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

Job Market Paper of Homa Taheri*

Abstract

Agricultural insurance subsidies are a major component of farm policy in the United States. Subsidies encourage farmer participation, facilitate risk management, and stabilize rural incomes. However, farmers often share in conversations that subsidies distort their production decisions and affect crop diversity, and alter sensitivity of yields to extreme weather events. In contrast, existing empirical research on the U.S. Federal Crop Insurance Program (FCIP) generally documents limited effects of subsidies on crop yields. This apparent mismatch between practitioner experience and empirical findings motivates this paper.

In the first part of the paper, I re-examine the production effects of crop insurance subsidies by distinguishing between two margins of enrollment: the external margin, measured by the percentage of planted acreage insured, and the internal margin, measured by insurance liability per insured acre. To identify causal effects, I develop instrumental variables that capture (i) variation in the ease of purchasing insurance, (ii) the historical reliability of weather-related news consumed by local farmers, and (iii) the past performance of farmers. I further account for spatial proximity across counties to capture peer effects in insurance adoption. To examine dynamic responses to policy shocks, I employ the local projection method to trace the medium- and long-run effects of subsidy changes on production outcomes. [Results to be added]

In the second part of the paper, I move from positive analysis to a normative question: what is the welfare-maximizing level of crop insurance subsidies? I address this question using a sufficient statistics framework commonly applied in public finance and international trade. This approach allows me to quantify the trade-off between risk-sharing benefits and production distortions implied by the estimated behavioral responses. [Results to be added]

Together, the results contribute to understanding how agricultural insurance design affects production behavior and inform ongoing policy debates about the appropriate level of federal support for climate risk management in agriculture.

Keywords: Crop Insurance, Local Projections, Climate Risk, Moral Hazard, Policy Evaluation, Spatial Propagation

JEL Codes: Q18 and Q58

*Homa Taheri is a Ph.D. candidate in Economics at Indiana University Bloomington, a researcher at the International Food Policy Research Institute (IFPRI), and a visiting Scholar at the Center for Spatial Data Science (CSDS) at the University of Chicago. Contact: htaheri@iu.edu. I am deeply grateful to Ahmad Lashkaripour for his guidance and support throughout this project. I thank the participants at the 90th and 89th Annual Meetings of the Midwest Economics Association, the 3rd Hoosiers Conference, and my colleagues at CSDS for valuable comments and feedback. Numerous staff members at the U.S. Department of Agriculture, especially at the Risk Management Agency, provided essential data assistance and institutional guidance. This project would not have been possible without their continued support. 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).

"*Cultivators of the earth are the most valuable citizens. They are the most vigorous, the most independent, the most virtuous, and they are tied to their country and wedded to it's liberty and interests by the most lasting bonds.*" - Thomas Jefferson, Letter to John Jay, August 23, 1785 (Jefferson 1785).

1 Introduction

Crop insurance transfers climate risk from farmers to the public sector through a heavily subsidized program administered by the U.S. Department of Agriculture (USDA). Under the Federal Crop Insurance Program (FCIP), farmers pay a subsidized premium. In return, the USDA provides indemnity payments when realized yields (or revenue, depending on the policy) fall below guaranteed levels. From the farmer's perspective, the expected net return from insurance depends on two factors: the subsidy rate and the actuarial performance of the policy. Because subsidies change both expected returns and the distribution of risk, they can affect production decisions. In particular, subsidies may influence crop choice, land allocation, and investment in adaptation.

In my conversations with the members of the Chicago Farmers[1], I often hear them sharing their experience with the FCIP: how it discourages growing produce other than the covered produce, and how they know of some farmers who have benefited from the program for financing their risky investment in some form of practice they did not have any prior experience with; examples of untended effects of policies, and moral hazard of the beneficiaries. Back to my office, and in the literature, I have come across a large body of literature that has studied the incidents of moral hazard and adverse selection of farmers from the point of view of use of inputs such as irrigation, fertilizers and pesticides, as well as decisions such as the area planted and harvested, and crop yield. In all cases except the very last one, other scholars have found evidence on negative effect of enrollment in crop insurance on farming practices. When it comes to the effect on crop yield in the US, Roberts, E. J. O'Donoghue, and Key (2007) reports little effect on yield distributions. This contrast between effects on input use and outcomes, raises the central question of my paper.

The first part of this paper addresses this question that if crop insurance alters production incentives, why do we observe limited effects on yields? Are traditional yield-based analyses missing important behavioral responses operating through crop allocation, risk exposure, or adaptation margins? In order to solve the puzzle of yield effects of enrollment in crop insurance, I first divide the task into extensive and intensive margins. At the extensive margin, I study the effect of an extra unit of insured acre per planted acre on the yield outcome of farmers. At the intensive margin, I study the effect of an extra US Dollar liability of farmers to the FCIP per insured acre on the yield outcome of farmers. To identify causal effects, I develop instrumental variables that capture (i) variation in

[1] https://chicagofarmers.org/

the ease of purchasing insurance, (ii) the historical reliability of weather-related news consumed by local farmers, and (iii) the past performance of farmers. I further account for spatial proximity across counties to capture peer effects in insurance adoption. To examine dynamic responses to policy shocks, I employ the local projection method to trace the medium- and long-run effects of subsidy changes on production outcomes. This is the first empirical analysis of long term effects of Farm Bills in literature. [Results to be added]

In the second part of the paper, I move from positive analysis to a normative question: what is the welfare-maximizing level of crop insurance subsidies? I address this question using a sufficient statistics framework commonly applied in public finance and international trade. This approach allows me to quantify the trade-off between risk-sharing benefits and production distortions implied by the estimated behavioral responses. I model the optimal level of subsidies for two cases of no land relocation across crops (Model A), and with relocation of acreage across crops (Model B) and recover the welfare effects of premium subsidies using estimable demand elasticities and fiscal parameters, without fully modeling production or consumption. The closest prior work to my analysis is the Ph.D. dissertation of Jialing Yu (2017, ch. 4), that characterizes the marginal welfare effect consequences and optimal policy design for the FCIP. Jialing Yu (2017) reports that under current policy settings, the net marginal welfare effects of coverage levels and subsidy rates are negative, and both are greater than their optimal levels. [Results to be added]

Prior research on yield effects of the FCIP very extensively focuses on the differential performance of insured and unsured farmers when in some form or another exposed to the consequences climate change. One exception is Roberts, E. J. O'Donoghue, and Key (2007) who studied the yield effects of crop insurance by by comparing how crop yields change for those cycling into insurance relative to those in the same county who were insured in both periods, uninsured in both periods, or who cycled out of insurance. The authors found that both buy-up and CAT insurance have little effect on yield distributions.

Instances of the predominant line of work that focuses on the responsiveness of insured crops to the changing climate, are studies such as Annan and Schlenker (2015) that show the highly subsidized crop insurance program of the USDA makes farmers more sensitive to fluctuations in extreme heat. The authors find that the sensitivity to extreme heat is 67% larger for insured corn than for uninsured corn, and 43% higher for soybeans than for uninsured soybeans. They interpret this finding as evidence that crop insurance subsidies can discourage costly adaptive investments, thereby increasing vulnerability to heat stress. Their results highlight a central tension in agricultural policy. Although crop insurance is designed to mitigate risk, it can also inadvertently reduce resilience to extreme weather. Subsequent work reinforces this concern. Regmi, Tack, and Featherstone (2023) finds that insured dryland corn is approximately 38% more sensitive to extreme heat than uninsured corn. These findings suggest that subsidies on the intensive margin not only reduce incentives to adapt but can also systematically increase the exposure of insured crops to weather-induced yield losses. Other studies emphasize the ways in which crop insurance subsidies shape production behavior and, in turn, exposure to weather risk. Jisang Yu, A. Smith, and Sumner (2018) estimate that a 10% increase in premium subsidies raises insured crop acreage by about 0.43%. By incentivizing greater insured acres, subsidies expand the area of production that is directly exposed to extreme weather events. This suggests that on the extensive margin, subsidies affect not only adaptation decisions but also the spatial scale of exposure to risk.

Some scholars have identified the role of Insurance subsidies on resilience through innovation and the adoption of technology. Chemeris, Liu, and Alan P. Ker (2022) find that while subsidies can encourage innovation in some contexts, they may simultaneously weaken crop yield resiliency under climate change. Their analysis indicates that the relationship between insurance, innovation, and weather risk is complex: subsidies provide stability in the short run but may undermine incentives for longer-term resilience strategies.

In this dissertation I use the case of American farmers for this goal, but implications of this study are by no means limited to the US, since subsidy provision for farmers is a common practice worldwide (see Appendix E) for more information on subsidies in other countries.

This paper contributes to a few lines of research. 1) The effect of premium subsidies on demand for crop insurance is a well-studied question in the literature (E. O'Donoghue 2014; Babcock and Hart 2005). 2) Sensitivity of crop yield to extreme weather in the Midwest. 3) Social welfare effects of crop insurance (Chung 2014), 4) crop insurance as a strategy for adapting to climate change(Falco et al. 2014) and 5) Method for decreasing income risk and income hedging of farmers (Walters and Preston 2018).

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of the FCIP. Section 3 introduces the data. Section 4 discusses the methodology. Sections on empirical framework, results, and conclusion are currently available as an unpolished format on page 48 and onward.

2 Background

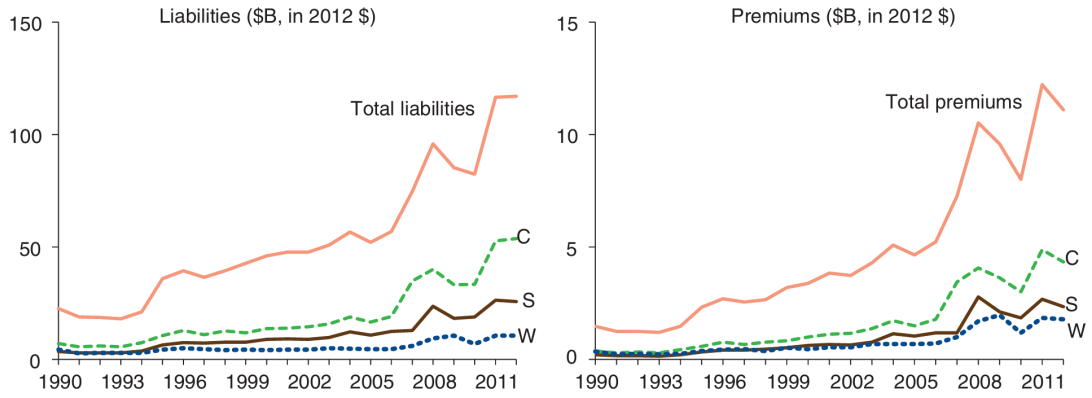
2.1 Institutional Background

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. 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 2
Growth in participation in crop insurance: alternate measures, normalized to 2012 dollars



Note the scale of the y-axes—a factor of 10 difference between liabilities and premiums.
C = corn; S = soybeans; W = wheat. Premiums and liabilities were normalized for these charts using the Consumer Price Index.
Source: USDA, Economic Research Service using data from USDA, Risk Management Agency, Summary of Business files, 1990-2012.

Figure 1: Growth in participation in crop insurance:alternate measures, normalized to 2012 dollars

2.2 The Premiums, Subsidies and Types of Protection

Ain't no one gonna want to read that. Radically cut down on this and integrate in your data section, or put this in an appendix. 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.

Since insurance payouts are a convex function of farm yield, the increase in variability adds to the cost of providing aggregate insurance.

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.

2.3 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 2 and 3.

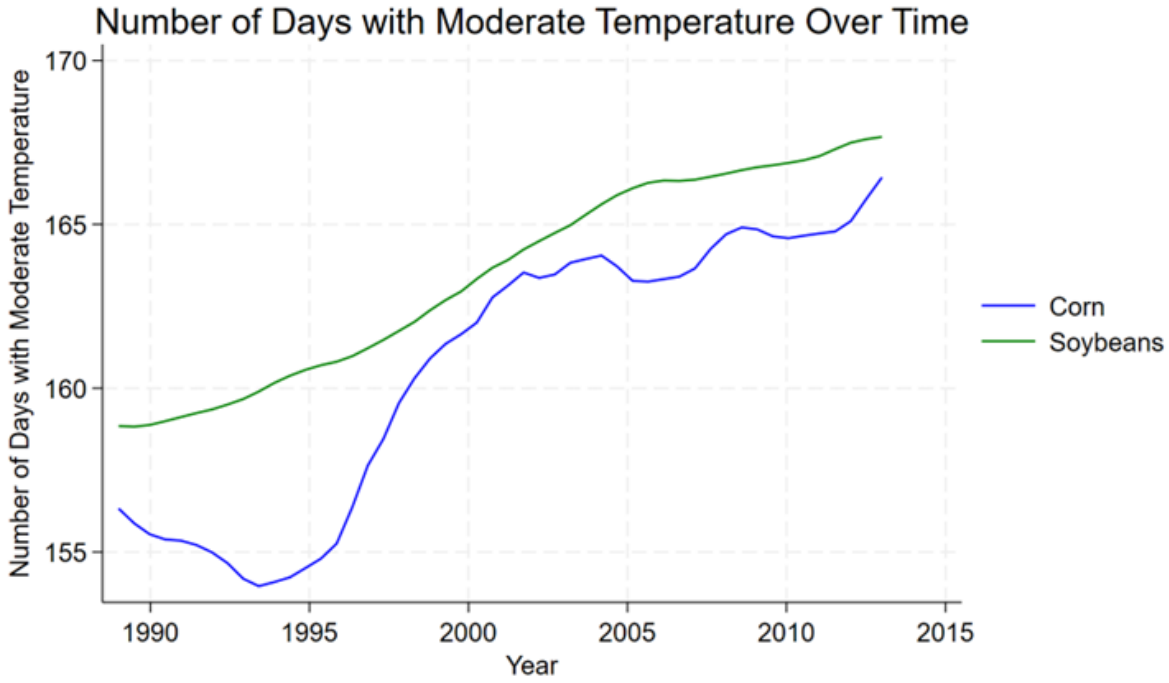


Figure 2: Number of days with moderate temperature (1989-2013)

1) Soybeans replenish soil nitrogen, and that is typically the reason why they are grown along with (or in rotation with) corn. We discuss this in our forthcoming JHR article (Bellemare, Liu, and Hadrich 2025) 2) Footnotes go after punctuation. 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

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 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.

Table 2: Means and Standard Deviations by Period, Part I

	1989–1990	1991–1996	1997–2002	2003–2008	2009–2013	1989–2013
Yield	76.66 (41.82)	78.18 (43.51)	85.35 (47.77)	93.98 (56.23)	95.20 (56.34)	87.24 (51.06)
Planted acreage	71,658.81 (54,830.23)	73,423.13 (55,373.44)	75,805.02 (55,212.94)	76,439.39 (55,931.45)	79,346.80 (57,386.19)	75,843.17 (55,890.01)
Insured acreage	30,247.99 (36,008.09)	33,538.87 (38,072.57)	50,766.59 (45,571.02)	58,242.57 (48,118.23)	65,843.22 (51,635.92)	50,294.75 (47,052.89)
Total premium	226537.57 (276287.73)	227806.09 (270828.00)	521716.66 (539847.19)	1.52e+06 (1.72e+06)	2.82e+06 (2.40e+06)	1.15e+06 (1.72e+06)
Subsidy	54,197.51 (65,328.70)	87,513.44 (111292.94)	259117.44 (261417.51)	825114.20 (901471.99)	1.65e+06 (1.36e+06)	627431.93 (970032.63)
Hot degree days	16.64 (12.99)	16.83 (16.35)	21.32 (19.00)	18.76 (20.08)	28.42 (29.56)	20.76 (21.40)
Moderate degree days	155.39 (20.04)	155.47 (27.35)	164.14 (26.95)	163.71 (26.66)	165.98 (28.14)	161.81 (27.12)
Precipitation	60.21 (12.47)	62.55 (16.21)	58.98 (10.72)	58.20 (13.85)	60.24 (16.12)	59.95 (14.23)
Precipitation squared	3,783.60 (1,473.70)	4,179.83 (2,360.22)	3,597.17 (1,277.73)	3,584.22 (1,784.62)	3,893.36 (2,160.20)	3,800.37 (1,905.06)
Liability per acre	137.00 (36.20)	124.15 (40.87)	114.48 (34.24)	166.11 (86.78)	422.03 (196.38)	193.32 (155.14)
Liability per acre (CAT)	0.00 (0.00)	24.54 (33.84)	66.36 (15.02)	89.07 (41.39)	152.61 (69.14)	75.23 (63.05)
Liability per acre (Buy-up)	0.00 (0.00)	57.22 (79.00)	156.20 (35.55)	213.97 (102.54)	455.66 (178.42)	197.53 (178.91)
Fraction insured (Any)	0.36 (0.22)	0.43 (0.27)	0.37 (0.18)	0.19 (0.13)	0.50 (0.35)	0.36 (0.26)
Fraction insured (CAT)	0.00 (0.00)	0.15 (0.22)	0.17 (0.12)	0.08 (0.09)	0.04 (0.07)	0.11 (0.15)
Fraction insured (Buy-up)	0.00 (0.00)	0.10 (0.18)	0.18 (0.14)	0.10 (0.09)	0.46 (0.35)	0.19 (0.24)
Price received	150.07 (70.87)	134.63 (57.28)	115.64 (49.34)	119.27 (54.85)	155.73 (57.98)	131.41 (58.18)
Harvested acreage	0.07 (0.05)	0.07 (0.05)	0.10 (0.08)	0.09 (0.07)	0.09 (0.08)	0.09 (0.07)
Irrigated planted acreage	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Non-irrigated planted acreage	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Irrigated harvested acreage	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Non-irrigated harvested acreage	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Observations	2,102	6,878	7,486	7,485	6,069	30,020

Means shown; standard deviations in parentheses.

Table 3: Means and Standard Deviations by Period, Part II

	1989–1990	1991–1996	1997–2002	2003–2008	2009–2013	1989–2013
Share insured any $\geq$ 10%	. (.)	. (.)	. (.)	. (.)	. (.)	. (.)
Share premium subsidy $\geq$ 10%	. (.)	. (.)	. (.)	. (.)	. (.)	. (.)
Share female	. (.)	. (.)	. (.)	. (.)	. (.)	. (.)
Fertilizer expenditure per acre	. (.)	. (.)	36.34 (18.22)	70.82 (39.45)	126.38 (65.89)	67.13 (54.48)
Share minority producers	. (.)	. (.)	. (.)	. (.)	. (.)	. (.)
Cattle inventory	0.01 (0.01)	0.01 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
Cattle per acre	0.06 (0.13)	0.06 (0.13)	0.04 (0.10)	0.05 (0.10)	0.02 (0.07)	0.05 (0.11)
Share acres owned	. (.)	. (.)	0.42 (0.11)	0.41 (0.11)	0.42 (0.10)	0.42 (0.11)
Driving distance to USDA office (km)	8.25 (7.11)	8.46 (8.20)	8.77 (9.06)	8.74 (8.98)	8.69 (8.82)	8.64 (8.67)
Weather stations in 1989	2.09 (1.14)	2.08 (1.14)	2.09 (1.15)	2.09 (1.14)	2.09 (1.13)	2.09 (1.14)
Tariff exposure	0.07 (0.07)	0.08 (0.07)	0.09 (0.07)	0.09 (0.05)	0.07 (0.04)	0.08 (0.06)
Farm income (real)	. (.)	. (.)	6.00 (3.83)	9.14 (6.38)	19.77 (15.79)	10.56 (10.55)
Net cash income (real)	. (.)	. (.)	12.30 (22.52)	67.31 (58.25)	89.35 (75.62)	51.93 (62.44)
CCC net payments (real)	. (.)	. (.)	10.25 (16.53)	90.47 (178.13)	17.96 (44.97)	42.40 (118.26)
Observations	2,102	6,878	7,486	7,485	6,069	30,020

Means shown; standard deviations in parentheses.

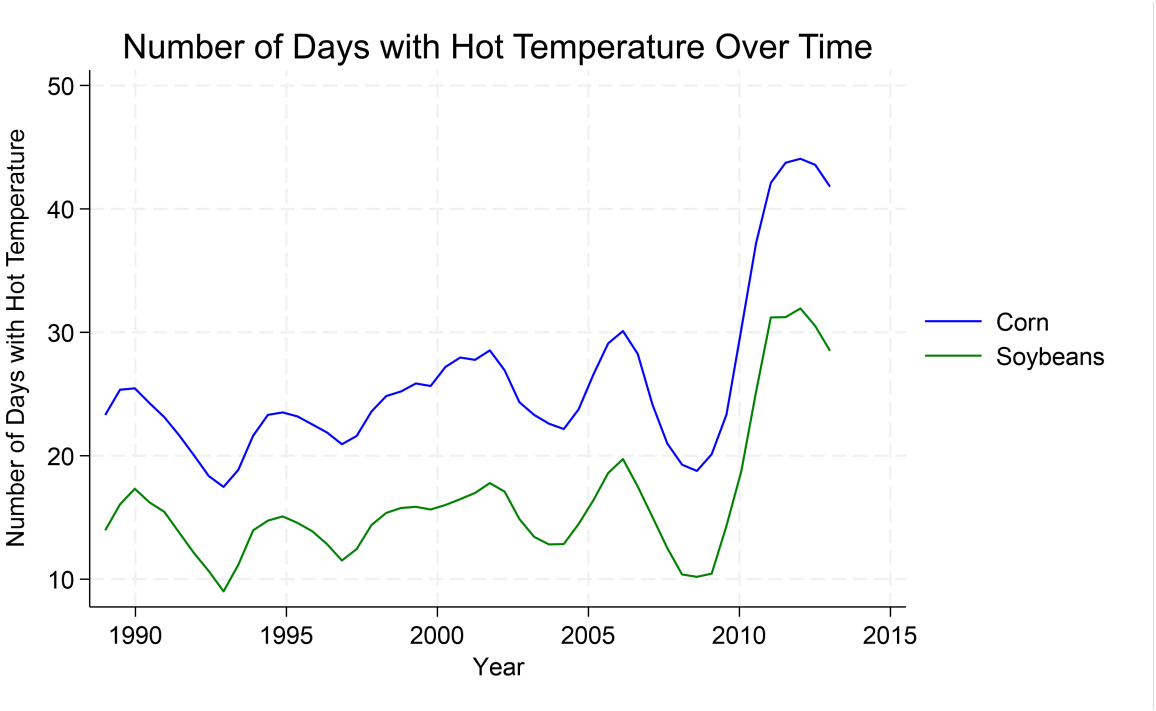


Figure 3: Number of days with extreme temperature (1989-2013)

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.

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.

Education Level of Study Area

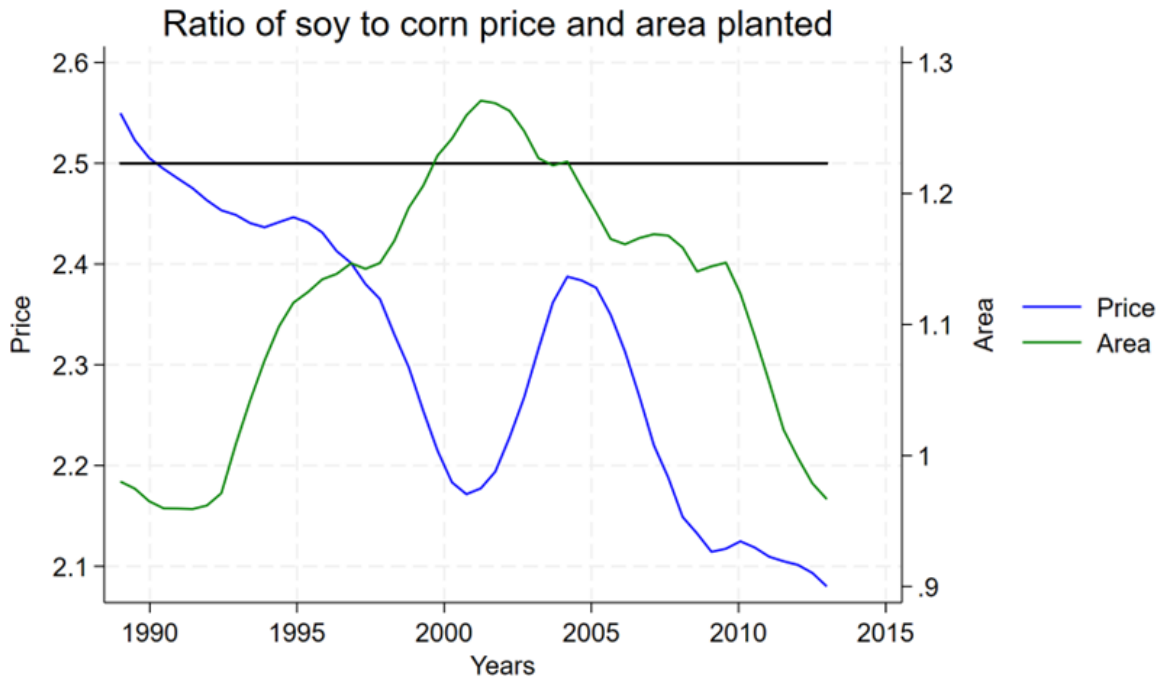


Figure 4: Left axis: Received price of farmers (blue) for soybeans relative to corn. Right axis: Relative area planted for soybeans relative to corn (green). The black horizontal line shows the rule of thumb threshold of switching from corn to soybeans, as known in the farming community.

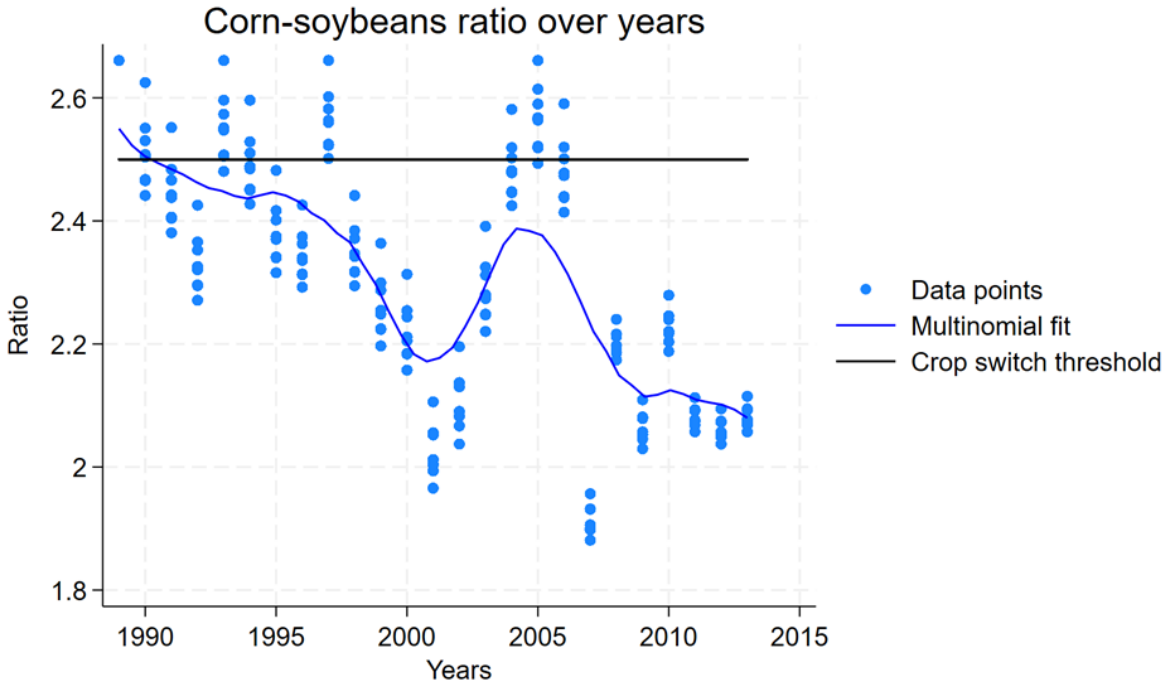


Figure 5: Relative price of soybeans to corn in eight Midwest states and the crop switching threshold.

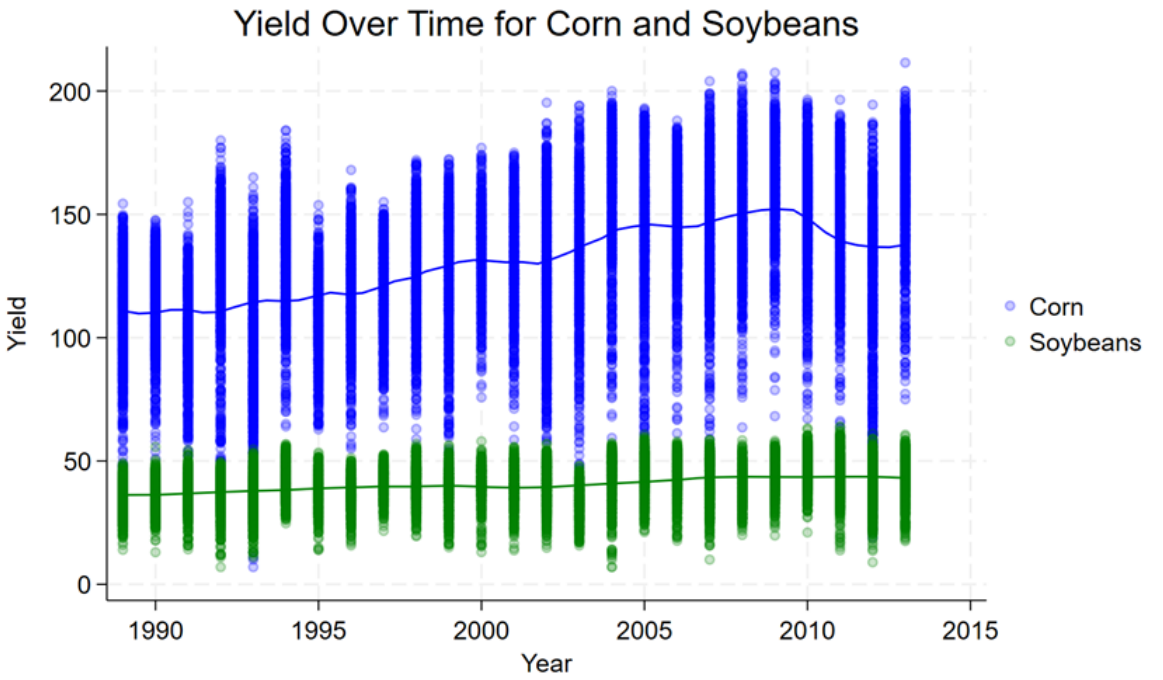


Figure 6: Corn yield is higher and has higher variance compared to soybeans.

