty. So, I focus on data for years beginning in 1996.

of crop insurance premium subsidies and crop acreage. Their regression-based expected price combines two ingredients:

1. Received price: the state-level average marketing-year price received by farms for each crop, from the USDA National Agricultural Statistics Service (NASS); and

2. Futures price: a pre-planting average of the relevant commodity futures contract. The futures prices were retrieved from Yahoo Finance's public chart APT endpoint [3]

Jisang Yu, A. Smith, and Sumner (2018) combine the two by regressing the state received price on the futures price and state dummies, and use the fitted values as the expected price. Using their method, I extended the series from its original window (1989-2014) through 2025[4]. See Appendix L for more details.

5- Weather Station Locations and Activity Periods I downloaded the weather station locations and activity periods from NOAA's Global Historical Climatology Network–Daily (GHCN-Daily) dataset (Menne, Durre, Korzeniewski, et al. 2012; Menne, Durre, Vose, et al. 2012; NOAA National Centers for Environmental Information 2025c). I retrieved the Station coordinates from the GHCN-Daily station metadata file (NOAA National Centers for Environmental Information 2025b), and station-year availability from the GHCN-Daily inventory file (NOAA National Centers for Environmental Information 2025a). I then defined the density of weather observation stations as the "number of stations per county" in baseline year of 1989 based on this collected data sets.

NOAA National Centers for Environmental Information (2025b) reports 129,658 weather stations, out of which 75,847 (58.50 percent of total in the world) are based in the United States. Out of these, 14,217 weather stations (10.97 percent of total in the

[3] This endpoint (`query1.finance.yahoo.com/v8/finance/chart/`) is not officially documented or supported by Yahoo; there is no vendor reference manual for it. It is, however, a long-standing public endpoint that is widely used and documented by the open-source `yfinance` Python library, which is the de facto reference for programmatic access to Yahoo Finance data (Aroussi 2019).

[4] One might ask why the regression and not the raw received price? To answer to this question, notice that the acreage decision is made at planting, on the basis of the price a farmer expects, whereas the realized NASS received price is determined at/after harvest and is jointly determined with quantity (more acreage → more output → lower realized price). Using the realized price as a regressor would therefore induce simultaneity bias and mistime the decision. The pre-planting futures price is predetermined and forward-looking; projecting the received price onto it (a) isolates the expectation-relevant, exogenous component of price and (b) through the state fixed effects, encodes each state's average basis (the persistent local cash-to-futures gap), so the fitted value is a state-level expected price in received-price units.

US) are located in 8 midwest States of this current research. Limiting the time scope to the 1989 afterwards, I find 12,623 of the weather stations (88.79 percent of midwest weather stations) have been active sometime in years 1989 onwards. Figure 2 shows the change in the number of active weather stations over the 1980-2025 period in the 8 midwest states. Figure 3 shows county-level variance in the number of weather stations over time.

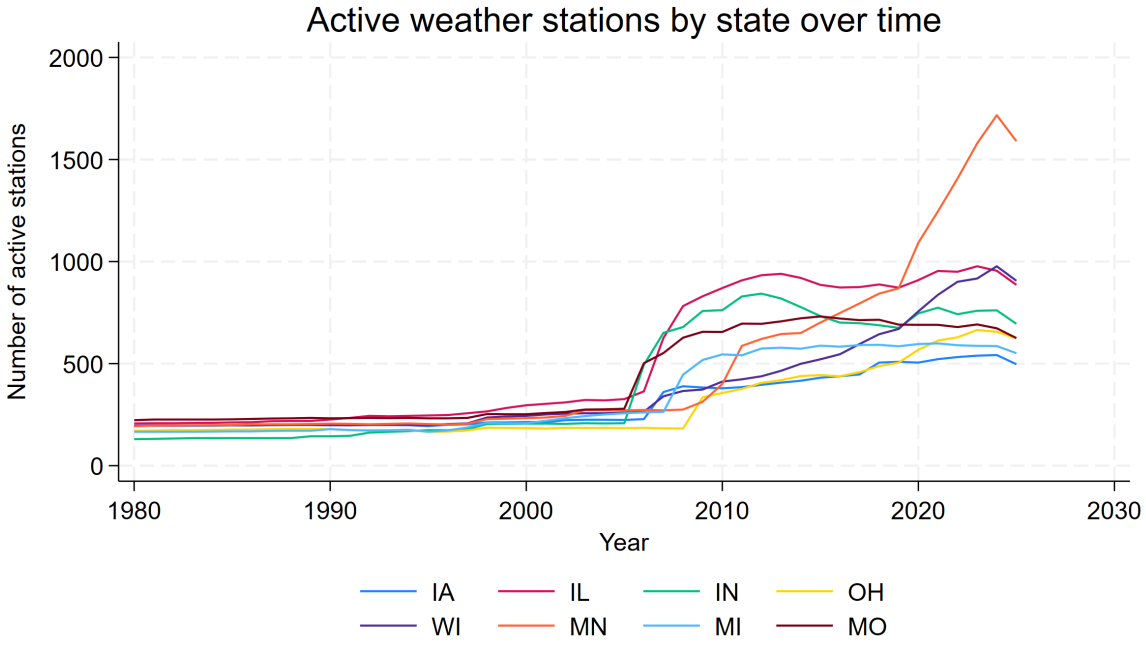


Figure 2: Change in number of weather stations over time

6- Census block level boundary and population data To construct population-weighted county centroids for the 2000, I use block-level population and census boundary data from the 2000 U.S. Census, obtained through the IPUMS National Historical Geographic Information System (NHGIS) (Schroeder et al. 2025a). These data allow for precise spatial weighting of population within each county, greatly improving the accuracy of location-based measures such as travel distance to USDA offices.

7- Access to Marketing Assistance Loans from Agricultural census. I construct county–year measures of Commodity Credit Corporation (CCC) marketing assistance loan exposure from United States Department of Agriculture, National Agricultural Statistics Service (2025). I defined net CCC loan position as loans received minus loans repaid, which proxies for county-level access to short-term agricultural credit. Net CCC loan receipts are transformed using the inverse hyperbolic sine to

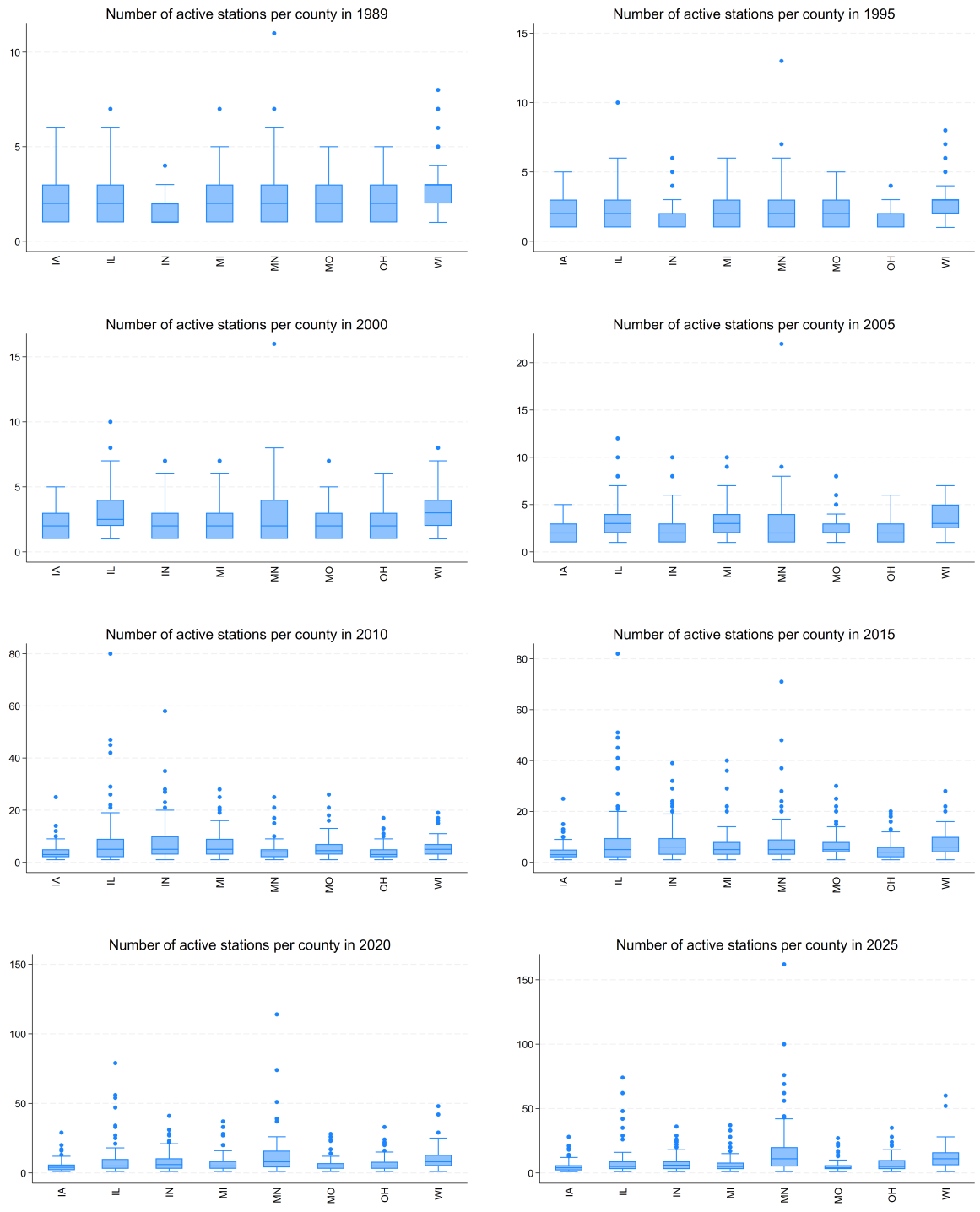


Figure 3: Distribution of the number of weather stations per county by state (1989–2025).

accommodate zero values and reduce right skewness. This data was reported only in Census years. I constructed Intercensal county-level value using Friedman's related-series method (Friedman 1962, Method 2, p. 732). For more details on this approach, see Appendix M. Table 2 provides summary statistics of the data used.

Table 2: Summary statistics for the variables used in the paper

Variable	N	Mean	S.D.	p10	p90
Panel A. Outcomes					
Yield, corn (bu/planted acre)	17,940	141.35	39.17	88.19	189.22
Yield, soybeans (bu/planted acre)	17,635	45.14	9.96	32.11	57.72
Yield, winsorized, corn (bu/acre)	17,940	141.38	38.70	88.19	189.22
Yield, winsorized, soybeans (bu/acre)	17,635	45.15	9.78	32.11	57.72
ln yield, corn [outcome]	17,940	4.91	0.32	4.48	5.24
ln yield, soybeans [outcome]	17,635	3.78	0.23	3.47	4.06
Panel B. Treatments (endogenous)					
Buy-up insured share [extensive]	35,575	0.68	0.25	0.28	0.95
Liability per insured acre (USD)	35,575	360.25	231.85	87.29	674.95
Liability per acre, winsorized [intensive]	35,575	359.62	230.16	87.29	674.95
Panel C. Controls					
Moderate degree days	35,575	165.41	27.03	130.09	201.92
Extreme-heat degree days	35,575	21.99	21.60	2.88	49.80
Precipitation (cm)	35,575	60.37	13.97	42.91	78.23
Precipitation squared	35,575	3839.99	1803.07	1840.84	6120.25
ln expected price (2013 USD)	35,575	5.50	0.45	4.92	6.05
Panel D. Program and scale					
Planted acres	35,575	77743.39	56289.93	13600.00	154500.00
Premium per insured acre (USD)	35,575	32.41	21.15	8.19	61.68
Subsidy rate	35,575	0.57	0.11	0.46	0.69
Actuarial base rate	35,566	0.06	0.05	0.01	0.13

4 Empirical Strategy

4.1 A Model of Enrollment, Inputs, and Yield

In this section, I formalize the idea of impact of insurance subsidy on production, and show why the moral hazard hypothesis has to be tested on higher moments of yield rather than on its mean alone. Consider a farmer in county i growing crop c in year t. Before the season, the farmer chooses buy up enrollment D_{ict} and a vector of yield-relevant inputs x_{ict}, which collects purchased inputs such as fertilizer together with adaptation effort. Yield is

$$y_{ict} = f_c(x_{ict}, \omega_{ict}) + \varepsilon_{ict}, \qquad \text{with} \qquad \frac{\partial^2 f_c}{\partial x \, \partial \omega} \neq 0, \tag{1}$$

where ω_{ict} is realized growing-season weather, measured empirically by the climate controls in $\mathbf{X}_{ict}$, and ε_{ict} is an idiosyncratic shock. The cross partial restriction (1) represents that inputs and adaptation change how sharply yield responds to weather, so x shifts the entire conditional distribution of y and not only its mean. Without it, effort could not affect heat sensitivity, variance, or downside risk, and the hypotheses of Section 7 would have no counterpart in the model. Let $s_{ict} \in [0,1]$ be the premium subsidy rate and $P_{ict}(D)$ the actuarially rated premium, so the farmer pays $(1 - s_{ict})P_{ict}(D_{ict})$. Buy-up coverage sets a yield guarantee $\bar{y}_{ict}(D)$ that is increasing in D and pays an indemnity when realized yield falls short of it:

$$I_{ict} = p^g_{ict} \max\{0,\ \bar{y}_{ict}(D_{ict}) - y_{ict}\}, \tag{2}$$

where p^g is the insured price. Because the payout depends on realized yield, it depends on the inputs that produce that yield: raising effort raises yield and therefore reduces the expected indemnity. That is the moral-hazard channel. Equation (2) is written as a yield guarantee. The treatment measured in the data is enrollment in buy-up coverage, and it includes both the yield-protection and revenue-protection policies, with prevalence of revenue insurance policies in more recent years. Writing the contract in its yield form is the conservative choice for a moral-hazard test, since under revenue protection, the trigger is revenue rather than yield, and the harvest-price provision raises the guarantee when prices rise. Section 7.1 returns to the shift toward revenue policies as one reason a null today can coexist with distortion under the earlier, yield-based regime captured by earlier studies. Realized profit is

$$\pi_{ict} = p_{ict}\, y_{ict} - k(x_{ict}) - \zeta(x_{ict}; F_{ict}) - (1 - s_{ict})P_{ict}(D_{ict}) + I_{ict}, \tag{3}$$

with p_{ict} the output price, $k(\cdot)$ the direct cost of inputs, and $\zeta(\cdot)$ the cost of financing those inputs through the season, decreasing in the farmer's financial position F_{ict} (collateral, borrowing capacity, and cash on hand). The function ζ is where the credit channel lives: insurance stabilizes income and serves as collateral, so enrollment raises F and lowers the shadow cost of carrying inputs. F is a state variable of the model, and its empirical counterparts are Commodity Credit Corporation borrowing and fertilizer spending.

The farmer chooses (D_{ict}, x_{ict}) to maximize $\mathbb{E}\big[U(\pi_{ict}) \mid \mathcal{J}_{ict}\big]$, where U satisfies $U' > 0 > U''$, so the farmer is risk averse, and $\mathcal{J}_{ict}$ is the information available at planting and sign-up: expected prices, baseline climate risk, and financial position. At an interior

optimum the two first-order conditions are

$$\mathbb{E}\Big[U'(\pi_{ict})\Big\{\big(p_{ict} - p^g_{ict}\mathbf{1}\{y_{ict} < \bar{y}_{ict}\}\big)\frac{\partial f_c}{\partial x} - \frac{\partial k}{\partial x} - \frac{\partial \zeta}{\partial x}\Big\}\Big] = 0, \tag{4}$$

$$\mathbb{E}\Big[U'(\pi_{ict})\Big\{p^g_{ict}\mathbf{1}\{y_{ict} < \bar{y}_{ict}\}\,\bar{y}'_{ict}(D_{ict}) - (1 - s_{ict})\frac{\partial P_{ict}}{\partial D}\Big\}\Big] = 0. \tag{5}$$

Three readings of these conditions organize the rest of the paper. First, (5) gives $\partial D^*/\partial s > 0$: a higher subsidy lowers the marginal cost of coverage and raises enrollment. That response is the demand slope estimated and used in Section 9. Second, (4) contains the incentive cost of insurance. In states where the guarantee binds, the marginal product of effort is valued at $p - p^g$ rather than at p, because higher yield reduces the indemnity: insurance confiscates part of the return to effort in precisely the states where effort matters most. Enrollment therefore lowers x in those states, which lowers mean yield, and, since $f_{x\omega} \neq 0$, the reduction is concentrated in bad-weather realizations. So, moral hazard predicts not only a lower mean but yields that are more sensitive to extreme heat and more dispersed with deeper downside losses. Third, the credit channel pushes the other way, and it appears in the same condition: enrollment raises $F(D)$, which lowers $\partial\zeta/\partial x$ and raises x. With higher extra inputs, the weather response becomes weaker, so heat sensitivity falls rather than rises. Equation (4) therefore contains both forces, the indemnity pulling effort down and the financing term pushing it up. The mean effect of enrollment is therefore the sum of two opposing forces and is ambiguous in sign, which is why the moment tests of Section 7 do the discriminating work: only the moral-hazard channel makes yields more heat-sensitive and more downside-exposed.

Solving (5) for D^*_{ict}, substituting the induced input choice $x^*(D^*_{ict}, \mathbf{X}_{ict}, F_{ict})$ into (1), and taking logs delivers the estimating equation of the next subsection, in which log yield is regressed on enrollment with climate and price controls and county, crop, and year fixed effects. The coefficient β on D is a composite of the two channels, net of one another. The model also indicates where identifying variation can come from. The subsidy rate s is set by statute and moves nearly all counties at once, which is the collinearity that defeats the dynamic exercise in Section 8, so it cannot serve as county-level variation. What does vary locally is the cost of participating: the county's actuarial base rate scales $P_{ict}(\cdot)$, and predetermined access to insurance sales and to weather records shifts the transaction cost of enrolling. Both enter (5) without entering the production function (1), which is the exclusion restriction the instruments of the next subsection impose. Finally, (5) also defines the farmer's willingness to pay for coverage; Section 9 works with the demand curve that this condition traces out, so the

welfare analysis values coverage at the same margin the farmer does.

4.2 From the Model to the Estimating Equations

Insurance take-up is endogenous: riskier crops and counties face higher premium rates, and coverage choices respond to expected returns. I therefore estimate the yield equation by two-stage least squares,

$$\ln y_{ict} = \beta\, D_{ict} + \mathbf{X}'_{ict}\theta + \alpha_i + \phi_c + \delta_t + u_{ict}, \qquad D_{ict} = \boldsymbol{\pi}'\mathbf{Z}_{ict} + \mathbf{X}'_{ict}\rho + \alpha_i + \phi_c + \delta_t + e_{ict}, \quad (6)$$

where D_{ict} is an enrollment measure, $\mathbf{X}_{ict}$ the climate and price controls, $\alpha_i, \phi_c, \delta_t$ county, crop, and year fixed effects, and $\mathbf{Z}_{ict}$ the excluded instruments (base rate $\times$ access shifters). Identification comes from within-county, within-year variation in the cost of participating that is plausibly unrelated to contemporaneous agronomic shocks.

To test moral hazard directly (Section 7) I add interactions of D with the climate variables; because $D \times C_k$ is also endogenous, I instrument it with $\mathbf{Z} \times C_k$ (the instrument interacted with the exogenous climate). To study peer effects (Section 8) I form a $k = 9$ nearest-neighbor spatial weight matrix W over county-crop units and instrument neighbors' enrollment WD with neighbors' instruments $W\mathbf{Z}$ and $W^2\mathbf{Z}$.

Throughout I report the Kleibergen-Paap (KP) first-stage F as a measure of instrument strength and the Hansen J test of over-identifying restrictions as a check on validity.

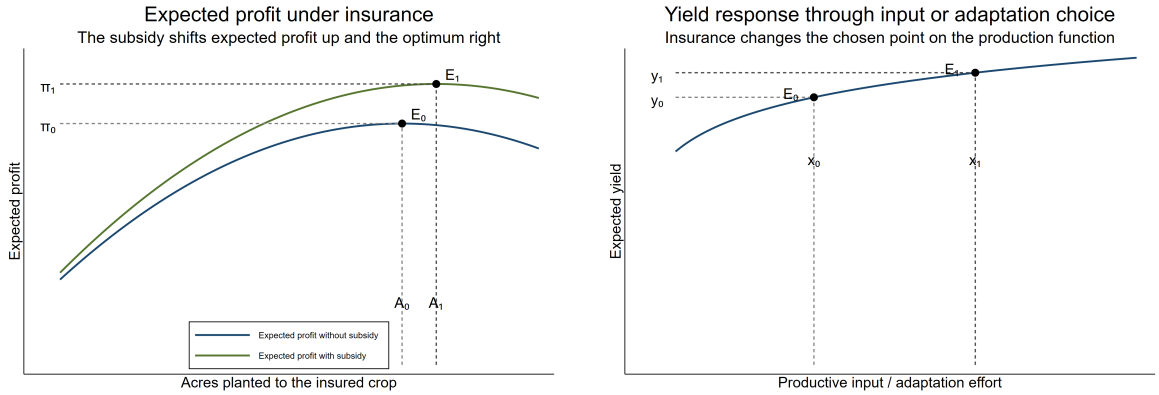
4.3 Hypotheses

Empirical hypotheses. The first three hypotheses examine whether subsidized insurance produces moral hazard in crop production. First, if insurance weakens production incentives, greater enrollment should lower mean yield; if instead it relaxes credit and liquidity constraints, greater enrollment should raise crop yield. The sign of the mean-yield effect therefore distinguishes the two mechanisms. Second, if insurance reduces incentives to adapt, greater enrollment should make yields more sensitive to extreme heat, as in Annan and Schlenker (2015). Third, if insurance encourages risk-taking, greater enrollment should increase yield variance or downside losses, especially under extreme heat.

The fourth hypothesis concerns the credit mechanism. If insurance relaxes financing constraints by stabilizing income or improving access to credit, greater enrollment should raise borrowing and shift input use. I examine this prediction in Section 6. I

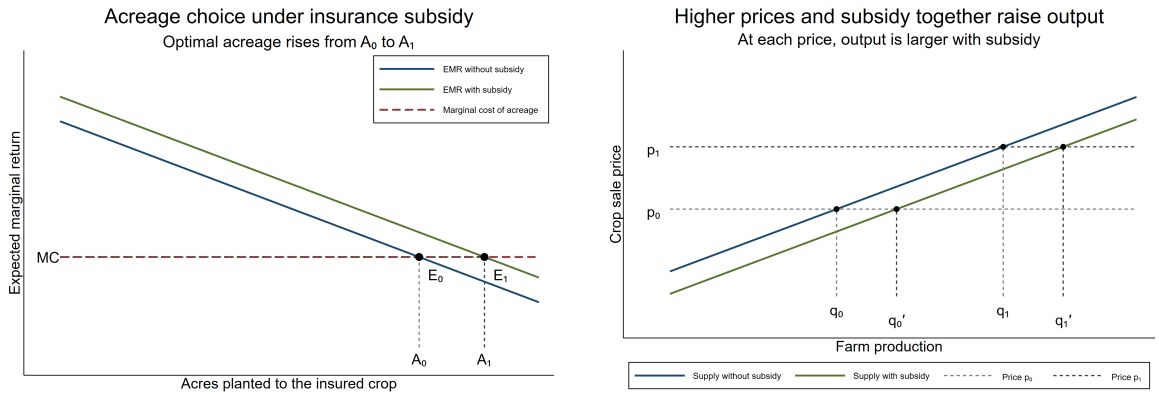
also ask whether enrollment spreads across neighboring counties through peer effects. Because neighboring counties share weather, soil, prices, and local institutions, spatial correlation alone is not evidence of causal diffusion. In Section 8, I examine whether neighbors' enrollment has an effect beyond these shared conditions.

Conceptual illustration. Figures 4-8 illustrate these channels in a stylized model of a representative farmer. Figures 4 and 5 show how a premium subsidy raises expected profit and shifts the optimal acreage and input choices: the same margins that could carry moral hazard (less care, more risk) also carry the credit-channel gain (more inputs, more adaptation), which is why the sign of the mean yield effect is an empirical question rather than a theoretical one. Figure 6 and Figure 8 distinguish the subsidy channel, an outward shift of supply, from the price channel, a movement along a given supply curve; this distinction is why expected output prices enter as controls in every specification. Figure 7 collects the channels in a single summary.



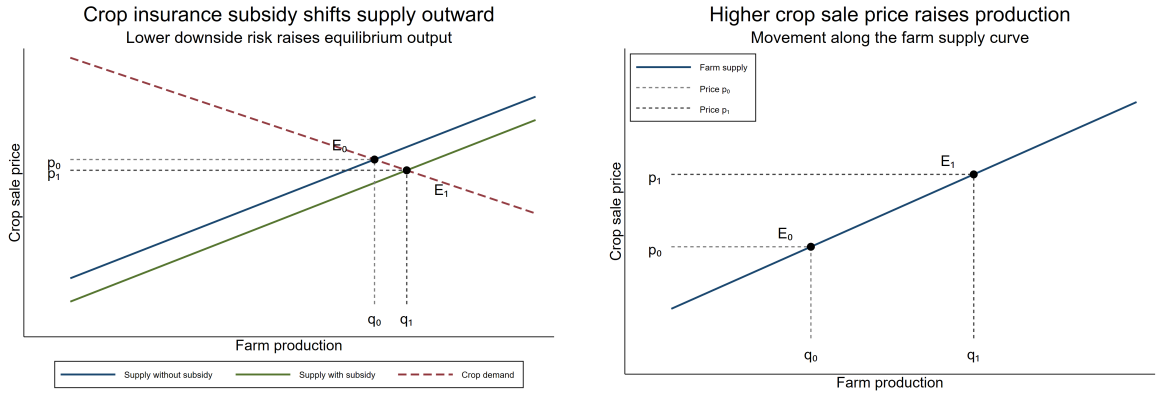
(a) Expected profit under the parametric model (b) Yield response under the parametric model

Figure 4: Panel (a): insurance raises expected profit and shifts the optimum acreage rightward, from A_0 to A_1. Panel (b): the subsidy raises the chosen input or adaptation effort, from x_0 to x_1; it need not shift the agronomic yield function itself, but can raise expected yield by changing input, technology, or adaptation choices. *Source:* Author's illustration.



(a) Acreage choice under the parametric model (b) Combined comparative statics

Figure 5: Panel (a): the subsidy raises the expected marginal return of acreage and shifts optimal acreage rightward, from A_0 to A_1. Panel (b): at each output price p_0 and p_1, the subsidy implies more output, from q_0 to q_0' and from q_1 to q_1'. *Source:* Author's illustration.



(a) Supply shift under the parametric model (b) Price and production effects

Figure 6: Conceptual distinction between subsidy-driven supply shifts and price-driven output responses. Panel (a): the crop insurance subsidy shifts supply outward by reducing downside risk, so equilibrium output rises. Panel (b): output also responds to a higher crop price through movement along the supply curve, from q_0 to q_1. *Source:* Author's illustration.

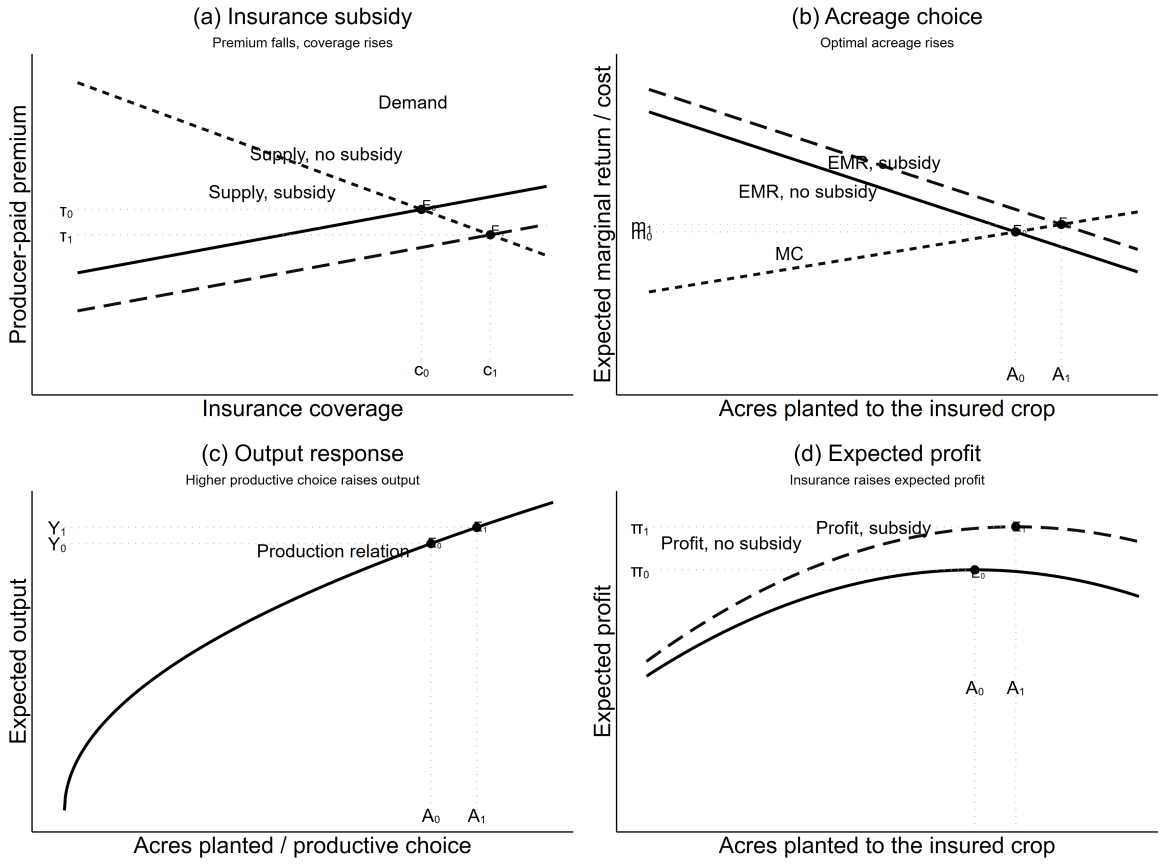


Figure 7: Four-panel summary of the main conceptual channels through which crop insurance subsidies may affect farm behavior: lower producer-paid premiums, increased acreage planted to insured crops, changes in production or adaptation choices, and higher expected profit. *Source:* Author's illustration.

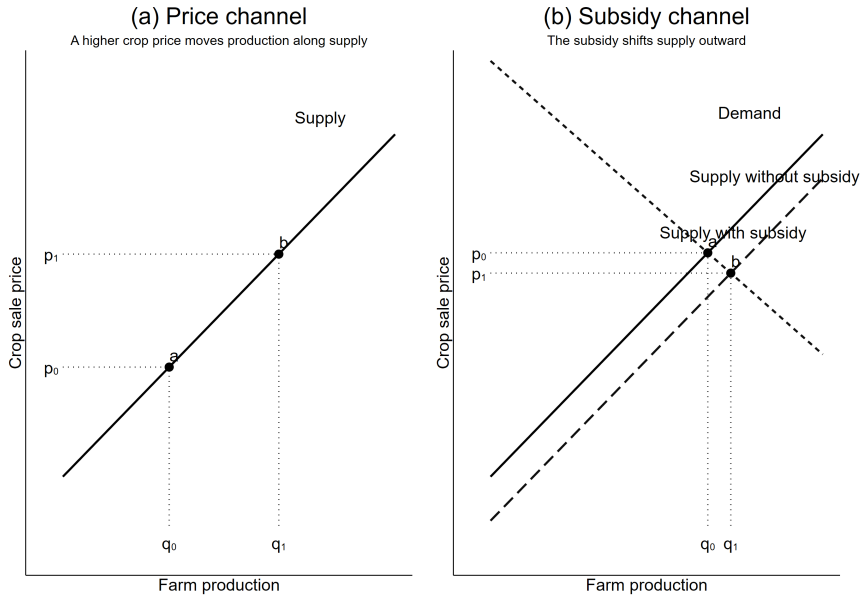


Figure 8: Two-panel comparison of the price channel and the subsidy channel. Panel (a): higher crop prices increase production through movement along a given supply curve. Panel (b): crop insurance subsidies instead shift supply outward, altering equilibrium output and price. *Source:* Author's illustration.

5 Identification Strategy

The setup of this paper has endogeniety problem, as a riskier crop or a riskier county has a higher premium rate for a given crop insurance choice (Keith H. Coble and Barnett 2013). As a result, farmers tend to plant less risky crops and plant less in riskier counties, ceteris paribus. Besides this, when considering the heterogenous effect of policy on each unit, the two-way fixed effects estimators can perform very poorly, resulting in it's incompletent estimate of average or weighted average of units coefficient. The only exception is when some units are left untreated (Sun and Shapiro 2022, p. 195). Since in RMA Summary of Business data on crop insurance, no county-year observation has a value of insurance enrollment equal to zero, the TWFE presented in the Section **??** are only valid in absence of treatment effect heterogeneity [5]. I use an instrumental variables framework that provides plausibly exogenous variation in intensive and extensive margins. The instrument for extensive margin is the driving distance between USDA county offices and population-weighted county centroid. The instrument for intensive margin is the number of weather stations in each county. Given the borderline strength of my instrumental variables, I present my IV results as complementary to those of my

[5] I borrowed this understanding of implications of Federal-level agricultural policies and crop prices from Bellemare, J.-Y. Liu, and Hadrich (2025, p. 10)

preferred core TFWE specifications.

5.1 Instruments

I considered several candidate instruments based on administrative access, historical measurement infrastructure, and pre-period climate exposure. Appendix F summarizes these candidates and the diagnostic reasons for excluding some from the preferred specification. The preferred instrument is selected based on its institutional rationale, first-stage strength, and performance in pre-period balance tests, rather than on the magnitude or statistical significance of the second-stage estimate.

Intuitively, the instruments capture whether statutory base rates have different effects on insurance enrollment in counties with different levels of pre-existing information and administrative access: counties with more or fewer weather stations, and counties closer to or farther from USDA offices.

1- Shortest driving distance between the population-weighted County centroids and USDA offices as an instrument in year 2000. To measure geographic access to agricultural service infrastructure, I construct a county-level measure of driving distance to the nearest USDA offices. First, I compute population-weighted county centroids using Census 2000 block-level population counts from IPUMS NHGIS Schroeder et al. (2025b). For each census block, the geometric centroid is calculated, and county-level centroids are obtained as the population-weighted average of block centroids within each county, ensuring that the resulting location reflects the spatial distribution of residents rather than land area.

To approximate travel distance along the road network, I use the U.S. Census Bureau's TIGER/Line road network files from the year 2000 (Bureau 2000). Road segments are converted into a graph representation in which nodes correspond to segment endpoints and edges represent road links weighted by their physical length. For each population-weighted county centroid and each USDA office location, the nearest road network node is identified. Driving distance is then computed as the shortest-path distance along the road network between the centroid's nearest node and the nearest node associated with a USDA office.

All spatial layers are projected to a common projected coordinate reference system (NAD83 / USA Contiguous Albers Equal Area) prior to distance calculations to ensure consistency and metric accuracy. Distances are computed separately for each state to reduce computational complexity, and for each county the minimum driving distance

across all USDA offices within the state is retained. The resulting measure captures effective geographic accessibility to USDA offices while accounting for both population location and the structure of the historical road network.

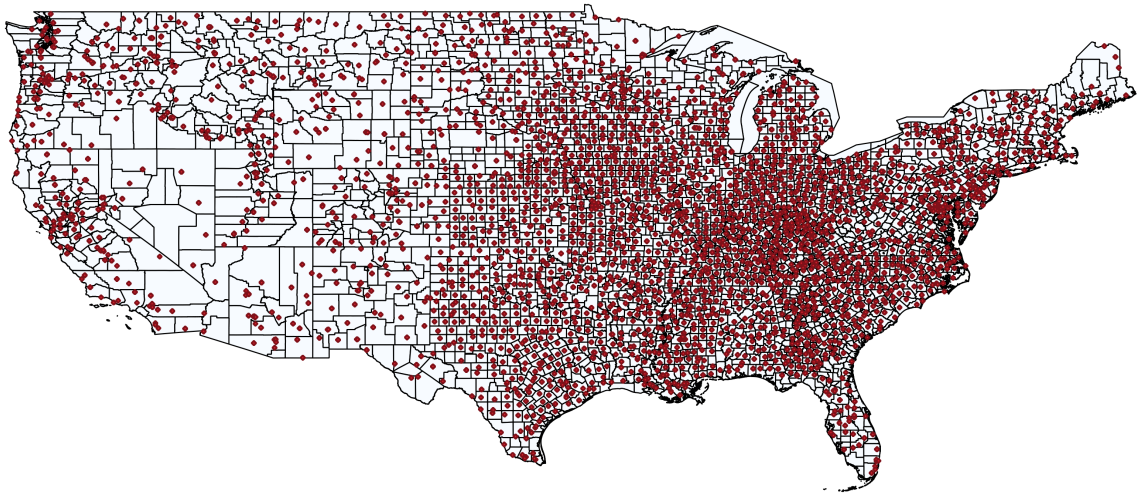
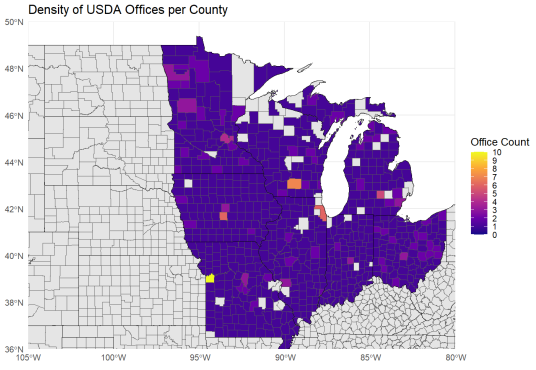


Figure 9: Locations of USDA offices in 2025 (based on 1990s county boundaries and 2025 USDA office locations)

2- Density of weather observation stations. I exploit spatial variation in the historical density of weather observation stations[6]. The underlying economic mechanism is that a denser local observation network improves the accuracy of area, as level weather measures used in insurance rating and indemnity determination, thereby reducing basis risk and measurement error in the insurance contract. Lower basis risk increases the expected accuracy of indemnity payments from the farmer's perspective, making crop insurance more attractive and increasing insurance take up. Importantly, this channel operates through the quality of the insurance contract and the information used in risk assessment, rather than through any direct effect on agronomic decisions or crop yields themselves. Figure 12 shows the number of weather stations in the study area in the base year of 1989.

[6]The idea of using this IV was a result of attending a paper presentation by Dr Vaibhav Anand at the ASSA 2026. I thank Dr Vaibhav Anand for the valuable discussion about the range of available data resources on weather stations in the US and beyond.



(a) Density of USDA offices



(b) USDA offices and their locations relative to roads

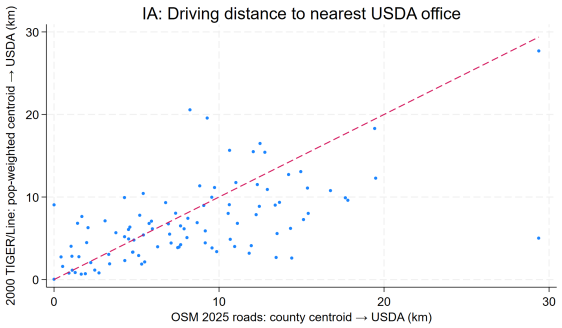
Figure 10: The condition of access to USDA offices in the Midwest. Base map is from OpenStreetMap contributors (2025), county boundaries based on the 2000 census (Bureau 2000) and USDA office locations if from USDA Farm Service Agency (2025).

3- Bartik Instrument The Federal Crop Insurance Program's statutory premium-subsidy rate $s_k(t)$ is set by Congress for each coverage-level tier $k \in \{50, 55, \ldots, 85\}\%$ and each year t. As a right-hand-side variable it has two attractive features: it is exogenous (federal law, not chosen by farmers) and it changed sharply in identifiable steps: the 1994 Federal Crop Insurance Reform Act, the 2000 Agricultural Risk Protection Act (ARPA), and subsequent Farm-Bill adjustments.