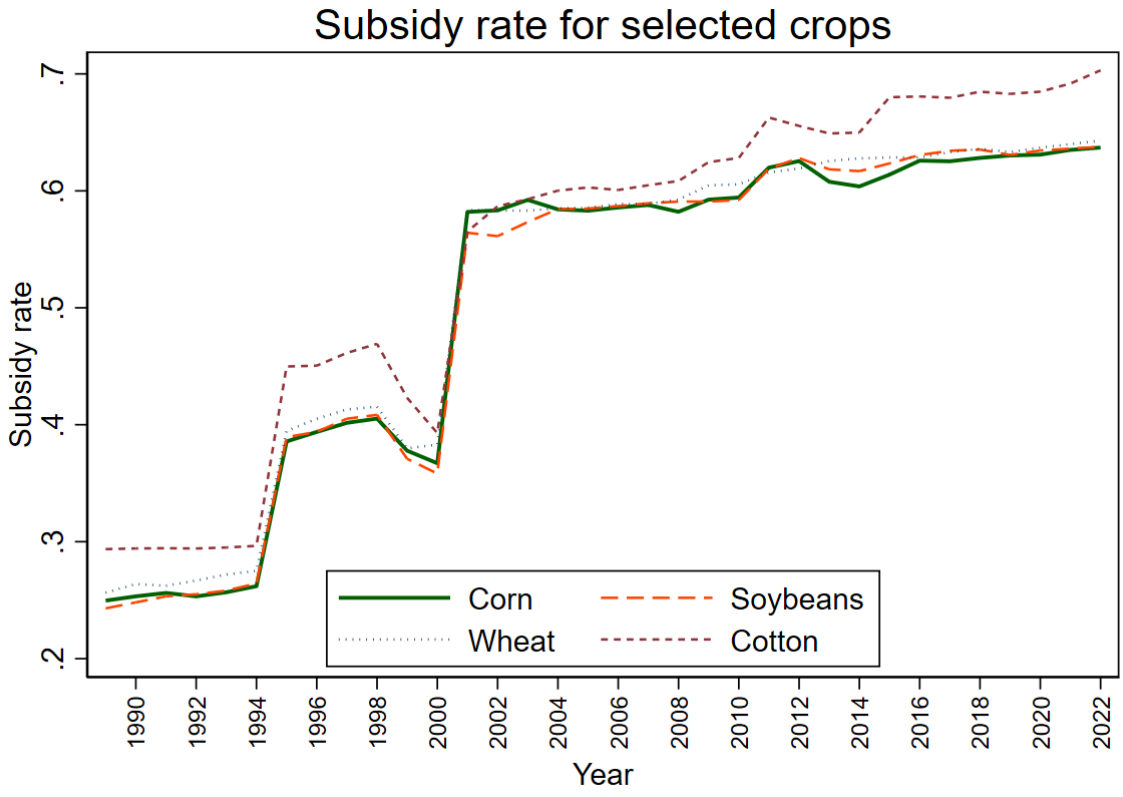
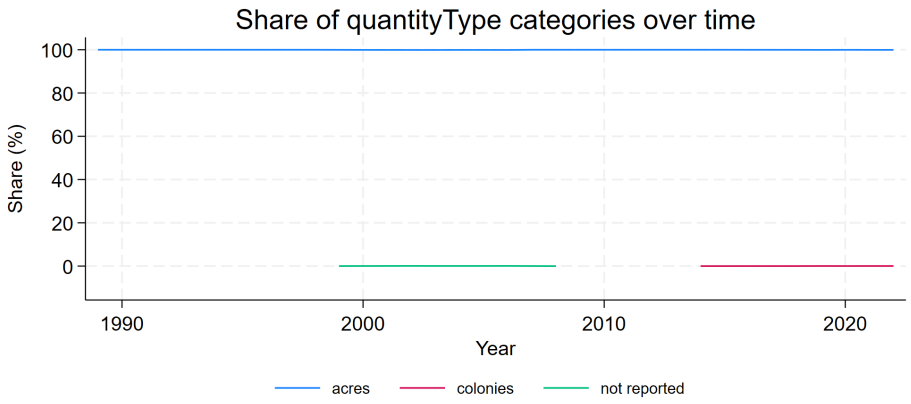
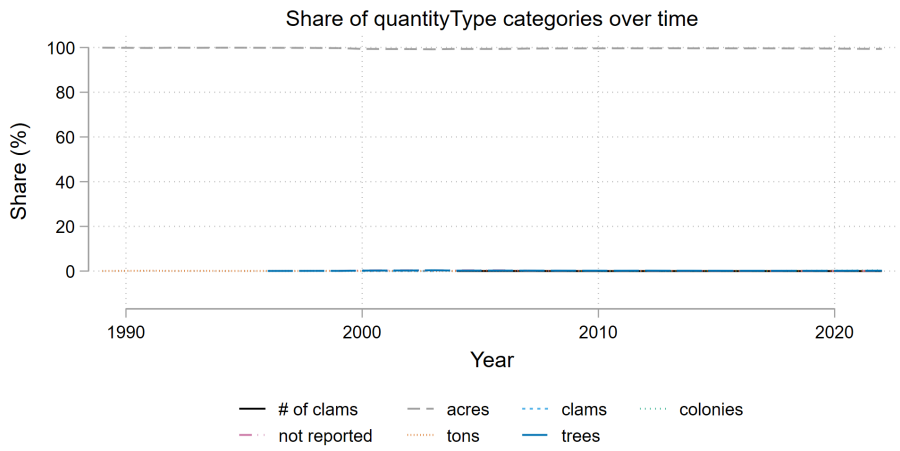


Figure 7: Premium subsidies in yield based on revenue based insurance in all States for corn, soybeans, wheat and cotton

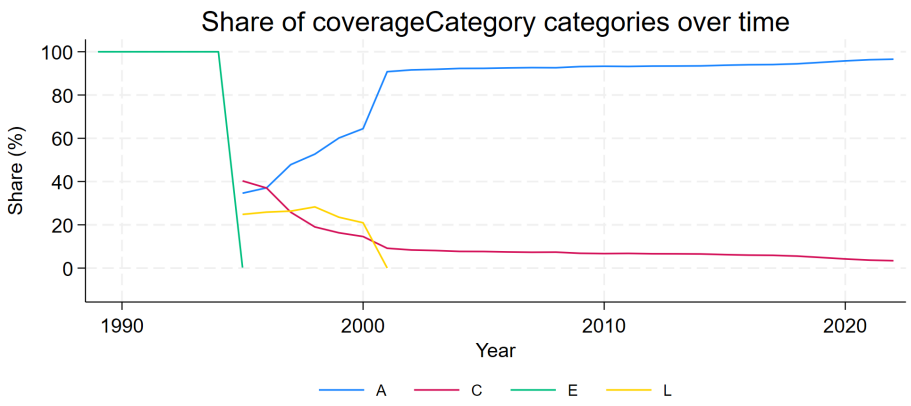


(a) Midwest

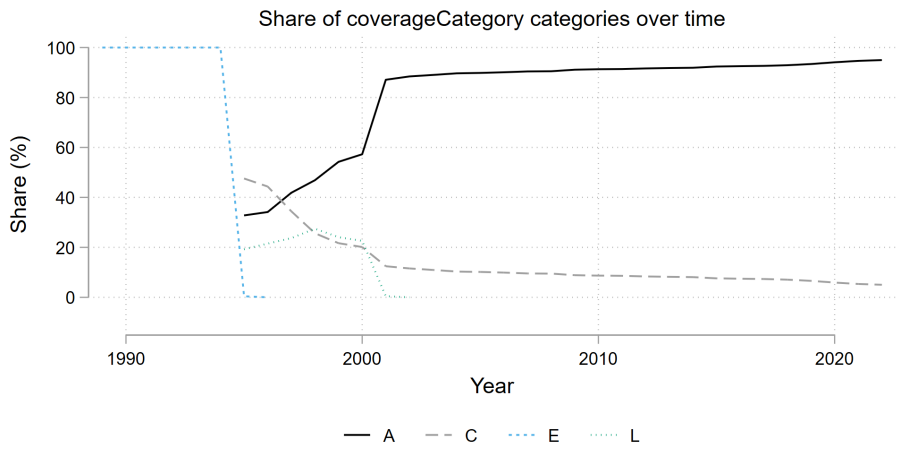


(b) Contingent United States

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

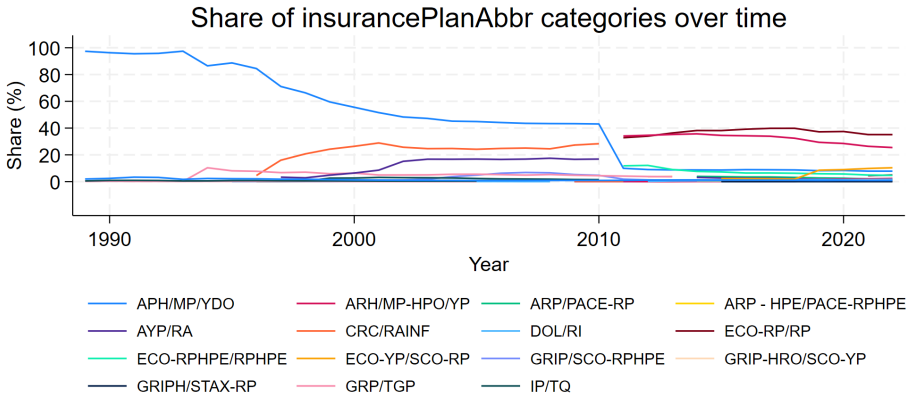


(a) Midwest

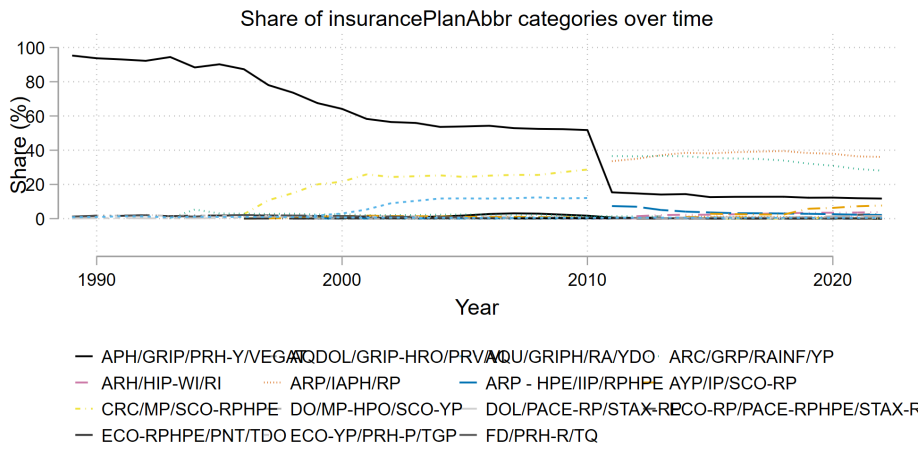


(b) Contingent United States

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

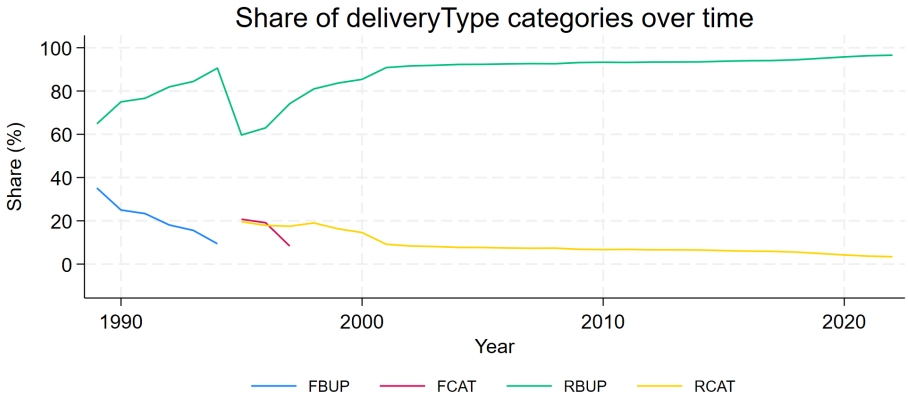


(a) Midwest

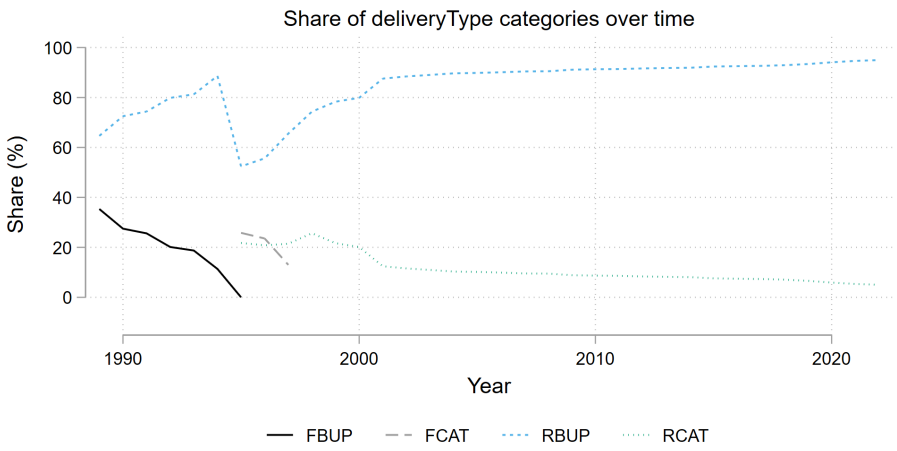


(b) Contingent United States

Figure 10: Share of insurancePlanAbbr categories over time. Data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).

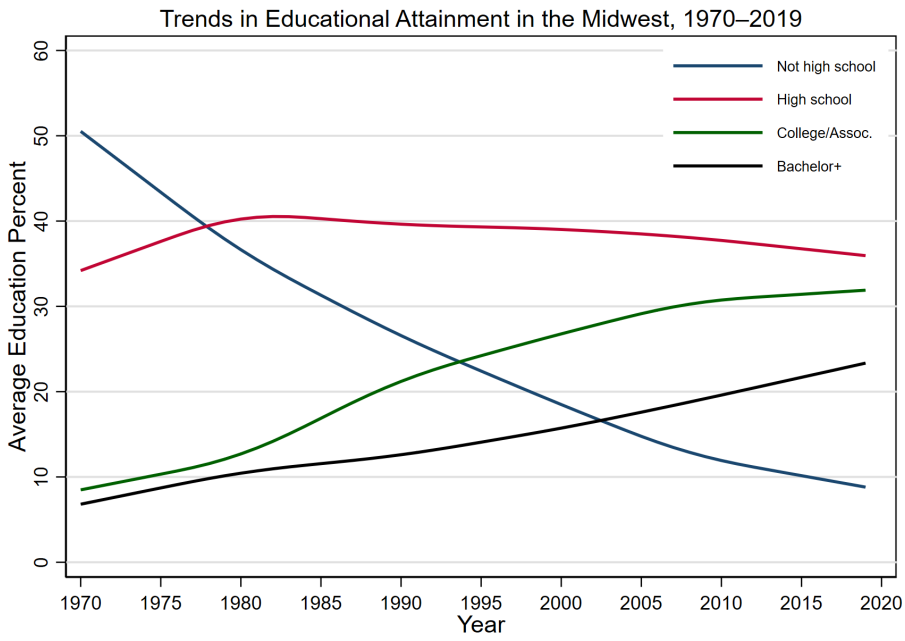


(a) Midwest

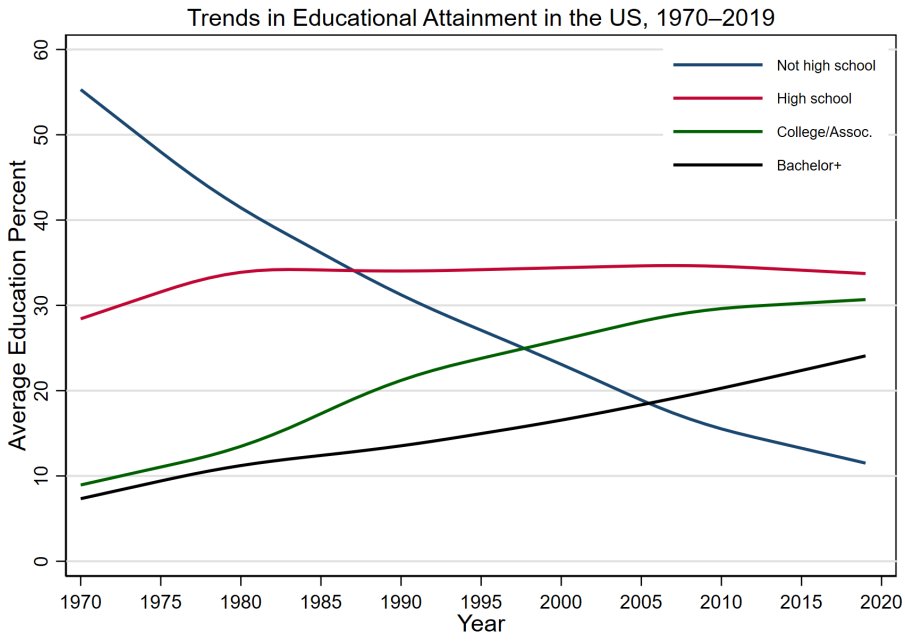


(b) Contingent United States

Figure 11: 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 12: Percent of adults age 25 and older at different levels of education in the Midwest. Data is from USDA Economic Research Service (2025).

3 Data

Demand for Insurance I follow Barry K Goodwin (1993, p. 427) read on Jan 21, 2026 and use two measures for capturing the demand for crop insurance: ratio of insured to planted acres, and liability per planted acre. Liability is the total value of indemnities that the insurer would have to pay in the event of a complete loss. It equals the product of insured acres, price election, yield guarantee and the FCIC's estimated insurable yield (Barry K Goodwin 1993, p. 427). I call the first measure the external margin, and the second measure, the internal margin of insurance enrollment.

I used databases developed by Annan and Schlenker (2015), Jisang Yu, A. Smith, and Sumner (2018), Tsiboe and Turner (2023a) and Tsiboe and Turner (2023b) in this research. I limited the scope of paper to barley, corn, cotton, oats, rice, sorghum, soybeans, and wheat, in eight Midwest states of Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin, over the years 1989 to 2013. This choice of geography and crops was due to the fact that these states are major producers of a wide range of food and animal feed for national consumption and international trade. We also focused on corn and soybeans, as these two main commodity crops are grown on 75% of the arable land in the region (Angel et al. 2018).

Annan and Schlenker (2015) combines the publicly available data on county-level insurance coverage (RMA) with data on county-level acres planted and yields for corn and soybeans (USDA National Agricultural Statistical Service). The two data sources derive county aggregates in different ways: the RMA data is the sum of all actual contracts for a given year in the county. The NASS data on the other hand, uses a representative subsample of all farms in the June Agricultural Survey of each year to approximate the planting area in a county and so has sampling noise. As a result, the fraction insured, the ratio of the insured planted area to the total planting area can be larger than one if the NASS estimate in the denominator is an underestimate of the actual planting area. The fraction of land insured and weather variables are from this data set. The weather variables capture moderate heat (degree days between 10-29 degrees Celsius for corn, and 10-30 degrees Celsius for soybeans), extreme heat (degree days above 29 degrees Celsius for corn and above 30 degrees Celsius for soybeans), amount of precipitation in centimeters, and its squared value, during the months of April to September. All parameters are annual and specific to counties. I used this data set as the basis for building the main data set of analysis, and only focused on counties, crops and years covered in (Annan and Schlenker 2015). With this choice, I eliminated the step for calculating the number of moderate degree days, number of heating degree days and temperature for other counties, crops and years.

Jisang Yu, A. Smith, and Sumner (2018) uses the annual county-level information on crop acreage and crop insurance characteristics for major field crops from surveys of the NASS and the RMA's Summary of Business. For price data, the authors have used two prices: 1) futures price data as reported by the Commodity Research Bureau and then deflated each price by the Producer Price Index from the Bureau of Labor Statistics using May 1986 as the base year (U.S. Department of Labor, Bureau of Labor Statistics 2016). 2) Received price, which is the State-level average received price by farms for each crop from NASS. The authors combined the information from both sources and constructed a price variable called "expected price" using both futures prices and average prices received by farms[2].

[2]To do so, they regressed the state level annual prices received by farms on state dummies and futures prices for 1989 2014. They used the predicted values from this regression as the expected prices.

The county-crop level subsidy per dollar of liability and the expected price of crops are from this data set.

Tsiboe and Turner (2023a) has used the summary of business and the cause of loss data sets of the RMA. The authors constructed a county-crop panel dataset from 1948 to 2020 that includes FCIP loss experience information on liability insured by producers, total premium, premium subsidy, and indemnity paid to farmers. The instrumental variable is from this data set.

Table 4 shows the summary statistics of the variables used in this study. There are observations without a specific county, recorded as "other" in the data set. Since they are not tractable through years, we dropped those observations. It made the panel data set unbalanced.

Table 4: Means and Standard Deviations of County–Crop Panels

Variables	**1989–1993**	**2010–2013**	**1989-2013**
Yield	74.35	92.68	86.34
	(41.40)	(54.66)	(50.39)
Planted Acreage	64,801.05	77,085.56	71,705.23
	(54,860.52)	(58,295.91)	(56,463.10)
Insured Acreage	21,388.70	64,330.86	46,947.02
	(29,829.91)	(52,459.90)	(47,008.81)
Fraction Insured (%)	26.94	80.15	59.99
	(20.13)	(13.70)	(26.61)
Subsidy per Liability	0.02	0.06	0.05
	(0.01)	(0.03)	(0.03)
Hot Degree Days	17.47	33.75	21.09
	(18.58)	(30.79)	(21.72)
Moderate Degree Days	156.52	170.50	162.02
	(29.47)	(27.28)	(27.81)
Precipitation	0.62	0.60	0.60
	(0.17)	(0.16)	(0.14)
Observations	6,568	5,007	32,289

Notes: Means shown; standard deviations in parentheses. Periods correspond to the early and late years in the study period. Variables are county–crop level.

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 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 **??** shows the change in the number of active weather stations over the 1980-2025 period in the 8 midwest states.

Actuarial Data I downloaded the actuarial data of the USDA insurance program from the "Insurance Offer Rate Reports" of U.S. Department of Agriculture, Risk Management Agency (2025). This website reports multiple data: the sub-county rate reports, have data on the base premium rate and rate method code (A, E, C, R and T for actuarial experience, experience blended, county rate, regional rate and transitional rate) by state of irrigation and organic production.

Census block level boundary and population data To do: to change the numbers to 2000 and refer to the correct citations. 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.

$$\text{Population-weighted centroid} = (\bar{x}, \bar{y}) = \left(\frac{\sum_{i=1}^{n} x_i p_i}{\sum_{i=1}^{n} p_i}, \frac{\sum_{i=1}^{n} y_i p_i}{\sum_{i=1}^{n} p_i} \right)$$

where:

- (x_i, y_i) denotes the geographic centroid of census block i,
- p_i is the population of census block i,
- n is the total number of census blocks in the county.

Bulk Agricultural Census and Survey Datasets (Ag census) The USDA National Agricultural Statistics Service (NASS) provides County-, State- and National-level data on both agricultural surveys and quinquennial agricultural censuses for years 1866 onward(United States Department of Agriculture, National Agricultural Statistics Service 2025). This data reports statistics at annual, monthly, seasonal and occasional levels. I downloaded five data sets on crops, demographics, environmental, economics and animals, retained the county-level annual data for years 1989 onward, and appended these data files for further analysis. In census years, this data set reports two lines for multiple measures (i.e. area planted and area harvested), one for census and one for survey. To avoid multiple counting, I use the census reported values by setting the corresponding values of survey data equal to zero (source description = *survey*). I started off with analysis for corn and soybeans, and due to data availability, I expanded the work to wheat, cotton, forage production, corn, sorghum, peanuts, soybeans, and barley (since Jisang Yu, A. Smith, and Sumner (2018) reported values for insurance liability and crop prices for these crops).

I controlled for diversification and savings, as Wright and Hewitt (1994) suggest that the perceived demand for agricultural insurance may be overstated, because farmers can use to cushion the impact of production shortfalls on consumption. I used the following variables from the census in this analysis 6 and 7.

Operators' Work Experience from Ag census: I construct county-year measures of operator experience by computing the share of operators with at least ten years of experience, separately for

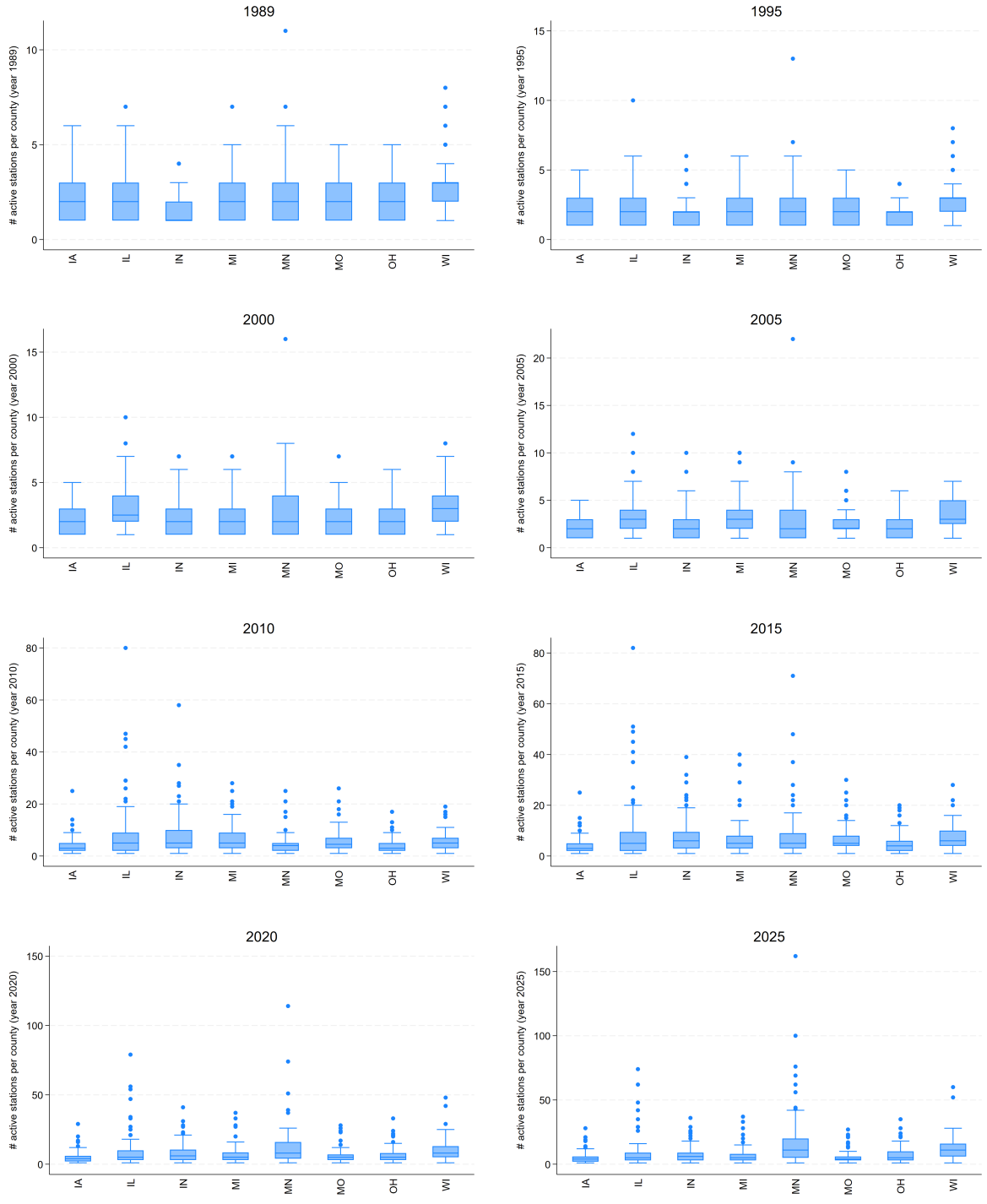


Figure 13: Distribution of the number of weather stations per county by state (1989–2025).

experience on any operation and on the present operation. Since operators' experience applies to all crop they grow, I defined this measure ay the county-year level, and not at he county-crop-year level.

Income Measures from Ag census. I construct two county–year measures of farm income. The primary measure is net cash farm income, defined as cash receipts minus cash expenses, which captures realized cash flow but can take zero or negative values in some county–year observations. To retain these observations and limit the influence of extreme values, I transform net cash farm income using the inverse hyperbolic sine (IHS), which is well defined for zero and negative values and approximates the logarithmic transformation for large positive values (Bellemare and Wichman 2020). Net cash farm income serves as the baseline income measure, while farm-related income is used in robustness analyses.

Access to Marketing Assistance Loans from Ag census. I construct county–year measures of Commodity Credit Corporation (CCC) marketing assistance loan exposure, defining 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 accommodate zero values and reduce right skewness.

Fertilizer Use from Ag census. I construct a county–year measure of fertilizer use by dividing total fertilizer expenditures by total planted acreage, yielding fertilizer expenditure per acre planted. This normalization accounts for differences in county size and serves as a proxy for input intensity. I apply the inverse hyperbolic sine transformation to reduce skewness and retain zero values.

Female Producers from Ag census. I construct a county-year measure of female representation among farmers. I aggregate the number of female and male producers to the county-year level and define the share of female producers as the number of female producers divided by the total number of producers in the county-year. The Census reports data on number of operators and principals as well, that I ignored in my calculations.

Minority Producers from Ag census. I construct a county–year measure of racial and ethnic composition among farmers using producer counts from the Census of Agriculture. I aggregate the number of producers identifying as Black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, multi-race, or Hispanic and define the share of minority producers as this total divided by the total number of producers in the county-year. This variable captures persistent cross-county differences in farmer demographics and is used in levels.

Cattle Intensity from Ag census. I control for livestock intensity using two measures of county–year cattle inventory at level and per planted acre, which captures the prevalence of mixed crop–livestock farming systems that may influence both insurance participation and crop outcomes.

Land Tenuring from Ag census. I construct a county-year measure of land tenure using data on acres operated by ownership status. To do so, I compute the share of operated farmland that is owned by operators as the ratio of acres owned to the total acres operated (owned plus rented from others). This variable captures persistent differences in land tenure and investment incentives across counties and is merged into the county–crop–year panel as a time-varying county-level control.

Construction of Intercensal County-Level Economic Controls for Level Variables As Table 5 shows, several county-level economic controls from the Census of Agriculture are observed only in census years (e.g., 1992, 1997, 2002, 2007, 2012, and 2017). 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 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

Table 5: County-Level Controls Available Only in Census Years

Variable	Source	Level	Years
Area planted	Ag census/survey	County-crop-year	1989–2024
Area harvested	Ag census/survey	County-crop-year	1989–2024
Planted area irrigated	Ag census/survey	County-crop-year	1989–2024
Planted area non-irrigated	Ag census/survey	County-crop-year	1989–2024
Harvested area irrigated	Ag census/survey	County-crop-year	1989–2024
Harvested area non-irrigated	Ag census/survey	County-crop-year	1989–2024
Share with ≥ 10 years on any operation	Ag census	County-year	2017, 2022
Share with ≥ 10 years on present operation	Ag census	County-year	2017, 2022
Net cash farm income (IHS)	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022
Farm-related income (log)	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022
CCC net payments	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022
CCC net payments (IHS)	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022
Share of female producers	Ag census	County-year	2017, 2022
Fertilizer expenditure per acre	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022
Fertilizer expenditure per acre (IHS)	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022
Share of minority producers	Ag census	County-year	2017, 2022
Cattle per planted acre	Ag survey	County-year	1989–2022
Cattle per planted acre (IHS)	Ag survey	County-year	1989–2022
Share of acres owned	Ag census	County-year	1997, 2002, 2007, 2012, 2017, 2022

Notes: Variables listed in this table are observed only in quinquennial Census years. When both Census and survey values are reported for a measure, I retain the Census-reported value to avoid double counting. Area planted and harvested, including irrigated and non-irrigated area, are for the selected crops (barley, corn, cotton, peanuts, sorghum, soybeans, and wheat). All other variables are for all crops.

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.

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 2026c).
- **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 2026b). This preserves 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 2026a) , 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.** Intercensal values are scaled using annual fertilizer price indices reported by the USDA ERS Fertilizer Use and Price dataset. To construct a state-crop level measure of fertilizer use intensity, I use data from the USDA Economic Research Service (ERS) *Fertilizer Use and Price* dataset (U.S. Department of Agriculture, Economic Research Service 2026d). Specifically, I rely on 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 and t index year. For each nutrient $j \in \{N, P, K\}$, the dataset provides:

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

  Because application rates are reported per fertilized acre, I convert them into an average intensity per planted acre by scaling by the share of acreage treated:

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

  where A_{sjt} is expressed as a fraction. The composite fertilizer intensity for crop $c \in \{\text{corn}, \text{soybeans}\}$ in state s and year t is then constructed as:

  $$F_{sct} = I_{sNt} + I_{sPt} + I_{sKt} = \sum_{j \in \{N,P,K\}} A_{sjt} R_{sjt}.$$

  This measure captures the total pounds of primary nutrients applied per planted acre, integrating both the extensive margin (acreage receiving fertilizer) and the intensive margin (application rate conditional on use). All nutrient quantities are measured in pounds of nutrient per acre. The resulting composite measure provides a state–crop level indicator of fertilizer input intensity suitable for benchmarking county-level fertilizer expenditure or input-use variables.

- **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).

Construction of Intercensal County-Level Economic Controls for Share Variables Several county-level structural controls are expressed as shares bounded between zero and one, including: (i) the share of producers with at least ten years of experience on any operation, (ii) the share of producers with at least ten years on their present operation, (iii) the share of female producers, (iv) the share of minority producers, and (v) the share of operated acres owned by operators. To construct annual intercensal series while respecting the unit-interval constraint, I implement temporal disaggregation in logit space. To benchmark each share variable, I employ an economically relevant annual state-level indicator that captures systematic variation in the underlying structural margin. Let S_{cT} denote the observed county-level share in census year T, and let Z_{st} denote the corresponding annual state-level indicator. The share is first mapped into logit space:

$$\tilde{S}_{cT} = \log\left(\frac{S_{cT}}{1 - S_{cT}}\right).$$

Intercensal dynamics are imposed proportionally in logit space using the state-level indicator:

$$\tilde{S}_{ct} = \tilde{S}_{cT} + (\log Z_{st} - \log Z_{sT}),$$

and the bounded share is recovered via the inverse-logit transformation:

$$S_{ct} = \frac{\exp(\tilde{S}_{ct})}{1 + \exp(\tilde{S}_{ct})}.$$

This procedure preserves the unit-interval constraint while allowing county-level shares to evolve in proportion to economically meaningful state-level trends.

The benchmark indicators are selected as follows:

- **Share of producers with at least ten years of experience on any operation (*share_any_ge10*).** Benchmarked to the annual state-level average age of principal operators from the USDA National Agricultural Statistics Service (NASS) Quick Stats database, which captures cohort aging and structural turnover dynamics in the farm sector (USDA National Agricultural Statistics Service 2026).
- **Share of producers with at least ten years on their present operation (*share_pres_ge10*).** Benchmarked to state-level farm tenure dynamics using annual state-level farm counts from NASS Quick Stats, reflecting entry and exit patterns that influence operator tenure distribution (USDA National Agricultural Statistics Service 2026).
- **Share of female producers (*share_female_cy*).** Benchmarked to annual state-level female principal operator shares derived from NASS demographic statistics, capturing secular trends in gender composition of farm operators (USDA National Agricultural Statistics Service 2026).
- **Share of minority producers (*share_minority_prod*).** Benchmarked to annual state-level minority operator shares from NASS demographic data, reflecting demographic shifts in agricultural participation (USDA National Agricultural Statistics Service 2026).
- **Share of operated acres owned by operators (*share_acres_owned*).** Benchmarked to annual state-level real estate asset values and land tenure indicators from the USDA Economic Research Service (ERS) Farm Income and Wealth Statistics, which capture trends in farmland ownership structure (**usda_ers_farmincome**).

By anchoring intercensal county dynamics to state-level demographic and structural trends, this approach preserves observed census benchmarks while avoiding linear interpolation and respecting the bounded nature of share variables.

Number of Heating Degree Days, Moderate Degree Days, Cooling Degree Days, and Rainfall I built up on the method of Annan and Schlenker (2015) and calculated the HDD and MDD for the period of study for the contiguous United States. Unlike the past efforts for capturing the impact of climate variables on agriculture output that only included the *cropping-season* values, I followed Farris et al. (2023) and studied the impact of these climate variables in over the year, and over the cropping season.

Table 6: Variables Included in the Regressions

Variable	Description
Insurance outcomes	
Share insured (any)	Share of planted acres insured under any crop insurance product.
Share insured (CAT)	Share of planted acres insured under catastrophic coverage.
Share insured (buy-up)	Share of planted acres insured under buy-up coverage.
Liability per acre	Total liability divided by insured acres.
Liability per acre (CAT)	Liability per insured acre for catastrophic coverage.
Liability per acre (buy-up)	Liability per insured acre for buy-up coverage.
Farmer-paid premium per acre	Producer-paid premium divided by insured acres.
Subsidy rate	Premium subsidy amount divided by total premium amount.
Coverage level	Selected insurance coverage level for buy-up contracts.
Agricultural production and land use	
Area planted	County–crop planted acreage.
Area harvested	County–crop harvested acreage.
Yield	Output per planted or harvested acre, depending on specification.
Revenue per acre	Crop revenue normalized by acreage.
Planted area irrigated	Irrigated planted acreage for the crop.
Planted area non-irrigated	Non-irrigated planted acreage for the crop.
Harvested area irrigated	Irrigated harvested acreage for the crop.
Harvested area non-irrigated	Non-irrigated harvested acreage for the crop.
Cattle per planted acre	County cattle inventory normalized by planted acreage.
Share of acres owned	Share of operated farmland that is owned.
Weather and climate controls	
Moderate degree days	Growing-season degree days in moderate temperature bins.
Extreme heat degree days	Growing-season degree days above the extreme-heat threshold.
Precipitation	Growing-season cumulative precipitation.
Precipitation squared	Square of cumulative precipitation, allowing nonlinear rainfall effects.
Flood exposure	Share of county area or cropland exposed to flooding.
Weather station density	Number of weather stations or related measure of local weather-data availability.
Prices, income, and government support	
Price received	Crop-year output price received by farmers.
Net cash farm income (IHS)	Inverse hyperbolic sine transformation of county-level net cash farm income.
Farm-related income (log)	Logarithm of county-level farm-related income.
CCC net payments	County-level Commodity Credit Corporation payments.
CCC net payments (IHS)	Inverse hyperbolic sine transformation of CCC net payments.
Fertilizer expenditure per acre	Fertilizer expenditures normalized by acreage.
Fertilizer expenditure per acre (IHS)	Inverse hyperbolic sine transformation of fertilizer expenditure per acre.
Loan deficiency payments	County-level loan deficiency payments or related USDA support measure.
Demographic and structural controls	
Educational attainment	Share or level of adult educational attainment in the county.
Share female producers	Share of producers identified as female.
Share minority producers	Share of producers identified as minority producers.
Experience on any operation	Share with at least 10 years on any farm operation.
Experience on present operation	Share with at least 10 years on the present operation.
Instruments and identification variables	
Distance to USDA office	Distance from county centroid to nearest USDA or service center office.
Lagged yield	Lagged county–crop yield, used in selected specifications.
Volatility factor	RMA volatility or rating-related adjustment used in premium-setting.
Expected yield / reference yield	Administrative yield measure entering premium determination.
Fixed effects and panel structure	
County fixed effects	Control for time-invariant county characteristics.
Year fixed effects	Control for aggregate shocks common across counties.
Crop fixed effects	Control for time-invariant crop-specific differences.
State-by-year fixed effects	Control for state-level time-varying shocks.
County trends	County-specific linear trends in selected specifications.

Notes: This table summarizes the main variables used across the regression specifications. Exact inclusion varies by model. Insurance outcomes are constructed from USDA Risk Management Agency Summary of Business data. Agricultural production, land use, and several structural controls are drawn from USDA NASS and Census of Agriculture sources. Weather controls are based on temperature and precipitation measures constructed at the county–year or county–crop–year level. Variables reported only in Census years are interpolated or otherwise incorporated as described in the text where applicable. IHS denotes the inverse hyperbolic sine transformation.

Table 7: Descriptive statistics of major explanatory variables to be updated

	All coverage		Buy-up		CAT	
Variable	Mean	SD	Mean	SD	Mean	SD
Crop insurance	0.022	0.148	0.015	0.122	0.120	0.325
Socioeconomic variables						
Age	57.890	14.462	57.816	14.458	59.017	14.491
Male	0.753	0.431	0.751	0.433	0.796	0.403
Reside in plot	0.748	0.434	0.772	0.420	0.441	0.497
Education	1.909	1.761	1.784	1.609	3.806	2.655
Dependence	1.411	1.219	1.371	1.214	1.940	1.176
Institutional variables						
Credit	0.084	0.277	0.060	0.237	0.394	0.489
Secure tenure (% hectares)	0.851	0.329	0.843	0.337	0.948	0.196
Extension services	0.220	0.414	0.226	0.419	0.141	0.348
Participation in organisations	0.204	0.403	0.184	0.388	0.457	0.498
Number insurance adopter	25.374	35.352	24.641	34.875	34.781	39.801
Farm size						
Yield (kg per hectare)	2,814.4	3,315.3	2,682.4	3,349.4	4,505.5	2,234.5
Total surface (hectares)	38.063	194.710	17.199	22.458	305.503	663.655
Capital	0.083	1.208	−0.092	0.585	2.317	3.226
Technology adoption						
Certified seed	0.213	0.409	0.180	0.384	0.628	0.483
Biological control	0.038	0.191	0.030	0.170	0.142	0.350
Modern irrigation	0.050	0.218	0.041	0.198	0.171	0.377
Environmental and location variables						
Rainfall (mm)	911.195	917.788	323.163	826.687	366.860	—
Soil quality (ratio)	0.283	0.189	0.283	0.186	0.289	0.223
Location variables						
North zone	0.007	0.081	0.005	0.073	0.024	0.153
Central zone	0.159	0.366	0.145	0.352	0.352	0.478
South zone	0.833	0.373	0.850	0.357	0.624	0.485
Observations	42,531		39,874		2,642	

Source: Authors' estimates based on 2007 Census data.

Summary of Business Data Many farmers insured under the Federal program, particularly in the mid 1990s, held only catastrophic coverage (CAT) for their crops. CAT coverage is fully subsidized by the government, requiring only a nominal administrative fee (typically $50 or $100, depending on the year) for each crop a farmer chooses to enroll in a given county, regardless of acreage total. Coverage provided by CAT, however, is very low, typically covering losses in excess of 50 percent of expected yield at 55 percent of the price. To focus on insurance more likely to influence behavior, we define the *insured* as those having coverage in excess of CAT coverage, generically referred to as "buy up" coverage in much of the literature. I followed the approach used by Roberts, E. J. O'Donoghue, and Key (2007) and considered all insurance coverage above the CAT level "insured" and considered CAT or no insurance as "uninsured." This paragraph above is also from Roberts, E. J. O'Donoghue, and Key (2007) and needs rewriting.

4 Methodology

I use six different means to measure participation in the crop insurance program: percent of planted acres enrolled, level of total liability per acre, total acres enrolled, number of acres enrolled in in buy-up coverage[3], level of total premiums, and level of total premiums per acre [4]. The level of total premiums is the first once actually included? it was on page 14 of this paper, level of total premiums per acre, level of total liability per acre and percent of planted area that is insured, capture the intensity of use of crop insurance. Large movements in crop prices can cause these variables to change from year to year while yields tend to trend upwards over time with technology advancement, making it difficult to compare a policy purchased in one year with a policy purchased in a subsequent year. For this reason, I normalized these measures throughout the the analysis to 2002 prices and yields. With this, if the value of total premiums or liabilities changed over time, it would be due to the underlying changes in the quantity of crop insurance demanded and would not be attributed to price or yield changes. The premiums, premiums per acre, and liabilities per acre were all multiplied by the ratio of 2002 expected revenues to 1997 expected revenues (expected yields are derived from a linear regression of normalized yields from 1966 through 2001 on a time dummy for 2002 and expected prices come from Chicago Board of Trade futures prices for the relevant crop's harvest price at the time of planting for the 2002 crop; similar calculations were used for the 1997 prices and yields) (E. O'Donoghue 2014)..

I defined a transfer payment multiplier for crop subsidies that government offers to the farmers. The transfer payment multiplier in the context of crop subsidies refers to the ratio of the total income generated in the recipient economy to the amount of subsidy transferred. In other words, it measures how much additional income is created for farmers (and sometimes the broader rural economy) for every unit of government subsidy provided. Multiplier at horizon h is

$$\mathbf{M}_h \equiv \Delta Y_{t+h} / \Delta T_t \tag{1}$$

where T_t is government transfers to farmers and Y is output (GDP, local earnings, etc.).

Choose of Variables I used Lasso method (Tibshirani 1996) for choosing the relevant variables out of the pool of covariates I had.

Should I have the insurance subsidy percentage as the shock, or the percentage changed compared

[3] those policies above the fully subsidized catastrophic coverage available to farmers

[4] In the earlier versions of this paper, I measured participation in crop insurance program only using the first two parameters. I expanded the list following the method of E. O'Donoghue (2014).

to the baseline year, say, 1989? Many policy changes have taked place over the time. How can I single out the effect of 1 policy out of all others? Define policy-episode shocks tied to legal changes (e.g., 1994, 2000, 2014 reforms), and measure each unit's exposure to that episode using formula-driven, pre-policy characteristics. Then run event-study or continuous-treatment DiD around those episodes.

The PROCAMPO program in Mexico is a well-documented example for transfer payment multiplier. Studies found that the income multiplier for recipients ranged from 1.5 to 2.6, meaning each peso transferred generated 1.5 to 2.6 pesos in total income for farm households, especially where liquidity constraints were significant 216. In China, the grain subsidy program was also analyzed for its multiplier effects, with findings indicating a high income multiplier and significant income promotion effects for farmers 1. These multipliers are typically higher for larger or better-endowed farms and in regions with more economic opportunities, reflecting the role of liquidity constraints and asset endowments 216. See F. Yi, W. Lu, and Y. Zhou (2016) and Sadoulet, De Janvry, and Davis (2001)

Dependent Variables The outcome of interest is the change in crop insurance demand measured one of six ways: percent of planted acres enrolled, level of total liability per acre, total acres enrolled, number of acres enrolled in in buy-up coverage, level of total premiums, and level of total premiums per acre.

Control Variables The control variables in this study that vary over time are

- The lagged change in the number of acres of a particular crop in the county,
- The lagged change in the number of acres run by female operators. Since it has been shown that the gender of the individual matters when it comes to dealing with risk (Halek and Eisenhauer 2001; Pawlowski, Atwal, and Dunbar 2008; Byrnes, D. C. Miller, and Schafer 1999).
- The 1-year lagged change in the ratio of yield policies to revenue policies. what does it mean? It is from page 18 of E. O'Donoghue (2014).
- The mean and variance of yield distribution: to explore any changes to the yield distribution, I use county-level yield data and detrends it and normalizes it to a base year's yield (1989) for each crop. This allows the yields to be compared over time and viewed as a distribution of yields for each year examined. I then constructed the mean and variance of these yield distributions and differenced to control for any changes in the yield distribution that may have induced changes in crop insurance participation. I adopted this treatment of the crop yield before using in the analysis from E. O'Donoghue (2014, pp.18-20).
- The change in a 3-year measure of lagged returns to crop insurance measured as total indemnities divided by total premiums paid by the farmer
- The difference in a 1-year lagged, actual-versus-expected revenue, differenced over time.

I used the last two sets of variables to capture the general state of affairs in the years (or year) leading up to the period examined, as recommended by E. O'Donoghue (2014, page 18). In terms of time-invariant control variables, because the regression is run at the county level, time-invariant county-level controls are controlled for through the differencing. Essentially, differencing acts as a county-level fixed effect, which eliminates the need to include variables that do not change over time, such as farmers' age, average farm size, average county size, and land quality. To do: variables to add

when dependent variable that is not a measure of change. The regression also includes region-by-year fixed effects This variable is from E. O'Donoghue (2014). I think region is from the Quickstats dataset that generate comparisons among counties within regions that were created based on soil and climatic attributes (crop reporting districts). Note that this is a fixed effect that captures trends that can vary by region. Any changes that differ across regions will be captured by these variables, such as weather shocks and price movements not picked up by other variables. Finally, the error, captures other unobserved factors affecting crop insurance demand, such as within-region weather variations.

A third group of control variables address the alternative methods for farmers to deal with risk, which might affect producers' willingness to buy crop insurance, including the PLC program and the ad-hoc disaster assistance. Inclusion of these variables sets this paper apart of papers like E. O'Donoghue (2014) that narrow the time window of study and limit the need for considering other concurrent policies.

Variables construction These paragraphs are copied from pages 20-21 of E. O'Donoghue (2014)The mean and variance of the yields are calculated using NASS yields collected from 1975 through 2002. For each county, yields are first detrended using a simple linear model, regressing the 27 years of data on a year variable. Following the study of Barry K Goodwin and Alan P Ker (1998), who found that the standard deviations of the yield tend to be proportional to the level of the average yield, normalized yields were created using the intercept, slope, and residuals from the regression in the following manner:

$$\bar{y}_t = y_{1989} \times (1 + \frac{e_t}{y_t}) \tag{2}$$

where $\bar{y}_t$ denotes the normalized yield for time t, e_t represents the residual from the regressionI think it means the detrending regression, and y_t is the predicted yield from the linear regression. With 27 years of data, equation 2 generates normalized yield observations for each county, allowing one to calculate a separate mean and variance for crop yields for each county. Distributions are constructed for both 1996 and 2001, and means and variances are constructed for each of these lagged years. They are then differenced to provide a change in the mean and variance of the detrended, normalized county-level yields.

The 3-year returns to crop insurance variable was constructed by dividing the indemnities by the premium paid by the producer for each insurance plan for the relevant crop and summing them together, weighted by their share of total acres enrolled in each plan. This is done for each of the 3 years preceding 1997 and 2002 (e.g., for 1997, the years 1996, 1995, and 1994 were used). The returns for the 3 years were then averaged to obtain two, 3-year average returns to crop insurance for each crop in each State (one for the years leading up to 1997; a second for the years leading up to 2002). These were then differenced and used in the regression.

If a producer experienced a loss in the year prior to that examined, he or she may be more inclined to enroll in crop insurance in the following year. To observe this over time, if a farmer fared worse in 2001 relative to 1996, he or she might find crop insurance more attractive in 2001 than in 1997. Therefore, to construct such a variable, the actual and expected revenues were calculated for 1996 and 2001. Actual revenues were generated using NASS price and yield data. Expected revenues were generated using the predicted yields from the detrending linear regression process discussed earlier and national level futures commodity prices (assuming away basis differences between counties). After the actual and expected revenues for 1996 and 2001 were constructed, they were differenced. The difference is the gain/loss for the year. The resulting gain/loss for 1996 was then subtracted from that

of 2001 to obtain a measure of relative gain/loss over time.

The crop insurance ratio used in the analysis was constructed using RMA data from the Summary of Business. The total premiums of all yield policies were summed up and divided by the total level of the sum of all yield policy and revenue policy premiums. Finally, when running the regressions, if a county had a change in total premiums per acre, change in total liabilities per acre, or a change in total insured acres that fell outside of four standard deviations from the mean of those variables, the observation (county) was omitted from the analysis. Results were also run (but not reported—available upon request) without omitting any variables, and results remained very similar.

Endogeneity I use the introduction of the premium subsidies as the source of price variation. This necessitates the careful examination of the relationship between the demand for crop insurance and the level of subsidies to derive causality. I am trying to identify a change in subsidy levels due to the policy change from the result of a change in demand and as a result, unrelated to the new policy. The second scenario is also probable, since subsidy levels are a function of the level of coverage demanded by the producer.

4.1 Instrumental Variable (IV)

I use an instrumental variables framework to decompose the change in crop yield into two channels: extensive margin of enrollment in crop insurance (by fraction of planted land insured), intensive margin of enrollment in crop insurance (amount of insurance liabilities of farmers). I control for climate variables and other demography variables in the regression as well. I use instruments that provide plausibly exogenous variation in intensive and extensive margins, and I allow the effects to vary for areas by the amount of exposure to extreme weather conditions. 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.

1- Exposure to the 1980s Drought I construct county-level measures of historical drought exposure from the NOAA monthly Palmer Drought Severity Index (PDSI). For each county-year, I first compute the growing-season PDSI as the average of June through September monthly values,

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

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

$$\text{DroughtSeverity}_{ct} = \max\{0,\ -2 - \text{GS_PDSI}_{ct}\}.$$

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:

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

Table 8: Covariate balance test, part I: weather covariates

	(1) MDD	(2) HDD	(3) Precip	(4) Precip^2
iv_drive_post1994	-0.064*	-0.193**	0.149**	20.731**
	(0.037)	(0.078)	(0.061)	(8.633)
iv_drive_post1996	0.012	0.192	0.005	2.006
	(0.030)	(0.136)	(0.071)	(10.589)
iv_drive_post2000	0.034	0.038	-0.132	-14.819
	(0.047)	(0.097)	(0.087)	(9.791)
iv_drive_post2002	-0.040	-0.174	0.127	13.539
	(0.057)	(0.113)	(0.138)	(15.371)
iv_drive_post2008	-0.017	0.173	0.020	3.938
	(0.054)	(0.172)	(0.109)	(15.713)
iv_stn1989_post1994	-0.071	-0.231	0.638	82.628
	(0.358)	(0.363)	(0.434)	(51.846)
iv_stn1989_post1996	-0.155	-0.192	-0.168	-29.210
	(0.233)	(0.346)	(0.554)	(69.436)
iv_stn1989_post2000	-0.170	-0.088	-0.550	-69.098
	(0.269)	(0.497)	(0.491)	(60.054)
iv_stn1989_post2002	0.238	-0.253	0.427	66.702
	(0.285)	(0.546)	(0.600)	(72.904)
iv_stn1989_post2008	0.307	0.715*	0.311	42.357
	(0.201)	(0.383)	(0.524)	(69.727)
iv_drought1980_post1994	0.269	0.418	-0.354	-45.513
	(0.275)	(0.373)	(0.623)	(71.635)
iv_drought1980_post1996	-0.202	-0.016	0.627	52.797
	(0.249)	(0.437)	(0.584)	(64.663)
iv_drought1980_post2000	-0.355	-0.226	0.294	43.018
	(0.233)	(0.310)	(0.311)	(41.412)
iv_drought1980_post2008	0.711**	-0.729	-0.647	-69.593
	(0.342)	(0.475)	(0.501)	(68.250)
iv_droughtcum8890_post1994	0.437	-0.031	-0.024	-5.367
	(0.334)	(0.343)	(0.558)	(60.584)
iv_droughtcum8890_post1996	-0.422*	-0.347	0.319	2.747
	(0.219)	(0.338)	(0.487)	(50.680)
iv_droughtcum8890_post2000	-0.079	0.533	0.247	60.527
	(0.241)	(0.324)	(0.301)	(40.294)
iv_droughtcum8890_post2008	0.185	-1.453**	-0.315	-49.989
	(0.266)	(0.594)	(0.444)	(60.591)
iv_droughtavg8890_post1994	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
iv_droughtavg8890_post1996	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
iv_droughtavg8890_post2000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
iv_droughtavg8890_post2008	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)
iv_droughtpeak8890_post1994	-0.829*	0.051	0.058	60.997
	(0.458)	(0.530)	(0.815)	(90.994)
iv_droughtpeak8890_post1996	0.727**	0.893	0.099	58.582

Table 9: Covariate balance test, part II: income and loan covariates

	(1) Farm income	(2) Net cash income	(3) CCC loans
iv_drive_post1994	0.000	0.000	0.000
	(.)	(.)	(.)
iv_drive_post1996	0.000	0.000	0.000
	(.)	(.)	(.)
iv_drive_post2000	0.002	-0.004	0.001
	(0.003)	(0.012)	(0.004)
iv_drive_post2002	-0.003	-0.011	-0.002
	(0.003)	(0.007)	(0.007)
iv_drive_post2008	-0.013**	0.005	0.030***
	(0.005)	(0.006)	(0.009)
iv_stn1989_post1994	0.000	0.000	0.000
	(.)	(.)	(.)
iv_stn1989_post1996	0.000	0.000	0.000
	(.)	(.)	(.)
iv_stn1989_post2000	-0.024	-0.020	0.036
	(0.018)	(0.064)	(0.039)
iv_stn1989_post2002	-0.038**	0.044	-0.058
	(0.019)	(0.048)	(0.055)
iv_stn1989_post2008	-0.017	0.019	-0.022
	(0.034)	(0.062)	(0.071)
iv_drought1980_post1994	0.000	0.000	0.000
	(.)	(.)	(.)
iv_drought1980_post1996	0.000	0.000	0.000
	(.)	(.)	(.)
iv_drought1980_post2000	0.010	0.025	-0.031
	(0.040)	(0.135)	(0.037)
iv_drought1980_post2008	-0.033	0.075	-0.071
	(0.034)	(0.140)	(0.066)
iv_droughtcum8890_post1994	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtcum8890_post1996	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtcum8890_post2000	0.012	-0.030	0.031
	(0.044)	(0.130)	(0.035)
iv_droughtcum8890_post2008	0.010	0.087	-0.120**
	(0.051)	(0.137)	(0.058)
iv_droughtavg8890_post1994	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtavg8890_post1996	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtavg8890_post2000	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtavg8890_post2008	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtpeak8890_post1994	0.000	0.000	0.000
	(.)	(.)	(.)
iv_droughtpeak8890_post1996	0.000	0.000	0.000

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

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.

1- Past years as an instrument. I used the change in subsidy rate between the previous two rounds of Farm Bills as an instrument for the immediate following Farm Bill. Take the example of 1994 and 1996 Farm Bills. In order to study the effect of changes in subsidy rates between 1995 and 1996, I instrument for subsidy change with the change in subsidy between the 1985 and 1994 Farm Bills, and use the difference between subsidy rates in 1994 and 1989 (first year in our data set). This is a valid instrument, because by the time farmers decide about crop insurance in crop year 1995, the decision about crop insurance in 1994 is already made. Therefore, the decision for 1994 crop year and its change from 1989 can be considered exogenous to the decision about quantity of crop insurance demanded in 1995.

2- Actuarial rating parameters of the USDA as an instrument. I used the actuarial rating parameters of the USDA RMA as instrument. These govern premium rates in a way that is plausibly exogenous to any individual farmer's crop insurance decisions (Woodard and J. Yi 2020; Tsiboe and Turner 2023b).

explain the instrument more. Is there a graph you make to illustrate the first stage? A standard test of an IV is a covariance balance test. Include it. Show you F-stats for your first-stage regressions.

The decision of farmers over the acreage of land to insure is a function of the insurance price (= premium rate) they face. This price itself is set as a function of risk for a given insurance pool[5], as well as the producer's production choices about the level of insurance coverage. This is because an insurance pool with a higher risk of initiating an indemnity payment has a higher premium rate for a given crop insurance choice. This creates a case for endogeneity and simultaneity problems. Farmers' decisions about production, insurance purchase, and the price they face for insurance are intertwined, through a link of unmeasurable "behavioral" characteristics of farmers.

To address this endogeneity problem, we adopt the identification strategy developed by Tsiboe and Turner (2023a) and Tsiboe and Turner (2023b). The authors rely on the rating framework used by RMA (i.e., continuous rating) and develop an instrumental variable (τ). The main idea is that FCIP premium rates are determined using exogenously set rating parameters. This lets us use the parameters to generate exogenous variation in premium rates that producers face when making crop insurance decisions. Tsiboe and Turner (2023a) and Tsiboe and Turner (2023b) estimate the parameters used by the RMA and then use these values as instrumental variables to resolve the endogeneity problem of farmer's decision about farming and purchasing crop insurance.

RMA updates rating parameters on a 3-year cycle using historic loss experience data from the previous 20 years starting from two crop years before the current crop year (Tsiboe and Turner 2023b). The previous yields reported by producers contribute to the producer's historical loss experience but are fixed and not under the control of producers once reported. Consequently, FCIP premium rates can be thought of as sequentially exogenous to producer insurance decisions. This assumption is plausible because the rating parameters are finalized in advance. This exogenous portion of PCIP premium rates is driven by the continuous rating parameter and the endogenous portion by the insured's choice of insurance unit, coverage level, and their historic production experience.

4- Density of weather observation stations. I exploit spatial variation in the historical density of weather observation stations. 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 14 shows the number of weather stations in the study area in the base year of 1989.

6- Shortest/mean/median driving distance between the County centroids and USDA offices as an instrument. Given the historical nature of the period of this study, and the requirement for in-person processing of the insurance contract by farmers., access to the USDA local offices could be affected by the proximity or distance of farmers from those offices. This is a classic "geographic access to public service" problem. To my knowledge, the only prior work that has studied the impact of distance from USDA offices on farmers' choices, is Diekmann et al. (2012). Diekmann et al.

[5]RMA defines an insurance pool as a group of producers formed by the unique combination of county, crop, crop type, and production practices (Tsiboe and Turner 2023b).

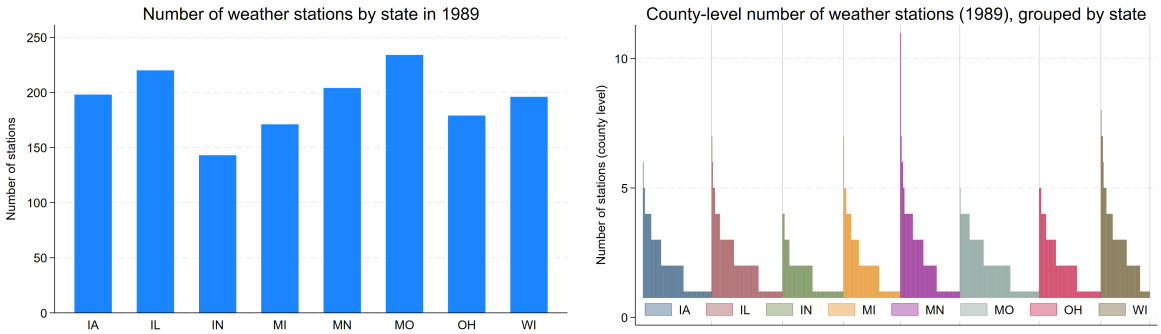


Figure 14: Number of weather stations by state in 1989 (authors' calculation using NOAA National Centers for Environmental Information (2025b) and NOAA National Centers for Environmental Information (2025a)).

(2012) used survey data on Ohio farmers to study the trade offs that farmers are willing to make when choosing between local agriculture extension offices based on distance to the local office, differences in educational programs offered, expertise level of staffing educators, and the level of local taxation required to support an extension office.

To quantify geographic access to USDA service centers, I construct a measure of the road travel distance between each county's geographical center and the nearest USDA office using the contemporary roads network[6]. First, I use 1990 county boundary shapefiles from the U.S. Census and compute the geographic centroid of each county. I pair these with the geocoded locations of USDA offices in the six Midwestern states of Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio and Wisconsin. For each state, I extract the full road network from OpenStreetMap using OSMnx (in Python) and build a state-specific drivable graph. I then project both county centroids and USDA offices onto the nearest nodes of that road network and compute the shortest road distance between each county and every USDA office using Dijkstra's algorithm (Dijkstra 1959) with distance as the weight. For computational stability, states are processed independently, and the resulting distances are saved separately before being merged into a unified dataset. The final variable, *distance to nearest USDA office*, captures minimum drivable kilometers and serves as a proxy for a county's access to USDA administrative services. How about calculating the minimum, mean and median driving distance and experiment with them? Because the IV with minimum driving distance turned out to be weak.