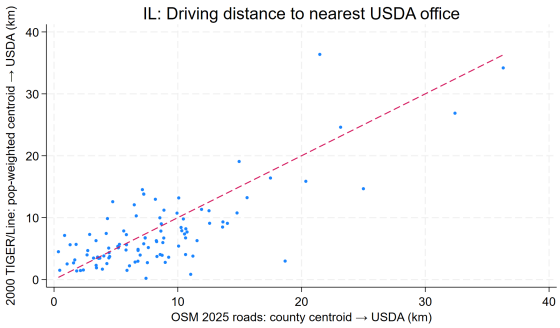
I construct a Bartik-style instrument (Bartik 1991; Goldsmith-Pinkham, Sorkin, and Swift 2020) that combines national changes in the statutory crop-insurance subsidy schedule with predetermined local exposure to different coverage levels. Let $\ell \in 50, 55, \ldots, 85$ index buy-up coverage levels, and let $s_{\ell t}$ denote the national statutory subsidy rate for coverage level ℓ in year t. For each county-crop pair (c, j), I compute pre-period exposure weights $w_{cj\ell,0}$, defined as the share of buy-up insured acres at coverage level ℓ in a baseline period before the main policy changes. The instrument is then

$$Z_{cjt} = \sum_{\ell} w_{cj\ell,0} s_{\ell t}. \tag{7}$$

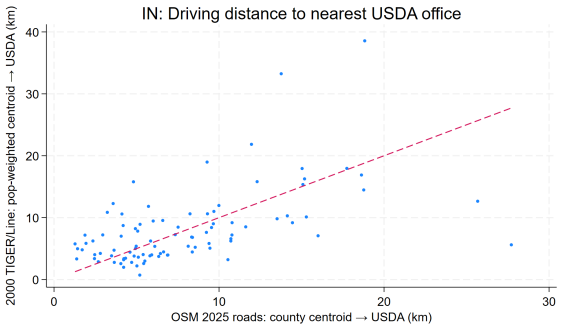
This variable measures the predicted statutory subsidy generosity facing a county-



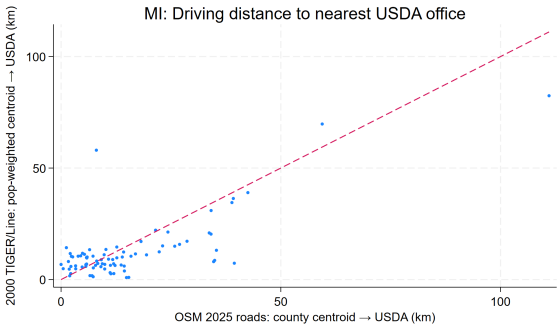
(a) Iowa



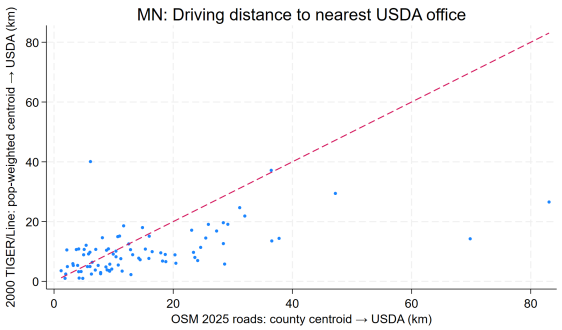
(b) Illinois



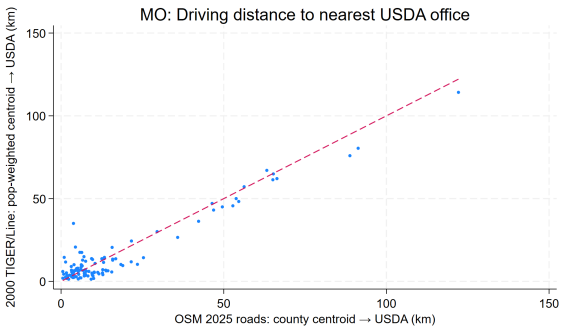
(c) Indiana



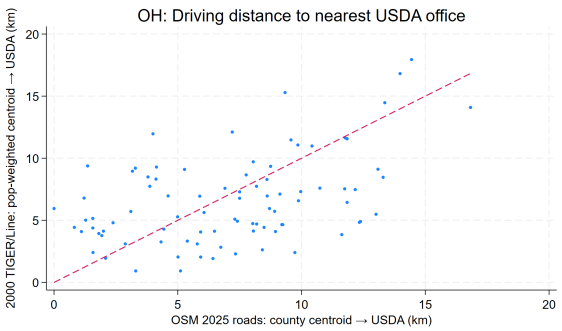
(d) Michigan



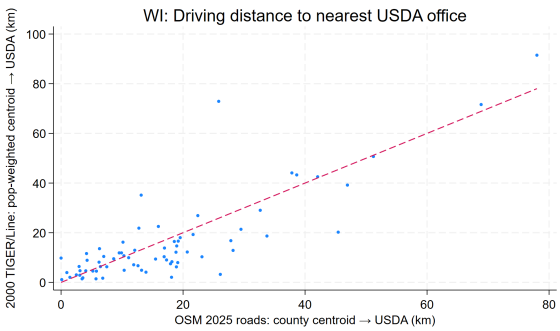
(e) Minnesota



(f) Missouri



(g) Ohio



(h) Wisconsin

Figure 11: Correlation between driving distance from county centroids to the nearest USDA office using the 2025 OpenStreetMap road network and from population-weighted county centroids using the 2000 Census TIGER/Line road network, for eight Midwest states.

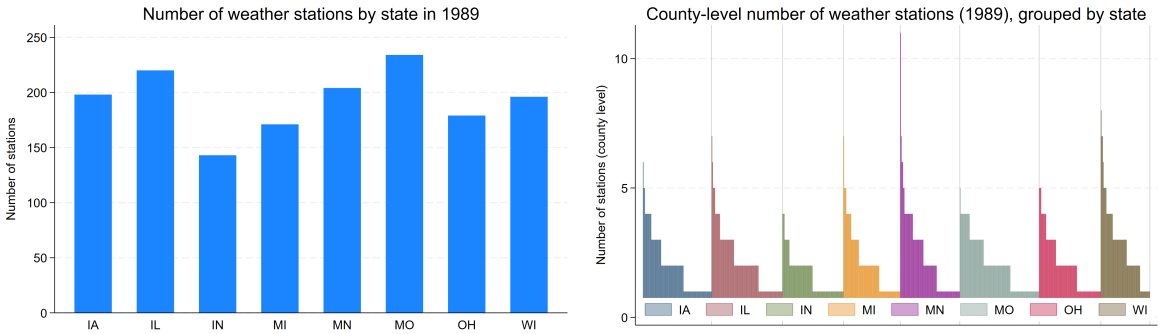


Figure 12: Number of weather stations by state in 1989 (authors' calculation using NOAA National Centers for Environmental Information (2025b) and NOAA National Centers for Environmental Information (2025a)).

crop pair, holding fixed its pre-period coverage-level composition. The identifying variation comes from the interaction of national statutory subsidy changes with predetermined local exposure to different coverage levels. Because the national subsidy schedule varies over time while baseline coverage-level shares vary across county-crop pairs, Z_{cjt} remains identified in specifications with county-crop and year fixed effects. The exclusion restriction is that, conditional on fixed effects and controls, baseline exposure to national statutory subsidy changes affects outcomes only through its effect on crop-insurance take-up or insurance intensity. This design follows the logic of shift-share or Bartik instruments, which combine aggregate shocks with predetermined local exposure shares. I create three baseline variants.

Table 3: Bartik subsidy instruments: baseline windows and coverage categories. Each variant fixes county-crop exposure shares over a different pre-period window, and the window determines the years for which those shares are predetermined. The strict variant draws only on pre-1995 data and is the construction used wherever a Bartik instrument is reported; the other two relax the window and enter as robustness checks.

Variable	Window	Coverage cat.	Predetermined for	Used
`subsidyBartik`	1995-2000	A (Buy-up)	Post-ARPA (2001+)	robustness check
`subsidyBartikPre`	1989-1995	E pre-95, A in 95	Post-1995	robustness check
`subsidyBartikPreStrict`	1989-1994	E only	Post-1994	main specification

Pooled net reported acres, aggregated within (county, crop) over each baseline window and normalised across tiers.

The bolded cross-county SDs at the 65% and 75% tiers are the source of the identifying variation in each Bartik instrument. Post-Reform (1995-2000) baseline concentrates mass at 65% (mean 0.71) with only 15% at 75%. Both pre-Reform baselines have substantially more mass at 75% (mean 0.21-0.25) and higher cross-county SD at 75% (0.22-0.24 vs 0.15).

Table 4: Post-Reform baseline coverage-tier shares, 1995-2000

Tier	Mean share	Cross-county SD	10th pctile	90th pctile
Corn ($N = 672$ counties with positive baseline buy-up)				
65%	0.711	**0.213**	0.389	0.969
70%	0.107	0.067	0.003	0.191
75%	0.151	**0.149**	0.002	0.378
80%	0.015	0.025	0.000	0.042
85%	0.016	0.028	0.000	0.049
Soybeans ($N = 646$ counties)				
65%	0.683	**0.190**	0.415	0.922
70%	0.117	0.072	0.030	0.187
75%	0.168	**0.143**	0.013	0.371
80%	0.015	0.022	0.000	0.040
85%	0.018	0.030	0.000	0.052

Table 5: Loose pre-Reform baseline coverage-tier shares, 1989-1995

Tier	Mean share	Cross-county SD	10th pctile	90th pctile
Corn ($N = 670$ counties)				
50%	0.109	0.154	0.006	0.298
65%	0.679	**0.211**	0.361	0.923
75%	0.212	**0.228**	0.003	0.581
Soybeans ($N = 643$ counties)				
50%	0.068	0.109	0.005	0.160
65%	0.698	**0.210**	0.387	0.932
75%	0.234	**0.216**	0.012	0.572

Table 6: Strict pre-Reform baseline coverage-tier shares, 1989-1994

Tier	Mean share	Cross-county SD	10th pctile	90th pctile
Corn ($N = 652$ counties)				
50%	0.129	0.175	0.009	0.339
65%	0.648	**0.216**	0.326	0.895
75%	0.223	**0.236**	0.005	0.614
Soybeans ($N = 625$ counties)				
50%	0.085	0.131	0.007	0.209
65%	0.667	**0.216**	0.352	0.913
75%	0.248	**0.223**	0.014	0.588

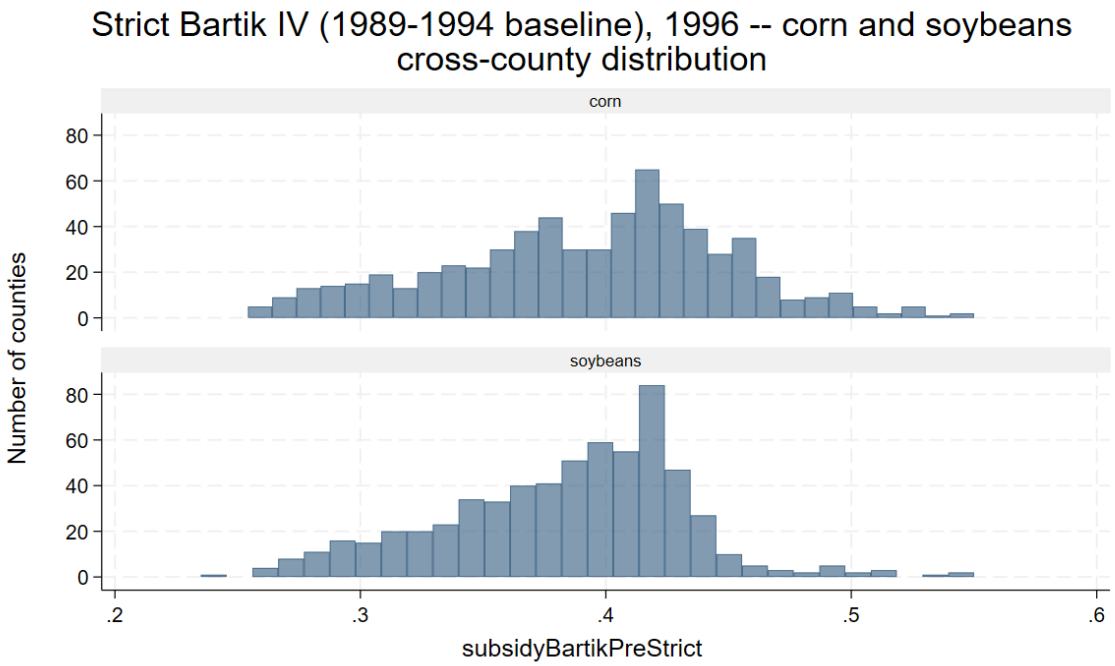


Figure 13: Cross-county distribution of the strict Bartik IV (`subsidyBartikPreStrict`) in 1996, by crop.

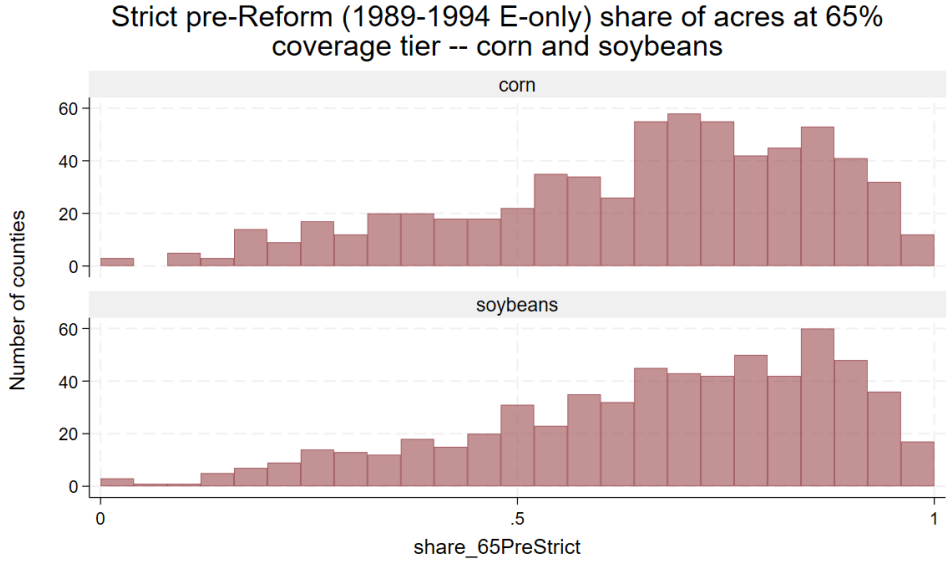
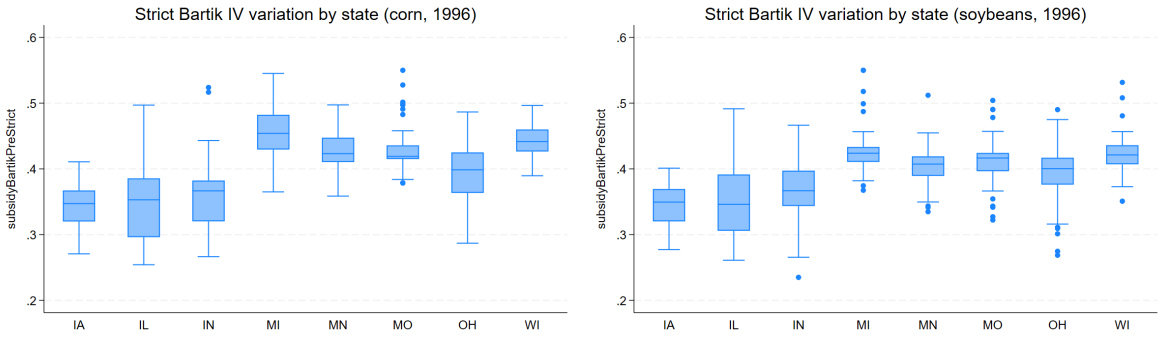


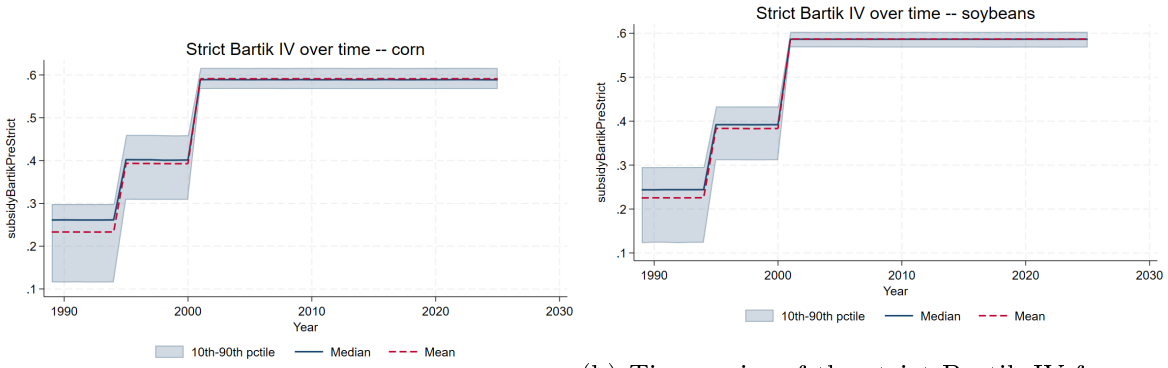
Figure 14: Cross-county distribution of the strict pre-Reform (1989-1994 E-only) share of acres at the 65% coverage tier, by crop.



(a) Strict Bartik IV in 1996, corn.

(b) Strict Bartik IV in 1996, soybeans.

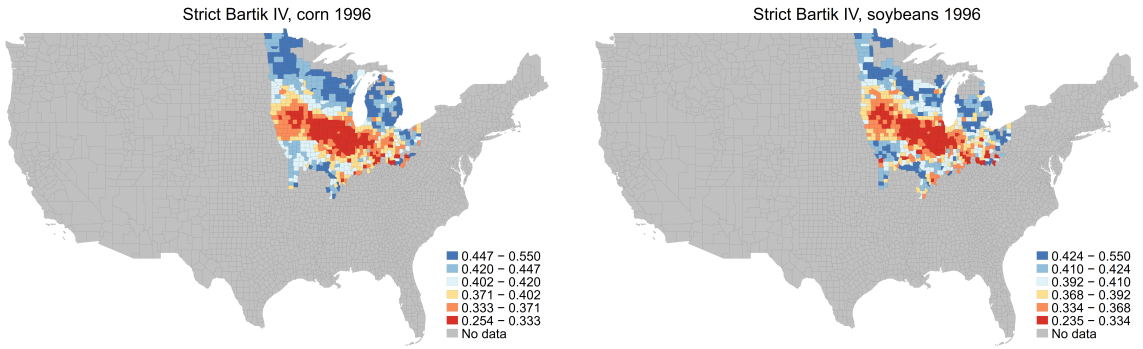
Figure 15: Cross-county distribution within each of the eight states. Baseline weights: 1989-1994 E-only. Each box = counties within state.



(a) Time series of the strict Bartik IV for corn.

(b) Time series of the strict Bartik IV for soybeans.

Figure 16: Strict Bartik IV over Time: yearly mean, median, and 10th-90th percentile band.



(a) Strict Bartik IV, corn 1996.

(b) Strict Bartik IV, soybeans 1996.

Figure 17: Strict Bartik IV map.

5.2 Covariate Balance

In this section, I check whether the instruments predict baseline covariates. I assess the strength of the instrumental variables by estimating first-stage regressions for both endogenous variables: (i) insurance participation, measured by the fraction of planted land that is insured, and (ii) insurance intensity, measured by subsidy per dollar of insured liability. Specifically, I estimate the following first-stage specification:

$$D_{ict}^{(k)} = \boldsymbol{\pi}_k' \mathbf{Z}_i + \mathbf{X}_{ict}' \boldsymbol{\beta}_k + \gamma_{s(i)} + \delta_t + \varepsilon_{ict}^{(k)}, \tag{8}$$

where $D_{ict}^{(k)} \in \{\text{fracInsured}_{ict},\ \ln(\text{fracInsured}_{ict}),\ \text{sublia}_{ict},\ \ln(\text{sublia}_{ict})\}$ denotes the endogenous insurance variable, and $\mathbf{Z}_i = (\text{drive_km_to_usda}_i,\ \text{stations_count1989}_i)$ is the vector of instrumental variables. The vector of controls $\mathbf{X}_{ict}$ includes predetermined weather and price variables:

$$\mathbf{X}_{ict} = (\text{ddayMod_100}_{ict}, \text{ddayHot_100}_{ict}, \text{prec_cm}_{ict}, \text{prec2_cm}_{ict}, \text{price_received}_{ict})\,.$$

State fixed effects $\gamma_{s(i)}$ absorb time-invariant institutional and geographic differences across states, while year fixed effects δ_t capture common macroeconomic and technological shocks. Identification therefore relies on within-state, cross-county variation over time. Standard errors are clustered at the county (FIPS) level, allowing for arbitrary correlation in regression residuals across crops and years within a county. Instrument relevance is assessed using joint first-stage tests of the null hypothesis that the coefficients on both instruments are equal to zero.

The station-count, station-density, and USDA-distance instruments show little evidence of systematic correlation with the climate covariates. The Bartik instrument is mostly balanced, but it shows a noticeable negative association with precipitation, suggesting that precipitation balance should be discussed and checked in robustness tests. Add explanation for the next table and figure.

Table 7: Covariate balance tests

	(1) MDD	(2) HDD	(3) Prec	(4) Prec2	(5) frac96	(6) liab96	(7) yld96	(8) plant96	(9) ins96	(10) lnPrice
Stations count 1996	0.200	-7.157***	-0.569	-164.831	-0.065*	1.621	3.819	12739.828	550.845	-0.031***
	(0.676)	(1.762)	(1.359)	(182.315)	(0.034)	(6.499)	(5.573)	(8164.501)	(4745.006)	(0.006)
Stations density 1996	1522.520	11124.937	22492.734**	2.28e+06*	176.645	-4.78e+04	-8.68e+04	-5.86e+07	-5.18e+06	43.434
	(5723.409)	(14275.907)	(8893.195)	(1.22e+06)	(336.892)	(45768.443)	(60175.729)	(3.88e+07)	(2.14e+07)	(46.577)
Distance USDA	0.137**	-0.208	-0.148	-15.629	-0.000	-0.117	0.036	-674.874	-216.313	-0.000
	(0.064)	(0.178)	(0.151)	(19.861)	(0.002)	(0.336)	(0.503)	(900.514)	(389.872)	(0.001)
Bartik (strict)	4.370**	-9.929**	-1.482	388.437	-0.395***	-118.406***	-73.507***	-27.632	4361.040	-0.083***
	(1.974)	(5.036)	(3.178)	(412.035)	(0.055)	(11.420)	(15.084)	(24996.508)	(13616.922)	(0.011)
Observations	33445	33445	33445	33445	33175	33175	33175	33175	33175	33445
R-squared	0.975	0.813	0.441	0.404	0.919	0.934	0.947	0.885	0.929	0.992
Joint F (IVs=0)	4.11	11.55	2.09	1.54	14.03	30.75	6.64	0.86	0.39	32.22
Joint p-value	0.003	0.000	0.080	0.189	0.000	0.000	0.000	0.487	0.812	0.000

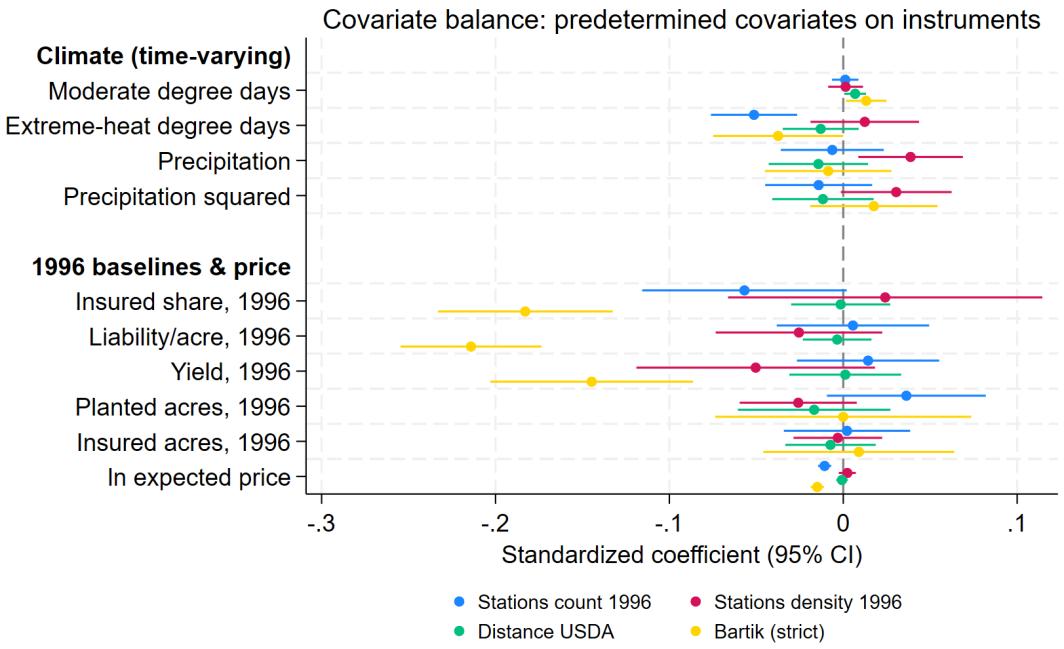


Figure 18: Covariate balance test

5.3 First Stage

In order to verify the relevance of two instruments, I did the first stage regression, with the endogenous variables on the left hand side and both instruments and the same control variables as the main regression model on the right hand side. Table 8 reports the results.

Table 8: First stage of IV regression

	(1) Buyup share	(2) Liability/acre	(3) ln(share)	(4) ln(liab)
Stations count 1996	-0.1333***	19.5128	-0.1463	-0.2804**
	(0.0372)	(22.4862)	(0.1568)	(0.1350)
Stations density 1996	32.4834	4.64e+05**	-878.3188	-2.32e+03**
	(297.6033)	(1.96e+05)	(1055.0837)	(992.2320)
Distance USDA	-0.0145***	6.4351***	-0.1470***	-0.1292***
	(0.0041)	(1.8353)	(0.0152)	(0.0131)
Bartik (strict)	-1.0751***	415.2338***	-6.0678***	-5.0489***
	(0.0977)	(59.7391)	(0.4051)	(0.3591)
Observations	33445	33445	33422	33422
R-squared	0.874	0.933	0.791	0.920
Joint F (IVs=0)	54.89	28.77	101.53	104.48
Joint p-value	0.000	0.000	0.000	0.000

5.4 Second Stage

The causal effect of insurance coverage on outcomes is estimated using two-stage least squares, instrumenting insurance participation and intensity with distance to USDA offices and historical weather-station density.

Let Y_{ict} denote the outcome of interest for county i, crop c, and year t. The second-stage equations are:

$$Y_{ict} = \beta_1 \widehat{\text{fracInsured}}_{ict} + \mathbf{X}'_{ict}\boldsymbol{\theta}_1 + \gamma_{s(i)} + \delta_t + u^{(1)}_{ict}, \tag{9}$$

$$Y_{ict} = \beta_2 \ln(\widehat{\text{fracInsured}})_{ict} + \mathbf{X}'_{ict}\boldsymbol{\theta}_2 + \gamma_{s(i)} + \delta_t + u^{(2)}_{ict}, \tag{10}$$

$$Y_{ict} = \beta_3 \widehat{\text{sublia}}_{ict} + \mathbf{X}'_{ict}\boldsymbol{\theta}_3 + \gamma_{s(i)} + \delta_t + u^{(3)}_{ict}, \tag{11}$$

$$Y_{ict} = \beta_4 \ln(\widehat{\text{sublia}})_{ict} + \mathbf{X}'_{ict}\boldsymbol{\theta}_4 + \gamma_{s(i)} + \delta_t + u^{(4)}_{ict}. \tag{12}$$

The fitted values $\widehat{\text{fracInsured}}_{ict}$, $\ln(\widehat{\text{fracInsured}})_{ict}$, $\widehat{\text{sublia}}_{ict}$, and $\ln(\widehat{\text{sublia}})_{ict}$ are obtained from the corresponding first-stage equations (??)–(??), which instrument insurance outcomes using $\text{drive_km_to_usda}_i$ and $\text{stations_count1989}_i$. The control vector $\mathbf{X}_{ict}$ includes predetermined weather and price variables:

$$\mathbf{X}_{ict} = (\text{ddayMod_100}_{ict}, \text{ddayHot_100}_{ict}, \text{prec_cm}_{ict}, \text{prec2_cm}_{ict}, \text{price_received}_{ict}) .$$

Table 11 reports the second stage of the two-stage least-squares regression under the *all-instrument* specification, which pools the three geographic instruments with the strict shift-share (Bartik) subsidy instrument. The dependent variable throughout is the by-crop winsorized log yield, and each column instruments a different measure of enrollment: the buy-up insured share (col. 1) and its log (col. 3) capture the *extensive* margin, while liability per acre (col. 2) and its log (col. 4) capture the *intensive* margin. All specifications absorb county, crop, and year fixed effects and cluster standard errors on county. On conventional diagnostics the specification looks healthy: the instruments are strong. The Kleibergen-Paap first-stage F ranges from 28.8 to 104.5, well above the usual weak-instrument thresholds and the Hansen J test never rejects the over-identifying restrictions (p between 0.19 and 0.34), so the four instruments appear mutually consistent. The point estimates give the familiar pattern: the extensive margin is positive (a higher buy-up share raises yield; significant at the 10% level in col. 1), while the intensive margin is small and imprecise.

These numbers should nonetheless not be read as the preferred estimate. The strong first stage is driven largely by the Bartik instrument, which on its own has an F above

Table 9: Second stage of the IV (2SLS) regression, geographic instruments only (preferred specification). The dependent variable is the by-crop winsorized *log* yield. Each column is a separate 2SLS in which "Insurance treatment" is a different endogenous regressor. (1) buy-up insured share, (2) liability per acre, (3) log share, (4) log liability-instrumented by the three geographic instruments (driving distance to the nearest USDA office, the 1996 weather-station count, and the 1996 station density per planted acre, each × the base rate). These are the instruments that pass the covariate-balance test, so these columns are the paper's preferred causal estimates: the extensive margin (cols. 1 and 3) is positive and significant, while the intensive margin is small and sign-unstable. All models absorb county, crop, and year fixed effects and cluster standard errors on county; the Kleibergen-Paap first-stage F and the Hansen J p-value (which never rejects) are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) share	(2) liab/acre	(3) ln(share)	(4) ln(liab)
Insurance treatment	0.559***	-0.001***	0.077**	0.082***
	(0.166)	(0.001)	(0.032)	(0.030)
Observations	34189	34189	34160	34160
First-stage KP F	25.22	16.27	66.59	79.50
Hansen J p-value	0.312	0.408	0.152	0.214

Table 10: Second stage of the IV (2SLS) regression, strict Bartik instrument only. The dependent variable is the by-crop winsorized *log* yield; columns instrument the same four enrollment measures as in Table 9, now with the single shift-share (Bartik) subsidy instrument. Because the model is just-identified (one instrument, one endogenous regressor), no Hansen J over-identification test is available. The Bartik is mechanically by far the strongest instrument-the Kleibergen-Paap first-stage F reaches 160-252, yet it delivers a precisely-estimated *zero* on every margin. Together with its failure of the covariate-balance test (it is correlated with baseline yield and insurance), this illustrates the paper's methodological point: first-stage strength is no guarantee of validity, and an invalid instrument yields a confident but spurious answer. All models absorb county, crop, and year fixed effects and cluster standard errors on county. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) share	(2) liab/acre	(3) ln(share)	(4) ln(liab)
Insurance treatment	-0.009	0.000	-0.004	-0.005
	(0.123)	(0.000)	(0.022)	(0.025)
Observations	34793	34793	34769	34769
First-stage KP F	160.12	51.06	251.90	229.01
Hansen J p-value	.	.	.	.

150; and although it passes the Hansen test when pooled here, it *fails* the covariate-balance test. It is significantly correlated with baseline yield and insurance (Figure **??**). Because the Bartik carries a spurious, near-zero yield relationship and receives heavy weight in the pooled first stage, it attenuates the extensive-margin coefficient from the geographic-only estimate of 0.559 (Table **??**) down to 0.175. This specification is therefore reported for comparison: it shows concretely that a mechanically powerful instrument can pass both the first-stage F *and* the over-identification test yet still bias the estimate when it violates covariate balance. The causal magnitude the paper relies on comes from the geographic instruments alone.

Table 11: Second stage of the IV (2SLS) regression, all-instrument specification. The dependent variable throughout is the by-crop winsorized *log* yield; logs are used so the coefficients are scale-free (semi-)elasticities comparable across corn and soybeans, whose bushel levels differ by roughly 3×. Each column is a separate 2SLS in which the row "Insurance treatment" is a different endogenous regressor. (1) buy-up insured share, (2) liability per acre, (3) log share, (4) log liability, instrumented by all four instruments (three geographic plus the strict Bartik). All models absorb county, crop, and year fixed effects and cluster standard errors on county; the Kleibergen-Paap first-stage F and the Hansen J p-value are reported. Because the strict Bartik fails covariate balance, the preferred estimates use the geographic instruments only; this pooled-instrument specification is reported for comparison. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(1) share	(2) liab/acre	(3) ln(share)	(4) ln(liab)
Insurance treatment	0.175*	-0.000	0.026	0.034*
	(0.097)	(0.000)	(0.017)	(0.020)
Observations	33445	33445	33422	33422
First-stage KP F	54.89	28.77	101.53	104.48
Hansen J p-value	0.269	0.343	0.187	0.221

Under the valid geographic instruments, the extensive margin enters positively and significantly at every functional form: a 0.1 increase in the insured share raises yield by about 6%, and the log-log elasticity is 0.077 (Table 9); the outcome throughout is the by-crop winsorized log yield, so these estimates are not driven by extreme crop-failure or bumper years. This is a local average treatment effect for counties whose enrollment responds to geographic access, and its *positive* sign is the opposite of the moral-hazard prediction. The most natural interpretation is a *credit/liquidity* channel: a crop-insurance contract stabilizes income and can serve as collateral, relaxing financing constraints. The *intensive* margin (liability per acre) is negative but small and, as the level-versus-log comparison shows, not sign-stable; I return to it below and treat it cautiously. In the welfare analysis of Section 9, this positive yield response is the

credit-externality term (β_Y) that turns the optimal subsidy from zero into a positive number, which is the first place the positive estimates feed the normative calculation.

Then I ask how much confidence we should have in the estimated causal effect if the instruments are weak. To answer to this question, I use the Anderson-Rubin (AR) test, which is one of the classic methods for inference that remains valid even when instruments are weak (so called *test inversion*[7]). To make it conceptually clear, suppose the true effect of increasing the buy-up insured share on log yields were 0.2. The AR test poses that If I remove that hypothesized effect from yields, do the instruments still predict the remaining variation? If they do, then 0.2 cannot be the true causal effect. If they do not, then 0.2 is consistent with the data.

[7] The 2018 NBER Summer School Lecture of Isaiah Andrews (Andrews 2018) is a great reference.

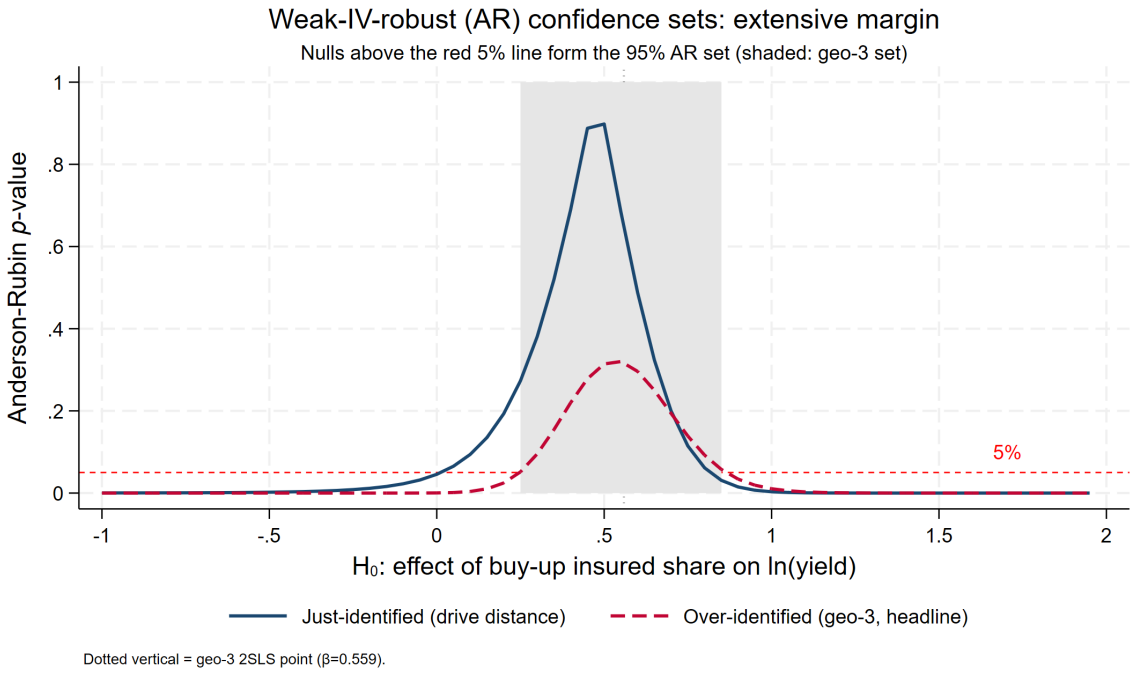


Figure 19: Anderson-Rubin (weak-instrument-robust) confidence sets for the extensive-margin effect of the buy-up insured share on log yield. Each curve traces the AR p-value for the null $H_0\colon \beta = \beta_0$ across candidate values β_0; the nulls *not* rejected at the 5% level (above the red dashed line) form the 95% AR confidence set, which remains valid even when instruments are weak. The solid line uses the single driving-distance instrument (just-identified, first-stage $F = 24.7$; AR set [0.05, 0.80]); the dashed line uses all three geographic instruments-the headline specification (AR set [0.25, 0.85], shaded). Both sets exclude zero and bracket the 2SLS point estimate of 0.559 (dotted vertical line): the positive effect of enrollment on yield is not an artifact of moderate instrument strength. The over-identified set is tighter but also reflects any disagreement among the three instruments; the just-identified set isolates pure weak-instrument uncertainty. Estimates control for climate and expected price, absorb county, crop, and year fixed effects, and cluster standard errors on county.

6 Mechanism

If the mechanism is credit, as discussed in 5.4, then enrollment should move farmers' financing and input decisions. In Table 12, I regress the USD expenditure on fertilier, net value of received CCC loans and net cash income on the buyup share, as instrumented before with the geographical IVs. The evidence is partial but pointed: enrollment sharply raises net ccc loan borrowing, consistent with insurance relaxing credit constraints.

It also lowers fertilizer expenditure. The negative fertilizer response places this paper in a long literature on insurance and input use that has split on both sign and

interpretation. Horowitz and Lichtenberg (1993) find insured Midwestern corn producers apply *more* nitrogen and pesticides, reading chemicals as risk-increasing inputs whose use expands once losses are covered; Vincent H. Smith and Barry K. Goodwin (1996), correcting for the endogeneity of the insurance decision itself, find the opposite: insured wheat growers spend less on chemical inputs, and Bruce A. Babcock and Hennessy (1996) obtain the same prediction from an expected-utility model in which fertilizer and insurance are substitute ways of managing risk; Barry K. Goodwin, Vandeveer, and Deal (2004) estimate generally modest input and acreage responses in the Corn Belt. My instrumented estimate sides with the Smith-Goodwin branch: enrollment reduces fertilizer expenditure. Two readings fit the rest of the evidence. The first is *self-insurance substitution*: farmers routinely over-apply nitrogen as a hedge against shortfall in favorable weather "insurance fertilizer", and formal coverage substitutes for that precautionary excess. This is an ex-ante input adjustment, and thus moral hazard in the technical sense, but a benign one: Section 7 shows directly that it does not raise heat sensitivity, variance, or downside risk. If anything, variance *falls*, exactly what trimming a risk-increasing input predicts in the Just and Pope (1978) framework. The second reading ties the fertilizer column to the borrowing column: expenditure is price times quantity, and insurance-backed access to cheaper formal credit lets farms escape high-cost dealer financing and buy off-peak, lowering measured expenditure without necessarily lowering application. On either reading, the input margin responds as a rational, better-financed farmer would..