Following standard IV diagnostics, I perform a *covariate balance test* test to verify whether the instrument (distance to USDA office) predicts predetermined covariates. I regress each covariate on the instrument with county and year fixed effects. Because the instrument is time-invariant and the specification includes county-crop fixed effects, all identifying variation is absorbed. Table 10 reports the results for this test. As expected, the instrument is omitted from these regressions, indicating that it has no residual correlation with any of the predetermined covariates. This supports the orthogonality condition required for instrument validity.

Figure 17 plots county-level averages of predetermined covariates against distance to the nearest USDA office. As expected, distance is correlated with cross-sectional features such as precipitation and historical crop prices: USDA offices tend to be located closer to high-production agricultural counties. However, these cross-sectional correlations are fully absorbed by county–crop fixed effects in the empirical specification. Indeed, the formal covariate-balance test shows that the instrument is omitted once fixed effects are included, indicating no remaining correlation with time-varying predetermined

[6] I am in the process of getting the historical road data, and will update the analysis with that once I receive it.

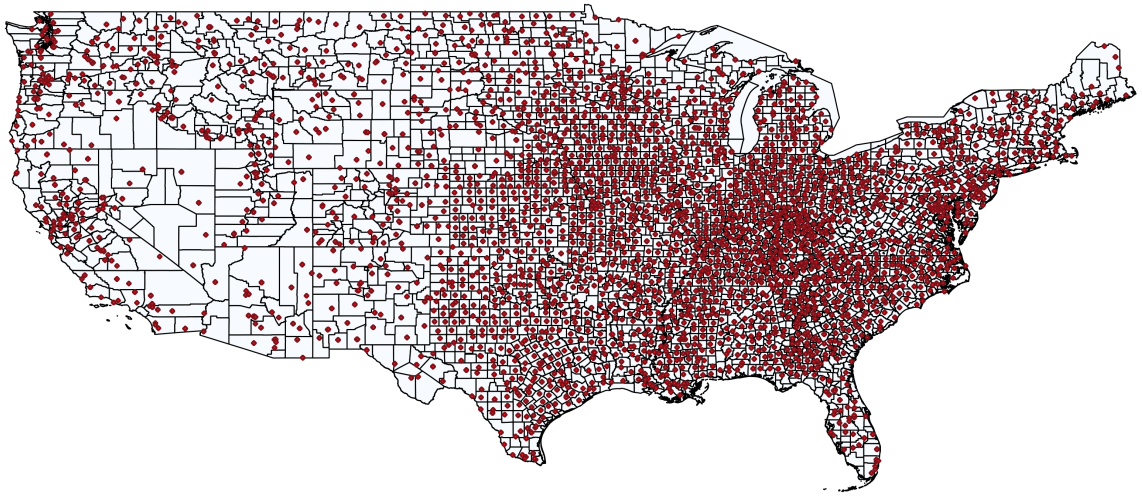
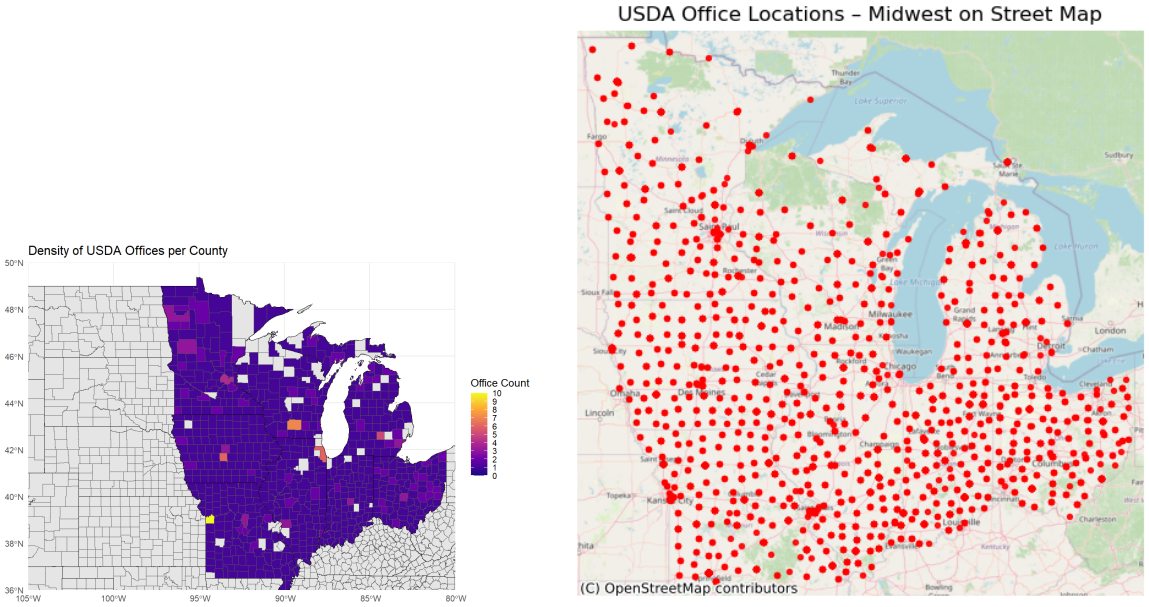


Figure 15: Locations of USDA offices in 2025 (based on 1990s county boundaries and 2025 USDA office locations)



(a) Density of USDA offices

(b) USDA offices and their locations relative to roads

Figure 16: 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).

Table 10: Covariate Balance: Distance to USDA Office and Predetermined Covariates

	Outcome: Predetermined Covariate X		
Covariate X	Coefficient on distance (km)	Std. error	Observations
$ddayMod_100$	0 (omitted)	–	29,990
$ddayHot_100$	0 (omitted)	–	29,990
$prec_cm$	0 (omitted)	–	29,990
$prec2_cm$	0 (omitted)	–	29,990
$price_received$	0 (omitted)	–	29,990
County–crop FE	Yes		
Year FE	Yes		
Cluster level	County–crop		

Notes: Each row reports results from a separate regression of the listed predetermined covariate on the instrument, distance (in km) to the nearest USDA office, controlling for county–crop and year fixed effects. Standard errors are clustered at the county–crop level. Because the instrument is time-invariant and perfectly collinear with the county–crop fixed effects, it is absorbed in all specifications and the coefficient on distance is omitted. This implies that the instrument has no residual explanatory power for the predetermined covariates after conditioning on fixed effects, consistent with covariate balance.

covariates. Therefore, identification relies on cross-sectional variation but is not threatened by omitted time-varying confounders.

The first-stage regression shows that distance to the nearest USDA office is statistically associated with insurance take-up ($\hat{\beta} = -0.091$, $p < 0.01$). However, the first-stage F-statistic is 7.60, below the conventional threshold of 10 for strong instruments (Staiger and Stock 1997). This indicates that although the instrument is relevant, it is relatively weak once state and year fixed effects and weather controls are included. Consequently, IV estimates using this instrument may suffer from weak-instrument concerns, and inference should be interpreted with caution.

Instrumental variable strength and validity: We assess the strength of our instrumental variables by estimating the first-stage regressions for both endogenous variables: subsidy per dollar of liability and the percentage of land insured.

$$ln(\Gamma_{ijt}) = \alpha_0 + \alpha_1 \tau_{ijt} + \alpha_2 EP_{ijt} + \alpha_3 Y_{ijt-1} + \sum_j \alpha_{j4} Time_{jt} + + \alpha_{j5} W_{ijt} + \alpha_{jw6} W_{ijt} \Gamma_{ijt} + \nu_{ij} + \epsilon_{ijt} \quad (3)$$

Fixed effects are at the county-crop level, and robust standard errors are clustered at the county-crop.

Report F-stats

Table 12 shows that across all four specifications, the RMA rating parameter (τ) is both statistically strong and economically meaningful. ??? Give me the F-stats, that is how people assess instrument strength. When the endogenous variable is sublia ??, the coefficient on τ lies between 0.062 and 0.068, highly significant at the 1% level. Similarly, when the dependent variable is fracInsured, the coefficient on τ ranges between 50 and 55, again precisely estimated and economically large. These results indicate that exogenous variation in actuarial rating rules translates directly into variation in subsidy incidence and acreage insured. This satisfies the relevance condition for instrumental variables and mitigates concerns about weak instruments. Moreover, the inclusion of weather interactions, lagged yields, and

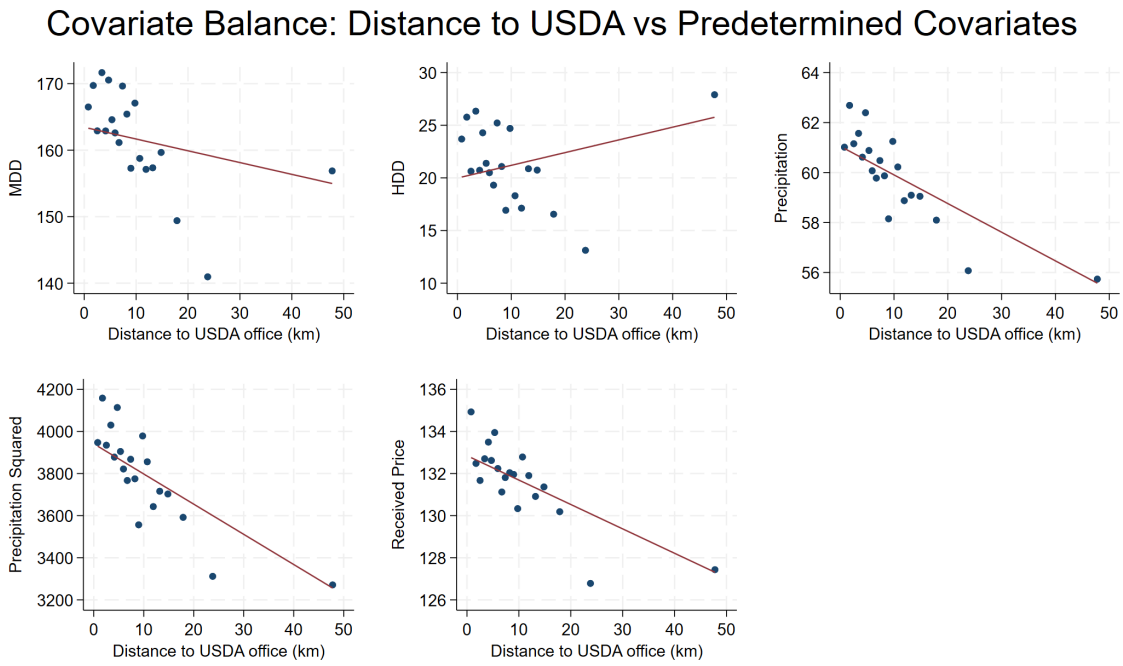


Figure 17: Covariate Balance: Distance to USDA vs Predetermined Covariates using the OpenStreetMap and simple county centroids

Table 11: First-stage Regression: Insurance Take-up on Distance

	(1) fracInsured
drive_km_to_usda	-0.0909** (0.0330)
ddayMod_100	-0.0327 (0.0223)
ddayHot_100	-0.0654*** (0.0166)
prec_cm	-0.0895 (0.0471)
prec2_cm	0.000486 (0.000323)
price_received	-0.0243*** (0.00547)
Constant	76.71*** (4.186)
Observations	30003

Standard errors in parentheses
State and year fixed effects included.
,
Standard errors clustered at the county level.
,
Instrument F-statistic (drive_km_to_usda): 7.60
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: First stage results of the instrument variable for the baseline (no weather variables), and the complete model.

VARIABLES	(1) sublia	(2) sublia	(3) fracInsured	(4) fracInsured
τ	0.0679***	0.0625***	55.06***	50.27***
	(0.00686)	(0.00688)	(5.060)	(5.026)
Last period yield	5.61e-05***	7.67e-05***	-0.00995**	-0.00514
	(4.73e-06)	(5.08e-06)	(0.00434)	(0.00472)
Received price	-3.87e-05***	-4.86e-05***	-0.114***	-0.111***
	(3.80e-06)	(4.01e-06)	(0.00312)	(0.00322)
MDD		-8.51e-05***		-0.0101*
		(6.56e-06)		(0.00570)
HDD		-6.21e-05***		0.0785***
		(7.03e-06)		(0.00538)
Precipitation		0.000186***		0.803***
		(3.51e-05)		(0.0332)
Squared precipitation		-1.37e-06***		-0.00573***
		(2.49e-07)		(0.000238)
Corntime	0.00202***	0.00208***	2.442***	2.406***
	(4.03e-05)	(4.10e-05)	(0.0293)	(0.0300)
Soybeantime	0.00209***	0.00217***	2.488***	2.475***
	(3.33e-05)	(3.38e-05)	(0.0262)	(0.0266)
Constant	0.0117***	0.0199***	40.76***	14.42***
	(0.00105)	(0.00169)	(0.778)	(1.503)
Observations	28,189	28,189	28,189	28,189
R-squared	0.570	0.579	0.611	0.620
Number of panelID	1,291	1,291	1,291	1,291

Notes: Robust standard errors in parentheses (*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.).
Model 1-2 have subsidy per dollar of liability (sublia) as endogenous variables, and models 3-4 have fraction of land insured (fracInsured) as endogenous variable.

crop-specific time trends ensures that identification is not confounded by omitted risk factors. Together, these results provide a credible empirical foundation for using τ as an instrument to estimate the causal effect of crop insurance subsidies on yield responses to weather extremes. You'll need to do a much better job of explaining what the IV does and why it satisfies the exclusion restriction. I am not convinced, and IV papers live (and die) by how good a job authors do of convincing readers that the IV is valid. I stopped reading here..

7- Shortest driving distance between the population-weighted County centroids and USDA offices as an instrument. 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.

Covariance Balance Test for this IV. To assess the plausibility of the exclusion restriction, I examine whether counties located farther from USDA offices differ systematically in predetermined characteristics (Table 13). Regressions of historical climate variables and crop prices on distance to the nearest USDA office, controlling for year fixed effects and clustering at the county level, reveal no statistically significant relationships for temperature exposure or output prices. While precipitation levels are modestly correlated with distance, this reflects well-known spatial climate gradients and is directly controlled for in all specifications. While precipitation is modestly correlated with distance, this pattern reflects smooth spatial climate gradients rather than institutional sorting, and precipitation is flexibly controlled for in all specifications. Consistent with these results, binscatter plots using long-run county–crop averages in Figure 19 reveal no economically meaningful nonlinearities or discontinuities. Together, these findings support the plausibility of the exclusion restriction. Overall, these results suggest that distance to USDA offices is not strongly associated with observable determinants of insurance demand, supporting its use as an instrument. Create an appendix for tables and move the results of balanced tests there.

Table 13: Covariate Balance Test for IV of Driving Distance Between Population-Weighted County Centroids and USDA Offices

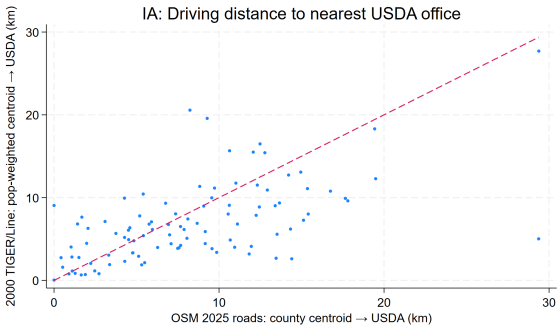
	MDD	HDD	Precipitation	Precipitation squared	Received price
drive_km_to_usda	-0.056	0.137	-0.079***	-10.307***	-0.069
	(0.207)	(0.124)	(0.023)	(2.930)	(0.047)
Observations	29616	29616	29616	29616	29616
R-squared	0.161	0.278	0.267	0.276	0.172

Standard errors in parentheses

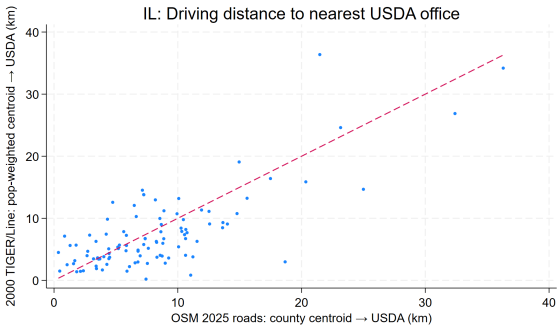
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Instrumental variable strength and validity: I assess the strength of this instrumental variable by estimating the first-stage regressions for both endogenous variables: subsidy per dollar of liability and the percentage of land insured.

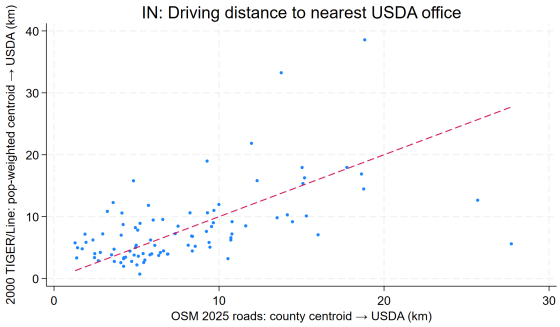
$$Old: ln(\Gamma_{ijt}) = \alpha_0 + \alpha_1 \tau_{ijt} + \alpha_2 EP_{ijt} + \alpha_3 Y_{ijt-1} + \sum_j \alpha_{j4} Time_{jt} + + \alpha_{j5} W_{ijt} + \alpha_{jw6} W_{ijt} \Gamma_{ijt} + \nu_{ij} + \epsilon_{ijt}$$



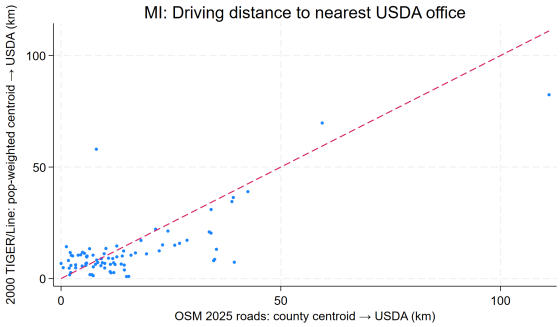
(a) Iowa



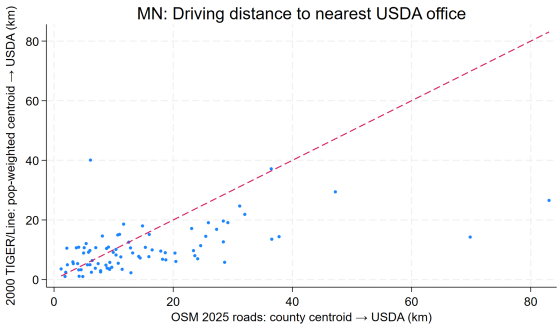
(b) Illinois



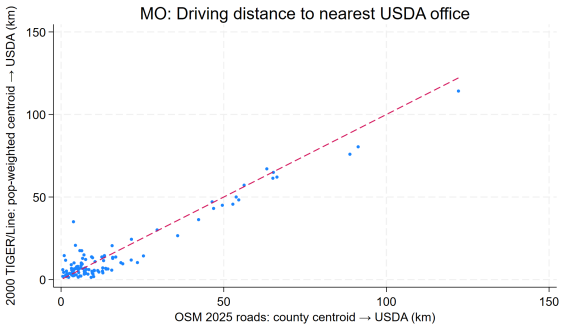
(c) Indiana



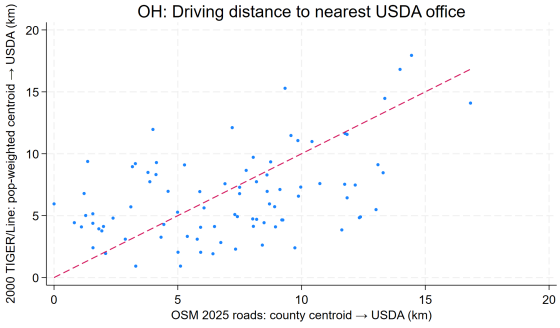
(d) Michigan



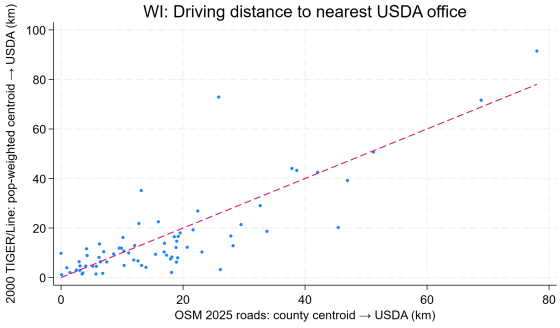
(e) Minnesota



(f) Missouri



(g) Ohio



(h) Wisconsin

Figure 18: 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.

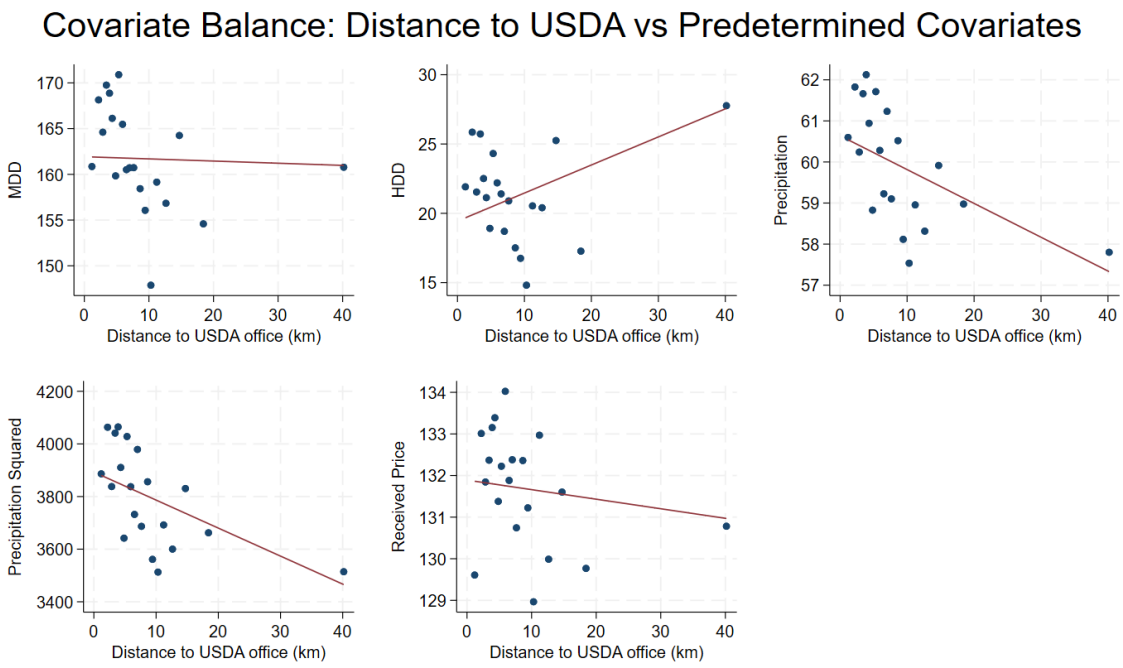


Figure 19: Long-run county–crop averages

$$\text{fracInsured}_{ict} = \pi_1 \,\text{drive_km_to_usda}_i + \mathbf{X}'_{ict}\boldsymbol{\beta}_1 + \gamma_{s(i)} + \delta_t + \varepsilon^{(1)}_{ict}, \tag{4}$$

$$\ln(\text{fracInsured}_{ict}) = \pi_2 \,\text{drive_km_to_usda}_i + \mathbf{X}'_{ict}\boldsymbol{\beta}_2 + \gamma_{s(i)} + \delta_t + \varepsilon^{(2)}_{ict}, \tag{5}$$

$$\text{sublia}_{ict} = \pi_3 \,\text{drive_km_to_usda}_i + \mathbf{X}'_{ict}\boldsymbol{\beta}_3 + \gamma_{s(i)} + \delta_t + \varepsilon^{(3)}_{ict}, \tag{6}$$

$$\ln(\text{sublia}_{ict}) = \pi_4 \,\text{drive_km_to_usda}_i + \mathbf{X}'_{ict}\boldsymbol{\beta}_4 + \gamma_{s(i)} + \delta_t + \varepsilon^{(4)}_{ict}. \tag{7}$$

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

Fixed effects are at the state and year level, in order to control for state-level time-invariant factors like institutions and year shocks like macro conditions, and state of technology. With this choice, my identification focuses on the 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.

Table 14 shows that the distance to USDA offices strongly predicts insurance liability but weakly predicts participation rates, indicating that the instrument primarily shifts the intensive margin of insurance coverage rather than the extensive margin. Accordingly, I focus on insurance liability as the endogenous regressor in the baseline IV specification.

Partialled-out first-stage analysis I next studied the relationship between the instrument and the set of endogenous variables after netting out all controls and fixed effects included in the first-stage regression using the partialled-out first-stage plots. Specifically, both the instrument (distance to the

Table 14: First-stage Regression: Insurance Take-up on Distance

	(1) fracInsured	(2) ln(fracInsured)	(3) sublia	(4) ln(sublia)
drive_km_to_usda	-0.0630	-0.00245*	0.000321***	0.00365*
	(0.0488)	(0.00118)	(0.0000861)	(0.00147)
ddayMod_100	-0.0430	-0.000648	-0.000123	0.00186
	(0.0283)	(0.000716)	(0.0000714)	(0.00121)
ddayHot_100	-0.0689***	-0.00246***	0.000271***	0.00386***
	(0.0175)	(0.000433)	(0.0000275)	(0.000507)
prec_cm	-0.105	-0.00402*	-0.000592***	-0.0103***
	(0.0614)	(0.00171)	(0.000127)	(0.00196)
prec2_cm	0.000578	0.0000281*	0.00000479***	0.0000819***
	(0.000418)	(0.0000124)	(0.000000847)	(0.0000130)
price_received	-0.0241***	-0.000690***	-0.0000186***	-0.000715***
	(0.00258)	(0.0000635)	(0.00000355)	(0.0000683)
Observations	29616	29616	29616	29616
R-squared	0.694	0.628	0.519	0.710
First-stage F	1.66	4.34	13.90	6.17

Standard errors in parentheses
State and year fixed effects included.
,
Standard errors clustered at the county (FIPS) level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

USDA office) and the endogenous variable are residualized with respect to state fixed effects, year fixed effects, and observable covariates, and the plot then relates these residuals to each other. As a result, the slope of the fitted line represents the conditional first-stage effect that identifies the IV, using only within-state, within-year variation. In Figure 20, the slopes are positive and steep for insurance liability (in both levels and logs), indicating a strong conditional association between distance and the intensive margin of coverage. In contrast, the slopes are negative but relatively flat and noisy for insurance participation (in both levels and logs), reflecting a much weaker conditional relationship once fixed effects and controls are removed.

Taken together, the partialled-out first-stage plots show that distance to USDA offices primarily affects the intensive margin of crop insurance coverage rather than the extensive margin. The instrument has substantial explanatory power for insurance liability, as evidenced by both the clear positive slopes and the corresponding strong first-stage F-statistics, while its relationship with participation rates is weak and imprecise. This visual and statistical evidence indicates that the instrument shifts how much insurance farmers carry more than whether they participate at all, supporting the choice of insurance liability as the endogenous regressor in the baseline IV specification and clarifying the economic channel through which distance operates.

Instrumental Variables Diagnosis for Two IVs of driving distance from population-weights county centroids and the number of weather stations Table 15 reports covariate balance tests in which predetermined weather and price variables are regressed on the full set of instruments. While some coefficients are statistically significant, the pattern is not systematic across covariates, and

magnitudes are modest. Given the spatial and historical nature of the instruments, some correlation with long-run agro-climatic conditions is expected. Our identifying assumption is conditional on fixed effects and realized weather controls, which absorb these channels in the main IV specification.

Table 15: Covariance Balance Test for the Instrumental Variables of Driving Distance Between Population-Weighted County Centroids and USDA Offices, and the Number of Weather Stations per County in 1989.

	MDD	HDD	Precipitation	Precipitation squared	Received price
drive_km_to_usda	-0.012	0.157	-0.075***	-9.746***	-0.063
	(0.206)	(0.124)	(0.023)	(2.897)	(0.047)
stations_count1989	-2.095**	-0.750	-0.234	-31.003	-0.438**
	(0.856)	(0.530)	(0.188)	(23.979)	(0.188)
Observations	29002	29002	29002	29002	29002
R-squared	0.171	0.281	0.270	0.279	0.172

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 21 provides a visual assessment of covariate balance by plotting predetermined weather and price variables against the instruments. Both instruments exhibit smooth relationships with long-run climatic conditions, reflecting their spatial and historical nature. Importantly, there is no evidence of sharp discontinuities or nonlinear patterns indicative of endogenous sorting. Combined with fixed effects and controls for realized weather shocks in the main specification, these results support the plausibility of the exclusion restriction.

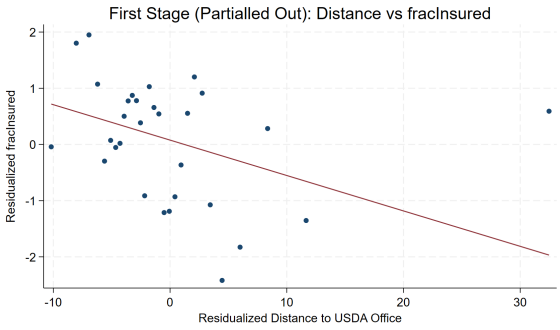
Instrumental variable strength and validity (First-stage of IV regression). I assess the strength of the instrumental variables by estimating first-stage regressions for both endogenous variables: (i) insurance participation, measured by the fraction of planted land that is insured, and (ii) insurance intensity, measured by subsidy per dollar of insured liability. Specifically, I estimate the following first-stage specification:

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

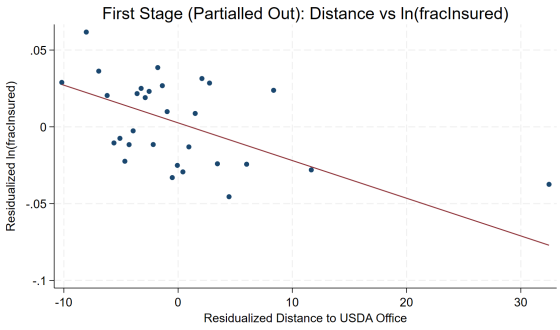
where $D_{ict}^{(k)} \in \{\text{fracInsured}_{ict},\ \ln(\text{fracInsured}_{ict}),\ \text{sublia}_{ict},\ \ln(\text{sublia}_{ict})\}$ denotes the endogenous insurance variable, and $\mathbf{Z}_i = (\text{drive_km_to_usda}_i,\ \text{stations_count1989}_i)$ is the vector of instrumental variables. The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

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

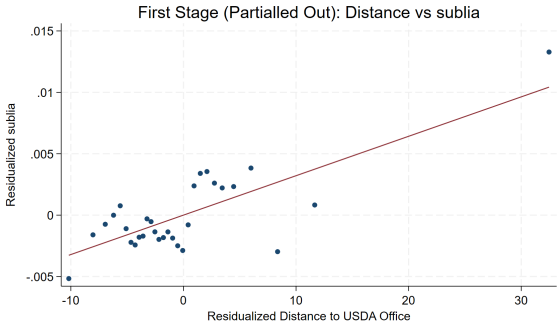
State fixed effects $\gamma_{s(i)}$ absorb time-invariant institutional and geographic differences across states, while year fixed effects δ_t capture common macroeconomic and technological shocks. Identification therefore relies on within-state, cross-county variation over time. Standard errors are 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.



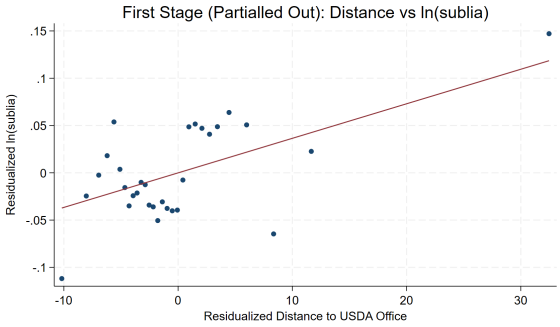
(a) Insurance participation (level)



(b) Insurance participation (log)



(c) Liability (level)



(d) Liability (log)

Figure 20: Partialled-Out First Stage: Distance to USDA Offices Predicts Insurance Liability More Strongly Than Participation

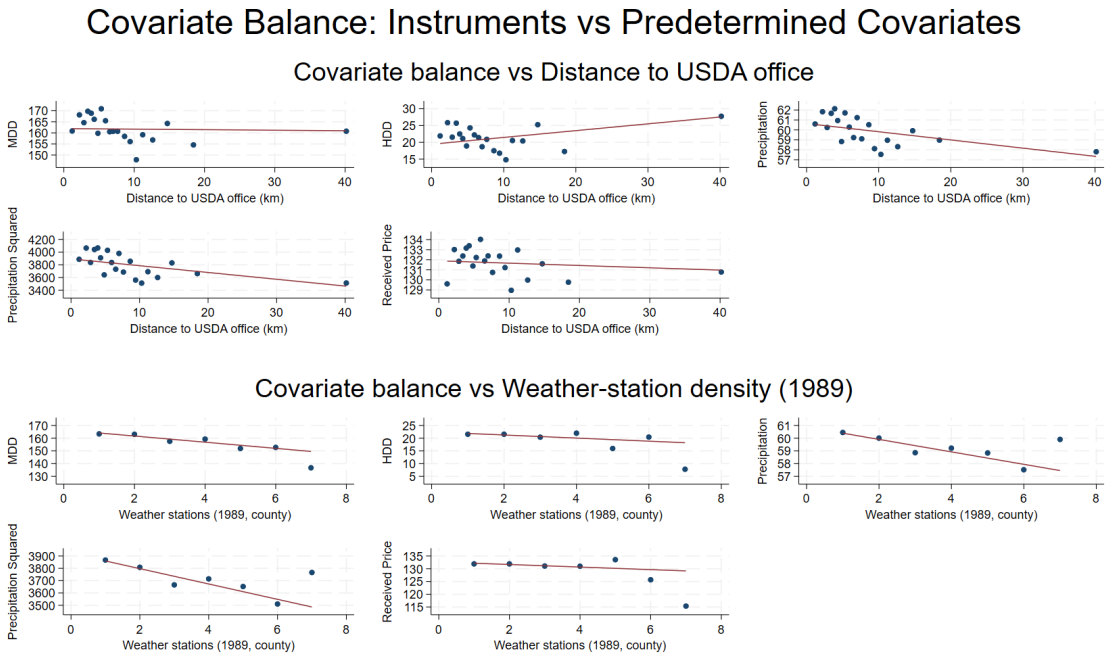


Figure 21: Results of Covariance Balance Tests: Instruments vs Predetermined Covariates

Weak-instrument robustness. Given the moderate first-stage strength for some participation outcomes, all inference is interpreted cautiously, and results are shown to be robust across alternative specifications and insurance intensity measures for which the instruments exhibit stronger relevance.

Table 16 reports the first stage regressions for insurance take-up and liability measures. The instruments exhibit stronger relevance for insurance liability than for fraction of planted land insured, with joint first-stage F-statistics ranging from 3.2 to 7.4 for liability outcomes. In contrast, relevance is weaker for participation outcomes, consistent with the interpretation that distance to USDA offices primarily affects coverage intensity rather than the extensive margin of insurance adoption.

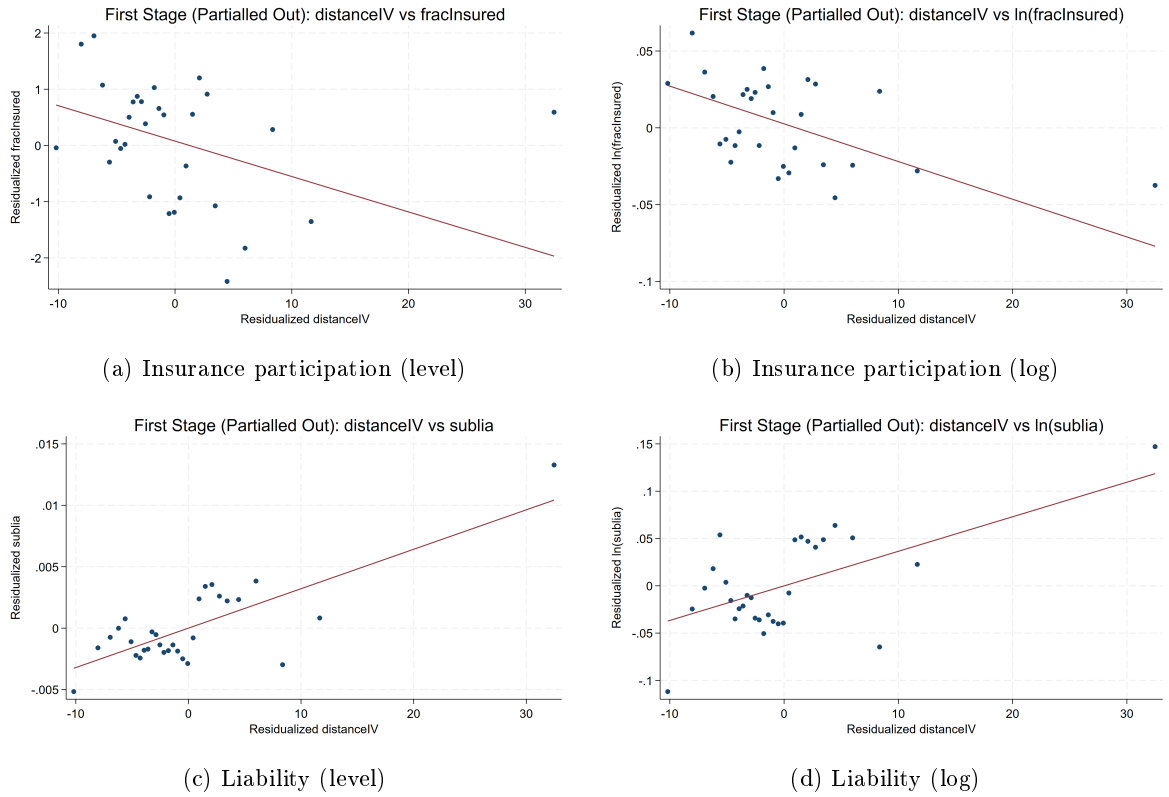
Table 16: First-stage Regressions: Insurance Take-up on Instruments

	(1) fracInsured	(2) ln(fracInsured)	(3) sublia	(4) ln(sublia)
drive_km_to_usda	-0.0565	-0.00237*	0.000313***	0.00353*
	(0.0498)	(0.00119)	(0.0000857)	(0.00145)
stations_count1989	-0.357	-0.00401	0.000549	0.00683
	(0.338)	(0.00746)	(0.000596)	(0.0114)
ddayMod_100	-0.0442	-0.000667	-0.000135	0.00163
	(0.0287)	(0.000728)	(0.0000705)	(0.00121)
ddayHot_100	-0.0702***	-0.00252***	0.000271***	0.00388***
	(0.0177)	(0.000441)	(0.0000279)	(0.000515)
prec_cm	-0.113	-0.00435*	-0.000612***	-0.0105***
	(0.0623)	(0.00174)	(0.000130)	(0.00199)
prec2_cm	0.000625	0.0000307*	0.00000491***	0.0000834***
	(0.000425)	(0.0000126)	(0.000000859)	(0.0000132)
price_received	-0.0244***	-0.000696***	-0.0000185***	-0.000710***
	(0.00261)	(0.0000641)	(0.00000361)	(0.0000693)
Observations	29002	29002	29002	29002
R-squared	0.694	0.628	0.523	0.713
First-stage F (joint IVs)	1.42	2.25	7.36	3.24

Standard errors in parentheses
State and year fixed effects included.
,
Standard errors clustered at the county (FIPS) level.
,
First-stage F is the joint test: drive_km_to_usda = 0 and stations_count1989 = 0.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Partialled-out first-stage analysis. Figures 22 and 23 present partialled-out first-stage relationships between the two instrumental variables—distance to the nearest USDA office and county-level weather-station density in 1989—and insurance outcomes. Across all panels, the fitted lines reflect systematic partial correlations between the instruments and insurance take-up and liability measures after controlling for weather covariates and state and year fixed effects. Distance to USDA offices exhibits a clear negative relationship with insurance participation and a positive relationship with insured liability, consistent with higher administrative or information costs discouraging participation but increasing per-policy liability among remaining participants. Weather-station density displays weaker but directionally consistent relationships, particularly for liability measures, suggesting that greater

availability of local weather information is associated with higher insured exposure. Importantly, these patterns persist after residualizing both the instrument and the endogenous variable, indicating that the observed relationships are not driven by confounding weather conditions or common state-level trends. Together, the figures provide visual evidence supporting the relevance condition for both instruments, while the absence of strong residual correlations with predetermined covariates (shown elsewhere) supports their plausibility as exogenous sources of variation in insurance outcomes.



(a) Insurance participation (level)

(b) Insurance participation (log)

(c) Liability (level)

(d) Liability (log)

Figure 22: First-stage relationships between distance to USDA offices and insurance outcomes. Each panel plots the residualized insurance outcome against the residualized distance to the nearest USDA office after partialling out weather controls and state and year fixed effects. The fitted line depicts the partial first-stage relationship implied by the Frisch–Waugh–Lovell theorem.

Second-stage specification. 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:

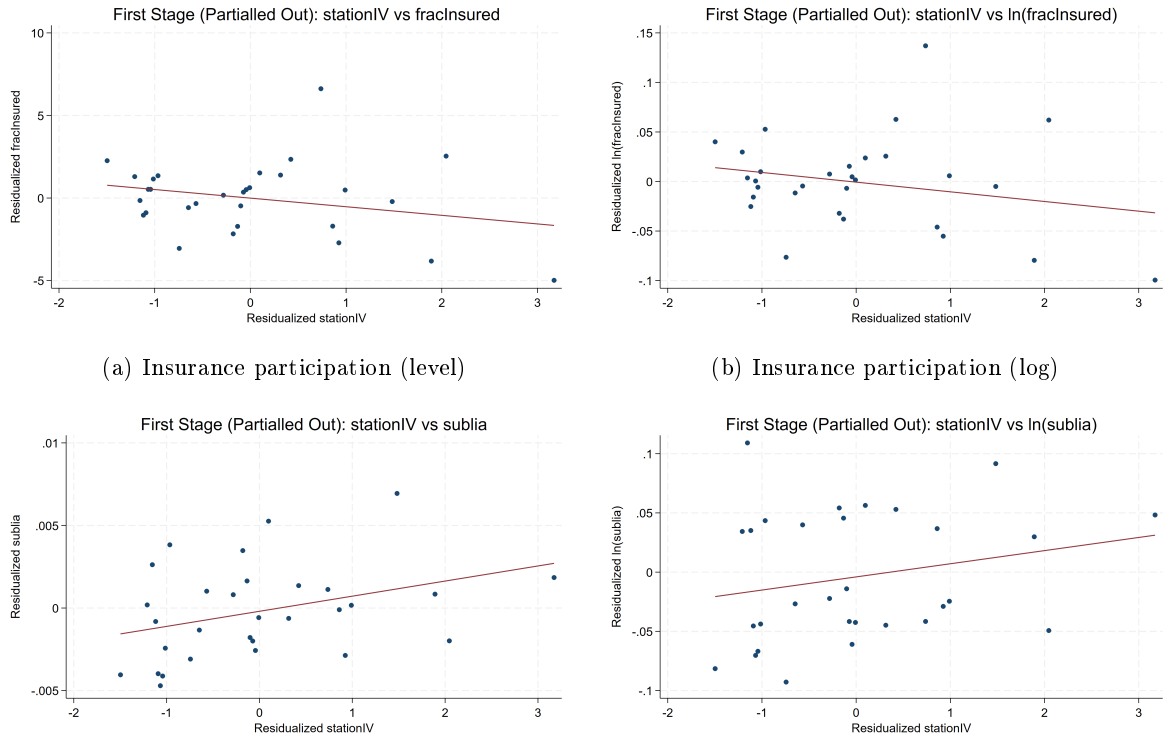
$$Y_{ict} = \beta_1 \widehat{\text{fracInsured}}_{ict} + \mathbf{X}'_{ict}\boldsymbol{\theta}_1 + \gamma_{s(i)} + \delta_t + u^{(1)}_{ict}, \tag{9}$$

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

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

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

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



(a) Insurance participation (level)

(b) Insurance participation (log)

(c) Liability (level)

(d) Liability (log)

Figure 23: First-stage relationships between weather-station density and insurance outcomes. Each panel plots the residualized insurance outcome against the residualized number of weather stations in the county (measured in 1989), after partialling out weather controls and state and year fixed effects. The fitted line represents the partial first-stage relationship implied by the Frisch–Waugh–Lovell theorem.

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:

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

State fixed effects $\gamma_{s(i)}$ absorb time-invariant institutional and geographic differences, while year fixed effects δ_t control for 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.

8- Population size in the past (Ciccone and hall 1996) as an IV Idea from Rossi-Handberg's class.

9- Immigrant population (card 2001) as an IV Idea from Rossi-Handberg's class.

10- Presence of land grant colleges (Moretti 2004b) as an IV Idea from Rossi-Handberg's class.

11- Previous reliability of the weather news on occurrence of hazardous events for farm produce

4.2 Baseline Empirical Methodology

I estimate a sequence of nested specifications to examine how sensitive the relationship between insurance coverage and yields is to alternative control sets. Let Y_{cst} denote either yield or log yield for county c, crop s, and year t, and let X_{cst} denote the insurance variable, measured either as the level or the log of the insured share. I begin with a parsimonious specification that includes only weather controls,

$$W_{cst} = \begin{bmatrix} \text{ddayMod}_{cst} \\ \text{ddayHot}_{cst} \\ \text{prec}_{cst} \\ \text{prec}^2_{cst} \end{bmatrix},$$

and estimate

$$Y_{cst} = \beta X_{cst} + W'_{cst}\gamma + \alpha_c + \lambda_k + \tau_t + \varepsilon_{cst},$$

where α_c, λ_k, and τ_t denote county, crop, and year fixed effects, respectively.

I then augment the model with a richer set of contemporaneous economic controls,

$$Z_{cst} = \begin{bmatrix} \text{price_received}_{cst} \\ \text{ihs_income_netcash}_{cst} \\ \text{ihs_ccc_net}_{cst} \end{bmatrix},$$

so that the second specification becomes

$$Y_{cst} = \beta X_{cst} + W'_{cst}\gamma + Z'_{cst}\eta + \alpha_c + \lambda_k + \tau_t + \varepsilon_{cst}.$$

Next, I allow weather effects to vary with insurance coverage by interacting each weather variable

with the insurance measure:

$$Y_{cst} = \beta X_{cst} + W'_{cst}\gamma + (X_{cst} \cdot W_{cst})' \theta + \alpha_c + \lambda_k + \tau_t + \varepsilon_{cst}.$$

This specification captures the possibility that insurance changes farmers' exposure or response to weather shocks.

Finally, I estimate the fully saturated version that combines the interaction terms with the additional economic controls:

$$Y_{cst} = \beta X_{cst} + W'_{cst}\gamma + (X_{cst} \cdot W_{cst})' \theta + Z'_{cst}\eta + \alpha_c + \lambda_k + \tau_t + \varepsilon_{cst}.$$

Across all specifications, standard errors are clustered at the county level.

5 Baseline causal Effects of Subsidy Shocks

In different models, I gradually add fixed effects for county, crop-specific time trend (or time trend), State-specific time trend, lagged rainfall control (for annual or cropping period).//

I measured the impact of farm bills using four measures, all endogenously set, and hence, need instrumental variable for deriving their causal effects. The four factors are: 1) the effective subsidy rates experienced by farmers; 2) the subsidy amount per $100 Liability, both calculated in Section 3; 3) The liability of farmers, from Jisang Yu, A. Smith, and Sumner (2018); and 4) The percent of planted land that has insurance coverage. I used these measures in all sections of the paper.

Econometric Model Specification To study the sensitivity of crop yield to crop insurance subsidies, we specify changes in yield as a dynamic panel model. In this paper, sensitivity refers to the responsiveness of yields (measured in log terms) to a one percentage point change in the measure of intensity of insurance enrollment, conditional on weather interactions and other controls.

I modeled the panel data using two-way clustering as formulated by Cameron, Gelbach, and D. L. Miller (2011). The panel data allows me to correct standard errors for serial correlation and heteroskedasticity by exploiting both the cross-sectional and time series dimensions. Clustering by unit accounts for serial correlation within units over time, while clustering by time accounts for simultaneous correlation across units. These adjustments are nonparametric and they rely on empirical covariances rather than assuming a specific correlation structure.

In modeling, I first followed the basic model of Deschênes and Greenstone (2007). The dependent variable is the yield Y_ijt, for county i, crop j, and year t. We capture the intensity of enrollment Γ_{ijt} with two parameters of subsidy per dollar of liability and percentage of land insured, and have them in the model once at a time. The weather-related variables are the number of heating degree days (HDD), number of moderate degree days (MDD), amount of precipitation in centimeter, and its squared value, all denoted collectively as W_{ijt}. The control variables are yield in the previous period Y_{ijt-1}, expected price EP_{ijt}, crop-specific time trends $Time_{jt}$, and county-crop fixed effects ν_{ij}. Why not year FEs (instead of crop-year FE)? Also, the presence of lagged Y (why not lagged $ln(Y)$?) on the RHS will cause its own set of problems—this was the starting observation for Nerlove developing random effects back in the day.. As Deschênes and Greenstone (2007) put, it is impossible to estimate the effect of the long-run climate averages in a model with county fixed effects, because there is no

temporal variation in $\overline{W}_{ic}$. Consequently, I replace the climate variables with annual realizations of weather, W_{ict}.

$$ln(Y_{ijt}) = \beta_0 + \beta_1 \Gamma_{ijt} + \beta_2 EP_{ijt} + \beta_3 Y_{ijt-1} + \sum_j \beta_{j4} Time_{jt} + \beta_{j5} W_{ijt} + \beta_{jw6} W_{ijt} \Gamma_{ijt} + \nu_{ij} + u_{ijt} \quad (13)$$

u_{ijt} denotes random errors Just call it an error term with mean zero. "Random" can imply iid, which it is not. Our main coefficients of interest are β_{jw6}, but we have also added Γ_{ijt} and W_{ijt} separately. According to Balli and Sørensen (2013), in a regression with interaction terms, the main terms should always be included Not sure—Pearl and also Angriest and Pischke show that identification by blocking backdoor paths is nonparametrically identified. In other words, if you want E(Y|X,W), you just need to control for X and W—doesn't matter if they're interacted or not, squared, etc. provided both are included somehow.; otherwise, excluding them would cause omitted variable bias.

The county-crop fixed effect captures unobserved heterogeneity across counties for each crop and across crops for each county. The crop-specific time trends capture unobserved heterogeneity across crops for each year. The instrumental variables ?? You haven't discussed them yet and the Γ_{ijt} variable trended upwards ? over time. To prevent detecting a spurious relation Granger and Newbold (1974), we include crop-specific trends in our models to mitigate this potential omitted variable bias.

We also have the expected price of farmers, since we expect farmers' planting decisions (such as input use and effort) to respond to the expected crop prices. The lagged yield controls farm productivity heterogeneity (e.g., soil quality, farmer ability, technology adoption, etc.) that persists over time. It can also proxy the farmer's expectations about future yield. But including the lag of the dependent variable could cause a potential "Nickell bias" (Nickell 1981). Our study covers the span of 25 years and has 1382 unique county-crop panel identifiers. Since the Nickell bias is proportional to the inverse of the number of time periods, we expect the bias to be small, on the order of 4 percent. In addition, our interest is to identify the causal relationship among the crop yield, crop insurance premium subsidy, and a measure of insurance enrollment (subsidy per dollar of liability or percentage of land insured), not the dynamic relationship between the current yield and its lag. The bias in the coefficient of the premium subsidy is less problematic when the covariance of the demeaned premium subsidy variable and the demeaned lagged dependent variable is small.

Endogeneity of Γ A riskier crop or a riskier county has a higher premium rate for a given crop insurance choice (Coble and Barnett 2013). As a result, farmers tend to plant less risky crops and plant less in riskier counties, ceteris paribus. Riskiness is positively correlated with Γ_{ijt} and negatively correlated with Y_{ijt}. Thus, omitting the "riskiness" variable that has a positive correlation with Γ_{ijt}, would cause a downward bias in the OLS estimate for coefficient of Γ_{ijt}. A second source of endogeneity is the dependence of Γ_{ijt} on the choice of coverage level. Although the sign of relationship is not determined. Appendix C has derivations of the sign of both sources of bias. Okay but who cares about the sign of the bias? If it's biased, it's biased.

5.1 Results- Baseline causal Effects of Subsidy Shocks

5.1.1 Subsidy per dollar of liability as the measure of intensity of insurance enrollment

Table 17 presents the results for the case of subsidy per dollar of liability (sublia) as a measure of the intensity of the insurance enrollment. In the OLS specifications (columns 1 and 3), the coefficient on subsidies is positive and significant, suggesting that higher subsidies are associated with higher yields. When weather interactions are introduced (column 3), the results show that the yield effect of subsidies depends on weather conditions. The interaction between subsidies and moderate degree days) is positive and statistically significant, indicating that subsidies amplify yield benefits of moderate heat exposure, which typically reflects favorable growing conditions. By contrast, the interaction with extreme heat is negative, consistent with the idea that subsidies exacerbate farmers' vulnerability to high temperature stress, possibly by discouraging long-term adaptive investments. The interaction with linear precipitation is positive, suggesting that subsidies strengthen the productivity benefits of favorable rainfall. However, the interaction with squared precipitation is negative, pointing to declining marginal returns under excessive rainfall.

The IV estimates (columns 2 and 4), which instrument for subsidies to address endogeneity, tell a more cautious story. The main effect of subsidies turns negative and significant, implying that after accounting for selection bias, subsidies are associated with lower yields on average. This aligns with the concern that subsidies reduce incentives for costly adaptation. Nonetheless, the interaction terms confirm that weather conditions mediate these effects. Subsidies heighten yield benefits under moderate temperatures and sufficient rainfall but increase losses under extreme heat or excessive precipitation variability. This pattern underscores that while subsidies provide short-run insurance against income shocks, they may reduce long-run resilience by altering the incentives for adaptive behavior.

Marginal effects at representative values, MERs (Williams 2012): The post-estimation marginal analysis of model (4) reports the result of one percent point increase in sublia on the yield at different levels of each weather variable, holding other variables as their average values[7]. Figure 24 reports that the effect of sublia on yield is negative across the distribution of hot degree days. While the impact is modest and mostly insignificant at lower quantiles, it becomes larger and statistically significant at higher heat levels, reaching yield losses of about 7–9 percent per additional percentage point of adoption. At the 95th percentile and average of hot degree days, a one percentage point increase in sublia causes 8.6 and 6.8 percent decrease in yield.

Figure 25 reports the marginal effects of sublia on yield at different values of number of moderate degree days. The marginal effect of sublia on yield shifts from negligible at low values of moderate degree days to positive and statistically significant at higher values. At the 90th and 95th percentiles, one percentage point increase in sublia adoption results in a 4.1-4.6 percent increase in yield, and at the 99th percentile the effect rises to about 5.4 percent.

Figure 26 presents the marginal effects of sublia on yields at different levels of precipitation. The results indicate that higher subsidies are consistently associated with lower yields, and the adverse effect intensifies as precipitation increases. At the 1st percentile of precipitation, a one percentage point

[7] Appendix E reports tables of marginal analysis for this and the subsequent measures.

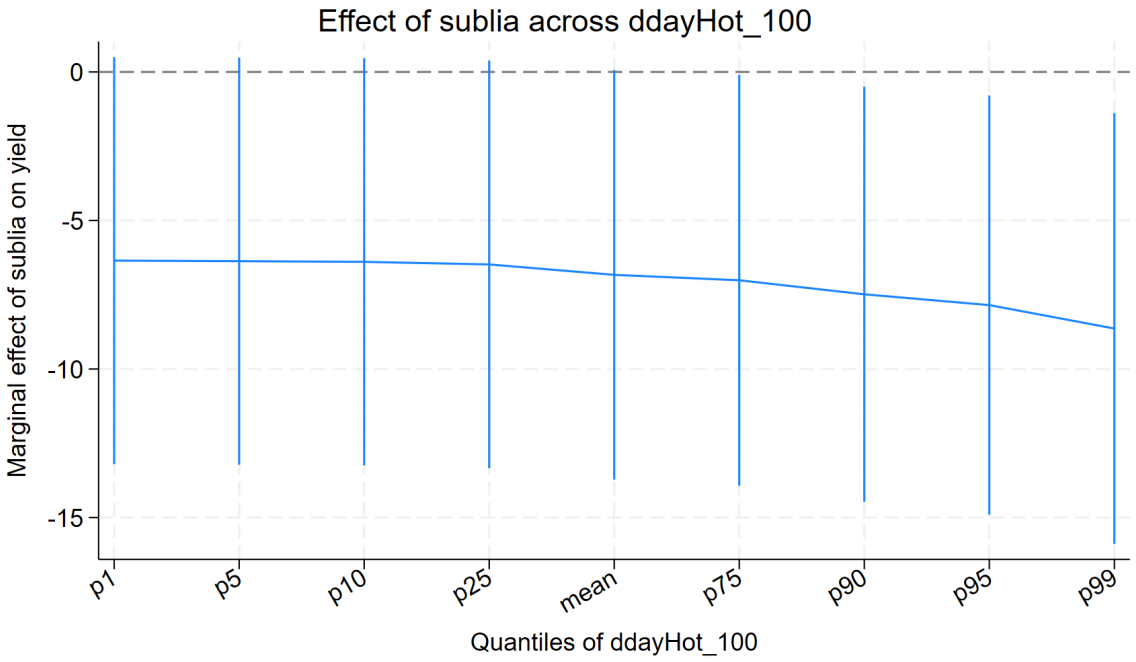


Figure 24: The overall marginal effects of change in sublia across distribution of heating degree days, considering the standalone sublia variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

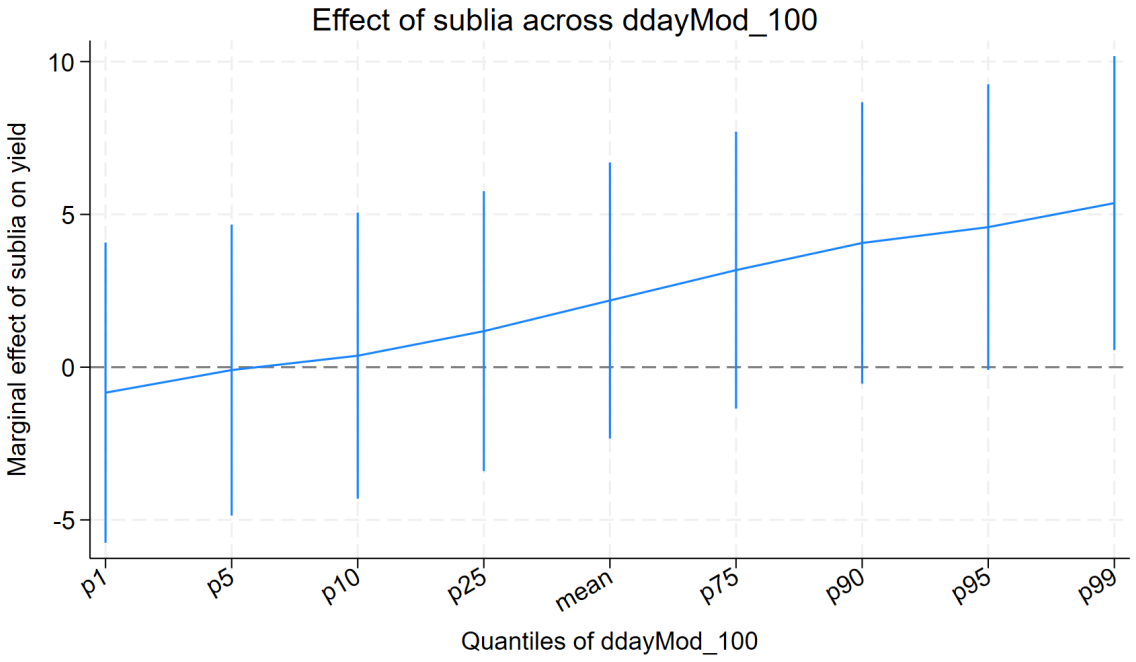


Figure 25: The overall marginal effects of change in sublia across distribution of moderate degree days, considering the standalone sublia variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

Table 17: Estimation of Equation (1) for subsidy per dollar of liability as endogenous variable.

	(1) *log_yield*	(2) *log_yield*	(3) *log_yield*	(4) *log_yield*
Received price	0.00034***	0.00011	0.00026***	0.00015**
	(0.000035)	(0.000072)	(0.000035)	(0.000077)
Last period yield,	0.00031***	0.00052***	0.00032***	0.00050***
	(0.000056)	(0.000082)	(0.000057)	(0.000094)
sublia	0.25**	-3.14***	5.56***	-6.34*
	(0.11)	(0.91)	(0.73)	(3.50)
MDD	0.0055***	0.0053***	0.00558***	0.0031***
	(0.00015)	(0.00018)	(0.00019)	(0.00059)
HDD	-0.0085***	-0.0088***	-0.0088***	-0.0076***
	(0.00016)	(0.00017)	(0.00022)	(0.00055)
Precipitation	0.013***	0.013***	0.019***	0.020***
	(0.00051)	(0.00058)	(0.00092)	(0.00022)
Squared precipitation	-0.00011***	-0.00011***	-0.00016***	-0.00019***
	(4.1e-06)	(4.6e-06)	(6.9e-06)	(0.000016)
Sublia × HDD			0.0068*	-0.023**
			(0.0036)	(0.0099)
Sublia × MDD			-0.011***	0.052***
			(0.0028)	(0.016)
Sublia × precipitation			-0.14***	-0.17***
			(0.018)	(0.049)
Sublia × sq.precipitation			0.0012***	0.0018***
			(0.00013)	(0.00037)
Corntime	0.014***	0.021***	0.014***	0.018***
	(0.00030)	(0.0018)	(0.00028)	(0.0019)
Soybeantime	0.0079***	0.015***	0.0082***	0.012***
	(0.00035)	(0.0020)	(0.00033)	(0.0022)
Constant	2.97***	3.07***	2.76***	3.27***
	(0.027)	(0.039)	(0.043)	(0.14)
Observations	30,500	28,189	30,500	28,189
R-squared	0.422		0.428	
Number of panelID	1,364	1,291	1,364	1,291

Notes: Robust standard errors in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

County–crop fixed effects included; errors clustered at the county–crop level. Model 1: no IV, no weather interaction; model 2, with IV, no weather interaction; model 3, no IV, with weather interaction; and model 4, with IV, with weather interaction.

increase in subsidy reduces yield by about 12.3 percent, whereas at the 99th percentile the reduction reaches nearly 24.9 percent.