The negative coefficient on *net cash income* (column 3) deserves its own reading rather than being lumped with fertilizer, because three forces push it down without contradicting the credit channel. First, a mechanical one: buy-up coverage is paid for *every* year, while indemnities arrive only in loss years, so enrollment raises annual premium outlays and depresses measured net cash income in ordinary years even when the expected net transfer to the farm is positive. Second, a behavioral one that is in fact the credit channel at work: a farm that uses insurance to support CCC borrowing takes on debt service, and if the borrowing finances investment it converts current cash income into capital formation that raises output later. A short-run decline in *cash* income during a leverage-and-invest phase is what the financing story predicts, not evidence against it-the mirror image of the borrowing response in column 2. Third, a statistical one: this is the weakest column, with an imputed county-level outcome and the only over-identification rejection in the table ($p = 0.01$), so the instruments disagree here in a way they do not for borrowing. I therefore read the yield effect as consistent with a *financing/collateral* channel borrowing clearly responds, and the cash-income decline is consistent with debt service and investment, while stopping short of identifying the

specific input margin through which it raises output. Direct farm-level data on borrowing and input use is the natural way to close this gap. The complementary question is whether the rival interpretation-moral hazard, leaves any trace at all; the next section tests it on the moments where it should appear.

Table 12: The credit channel: financing and input outcomes on the buy-up share (2SLS, geographic instruments).

	(1) ln Fertilizer exp.	(2) ln CCC loans	(3) ln Net cash income
Buy-up insured share	-5.615***	5.950***	-1.866**
	(0.901)	(0.846)	(0.782)
Observations	21375	29434	30088
First-stage KP F	26.02	34.74	26.58
Hansen J p-value	0.116	0.450	0.010

7 The Competing Interpretation: Is It Moral Hazard?

In this section, I study the case for moral hazard of farmers. The idea is that a farmer protected against losses may invest less in defending her crop, skipping costly adaptation and accepting more exposure to bad weather. Such behavior may not appear in average levels of yield, but rather in how yields respond to extreme conditions. This section therefore tests for moral hazard on the higher moments by studying the sensitivity of yield at level, variance and downside losses to extreme heat. Throughout this section, the outcome is the raw log yield, and not the by-crop winsorized measure I used in Section 5.4.

Sensitivity to Extreme Heat I extend (6) with interactions of the margin and all four climate variables, instrumenting each $D \times C_k$ with $Z \times C_k$. Moral hazard predicts a negative coefficient on $D \times \text{HDD}$ (enrollment amplifies heat damage). Table 13 reports the interaction regressions for both margins and Figure 20 the marginal effect of the margin on yield across the range of extreme heat. For the extensive margin, the marginal effect at sample mean climate is 0.83 and *flat* in heat: the heat interaction is essentially zero ($\hat{\gamma}_{\text{HDD}} = -0.0005$, s.e. 0.0031), and the only significant interaction, precipitation squared, bends the profile trivially (left panel of Figure 20). Enrollment does not raise yield's heat sensitivity. For the intensive margin the average marginal effect is *negative* (-1.5×10^{-3} per dollar of liability), about -1.7×10^{-3} at low heat and shrinking toward zero at the top of the heat distribution (right panel).[8] Its heat interaction is positive

[8] "Low heat" means the 10th percentile of extreme-heat degree days in the estimation sample

and significant though tiny, meaning the negative effect *attenuates* as heat rises, again the opposite of the moral-hazard prediction that insurance amplifies heat damage.

Table 13: Second-stage IV with margin × climate interactions, outcome ln(yield). Left: extensive margin. Right: intensive margin. The climate variables are centered at their sample means inside the interactions, so the coefficient on each margin is its marginal effect at sample-mean climate.

	(1) ln(yield)
fracInsuredBuyup	0.8327***
	(0.2528)
D x MDD	-0.0039
	(0.0027)
D x HDD	-0.0005
	(0.0031)
D x Precip	0.0106
	(0.0097)
D x Precip2	-0.0001*
	(0.0001)
Observations	34244
First-stage KP F	8.54
Hansen J p-value	0.133

	(1) ln(yield)
liabBuyup	-1.47e-03**
	(5.97e-04)
D x MDD	-1.30e-05**
	(5.80e-06)
D x HDD	1.23e-05**
	(5.81e-06)
D x Precip	-9.20e-06
	(9.01e-06)
D x Precip2	4.11e-08
	(7.99e-08)
Observations	34244
First-stage KP F	8.72
Hansen J p-value	0.377

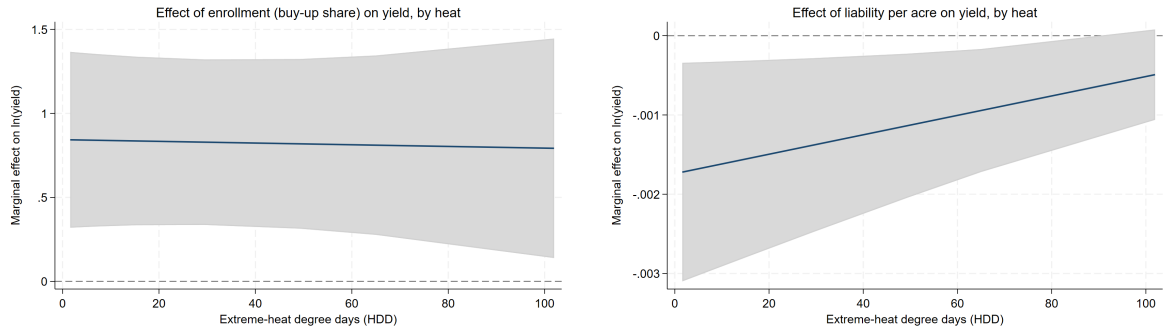


Figure 20: Marginal effect on ln(yield) across extreme-heat degree days (95% CI), other climate at means. Left: extensive margin (buy-up share); right: intensive margin (liability/acre).

Higher Moments of Crop Yield Moral hazard may instead raise the spread of yield. I construct the crop yield surprise by first subtracting the fitted insurance controbition from the log yield.

(HDD = 2.9, against a sample mean of 22.0). Because the interactions are centered, the marginal effect of liability at heat level HDD is $\hat{\theta}_D + \hat{\gamma}_{\text{HDD}}\,(\text{HDD} - \overline{\text{HDD}})$, with the other climate terms held at their sample means, where their centered interactions vanish. Plugging in the Table 13 estimates: $-1.47\times10^{-3}+1.23\times10^{-5}\times(2.9-22.0) \approx -1.7\times10^{-3}$ (s.e. 7.0×10^{-4}). The same formula traces the rest of the profile; at the 99th percentile (HDD = 102.5) it gives $-1.47\times10^{-3}+1.23\times10^{-5}\times(102.5-22.0) \approx -0.5\times10^{-3}$, the near-zero value at the right edge of Figure 20.

$$\tilde{y}_{ict} \;=\; \ln y_{ict} \;-\; \hat{\theta}_D\, D_{ict} \;-\; \sum_k \hat{\theta}_k \left(D_{ict} \times \tilde{C}_{k,ict}\right), \tag{13}$$

where $\tilde{C}_k$ are the mean-centered climate variables and the $\hat{\theta}$ are the 2SLS estimates from Table 13. Then I find the rrror term from the regression of the fitted values from Equation 13 on climate variables, price, and the fixed effects over the same estimation sample as the mean model,

$$u^*_{ict} \;=\; \tilde{y}_{ict} \;-\; \mathbf{X}'_{ict}\hat{\beta} \;-\; \hat{\alpha}_i \;-\; \hat{\phi}_c \;-\; \hat{\delta}_t. \tag{14}$$

I create the measures of dispersion, u^{*2} and downside deviation $\max(0, -u^*)$, from u^*. Table 14 reports the results and Appendix G collects the estimation details, including the bootstrap algorithm behind the bracketed standard errors. Table 14 shows that the higher moments results as captures by the coefficient of Coverage are not significant. The HDD interactions are negative in every column. Under the preferred bootstrap SE they are significant for the squared surprise on both margins, and for the share the implied difference in the marginal effect between high and low heat (the Diff. in marginal effect row of column (1)) is -0.20 (bootstrap s.e. 0.09). In other words, risk effects become more negative as heat rise; coverage lowers yield dispersion and downside risk under extreme heat rather than amplifying them.

Results of Tables 13 and 14 together show that enrollment in crop insurance raises mean yield, no margin becomes more exposed to heat, and there is no detectable level effect on the risk moments in either direction. For the welfare calculation what matters is the absence of the adverse response: the moral-hazard term that enters the Baily-Chetty optimal-subsidy formula as the behavioral cost is approximately zero, which is a central reason a substantial subsidy can be optimal (Section 9).

7.1 Reconciling the Null Results with the Prior Evidence of Moral Hazard

Given the earlier discussion in the 1 about the earlier literature on the impact of enrollment in crop insurance and crop yield, the question is how can this null coexist with Annan and Schlenker (2015), who find insured corn 67% and insured soybeans 43% more sensitive to extreme heat than their uninsured counterparts, or with Regmi, Tack, and Featherstone (2023), who estimate insured dryland corn roughly 38% more heat-sensitive? One major deviation I had from the prior work, has been using instru-

Table 14: Higher-moment moral-hazard tests. Each column is a separate 2SLS defined by an outcome and a margin: columns (1)-(2) use the squared yield surprise, whose conditional expectation is the conditional second moment of the surprise, and columns (3)-(4) the downside deviation $\max(0, -u^*)$, a lower partial moment; "share" columns treat the buy-up share as the coverage measure D, "liability" columns liability per acre. Each interacted climate term is centered, so the Coverage row is the average marginal effect of D over the sample climate distribution. The numbers in parentheses give analytic county-clustered standard errors and square brackets give county-cluster bootstrap standard errors from 999 replications of the entire two-step procedure. Stars are based on the coefficient divided by its bootstrap standard error, compared to standard-normal critical values.

	(1) Sq. surprise share	(2) Sq. surprise liability	(3) Downside dev. share	(4) Downside dev. liability
Coverage (D)	-0.098895	3.24e-04	0.067469	1.04e-05
	(0.154252)	(3.16e-04)	(0.159624)	(3.49e-04)
	[0.136554]	[2.57e-04]	[0.052233]	[1.07e-04]
D x MDD	0.003591*	6.19e-06*	0.000436	9.13e-07
	(0.002291)	(4.33e-06)	(0.001851)	(3.64e-06)
	[0.001963]	[3.23e-06]	[0.000571]	[7.72e-07]
D x HDD	-0.003193**	-5.26e-06**	-0.000521	-2.38e-07
	(0.001374)	(3.71e-06)	(0.001787)	(3.34e-06)
	[0.001509]	[2.68e-06]	[0.000525]	[7.51e-07]
D x Precip	-0.014554	-1.40e-05**	-0.002534	-6.61e-06*
	(0.012243)	(7.40e-06)	(0.006941)	(5.07e-06)
	[0.010021]	[6.96e-06]	[0.003175]	[3.40e-06]
D x Precip2	0.000096	8.38e-08*	0.000025	5.85e-08**
	(0.000077)	(4.19e-08)	(0.000049)	(4.66e-08)
	[0.000066]	[4.63e-08]	[0.000023]	[2.87e-08]
Post-estimation linear combination (not a regressor):				
Diff. in marginal effect (HDD p95 minus p10)	-0.198293**	-3.26e-04**	-0.032335	-1.48e-05
	(0.085346)	(2.31e-04)	(0.110992)	(2.08e-04)
	[0.091303]	[1.66e-04]	[0.032013]	[4.63e-05]
Observations	34244	34244	34244	34244
First-stage KP F	8.54	8.72	8.54	8.72
Min. SW conditional F	15.99	16.72	15.99	16.72
Hansen J p-value	0.517	0.760	0.054	0.421
Successful bootstrap reps	999	999	999	999

mental variable framework and taking into account the endogeneity of the financial and farming behavior of farmers. By instrumenting enrollment with predetermined access costs of distance to USDA offices and weather-station density, interacted with the insurance base rate, that shift the cost of participating but are unrelated to a farm's heat vulnerability.

The second point is that the sample period and contract mix have changed. My panel data set covers years 1996-2025. During this period, the program's composition shifted toward revenue-protection policies, enterprise units, and higher coverage levels following ARPA (2000) and subsequent Farm Bills. If contract design has become less distortionary-revenue policies with harvest-price provisions change the incentive to abandon a damaged crop-then a null today is compatible with genuine moral hazard under the earlier, yield-based contract regime. The extended window also adds the record-heat seasons of recent years, so the null is not for lack of identifying heat variation.

8 The Peer Effects

This section studies whether enrollment in the Federal Crop Insurance Program (FCIP) exhibits peer effects and whether such spatial interactions matter for agricultural yields. When decisions are complex and information is costly to acquire, imitation can be a rational response; Conlisk (1980)read April 9, 2026 formalizes this logic in a model of costly optimization. In agriculture, social learning and knowledge spillovers have been widely documented for the adoption of production technologies and management practices. Examples are Foster and Rosenzweig (1995) and Munshi (2004) for the spillover effects of adopting high-yielding seed varieties, Zhou et al. (2020) studied it for application of pesticides, and Adegbola and Gardebroek (2007) for seed storage technologies. By contrast, much less is known about whether participation in crop insurance diffuses through neighboring farmers and communities. Much of the empirical crop insurance literature implicitly treats farmers' insurance decisions as independent across space. If enrollment decisions are instead spatially correlated through learning, imitation, or common local networks, standard non-spatial models risk conflating own effects with peer effects.

Spatial econometric methods are therefore useful for studying crop insurance enrollment and its relationship with yields. Farmers operate in geographically connected environments in which information, beliefs, and production conditions spill across county boundaries. Insurance participation may diffuse through peer effects, extension net-

works, and shared perceptions of risk. At the same time, yields themselves are shaped by spatially correlated processes, including weather, soil conditions, pest pressure, and local agronomic norms. As a result, both the treatment variable (insurance uptake) and the outcome variable (yield) may display substantive spatial dependence. Ignoring this structure can produce misleading estimates of the relationship between insurance and yields, because neighboring counties' outcomes and covariates may affect local outcomes directly. In addition, some spatial dependence may arise from non-behavioral sources, such as geographic aggregation and shared ecological gradients across nearby counties. These considerations motivate the use of spatial regression models to distinguish direct effects from spillover effects and to account for spatial dependence in the data.

Figures 21 and 22 provide descriptive evidence consistent with this concern. For both corn and soybeans, the county-level choropleth maps reveal clear spatial clustering in yields, insurance enrollment, and liability per acre. Yield levels are not randomly distributed across space: counties with relatively high yields tend to be located near other high-yield counties, while lower-yield areas also appear in geographically contiguous clusters. The same is true for insurance variables. The fraction of planted acreage enrolled in buy-up insurance and liability per acre both exhibit strong regional concentration rather than isolated county-level variation. The maps also show substantial temporal change. Between 1989 and 2010, average yields generally increase, while liability per acre rises sharply and becomes more spatially concentrated. Insurance take-up is very limited in the earliest year and becomes much more pronounced in later years, again in clustered geographic patterns. These visual patterns suggest that insurance participation and production outcomes evolve jointly across neighboring counties, reinforcing the case for models that allow for spatial interaction.

Ignoring spatial lags in the crop insurance-yield relationship amounts to omitting relevant neighboring determinants of farm outcomes. Farmers' yields and insurance decisions are spatially correlated because nearby counties face related weather shocks, similar agroecological conditions, and overlapping information environments. If a model omits neighboring counties' insurance uptake or other neighboring covariates even though they affect local yields, the resulting omitted-variable problem can bias ordinary least squares estimates. More generally, failure to account for spatial dependence can distort coefficient magnitudes, standard errors, and statistical significance, leading to misleading conclusions about whether insurance improves or distorts production outcomes. Even in cases where the problem takes the form of spatial error correlation rather than omitted spatial lags, ignoring spatial dependence undermines inference by making the estimator inefficient and standard errors unreliable.

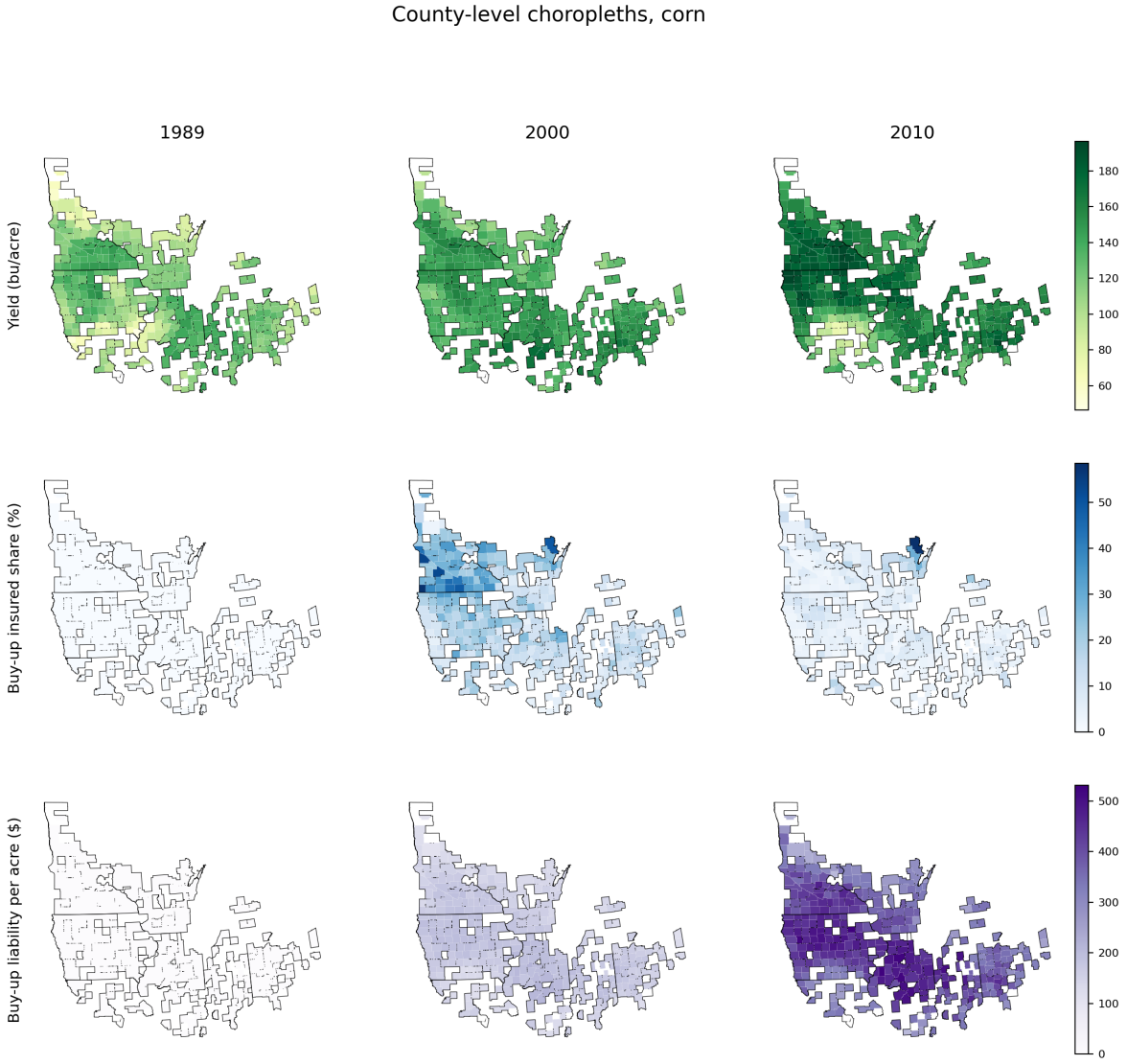


Figure 21: County-level choropleth maps for corn in 1989, 2000, and 2010. The maps display the spatial distribution of yield, the share of planted acreage enrolled in buy-up insurance, and buy-up liability per acre across the study region. The figure shows substantial geographic clustering in all three variables and notable changes over time, particularly the increase in liability per acre and the broader diffusion of buy-up insurance after 1989. *Source:* Author's calculations.

Weight matrix. My baseline specification uses a row-standardized Queen-contiguity matrix defined over county-crop units in the study region. Let W denote an $N \times N$ spatial weights matrix such that $w_{ij} = 1$ if two county-crop units are associated with counties that share a common border or vertex, and $w_{ij} = 0$ otherwise. I row-

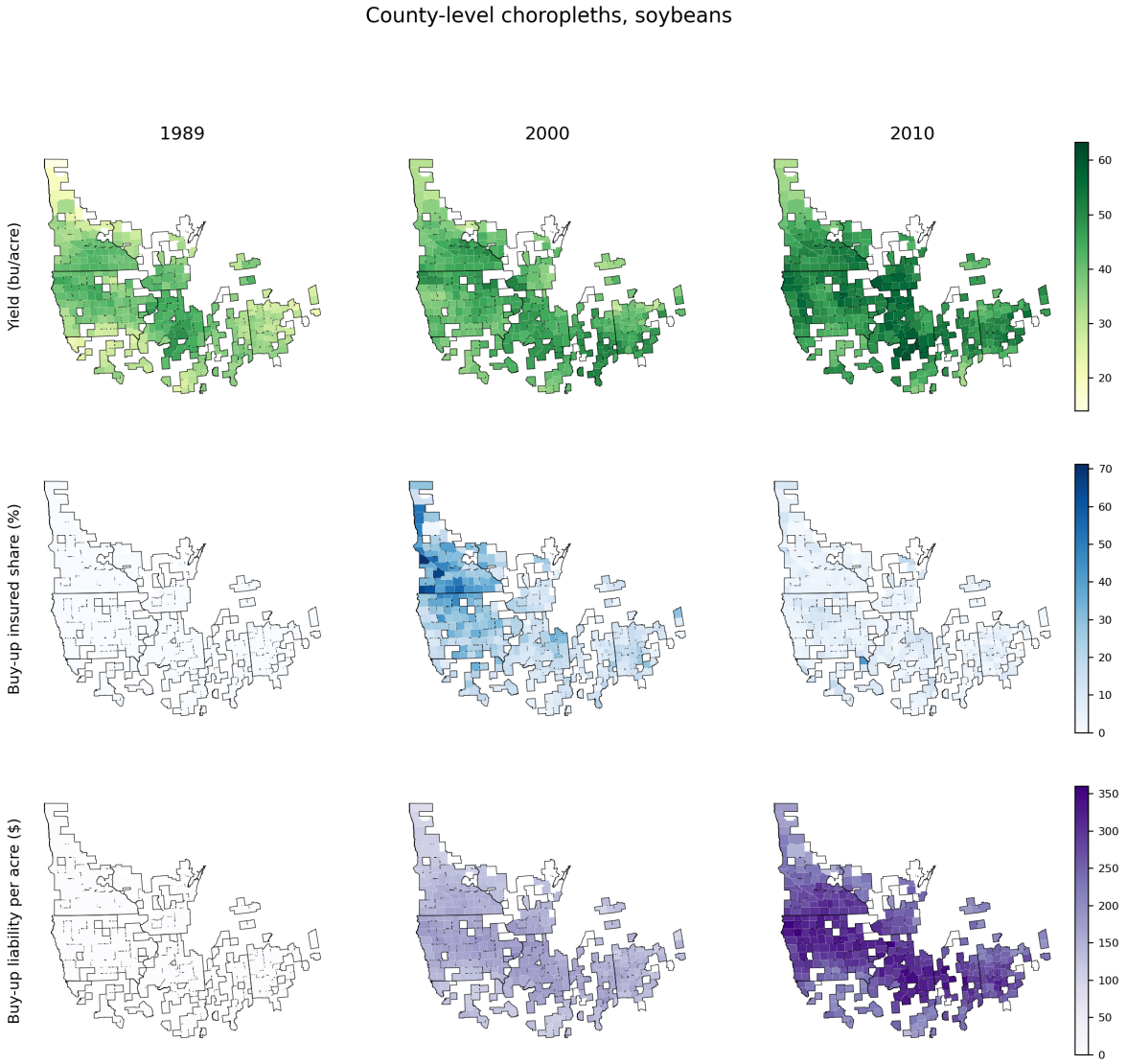


Figure 22: County-level choropleth maps for soybeans in 1989, 2000, and 2010. The maps display the spatial distribution of yield, the share of planted acreage enrolled in buy-up insurance, and buy-up liability per acre across the study region. As with corn, the figure reveals clear geographic clustering and substantial temporal change, indicating that both production outcomes and insurance variables are spatially patterned rather than randomly distributed across counties. *Source:* Author's calculations.

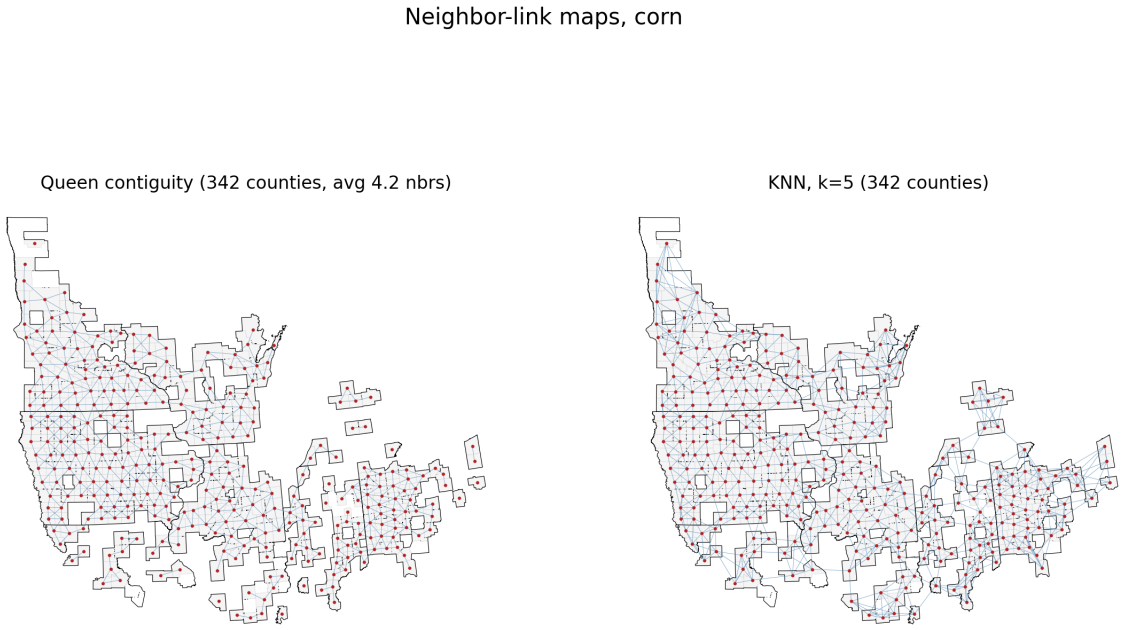
standardize the matrix so that each row sums to one:

$$\tilde{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}. \tag{15}$$

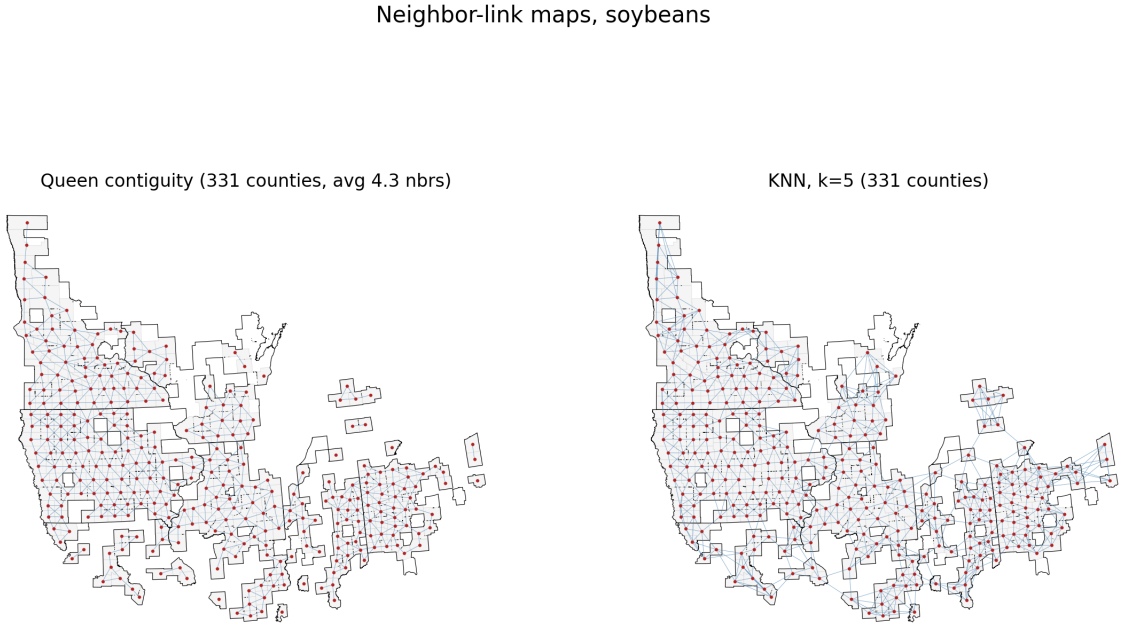
This makes each spatial lag interpretable as a weighted average of neighboring values. Because each county has only a limited number of contiguous neighbors, the matrix is sparse. As a robustness check, I also estimate analogous models using a row-standardized k-nearest-neighbors (KNN) weights matrix, in which each county-crop unit is linked to its k geographically closest neighbors regardless of administrative borders. Unlike Queen contiguity, the KNN matrix imposes a fixed number of neighbors for each unit and ensures that spatial interactions are defined even for counties with relatively few contiguous neighbors. Comparing results across the Queen and KNN specifications helps assess whether the estimated spillovers are driven by the particular notion of neighborhood imposed by contiguity, or whether they are robust to an alternative distance-based definition of local spatial interaction. Figure 23 illustrates the structure of the spatial links underlying the weights matrices for corn and soybeans. In each crop-specific panel, the left map displays the Queen-contiguity network, where counties are connected only when they share a border or vertex, while the right map displays the k-nearest-neighbors network with $k = 5$, where each county is connected to its five closest counties based on centroid distance. The figure makes clear that the Queen matrix reflects administrative adjacency and therefore produces an irregular number of neighbors across counties, whereas the KNN matrix generates a more uniform connection structure by imposing the same number of neighbors for each unit.

Appendix D reports results for the KNN sensitivity analysis for the spatial panel regressions. My analysis shows that for fraction of planted land that is insured, the k minimizing the average focal-coefficient standard error across the six yield models is $k = 9$, while the k minimizing the combined precision-stability rank score is $k = 13$. Also for the insurance liability per acre, the k minimizing the average focal-coefficient standard error across the six yield models is $k = 11$, while the k minimising the combined precision-stability rank score is $k = 2$. Given these results, I proceeded with k minimizing the average focal-coefficient standard errors, and modeled the spatial effects with $k = 9$ for percent of planted land that is insured, and $k = 11$ for insurance liability per acre.

I use the results of this section to pin down the social multiplier for calculating the optimal-subsidy. If enrollment diffuses across neighbors, a subsidy that raises one county's take-up also raises its neighbors', amplifying both the participation response



(a) Corn



(b) Soybeans

Figure 23: County-level neighbor-link maps for corn and soybeans. Maps show the spatial links used to construct the weights matrix on unique county polygons. In each panel, the left map uses the Queen contiguity matrix, while the right map uses k-nearest neighbors with $k = 5$, based on county centroids. Line segments connect each county centroid to its neighbors. *Source:* Author's calculations.

and the credit benefit, and hence the optimal subsidy. In this section, I documents the magnitude of enrollment diffusion across neighboring counties, without any success in identifying causality. I use a $k = 9$ nearest-neighbor weight matrix and instrument neighbors' enrollment WD with neighbors' and neighbors-of-neighbors' geographic instruments (Bramoullé, Djebbari, and Fortin 2009). I only focus on the extensive margin, the buy-up share, rather than liability per acre, as peer influence in insurance is an adoption mechanism, with neighbors learning from one another whether coverage is worth having rather than how many dollars of liability to place on an enrolled acre (Cai, Janvry, and Sadoulet 2015). The fact that the welfare calculation of Section 9 moves enrollment only through the insured share $q(s)$, makes the extensive margin the only channel through which a social multiplier be considered.

Since the spatial weights require a rectangular panel, which retains 673 of the 1,332 county-crop units, and that half is not a random one: retained units average 19% higher yield, a 17% higher buy-up share, and 28% fewer extreme-heat degree days than those dropped, and the retained set tilts toward the Corn Belt core (Iowa and Minnesota keep 84% and 71% of their units, Michigan and Missouri 16% and 20%), so the estimates below describe a cooler, better-insured core of the region rather than the full sample.

$$D_{ict} = \gamma\,(WD)_{ict} + \mathbf{Z}'_{ict}\beta + \mathbf{X}'_{ict}\theta + (W\mathbf{X})'_{ict}\psi + \alpha_i + \phi_c + \delta_t + \varepsilon_{ict}. \tag{16}$$

The object of interest is the endogenous peer effect γ on neighbors' enrollment (WD), instrumented by $W\mathbf{Z}$ and $W^2\mathbf{Z}$; $\mathbf{Z}$ is the county's own geographic access, an exogenous own effect; and ($W\mathbf{X}$) are contextual effects of neighbors' climate. Correlated effects, the shared weather, soil, and price shocks that make neighbors behave alike for reasons unrelated to diffusion, are absorbed by $\mathbf{X}$, $W\mathbf{X}$, and the fixed effects. The design's substantive restriction is thus that $W\mathbf{Z}$ is used as an instrument rather than entered as a regressor, which imposes that neighbors' access shifts a county's enrollment only through neighbors' enrollment. That restriction is what the over-identification test rejects below, so the failure is informative about the exclusion restriction itself and not only about instrument strength.

It turns out that enrollment is very strongly spatially clustered: the raw correlation between own and neighbors' buy-up share is 0.91, and the naive OLS peer coefficient is 0.92 (Table 15), consistent with strong peer diffusion, though neither statistic separates diffusion from *correlated effects* such as shared weather, soil, and prices.

The causal peer effect is not identified in the setup of my paper. On the balanced spatial subsample the instruments weaken sharply (own-enrollment KP $F = 3.2$ versus ≈ 25 in the full sample). The spatial-IV peer estimate is $\hat{\gamma} = 0.80$, but it rests on a weak

Table 15: Peer effects in FCIP enrollment (balanced spatial subsample, KNN $k = 9$).

	(1) Yield: IV (baseline)	(2) Enroll: OLS	(3) Enroll: spatial-IV	(4) Yield: spatial-IV
Own enrollment (D)	0.317			0.158
	(0.495)			(0.156)
Peer enrollment (W·D)		0.920***	0.799***	0.041
		(0.027)	(0.153)	(0.199)
Observations	19091	19091	19091	19091
First-stage KP F	3.19		7.97	2.22
Hansen J p-value			0.025	0.000

first stage ($F = 8.0$, below the rule-of-thumb 10) and *fails* the over-identification test ($p = 0.03$), which points at the no-contextual-effect restriction that (16) imposes on $W\mathbf{Z}$; the joint yield model is weaker still ($F = 2.2$, over-identification $p < 0.001$). Neither delivers a credible causal magnitude. A dynamic local projection off Farm-Bill shocks, over a conventional five-year horizon fixed in advance, yield similar results: subsidy shocks are spatially near-common (own/neighbor correlation 0.96), collinearizing the own and peer channels (Figure 24). Extending it to ten years does not separate them, since the shock correlation is unchanged and the forward window costs roughly a quarter of the estimation sample by $h = 10$, so the longer projections are uninformative about the peer channel. The social multiplier is thus a real feature of the data but of unknown magnitude; identifying it, with an idiosyncratic shock or farmer-level network data (Cai, Janvry, and Sadoulet 2015), is exactly what a fully closed welfare calculation would need, and I flag it as the main missing ingredient rather than force a number. Section 9 bounds its influence instead: treated as a calibrated parameter over the range the diffusion literature supports, the multiplier moves the optimal subsidy by at most two percentage points.

The local-projection design. The dynamic exercise deserves its methodological statement. Local projections estimate the impulse response

$$\mathcal{R}_{s\to y}(h,\delta) \equiv \mathbb{E}[y_{t+h} \,|\, s_t = s_0 + \delta;\, \mathbf{x}_t] - \mathbb{E}[y_{t+h} \,|\, s_t = s_0;\, \mathbf{x}_t]\,, \qquad h = 0, 1, \ldots, H, \quad (17)$$

by a separate regression at each horizon h, a semiparametric method borrowed from the macroeconomics literature that imposes weak assumptions on the shape of the response, at some efficiency cost relative to models with more structure such as a VAR (Jordà 2023; Jordà and Taylor 2025). Here the dose δ is the Farm-Bill subsidy shock, the change in the county-crop effective subsidy rate, whose movements are driven by the

statutory schedule; because the shock is a legislated price change rather than a chosen quantity, I read the projections as policy-event responses. When the treatment is instead an endogenous choice, the LP-IV variant of Jordà, Schularick, and Taylor (2015) applies, and its exclusion requirement takes the dynamic form of lead-lag exogeneity, $\mathbb{E}(\nu_{t+h}\,\mathbf{z}_t) = 0$ at *every* horizon h rather than only on impact (Jordà 2023). Each projection conditions on the pre-shock outcome, the lag augmentation that guards inference when lagged outcomes sit among the regressors (Montiel Olea and Plagborg-Møller 2021). Two disciplines then shape the reading of Figure 24. First, impulse-response coefficients are correlated across horizons, so pointwise bands cannot support statements about the shape of the path, and joint tests across horizons are the appropriate summary of overall significance (Jordà 2009). Second, the spatial layer follows the spatial local projections of Brady (2011) and Gonne and Hubert (2020): spatial dependence is an omitted-variable problem, and entering neighbors' weighted shock alongside the own shock, a spatially lagged regressor in the SLX tradition of Florax and Folmer (1992) and Halleck Vega and Elhorst (2015), models the cross-county propagation that fixed effects alone would leave in the error term. It is exactly this joint own-plus-neighbors specification that the 0.96 correlation between the two shocks renders collinear, which is why only the own-shock path is read as informative.

On the economic channel behind the diffusion, peer effects in a purchase decision can operate through preferences, the keeping-up-with-the-Joneses consumption externality of Gali (1994), or through learning from peers (Moretti 2011). Both are plausible for insurance, and the designs above cannot separate them, for the same reasons they cannot pin down γ. One feature of the setting nonetheless tilts the interpretation toward the preference channel: the program is three decades old and buy-up coverage has long been the norm in these counties, so learning about whether coverage is worth having should largely have run its course, yet the clustering remains strong. If the preference reading is right, part of any peer-induced enrollment surplus is positional, which reinforces the caution in Section 9 that the multiplier band's surplus valuation is generous.

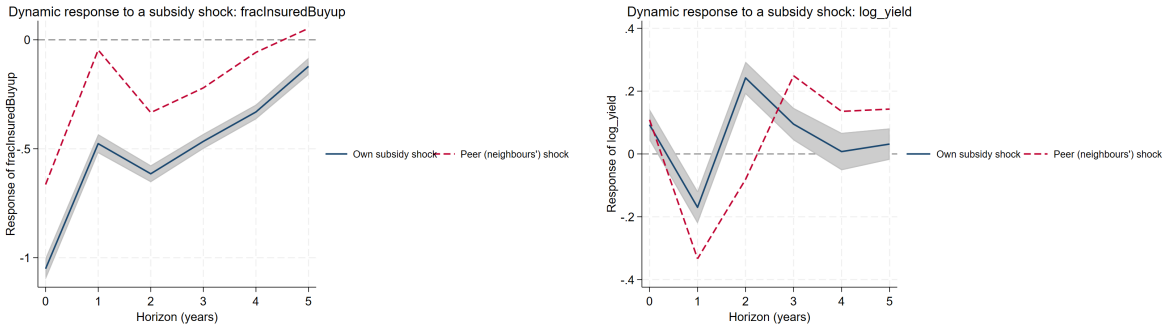


Figure 24: Dynamic response of own enrollment (left) and log yield (right) to subsidy shocks (KNN $k = 9$). The horizontal axis is the horizon h in years, and each point is a *separate* local projection of the outcome at $t + h$ on the shock at t, conditional on the pre-shock outcome, so the coefficient is the cumulative response relative to the year before the shock. The solid line is the county's *own* subsidy shock, estimated with the own shock alone, and the shaded band is its 95% confidence interval. The dashed line is the coefficient on *neighbors'* shocks from the joint specification containing both shocks, and is plotted without a band. Because own and neighbors' shocks correlate at 0.96, the two channels are not separately identified: the dashed path should be read as a diagnostic of that collinearity, not as an estimated peer effect, and only the own-shock path is informative.