Figure 27 presents the marginal effects of sublia on crop yield across quantiles of squared precipitation. The results indicate a nonlinear relationship: at lower precipitation squared levels, the marginal effect of sublia on yield is negative but statistically insignificant, suggesting no robust adverse effect in drier or moderately wet conditions. However, as values increase toward the upper tail of the distribution, the marginal effect becomes positive and statistically significant at the 99th percentile, where a one-percentage-point increase in sublia causes a 14.1 percent increase in yield.

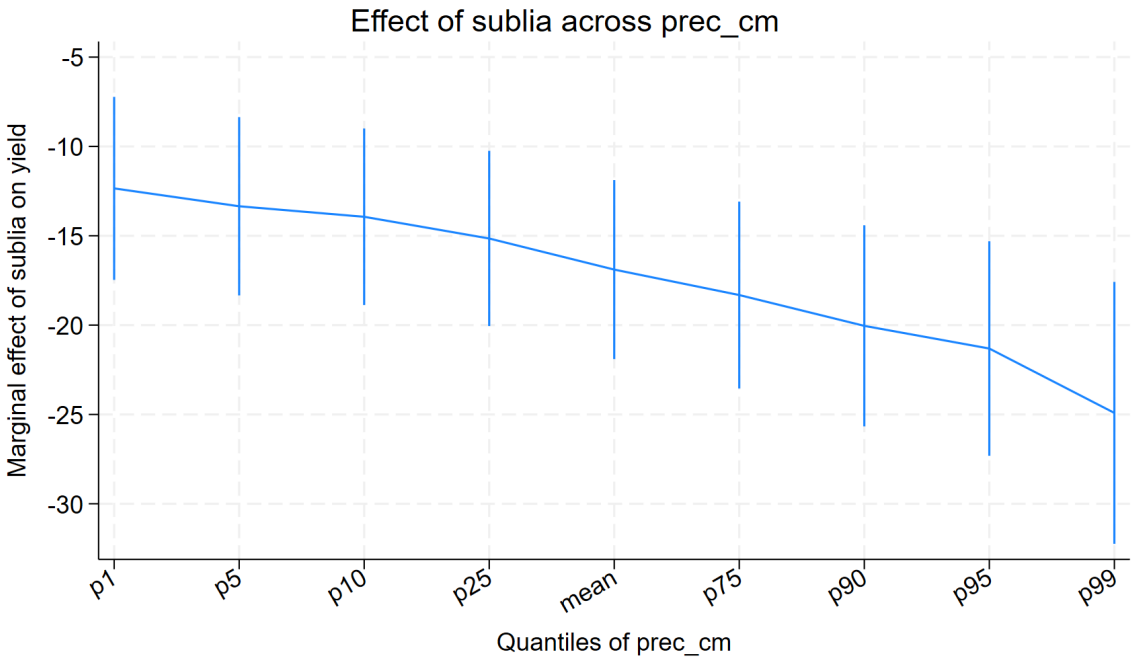


Figure 26: The overall marginal effects of change in sublia across distribution of precipitation, considering the standalone sublia variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

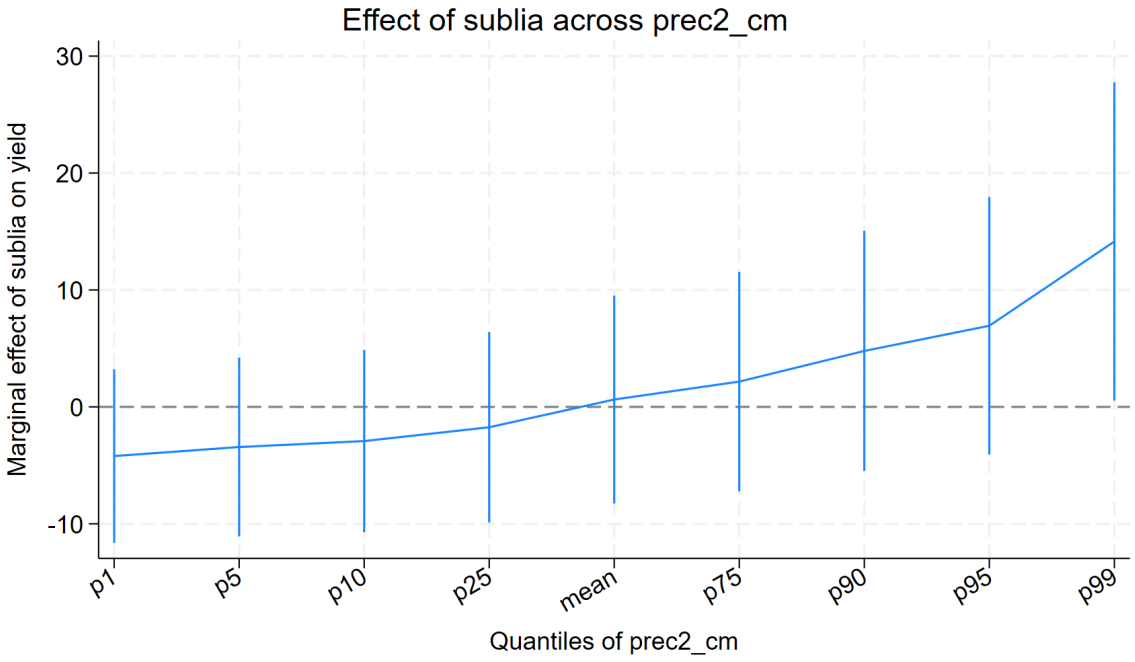


Figure 27: The overall marginal effects of change in sublia across distribution of precipitation squred, considering the standalone sublia variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

5.1.2 Fraction of planted land insured as the measure of intensity of insurance enrollment

Table 18 shows the results for the case of fracInsured (fraction of land insured) as the measure of the intensity of insurance enrollment. This table presents fixed-effects regressions of log yield on fracInsured, weather variables, and their interactions. Columns (5) and (7) report OLS estimates, while columns (6) and (8) use an IV strategy to address endogeneity of insurance take-up. Columns 1 shows essentially no average association between the insured share of acreage and log yields. When we instrument FracInsured with the rating-based parameter (column 2), the coefficient turns negative and statistically significant, indicating that at average weather conditions a 1 pp increase in insured acreage lowers yields on the order of a few tenths of a percent. This pattern is consistent with selection into more marginal land and/or weakened adaptation incentives when a larger share of acreage is covered. Allowing for heterogeneous effects across the weather distribution (columns 3-4) changes the picture. With interactions, the OLS coefficient becomes positive, but the IV main effect loses precision, while the interaction terms are informative: in the IV specification, the effect of insurance becomes more favorable as hot-degree days increase and less favorable as moderate-temperature days increase; interactions with precipitation are small and not robust. Taken together, the results suggest that a higher insured share is not on average productivity-enhancing, but it reallocates risk across states of nature by penalizing yields in benign conditions while partially offsetting losses in hot years. This heterogeneity reconciles the negative average IV effect with the mixed OLS patterns and is consistent with a portfolio/exposure channel: expanding insured acreage encourages cultivation on riskier margins and dampens baseline adaptation, yet coverage cushions outcomes when heat stress is severe.

Marginal effects at representative values: The post-estimation marginal analysis of model (8) reports the outcome of one percent point in fracInsured on yield at different levels of weather variables, holding other variables as their average values. Figures 28 and 29 present the marginal effects of crop insurance coverage, measured as the share of planted land insured (fracInsured), on yields across the distribution of hot degree days and moderate degree days, respectively. The results suggest that the marginal impact of insurance coverage on yield is positive across all quantiles, but statistically insignificant at conventional levels. Figure 28 shows a relatively flat profile, with modest increases in the effect of insurance as heat exposure intensifies, particularly at the upper tail (95th–99th percentiles).

Figure 30 report the marginal effect of fracInsured on yield across the precipitation distribution. The results show that insurance consistently has a positive effect on yield, with the magnitude of the effect increasing as precipitation increases. At the 1st percentile of precipitation, a one percentage point increase in insurance coverage raises yield by about 0.13 percent, and this effect strengthens to about 0.24 percent at the 99th percentile. The figure suggests that the yield-enhancing role of insurance is more pronounced under wetter conditions.

Estimates of the marginal effect of fracInsured at different percentiles of squared precipitation in Figure 31 show that the marginal effect of insurance on yields is slightly positive at lower quantiles of squared precipitation, but not statistically significant. As precipitation squared increases, the effect decreases and turns slightly negative at the upper quantiles (p95–p99), suggesting diminishing or even adverse returns to insurance under very high levels of precipitation variability. The confidence intervals overlap zero across the entire range, indicating no robust evidence of significant effects.

Taken together, these results highlight a dual role of insurance subsidies. While they buffer yields under moderate stress and average rainfall, they may simultaneously exacerbate sensitivity to extreme

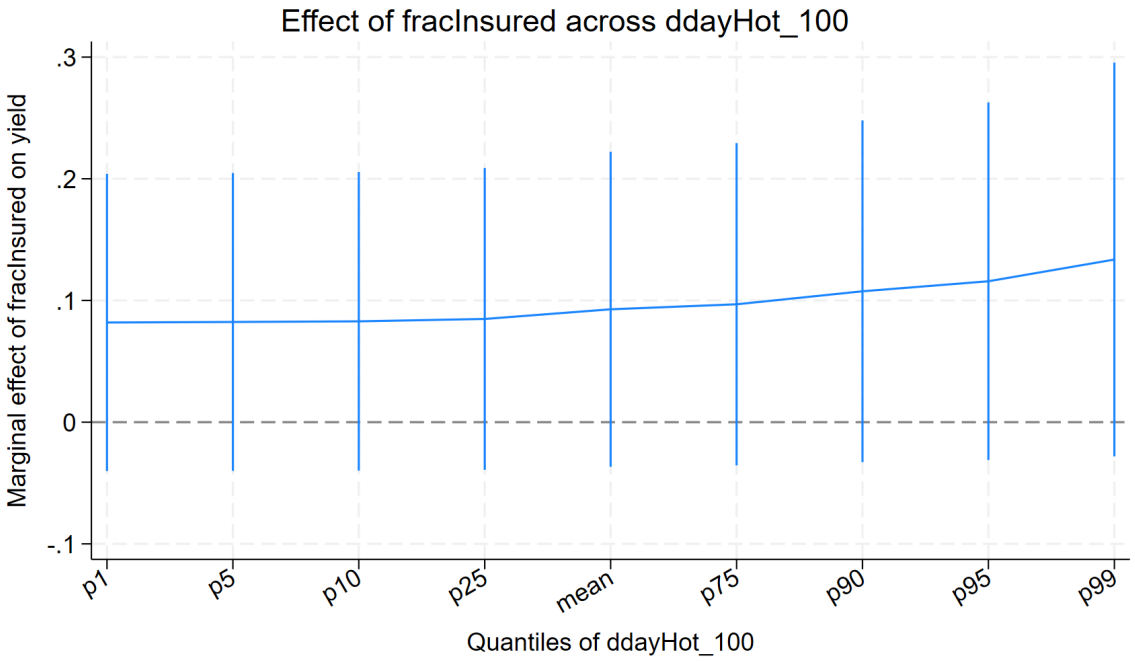


Figure 28: The overall marginal effects of change in fracInsured across distribution of number of heating degree days, considering the standalone fracInsured variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

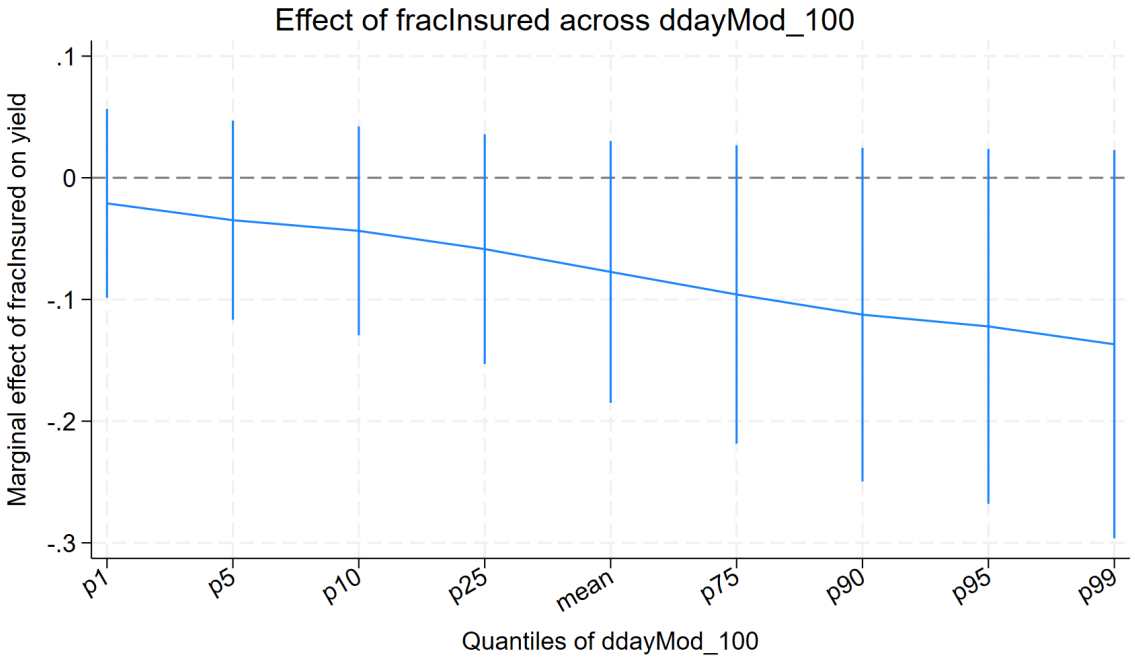


Figure 29: The overall marginal effects of change in fracInsured across distribution of number of moderate degree days, considering the standalone fracInsured variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

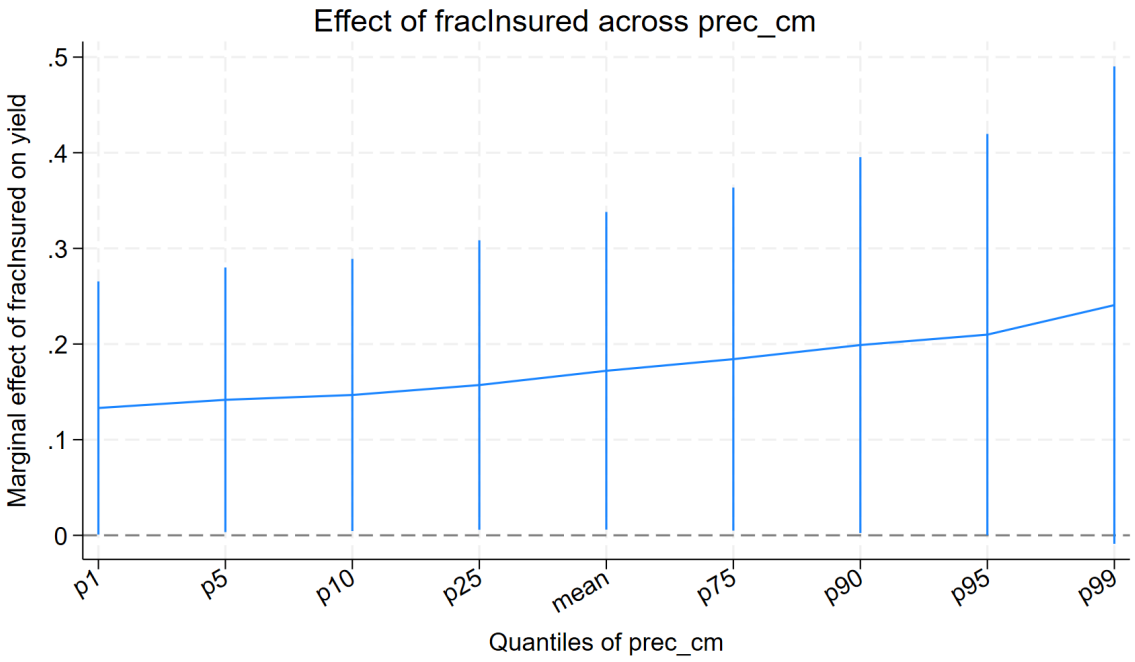


Figure 30: The overall marginal effects of change in fracInsured across distribution of amount of precipitation, considering the standalone fracInsured variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

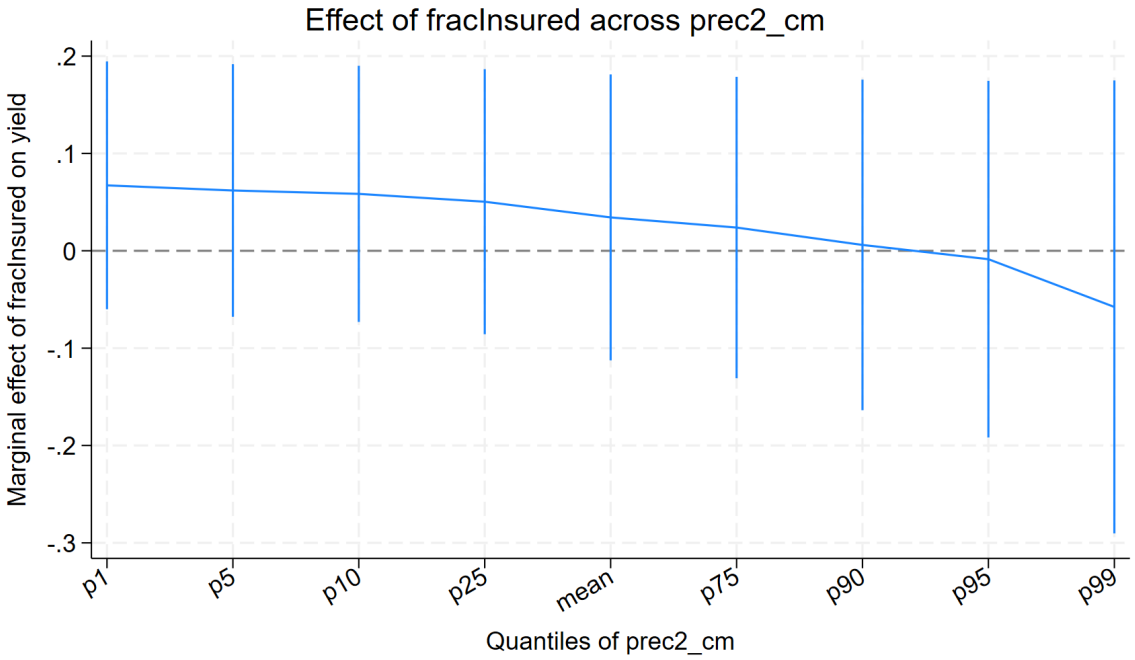


Figure 31: The overall marginal effects of change in fracInsured across distribution precipitation squared, considering the standalone fracInsured variable as well as its interaction with weather variable in the regression. All other variables are at their mean values.

Table 18: Estimation of Equation (1) for fracInsured as endogenous variable.

	(5) *log_yield*	(6) *log_yield*	(7) *log_yield*	(8) *log_yield*
Last period yield,	0.00032***	0.00026***	0.00033***	-0.00073
	(0.000057)	(0.000057)	(0.00057)	(0.00069)
FracInsured	-0.000075	-0.0039***	0.0061***	0.082
	(0.000071)	(0.0011)	(0.00073)	(0.062)
MDD	0.0054***	0.0055***	0.0058***	0.060**
	(0.00014)	(0.00016)	(0.00025)	(0.029)
HDD	-0.0086***	-0.0083***	-0.0089***	-0.037**
	(0.00016)	(0.00021)	(0.00036)	(0.015)
Precipitation	0.013***	0.016***	0.023***	-0.043
	(0.00052)	(0.0011)	(0.0011)	(0.067)
Squared precipitation	-0.00011***	-0.00013***	-0.00017***	0.00041
	(4.1e-06)	(8.0e-06)	(8.0e-06)	(0.00044)
FracInsured × HDD			2.8e-06	0.00053*
			(4.8e-06)	(0.00029)
FracInsured × MDD			-2.6e-06	-0.00098*
			(3.1e-06)	(0.00054)
FracInsured × precipitation			-0.00016***	0.0015
			(0.000016)	(0.0012)
FracInsured × sq.precipitation			9.8e-07***	-0.000012
			(1.2e-07)	(7.8e-06)
Received price	0.00030***	-0.00017	0.00032***	-0.0040
	(0.000035)	(0.00015)	(0.000036)	(0.0026)
Corntime	0.015***	0.024***	0.014***	0.067**
	(0.00027)	(0.0027)	(0.00027)	(0.031)
Soybeantime	0.0087***	0.018***	0.0086***	0.079**
	(0.00028)	(0.0028)	(0.00027)	(0.039)
Constant	2.98***	3.07***	2.57***	-2.63
	(0.026)	(0.037)	(0.054)	(3.79)
Observations	30,500	28,189	30,500	28,189
R-squared	0.421		0.426	
Number of panelID	1,364	1,291	1,364	1,291

Notes: Robust standard errors in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

County–crop fixed effects included; errors clustered at the county–crop level.

weather, consistent with the hypothesis that subsidized insurance reduces incentives for long-term resilience investments.

5.1.3 Performance of the IV in Estimations

To check the relevance of the instrument for model (2), we conducted the Kleibergen-Paap rk LM statistic of under-identification, which reports the value of 69.457, with Chi-sq(1) P-val = 0.0000. This highly significant result rejects the null hypothesis of under-identification. Second, we checked the strength of the instrument and calculated the Kleibergen-Paap rk Wald F statistic, which results in values 413.118 and 82.474, respectively. Considering the simulated thresholds of maximum tolerated size distortion for F-statistic developed by Stock and Yogo (2002), our F statistics are well above the

10% Stock-Yogo threshold of 16.38. This shows that our instrument is strong, even in the strictest category of Stock and Stock and Yogo (2002). In model (4), the under-identification test (Kleibergen-Paap rk LM statistic) reports value of 72.737 with Chi-sq(1) P-val = 0.0000. Again, this statistic rejects the null hypothesis of under-identification when including the weather variable interactions in the model. The Kleibergen-Paap rk Wald F statistic for the strength of instrument takes value of 16.347 is almost the 10% Stock-Yoko threshold of 16.38. The instrument is adequate but only just strong enough under the strictest criterion, leaving them robust to weak-IV inference.

In model (6), the under-identification test (Kleibergen-Paap rk LM statistic) reports value of 94.095 with Chi-sq(1) P-val = 0.0000, rejecting the null hypothesis of irrelevance of instrument. The Kleibergen-Paap rk Wald F statistic for weak instrument had value od 100.058, well above the the 10% Stock-Yogo threshold of 16.38. This proves the strength of the instrument. In model (8), the under-identification test (Kleibergen-Paap rk LM statistic) reports 3.897 with Chi-sq(1) P-val = 0.0484. This low static shows weaker results, but it still rejects the null hypothesis of irrelevance at the 5% level. But the Kleibergen-Paap rk Wald F statistic for weak instrument has value of 0.801, which is far below of even the 25% threshold (loosest threshold) of Stock-Yoko that is 5.53. This means that the instrument is weak for the specification of model (8): fraction land insured as the endogenous variable, and interaction of fraction insured with the weather variables included in the regression.

5.2 Sensitivity Analysis

Table 19 reports the IV–GMM estimates for corn when using sublia as a measure of intensity of insurance enrollment (columns 9 and 10), and fracInsured as a measure of intensity of enrollment (columns 11 and 12). Table 19 demonstrates a robust negative association between intensity of insurance enrollment and yield outcomes, consistent with the hypothesis that subsidies may reduce incentives for costly adaptive investments. In the baseline specification without interactions (columns 9 and 11), the coefficients on sublia and fracInsured are negative and statistically significant, with a strong first stage (Kleibergen–Paap F = 63.402), giving confidence in the validity of the estimates. When interactions with weather variables are introduced (columns 10 and 12), the results show that subsidies amplify losses under extreme heat (negative $subliaHDD$ and $fracInsuredHDD$) while providing modest resilience under moderate heat (positive $subliaMDD$ and $fracInsuredMDD$). The precipitation terms reveal a convex relationship, where subsidies are associated with lower yields at moderate rainfall levels (negative $subliaprecipitation$) but this penalty diminishes at higher precipitation (positive $subliaprec^2$). All coefficients in column 12 are insignificant.

Table 20 shows the IV_GMM results for soybeans. In this case, the evidence is weaker and less consistent, reflecting both crop-specific agronomic differences and weaker first-stage identification. In the baseline IV specification, enrollment intensities again display a negative association with yield in models 13-15, but the instrument strength is marginal in models 13, 14 and 16 (Kleibergen–Paap F = 6.409 in model 13, 3.480 in model 15, 11.717 in model 15 and 0.526 in model 16), making these estimates a case for a weak IV. When interactions are added, the coefficients on subsidy-weather terms are small and statistically insignificant, suggesting that insured soybean yields do not exhibit the same heightened sensitivity to heat and precipitation as observed for corn. Instead, the positive direct effect of insured acreage (fracInsured) in model 16 may reflect management choices or residual selection into

Table 19: Results for crop of corn, for models with and without weather interaction.

	(9) *log_yield*	(10) *log_yield*	(11) *log_yield*	(12) *log_yield*
Sublia	-3.02***	-1.97		
	(1.13)	(4.28)		
Sublia × HDD		-0.044***		
		(0.016)		
Sublia × sq.precipitation		0.0024***		
		(0.00044)		
Sublia × MDD		0.053*		
		(0.028)		
Sublia × precipitation		-0.27***		
		(0.057)		
Last period yield,	0.00047***	0.00041***	0.00022***	0.0012
	(0.000080)	(0.000087)	(0.000067)	(0.00099)
MDD	0.0050***	0.0024**	0.0053***	-0.017
	(0.00030)	(0.0010)	(0.00024)	(0.033)
HDD	-0.0095***	-0.0069***	-0.0091***	-0.0032
	(0.00022)	(0.00080)	(0.00029)	(0.021)
Precipitation	0.017***	0.029***	0.019***	0.087
	(0.00085)	(0.0028)	(0.0013)	(0.11)
Sq. precipitation	-0.00015***	-0.00025***	-0.00016***	-0.00041
	(6.6e-06)	(2.1e-05)	(9.1e-06)	(0.00061)
Received price	-0.0011***	-0.00088**	-0.0010***	0.0037
	(0.00026)	(0.00035)	(0.00022)	(0.0055)
Corntime	0.023***	0.019***	0.024***	-0.016
	(0.0025)	(0.0032)	(0.0030)	(0.049)
FracInsured			-0.0034***	-0.0015
			(0.0012)	(0.019)
FracInsured × HDD				-0.00018
				(0.00040)
FracInsured × sq.precipitation				4.6e-06
				(0.000011)
FracInsured × MDD				0.00046
				(0.00063)
FracInsured × precipitation				-0.0012
				(0.0020)
Constant	3.71***	3.74***	3.67***	2.93**
	(0.065)	(0.18)	(0.049)	(1.33)
Observations	14,281	14,281	14,281	14,281
Number of panelID	653	653	653	653

Notes: Robust standard errors in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

County–crop fixed effects included; errors clustered at the county–crop level.

coverage rather than a robust weather interaction channel. Overall, while the soybean regressions preserve the expected signs on weather controls (moderate heat positive, extreme heat negative, precipitation positive, precipitation squared negative) the insurance interactions provide little systematic evidence. These results suggest that the subsidy-weather sensitivity mechanism is more pronounced in corn production, while soybean responses remain comparatively muted.

Table 20: Yield regressions (dependent variable: log(yield))

	(13) *log_yield*	(14) *log_yield*	(15) *log_yield*	(16) *log_yield*
Sublia	-15.9**	-67.8		
	(6.85)	(43.9)		
Sublia × HDD		-0.022		
		(0.028)		
Sublia × sq.precipitation		-0.00077		
		(0.0019)		
Sublia × MDD		0.19*		
		(0.11)		
Sublia × precipitation		0.24		
		(0.30)		
Last period yield,	0.00049	0.00027	0.0018***	-0.00080
	(0.00031)	(0.00049)	(0.00059)	(0.00081)
MDD	0.0043***	-0.0034	0.0051***	0.0080
	(0.00054)	(0.0048)	(0.00026)	(0.019)
HDD	-0.0085***	-0.0095**	-0.0061***	0.048
	(0.00061)	(0.0037)	(0.00047)	(0.054)
Precipitation	0.012***	0.0032	0.021***	0.13
	(0.0013)	(0.012)	(0.0043)	(0.10)
Squared precipitation	-0.000096***	-0.000057	-0.00016***	-0.00054
	(9.0e-06)	(0.000069)	(0.000032)	(0.00070)
Received price	-0.00012	-0.00029	-0.0011*	0.0033*
	(0.00026)	(0.00043)	(0.00059)	(0.0017)
Soybeantime	0.043***	0.063*	0.041***	-0.036*
	(0.015)	(0.034)	(0.012)	(0.021)
FracInsured			-0.014***	0.12**
			(0.0050)	(0.060)
FracInsured × HDD				-0.00077
				(0.00080)
FracInsured × sq.precipitation				5.8e-06
				(0.000012)
FracInsured × MDD				-0.00080
				(0.00031)
FracInsured × precipitation				-0.0017
				(0.0017)
Constant	2.85***	4.82***	2.72***	-5.13
	(0.20)	(1.52)	(0.12)	(3.60)
Observations	13,908	13,908	13,908	13,908
Number of panelID	638	638	638	638

Notes: Robust standard errors in parentheses.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

County–crop fixed effects included; errors clustered at the county–crop level.

Comparing the two crops highlights an important asymmetry in the interaction between crop insurance subsidies and weather risk. Corn yields show a clear and robust pattern whereby subsidies increase vulnerability to extreme heat and excessive rainfall but provide some protection under moderate heat. By contrast, soybean yields display weaker and less systematic responses, with interaction terms generally imprecise and instrument strength considerably lower. This divergence likely reflects both agronomic differences and differences in insurance participation patterns across crops, as corn is

more heat- and water-sensitive than soybeans (Licht and Archontoulis 2017). The results therefore suggest that the disincentive-to-adapt channel documented in the broader literature is particularly relevant for corn, while the effects for soybeans are more modest and less conclusive.

5.3 Counterfactual Cost of Subsidized Insurance for Counties Exposed to Extreme Heat

In this section, we calculate the financial burden of excessive sensitivity of crop yield for counties located in hot areas. To do so, we set the crop yield of counties exposed to extreme heat (Y^h) to the average crop yield of counties exposed to 0–50% extreme heat (Y^l), and compute the crop-specific cost of the shortfall for counties located in hot regions in each year:

$$C^F_{jt} = \sum_i (Y^l_{jit} - Y^h_{jit}) \times A^h_{jit} \times EP_{jit} \tag{14}$$

Here, A^h_{jit} is the area planted in counties exposed to extreme heat, and summation is over counties.

6 Peer Effects of Enrollment in Federal Crop Insurance Program

In this section, I focus on the peer effects of enrollment in FCIP. In situations that the amount of information needed to make a decision is high, decision makers do not have experience or guidance on how to proceed with financial decisions, it is likely to emulate others. This influence of others on an individual's decision making is called *peer effect* (Manski 1993). Presence of social learning and knowledge spillover among farmers is well studied when it comes to production technology (Foster and Rosenzweig 1995; Munshi 2004; L. Zhou et al. 2020; Adegbola and Gardebroek 2007). However, little is known as to whether, and to what extent, enrollment in agricultural insurance can diffuse and increase the enrollment of other farmers. Studies have frrequently assumed that farmers' decision to purchase crop insurance are independent of that of other farmers. In this case, there will be bias in the ordinary least square model (OLS) because of unobserved endogeneity.

Using spatial econometrics for modelling the peer effects of farmers has been previously used in studies on pesticide use (L. Zhou et al. 2020; Adegbola and Gardebroek 2007).

To study the impact of crop insurance enrollment on agricultural yields, spatial modeling is essential because the underlying economic processes exhibit substantive spatial dependence. Farmers operate in spatially connected environments where decisions, risks, and outcomes spill across boundaries: adoption of insurance can influence neighboring farmers through peer effects, information exchange, or shared perceptions of risk. Moreover, yields are the equilibrium outcome of spatial interaction processes, including pest pressure, weather patterns, soil characteristics, and local management norms; factors that are documented in Meena et al. (2024) read Jan 21, 2026. Ignoring these channels leads to biased estimates because spatial correlation in both the treatment (insurance uptake) and the outcome (yield) violates the independence assumptions of standard regression approaches. Additionally, some spatial dependence arises from non-behavioral sources, such as data aggregation across counties or regions that share ecological and climatic gradients. For these reasons, spatial regression models (like the spatial lag and spatial autoregressive models) are necessary to correctly identify the impact

of insurance enrollment on yield while accounting for the spatial structure of decisions and outcomes [8].

Ignoring spatial lags in the relationship between crop insurance enrollment and agricultural yields effectively means ignoring the substantive spatial interactions that shape farm outcomes. Farmers' yields and insurance decisions are spatially correlated because they face shared weather shocks, soil characteristics, pest pressures, and information networks. If a model fails to incorporate this spatial dependence, it creates an omitted variable problem: the influence of neighboring counties' insurance uptake or yields is left out of the specification even though it systematically affects outcomes. As a result, ordinary least squares estimates become biased and inconsistent, since the spatially lagged variables are correlated with the error term. This can lead to incorrect conclusions—incorrect coefficient magnitudes, wrong signs, misleading standard errors, spurious statistical significance, and poor model fit. For policy questions, such as assessing whether insurance promotes or distorts yield outcomes, these errors can substantially misguide interpretation and policy design.

In the crop insurance–yield relationship, ignoring spatial error correlation does not bias the OLS coefficient on insurance enrollment, but it does make the estimator inefficient. Inefficiency means that although OLS still delivers the correct average effect, it fails to use all the information in the data because the residuals contain spatial structure that the model ignores. As a result, the estimated standard errors are too small or too large, confidence intervals are misleading, and statistical significance tests become unreliable. In practice, this means we we think insurance has a precisely estimated effect on yield when in fact the true uncertainty is much larger due to spatially correlated shocks (such as common weather events or pest outbreaks) that affect clusters of neighboring counties. Thus, inefficiency undermines inference even when the point estimate itself is correct.

Spatial Cross-Regressive (SLX) model. Following the Spatial Cross–Regressive (SLX) model of Florax and Folmer (1992) and its modern treatment in Halleck Vega and Elhorst (2015), I model spatial spillovers by allowing only the covariates, but not the dependent variable, to be spatially lagged. Specifically, I estimate

$$y_{ict} = X_{ict}\beta + WX_{ict}\theta + \alpha_i + \delta_t + \varepsilon_{it}. \tag{15}$$

where y_{ict} denotes crop yield in county i, crop c and year t, X_{ict} is a vector of county-level covariates including insurance adoption, weather conditions, and prices, and W is a row-standardized spatial weights matrix. The term WX_{ict} captures spatial spillovers arising from neighboring counties' covariates. States fixed effects α_i absorb time-invariant heterogeneity across states, while year fixed effects δ_t control for common aggregate shocks. Because the instrumental variables are time-invariant, specifications with county or county-by-crop fixed effects would absorb the excluded instruments and are therefore not identified. I therefore use state and year fixed effects and cluster standard errors at the county level in both OLS and IV specifications all along.

This specification allows for flexible local spillovers in observable determinants of yields while avoiding the strong equilibrium assumptions and potential endogeneity associated with spatial lag or

[8] I appreciate Zhuyin Li and Pedro Amal at the Center for Spatial data Science of UChicago for our discussions. I thank Professor Rossi-Hansberg for letting me sit in his class on Spatial Economics

spatial error models. As emphasized by Halleck Vega and Elhorst (2015), the SLX model provides a transparent and robust framework for identifying spatial interactions when spillovers operate through covariates rather than outcomes.

Spatial Durbin model. For the robustness check, I also modeled the spatial propagation using the spatial Durbin model. The Spatial Durbin Model allows for endogenous spatial dependence through the spatial lag of the dependent variable while simultaneously modeling spillovers in covariates (LeSage and Pace 2009). Let the spatially lagged dependent variable be Wy, and spatially lagged explanatory variables be WX. Then I have:

$$y_{ict} = \rho W y_{ict} + X_{ict}\beta + WX_{ict}\theta + +\alpha_i + \delta_t + \nu_{it}. \tag{16}$$

where ρ captures endogenous spatial dependence, and all other terms are similar to the SLX model.

Weight Matrix I calculate time-invariant "internal-only" mid west weights, meaning that I only used counties in the study area of 8 states for the calculation. With this choice, I modeled the spillovers only within the study region. I modeled the spatial dependence using a contiguity-based spatial weights matrix. Let W denote an $N \times N$ spatial weights matrix constructed at the county level using "Queen contiguity", where $w_{ij} = 1$ if counties i and j share a common boundary or vertex, and $w_{ij} = 0$ otherwise. The matrix is row-standardized such that each row sums to one,

$$\tilde{w}_{ij} = \frac{w_{ij}}{\sum_j w_{ij}}, \tag{17}$$

so that spatial lags can be interpreted as weighted averages of neighboring outcomes. Because each county borders only a small number of neighbors, the resulting weights matrix is sparse, with the vast majority of entries equal to zero. This sparsity substantially reduces computational burden and is standard in spatial econometric applications using large geographic units such as U.S. counties.

As an alternative specification, I also modeled the spatial dependence using a "Rook contiguity" spatial weights matrix. If counties i and j share a common boundary (edge), the corresponding weight takes value of 1 and 0 otherwise. Compared to Queen contiguity, this definition imposes a stricter notion of neighborhood by excluding corner-touching units. I row-standardized this matrix as well, using the same method for Queen weight matrix.

Spatial Panel Data Setup I first create separate data sets for each crop, then balance each at the county-year level. Then I set it a panel at the county-crop-year level and ensure that geographic identifiers are consistent across the outcome/covariate file and the spatial weights. The spatial weights matrix W is time-invariant at the county level, row-standardized, and applied year-by-year to compute spatial lags of the regressors, WX_{ict}, by multiplying W with the county-level covariates within each year and crop, using the exact county ordering implied by W. Finally, I merge these spatially lagged covariates back to the long panel and estimate the SLX specification with state and year fixed effects and standard errors clustered at the county-crop level.

Spatial Regression Results I estimate the equations 15 and 16 for $y_{it} \in (\text{yield}_{it}, \ln(\text{yield}_{it}))$ and $\mathbf{X}_{it}$ contains the weather controls and two measures of enrollment in insurance; fraction of planted acres insured (level and log), and insurance liability per acre. I start with reporting simple OLS results,

absent spatial effects and instrumental variables, then report results for 2sls-IV and spatial models with and without IV. For the IV-method, I report results in two cases: 1) just identified: two endogenous variables - two IVs, and 2) over-identified: one endogenous variable - two IVs, one at a time, without including the other endogenous variable in the model.

7 Dynamic Effects of Upward Subsidy Shocks in Farm Bills

Comments from Christian Matthes (November 3, 2025):

1. Use long LP as it is the best practice.
2. Have subsidy of each year as the main variable, and amount of subsidy shocks in the previous Farm Bills as control variable.
3. He can join my committee, but will not write a recommendation letter.
4. The choice of spatial LP and any other configurations, depends on my IV variable, and whether it will cover it or not.

I borrowed the Local Projection (LP) method from the macroeconomics literature (Jordà (2023) and Jordà and Taylor (2025)) to study the dynamic effects of subsidies on farm outcomes. LP is a semiparametric method that imposes relatively weak assumptions on data and shape of the response. This choice comes at the cost that LP will be less efficient than models that impose more structure, such as Vector Autoregression (VAR). Using LP for estimating the dynamic effect of subsidies puts this paper in a recent literature on studying the dynamic effects of policy shocks like Baquie et al. (2025) that study the long-term effect of industrial policy instruments on competitiveness of the targeted product and Paredes et al. (2022) that studies the impact of introducing the tax incentive scheme for corporate R&D in firms from different R&D intensity sectors.

I am interested in how the Governmental intervention in the crop insurance market by providing subsidy affects the average outcome in the future relative to a baseline of no-intervention. The impulse response to the intervention is

$$\mathcal{R}_{s\to y}(h,\delta) \equiv \mathbb{E}[y_{t+h} \,|\, s_t = s_0 + \delta;\, \mathbf{x}_t] - \mathbb{E}[y_{t+h} \,|\, s_t = s_0;\, \mathbf{x}_t]\,, \qquad h = 0, 1, \ldots, H. \tag{18}$$

where δ is the size of the intervention, or *dose*, and s_0 is a baseline level. I calculated the heteroskedasticity-robust and serial correlation-robust standard errors within each cluster of county-crop (Arellano 1987).

When it comes to the selection of variables of past periods as explanatory variables (lag of variables), there is no well-established method to select how many lags to include in the LP (Jordà and Taylor 2025, p. 65). The selection criteria are invalid when residuals are autocorrelated, as in LPs for $h > 1$. However, since the LP in the first horizon h = 1 is equivalent to the corresponding equation in a VAR, a natural approach is to use information criteria to determine the lag length similar to VAR method for that case, and then use the same lag length at subsequent horizons (Jordà and Taylor 2025, p. 65).

Knowing this, I chose the optimal number of lags based on the VAR literature, and using the *pvarsoc* command in *STATA*. It reports the model overall coefficient of determination, Hansen (1982) J statistic,

its corresponding p-value, as well as the moment model selection criteria developed by Andrews and B. Lu (2001) based on the J statistic. Andrew and Lu's criteria are all based on Hansen's J statistic, which requires the number of moment conditions to be greater than the number of endogenous variables in the model (Abrigo and Love 2016). Tables 21 and 22 report the Modified Bayesian Information Criterion (MBIC), Modified Akaike Information Criterion (MAIC), and Modified Hannan-Quinn Information Criterion (MQIC) for corn and soybeans, respectively. These parameters penalize model complexity (e.g., the number of lags or instruments) to help select the optimal lag length or model specification in GMM-based systems like the panel VAR used in pvarsoc. In Tables 21 and 22, all MBIC, MAIC and MQIC decline monotonically from 1 to 3 lags and reach their minima at the third lag. Because the local projection at $h = 1$ is equivalent to the corresponding VAR, it is standard to select the dynamic order from a (P)VAR at $h = 1$ and carry that order to all horizons (as also recounted by Jordà and Taylor (2025)).

Table 21: Panel VAR lag order selection for corn (Sample: 1993–2012)

lag	CD	J	J pvalue	MBIC	MAIC	MQIC
1	0.8996546	1591.311	0	1479.255	1567.311	1537.693
2	0.9092503	1100.909	2.4e-232	1026.205	1084.909	1065.164
3	0.9202226	181.7747	3.10e-38	144.4229	173.7747	163.902
4	-0.9204785	.	.	.	.	.

No. of obs = 11361
No. of panels = 634
Ave. no. of T = 17.920

Table 22: Panel VAR lag order selection for soybeans (Sample: 1993–2012)

lag	CD	J	J pvalue	MBIC	MAIC	MQIC
1	0.9095104	1244.541	4.4e-259	1132.776	1220.541	1190.985
2	0.9152186	955.4655	6.1e-201	880.9558	939.4655	919.7617
3	0.9311757	199.2435	5.46e-42	161.9887	191.2435	181.3917
4	0.9009046	.	.	.	.	.

No. of obs = 11089
No. of panels = 619
Ave. no. of T = 17.914

In order to model the dynamic effects of crop subsidies, I consider a typical LP of the form

$$\ln(Y_{ci,t+h}) = \theta_h \Gamma_{ci,t} + \sum_{p=1}^{3} \rho_{h,p} \ln(Y_{ci,t-p}) + \sum_{p=1}^{3} \kappa_{h,p} \Gamma_{ci,t-p} + \boldsymbol{\gamma}_h' \mathbf{x}_{ci,t} + \beta_{ci} + \delta_t + \nu_{ci,t+h}, \quad h = 0, 1, \ldots, H. \tag{19}$$

where:

- $\ln(Y_{ci,t+h})$ denotes the log of the outcome variable (e.g., crop yield) for crop c in county i at horizon $t + h$.
- $\Gamma_{ci,t}$ is the endogenous policy or treatment variable of interest (e.g., insurance subsidy rate)

observed at time t.

- $\Gamma_{ci,t-p}$ for $p = 1, 2, 3$ are the lagged values of the endogenous variable. The optimal number of lags is from Tables 22 and 21.
- $\ln(Y_{ci,t-p})$ for $p = 1, 2, 3$ are the lagged dependent variables, included to control for dynamic persistence. The optimal number of lags is from Tables 22 and 21.
- $\mathbf{x}_{ci,t}$ is a vector of control covariates (e.g., weather variables, input prices, or crop shares).
- $\boldsymbol{\gamma}_h$ is the vector of coefficients on the control variables at horizon h.
- θ_h measures the contemporaneous response of the dependent variable at horizon h to the shock $\Gamma_{ci,t}$.
- $\rho_{h,p}$ and $\kappa_{h,p}$ capture the effects of the p-th lag of the dependent and endogenous variables, respectively.
- β_{ci} represents county-crop fixed effects that absorb time-invariant heterogeneity.
- δ_{ct} represents crop-time fixed effects that absorb aggregate crop-specific shocks common to all counties.
- $\nu_{ci,t+h}$ is the error term at horizon h.
- $h = 0, 1, \ldots, H$ indexes the projection horizon.

Given the endogenous nature of the treatment, we estimate the LP using instrumental variables method, LP-IV, introduced by Jordà, Schularick, and Taylor (2015) in their study of the effect of low interest rate environments on mortgage lending and house prices.

Using the IVs in the LP setting still requires satisfying two conditions of relevance and independence (as first evaluated in Section 5). The exogeneity condition in an LP setting is slightly different than usual due to the dynamic structure of the problem, and it takes the form of (Jordà 2023, sec. 7):

$$\text{Lead-lag exogeneity:} \quad E(\nu_{t+h}, \mathbf{z}_t^P) = 0 \quad \text{for all h.}$$

in which, ν_{t+h} is the error term in h^{th} period impulse response, and $\mathbf{z}_t$ is the vector of instrumental variables. Because lags of the endogenous variable is included in $\mathbf{x}_{ijt}$, potential incidental parameter biases could arise in short panels with a highly serially correlated endogenous variable. Per recommendation of Jordà (2023, p. 626), we used an Arellano-Bond estimator to address this serial correlation.

7.1 Results-Dynamic Effects of Upward Subsidy Shocks in Farm Bills

I used the *locproj* STATA package developed in Ugarte-Ruiz (2025b) and Ugarte-Ruiz (2025a) to model the LP for panel data in presence of endogeniety.

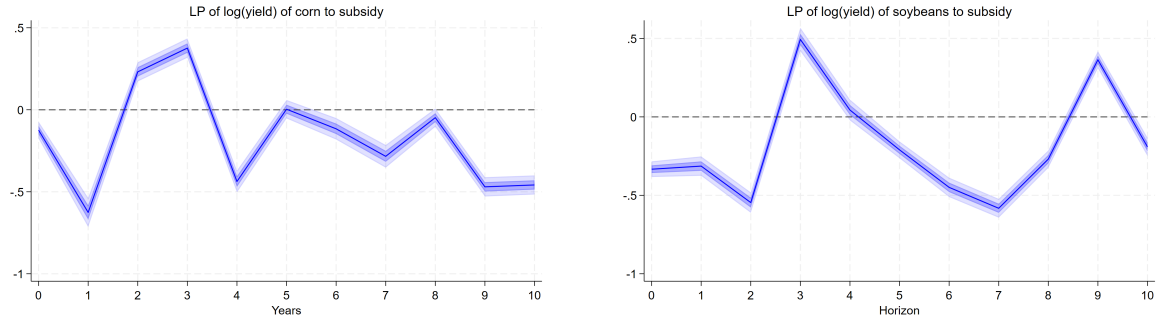
Figure 32 shows results for local project for a 10-year period. The graph shows two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ± 1 standard error band. Notice that since the impulse response coefficients are correlated with one another Jordà (2009) , I cannot make inferential statements about the shape of impulse response based on the error bands.

In order to asses the overall significance of the impulse response, I did a joint test of the null hypnosis that all the response coefficients are zero (as recommended by Jordà (2023)).[9]

Overall Significance of the Impulse Response Since the impulse response coefficients are correlated with one another Jordà (2009) , I cannot make inferential statements about the shape of impulse response based on the error bands. In order to asses the overall significance of the impulse response, I did a joint test of the null hypnosis that all the response coefficients are zero (as recommended by Jordà (2023)). I assess the overall significance of the impulse response by estimating a stacked local projection in which the shock is interacted with horizon dummies, and then conducting a joint Wald test of the null that all horizon-specific coefficients are zero (cf. Jordà (2009); Jordà (2023)). Specifically, letting $\widehat{\boldsymbol{\theta}} = (\hat{\theta}_0, \ldots, \hat{\theta}_H)'$ denote the estimated effects of $1\{H = h\} \times \text{subsidy}_t$ and $\widehat{\Sigma}$ their covariance matrix (estimated with clustered or Driscoll–Kraay standard errors), the test statistic

$$W = \widehat{\boldsymbol{\theta}}^\top \widehat{\Sigma}^{-1} \widehat{\boldsymbol{\theta}} \sim \chi^2_{H+1}$$

under $H_0 : \theta_0 = \cdots = \theta_H = 0$. Rejecting H_0 implies that the impulse response is jointly different from zero across horizons.



(a) Corn: LP of log(yield) to subsidy

(b) Soybean: LP of log(yield) to subsidy

Figure 32: Local projections responses (horizons 0–10) with controls and FE.

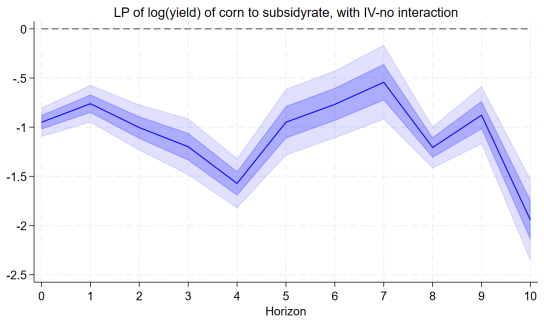
Results in presence of Instrumental Variable When considering an IV for resolving the endogeneity problem, the impulse response changes to 33.

Figures ?? and ?? show results for local project in 10-year period. The error bands provide a measure of estimation uncertainty. Notice that since the impulse response coefficients are correlated with one another Jordà (2009) , we cannot make inferential statements about the shape of impulse response based on the error bands. In order to asses the overall significance of the impulse response, we did a joint test of the null hypnosis that all the response coefficients are zero (as recommended by Jordà (2023)). to do: a joint test of the null hypothesis that all the response coefficients are zero has p-value = xx.[10]

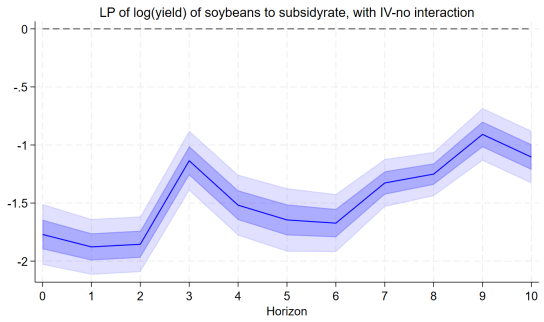
We also tested for structural differences before and after the 1994 reform of the Federal Crop Insurance Program, we interact the shock variable with a post-1994 dummy. When taking sublia as the measure of intensity of insurance enrollment, the resulting impulse response function (Figure 36)

[9] I thank Christian Matthes for his insights about applications of LP. I thank Zhuyin Li for the fruitful discussions about the prerequisites of using LP in research.

[10] I thank Christian Matthes for his insights about applications of LP. I thank Zhuyin Li for the fruitful discussions about the prerequisites of using LP in research.



(a) Corn: LP of log(yield) to subsidy



(b) Soybean: LP of log(yield) to subsidy

Figure 33: Local projections responses (horizons 0–10) with Instrumental Variable of Tsiboe and Turner (2023a) and controls and FE.

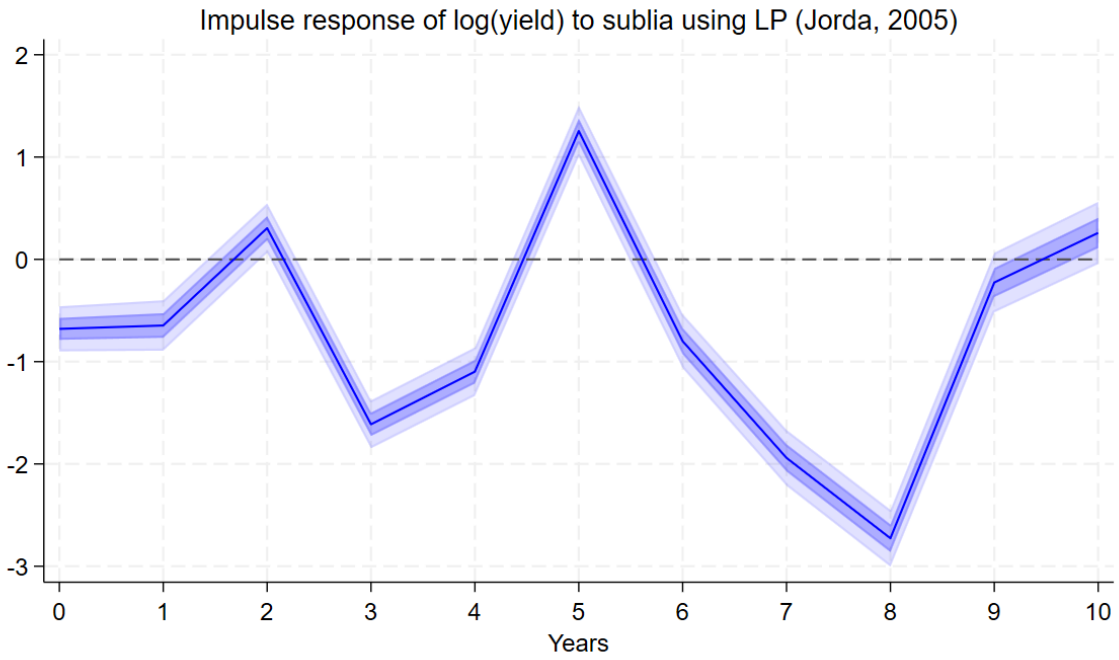


Figure 34: LP with two lag values of yield, sublia and price received as explanatory variables. The graph shows two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ±1 standard error band.

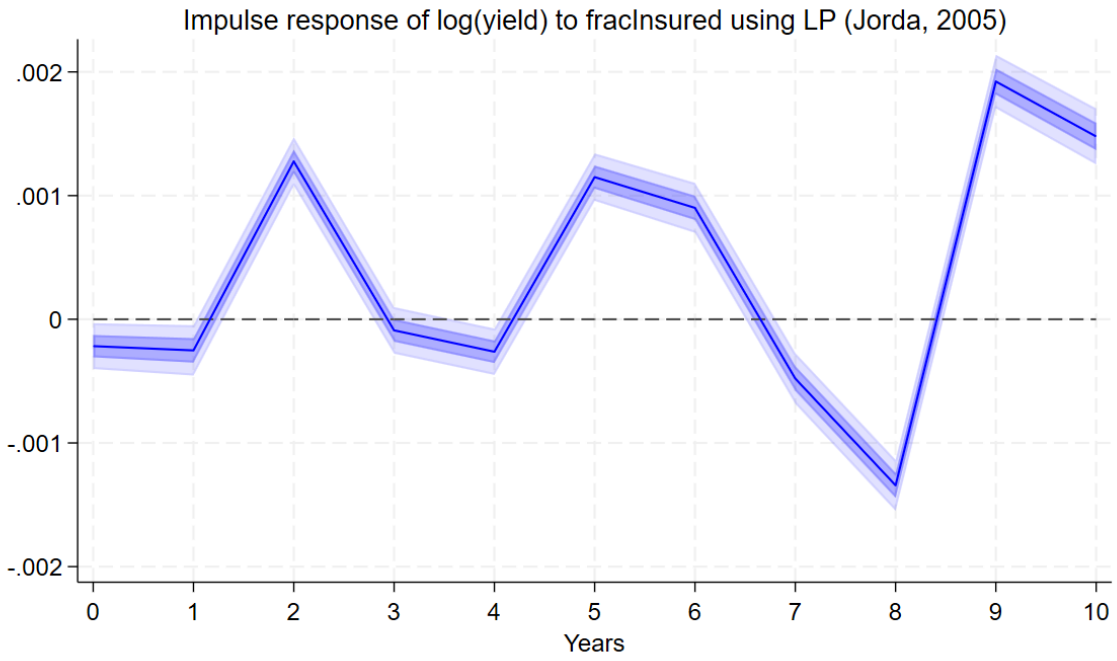


Figure 35: LP with two lag values of yield, fracInsured and price received as explanatory variables. The graph shows two two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ±1 standard error band.

shows that yield losses following subsidy increases are significantly larger and more persistent after 1994, consistent with greater behavioral distortion under the expanded subsidy regime. However, when taking fracInsured as the measure of intensity of insurance enrollment (Figure 37), the pattern is similar but slightly flatter and shifted upward, meaning the negative long-run yield response is weaker after 1994. That implies the 1994 reform may have reduced the marginal distortion of additional enrollment intensity by broadening participation and changing subsidy design.

Figure 38 presents impulse response functions of outcome to changes in sublia as a measure of insurance enrollment intensity. Top panel shows the impulse response of yields to changes in sublia before 1995 is negative and persistent, indicating that higher intensity was associated with slightly lower yields over subsequent years. This suggests limited short run adaptation and possible inefficiencies in production before the 1994 reform that made crop insurance participation mandatory. Middle panel shows that after 1995, the response remains negative but is smaller in magnitude and shorter in duration. This attenuation may reflect structural changes in the Federal Crop Insurance Program, such as broader participation and revised subsidy rates, which reduced the heterogeneity in responses across regions. Finally,,the bottom panel is the difference between post-1995 and pre-1995 responses isolating the policy induced change. The small positive coefficients after the fifth year imply that relative to the earlier period, post reform insurance subsidies had a somewhat less negative (or slightly more stabilizing) effect on yields. This suggests improved program design and risk management practices.

In the top panel of Figure 39, before 1995, a rise in the fraction of insured acreage is followed by a slight, statistically insignificant decline in yields, which then stabilizes. This negative effect may reflect moral hazard or adverse selection in the early years of the insurance program. Farmers who purchased insurance were often those facing higher production risks, and insurance may have reduced incentives to exert full effort or adopt yield-stabilizing practices. In the middle panel, after 1995, the

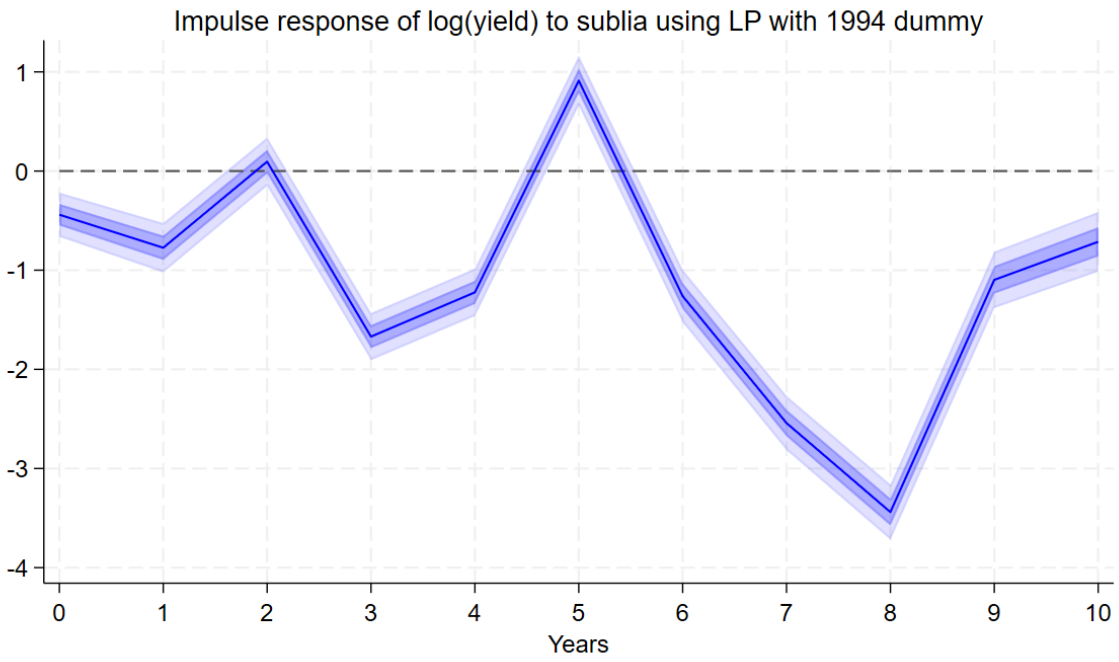


Figure 36: LP with two lag values of yield, sublia and price received as explanatory variables. The graph shows two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ±1 standard error band.

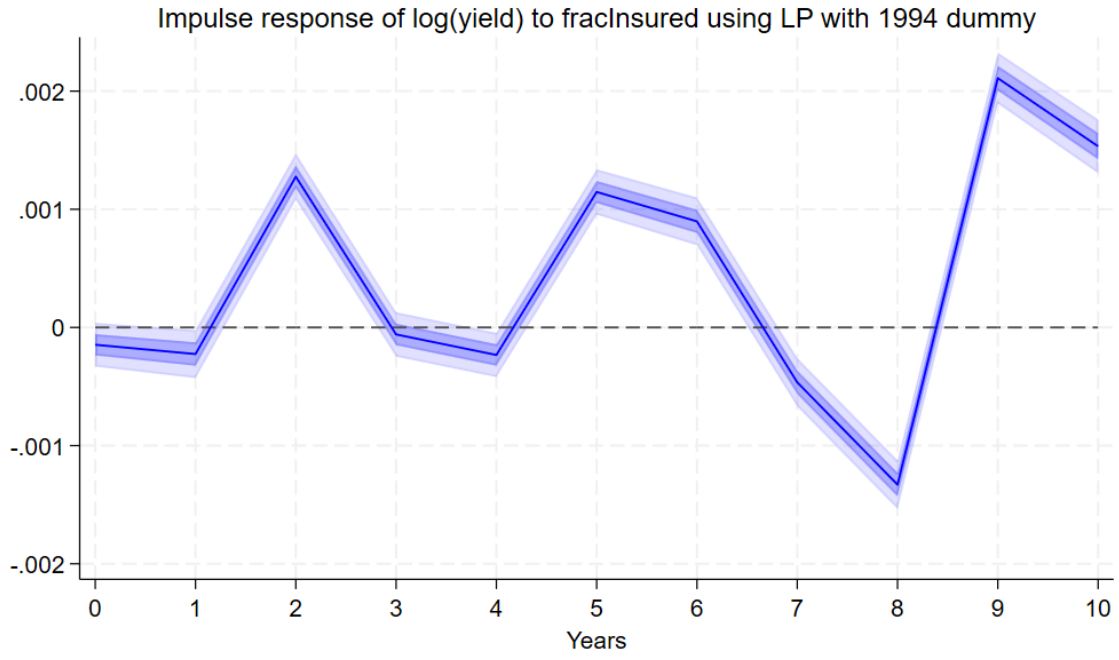


Figure 37: LP with two lag values of yield, fracInsured and price received as explanatory variables. The graph shows two two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ±1 standard error band.