9 The Optimal Subsidy: A Sufficient Statistics Approach

Having found no moral hazard and a positive credit-channel yield effect, I ask what these behavioral responses imply for the welfare-maximizing subsidy. I work in the sufficient-statistics tradition of social insurance (Baily 1978; Chetty and Finkelstein 2013), implementing it in the discrete-choice demand-surplus form: welfare is the log-sum consumer surplus of the take-up model (Small and Rosen 1981)[9], evaluated against the fiscal cost of public funds (Einav, Finkelstein, and Cullen 2010; Hendren and Sprung-Keyser 2020), in the crop-insurance application of Jialing Yu (2017, ch. 4) and Klosin and Solomon (2024). The utilitarian social planner chooses the buy-up premium subsidy rate s. Farmers choose between buy-up coverage and no buy-up for each planted acre. I estimate the log-odds of the logistic functional form by two-stage least

[9]Specifically their equation (5.9) (p. 127), the logit evaluation of the general compensating-variation formula in their equation (5.5): $(\Delta E)/N = -(1/\lambda)\big[\ln \sum_j \exp(W_j)\big]$, where the term inside the brackets is the log of the *sum* of exponentiated systematic utilities over the choice set (the *inclusive value*), the brackets denote evaluation at the initial and final points, and λ is the marginal utility of income. The welfare gain from a policy change is therefore $(1/\lambda)$ times the increase in the log-sum. In the binary application here the utility of forgoing buy-up coverage is normalized to zero, so the log-sum is $\ln(1 + e^V)$ with V the log-odds of take-up, and λ is the estimated premium coefficient $\hat{\beta}$.

squares. Welfare is measured relative to the observed baseline s_0 as

$$\Delta W(s) = \sum_{ict} A_{ict} \Big\{ \underbrace{[CS_{ict}(s) - CS_{ict}(s_0)]}_{\text{log-sum surplus}} - (1+\lambda) \underbrace{[C_{ict}(s) - C_{ict}(s_0)]}_{\text{subsidy outlay}} - \kappa \, [q_{ict}(s) - q_{ict}(s_0)] \Big\},$$

where the sum runs over county-crop-year cells and A_{ict} is planted acreage, so that the per-acre terms in braces aggregate to total welfare. CS is the extensive-margin log-sum consumer surplus per planted acre, $C = s\,P\,q$ the fiscal outlay per planted acre (P the gross premium per acre, q the buy-up share). λ is the marginal cost of public funds, and κ a per-insured-acre administrative cost. The objective is maximizing social welfare: a utilitarian planner with equal welfare weights sums farmers' participation surplus and the excess burden cost to taxpayers, so a subsidy dollar is an internal transfer that costs society the deadweight loss λ.

I simplify the model in three ways. First, I treat insurers as actuarially zero-profit agents and leave their administrative and operating subsidies and underwriting gains outside the calculation. Second, I exclude crop consumers from the welfare calculation: I value the output gain in Section 9.2 at each year's observed market price and hold that price invariant to the policy counterfactual. This assumption shuts down price pass-through. In reality, if the subsidy raised aggregate output enough to lower the prices, part of the gain would pass to consumers rather than remaining with farmers. Third, I do not distinguish landowners from operators: I do not consider any capitalization of the subsidy into land rents, so "farmers" includes landowners. Third, I do not distinguish landowners from operators: I do not trace any capitalization of the subsidy into land rents, so "farmers" includes landowners.

I set $\lambda = 0.25$, a conventional central value for the excess burden of U.S. taxation,[10] and $\kappa = 0$ in the baseline, and evaluate s on a grid over $[0, 0.95]$. All dollar magnitudes are deflated to 2025 dollars with the annual CPI, so welfare sums across years are in constant dollars, and the exercise runs on the paper's 1996-2025 panel restricted to observations with a nonmissing base rate, so the demand slope below is estimated on

[10] λ is the marginal excess burden of taxation, the *marginal cost of public funds* net of the dollar itself: because the subsidy is financed by distortionary taxes rather than lump-sum levies, raising one more dollar of revenue destroys λ dollars of private surplus beyond the dollar collected, so a dollar of subsidy outlay costs society $1 + \lambda$ dollars (Browning 1976). General-equilibrium estimates for the United States put λ roughly between 0.17 and 0.56 depending on the tax instrument (Ballard, Shoven, and Whalley 1985; Browning 1987), which makes 0.25 a central choice. The choice matters little here: s^* is unchanged across $\lambda \in \{0, 0.15, 0.25, 0.30\}$, and the benchmark corner $s^* = 0$ holds even at $\lambda = 0$, because by the envelope logic marginal enrollees value coverage at exactly what they pay, so subsidy-induced take-up generates no surplus while the outlay on it is strictly positive. Appendix H derives this formally.

exactly the sample over which welfare is aggregated. Relative to the canonical Baily (1978) formula, whose central cost term is the moral-hazard response of insured behavior, that term is *zero* here by the estimates of Section 7; this is exactly what licenses the reduced form above, in which welfare depends only on participation surplus and the fiscal cost of funds.

Estimating the price responsiveness of participation, the log-odds of the buy-up share on the farmer-paid premium with the actuarial base rate as the instrument, gives a demand slope with the expected sign: participation falls as the farmer-paid premium rises ($\hat{\beta} = 0.29$ per 2025 dollar).[11] Feeding it through the welfare simulation, Figure 25 shows that $\Delta W(s)$ is monotonically decreasing in s: the optimal buy-up subsidy is $s^* = 0$, robust across $\lambda \in \{0, 0.15, 0.25, 0.30\}$. Intuitively, it means that without a market failure for the subsidy to correct, raising s creates little value: the few farmers it draws into coverage were nearly indifferent to enrolling (hence the marginal farmer), so their new surplus is small, while the government must pay the higher subsidy on every acre that was already enrolled. The cost outweighs the small gain at the margin at every subsidy level, so the roughly 57% subsidy in force (measured in Section 9.1) exceeds the extensive-margin optimum of zero. This result aligns with Jialing Yu (2017) that finds negative net marginal welfare effects of both coverage and subsidy rates.

This benchmark model is absent the credit-channel yield gain of Section 5. It also abstracts from the two classic rationales for insurance subsidies of correcting adverse selection and crowding out ad hoc disaster relief. In the next two subsections, I add these two aspects to the model.

9.1 Disaster Relief Crowd Out

The FCIP's original purpose of raising insurance subsidies in 1980s was to replace ad-hoc disaster relief of farmers damaged by unexpected event (Glauber 2013). In this section, I add that crowd-out. Let an uninsured acre that suffers a loss receives federal relief covering a share ρ of the loss; insuring it saves the government $\rho\,\mathbb{E}[\text{loss}]$. Government outlay per acre becomes $s\,P\,q + \rho\,P\,(1-q)$ (subsidy plus relief on the uninsured), where q is the buy-up share and I proxy the expected loss by the gross premium P. The crowd out mechanism is that raising the subsidy now yields a fiscal saving, as fewer uninsured acres would left to bail out, which in turn offsets its cost.

[11] The paper's main access instruments are deliberately not used here: the endogenous regressor is the *price*, and a demand curve requires a price shifter excluded from demand, which access shifters are not; the base rate is the supply-side shift component of the main interacted instruments. The first stage is strong and correctly signed (KP $F = 70.8$), OLS is wrong-signed as simultaneity predicts, and Appendix **??** reports the full validity battery (first stage, predeterminedness of the rates, lagged-rate instruments, acre weighting): no alternative moves the relief optimum outside [0.37, 0.40].

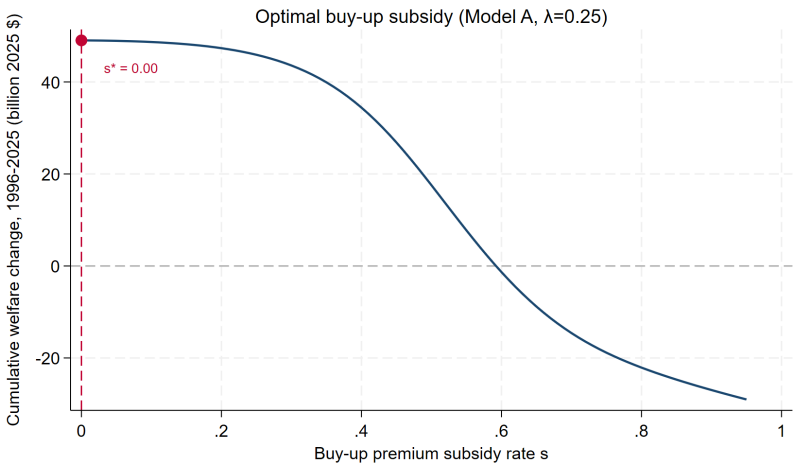


Figure 25: Welfare change relative to the baseline as a function of the buy-up premium subsidy rate ($\lambda = 0.25$; cumulative over 1996-2025, acre-weighted, billion 2025 dollars). Welfare compares each counterfactual uniform subsidy with the *observed* county-crop-year subsidy schedule, which is why the value at $s = 0$ is positive. The optimum is at $s^* = 0$.

This delivers an interior optimum that rises with ρ (Figure 26): $s^* = 0$ at $\rho = 0$, 0.27 at $\rho = 0.30$, 0.46 at $\rho = 0.50$, and 0.60 at $\rho = 0.70$. Intuitively $s^* \approx \rho$: the subsidy is worth paying up to the point where the marginal outlay equals the marginal relief it averts. The timing is ex ante: the subsidy is set before losses are realized and is chosen against expected relief $\rho P(1-q)$[12].

Calibrating ρ from disaster-payment data. Using the USDA ad-hoc disaster-relief data from the Farm Income and Wealth Statistics (U.S. Department of Agriculture, Economic Research Service 2026a) and the RMA cause-of-loss indemnities (U.S. Department of Agriculture, Risk Management Agency 2026b), I estimate the crowding-out effect. For the eight states over 1989-2024, ad-hoc disaster payments total \$23.0 billion against \$54.9 billion of insured disaster losses, so relief reimburses about $\hat{\rho} = 23.0/54.9 \approx 0.42$ per dollar of loss. Figure 26 shows what any value in this [0.36, 0.51] range implies for s^*. At the estimated $\hat{\rho} \approx 0.42$ the optimal buy-up subsidy is $s^* = 0.39$ (Figure 26). Sampling uncertainty barely moves it: a county-cluster bootstrap of the entire procedure puts a 95% interval of [0.37, 0.40] around the relief optimum, and the benchmark corner $s^* = 0$ obtains in every draw. The demand slope itself is much less precise across draws (95% interval [0.20, 0.41]), but the interior op-

[12]The mechanism may read backwards at first: the substitution here is fiscal, insurance replacing relief in the government budget rather than in the farmer's risk protection, so a larger relief liability calls for a larger subsidy to divert acres out of the costlier bailout channel, not a smaller one.

timum is pinned by the crowd-out rate rather than by β, so that imprecision hardly transmits (Appendix H shows why: β enters the first-order condition only through a small shading term below ρ). Relative to the \$8.1 billion cumulative gain (2025 dollars) that the optimum delivers over the observed baseline, setting the subsidy five points off forfeits 2–9%; ten points off forfeits about 12% below the optimum but 26% above it, the surface being steeper on the high side; and a uniform subsidy of 0.60, roughly the rate in force in the recent decade, does slightly worse than the status quo.

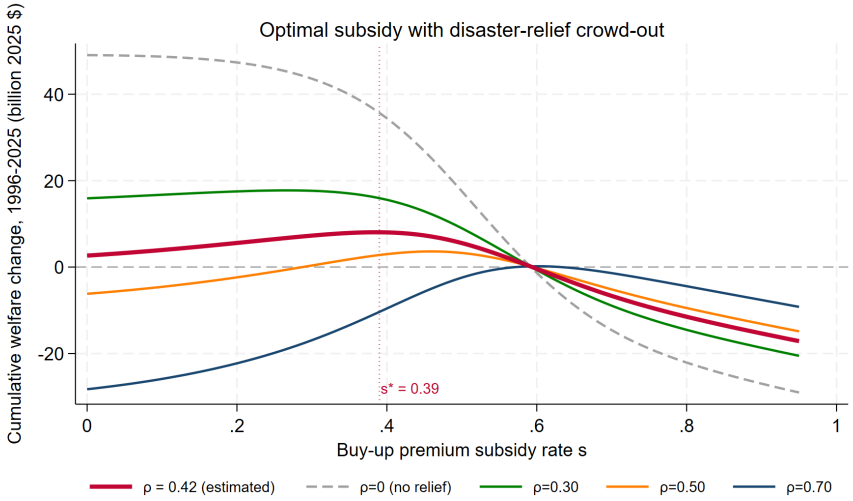


Figure 26: Welfare change vs. baseline across the subsidy rate, for several disaster-relief crowd-out rates ρ ($\lambda = 0.25$; cumulative over 1996-2025, acre-weighted, 2025 dollars). The optimum is interior and rises with ρ ($s^* = 0.27$, 0.46, 0.60 at $\rho = 0.30$, 0.50, 0.70). The thick highlighted curve is the estimated $\hat{\rho} = 0.42$; the dotted vertical line marks its optimum, the headline $s^* = 0.39$.

9.2 Internalizing the Credit Externality

When a subsidy raises the buy-up share from q_0 to $q(s)$, induced participation raises production by $\beta_Y = \partial \ln y / \partial q$. That shift moves *log* yield by $\beta_Y\,(q(s) - q_0)$, so yield itself scales by $\exp(\beta_Y\,(q(s) - q_0))$; valuing the extra output at baseline revenue R_0[13] gives the gain

$$R_0\big[\exp(\beta_Y\,(q(s) - q_0)) - 1\big] \quad \text{per planted acre,} \tag{18}$$

[13]Prices are the NASS marketing-year averages received by farmers (**usda_ers_sapf**), merged by crop and year (1996-2024 final values; 2025 is the latest WASDE forecast, as that marketing year is not yet complete). Realized rather than expected prices match the realized-output interpretation of the gain, and year-specific prices avoid anchoring the valuation to a single point of the commodity price cycle: the marketing-year average ranged from \$1.82 to \$6.89 per bushel for corn over the sample, so any one year's price is an arbitrary choice.

where β_Y is the extensive-margin yield effect estimated in Section 5 (≈ 0.56). For a small take-up change $\exp(x) - 1 \approx x$, so (18) reduces to the first-order revenue gain from the induced output, $R_0\,\beta_Y\,(q(s) - q_0)$. Figure 27 re-optimizes welfare with this term.

The optimum is extraordinarily sensitive to β_Y. With no credit externality ($\beta_Y = 0$) the optimum is $s^* = 0$, as before; but even a small yield effect tips it sharply-$\beta_Y = 0.10$ gives $s^* = 0.78$, and for $\beta_Y \geq 0.25$ the optimum reaches the policy ceiling. At the estimated $\beta_Y \approx 0.56$, s^* sits at the 95% corner. With this, I can conclude that the risk protection alone justifies no subsidy, whereas the credit externality this paper identifies can justify a very large one.

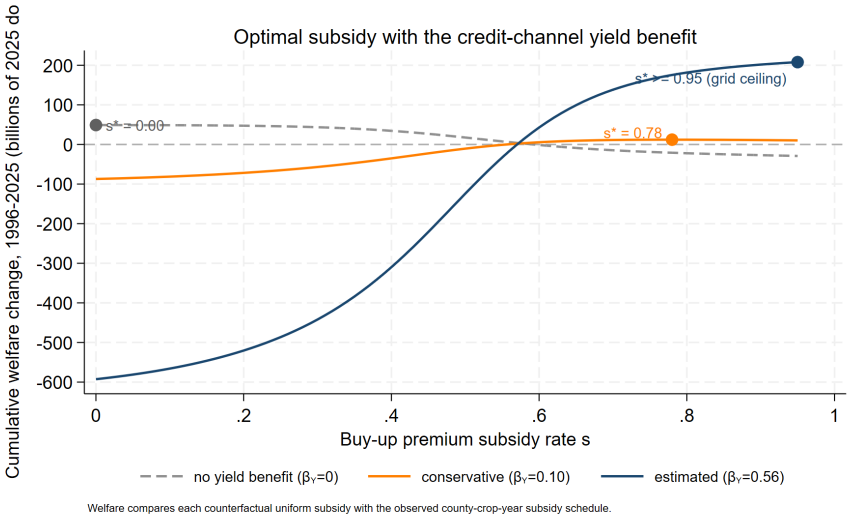


Figure 27: Welfare change vs. baseline across the subsidy rate, for three sizes of the credit-channel yield effect β_Y ($\lambda = 0.25$; cumulative over 1996-2025, acre-weighted, 2025 dollars): no yield benefit ($\beta_Y = 0$, optimum at $s^* = 0$), a conservative benefit ($\beta_Y = 0.10$, interior optimum $s^* = 0.78$), and the estimated benefit ($\beta_Y \approx 0.56$, optimum at the grid ceiling). Comparisons are against the observed county-crop-year subsidy schedule.

One important point to consider is that this calculation of yield is done without netting input costs, so it is an upper bound; the β_Y grid (0 to 0.56) is best read as spanning the net benefit. A second point is that β_Y is a complier LATE extrapolated across the participation range. Because the net benefit is not pinned down, I therefore follow Gruber (1997) and Diamond and Saez (2011) and report the whole relationship rather than one number. The most informative entry in that relationship is this: at a net benefit of 0.05-0.07, about one-tenth of the gross estimate, the optimal subsidy is 0.52-0.64, encompassing the actual rate of 0.57-0.60. After input costs, a net benefit of

0.05 equals the revenue from a 5% yield increase at full enrollment.

Table 16 reports the optimal subsidy under each rationale, in 2025 dollars and at $\lambda = 0.25$. Each row adds one rationale to the benchmark. With risk protection alone, the optimal subsidy is $s^{=}0$. Adding disaster-relief savings, using $\hat{\rho} = 0.42$, raises the optimum to 0.39. The estimate also rises with the assumed crowd-out rate: $s^{=}0.27$, 0.46, and 0.60 when $\rho = 0.30$, 0.50, and 0.70. Adding the credit benefit pushes the optimum to the top of the simulated grid. I report this as at least the grid ceiling because welfare is still rising at that point. The memo rows show the observed buy-up subsidy rate, weighted by premium dollars. It is 0.57 over the full sample and about 0.60 in 2020-2025. The comparison therefore depends on which benefits are included. With only risk protection and disaster-relief savings, the observed rate is above the estimated optimum ($0.60 > 0.39$). Once the credit benefit is included, the estimated optimum is above the observed rate. The current subsidy therefore falls between the estimates implied by the paper's main mechanisms. The remaining uncertainty comes mainly from the size of the net credit benefit. The next section examines the social multiplier, which is still held at its neutral value.

Table 16: Optimal buy-up premium subsidy s^* under each rationale ($\lambda = 0.25$, acre-weighted welfare in 2025 dollars). Each row adds one rationale to the benchmark; the two memo rows give the observed premium-weighted subsidy in force. The credit optimum is a corner at the simulated grid ceiling.

Scenario	Additional input	Subsidy rate
Benchmark (risk protection only)	none	$s^* = 0.00$
Add disaster-relief crowd-out	$\hat{\rho} = 0.42$	$s^* = 0.39$
Add credit externality	$\beta_Y \approx 0.56$	$s^* \geq 0.95$ (ceiling)
Observed subsidy, 1996–2025		0.57
Observed subsidy, 2020–2025		0.60

The missing social multiplier. . If enrollment diffuses across neighbors, a subsidy that raises one county's take-up also raises its neighbors', so the equilibrium participation response exceeds the individual response by a social multiplier m (Glaeser, Sacerdote, and Scheinkman 2003). Section 8 attempts to estimate that diffusion directly and concludes it is not credibly identified with county-level variation, so the calculations above set $m = 1$. Rather than leave the omission unexamined, I treat m the way ρ and β_Y are treated: as a calibrated parameter. Field experiments on insurance take-up and the technology-diffusion literature establish that adoption peer effects are real and material (Cai, Janvry, and Sadoulet 2015; Foster and Rosenzweig 1995; Conley and Udry

2010), motivating a sensitivity band of $m \in \{1.25, 1.5, 2\}$,[14] implemented by scaling the demand slope ($\beta \to m\beta$) in the enrollment response and the log-sum surplus alike; this values peer-induced enrollment on the same demand curve, a social-learning reading under which the surplus side of the band is, if anything, generous. The optimum barely moves: the benchmark stays at $s^* = 0$ for every m, the disaster-relief optimum at $\hat{\rho} = 0.42$ rises from 0.39 to 0.40 at $m = 2$, and the credit case at $\beta_Y = 0.10$ drifts from 0.78 to 0.75 across the band while remaining at the ceiling at the estimated β_Y. The reason is that the multiplier amplifies the participation surplus and the fiscal outlay nearly in proportion, so the interior optimum, which is pinned by the crowd-out rate ρ, is close to invariant. Two additional patterns are worth noting. First, the welfare gain at the relief optimum rises with (m), from $(8.1) billion at $(m = 1) to (14.4) billion at (m=5). This increase does not come from preventing more relief. Because the observed subsidy is above the relief-only optimum, the gain comes mainly from reducing excess subsidy spending. Part of the increase also depends on the assumed value of social learning. Second, (m) raises the relief optimum but lowers the interior credit optimum. With a production benefit, stronger diffusion increases participation at a lower subsidy, so further subsidy increases mainly go to farmers who are already insured.

10 Discussion and Conclusion

I revisit a puzzle in this paper: practitioners see distortion, empirical work sees small yield effect. To do so, I have implemented a research design that instruments both margins of insurance enrollment and tests the distortion hypothesis. The results point to three conclusions.

First, on the extensive margin, subsidized insurance raises mean yield, with an elasticity of about 0.08, the opposite of moral hazard. This is most consistent with a credit/liquidity channel. Second, I find no evidence that insurance increases production risk. Enrollment does not make yields more sensitive to extreme heat on either margin. After removing the mean effect of insurance from the yield surprise, neither margin increases yield variance or downside risk, and every heat interaction points away from moral hazard. Subsidized insurance therefore appears yield-enhancing. Third, when the

[14] The multiplier m is the ratio of the *equilibrium* take-up response to a subsidy, after diffusion across neighbors has worked through, to the *direct* individual response holding neighbors fixed. A hypothetical value m_0 thus means the general-equilibrium response is m_0 times the partial-equilibrium one, with $m_0 = 1$ the no-spillover benchmark. It maps to the peer coefficient γ of Section 8 by $m = 1/(1-\gamma)$, so the band $\{1.25, 1.5, 2\}$ corresponds to $\gamma \in \{0.2, 0.33, 0.5\}$. To fix ideas, $m = 2$ ($\gamma = \frac{1}{2}$) means a subsidy that would raise a county's own buy-up share by 5 percentage points in isolation raises it by 10 once neighbors also respond and the increases reinforce one another, half the total coming from the direct effect and half from diffusion.

focus is on risk protection alone, the extensive margin subsidy is not justified: $s^* = 0$. But once I add the program's disaster relief saving for the government, with ad-hoc relief reimbursing about $\hat{\rho} \approx 0.42$ per dollar of insured loss, the optimal buy-up subsidy rises to about 39%. Internalizing the credit-channel yield gain results in a higher optimal subsidy of ($s^* \geq 0.95$). These findings suggest that the observed subsidy is broadly justified by the mechanisms this paper identifies: crowd-out lifts the optimum from zero to 39%, and the credit externality raises it above the subsidy range that is currently in force. This result places crop insurance alongside flood insurance, for which Gruber and Solomon (2026) estimate an optimal subsidy of 52%. There, the subsidy reduces future FEMA spending and insures households against reclassification risk, while in crop insurance it reduces ad-hoc disaster relief and may relax credit constraints. In both settings, these benefits support a substantial positive subsidy.

Two questions remain open. The intensive margin is negative but its sign is unstable across different specifications, and it would require a design that isolates variation in coverage depth. Secondly, although peer effects is a first-order feature of the data, I could not derive causality with the available county-level variation. Identifying it will require an idiosyncratic shock or micro network data. Both are natural next steps for understanding how subsidized insurance reshapes production and for improving the welfare calculation.

A limitation of this research design is that it keeps the allocation of land across crops fixed. The panel is at the county-by-crop level, with yield measured per planted acre of that crop, so the analysis does not model insurance-induced reallocation of land between corn and soybeans, even though acreage responses to insurance are documented in the literature (Barry K. Goodwin, Vandeveer, and Deal 2004; Jisang Yu, A. Smith, and Sumner 2018). If enrollment moves marginal land into or out of a crop, part of the measured yield response could reflect the changing composition of planted land rather than production choices on given land, and the welfare calculation likewise values yield gains at the observed acreage rather than re-optimized acreage. Modeling the acreage margin jointly with enrollment is a natural extension.

Three additional factors limit the welfare exercise. First, the optimal-subsidy calculation does not model the farmer's choice of coverage level. Enrollment enters the policy problem only through the buy-up share, so the within-policy choice of coverage percentage is held fixed. Future work could bring this intensive margin directly into the planner's problem.

Second, in the spirit of the limitations that Chetty (2006) identifies for the optimal unemployment-benefit formula, the calculation abstracts from other fiscal externalities. The marginal cost of public funds, λ, captures the distortion from raising revenue, but the formula otherwise treats the crop insurance subsidy as the only distortionary policy in the economy. In practice, farm income and input use interact with commodity support programs, the tax system, and other subsidies. Behavioral responses that change those tax or subsidy bases would enter a more general welfare formula.

Third, the analysis measures enrollment using the buy-up share of planted acres and liability per insured acre. Other measures, including total insured acres, buy-up acres, total premiums, and premiums per acre, may capture different dimensions of participation (O'Donoghue 2014). Re-estimating the main results using this broader set of measures would show whether the conclusions depend on the definitions used here.

Finally, the conversations that prompted the paper deserve an answer. The farmers were right that insurance finances what would otherwise be unaffordable: enrollment raises Commodity Credit Corporation borrowing, which is the channel they described. What does not follow is the conclusion from it. Because what that financing buys, raises yields rather than decreasing it, and enrollment leaves yields no more sensitive to extreme heat, no more variable, and no more prone to downside losses. Their remaining concern, that insurance narrows what gets planted, is the one this design cannot address, and it is the natural next step.

11 References

Adegbola, Patrice and Cornelis Gardebroek (2007). "The effect of information sources on technology adoption and modification decisions". In: *Agricultural Economics* 37.1, pp. 55–65.

Andrews, Isaiah (July 2018). *Weak Instruments and What to Do about Them*. NBER Methods Lecture, YouTube video. National Bureau of Economic Research. URL: https://www.youtube.com/watch?v=SW-WT_1qzRc (visited on 07/14/2026).

Annan, Francis and Wolfram Schlenker (2015). "Federal Crop Insurance and the Disincentive to Adapt to Extreme Heat". In: *American Economic Review* 105.5, pp. 262–266.

Aroussi, Ran (2019). *yfinance: Download market data from Yahoo! Finance's API*. https://github.com/ranaroussi/yfinance. Python library. PyPI: https://pypi.org/project/yfinance/. Accessed June 2, 2026.

Babcock, Bruce A and Chad E Hart (2005). "Influence of the premium subsidy on farmers' crop insurance coverage decisions". In.

Babcock, Bruce A. and David A. Hennessy (1996). "Input Demand under Yield and Revenue Insurance". In: *American Journal of Agricultural Economics* 78.2, pp. 416–427.

Baily, Martin Neil (1978). "Some aspects of optimal unemployment insurance". In: *Journal of public Economics* 10.3, pp. 379–402.

Ballard, Charles L, John B Shoven, and John Whalley (1985). "General equilibrium computations of the marginal welfare costs of taxes in the United States". In: *The American Economic Review* 75.1, pp. 128–138.

Bartik, Timothy J (1991). *Who benefits from state and local economic development policies?* JSTOR.

Baskerville, Gregory L and Peter Emin (1969). "Rapid estimation of heat accumulation from maximum and minimum temperatures". In: *Ecology* 50.3, pp. 514–517.

Bellemare, Marc F, Jhih-Yun Liu, and Joleen Hadrich (2025). "Commodity Revenue Shocks and Mortality". In: *Journal of Human Resources*.

Brady, Ryan R (2011). "Measuring the diffusion of housing prices across space and over time". In: *Journal of Applied Econometrics* 26.2, pp. 213–231.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin (2009). "Identification of Peer Effects through Social Networks". In: *Journal of Econometrics* 150.1, pp. 41–55.

Brazil Ministry of Agriculture, Livestock and Food Supply (Nov. 2016). *Agricultural Risk Management in Brazil*. Presentation at the World Trade Organization. Accessed via WTO website. Geneva. URL: https://www.wto.org/english/tratop_e/agric_e/brazil91116_e.pdf.

Browning, Edgar K (1976). "The Marginal Cost of Public Funds". In: *Journal of Political Economy* 84.2, pp. 283–298.

— (1987). "On the marginal welfare cost of taxation". In: *The American Economic Review*, pp. 11–23.

Bureau, U.S. Census (2000). *Census 2000 TIGER/Line Files*. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html. Geographic boundary files in TIGER/Line ASCII format.

Bureau of Economic Analysis (2026). *Price Indexes for Gross Domestic Purchases*. Interactive Data Application, Table 1.6.4. URL: https://apps.bea.gov/iTable/?categories=survey&isuri=1&reqid=19&step=2#eyJhcHBpZCI6MTksInN0ZXBzIjpbMSwyLDMsM
=.

Cai, Jing, Alain De Janvry, and Elisabeth Sadoulet (2015). "Social networks and the decision to insure". In: *American Economic Journal: Applied Economics* 7.2, pp. 81–108.

Chemeris, Anna, Yong Liu, and Alan P. Ker (2022). "Insurance subsidies, climate change, and innovation: Implications for crop yield resiliency". In: *Food Policy* 108.

Chetty, Raj (2006). "A general formula for the optimal level of social insurance". In: *Journal of Public Economics* 90.10-11, pp. 1879–1901.

Chetty, Raj and Amy Finkelstein (2013). "Social insurance: Connecting theory to data". In: *Handbook of public economics*. Vol. 5. Elsevier, pp. 111–193.

Chintapalli, Prashant and Christopher S Tang (2022). "Crop minimum support price versus cost subsidy: Farmer and consumer welfare". In: *Production and Operations Management* 31.4, pp. 1753–1769.

Coate, Stephen (1995). "Altruism, the Samaritan's dilemma, and government transfer policy". In: *The American Economic Review*, pp. 46–57.

Coble, Keith H et al. (2010). "A comprehensive review of the rma aph and combo rating methodology: Final report". In: *Prepared by Sumaria systems for the risk Management agency*.

Coble, Keith H. and Barry J. Barnett (2013). "Why do we subsidize crop insurance?" In: *American Journal of Agricultural Economics* 95.2, pp. 498–504.

Conley, Timothy G and Christopher R Udry (2010). "Learning about a New Technology: Pineapple in Ghana". In: *American Economic Review* 100.1, pp. 35–69.

Conlisk, John (1980). "Costly optimizers versus cheap imitators". In: *Journal of Economic Behavior & Organization* 1.3, pp. 275–293.

Connor, Lawson and Ani L Katchova (2020). "Crop insurance participation rates and asymmetric effects on US corn and soybean yield risk". In: *Journal of Agricultural and Resource Economics* 45.1, pp. 1–19.

Davidson, Russell and James G. MacKinnon (2000). "Bootstrap Tests: How Many Bootstraps?" In: *Econometric Reviews* 19.1, pp. 55–68.

DeSimone, Bailey (Mar. 15, 2021). *History of the United States Farm Bill*. Law Library of Congress. URL: https://blogs.loc.gov/law/2021/03/history-of-the-united-states-farm-bill/ (visited on 03/16/2026).

Diamond, Peter and Emmanuel Saez (2011). "The Case for a Progressive Tax: From Basic Research to Policy Recommendations". In: *Journal of Economic Perspectives* 25.4, pp. 165–190.

Edwards, Chris (2025). *Department of Agriculture Timeline*. Accessed: 2025-11-03. Cato Institute. URL: https://www.downsizinggovernment.org/agriculture/timeline.

Einav, Liran, Amy Finkelstein, and Mark R Cullen (2010). "Estimating welfare in insurance markets using variation in prices". In: *The quarterly journal of economics* 125.3, pp. 877–921.

Florax, Raymond and Henk Folmer (1992). "Specification and estimation of spatial linear regression models: Monte Carlo evaluation of pre-test estimators". In: *Regional science and urban economics* 22.3, pp. 405–432.

Foster, Andrew D and Mark R Rosenzweig (1995). "Learning by doing and learning from others: Human capital and technical change in agriculture". In: *Journal of political Economy* 103.6, pp. 1176–1209.

Friedman, Milton (1962). "The interpolation of time series by related series". In: *Journal of the American Statistical Association* 57.300, pp. 729–757.

Gali, Jordi (1994). "Keeping up with the Joneses: Consumption externalities, portfolio choice, and asset prices". In: *Journal of Money, Credit and Banking* 26.1, pp. 1–8.

Glaeser, Edward L, Bruce I Sacerdote, and José A Scheinkman (2003). "The Social Multiplier". In: *Journal of the European Economic Association* 1.2-3, pp. 345–353.

Glauber, Joseph W (2004). "Crop insurance reconsidered". In: *American Journal of Agricultural Economics* 86.5, pp. 1179–1195.

— (2013). "The growth of the federal crop insurance program, 1990—2011". In: *American Journal of Agricultural Economics* 95.2, pp. 482–488.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020). "Bartik instruments: What, when, why, and how". In: *American Economic Review* 110.8, pp. 2586–2624.

Gonne, Nicolas and Olivier Hubert (2020). "Spatial distancing: air traffic, COVID-19 propagation, and the cost efficiency of air travel restrictions". In: *Covid Economics* 24, pp. 111–25.

Goodwin, Barry K (1993). "An empirical analysis of the demand for multiple peril crop insurance". In: *American journal of agricultural economics* 75.2, pp. 425–434.

Goodwin, Barry K., Monte L. Vandeveer, and John L. Deal (2004). "An Empirical Analysis of Acreage Effects of Participation in the Federal Crop Insurance Program". In: *American Journal of Agricultural Economics* 86.4, pp. 1058–1077.

Gruber, Jonathan (1997). "The Consumption Smoothing Benefits of Unemployment Insurance". In: *American Economic Review* 87.1, pp. 192–205.

Gruber, Jonathan and Adam Solomon (Jan. 26, 2026). *Optimal Flood Insurance in a Second-Best World: Fiscal Spillovers, Reclassification Risk and Moral Hazard*. Working paper. URL: https://adam-solomon.com/FloodPaper.pdf (visited on 03/09/2026).

Halleck Vega, Solmaria and J Paul Elhorst (2015). "The SLX model". In: *Journal of Regional Science* 55.3, pp. 339–363.

Hendren, Nathaniel and Ben Sprung-Keyser (2020). "A unified welfare analysis of government policies". In: *The Quarterly journal of economics* 135.3, pp. 1209–1318.

Horowitz, John K. and Erik Lichtenberg (1993). "Insurance, Moral Hazard, and Chemical Use in Agriculture". In: *American Journal of Agricultural Economics* 75.4, pp. 926–935.

Hsiang, Solomon (2016). "Climate econometrics". In: *Annual Review of Resource Economics* 8, pp. 43–75.

Jordà, Òscar (2009). "Simultaneous confidence regions for impulse responses". In: *The Review of Economics and Statistics* 91.3, pp. 629–647.

— (2023). "Local projections for applied economics". In: *Annual Review of Economics* 15.1, pp. 607–631.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor (2015). "Betting the house". In: *Journal of international economics* 96, S2–S18.

Jordà, Òscar and Alan M Taylor (2025). "Local projections". In: *Journal of Economic Literature* 63.1, pp. 59–110.

Just, Richard E. and Rulon D. Pope (1978). "Stochastic Specification of Production Functions and Economic Implications". In: *Journal of Econometrics* 7.1, pp. 67–86.

Klosin, Sylvia and Adam Solomon (2024). "Optimal Insurance Scope: Theory and Evidence from US Crop Insurance". In: *Manuscript in preparation*.

Koijen, Ralph SJ and Motohiro Yogo (2023). *Financial economics of insurance*. Princeton University Press.

LaPorte, Jonathan (Mar. 21, 2025). *The Corn–Soybean Ratio and Its Potential Impact on Farm Profits*. Michigan State University Extension. URL: https://www.canr.msu.edu/news/the-corn-soybean-ratio-and-its-potential-impact-on-farm-profits (visited on 09/21/2025).

Library of Congress (2026). *History of the United States Farm Bill*. Story Map. Library of Congress. URL: https://www.loc.gov/ghe/cascade/index.html?appid=

1821e70c01de48ae899a7ff708d6ad8b&bookmark=What+is+the+Farm+Bill (visited on 03/16/2026).

Mahul, Olivier and Charles J Stutley (2010). *Government support to agricultural insurance: challenges and options for developing countries*. World Bank Publications.

Menne, Matthew J, Imke Durre, Bryant Korzeniewski, et al. (2012). "Global historical climatology network-daily (GHCN-Daily), Version 3". In: *(No Title)*.

Menne, Matthew J, Imke Durre, Russell S Vose, et al. (2012). "An overview of the global historical climatology network-daily database". In: *Journal of atmospheric and oceanic technology* 29.7, pp. 897–910.

Mieno, Taro, Cory G Walters, and Lilyan E Fulginiti (2018). "Input use under crop insurance: the role of actual production history". In: *American Journal of Agricultural Economics* 100.5, pp. 1469–1485.

Mishra, Ashok K and Vijay P Singh (2010). "A review of drought concepts". In: *Journal of hydrology* 391.1-2, pp. 202–216.

Montiel Olea, José Luis and Mikkel Plagborg-Møller (2021). "Local projection inference is simpler and more robust than you think". In: *Econometrica* 89.4, pp. 1789–1823.

Moretti, Enrico (2011). "Social learning and peer effects in consumption: Evidence from movie sales". In: *The Review of Economic Studies* 78.1, pp. 356–393.

Munshi, Kaivan (2004). "Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution". In: *Journal of development Economics* 73.1, pp. 185–213.

NOAA National Centers for Environmental Information (2025a). *GHCN-Daily Station Inventory (ghcnd-inventory.txt)*. https://www.ncei.noaa.gov/pub/data/ghcn/daily/ghcnd-inventory.txt. Accessed January 07, 2026.

— (2025b). *GHCN-Daily Station Metadata (ghcnd-stations.txt)*. https://www.ncei.noaa.gov/pub/data/ghcn/daily/ghcnd-stations.txt. Accessed January 07, 2026.

— (2025c). *README for Daily Global Historical Climatology Network (GHCN-Daily)*. https://www.ncei.noaa.gov/pub/data/ghcn/daily/readme.txt. Accessed on January 07, 2026.

— (2026a). *Climate Divisional Database: Palmer Drought Severity Index (PDSI), County-Level, Version 1.0.0*. https://www.ncei.noaa.gov/pub/data/cirs/climdiv/climdiv-pdsicy-v1.0.0-20260305. Released March 5, 2026; accessed April 2, 2026.

— (2026b). *County-to-Climate-Division Crosswalk*. https://www.ncei.noaa.gov/pub/data/cirs/climdiv/county-to-climdivs.txt. Text file linking counties to climate divisions; accessed April 2, 2026.

O'Donoghue, Erik (2014). "The effects of premium subsidies on demand for crop insurance". In: *USDA-ERS economic research report* 169.

OpenStreetMap contributors (2025). *OpenStreetMap*. https://www.openstreetmap.org. [Online; accessed 24-November-2025].

Parliament of the Islamic Republic of Iran (May 1983). *Agricultural Products Insurance Law*. Official legislation. Adopted 21 May 1983 (31 Ordibehesht 1362); Published 26 July 1983 (4 Mordad 1362). URL: https://web.archive.org/web/20190629235755/http://rc.majlis.ir/fa/law/show/90745 (visited on 02/03/2025).

Plastina, Alejandro (2021). *Current Crop Insurance Policies*. First published 2014; last updated August 2021. Iowa State University Extension and Outreach, Ag Decision Maker. URL: https://www.extension.iastate.edu/agdm/crops/html/a1-48.html (visited on 05/25/2026).

Regmi, Madhav, Jesse Tack, and Allen M. Featherstone (2023). "Does crop insurance influence crop yield impacts of warming temperatures? A farm-level analysis". In: *Journal of the Agricultural and Applied Economics Association* 2.4, pp. 808–822.

Roberts, Michael J, Nigel Key, and Erik O'Donoghue (2006). "Estimating the extent of moral hazard in crop insurance using administrative data". In: *Review of Agricultural Economics* 28.3, pp. 381–390.

Roosevelt, Franklin D. (Feb. 18, 1937). *Message to Congress on Crop Insurance*. Online by Gerhard Peters and John T. Woolley. URL: https://www.presidency.ucsb.edu/node/209338 (visited on 12/15/2025).

Rosa, I. (2018). *Federal Crop Insurance: Program Overview for the 115th Congress*. Report R45193. Congressional Research Service.

Rosch, Stephanie (2021). "Federal crop insurance: A primer (no. r46686)". In: *Congressional Research Service*.

Ross, Tom and Neal Lott (2003). *A Climatology of 1980–2003 Extreme Weather and Climate Events*. Technical Report 2003-01. National Climatic Data Center, NOAA/NESDIS, U.S. Department of Commerce. URL: https://www.ncei.noaa.gov/monitoring-content/billions/docs/lott-and-ross-2003.pdf.

Schlenker, Wolfram and Michael J Roberts (2009). "Nonlinear temperature effects indicate severe damages to US crop yields under climate change". In: *Proceedings of the National Academy of sciences* 106.37, pp. 15594–15598.

Schroeder, Jonathan et al. (2025a). *IPUMS National Historical Geographic Information System: Version 20.0*. Dataset. Minneapolis, MN: IPUMS. DOI: 10.18128/D050.V20.0.

— (2025b). *IPUMS National Historical Geographic Information System: Version 20.0 [dataset]*. Dataset. Minneapolis, MN. DOI: 10.18128/D050.V20.0.

Service, USDA Economic Research (1992). "Weights, measures, and conversion factors for agricultural commodities and their products". In: *Agric. Handbook 697*. URL: https://www.ers.usda.gov/media/7780/ah-697.pdf?v=75540.

Small, Kenneth A. and Harvey S. Rosen (1981). "Applied Welfare Economics with Discrete Choice Models". In: *Econometrica* 49.1, pp. 105–130.

Smith, Vincent H and Joseph W Glauber (2012). "Agricultural insurance in developed countries: where have we been and where are we going?" In: *Applied Economic Perspectives and Policy* 34.3, pp. 363–390.

Smith, Vincent H. and Barry K. Goodwin (1996). "Crop Insurance, Moral Hazard, and Agricultural Chemical Use". In: *American Journal of Agricultural Economics* 78.2, pp. 428–438.

Snyder, Richard L (1985). "Hand calculating degree days". In: *Agricultural and forest meteorology* 35.1-4, pp. 353–358.

Sun, Liyang and Jesse M Shapiro (2022). "A linear panel model with heterogeneous coefficients and variation in exposure". In: *Journal of Economic Perspectives* 36.4, pp. 193–204.

U.S. Department of Agriculture, Economic Research Service (2026a). *Farm Income and Wealth Statistics*. Data product, February 5, 2026 release. URL: https://www.ers.usda.gov/data-products/farm-income-and-wealth-statistics.

— (2026b). *Farm Income and Wealth Statistics: Government Payments by Program*. Updated February 5, 2026. URL: https://data.ers.usda.gov/reports.aspx?ID=4050.

— (Feb. 5, 2026c). *Farm Income and Wealth Statistics: Net Cash Income*. Accessed February 18, 2026. URL: https://data.ers.usda.gov/reports.aspx?ID=4024#Pcda946151771493492d2191ecfd07a50_8_xA.

— (2026d). *Farm Income and Wealth Statistics: Net Farm Income (U.S. and States)*. Accessed February 17, 2026. URL: https://data.ers.usda.gov/reports.aspx?ID=4047.

— (2026e). *Fertilizer Use and Price*. Accessed February 2026. USDA Economic Research Service. URL: https://www.ers.usda.gov/data-products/fertilizer-use-and-price.

U.S. Department of Agriculture, Farm Service Agency (2026). *Loan Deficiency Payments*. USDA FSA Program Documentation. URL: https://www.fsa.usda.gov/resources/price-support/loan-deficiency-payments.

U.S. Department of Agriculture, National Agricultural Statistics Service (May 14, 2026). *Quick Stats Crops Dataset*. Version May 14, 2026. Large-dataset snapshot, `qs.crops_20260514.txt.gz`. URL: https://www.nass.usda.gov/datasets/ (visited on 06/16/2026).

U.S. Department of Agriculture, Risk Management Agency (2025). *Premium Calculation: Exhibit P11-1, Plan 01, 02, 03*. M-13 Handbook, Reinsurance Year 2025, Draft Version. Record Name: Acreage; Record Code: P11; release date June 6, 2024. URL: https://www.rma.usda.gov/sites/default/files/handbooks-m13/2025/draft/P11_1_PLAN_01_02_03_PREMIUM_CALCULATION.pdf (visited on 05/21/2026).

U.S. Department of Agriculture, Risk Management Agency (2026a). *Actuarial Data Master*. Public data archive. Annual Actuarial Data Master files by reinsurance year. URL: https://pubfs-rma.fpac.usda.gov/pub/References/actuarial_data_master/ (visited on 05/20/2026).

— (2026b). *Summary of Business: Cause of Loss Historical Data Files*. Annual colsom files, accessed 2026. URL: https://www.rma.usda.gov/tools-reports/summary-of-business/cause-loss.

U.S. Department of Agriculture, Risk Management Agency (RMA) (2025). *State/County/Crop Summary of Business*. U.S. Department of Agriculture. URL: https://www.rma.usda.gov/tools-reports/summary-of-business/state-county-crop-summary-business (visited on 11/08/2025).

U.S. Government Accountability Office (Sept. 2005). *Crop Insurance: Actions Needed to Reduce Program's Vulnerability to Fraud, Waste, and Abuse*. GAO Report GAO-05-528. U.S. Government Accountability Office. URL: https://www.gao.gov/assets/gao-05-528.pdf (visited on 05/26/2026).

United States Department of Agriculture, National Agricultural Statistics Service (2025). *Data – Large Datasets*. https://www.nass.usda.gov/datasets/. Accessed 16 March 2025.

United States Department of Agriculture, Risk Management Agency (Feb. 2008). *Rate Methodology Handbook: Actual Production History (FCIC-11010 RMH-APH)*. Handbook FCIC-11010. Federal Crop Insurance Corporation (FCIC). URL: https://www.rma.usda.gov/sites/default/files/topics/ratemethodology.pdf.

USDA Economic Research Service (Jan. 31, 2025). *Educational attainment for adults age 25 and older for the United States, States, and counties, 1970–2023. County-level Data Sets: Download Data*. URL: https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data (visited on 02/24/2025).

USDA Farm Service Agency (2025). *USDA Service Center Locator*. U.S. Department of Agriculture. URL: https://offices.sc.egov.usda.gov/locator/app (visited on 11/24/2025).

USDA National Agricultural Statistics Service (2026). *Quick Stats Database*. https://quickstats.nass.usda.gov/. Accessed February 20, 2026.

Wang, Ruixue, Roderick M Rejesus, and Serkan Aglasan (2021). "Warming temperatures, yield risk and crop insurance participation". In: *European Review of Agricultural Economics* 48.5, pp. 1109–1131.

Wright, Brian D and Julie A Hewitt (1994). "All-risk crop insurance: lessons from theory and experience". In: *Economics of agricultural crop insurance: theory and evidence*. Springer, pp. 73–112.

Yu, Jialing (2017). "Welfare Effects of the Federal Crop Insurance Program: The Sufficient Statistics Approach". In.

Yu, Jisang, Aaron Smith, and Daniel A Sumner (2018). "Effects of crop insurance premium subsidies on crop acreage". In: *American Journal of Agricultural Economics* 100.1, pp. 91–114.

Zhou, Li et al. (2020). "The peer effect of training on farmers' pesticides application: a spatial econometric approach". In: *China Agricultural Economic Review* 12.3, pp. 481–505.

Appendices

A Parameters of Interest

In this paper, I often refer to technical expressions in the context of agricultural insurance. To be on the same page, I provide an explanation for the major terms from a Handbook of the United States Department of Agriculture, Risk Management Agency (2008, p. 50).

Liability is a measure of the USDA's exposure to loss for a given farmer or group of farmers. Liability is the maximum amount of insurance coverage that a farmer has under their policy for a given crop and year. It represents the total insured value of the crop.

$$\text{Liability} = \text{Acres Planted} \times \text{Expected Yield} \times \text{Selected Coverage Level} \times \text{Base Price} \times \text{Price Election Perce} \tag{A.1}$$

It is important to note that

- Farmers do not pay the liability itself. They pay a premium minus the subsidy to receive insurance coverage.
- When a covered loss occurs, the insurer (usually a private company that provides the federal crop insurance) pays the farmer an indemnity.
- Indemnity payment is triggered when the actual value of production falls below the liability.
  - If actual revenue $<$ liability $\rightarrow$ indemnity payment.
  - If actual revenue $\geq$ liability $\rightarrow$ no indemnity payment.

Indemnity is the amount paid to a farmer who suffers a covered loss. Indemnity is paid when the value of production is less than the purchased liability. The amount of indemnity paid equals the liability minus the value of the production, where they are defined as follows:

$$\text{Indemnity} = \text{Liability} - \text{Value of Production} \tag{A.2}$$

In which, *value of production* is defined as

$$\text{Value of Production} = \text{Acres Planted} \times \text{Actual Yield} \times \text{Base Price} \times \text{Price Election Percentage} \tag{A.3}$$

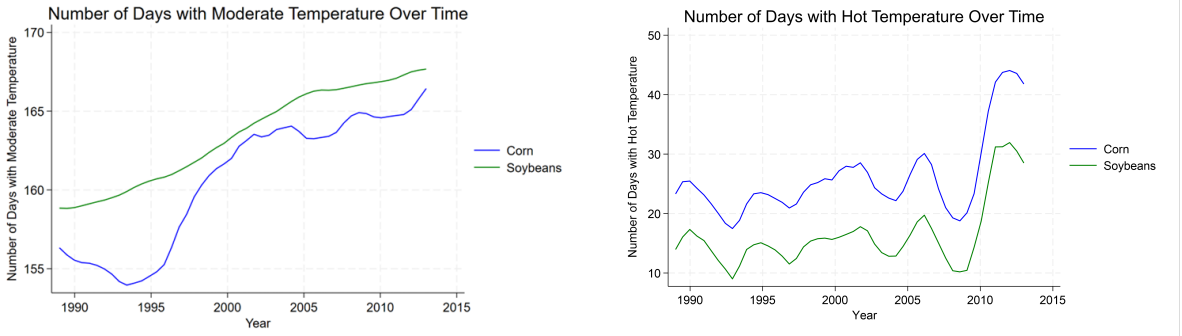
Finally, **premium**, is the price that farmers pay for insurance, and it equals

$$\text{Premium} = \text{Liability} \times \text{Premium Rate} \times \text{Adjustment Factor} \tag{A.4}$$

B Stylized Facts

To add: Statistics and map on adoption of different types of insurance coverage for corn and soybeans in different places, to justify the differences in experienced insurance subsidy.

During the 1989-2013 period, the weather has been relatively more conducive for producing soybeans, as captured by number of moderate days and extremely hot days (defined in Section 4) in Figures B.1a and B.1b.



(a) Number of days with moderate temperature (1989-2013)

(b) Number of days with extreme temperature (1989-2013)

Figure B.1: Temperature patterns over time for corn and soybeans, 1989-2013

Shifting from weather variables to production behavior of farmers, we know that a rule of thumb for farmers to decide on whether to grow corn or soybeans is to check the relative size of soybeans price to corn price and grow soybeans when this ratio is higher than 2.5 LaPorte (2025). Whereas Figure 3 that although in early 1990s we expected the farmers to have higher relative soybeans-to-corn acreage, they did not follow their rule of planting decisions. Figure 4 has a more detailed picture of relative prices of two crops over our study period.

Another observation is that corn yield has been higher in levels and variability compared to soybeans, suggesting that corn is a higher risk-higher return crop compared to soybeans.

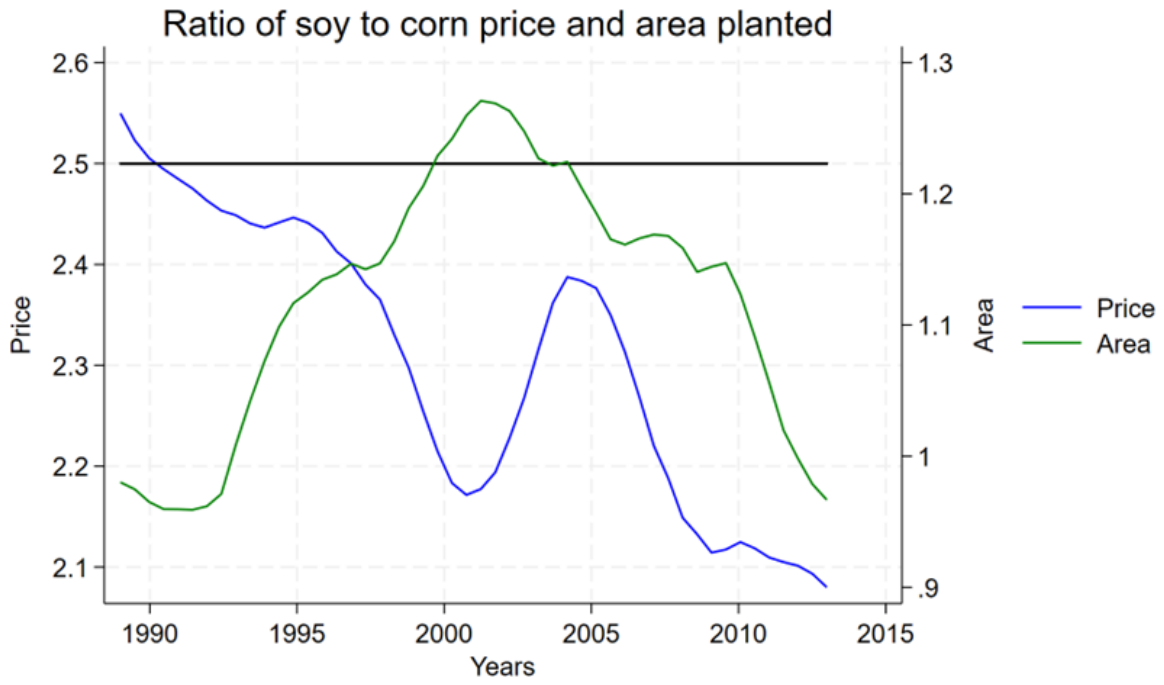


Figure B.2: Left axis: Received price of farmers (blue) for soybeans relative to corn. Right axis: Relative area planted for soybeans relative to corn (green). The black horizontal line shows the rule of thumb threshold of switching from corn to soybeans, as known in the farming community.

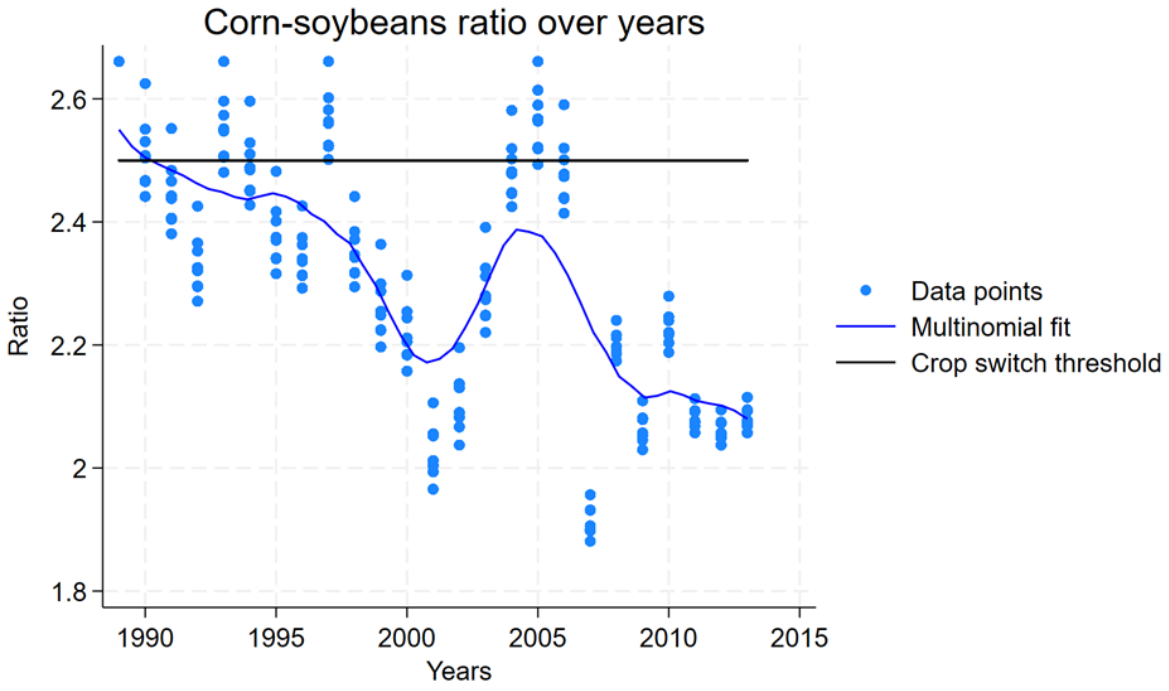


Figure B.3: Relative price of soybeans to corn in eight Midwest states and the crop switching threshold.

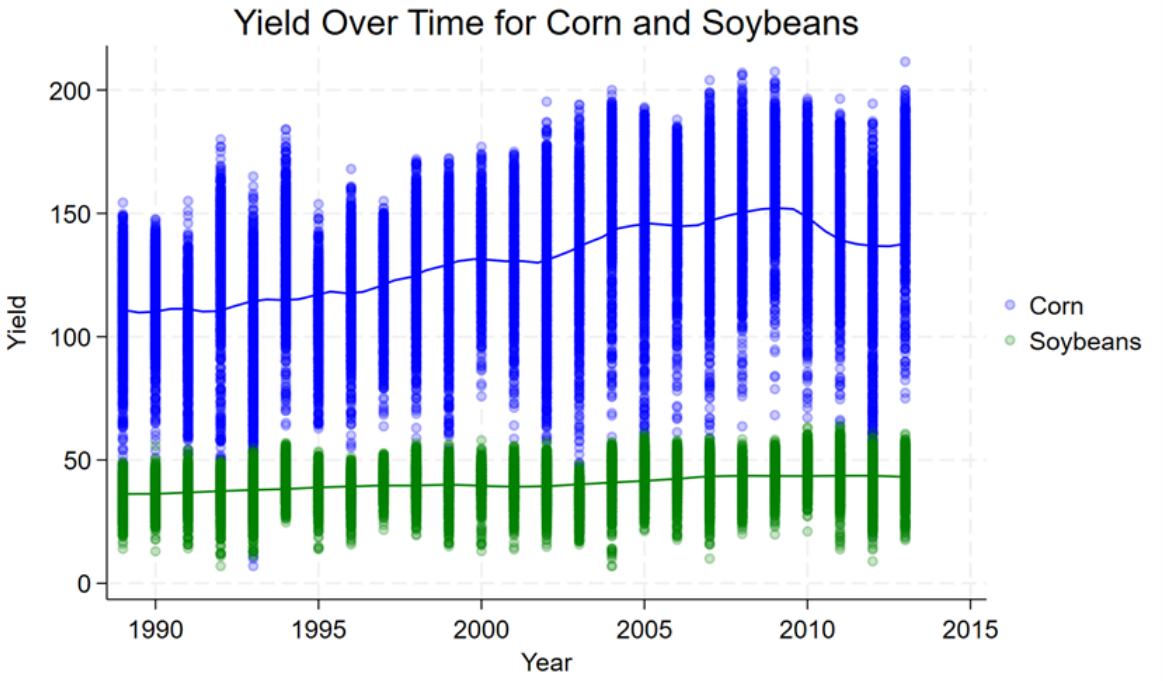


Figure B.4: Corn yield is higher and has higher variance compared to soybeans.

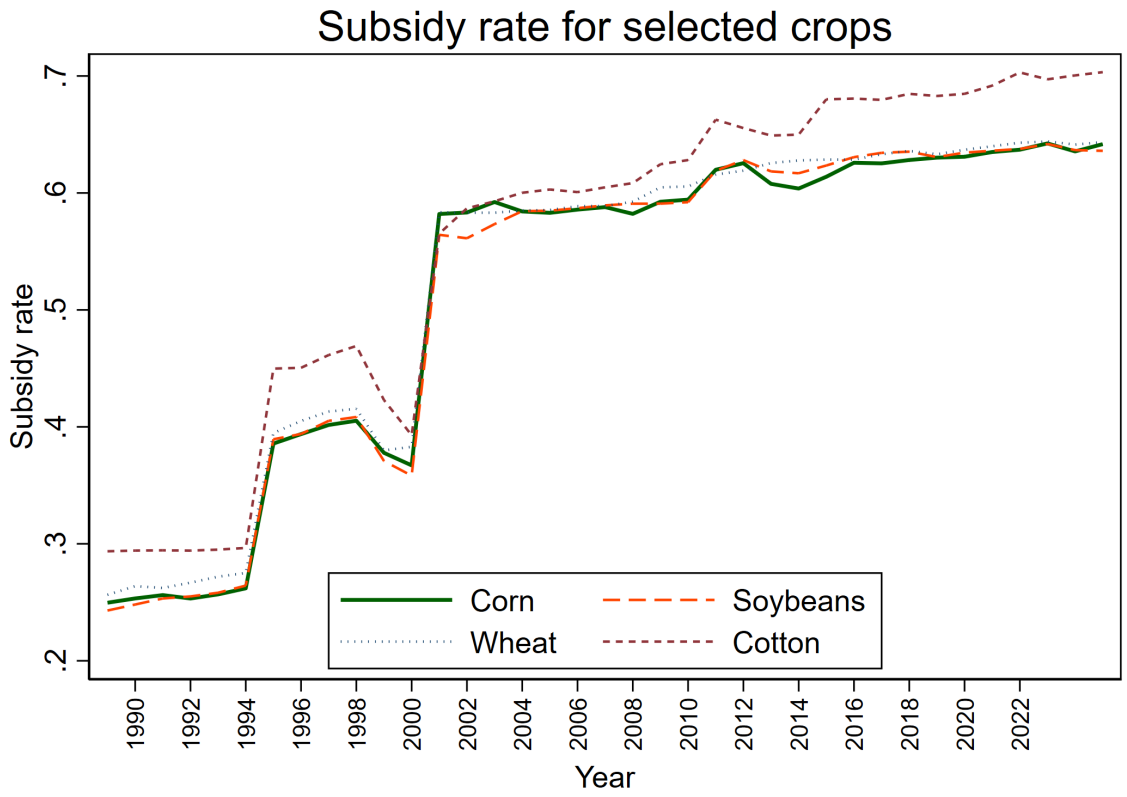
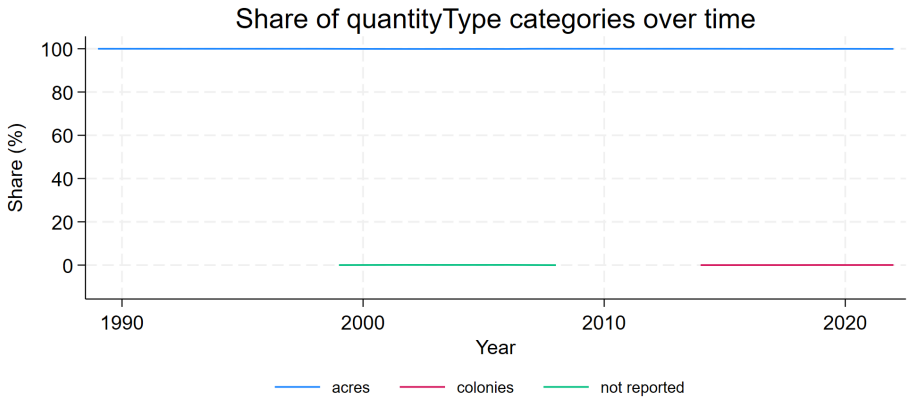
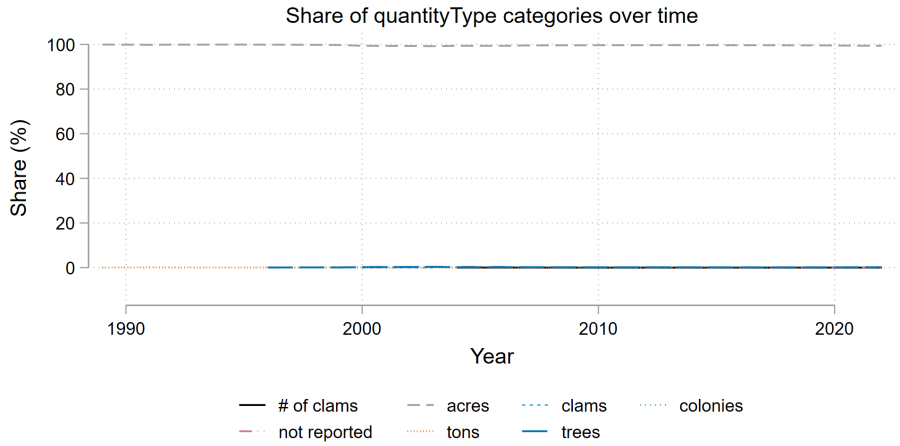


Figure B.5: Premium subsidies in yield based on revenue based insurance in all States for corn, soybeans, wheat and cotton

Type of Insurance Enrollment To do: Have graphs of same type to have same color codes, and decrease the empty gap between figures and their caption. I need to clean them in paint.



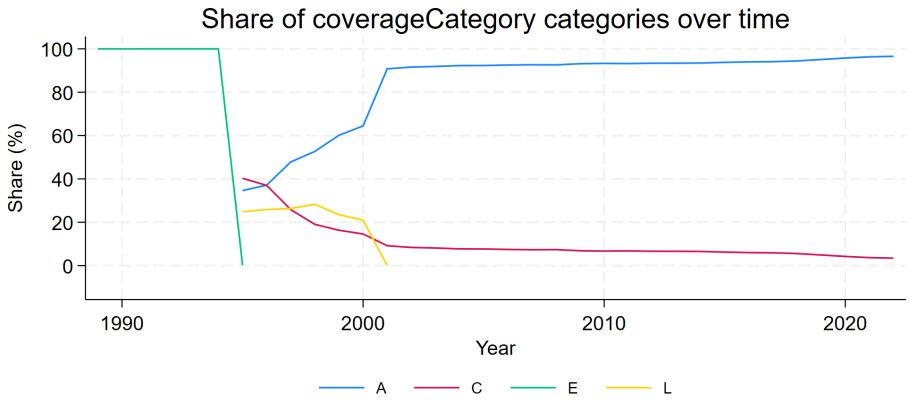
(a) Midwest



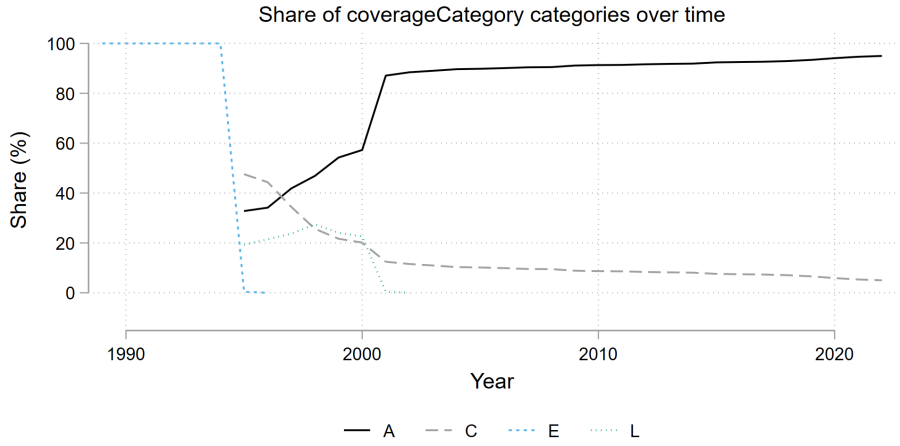
(b) Contingent United States

Figure B.6: Share of quantity type categories over time. Data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).

Education Level of Study Area

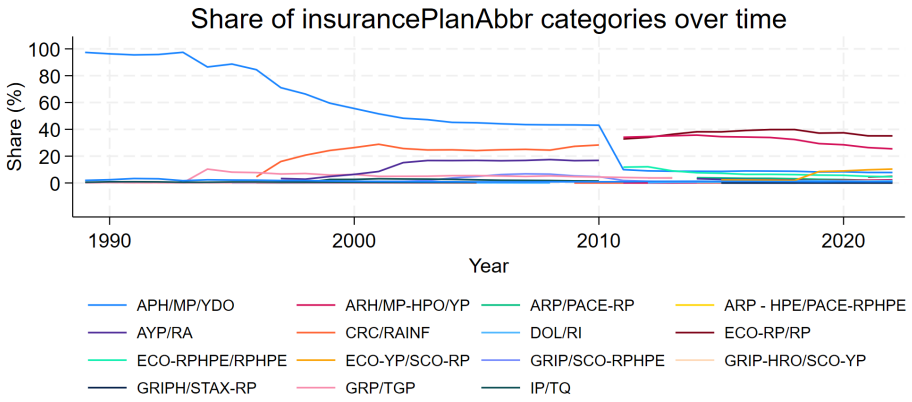


(a) Midwest

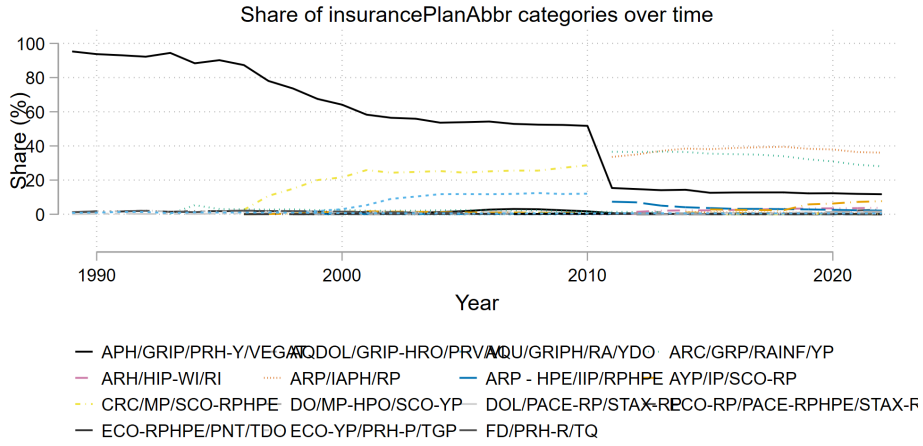


(b) Contingent United States

Figure B.7: Share of quantity type categories over time. Data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).

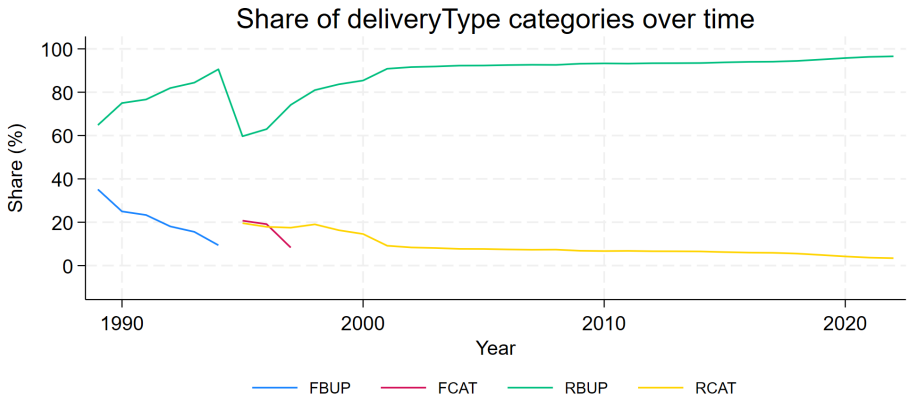


(a) Midwest

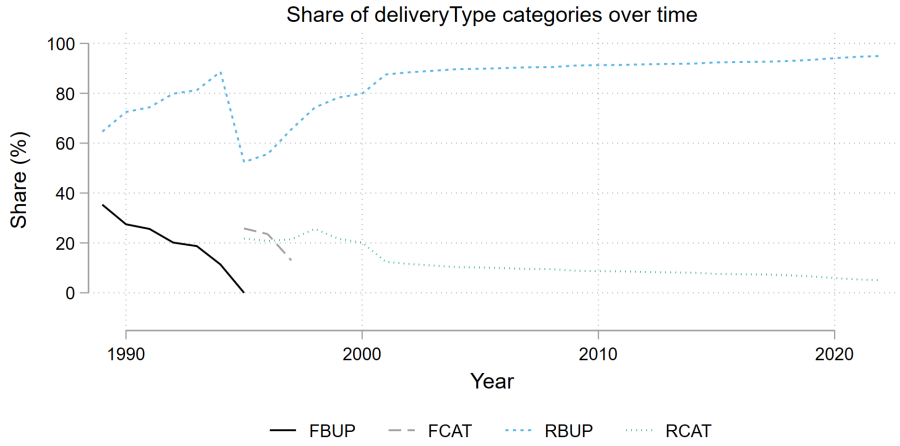


(b) Contingent United States

Figure B.8: Share of insurancePlanAbbr categories over time. Data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).

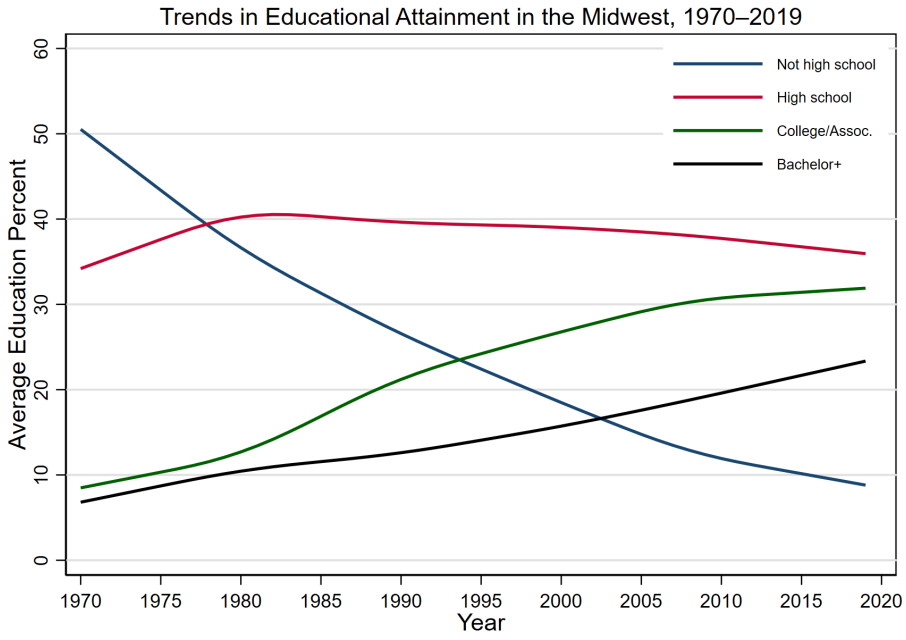


(a) Midwest

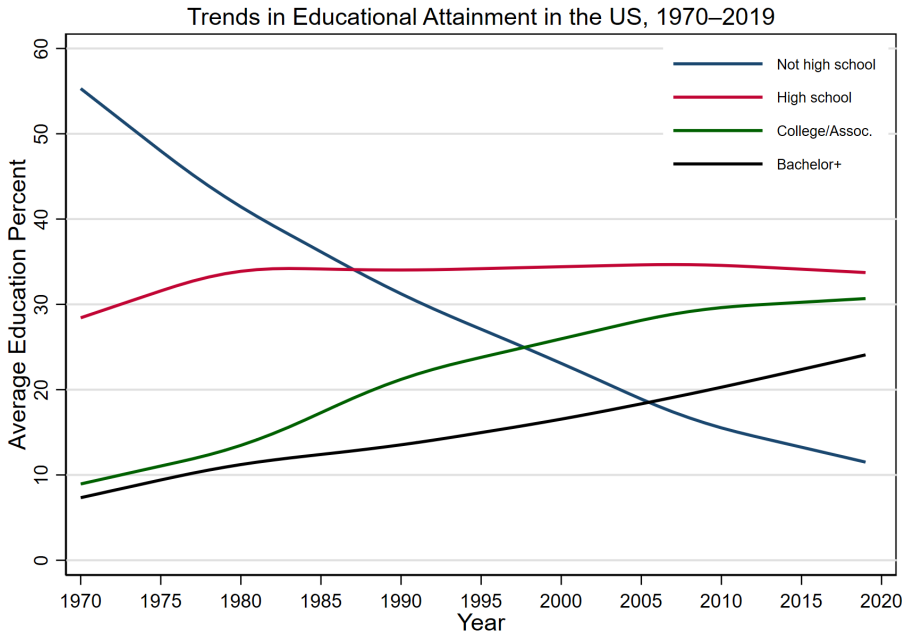


(b) Contingent United States

Figure B.9: Share of deliveryType categories over time. Data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).



(a) Midwest



(b) Contingent United States

Figure B.10: Percent of adults age 25 and older at different levels of education in the Midwest. Data is from USDA Economic Research Service (2025).

C Crop Insurance Subsidies in the World

United States is not alone in its provision of governmental crop insurance. Japan started its public crop insurance in 1939 as a rent insurance, then developed into yield insurance in 1947. In Sweden, compulsory crop insurance has been in effect since 1961 (Wright and Hewitt 1994). Brazil started its public crop insurance in 1973 (Brazil Ministry of Agriculture, Livestock and Food Supply 2016). Mexico began its all-risk crop insurance in 1964. Philippines had a credit insurance type scheme in 1970s, which evolved later into crop insurance in 1981. Sri Lanka's crop insurance has been in operation since 1958. In Mauritius, it began in 1945. India began a nation-wide crop insurance in 1985 (Wright and Hewitt 1994). Iran also started its crop insurance in 1984 (Parliament of the Islamic Republic of Iran 1983). The next three paragraphs are from Chintapalli and Tang (2022).

To sustain small farmers' operations, governments of developing countries offer various programs to support and protect their income. A common farmer-support program is input cost subsidies, wherein a government offers financial assistance to farmers by bearing a fraction of the farmers' cultivation costs. Input subsidy programs (ISPs) are adopted by many emerging economies such as Africa and India. In Africa, smallholder farmers are issued coupons that they can redeem for input packets from both government and private vendors (Jayne et al., 2018). In India, central and individual state government subsidies for seeds, irrigation, power, and fertilizers are offered to farmers (see Hemming et al., 2018, and the references therein).

Besides cost subsidy, an alternative support program is the credit-based minimum support price (MSP; also known as deficiency payment). Under a credit-based MSP, a government will not take any possession of the crop; instead, it will provide a monetary credit to farmers cultivating the crop if the prevailing market price of the crop is below the prespecified MSP. For example, in India, credit-based MSPs are currently offered for eight crops (mostly oilseeds and legumes) for which the Indian government takes no (or very low) possession of crop (Bera, 2017).

Cost subsidies can occur in different forms; for instance, the governments in Mali, Ghana, Nigeria, and other such African countries offer subsidies for seeds and fertilizers to smallholder farmers (Brooks & Wiggins, 2010; Jayne & Rashid, 2013; Jayne et al., 2018). In India, subsidies are offered not only on direct inputs such as seeds and fertilizers (Prasad, 2016) but also on other resources such as power and irrigation by both central and state governments (Bardhan & Mookherjee, 2011; Fan et al., 2008). Xu

et al. (2009) find that fertilizer subsidies may reduce or surge the commercial demand for fertilizers in the area depending on whether the private sector is active or inactive. Chibwana et al. (2012) show that input subsidies can entice farmers to allocate more land to those subsidized crops in Malawi. We refer the readers to Hemming et al. (2018) for a comprehensive review of the various types of input subsidies provided to smallholder farmers.

Mahul and Stutley (2010): Agricultural (crop and livestock) insurance is currently available in more than 100 countries, either as well-developed programs or pilots (figure 1.3; see also appendix A). The majority of high-income countries (58 percent) have well-established agricultural insurance markets. In contrast, only 35 percent of low- and middle-income countries offer such products and programs. The availability of agricultural insurance is particularly low in low-income countries (8 percent). Pilot programs, which reach only a limited number of farmers and herders, are being implemented in various forms (including named-peril crop insurance, index-based crop insurance, or livestock insurance) in eight middle-income and eight low-income countries (table 1.4). to add: map of availability of agricultural insurance in different countries with four categories: yes, no, pilot and unknown.

World Bank conducted a survey of insurance companies in 65 countries in 2008. The key objectives of the survey were to document international experience with public and private agricultural insurance in developed and developing economies and to examine the ways in which governments support agricultural insurance. The focus was on crop and livestock insurance and excluded the forestry or aquaculture insurance.

D Choice of Best Neighbors in KNN Spatial Weights

The KNN weights matrix W_K connects each county-crop panel unit to its k nearest neighbours. As k grows, W_K averages over a wider spatial window: local signals are attenuated but sampling noise in row sums falls. The sweep is used to check whether the sign, magnitude and precision of the focal insurance coefficient (and of the SAR spatial parameter ρ) are robust to the chosen k.

We score each k by two summaries of the six yield models (SAR/SDM/SLX $\times$ OLS/IV):

- $\bar{SE}_{\text{focal}}(k)$ - the mean standard error of the focal coefficient across the six models. Lower is better (more precise).
- $|\Delta\bar{\beta}_{\text{focal}}(k)| = |\bar{\beta}(k) - \bar{\beta}(k-1)|$ – mean absolute change in the focal point estimate

from the previous k. Lower is better (more stable).

Taking the simple average of the within-sweep ranks of these two quantities yields a composite *rank score*; the k with the lowest score balances precision and stability. The ρ_{SAR} columns show the estimated spatial autoregressive parameter under OLS (`spml`) and IV (`spgm`). A ρ that stabilises past some $k^\star$ indicates that further enlarging the neighbour set no longer adds information, and $k^\star$ is a sensible choice.

Quantitative summary:

- For focal = `fracInsuredBuyup`, min-$\bar{SE}$ k is 9; composite rank-score k is 13.
- For focal = `liabilityperacreBuyup`, min-$\bar{SE}$ k is 11; composite rank-score k is 2.

Because the two rules can disagree (precision prefers larger k, stability often prefers middling k), we report both and use the composite score as the recommended default. Per-focal-variable diagnostic tables appear in the corresponding section below.

D.1 Coefficients across k for Fraction Insured (Buyup)

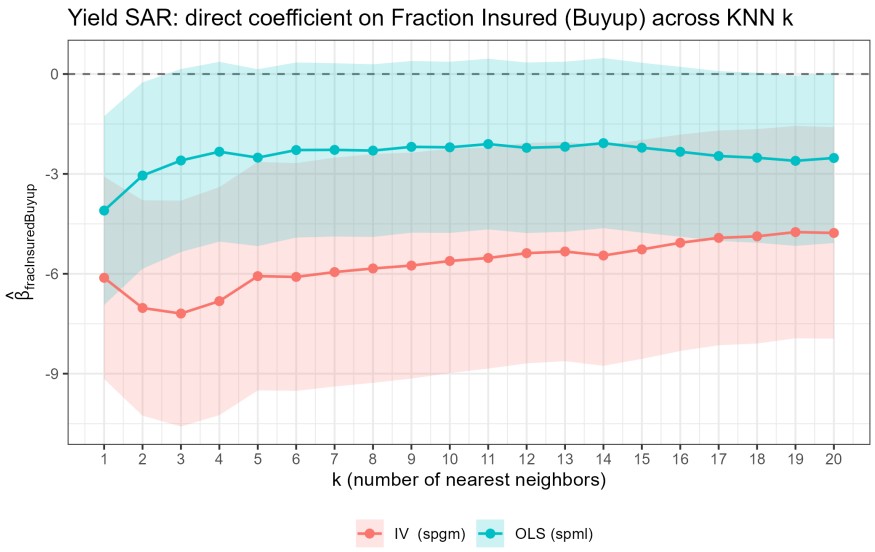


Figure D.1: Yield SAR: direct coefficient on `fracInsuredBuyup` across KNN k with 95% CIs.

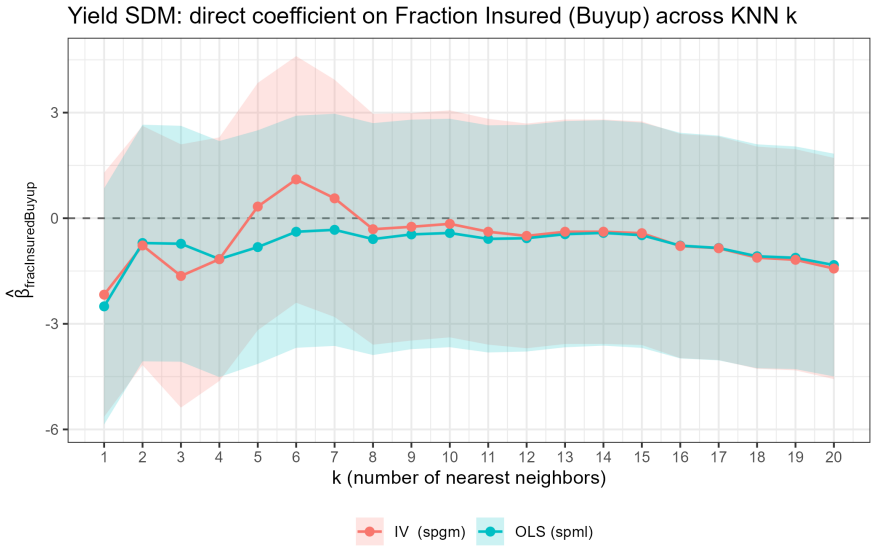


Figure D.2: Yield SDM: direct coefficient on `fracInsuredBuyup` across KNN k.

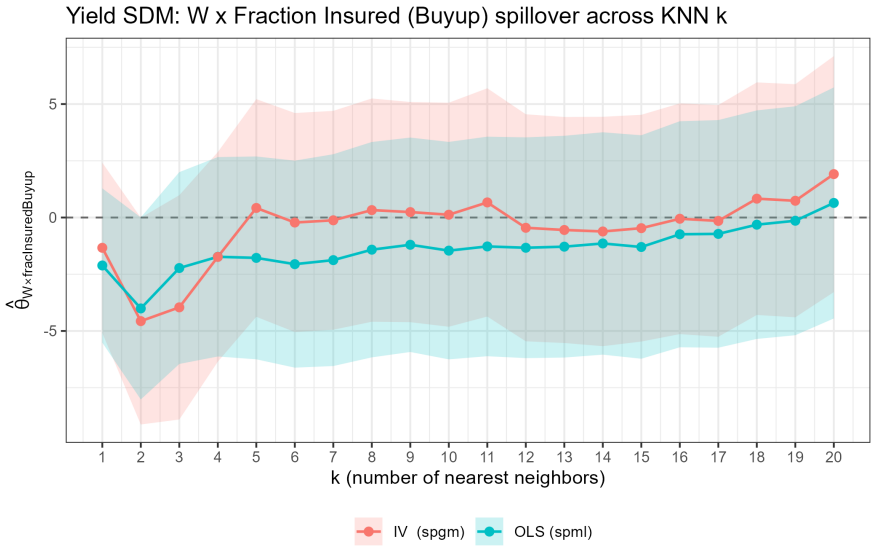


Figure D.3: Yield SDM: $W\times$`fracInsuredBuyup` spillover across KNN k.

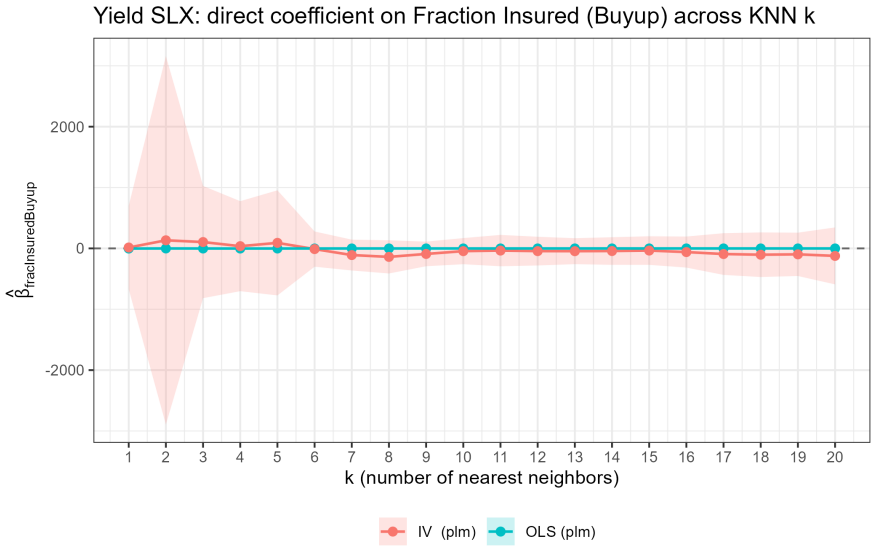


Figure D.4: Yield SLX: direct coefficient on `fracInsuredBuyup` across KNN k.

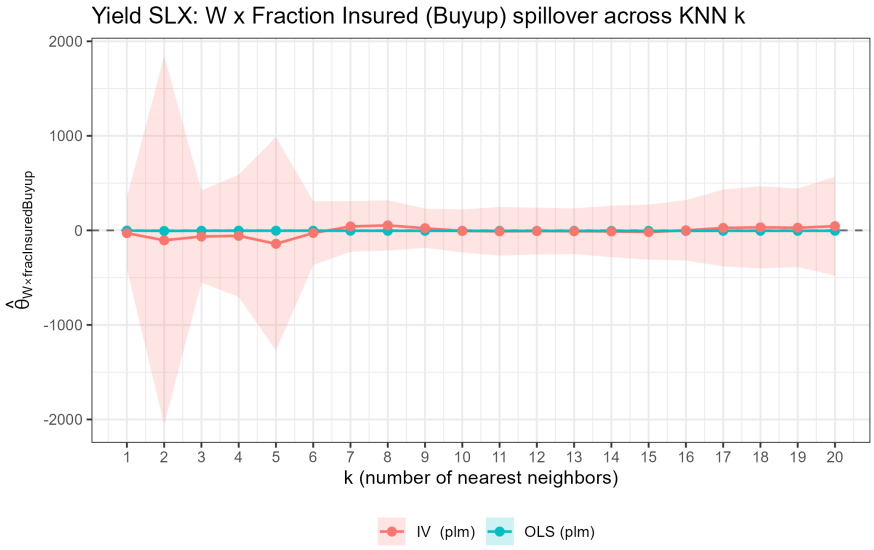


Figure D.5: Yield SLX: $W\times$`fracInsuredBuyup` spillover across KNN k.

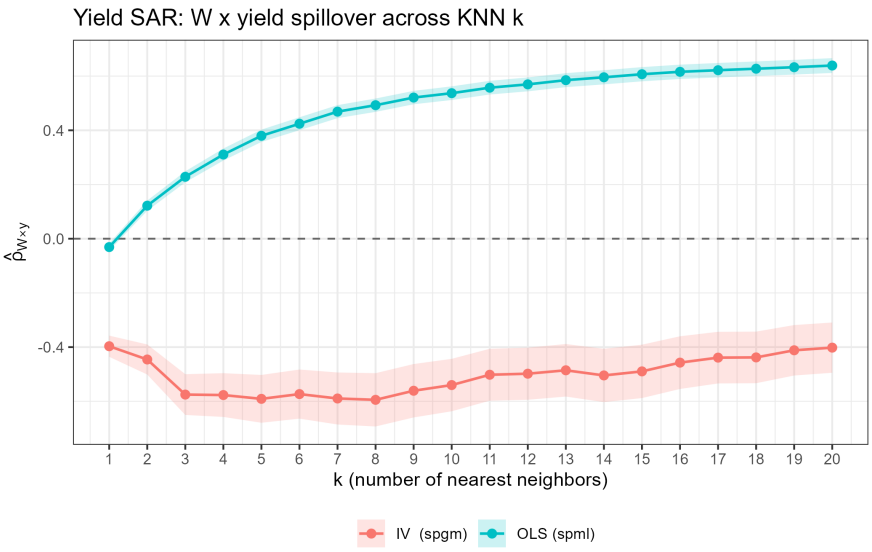


Figure D.6: Yield SAR: $W\times y$ (yield) spillover, i.e. the spatial autoregressive parameter ρ, across KNN k with 95% CIs.

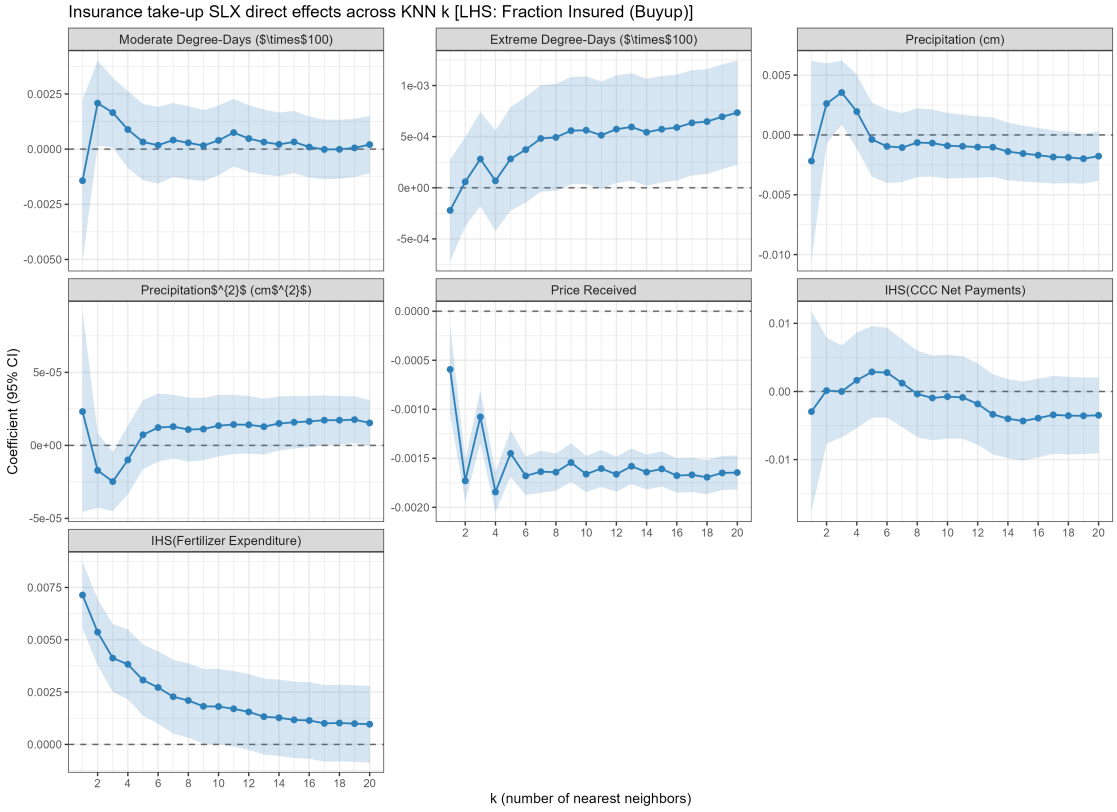


Figure D.7: Insurance take-up SLX *direct* effects across KNN k, faceted by covariate. LHS = `fracInsuredBuyup`.

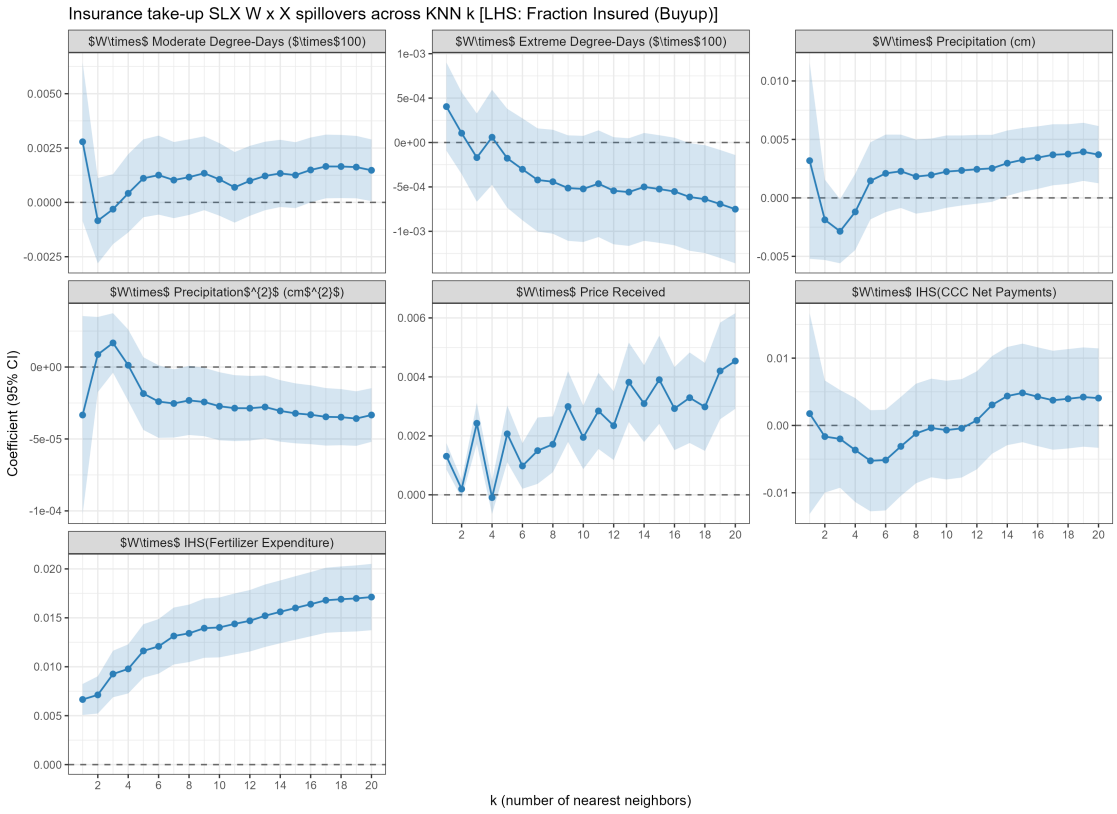


Figure D.8: Insurance take-up SLX *indirect* ($W \times X$) spillovers across KNN k, faceted by covariate. LHS = `fracInsuredBuyup`.

D.2 Best-k diagnostic

Table D.1: Best-k diagnostics for focal = `fracInsuredBuyup`. The *rank score* averages the rank of $\bar{SE}_{\text{focal}}$ and $|\Delta\bar{\beta}_{\text{focal}}|$; lower is better.

k	$\rho_{\text{SAR,OLS}}$	$\rho_{\text{SAR,IV}}$	$\bar{\beta}_{\text{focal}}$	$\bar{SE}_{\text{focal}}$	#sig. (of 6)	rank score
1	-0.0309	-0.3967	-0.4610	59.4909	2	—
2	0.1219	-0.4457	20.0693	258.6899	2	19.5
3	0.2284	-0.5751	15.0843	79.7885	1	14.5
4	0.3106	-0.5768	3.8127	64.2617	1	16.5
5	0.3797	-0.5908	13.3495	74.8081	1	16.5
6	0.4242	-0.5733	-3.4351	25.9938	1	14.5
7	0.4689	-0.5895	-19.8089	22.9256	1	12.0
8	0.4926	-0.5945	-24.9445	24.5566	1	10.5
9	0.5208	-0.5610	-16.7896	18.5435	1	7.5
10	0.5370	-0.5400	-9.0863	19.6018	1	8.0
11	0.5572	-0.5019	-7.5885	23.2630	1	7.5
12	0.5695	-0.4981	-9.0449	21.4002	1	5.5
13	0.5850	-0.4855	-9.1603	19.4729	1	1.5
14	0.5956	-0.5045	-8.7189	20.6302	1	3.0
15	0.6069	-0.4895	-7.4526	21.2956	1	4.5
16	0.6158	-0.4572	-11.6754	23.0097	1	8.0
17	0.6217	-0.4388	-17.1098	30.4542	1	12.0
18	0.6272	-0.4380	-19.0674	32.5010	1	10.5
19	0.6327	-0.4117	-18.0254	31.5889	2	8.0
20	0.6389	-0.4019	-22.3933	41.1008	1	12.0

For `fracInsuredBuyup`, the k minimising the average focal-coefficient standard error across the six yield models is $k = 9$, while the k minimising the combined precision–stability rank score is $k = 13$.

D.3 Coefficients across k for Insurance Liability per acre (Buyup)

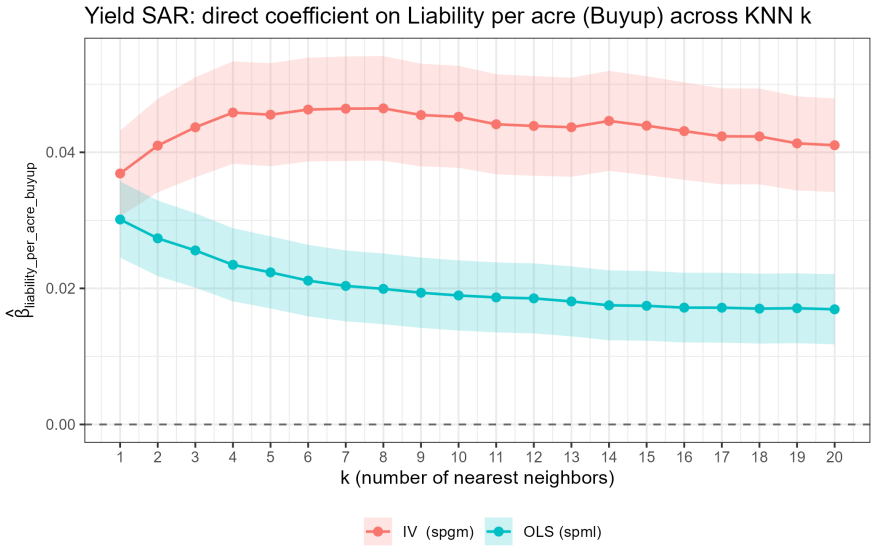


Figure D.9: Yield SAR: direct coefficient on `liability_per_acre_buyup` across KNN k with 95% CIs.

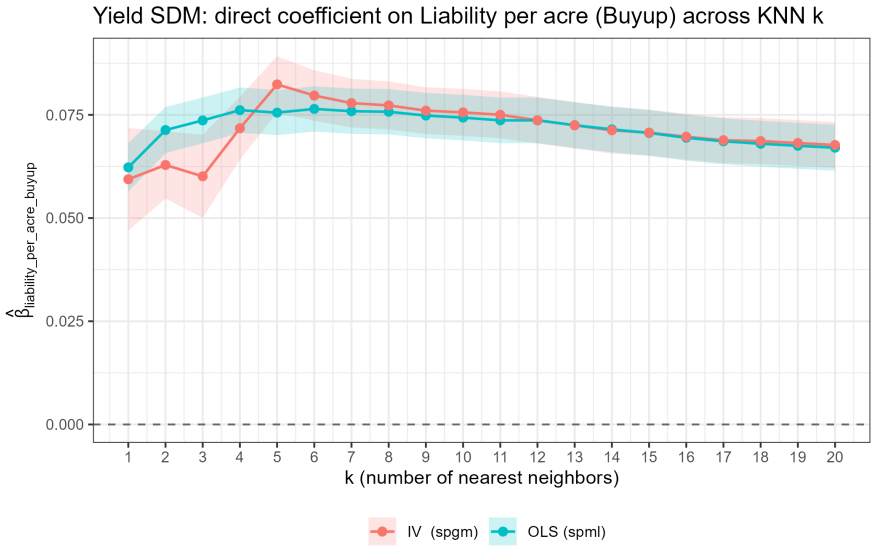


Figure D.10: Yield SDM: direct coefficient on `liability_per_acre_buyup` across KNN k.

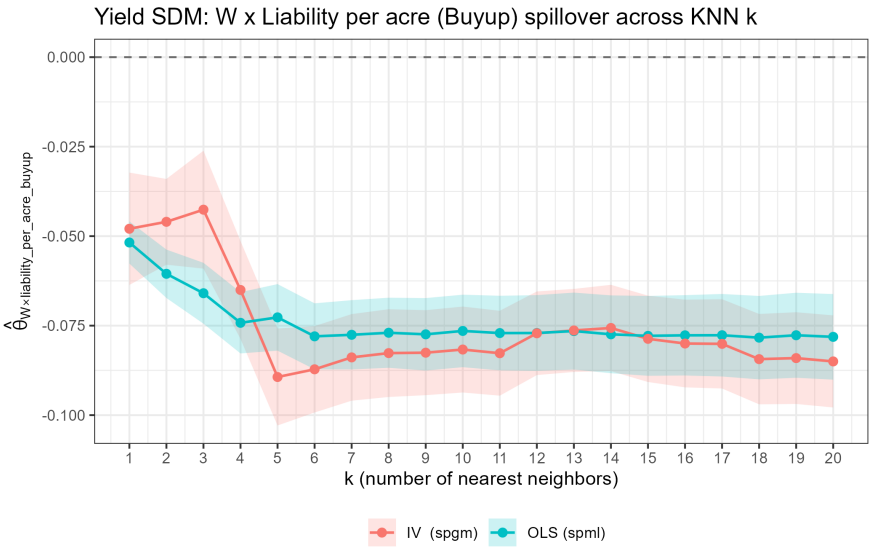


Figure D.11: Yield SDM: $W\times$`liability_per_acre_buyup` spillover across KNN k.

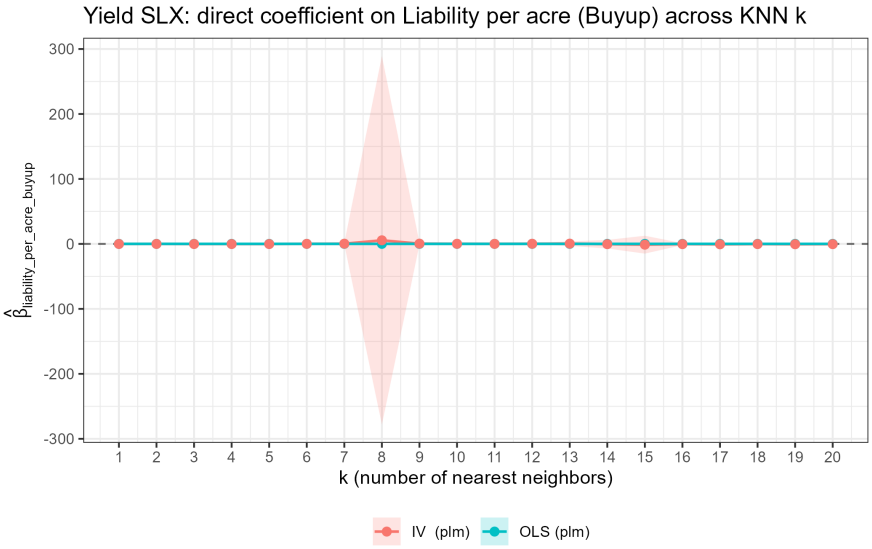


Figure D.12: Yield SLX: direct coefficient on `liability_per_acre_buyup` across KNN k.

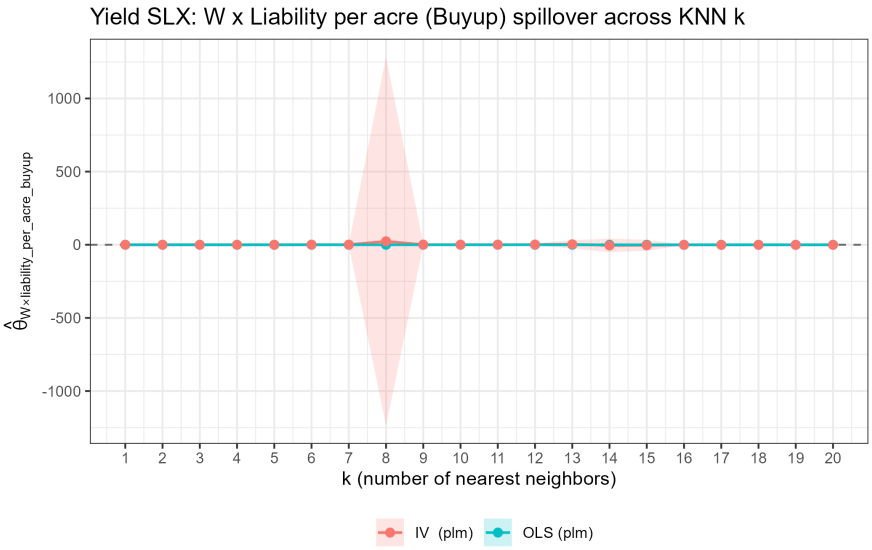


Figure D.13: Yield SLX: $W\times$`liability_per_acre_buyup` spillover across KNN k.

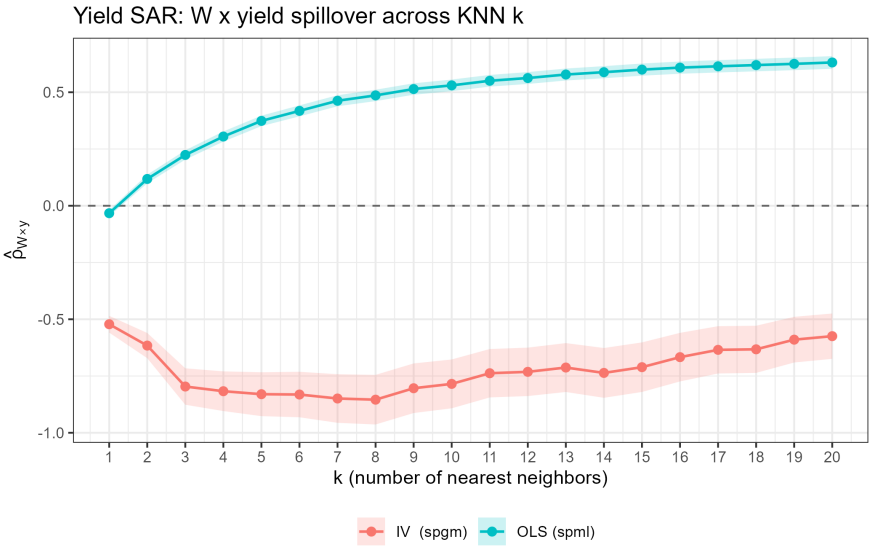


Figure D.14: Yield SAR: $W\times y$ (yield) spillover, i.e. the spatial autoregressive parameter ρ, across KNN k with 95% CIs.

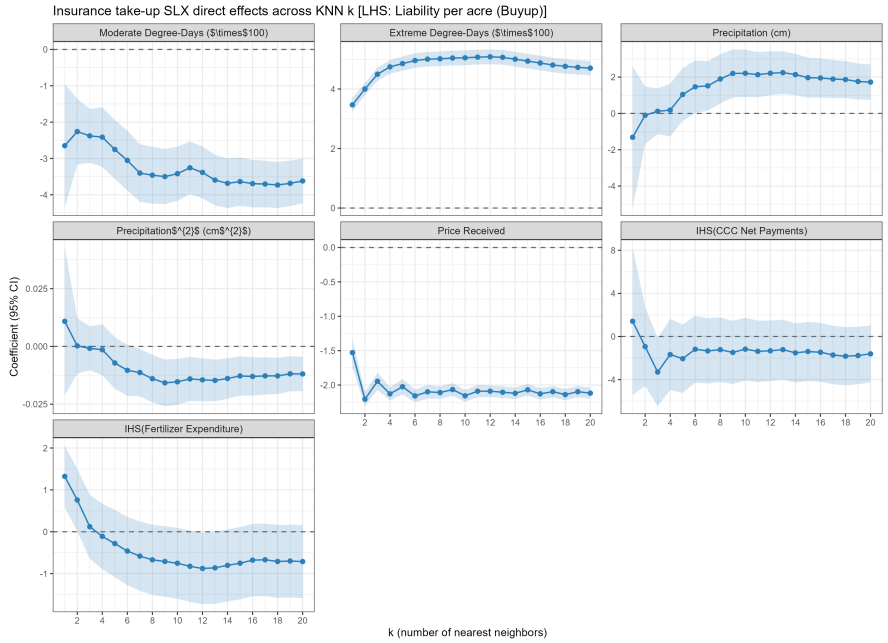


Figure D.15: Insurance take-up SLX *direct* effects across KNN k, faceted by covariate. LHS = `liability_per_acre_buyup`.

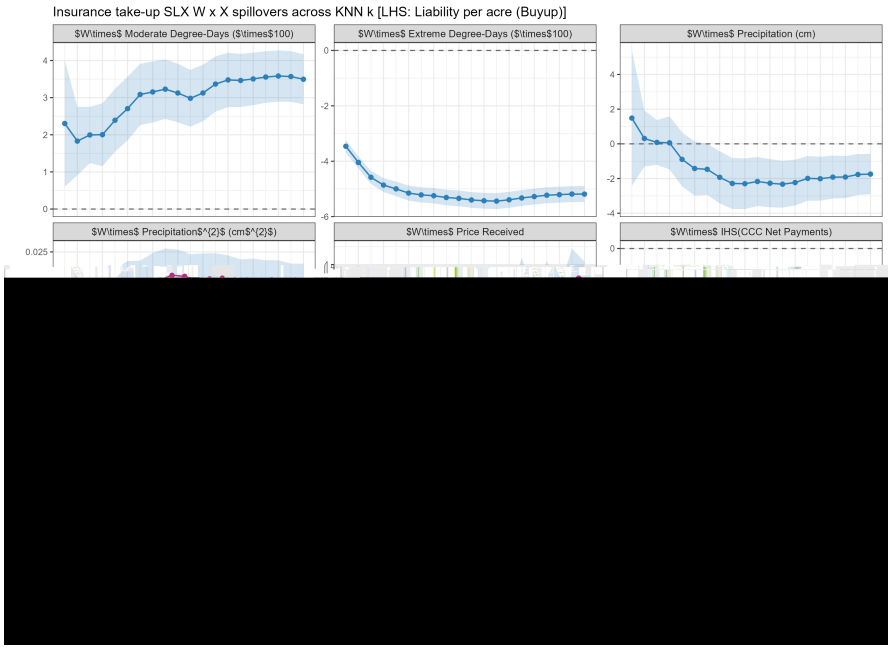


Figure D.16: Insurance take-up SLX *indirect* ($W \times X$) spillovers across KNN k, faceted by covariate. LHS = `liability_per_acre_buyup`.

D.4 Best-k diagnostic

Table D.2: Best-k diagnostics for focal = `liability_per_acre_buyup`. The *rank score* averages the rank of $\bar{SE}_{\text{focal}}$ and $|\Delta\bar{\beta}_{\text{focal}}|$; lower is better.

k	$\rho_{\text{SAR,OLS}}$	$\rho_{\text{SAR,IV}}$	$\bar{\beta}_{\text{focal}}$	$\bar{SE}_{\text{focal}}$	#sig. (of 6)	rank score
1	-0.0328	-0.5219	0.0211	0.0837	5	—
2	0.1180	-0.6159	0.0329	0.0504	5	2.5
3	0.2238	-0.7963	0.0128	0.1189	5	8.0
4	0.3050	-0.8169	0.0137	0.0897	5	3.5
5	0.3739	-0.8301	0.0120	0.1173	5	6.0
6	0.4180	-0.8316	0.0608	0.0773	5	9.0
7	0.4625	-0.8491	0.1035	0.0973	5	9.5
8	0.4861	-0.8542	0.9733	24.1059	5	19.5
9	0.5142	-0.8038	0.1148	0.1053	5	13.5
10	0.5303	-0.7847	0.0735	0.0588	5	7.0
11	0.5506	-0.7377	0.0553	0.0485	5	2.5
12	0.5627	-0.7316	0.0843	0.1043	5	7.5
13	0.5781	-0.7127	0.1143	0.2979	5	12.5
14	0.5884	-0.7363	-0.0244	0.5590	5	17.5
15	0.5997	-0.7111	-0.1529	1.1686	5	17.0
16	0.6085	-0.6669	-0.0155	0.1265	5	14.5
17	0.6143	-0.6347	-0.0562	0.2267	5	13.0
18	0.6197	-0.6324	-0.0075	0.1199	5	12.5
19	0.6251	-0.5899	-0.0423	0.1971	5	12.0
20	0.6312	-0.5745	-0.0190	0.1385	5	10.0

For `liability_per_acre_buyup`, the k minimising the average focal-coefficient standard error across the six yield models is $k = 11$, while the k minimising the combined precision–stability rank score is $k = 2$.

E NASS Crop Data

NASS class conventions used here. NASS classifies corn and soybeans asymmetrically. For corn, NASS publishes planted acres only at the parent-commodity level (corn) because at planting time the eventual use (grain vs. silage) is unknown; harvested acres, production, and yield are then published separately by utilization class (corn, grain and corn, silage). For soybeans, there is no utilization split—all three series exist only at the parent level (soybeans). We therefore keep:

Measure	short_desc kept
planted acres	corn – acres planted, soybeans – acres planted
harvested acres	corn, grain – acres harvested, soybeans – acres harvested
production (bu)	corn, grain – production, measured in bu, soybeans – production, measured in bu

This matches the construction NASS itself uses when it publishes the county-level yield series: yield is built as production divided by harvested-grain acres for corn and as production divided by harvested acres for soybeans. corn, silage is excluded.

Verifying the corn, grain convention. The choice is mechanically forced by how nass structures its data. Across the eight Midwest states for 1980–2025 at survey/county level, the only corn yield series nass publishes in bushels per acre is corn, grain - yield, measured in bu/acre (69,905 county-year cells in our scan, plus 7,015 cells of the secondary corn, grain - yield, measured in bu/net planted acre series introduced in recent years). The only other corn yield series is corn, silage - yield, measured in tons/acre (19,149 cells)- reported in different units (tons rather than bushels per acre), so grain and silage yields cannot even be summed or averaged together. There is no bare corn – yield series at the county level. Therefore, any panel of county-level corn yields in bushels per acre is, by construction, a panel of corn, grain yields.

Annan and Schlenker (2015) regress log corn and soybean yields on weather in their Tables A1–A10, and their data section cites the National Agricultural Statistics Service (1981–2013) as the source. Their corn yield variable is therefore the corn, grain series by the mechanical argument above. Indeed, when we previously held a copy of the original data-withIV-filled-missing.dta, the corn yield column ranged $\approx$ 100-170 bu/ac for Midwest counties, exactly matching the NASS published values for corn, grain - yield, measured in bu/acre. The same corn, grain convention is standard throughout the related agricultural economics literature on weather and yields, including Schlenker and Roberts (2009), Deschênes and Greenstone (2007), Burke and Emerick (2016), and Roberts, Schlenker, and Eyer (2013); see citations below.

F Candidate Instrumental Variables Considered

1- Historical drought exposure in 1980s This instrument captures the role of climate through a belief-based channel (Hsiang 2016, pp. 45–46). I construct county-level measures of historical drought exposure using the NOAA monthly Palmer Drought Severity Index (PDSI) data set (NOAA National Centers for Environmental Information 2026a). PDSI combines information on precipitation, temperature, and local soil water-

holding capacity, making it a standard indicator of cumulative moisture conditions. I use this measure for three reasons. First, PDSI is among the most widely used regional drought indices in the literature (Mishra and Singh 2010, p. 208). Second, because PDSI is designed to capture persistent moisture deficits over time rather than short-lived weather fluctuations, it is especially well suited to agricultural applications (Mishra and Singh 2010, p. 208). Third, while PDSI has known limitations—including its limited treatment of precipitation forms other than rainfall—these drawbacks are less consequential in this setting because the drought episode of interest occurred during summer growing seasons.

A practical issue in using the NOAA county PDSI files is that the county identifiers embedded in the raw dataset are not directly the standard postal county FIPS codes used in most agricultural and demographic data. Instead, the NOAA file reports county identifiers in the NCDC coding system. To harmonize the drought data with my main county-level panel, I use NOAA's `county-to-climdivs` crosswalk (NOAA National Centers for Environmental Information 2026b), which provides, for each county, both the standard postal FIPS code and the corresponding NCDC county identifier, together with the assigned climate division. I merge the raw PDSI file to this crosswalk using the NCDC county identifier and then recover the standard five-digit county FIPS code from the `POSTAL_FIPS_ID` field. This procedure ensures consistency between the NOAA drought series and the county identifiers used in the rest of the analysis. Because NOAA climate divisions and county boundaries do not always align perfectly, NOAA assigns counties to climate divisions based on county centroids; the crosswalk therefore provides the official mapping needed to link the county PDSI series to standard county geographies.

For each county-year, I first compute the growing-season PDSI as the average of June through September monthly values,

$$\text{PDSI}_{ct}^{GS} = \frac{1}{4}\left(\text{PDSI}_{c,t,\text{Jun}} + \text{PDSI}_{c,t,\text{Jul}} + \text{PDSI}_{c,t,\text{Aug}} + \text{PDSI}_{c,t,\text{Sep}}\right).$$

Because more negative PDSI values indicate drier conditions, I define annual drought severity relative to the conventional moderate-drought threshold of -2 as

$$\text{DroughtSeverity}_{ct} = \max\{0,\ -2 - \text{PDSI}_{ct}^{GS}\}.$$

This definition sets drought severity equal to zero in non-drought years and assigns larger positive values to counties experiencing more severe growing-season drought. I then use this annual measure to construct several county-level historical exposure

variables. Exposure to the 1980 drought is defined as $\text{DroughtSeverity}_{c,1980}$. For the late-1980s drought episode, I compute cumulative, average, and peak drought severity over 1988–1990:

$$\text{CumDrought}_{c,1988\text{–}1990} = \sum_{t=1988}^{1990} \text{DroughtSeverity}_{ct},$$

$$\text{AvgDrought}_{c,1988\text{–}1990} = \frac{1}{3} \sum_{t=1988}^{1990} \text{DroughtSeverity}_{ct},$$

and

$$\text{PeakDrought}_{c,1988\text{–}1990} = \max_{t \in \{1988,1989,1990\}} \text{DroughtSeverity}_{ct}.$$

These variables capture distinct dimensions of historical drought risk: one-year exposure in 1980, cumulative drought intensity over 1988–1990, average annual intensity during that period, and the worst single drought year within the late-1980s episode.

As a robustness check, I also construct a historical drought-exposure measure using county-level monthly precipitation from NOAA's nClimDiv county archive. Let $PCPN_{cmt}$ denote precipitation in county c, month m, and year t, and let $\overline{PCPN}_{cm}$ denote the county-specific long-run normal for month m. I define drought exposure during the 1980 drought window as the cumulative June–September precipitation deficit relative to normal:

$$\text{PrecipDeficit}_c^{1980} = \sum_{m \in \{\text{Jun,Jul,Aug,Sep}\}} \left(\overline{PCPN}_{cm} - PCPN_{cm,1980} \right).$$

Larger values of $\text{PrecipDeficit}_c^{1980}$ therefore indicate more severe drought, since they correspond to a larger rainfall shortfall relative to the county's normal summer precipitation. I then interact this baseline exposure with post-reform FCIP policy indicators in the panel. Relative to PDSI, this measure is more transparent because it is directly based on rainfall shortfalls, although it is narrower conceptually because it captures only precipitation rather than the broader moisture-balance conditions summarized by the Palmer drought indices. The June–September 1980 window follows NOAA's dating of the 1980 Drought/Heat Wave (Ross and Lott 2003, p. 6), and the county-level precipitation data are obtained from NOAA's nClimDiv bulk archive. Since agents faced with climate change events for a longer time period are more likely to change their believes (Hsiang 2016, p. 51), I focus on measures of *persistence of exposure* as the primary measures of drought.

Table F.1: Candidate instruments considered

Candidate instrument	Economic rationale	Diagnostic concern	Decision
Base rate × distance to USDA office	Captures differential exposure to administrative access costs when statutory rates change	First stage is positive/weak in some specifications; balance concerns for climate variables	Reported as alternative specification / not used in preferred model
Base rate × 1989 weather-station density	Captures differences in local measurement infrastructure relevant for insurance pricing	Stronger balance concerns; may directly proxy local climate information quality	Not used in preferred specification
Base rate × 1980s drought exposure	Captures pre-period risk exposure interacted with later premium variation	May violate exclusion restriction if historical drought predicts future adaptation or yield trends	Used only as robustness / not used as preferred IV

G Estimation details for the higher-moment tests

This appendix consolidates the procedural details behind Section 7's variance and downside-risk tests (Table 14).

Construction of the purged surprise. For each margin D (buy-up share or liability per acre), the surprise is built in two steps, per equations (13)–(14) in the main text. Step one estimates the IV interaction model of Section 7 by 2SLS and subtracts the fitted insurance contribution, $\hat{\theta}_D D + \sum_k \hat{\theta}_k (D \times \tilde{C}_k)$, from log yield. Step two residualizes the purged outcome on the climate and price controls and the county, crop, and year fixed effects. Every step is restricted to the estimation sample of the step-one model, so observations excluded there (for example through missing instruments) cannot re-enter later steps. The surprise is margin-specific: each margin's risk regressions use the surprise purged of that margin's own mean effect.

Centering convention. The climate variables entering the interactions are demeaned using a common analysis sample (observations with non-missing yield and both coverage margins) rather than each model's exact estimation sample. This is a reparameterization choice only: interaction coefficients and all test statistics are invariant to it.

Two precisions on how the Coverage coefficient should be read. First, because every interacted term is demeaned, the coefficient equals the *average marginal effect* of coverage over the sample climate distribution, since the interaction terms average to zero. Second, precipitation and its square are centered *separately*, so the point at which all centered terms vanish requires $P = E[P]$ and $P^2 = E[P^2]$ simultaneously and therefore does not correspond to a single precipitation realization; the average-marginal-effect reading is the exact one. An equivalent reparameterization that centers precipitation first and then squares the centered variable would instead make the coefficient the effect at exactly $P = E[P]$, at the cost of the average-effect property; it spans the same model and is available as a variant.

Risk regressions and diagnostics. The outcomes u^{*2} and $\max(0, -u^*)$ enter the same interacted 2SLS design as the mean model (five endogenous terms, ten excluded instruments). Because a single joint Kleibergen–Paap statistic understates identification with several endogenous regressors, the table reports the minimum Sanderson–Windmeijer conditional F; the full set is (share margin) 16.2 for D, 43.3 for $D\times$MDD, 37.6 for $D\times$HDD, 16.0 for $D\times$precipitation, and 18.0 for $D\times$precipitation^2, and (liability margin) 16.7, 27.8, 43.8, 25.2, and 26.5 respectively; the sets are identical across the two outcomes within a margin because the first stage is the same. The weakest term is thus the precipitation interaction, while the central $D\times$HDD term is among the strongest. The risk-equation Hansen J statistics treat the generated outcome as observed and are reported as indicative only.

Bootstrap algorithm. The bracketed standard errors in Table 14 come from a county-cluster bootstrap of the entire procedure, with 999 replications and counties (the clustering level) as the resampling unit. Within every draw the algorithm: (i) resamples counties with replacement and assigns new cluster identifiers; (ii) recomputes the climate centering means on the drawn sample; (iii) re-estimates the IV mean model; (iv) subtracts the fitted insurance contribution; (v) residualizes on the mean model's estimation sample; (vi) rebuilds the two risk outcomes; (vii) re-estimates the risk regressions and the high-versus-low-heat contrast. Draws in which any step fails return missing and are counted; all 999 replications succeeded for every statistic. Percentile 95% intervals from the same draws are reported in the replication files and quoted in the text where they matter. The replication count follows the convention of choosing B so that $\alpha(B+1)$ is an integer (Davidson and MacKinnon 2000): with $B = 999$, the percentile interval endpoints fall exactly on the 25th and 975th order statistics of the sorted draws, so no interpolation between draws is needed, and 999 is large enough that simulation noise in the bootstrap standard errors is negligible relative to the standard

errors themselves.

Trimming check. As a guard against data errors in the extreme tails, the share-margin risk regressions are re-estimated after dropping the most extreme 0.5% of $|u^*|$. The level coefficients are sensitive to this trimming while the heat interactions are not, which is why the section's conclusions rest on the interaction pattern.

H Why the benchmark optimum is a corner and the relief optimum tracks the crowd-out rate

This appendix derives the first-order conditions behind Section 9, making precise the envelope argument invoked there and explaining two empirical regularities: the benchmark corner $s^* = 0$ at every cost of public funds, and the near-invariance of the relief optimum to the demand slope found by the bootstrap and the instrument-choice checks.

Implementation of the welfare simulation. For each county–crop–year cell i, the baseline demand index is *calibrated* to reproduce the observed buy-up share exactly,

$$odds_{0i} = \ln\frac{q_{0i}}{1-q_{0i}},$$

with shares clipped to [0.001, 0.999] so the log-odds exists at the boundary; all cell-level heterogeneity in preferences and access is thereby absorbed into the intercept. A counterfactual uniform subsidy s changes cell i's farmer-paid premium by $-P_i(s - s_{0i})$, so the counterfactual index, participation, and surplus are

$$odds_i(s) = odds_{0i} + \beta P_i\,(s - s_{0i}), \qquad q_i(s) = \frac{1}{1+e^{-odds_i(s)}}, \qquad CS_i(s) = \frac{1}{\beta}\ln\!\left[1+e^{odds_i(s)}\right],$$

with β the IV demand slope. Fiscal outlay per acre is $sP_iq_i(s)$ in the benchmark and $sP_iq_i(s) + \rho P_i\big(1-q_i(s)\big)$ in the disaster-relief model, each differenced against its value at the observed s_{0i}. These objects are aggregated with planted-acre weights as in the text and evaluated on the grid $s \in \{0, 0.01, \ldots, 0.95\}$; s^* is the grid argmax. The bootstrap repeats this entire construction, including re-estimating β, within every county-cluster draw.

The envelope property of the log-sum. For one acre, surplus is $CS(s) = (1/\beta)\ln\big(1+e^{v(s)}\big)$ with $v(s) = v_0 + \beta P(s - s_0)$. Differentiating,

$$\frac{dCS}{ds} = \frac{1}{\beta}\cdot\frac{e^{v}}{1+e^{v}}\cdot\beta P = q(s)\,P.$$

The behavioral response $q'(s)$ does not appear: a marginal subsidy increase raises surplus only through the mechanical price cut handed to the q acres already enrolled. This is the envelope theorem in its discrete-choice form (the derivative of expected maximum utility with respect to a price equals demand; it is how the compensating-variation formula of Small and Rosen (1981) is obtained), and its economic content is that the enrollees *induced* by the subsidy were indifferent, so their enrollment creates no first-order private value.

Benchmark. With fiscal cost $C(s) = sPq(s)$,

$$\frac{dW}{ds} = qP - (1+\lambda)\big[qP + sPq'(s)\big] = -\lambda qP - (1+\lambda)\,sP\,q'(s) \;\le\; 0,$$

with equality only at $s = 0$ when $\lambda = 0$. The mechanical terms are a transfer that cancels up to the tax distortion λqP, and the behavioral term is a pure cost: each induced acre brings no surplus but costs the government sP. Hence $s^* = 0$ for every $\lambda \ge 0$, including $\lambda = 0$, as the text asserts.

Disaster-relief crowd-out. Government outlay becomes $G(s) = sPq + \rho P(1-q)$, so $G'(s) = qP + (s-\rho)Pq'(s)$: each induced acre now also *saves* the expected relief ρP. The welfare derivative is

$$\frac{dW}{ds} = -\lambda qP - (1+\lambda)P\,q'(s)\,(s-\rho),$$

and the interior optimum solves

$$s^* = \rho - \frac{\lambda}{1+\lambda}\cdot\frac{q}{q'(s^*)}, \qquad q' = \beta Pq(1-q) \text{ for the logit,}$$

for a representative acre (the computed optimum solves the acre-weighted sum of these conditions across heterogeneous cells). Three properties follow. First, s^* sits *just below* ρ: the subsidy is pushed up to the point where the marginal outlay equals the marginal relief averted, shaded down by the tax distortion on the inframarginal base. Second, the demand slope enters only through the shading term $\frac{\lambda}{1+\lambda}/\big(\beta P(1-q)\big)$, which is small and shrinks as β rises; this is why the bootstrap interval for s^* is far tighter than that

for β, and why the lagged-instrument ($\hat{\beta} = 0.33, 0.41$) and acre-weighted ($\hat{\beta} = 0.20$) alternatives move the optimum only between 0.37 and 0.40. Third, as $\beta \to \infty$, $s^* \to \rho$. Evaluated at sample means ($P \approx \$45$ per acre in 2025 dollars, $q \approx 0.68$, $\hat{\beta} = 0.29$, $\lambda = 0.25$), the shading term is about 0.05, so the formula predicts $s^* \approx 0.37$, within two points of the exact acre-weighted grid optimum 0.39; and it reproduces the ordering of every robustness check (s^* rising from about 0.35 to 0.39 as β rises from 0.20 to 0.41, always below $\hat{\rho} = 0.42$). loso.tex

I USDA RMA Base and Reference Rates for Insurance Premium

In the modern era, A01010 carries two parallel rate concepts, only one of which is populated for any given insurance offer: base rate and reference rate. Base rate is used for crops priced via a lookup table: specialty, perennial, PRF/forage, nursery, and most area-based plans. The producer rate is read directly from this field. Reference rate on the other hand is used for crops priced via continuous rating: corn, soybeans, wheat, cotton, grain sorghum, barley, oats, peanuts, dry beans, hemp. The producer rate at yield y is

$$r(y) \;=\; \texttt{reference_rate} \;\times\; \big(y \,/\, \texttt{reference_amount}\big)^{\texttt{exponent_value}}$$

where the exponent value is a non-positive value reported in modern A01010. Across the 9.67M continuous-rated offers in this panel, the median exponent is -0.97 and the 10^{th}–90^{th} percentile range is $[-1.90, 0.00]$, with crop-level medians of -0.89 (corn), -1.29 (soybeans), -1.06 (wheat), -0.80 (cotton). This negative exponents mean producers with above-reference yields pay a lower rate than the reference rate. An exponent of exactly 0 collapses the formula to a flat rate. The exponent is RMA's empirical elasticity of the loss-cost ratio with respect to yield, estimated per county-crop group from a long history of unit-level losses and smoothed across nearby counties. Specifically, it is the parameter β_1 in the unit-level regression in Keith H Coble et al. (2010, Eq. 6.2, p. 64):

$$LCR_i \;=\; \beta_0 \cdot \big(\text{Rate Yield}_i \,/\, \text{Reference Yield}\big)^{\beta_1} + \varepsilon$$

A01010's `reference_rate` is therefore the rate *at the reference yield*, at the reference coverage level (uniformly 0.65 for these crops).

Both rates are stored as dollars of premium per dollar of liability, *before* the federal premium subsidy and *before* any coverage-level adjustment from A01040. They are also

reported at the reference coverage level of 0.65. To take this 0.65 reference rate to the rate a producer actually faces, one would need to do coverage-level adjustment (U.S. Department of Agriculture, Risk Management Agency 2025, p. 13)[15]:

$$r(c,u,y) \;=\; r_{\text{base}}(y) \;\times\; \underbrace{\text{CLDF}(c)}_{\text{coverage-level differential}} \;\times\; \underbrace{\text{URF}(u)}_{\text{unit residual factor}}, \qquad \text{CLDF}(0.65) \equiv 1.$$

where the terms are:

$r(c,u,y)$ the producer's premium rate (dollars of premium per dollar of liability) for a unit with approved yield y that elects coverage level c and unit structure u. This is the quantity that, after applying the federal subsidy share, the producer actually pays on.

$r_{\text{base}}(y)$ the yield-adjusted base rate from step 1 — i.e. the A01010 `reference_rate` re-priced to the unit's own yield y, but still at the reference coverage level 0.65 and at the most disaggregated (basic) unit.

c the elected *coverage level*, $c \in \{0.50, 0.55, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85\}$ — the share of the yield (or revenue) guarantee the producer chooses to insure.

$\text{RDF}(c)$ the *Rate Differential Factor* (RMA's name): the multiplicative factor that scales the rate from the reference level 0.65 to the elected level c. By construction $\text{RDF}(0.65) = 1$; higher coverage is riskier per dollar of liability, so $\text{RDF}(c) > 1$ for $c > 0.65$ and < 1 for $c < 0.65$. Published in ADM record A01040 (Coverage Level Differential).

u the *unit structure* the producer elects: basic (BU), optional (OU), enterprise (EU), or whole-farm (WU). Finer units (BU/OU) pool less risk and carry a higher rate; coarser units (EU/WU) pool more and carry a lower rate.

$\text{URF}(u)$ the *Unit Residual Factor* (RMA's name; the exhibit also defines an Enterprise Unit Residual Factor and Whole Farm Unit Residual Factor): the multiplicative factor for the elected unit structure relative to the basic-/optional-unit reference. Also published in ADM record A01040.

Figure I.1 gives the full column layout of a modern A01010 row and shows where the two rate concepts sit (column 36 for `base_rate`; column 27 for `reference_amount`, 29 for `reference_rate` and 30 for `exponent_value`).

[15] *Current Year Base Premium Rate'* row, which reads, verbatim,

```
Round(Current Year Base Rate * Rate Differential Factor * Unit
Residual Factor, 8)
```

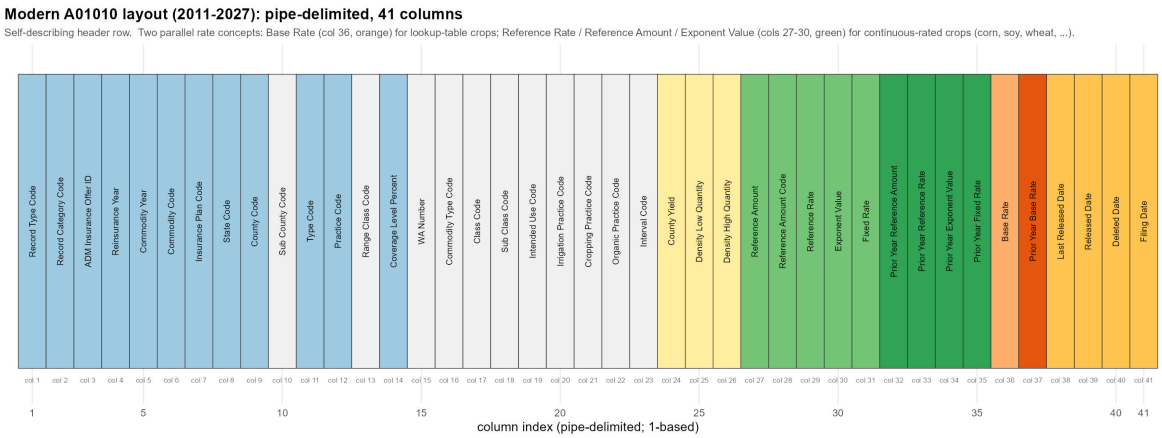


Figure I.1: Modern A01010 layout (2011–2027): pipe-delimited text with a self-describing header row, 41 columns. Colour key: blue = identifier keys (commodity, location, plan, type, practice, coverage level); grey = optional qualifiers; light yellow = yield / density; light green = continuous-rating block (`reference_amount`, `reference_rate`, `exponent_value`, plus the `fixed_rate` loading) and its prior-year counterparts in dark green; light orange = lookup-table `base_rate` (with prior-year counterpart in dark orange); gold = audit-trail dates. The pipeline reads columns 6–12 plus 14 as the offer key and columns 27, 29, 30, 36 as the rate primitives.

The legacy era (1996-2010) uses three different fixed-width record layouts; year 1996, year 1997, and years 1998–2010. Figures I.2, I.3, and I.4 depict these layouts, respectively. I reframed these data into the modern A01010 schema. Every legacy year's directory on RMA's public ADM mirror contains a small companion archive alongside the data zip:

https://pubfs-rma.fpac.usda.gov/pub/References/actuarial_data_master/{YYYY}/{YY}admrec.zip

For example, for 2005, this is `05admrec.zip`, which unpacks into three plain-text files: a table of contents (2005V7TO.TXT), a record layout document (2005V7RE.TXT) giving for every field the byte offset, length, and COBOL picture clause, and a changes-vs-prior-year file (2005V7CH.TXT). This is the canonical and only documentation needed to read `R1{YY}YTD`. By pulling the `admrec.zip` for 1996, 1997, 1998, 1999, 2000, 2005 and 2010 and diffing the field offsets, I discovered three distinct layouts within the legacy era. Figures I.2–I.4) plus **one sub-record overlay** within layout B (Figure I.5) illustrates the layouts.

Layout A1 (1996). 2 digits crop-year, no insurance plan ID (I.2).

Layout A2 (1997). 2 digits crop-year, with insurance plan ID (I.3).

Layout B (1998–2010). 4 digits crop-year, with insurance plan ID (I.4).

- *Continuous-rating overlay* in years 2001 on wards when record type suffix is letter C. In this case, the same header (bytes 1–23) is followed by a different payload in bytes 24–44: `REFERENCE-YIELD` + `REFERENCE-RATE` + `EXPONENT` + `EXPONENT-SIGN` + `FIXED-RATE-LOAD`. This is how corn, soybeans, wheat, sorghum, etc. are priced from the 2001 reinsurance year onward and is the legacy-era analogue of A01010's modern `reference_rate` columns (Figure I.5).

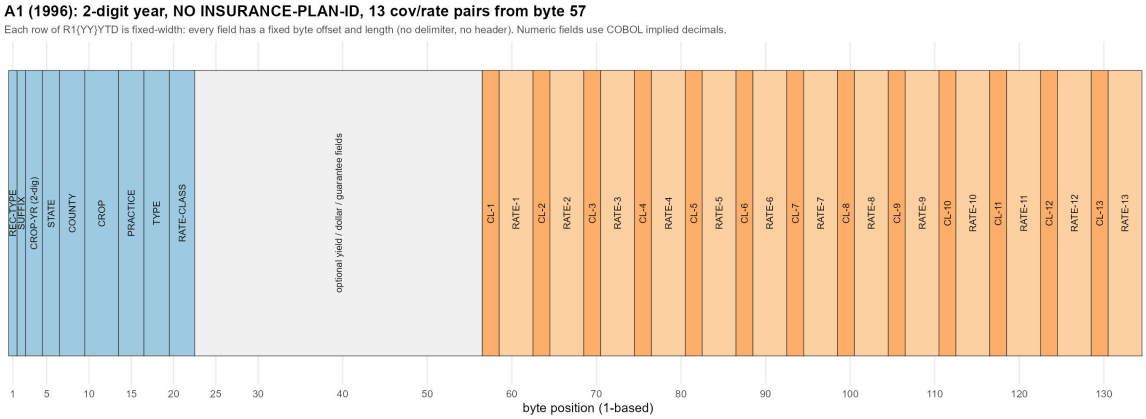


Figure I.2: Layout A1 (1996). Colour key: light blue = shared header field; grey = optional yield / dollar / guarantee payload; alternating orange = the inline cov/rate pairs.

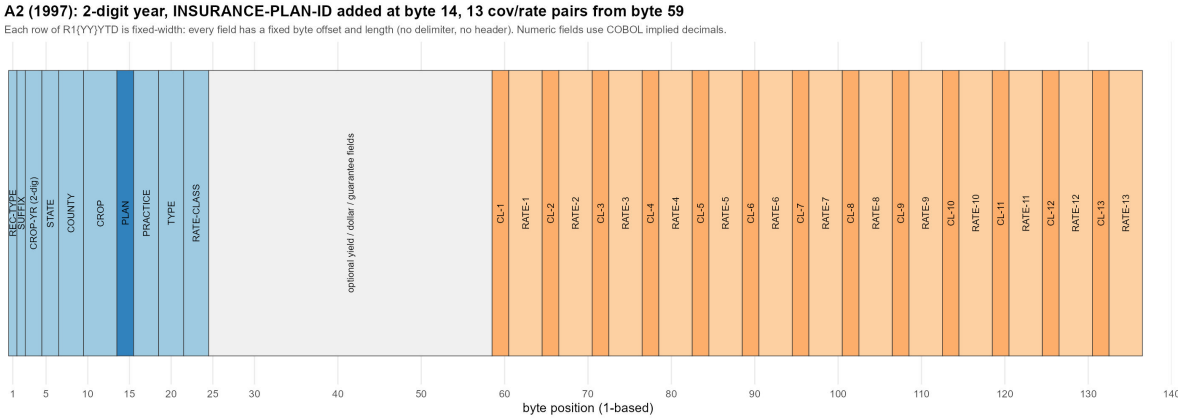


Figure I.3: Layout A2 (1997). insurance plan ID added at byte 14 (darker blue).

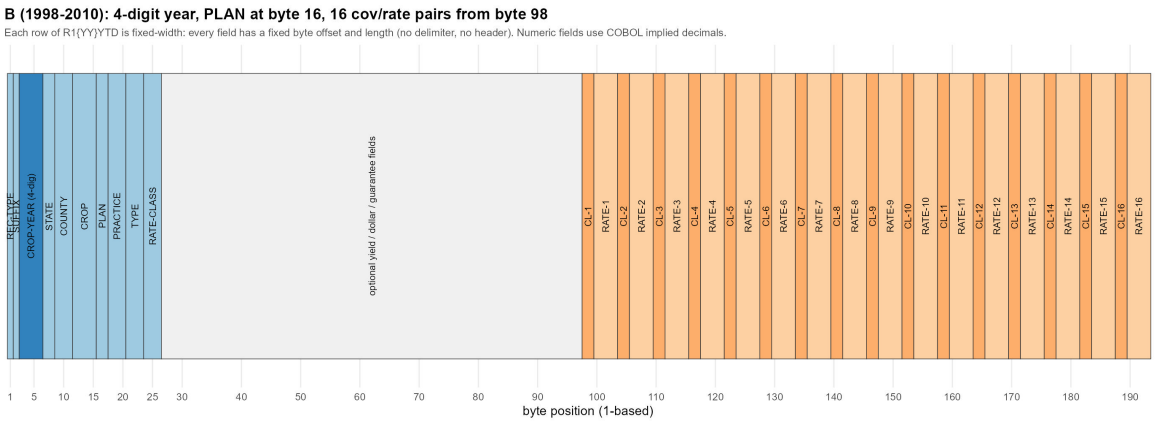


Figure I.4: Layout B (1998–2010). This layout is stable for 13 reinsurance years (1998 through 2010).

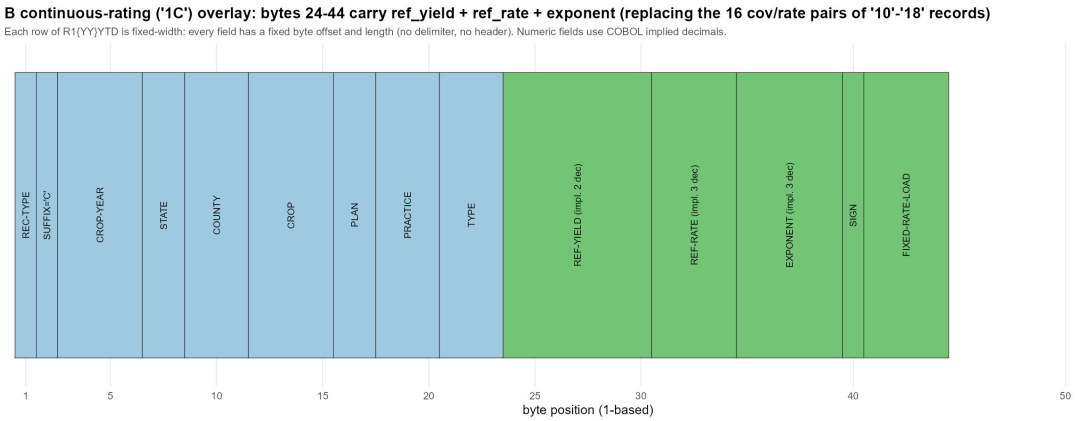


Figure I.5: Layout B continuous-rating for crops like corn and soybeans.

Structural break at 2001. Continuous rating for the major field crops (corn, soybeans, wheat, sorghum, etc.) was introduced around the 2001 reinsurance year as a new sub-record type, shown in Figure I.5). Before 2001, A01010 stored a vector of base rates stratified by *yield band* (low–high yield interval) inside the standard record. After 2001, those crops migrated to a single `reference_rate` per cell plus a continuous-rating formula $r(y) = \texttt{reference_rate} \cdot (y/\texttt{reference_amount})^{\texttt{exponent_value}}$.

For the legacy ADM years 1996–2010, I converted the fixed-width RMA record-type-1 files into the modern A01010 BaseRate schema using a custom Node.js parser. The parser applies year-specific column layouts, extracts standard 65% coverage base rates and continuous-rating reference parameters where available, and writes pipe-delimited files with the same 41-column header as modern A01010 files.

Statutory Subsidy Rates Set by the Congress Over the course of 1989-2027, Congress authorized changes in FCIP subsidies through its legislation. To turn an

A01010 rate (gross of subsidy, what RMA charges the offer) into a producer-paid rate (what the farmer actually writes a check for) we need the federal premium subsidy share for that year, coverage level, and unit structure:

$$\text{rate}^{\text{net}}_{c,k,t,\ell,u} \;=\; \text{rate}_{c,k,t,\ell} \;\cdot\; \big(1 - s_{t,\ell,u}\big).$$

The subsidy share s is set by Congress and applies uniformly to every county and crop in the country.

Data source of 2011–2027 (modern): machine-readable. Lives inside every `$Y_ADM_YTD.zip` as `$Y_A00070_SubsidyPercent_YTD.txt`, a pipe-delimited table keyed by (insurance plan, unit structure code, coverage level percent, coverage type code), with subsidy percent as the value.

Data source of 1989–2010 (pre-ADM / legacy): not in the ADM at all. Hand-coded into `tools/subsidy_pre2011.csv` directly from statute, with a citation to the relevant section of 7 U.S.C. § 1508 and the Public Law that enacted each regime (1980 Federal Crop Insurance Act, 1994 Federal Crop Insurance Reform Act, 2000 Agricultural Risk Protection Act, 2008 Farm Bill).

Construction of the pre-2011 table. Because the legacy ADM does not publish the premium subsidy schedule as data, I assembled the 1989–2010 portion was assembled manually. The statutory subsidy share is transcribed from the Federal Crop Insurance Act — 7 U.S.C. § 1508, with buy-up subsidies in subsection (e) and catastrophic coverage in subsection (b)(5) — as amended by the Public Law that established each regime: the 1980 FCIA (P.L. 96-365), the 1994 Reform Act (P.L. 103-354), the 2000 ARPA (P.L. 106-224), and the 2008 Farm Bill (P.L. 110-246). Every row carries its citation in a dedicated column. I then cross-checked the transcribed values two ways: against secondary compilations (Plastina 2021; O'Donoghue 2014; Rosch 2021) and, at the 2011 era boundary, against RMA's own machine-readable A00070 file, whose 2011 buy-up figures coincide with the ARPA values coded here.

- **1980-Act era, 30% (1989–1994):** O'Donoghue (2014, p. 3) states the 1980 Act offered subsidies covering up to 30 percent of the total premium, matching the 0.30 coded at the 50–65% levels.
- **Catastrophic (CAT) = 100%:**O'Donoghue (2014, p. 3) describes CAT as the fully subsidized tier created by the 1994 Reform Act, first available 1995 — matching the 1.00 coded from 1995 on and the absence of CAT before 1995.

- **2011 era boundary:** The data source changes at the 2010/2011 seam: 1989–2010 is hand-coded from statute, while 2011–2027 is read directly from RMA's machine-readable A00070 file. If the two sources are mutually consistent, the last hand-coded year (2010) and the first machine-read year (2011) should report the *same* subsidy schedule, since no legislation changed rates between them. They do at every coverage level and for both unit structures, confirming the two halves of the panel stitch together with no artificial jump at the seam.

- **1994-Act buy-up schedule (1995–2000): 0.55/0.46/0.38/0.417/0.32/0.235/0.17/0.13 at 50–85%:** confirmed against two independent sources.
  (a) U.S. Government Accountability Office (2005, p. 17) lists the pre-ARPA basic/optional subsidy as 55/46/38/42/32/24/17/13 across the 50–85% levels (whole percentages) — matching every cell coded here to rounding, including the deliberately non-monotone 65% "buy-up bump" (42% at 65% exceeds 38% at 60%, the 1994 Act's device to pull producers from CAT into buy-up coverage). (CRS R46686 does *not* corroborate this; its only subsidy table, Table 4 on p. 23, is the *current* schedule and contains no historical pre-ARPA figures. Likewise CRS R40532 was not independently verified to contain the pre-ARPA table.)
  (b) Bruce A Babcock and Hart (2005, p. 4) gives the precise pre-ARPA subsidy at two coverage levels in narrative form (no full table): "premium subsidy rates were increased from 41.7 to 59 percent for 65 percent coverage and from 23.5 to 55 percent for 75 percent coverage." These precise figures (0.417, 0.235) confirm GAO's rounded 42% / 24% and are what is coded in `subsidy_pre2011.csv` at the 65 and 75% cells; the remaining cells use GAO's whole-percent figures (the 70% precise value is not stated in either source and is kept as 0.32; the published 0.55/0.46/0.38/0.17/0.13 at 50/55/60/80/85% likewise follow GAO).
  Three further notes: (i) Contrary to a common impression, the 80% and 85% buy-up levels were *not* created by ARPA — they existed pre-ARPA with very low subsidies (17% and 13%); ARPA raised them to 48% and 38%. (ii) Pre-ARPA, the dollar subsidy was kept constant for all coverage levels between 65% and 85%, so the declining percentage subsidies (42, 32, 24, 17, 13) are inversely proportional to the rising premium rate. Below 65%, the dollar subsidy fell, so the percentage rose (55, 46, 38 at 50, 55, 60). (iii) The GAO table footnote flags that for the **1999 and 2000 crop years**, the *actual* subsidy producers received was higher than the statutory baseline shown: emergency supplemental acts provided an additional 30% premium discount in 1999 and 25% in 2000. Babcock & Hart (p. 6) corroborate the 1999 discount (citing 25%). These ad hoc discounts are *not* coded as separate rows in `subsidy_pre2011.csv`, but they are flagged

in the `notes` column for 1999 and 2000. Users doing demand or producer-paid-rate analysis for those years should apply the emergency discount on top of the statutory subsidy, or work directly from SOB realized subsidy / premium.

Statutory regimes The schedule has four discrete regimes within 1989–2027, each established by a specific piece of legislation:

Regime	Years	Authority
FCIA_1980	1989–1994	P.L. 96-365 (Federal Crop Insurance Act of 1980); 30% of premiu
FCIRA_1994	1995–2000	P.L. 103-354 (1994 Federal Crop Insurance Reform Act); 7 U.S.C
ARPA_2000	2001–2008	P.L. 106-224 (Agricultural Risk Protection Act of 2000); 7 U.S.C
Farm_Bill_2008	2009–2010	P.L. 110-246 (Food, Conservation, and Energy Act of 2008); add
ADM_A00070 (pre-OBBBA)	2011–2025	RMA ADM A00070 file; buy-up schedule unchanged from ARPA
ADM_A00070 (OBBBA)	2026–2027	2025 reconciliation "One Big Beautiful Bill Act"; +5 pts most cov

Table I.1: Statutory regimes underlying the subsidy schedule. The 1980-Act era (1989–1994) is provided for SOB-based work only — no ADM rate exists to multiply against before 1996. All buy-up subsidies live in 7 U.S.C. § 1508(e); the catastrophic (CAT) level (created by the 1994 Reform Act, first available 1995) is at 100% subsidy per 7 U.S.C. § 1508(b)(5)(B).

Which legislation did and did not move the schedule. Not every Farm Bill changed the premium-subsidy percentages. The percentage changes occurred at five points: the 1980 FCIA, the 1994 Reform Act, the 2000 ARPA, the 2008 Farm Bill (enterprise units), and the 2025 reconciliation act (effective 2026). Other laws touched crop insurance without changing the core individual-plan buy-up percentages: the 1990 Farm Bill; the 1996 FAIR Act (created RMA, ended the linkage requirement); the 2002 Farm Bill (kept the ARPA schedule); the 2014 Farm Bill (added new products — SCO at 65%, STAX — and the beginning-farmer add-on, but left YP/RP/RP-HPE buy-up at ARPA/2008 levels); and the 2018 Farm Bill. This regime-by-regime history is documented secondarily in O'Donoghue (2014, ERS ERR-169, subsidy-rate table by coverage level and regime), the ERS "Title XI: Crop Insurance Program Provisions" farm-bill-by-farm-bill summary, and Glauber (2013); see Section **??**. These predate the 2025 reconciliation act, so the 2026 change is sourced separately to RMA Manager's Bulletin MGR-25-006. For the modern era this is handled automatically: because 2011–2027 subsidies are read directly from the RMA A00070 file, any legislative change is reflected in the data without us re-coding it — which is why the schedule shows the 2026 step from 0.59 to 0.64 at the 65% level (the OBBBA increase) with no manual intervention.

Discrete shocks for identification. Two large, plausibly exogenous subsidy jumps anchor empirical work: the **ARPA 2000** jump (0.42 → 0.59 at the 65% buy-up level,

2000→2001) and the **2008 Farm Bill** enterprise-unit introduction (a ≈0.21 jump at 65% for producers who could form enterprise units). The **2026 OBBBA** +5-point increase adds a third. Having more than one shock allows the Borusyak–Hull–Jaravel shock-orthogonality approach rather than leaning on a single break.

In Figure I.6, each dot is one county-year (a mean across whatever type/practice combinations are filed); the heavy black line is the unweighted across-county mean by year. Corn and soybeans are shown on a common scale to make the side-by-side comparison direct. Figure I.7 shows the statutory federal premium subsidy schedule that applies in the midwest 8 (and, in fact, nationally — the schedule is set by Congress and uniform across counties). Two panels stack the basic / optional unit schedule (which exists across the whole 1989–2027 window) and the enterprise unit schedule (first available in 2009, per the 2008 Farm Bill). Each line is one coverage level (50–85%). The discrete steps correspond to legislative changes: 1980 FCIA (30% baseline up to 65%), 1994 Reform Act (effective 1995), 2000 ARPA (effective 2001), 2008 Farm Bill (adds the enterprise panel from 2009), and the 2025 OBBBA reconciliation act (+5pp most levels, effective 2026).

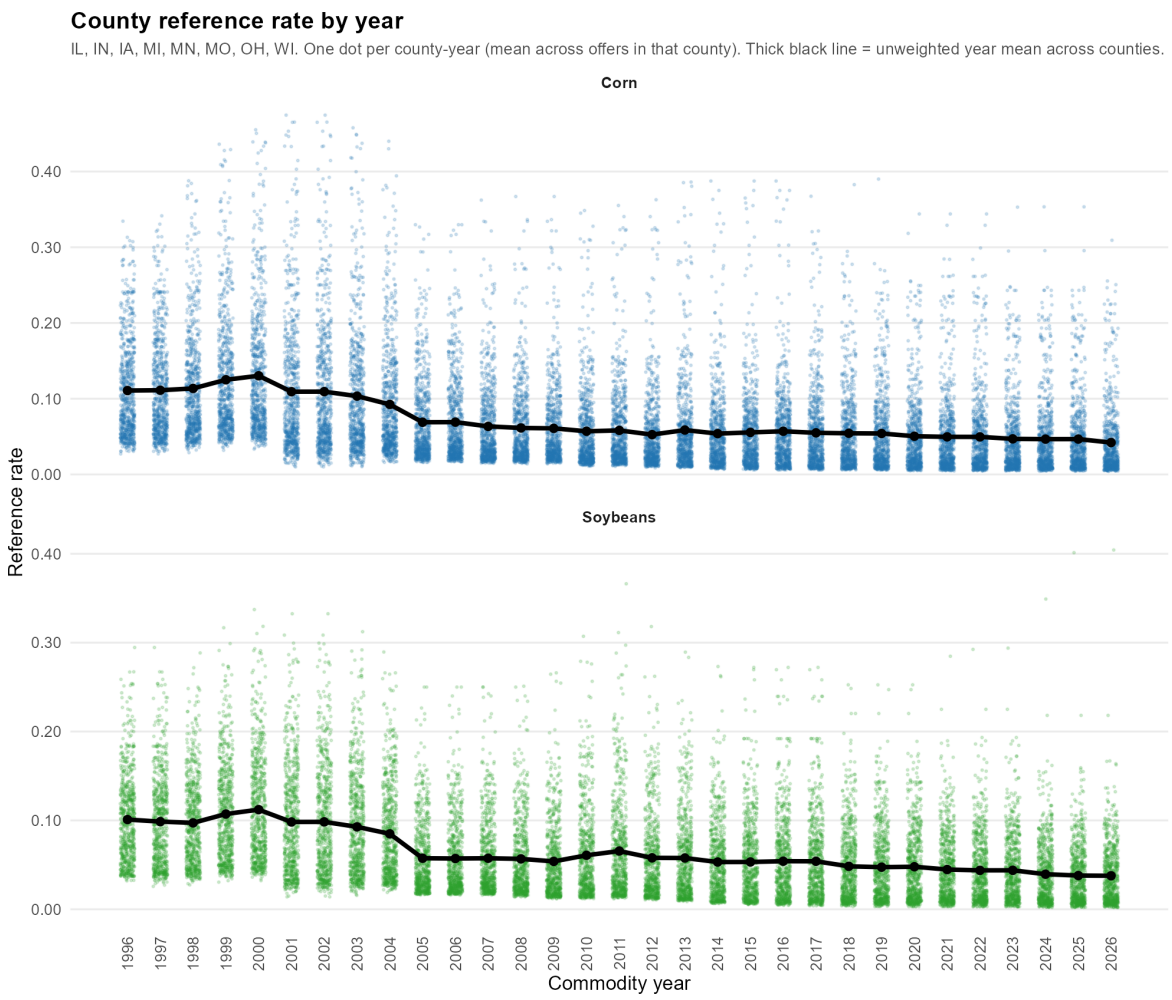


Figure I.6: Corn and soybeans published rate by year, midwest 8, 1996–2026, on a common scale. Rate concept switches from base rate (1996–2000) to continuous-rating reference rate (2001–2026).

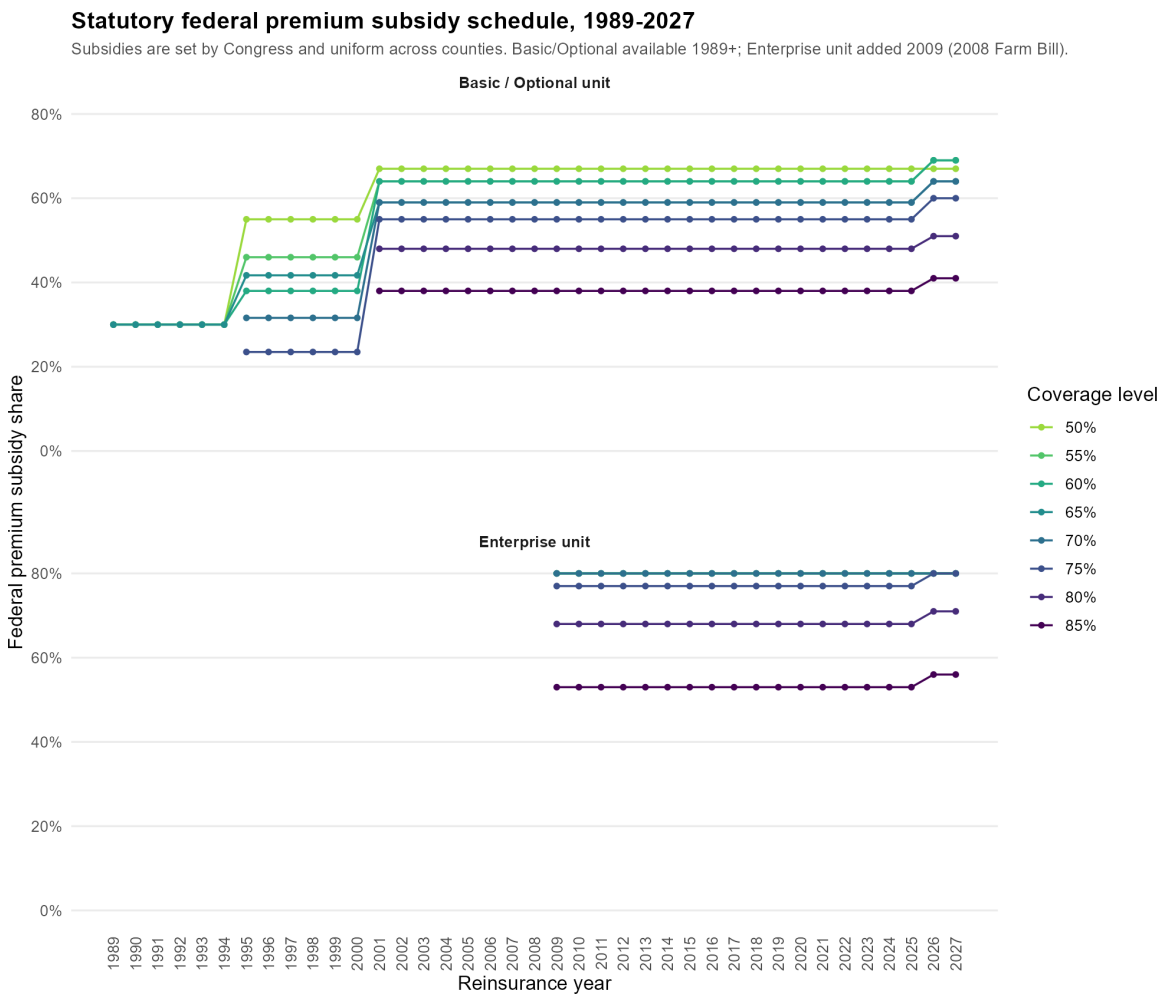


Figure I.7: Statutory federal premium subsidy schedule, 1989–2027, by coverage level and unit structure. Basic/Optional unit (top) covers the full window; Enterprise unit (bottom) starts in 2009. The visible vertical jumps trace the legislative breaks (1994 Reform Act, 2000 ARPA, 2008 Farm Bill, 2025 OBBBA). Federal subsidies are set by Congress and uniform across counties, so the schedule shown applies in all midwest 8 states.

J Climate Data

NCEI publishes nClimGrid-Daily in two release tiers, distinguished by the filename suffix:

Suffix	Content	Released
`-cty-prelim.csv`	Preliminary daily values; no monthly bias-correction yet; subject to revision.	Within days of the observation date.
`-cty-scaled.csv`	**Final** daily values, *bias-corrected* so that the daily series reconciles with NOAA's separately and more carefully quality-controlled monthly product (nClimGrid-Monthly): the monthly sum / mean of the scaled dailies matches the monthly product to within rounding.	Approximately 30–60 days after month end, once the monthly QC pipeline has run.

Every month in our 1989–2025 window is well past NOAA's QC lag, so all 666 downloaded files carry the `-cty-scaled.csv` suffix. The `scaled` suffix in the filename is therefore the user's guarantee that the daily values are the final, calibrated release – temporally consistent with the monthly nClimGrid product and stable across future re-runs of the download script.

(For reference, the only growing-season month presently available only as `prelim` would be the most recent month or two; e.g. as of late May 2026, May 2026 was `prelim` while April 2026 was already `scaled`. This was one of the reasons we elected not to attempt a partial-season 2026 observation.)

County identifier convention. nClimGrid-Daily does *not* use the standard 5-digit county FIPS code; it uses an NWS climate-division state code (alphabetical: 01 = AL, 02 = AZ, ..., 47 = WI, 48 = WY) as the first two digits, concatenated with the county FIPS sub-code in the last three digits. The do-file therefore includes an explicit `nws_to_fips` subroutine that converts the nClimGrid 5-digit id to the standard FIPS code before any geographic restriction or merge.

Validation against the 1989–2013 published panel In order to check the calculations, I compare the results with that of Annan and Schleker (2015) in year 2013. Results are as follows:

	Published (Annan–Schlenker)		New (this project)		Difference	
Variable	Mean	SD	Mean	SD	Mean diff	Corr.
`ddayMod_100`	161.67	27.06	161.22	27.36	−0.45	0.994
`ddayHot_100`	20.69	21.38	21.92	22.13	+1.23	0.989
`prec_cm`	59.97	14.24	60.40	14.62	+0.44	0.982
`prec2_cm`	3803.1	1905.7	3862.3	1973.9	+59.3	–

All four variables agree to better than 1% in mean for the two "thick" variables (`ddayMod_100` and `prec_cm`); the small-mean `ddayHot_100` is biased upward by $\approx +1.2$ C-days ($\sim$6% of its mean). The systematic positive bias in `ddayHot_100` is consistent with the well-known difference between county-averaged temperature inputs (as in nClimGrid's county CSVs, which we use) and a grid-then-aggregate degree-day computation (as in Schlenker–Roberts, which the published panel uses): the latter captures within-county hot-tail variability that flattens out slightly when temperatures are averaged before integration.

The excel files are located at https://www.ncei.noaa.gov/data/nclimgrid-daily/access/averages/<YYYY>/<var>-<YYYYMM>-cty-scaled.csv, where `<var>` is one of `tmin`, `tmax`, `prcp`. Each file holds 3,107 rows (one per US county) and 37 columns: `cty`, `county_id`, "`ST: County`", `year`, `month`, variable name, and 31 daily values (days that do not exist in a 28/29/30-day month are padded with `-999.99`). Temperatures are in degrees Celsius and precipitation in millimetres.. I focused on months of April–September of 1989-2025, following the convention in the literature (Annan and Schlenker 2015).

Using this data, I approximate the temperature trajectory $T(t)$ by a sine wave between the daily minimum $T_{\min}$ and the daily maximum $T_{\max}$. This sine-wave integration of daily degree-day exposure, is formulated by Baskerville and Emin (1969) as:

$$T(t) = M + W\sin(t), \qquad t \in [0, 2\pi), \tag{J.1}$$

where $M = (T_{\min} + T_{\max})/2$ is the daily mean and $W = (T_{\max} - T_{\min})/2$ is the half-amplitude. The day-long degree-day exposure above a threshold τ is

$$\mathrm{dd}(\tau) \;=\; \frac{1}{2\pi}\int_0^{2\pi} \max\big(T(t) - \tau,\, 0\big)\, dt. \tag{J.2}$$

Writing $D \equiv M - \tau$, this integral has the closed form (Snyder 1985):

$$\mathrm{dd}(\tau) \;=\; \begin{cases} 0, & D \le -W \quad (T_{\max} \le \tau) \\ D, & D \ge \;\; W \quad (T_{\min} \ge \tau) \\ \dfrac{D}{2} + \dfrac{D\,\arcsin(D/W)}{\pi} + \dfrac{\sqrt{W^2 - D^2}}{\pi}, & |D| < W. \end{cases} \tag{J.3}$$

Degree-day exposure within an interval (a, b) is obtained by differencing two thresholds:

$$\mathrm{dd}_{(a,b)} \;=\; \mathrm{dd}(a) - \mathrm{dd}(b). \tag{J.4}$$

The flat-day case $W = 0$: When $T_{\min} = T_{\max}$ (a constant-temperature day) the half-amplitude W is zero. Equation (3) still handles this correctly: the first row collapses to $D \leq 0 \Rightarrow \mathrm{dd}(\tau) = 0$, the second row to $D \geq 0 \Rightarrow \mathrm{dd}(\tau) = D$, and the interior case $|D| < W$ becomes the empty set. The two boundary rows therefore reduce to the compact summary

$$\mathrm{dd}(\tau) \;=\; \max(0, D) \qquad \text{when } W = 0,$$

which is the value any consistent implementation must return.

For each county-day in the growing season (April–September), I compute dd(10), dd(29), dd(30) – each one a "degrees-above-threshold" quantity from equation J.2 – and then form the moderate and extreme aggregates as:

Variable	Corn definition	Soybeans definition
Moderate (MDD)	dd(10) − dd(29)	dd(10) − dd(30)
Extreme (HDD)	dd(29)	dd(30)

The two forms in the table are not parallel and the asymmetry is intentional. For moderate heat we want the *interval* 10–29°C (corn) or 10–30°C (soybeans), so MDD uses the differenced form $\mathrm{dd}_{(10,29)} = \mathrm{dd}(10) - \mathrm{dd}(29)$ – exposure above 10°C with the harmful tail above 29°C removed. For extreme heat we want *everything above* 29°C (corn) or 30°C (soybeans) with no upper cap, since hotter temperatures are simply more damaging; HDD therefore uses the above-threshold form dd(29) or dd(30) directly, unsubtracted. Figure J.1 illustrates the trend in exposure to moderate temperature, hot temperature and precipitation for corn and soybeans in the study area over the course of 1989-2025. It shows that the exposure to moderate and hot temperature has increases, and to precipitation has stayed almost the same. to do: add reference to these thresholds from the earlier versions of the paper draft.

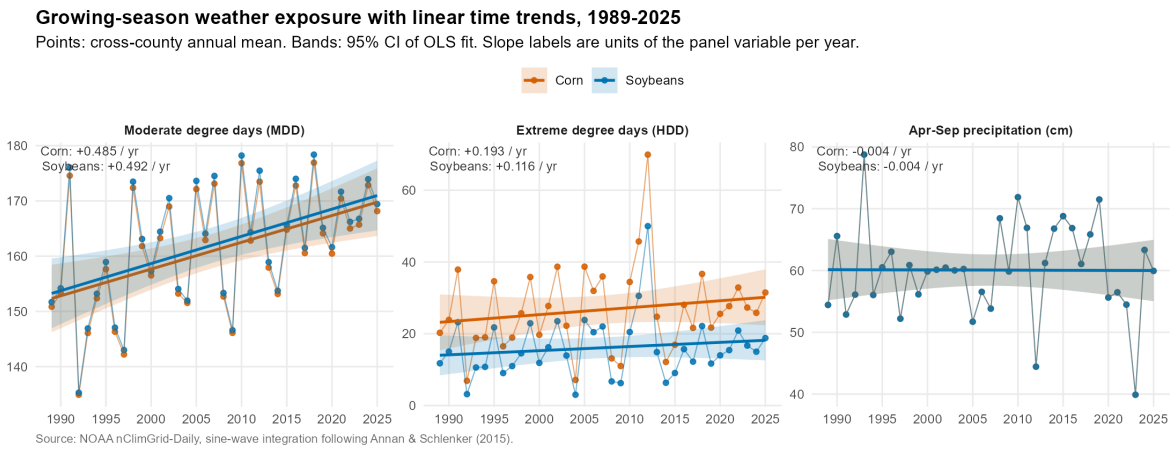


Figure J.1: Growing-season weather exposure and linear time trends, US Midwest, 1989–2025.

Notes: Each panel plots an annual cross-county mean of one growing-season (April–September) weather aggregate for corn (orange) and soybeans (blue), constructed for a balanced panel of 738 US Midwest counties (Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, Wisconsin). *Moderate degree days* (MDD) are the sine-wave-integrated daily temperature exposure within the $(10, 29)\,^{\circ}$C interval for corn and $(10, 30)\,^{\circ}$C for soybeans, summed across the 183-day season. *Extreme degree days* (HDD) are the analogous exposure *above* $29\,^{\circ}$C (corn) or $30\,^{\circ}$C (soybeans). *April–September precipitation* is the seasonal-total precipitation in cm and is identical across crops by construction (the corn and soybean lines overlap in the right panel). Points are the yearly mean across counties; solid lines are OLS fits of the annual mean on year; shaded ribbons are the 95% confidence interval of the fitted line. In-panel labels report the estimated slope in panel-variable units per year. Slopes are computed on 37 annual observations. Daily temperatures and precipitation come from NOAA nClimGrid-Daily Version 1 (Durre et al., 2022); degree-day exposure follows the sine-wave integration of Baskerville and Emin (1969), as implemented in Schlenker and Roberts (2009) and Annan and Schlenker (2015). to do: update the references in the notes.

K Area Planted and Area Insured

The full crops data set contains approximately 23.7 million observations. I retain observations for which `SOURCE_DESC` is `SURVEY`; `AGG_LEVEL_DESC` is `COUNTY`; `STATISTICCAT_DESC` is one of `AREA PLANTED`, `AREA HARVESTED`, or `PRODUCTION`; and `SHORT_DESC` is one of the following: `CORN - ACRES PLANTED`, `CORN, GRAIN - ACRES HARVESTED`, `CORN, GRAIN - PRODUCTION, MEASURED IN BU`, `SOYBEANS - ACRES PLANTED`, `SOYBEANS - ACRES HARVESTED`, or `SOYBEANS - PRODUCTION, MEASURED IN BU`. I additionally require `DOMAIN_DESC` to be `TOTAL`, `FREQ_DESC` to be `ANNUAL`, and `REFERENCE_PERIOD_DESC` to be `YEAR`.

For corn, I use `CORN - ACRES PLANTED` rather than `CORN, GRAIN - ACRES PLANTED`. NASS records planted acreage before the eventual use of the crop is known, so the grain–silage distinction appears only for harvested acreage, production, and yield. This treatment follows the convention used by Schlenker and Roberts (2009) and Annan and Schlenker (2015).

with county-level insured acreage and dollar amounts from the USDA Risk Management Agency's Summary of Business at the state–county–crop level, including coverage-level detail (U.S. Department of Agriculture, Risk Management Agency (RMA) 2025).

Using this data, I define crop yield as bushels per *planted* acre:

$$\texttt{crop_yield} = \frac{\texttt{production_bu}}{\texttt{areaPlanted}}.$$

This is a yield-per-planted-acre measure: it folds harvest failure (drought-prevented planting, silage diversion, abandoned acres) directly into the yield, so a county that planted 100 k acres but only harvested 80 k acres is penalised on the same denominator as one that harvested all 100 k. Given the focus of my study, this is the right measure of crop yield. This deviates from the work of Annan and Schlenker (2015) that use NASS's *published* yield series at the county level—`CORN, GRAIN - YIELD, MEASURED IN BU/ACRE` and `SOYBEANS - YIELD, MEASURED IN BU/ACRE`—which NASS computes as `production_bu / area_harvested`. That is the standard agronomic yield: bushels per acre that actually made it to harvest. The two measures are linked by the harvest ratio,

$$\texttt{crop_yield} = \underbrace{\frac{\texttt{production_bu}}{\texttt{area_harvested}}}_{\text{NASS / Annan-Schlenker yield}} \times \underbrace{\frac{\texttt{area_harvested}}{\texttt{areaPlanted}}}_{\texttt{harvest_ratio}}.$$

Empirically in our 8-state panel the two measures differ by the average harvest ratio: $\texttt{crop_yield}_{\text{corn}} \approx 131$ bu/planted-acre versus NASS yield ≈ 143 bu/harvested-acre (ratio ≈ 0.916, matching mean corn harvest ratio); soybeans ≈ 43.1 vs 43.7 (ratio ≈ 0.988). Substantively, the `crop_yield` variable here is the right choice when you want

the yield response to weather/policy to *include* the harvest-failure margin; the Annan-Schlenker NASS-published series is the right choice when you want yields conditional on harvest having occurred. The two will give different answers in years with widespread prevented-planting (e.g. 2019 in the eastern Corn Belt).

When reporting the results, I convert values in bushels to kilograms using the crop-specific statutory U.S. bushel weight:

$$\texttt{production_kg} \;=\; \texttt{production_bu} \times \begin{cases} 56\,\text{lb/bu} \times 0.45359237\,\text{kg/lb} = 25.4012\,\text{kg/bu} & \text{(corn)} \\ 60\,\text{lb/bu} \times 0.45359237\,\text{kg/lb} = 27.2155\,\text{kg/bu} & \text{(soybeans).} \end{cases}$$

The 56 lb/bu (corn) and 60 lb/bu (soybeans) figures are the legal weight per measured bushel for these commodities under the U.S. grain standards and have been the standard NASS reporting conversion for decades (see Service (1992, Table 5) under the section headings *56-pound bushel of shelled corn, rye, sorghum grain, and flaxseed* and *60-pound bushel of wheat, white potatoes, and soybeans.*

L Expected Price Construction

For each crop separately, let p_{st} be the NASS state-level marketing-year price received in state s and crop year t, and let f_t be the pre-planting (January–March) average daily settlement of the crop's futures contract for that crop year (in dollars per native unit). The futures price is common at the national level. Jisang Yu, A. Smith, and Sumner (2018) estimate, by crop,

$$p_{st} = \alpha_s + \beta f_t + \varepsilon_{st}, \tag{L.1}$$

with state fixed effects α_s, and define the expected price as the fitted value

$$\widehat{p}_{st} \;=\; \widehat{\alpha}_s + \widehat{\beta} f_t. \tag{L.2}$$

Why the fitted value is an expected price The construction rests on a simple statistical fact: an OLS fitted value is, by definition, the minimum-mean-squared-error *linear predictor* of the dependent variable given the regressors. If the regressors are exactly the information available to the farmer at *planting*, then, under $E[\varepsilon_{st} \mid f_t,\ s] = 0$,

$$\widehat{p}_{st} \;=\; E\big[p_{st} \mid f_t,\ \text{state } s\big].$$

That is, Eq. (L.1) decomposes the realized harvest-time received price into a part that was *forecastable* at planting and a part that was not:

$$\underbrace{p_{st}}_{\text{realized}} = \underbrace{\widehat{p}_{st}}_{\text{expected at planting}} + \underbrace{\varepsilon_{st}}_{\text{post-planting surprise}}.$$

The residual ε_{st} absorbs the weather, demand, and harvest shocks that no one could have known in the spring; the fitted value keeps the anticipated component. The two terms of $\widehat{p}_{st}$ in Eq. L.1 each have a concrete reading:

- $\widehat{\beta}\, f_t$ is the **forward-looking forecast, rescaled.** The pre-planting futures price is the market's forecast of the harvest-time price and is in the farmer's information set at planting, but it is quoted for a specific contract, delivery point, and unit—not the cash price a farm in a given state actually receives. The slope $\widehat{\beta}$ estimates the empirical *pass-through* from the national futures quote to the state cash price (it need not equal one, given storage, quality, and the fact that a marketing-year average differs from a single futures number).

- $\widehat{\alpha}_s$ is the **basis.** The state fixed effect captures the persistent average gap between the national futures price and state s's cash price (transport, quality mix, local supply and demand), which is what makes the expected price state-specific.

So $\widehat{p}_{st}$ = (state's average basis) + (pass-through of the market's current forecast). This is preferable to the two obvious alternatives. Relative to the *realized* received price, $\widehat{p}_{st}$ removes the endogenous, supply-driven component: the realized price is set at harvest jointly with quantity (more acreage → more output → lower price), so using it as a regressor in an acreage equation induces simultaneity bias and mistimes the decision; that component lives in ε_{st}, not in $\widehat{p}_{st}$. Relative to the *raw futures* price, the regression calibrates a national, wrong-unit quote into a state-level cash-price object via $\widehat{\beta}$ and $\widehat{\alpha}_s$.

The fitted value is then converted to dollars per metric ton using the crop conversion factors in Table L.1, deflated to May 1986 dollars with the PPI, and logged:

$$P_{st}^{1986} \;=\; 100 \cdot \frac{\widehat{p}_{st}/c_{\text{crop}}}{\text{PPI}_t}, \qquad \ln P_{st}^{1986} \;=\; \ln\!\big(P_{st}^{1986}\big). \tag{L.3}$$

This $\ln P_{st}^{1986}$ is exactly the variable `lnprice_received` in the YSS replication data and the $\ln P_{ict}^{1986}$ that enters the dissertation's constant-dollar conversion below.

The conversion factors c_{crop} in Table L.1 are taken from the YSS replication cleaning code (`4_Data_cleaning_final_dataset.do`; Yu, Smith, and Sumner 2018), with one correction noted below. Each factor is the mass of one native marketing unit expressed in metric tons, i.e. a standard weights-and-measures conversion. These standard

agricultural weights and U.S.-to-metric factors are documented in USDA Economic Research Service (1992), *Agricultural Handbook 697*: the commodity bushel weights in its Tables 5–6 and the hundredweight definitions in its Table 2 (p. 5).[16]

Table L.1: Native-unit-to-metric-ton conversion factors c_{crop} (from YSS cleaning code) and the futures contract used for the 1989–2025 extension.

Crop	Code	Native unit	c_{crop}	Extension futures
Wheat	11	\$/bu	0.027	KC HRW wheat (KE=F)
Rice	18	\$/cwt	0.0454	Rough rice (ZR=F)
Cotton	21	\$/lb	0.0005	Cotton no. 2 (CT=F)
Corn	41	\$/bu	0.025	Corn (ZC=F)
Sorghum	51	\$/cwt	0.0454	Corn (ZC=F)
Soybeans	81	\$/bu	0.027	Soybeans (ZS=F)
Barley	91	\$/bu	0.021	Corn (ZC=F)

Note: Factors are the mass of one native marketing unit in metric tons, from Yu, Smith, and Sumner (2018); the underlying standard weights are from USDA ERS (1992), Agricultural Handbook 697, Table 2 and Tables 5–6. Rice and grain sorghum are reported in dollars per hundredweight (cwt). The YSS cleaning code used $c = 0.051$, which corresponds to the 112-lb "long" hundredweight (USDA ERS 1992, Table 2, p. 5); here it is **corrected** to the 100-lb U.S. hundredweight, $c = 0.0454 = 100\,\text{lb} \times 0.453592\,\text{kg/lb}$ MT. This correction is a constant $0.051/0.0454 \approx 1.12$ rescale applied to rice and sorghum in *every* year (including 1989–2014), and raises their mean constant-2013 prices from \$263.6 to \$296.4 and from \$140.1 to \$157.6 per MT, respectively; all other crops are unchanged.

Following YSS, sorghum and barley use the corn futures contract (no dedicated, liquid futures exist for them). In the original paper wheat used a combination of Kansas City winter and Minneapolis spring contracts; for the extension I use a single KC hard-red-winter contract (the dominant U.S. wheat futures), averaged over the same January–March window. The futures leg is a level/variation proxy: because α_s absorbs the level and the series is calibrated to the original over the overlap (below), the precise choice affects only second-order within-year variation.

M Construction of Intercensal County-Level Economic Controls for Level Variables

I adjust these series for annual time using the census years observations as benchmarks. To do so, I interpolate intercensal county values with Friedman's related-series

[16] The bushel and pound factors match the standard weights exactly (e.g. $56\,\text{lb} \times 0.45359\,\text{kg/lb} = 25.4\,\text{kg} = 0.0254$ MT ≈ 0.025). USDA ERS (1992, Table 2, p. 5) lists *both* hundredweights: $100\,\text{lb} = 1$ (avoirdupois) cwt and $112\,\text{lb} = 1$ gross or "long" cwt. The cwt factor used by YSS, 0.051, corresponds to the 112-lb long hundredweight rather than the 100-lb U.S. hundredweight; I correct rice and sorghum to the 100-lb U.S. hundredweight (0.0454 MT) uniformly across all years, as detailed in the note to Table L.1.

method (Friedman 1962, Method 2, p. 732) and use the annually observed state-level benchmark to trace within-period movement. For bounded share variables, I transform the county series into logit space before interpolation so that the imputed values remain inside the unit interval after I transform them back. For strictly positive economic variables, I apply the proportional, logarithmic version of Friedman's method so that each county inherits the annual pattern of the benchmark series while still matching its observed census endpoints. I deflate each monetary variable and its corresponding state-level benchmark using a common price index appropriate to that variable category. When a variable-specific deflator is available, I use it; otherwise, I use a general deflator to express values in constant dollars. I apply the same deflator to the county series and its related benchmark so that Friedman's interpolation operates on comparable real magnitudes. I express monetary values in constant 2013 dollars. Specifically, I deflate each nominal county series and its corresponding state-level benchmark using an annual price index normalized to the 2013 level. I convert nominal income and loan variables to constant dollars using the BEA Gross Domestic Purchases Price Index, which measures overall inflation in prices paid within the U.S. economy. I then apply Friedman's logarithmic related-series interpolation to the resulting real series. I deflate nominal income and loan variables using the Bureau of Economic Analysis (BEA) *Gross Domestic Purchases* annual price index from Table 1.6.4, *Price Indexes for Gross Domestic Purchases* (row 1; index numbers, 2017 = 100), and express values in constant 2013 dollars (Bureau of Economic Analysis 2026).

$$X_t^{\text{real}} = X_t^{\text{nominal}} \times \frac{P_{2013}}{P_t},$$

Formally, let Y_{cT} denote the observed county-level value in census year T, and let X_{st} denote an annual state-level (or national-level) indicator observed for all years t. For intercensal years t between two census benchmarks, the annualized county value is constructed as:

$$Y_{ct} = Y_{cT} \times \frac{X_{st}}{X_{sT}},$$

where X_{sT} denotes the indicator value in census year T. This proportional scaling preserves the observed census benchmark while allowing annual variation to follow economically meaningful macro trends. When two surrounding census benchmarks are available, I interpolate growth proportionally within the intercensal interval while ensuring consistency at both endpoints. This approach avoids imposing arbitrary linear trends and mitigates the risk of extrapolation-induced distortions.

Let $\tilde{y}_{ct}$ denote the filled county-level value for county c in year t, where the fill is performed in inverse hyperbolic sine (IHS) space. For years between two observed anchor years t_0 and t_1, I construct

$$\tilde{y}_{ct} = y_{ct_0} + \lambda_t(y_{ct_1} - y_{ct_0}) + \left[B_{st} - \left(B_{st_0} + \lambda_t(B_{st_1} - B_{st_0})\right)\right], \qquad \lambda_t = \frac{t - t_0}{t_1 - t_0}.$$

The first component is the county-specific linear interpolation between the two anchor values, while the second component adjusts this interpolated path by the deviation of the state benchmark from its own linear trend. Thus, the imputed county series matches observed county endpoints while also incorporating annual movements in the benchmark series. The resulting filled values are then transformed back to levels. Figure M.2 show the filled values and compares them with the original values.

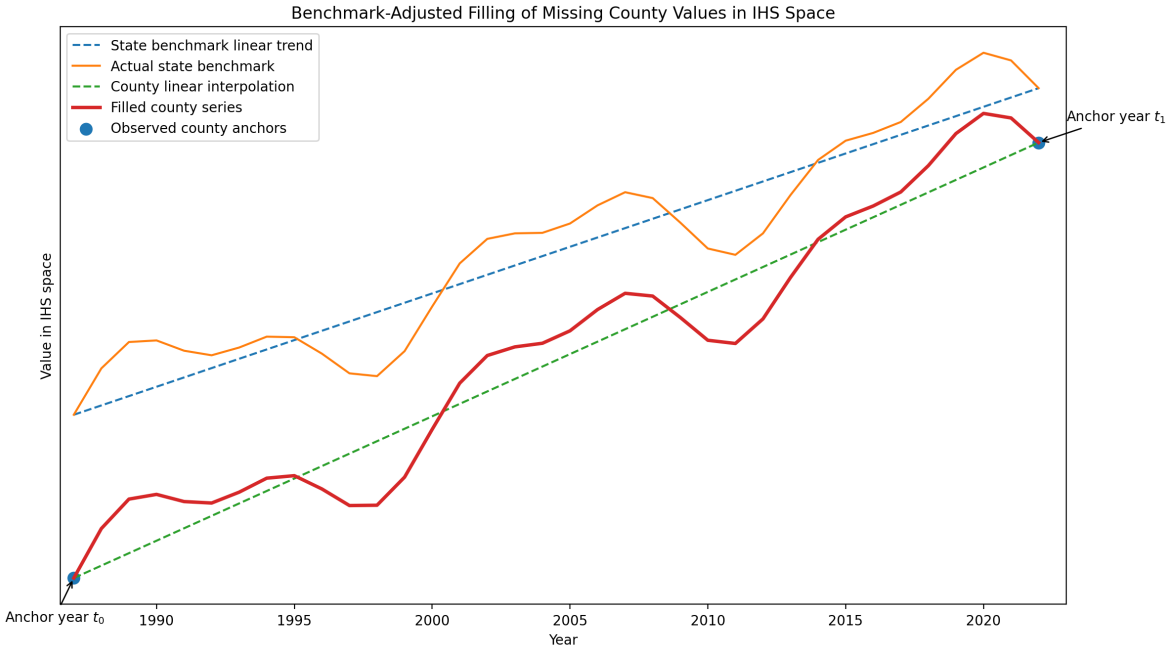
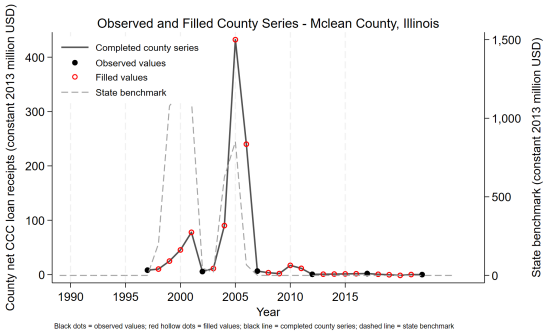


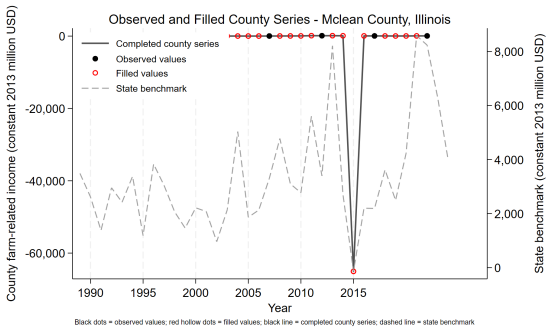
Figure M.1: Benchmark-adjusted method for filling the missing values using the method of (Friedman 1962, Method 2, p. 732).

The choice of indicator series is variable-specific and grounded in economic relevance:

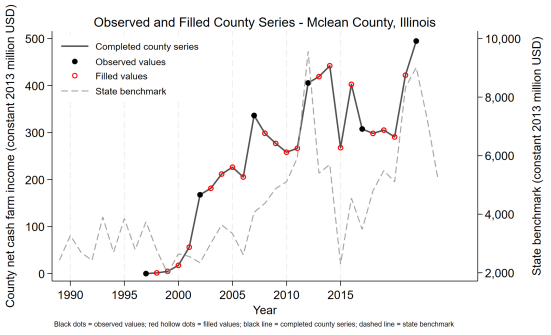
- **Farm-Related Income.** Intercensal values are benchmarked to annual state-level Net Farm Income from the USDA Economic Research Service (ERS) Farm Income and Wealth Statistics (U.S. Department of Agriculture, Economic Research Service 2026d).
- **Net Cash Farm Income.** Annual growth is scaled using state-level Net Cash Farm Income from the USDA ERS Farm Income and Wealth Statistics (U.S. Department of Agriculture, Economic Research Service 2026c). This preserves



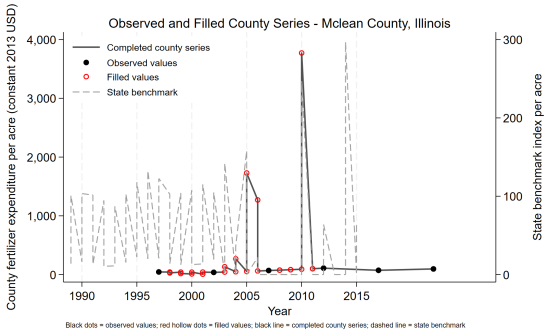
(a) Net CCC loan receipts



(b) Farm-related income



(c) Net cash farm income



(d) Fertilizer expenditure per acre

Figure M.2: Observed and filled county series for McLean County, Illinois. Black dots denote observed values, red hollow dots denote filled values, the solid black line denotes the completed county series, and the dashed line denotes the state benchmark.

alignment with national accounting definitions of cash receipts and production expenses.

- **Commodity Credit Corporation Loans.** County-level net Commodity Credit Corporation (CCC) loan receipts are benchmarked to annual state-level Marketing Loan Gains (MLG) and Loan Deficiency Payments (LDP) (U.S. Department of Agriculture, Economic Research Service 2026b) , which together capture the cyclical intensity of CCC marketing assistance loan activity. Marketing Assistance Loans (MALs) are nonrecourse loans provided to producers at harvest time that allow them to obtain interim financing using eligible commodities as collateral. These loans enable producers to delay sale when market prices are seasonally low and repay the loan when market conditions improve (U.S. Department of Agriculture, Farm Service Agency 2026). Marketing Loan Gains occur when producers repay MALs at a rate below the statutory loan rate when market prices fall below support levels, generating an implicit subsidy equal to the difference between the loan rate and the repayment rate. Loan Deficiency Payments are direct payments made to producers who, although eligible for a marketing assistance loan, elect to forgo the loan and instead receive a payment reflecting the same price-support benefit (U.S. Department of Agriculture, Farm Service Agency 2026). Because both MLGs and LDPs are triggered by commodity prices falling below loan rates, they closely track periods of elevated CCC marketing loan utilization. While these measures capture the subsidy component rather than gross loan principal extended, they provide a policy-consistent annual indicator suitable for temporally disaggregating census-based county CCC loan receipts.
  Because Marketing Loan Gains and Loan Deficiency Payments are frequently zero in years when market prices exceed loan rates—and may be negative in some years due to accounting adjustments, net recoveries, or revisions—I construct the state-level benchmark indicator using the nonnegative component of each series:

$$B_{st}^{L} = \max(0, \mathrm{LDP}_{st}) + \max(0, \mathrm{MLG}_{st}),$$

  where s indexes states and t indexes years. This construction ensures that (i) zero-payment years remain zero, and (ii) negative accounting corrections do not reverse the economic signal of CCC program intensity. The resulting series captures the cyclical activation of CCC marketing loan programs while avoiding distortions from administrative adjustments.

- **Fertilizer Expenditure per Acre.** To construct a state–crop–year benchmark for fertilizer expenditure per acre, I combine information on fertilizer use inten-

sity from the USDA Economic Research Service (ERS) *Fertilizer Use and Price* dataset with the annual fertilizer price index reported in Table 8 (U.S. Department of Agriculture, Economic Research Service 2026e). Specifically, I use Tables 9–14 for corn and Tables 21–26 for soybeans, which report (i) the percentage of acreage receiving each major nutrient and (ii) the application rate per fertilized acre for nitrogen, phosphate, and potash.

Let s index state, $c \in \{\text{corn}, \text{soybeans}\}$ index crop, t index year, and $j \in \{N, P, K\}$ index nutrients. For each nutrient j, the dataset provides:

- A_{sjct}: the share of crop acreage receiving nutrient j (Tables 9, 11, 13 for corn; Tables 21, 23, 25 for soybeans),
- R_{sjct}: the application rate of nutrient j in pounds per fertilized acre (Tables 10, 12, 14 for corn; Tables 22, 24, 26 for soybeans).

Because application rates are reported per fertilized acre, I first convert them into average nutrient use per planted acre by scaling by the treated acreage share:

$$I_{sjct} = A_{sjct} \times R_{sjct},$$

where A_{sjct} is expressed as a fraction rather than a percentage. The composite fertilizer-use measure for crop c in state s and year t is then:

$$F_{sct} = I_{sNct} + I_{sPct} + I_{sKct} = \sum_{j \in \{N,P,K\}} A_{sjct} R_{sjct}.$$

This measure captures the total pounds of primary nutrients applied per planted acre, combining both the extensive margin (the share of acreage receiving fertilizer) and the intensive margin (the application rate conditional on use). All nutrient quantities are therefore measured in pounds per planted acre.

To convert this quantity-based measure into a fertilizer expenditure index per acre, I scale the composite use measure by the annual fertilizer price index from Table 8, denoted by P_t, where the index is normalized to 2011=100:

$$E^{\text{fert}}_{sct} = F_{sct} \times \frac{P_t}{100}.$$

Equivalently,

$$E^{\text{fert}}_{sct} = \left(\sum_{j \in \{N,P,K\}} A_{sjct} R_{sjct} \right) \times \frac{P_t}{100}.$$

This yields a state–crop–year fertilizer expenditure index per acre that combines cross-state and cross-crop variation in fertilizer application intensity with common annual movements in fertilizer prices. Because Table 8 provides an index rather than fertilizer prices in dollars, E_{sct}^{fert} should be interpreted as an expenditure *index*, not as a dollar-valued measure of fertilizer spending per acre. Intercensal values are not filled at this stage; the index is constructed only for years in which the underlying ERS fertilizer-use information is observed.

- **Cattle Inventory at Level and per Acre.** Annual variation is benchmarked to state-level cattle inventory series from the USDA National Agricultural Statistics Service (NASS) Quick Stats database (USDA National Agricultural Statistics Service 2026).