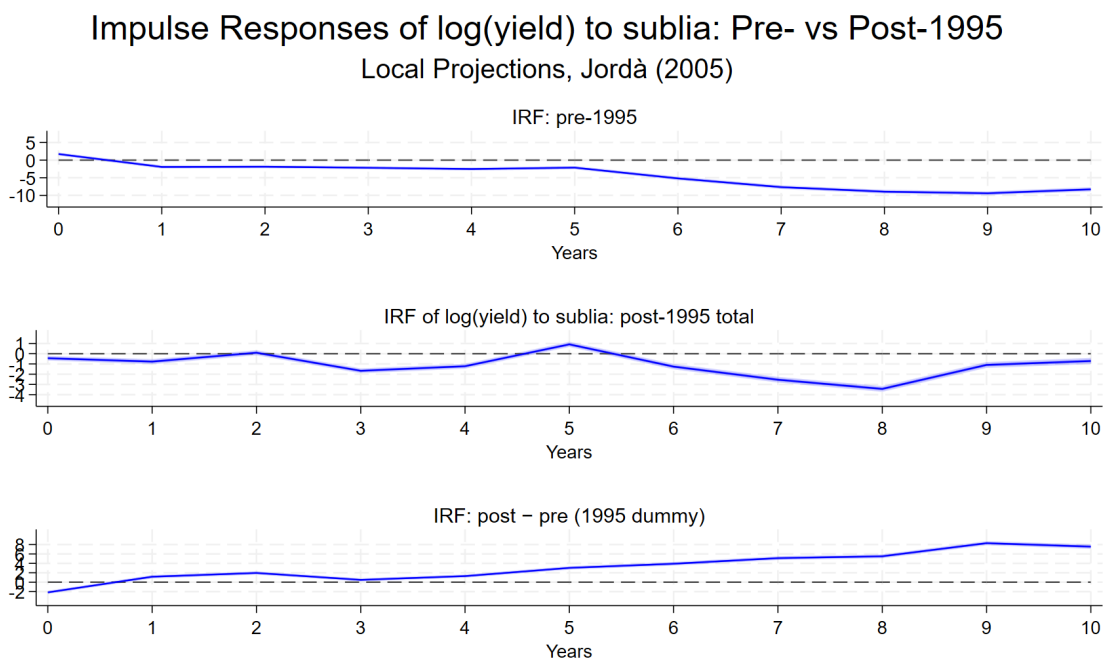


Figure 38: LP with two lag values of yield, fracInsured and price received as explanatory variables. The graph shows two two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ±1 standard error band.

yield response remains small and near zero, but it becomes less negative and more short-lived. This suggests that as crop insurance coverage expanded and the program matured, its distortive effects on yield outcomes diminished. Broader participation and improved underwriting likely mitigated selection biases and strengthened the link between insurance and risk management rather than moral hazard. Finally, in the bottom panel, the differential effect shows a slightly positive and persistent response, implying that, relative to the pre-reform period, yields became modestly higher following increases in insurance coverage. This shift is consistent with improved program design, such as tighter compliance, data-driven premium rating, and the diffusion of adaptive technologies, that enhanced farmers' ability to manage risk without reducing productivity.

7.2 Event Study of Farm Bills as Policy Shocks

In this section, I modified the LP by defining the percent change in county-crop level subsidies between two subsequent 1994, 1996, 2000, 2008, 2014 and 2018 Farm Bills as policy shocks. I used 3 lags for the dependent and shock variables as suggested in Tables 23 and 24.

The two graphs in Figure 40 display local projection (LP) impulse responses of crop yields to subsidy shocks associated with successive U.S. Farm Bills. For corn, the impulse response fluctuates around zero, with alternating positive and negative coefficients across horizons, and modest confidence intervals—indicating that yield responses to subsidy shocks are small and statistically weak. This oscillation could reflect short-term adjustments in production practices, insurance coverage, or acreage allocation following the enactment of new Farm Bill provisions, but without a clear long-term trend. For soybeans, the estimated effects are even smaller, with a mostly flat response and only a slight uptick around horizon 8, suggesting limited yield sensitivity to subsidy changes. The narrow confidence intervals imply precise but near-zero effects. Taken together, these patterns suggest that

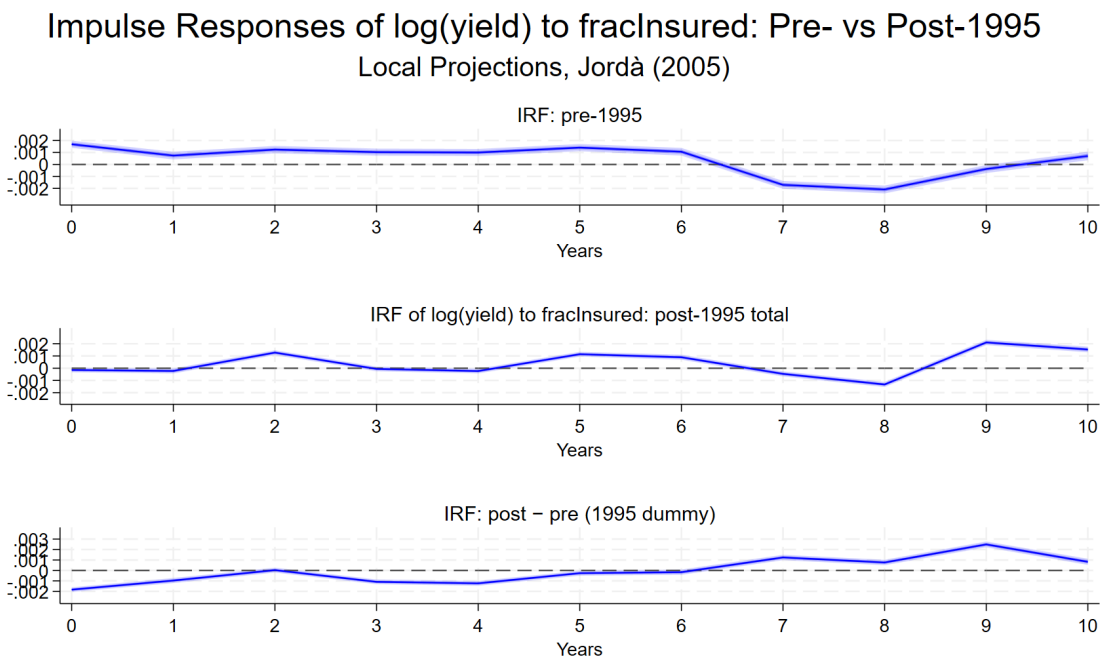


Figure 39: LP with two lag values of yield, fracInsured and price received as explanatory variables. The graph shows two two point-wise confidence intervals at the 95% and at the 66%, with the later to roughly shows the ±1 standard error band.

Table 23: Panel VAR lag order selection for corn with FB shocks (Sample: 1993–2012)

Lag	CD	J	J pvalue	MBIC	MAIC	MQIC
1	0.1137823	1751.253	0	1639.246	1727.253	1697.646
2	0.0256316	1054.1	3.1e-222	979.4286	1038.1	1018.362
3	-0.2004721	489.9997	9.7e-105	452.6638	481.9997	472.1304
4	0.4324079	.	.	.	.	.

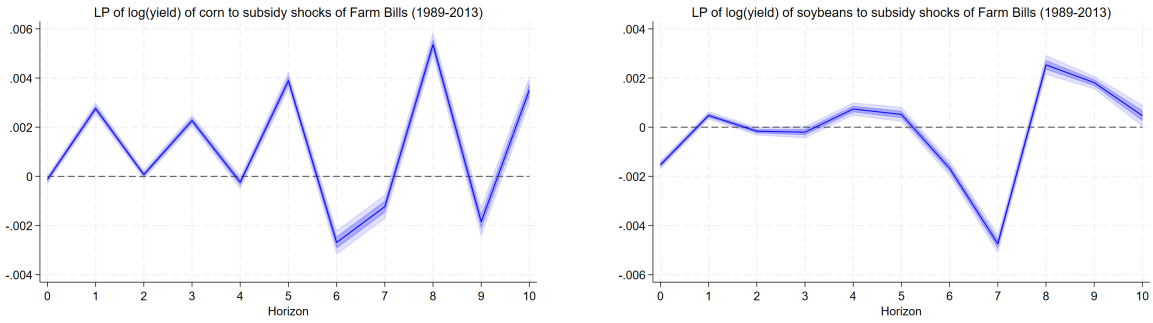
No. of obs = 11361
No. of panels = 633
Ave. no. of T = 17.877

Table 24: Panel VAR lag order selection for soybeans with FB shocks (Sample: 1993–2012)

Lag	CD	J	J pvalue	MBIC	MAIC	MQIC
1	0.2900373	1064.687	2.3e-220	952.9978	1040.687	1011.148
2	0.3108157	718.0961	9.1e-150	643.6364	702.0961	682.403
3	0.2879469	268.5619	6.51e-57	231.3321	260.5619	250.7154
4	0.4176169	.	.	.	.	.

No. of obs = 11020
No. of panels = 619
Ave. no. of T = 17.803

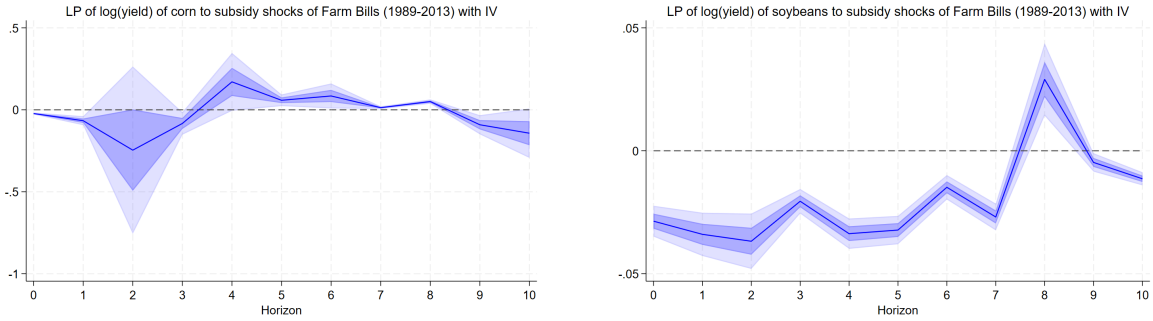
in the Midwest, while Farm Bill subsidy changes affected insurance incentives and risk coverage, did not translate into persistent yield-level responses. Instead, the observed fluctuations likely capture short-run adaptation and risk-management behavior rather than a causal impact of subsidy reform on productivity.



(a) Corn (b) Soybeans

Figure 40: Local projections for corn and soybeans under Farm Bill shocks without IV, without interaction with climate variables, with three lags of log(yield) and subsidy shocks.

In the next step, I used the IV developed by Tsiboe and Turner (2023a) for running the LP-IV as discussed by Jordà, Schularick, and Taylor (2015).



(a) Corn (b) Soybeans

Figure 41: Local projections for corn and soybeans under Farm Bill shocks with textcitetsiboe2023a 's IV

Instrument Relevance. To verify the strength of the instrument `tau_final`, we estimate the first-stage regression of the endogenous variable of shock to subsidies in the Farm Bills on the instrument and all control variables under a fixed-effects specification:

$$\texttt{fbshock}_{it} = \alpha_i + \gamma_t + \pi\, \texttt{tau_final}_{it} + \mathbf{X}'_{it}\boldsymbol{\delta} + \varepsilon_{it},$$

where α_i denotes county fixed effects and $\mathbf{X}_{it}$ includes price, weather, and lagged yield controls. Instrument relevance requires that $\pi \neq 0$, meaning that the instrument is strongly correlated with the endogenous regressor after conditioning on the controls. The coefficient on `tau_final` is large and highly significant ($t = 14.47$, $p < 0.001$), indicating strong correlation between the instrument and the

Table 25: First-stage fixed-effects regression of `fbshock` on `tau_final` and controls

	Coefficient	Std. Err.	t-stat	p-value
`tau_final`	64.338	4.446	14.47	0.000
`price_received`	0.497	0.007	62.77	0.000
`ddayMod_100`	-0.098	0.007	-13.67	0.000
`ddayHot_100`	-0.197	0.008	-25.51	0.000
`prec_cm`	-0.245	0.044	-5.51	0.000
`prec2_cm`	0.002	0.000	6.10	0.000
`corntime`	-0.340	0.029	-11.89	0.000
First-stage F-statistic: $F = 209.38$				

policy shock variable. The corresponding first-stage F-statistic is computed as:

$$F = \left(\frac{\hat{\pi}}{SE(\hat{\pi})}\right)^2 = \left(\frac{64.34}{4.45}\right)^2 = 209.38,$$

which far exceeds the conventional weak-instrument threshold of $F > 10$ (Staiger and Stock 1997), confirming that `tau_final` is a highly relevant instrument for `fbshock`. The main Stata output is summarized in Table 25.

This strong first-stage relationship ensures that `tau_final` provides meaningful exogenous variation in `fbshock`, thereby satisfying the relevance condition for valid instrumental-variable estimation.

8 Dynamic Spatial Effects of Farm Bills

In section 7, I used county fixed effect to absorb all time-invariant spatial heterogeneity like soil quality in modeling the local projection. But the model still assumed that after controlling for fixed effect, the error term is spatially independent across counties. It means that any correlation between neighboring counties' shocks, such as drought spillovers, policy diffusion or local markets, is treated as noise, and hence, not modeled.

In order to bring in information about the spatial propagation, spillovers, or cross-county feedback, I add an extra layer spatial analysis to the model in this section. I follow the spatial local projection method applied by Brady (2011) and Gonne and Hubert (2020) for studying the dynamic effect of subsidy shocks, taking into account the spatial overlaps. I define the spatial autoregressive model by including the spatially lag dependent variable as a regressor, and call it hereafter the *spatial regressor*.[11] This spatially lag variable captures the direct effect of the spatial interaction in determining the dependent variable. As Brady (2011, p. 215) put, spatial correlation can be thought of an omitted variable problem, and adding a regressor that captures the spatial correlation between regions, forms the spatial autoregressive model.

Spatial Weight Matrix Brady (2011, p. 215) defines the spatial weight matrix W by contiguity, and assumes that the spatial propagation does not fade with distance. This simplifying assumption is common in the spatial literature and helps with simplifying the estimation of spatial models (Anselin and Bera 1998).

[11] Brady (2011, p. 215) introduces this style of simplifying the expressions, to prevent confusion with the common terms in the time series analysis.

I use a row-standardized inverse-distance matrices; $w_{ii} = 0$. Common distance-decay choices choose one and expand it:

$$\text{Power (inverse distance):} \quad w_{ij} \propto d_{ij}^{-p}, \qquad \text{Negative exponential:} \quad w_{ij} \propto e^{-\phi d_{ij}}.$$

Using multiple decays ("*NSLX*") is possible by including separate $W^{(1)}$ and $W^{(2)}$ built with different (p, ϕ).

SLX model Following the Spatial Cross–Regressive model of Florax and Folmer (1992) and its modern treatment in Halleck Vega and Elhorst (2015), I exclude a spatial lag of the dependent variable and allow only spatially lagged regressors. For horizon h in a local-projection setting:

$$y_{i,t+h} - y_{i,t-1} = \alpha_i + \tau_t + \beta_h D_{it} + \theta_h (WD_{it}) + \boldsymbol{\gamma}_h^\top \mathbf{X}_{it} + \boldsymbol{\eta}_h^\top (W\mathbf{X}_{it}) + \varepsilon_{i,t+h} \tag{20}$$

in which, WD_{it} captures spatial spillovers of the (endogenous) policy/treatment; and $W\mathbf{X}_{it}$ are optional spatial lags of exogenous controls. α_i and τ_t are unit and time fixed effects; errors are $\varepsilon_{i,t+h}$.

I also estimate the dynamic spatial marginal effects presented in Equation 21 per (Anselin 2024d)'s specification. This is a methodological contribution of this paper. check it with Luc Anselin.

$$\mathbf{IMP} = \frac{\partial \mathbb{E}(\mathbf{Y})}{\partial \mathbf{x}_k'} = \begin{bmatrix} \frac{\partial \mathbb{E}(y_1)}{\partial x_{1k}} & \frac{\partial \mathbb{E}(y_1)}{\partial x_{2k}} & \cdots & \frac{\partial \mathbb{E}(y_1)}{\partial x_{nk}} \\ \frac{\partial \mathbb{E}(y_2)}{\partial x_{1k}} & \frac{\partial \mathbb{E}(y_2)}{\partial x_{2k}} & \cdots & \frac{\partial \mathbb{E}(y_2)}{\partial x_{nk}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbb{E}(y_n)}{\partial x_{1k}} & \frac{\partial \mathbb{E}(y_n)}{\partial x_{2k}} & \cdots & \frac{\partial \mathbb{E}(y_n)}{\partial x_{nk}} \end{bmatrix} = \mathbf{S}^{-1}\boldsymbol{\beta}_k, \tag{21}$$

8.1 Specification Tests of the Dynamic Spatial Estimation

This section assesses whether the empirical model satisfies the assumptions required for valid inference and, where violations are likely, documents the robust procedures I use.

Null Hypothesis: There is no spatial spillover from the insurance subsidy policy.

$$\text{Null Hypothesis:} \quad H_0 : \theta_h = 0$$

It means the spatially lagged treatment WD_{it} does not affect outcomes in other counties (no cross-county policy spillovers).

One common specification test in the spatial modeling literature is Moran's I (Moran 1950). In the context of the spatial local projection, Moran's I serves as a useful but limited specification diagnostic. A failure to reject the null of zero spatial autocorrelation in residuals suggests that the model captures the main spatial structure adequately. However, rejecting the null must be interpreted cautiously, because Moran's I does not identify which spatial process is misspecified, and it could arise from a missing spatial lag of the dependent variable, a spatial error process, or omitted spatially correlated regressors (as in an SLX-type model). Moreover, Moran's I has power against many alternative forms of misspecification, including non-normality and heteroskedasticity, so rejection alone does not confirm spatial dependence. Thus, while Moran's I is a good general misspecification test, it is not diagnostic

in isolation (Anselin 2024e).

Due to this limitation, I used the Lagrange Multiplier (LM) Tests for specification. In spatial specification, one-directional tests like LM error test, LM lag test, or LM SLX (aka LM-WX) tests, have power against other alternatives. It means they reject they Null, when the cause for rejecting Null is not what researchers are actually looking for. This happens beccause these tests are affected by the presence of other alternatives. For example, it is possible to have a situation that the real case is an *error* case, researchers run a *lag* test, and it's get rejected; and researcher conclude that it is a lag, although it is not. This is where this notion of robust test comes into play (Anselin 2024a). What is the written formal reference for this paragraph? To address this possible misinterpretation, Bera and Yoon (1993) proposed a correction of LM tests under local misspecification, and calls it *robust LM-test*. This test is robust to local[12] missepecification under another alternative. Still this test suffers from a phenomenon called *Two out of Three Problem* (Anselin 2024b), which means it is not possible to develop three-way LP tests for lag, error and SLX (aka WX).

Summary Impact Measures Following LeSage and Pace (2009), I estimated two dynamic spatial summary measures as follows:

- **ADI** = Average Direct Impact of x_i on y_i
  - Average of the trace of the impact matrix IMP
  - Formula:
$$ADI = \frac{1}{n}\text{tr}(S^{-1}\beta_k) = \frac{1}{n}\sum_{i=1}^{n}\text{IMP}_{ii}$$

- **ATI** = Average Total Impact
  - Average of all elements of the impact matrix IMP
  - Multiplier: $\frac{1}{1-\rho}$
  - Formula:
$$ATI = \frac{1}{n}\mathbf{1}_n' S^{-1}\mathbf{1}_n\beta_k = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\text{IMP}_{ij}$$

8.2 Choice of Spatial Model

- I did not use a Spatial Lag Model, as it is a cross-sectional model (Anselin 2024c).
- I did not use Spatial Error Model, as the data set is so large, that it is not prone to this error. for the spatial correlation of error terms in measurement. Spatial structure in the covariance matrix. Use HAC standard errors with OLS.
- Very useful: SLX model, rediscovered by Vega and Elhorst (2015) of the Spatial Cross-Regressive Model (Florax and Folmer, 1992). Only spatially lagged endogenously variable. No spatially lagged dependent variable. Use of inverse of distance. Distance Decay graphs with NSLX parameters: inverse distance power function and negative exponential distance.
- Classic Spatial Durbin Model : has spill-over effects

[12] Here local means for small amounts of error in misspecification.

- General Nesting Spatial Model (GNS), by Elhorts (2014): the specification doesn't work. We cannot meaningfully estimate it.

I used the specification test proposed in Bera, Doğan, et al. (2020) for the specification tests of the spatial panel data model.

8.3 Price Volatility

RMA uses a measure of price volatility based on the Black-Scholes Model. This model provides a formula that translates options prices (the amount the market charges to 'lock-in' a future price) into an implied volatility of the price of the commodity. This price volatility is used in the calculation of RMA's premium rates for revenue coverage. The result is that the premium rate RMA charges to lock-in a future (harvest time) price through crop insurance is equivalent to what the market charges to lock in a price through an options contract (United States Department of Agriculture, Risk Management Agency 2025).

9 Optimal Level of Subsidy

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)read jan. 21.2026. In this section, I address the economic efficiency of the subsidized crop insurance, considering the costs and benefits of its provision. On the one hand, government provision of insurance in the context of market imperfection and high correlated risk facilitates insurance that private markets do not otherwise provide. This is the case, as according to Quiggin (1994) read Jan 2026, one problem that distinguishes crop insurance from all other types of insurance, is small potential gain from pooling of similar risks; a problem that makes it financially infeasible to provide for the private sector. While in cases like life insurance, individual risks are uncorrelated, all farms in each district are likely to experience adverse conditions at the same time, and correlations are likely to be high even for large regions. On the other hand, the asymmetric information problems may generate costs that outweigh the benefits. Social costs might arise if farmers alter their behavior in ways that lower expected profits net of insurance indemnities.

To evaluate the welfare impact of the subsidy, I quantify the costs and benefits of farms enrolling in different percentages of insurance coverage. The primary benefit is better income smoothing as more protection is offered against uncertainty. The primary cost is the financial cost of subsidies. However, secondary benefits and costs may arise if diversification affects costs and expected yields. Taking them into account is similar to the job market paper of Adam Solomon (Klosin and Solomon 2024)[13]. The optimal policy is a simple variant of Baily (1978)'s formula for the optimal level of social insurance: the optimal subsidy level equates the marginal gains from a smoother consumption path with the marginal cost, measured by the behavioral response in effort.

I estimate the optimal level of crop insurance subsidies using various measures of demand in the vein of Barry K Goodwin (1993) and E. O'Donoghue (2014). In the empirical literature, Hennessy

[13] Adam Solomon generously shared with me the replication package of his job market paper before its publication. I thank him for his trust and support.

and Babcock (1997), Wang et al. (1998), Adhikari, Knight and Belasco (2012) estimate the welfare benefits from risk protection provided by the crop insurance (Jialing Yu 2017).

In this section, I first calculate the optimal level of subsidy using the a reduced form sufficient statistics. Similar to the previous parts of this paper, I use county-level data in the analysis. This means treating each county as a single farm, as notes by Barry K Goodwin (1993).

9.1 Sufficient Statistics

In this section, I apply the sufficient statistics approach (Baily 1978; Chetty and Finkelstein 2013; Lashkaripour 2021; Lashkaripour and Lugovskyy 2023) to calculate the optimal level of crop subsidy. Applying sufficient statistics would allow this study to estimate the welfare effects of policy changes in terms of estimable elasticities, and performing policy evaluation without making parametric assumptions or estimating the structural primitives of fully specified models. However, this modeling is possible with two assumptions: first, the policy change is small. This would allow using first-order approximations to be precise. And second, government policy is the only wedge between the actual equilibrium and the first-best equilibrium. This means there are no externalities or internalities other than of government that would change by behavioral responses to policy reforms (Kleven 2021, page 517). Working with sufficient statistics also limits the scope to analyze marginal changes in policy. An alternative traditional approach to finding the optimal point is to identify the structural primitives and perform welfare analysis by simulating alternative policies. In principle, that structural approach is an ideal model, but in practice, it is difficult to capture all the primitives of the model (Chetty and Finkelstein 2013, p. 143). Hence, I work with sufficient statistics rather than a fully specified structural model.

The closest prior work to my analysis is the Ph.D. dissertation of Jialing Yu (2017, ch. 4), that characterizes the marginal welfare effect consequences and optimal policy design for the FCIP. Jialing Yu (2017) reports that under current policy settings, the net marginal welfare effects of coverage levels and subsidy rates are negative, and both are greater than their optimal levels. However, if only consider the welfare of the farmers, the marginal welfare benefit of risk protection approximately equals to the welfare cost of moral hazard, indicating that the risk protection level seems to be appropriate in this case. The innovative use of Bootstrap method to introduce uncertainties in parameters, stands out in Jialing Yu (2017)'s work.

In modeling the welfare implications of insurance subsidies and farmers' choices about them, I take into account that most US corn and soybean acres are insured under revenue-based policies, and changes in insurance design or subsidies can shift coverage levels and total insured acreage. However, the relative share of corn vs. soybeans is only modestly affected, as both crops are similarly responsive to insurance incentives (Maisashvili, H. L. Bryant, and Jones 2020; Maisashvili, Fischer, and H. Bryant 2023; Jisang Yu, A. Smith, and Sumner 2018). Some studies note that insurance may encourage expansion onto lower-quality land, potentially affecting crop rotation and environmental outcomes, but the overall impact on crop mix remains small (Claassen, Langpap, and Wu 2017; Jisang Yu, A. Smith, and Sumner 2018).

Given this, I start off with modeling the insurance enrollment without an acreage substitution between corn and soybeans (model A). I estimate robust, IV-backed crop-specific demand for insurance;

and compute the optimal subsidy rate s^*, with sensitivity check on λ, administrative costs. Then I run the quick upper-nest test. If α_k are small and counterfactual mix shifts are less than 1 pp, I explain why nested logit is necessary for my question. If unexpectedly, α_k are meaningful or policy scenarios target one crop asymmetrically, then I promote the nest model.

9.2 Model A (v6): Joint Intensive–Extensive Margins Crop-Specific Buy-Up Demand, Coverage Intensity, and Optimal Subsidy

This section develops a crop-specific model of buy-up crop insurance demand and welfare grounded in the sufficient-statistics approach to social insurance (Baily 1978; Chetty and Finkelstein 2013) and a random-utility foundation for discrete-choice insurance participation (Berry and Haile 2021, sec. 3.2). The goal is to quantify the welfare effects of premium subsidies using estimable demand responses and fiscal parameters, without fully modeling production or consumption. In Model A(v6), I simulate how subsidies jointly affect (i) buy-up participation and (ii) liability per insured acre. The welfare accounting values participation changes using the extensive-margin log-sum surplus. I report changes in coverage intensity and total liability as complementary mechanism outcomes. In my analysis, I leave out a full welfare valuation of the intensive margin. That analysis would require modeling discrete coverage choices (or imposing an explicit valuation of liability conditional on participation), and could be a subject for later work.

Each planted acre faces a discrete choice between enrolling in buy-up coverage and remaining uninsured. This is a simplifying assumption, as in reality, farmers can enroll in one of Catastrophic Coverage (CAT) or Limited Coverage as well. But since buy-up accounts for about 80 percent of contracts, and also farmers do not have much space for maneuver over CAT policies, it is acceptable. Under i.i.d. Type-I extreme value preference shocks, the acre-level participation decision implies a logistic form for the county–crop–year buy-up share. I estimate this share equation using a fractional logit specification (Papke and Wooldridge 1996), which is well suited to outcomes in $[0, 1]$ and accommodates fixed effects and clustered inference. Conditional on enrolling, insured acres choose coverage intensity; I summarize this intensive margin by liability per insured acre and model it with a reduced-form log specification that maps into counterfactual liability changes under alternative subsidy rates.

Markets, premiums, and outcomes. Counties are indexed by c, years by t, and crops by $k \in \{\text{corn}, \text{soybeans}\}$. The potential market size in county–year–crop (c, t, k) is planted acreage A_{ctk}. Let P_{ctk} denote the gross premium per insured acre and let $s \in [0, 1)$ denote the subsidy rate. The farmer-paid premium per acre is $(1 - s)P_{ctk}$. Let $Q_{ctk}(s)$ denote buy-up insured acreage, with buy-up share

$$S_{ctk}(s) \equiv \frac{Q_{ctk}(s)}{A_{ctk}}.$$

Let $L_{ctk}(s)$ denote buy-up *liability per insured acre*. Total buy-up liability in the market is

$$\mathcal{L}_{ctk}(s) \equiv Q_{ctk}(s) \cdot L_{ctk}(s) = A_{ctk}\, S_{ctk}(s)\, L_{ctk}(s).$$

Farmer-paid premiums and empirical proxies. The farmer-paid premium per acre under subsidy rate s is $(1-s)P_{ctk}$. In the empirical implementation, I measure P_{ctk} by the average gross premium per insured acre,

$$P_{ctk} \;\approx\; \frac{\text{total premium payments}_{ctk}}{\text{insured acres}_{ctk}}.$$

I summarize the relevant price faced by farmers using the log farmer-paid premium,

$$\ln \text{FP}_{ctk}(s) \;\equiv\; \ln\big((1 - s)P_{ctk}\big),$$

which maps directly to the policy counterfactual: varying s shifts the farmer-paid premium through $(1 - s)$ holding P_{ctk} fixed.

Extensive margin: buy-up take-up. I model buy-up enrollment using a fractional logit (Papke and Wooldridge 1996). Let $v_{ctk}(s)$ denote the systematic component of the logit index. The market share satisfies

$$S_{ctk}(s) = \Lambda\big(v_{ctk}(s)\big) = \frac{\exp(v_{ctk}(s))}{1+\exp(v_{ctk}(s))}, \tag{22}$$

with index

$$v_{ctk}(s) = \alpha_{ck} + \tau_{tk} - \theta_k \ln \mathrm{FP}_{ctk}(s) + X'_{ctk}\beta_k, \tag{23}$$

where α_{ck} are county–crop fixed effects and τ_{tk} are year–crop fixed effects. The vector X_{ctk} includes climate controls (degree days and precipitation), and I cluster standard errors at the county level. The parameter $\theta_k > 0$ captures the sensitivity of buy-up take-up to the farmer-paid premium.

Intensive margin: liability per insured acre. Conditional on buy-up enrollment, I model coverage intensity using log liability per insured acre:

$$\ln L_{ctk}(s) = \mu_{ck} + \xi_{tk} - \psi_k \ln \mathrm{FP}_{ctk}(s) + X'_{ctk}\gamma_k + \omega_{ctk}, \tag{24}$$

where (μ_{ck}, ξ_{tk}) are county–crop and year–crop fixed effects and X_{ctk} matches the control set in the extensive-margin specification. The parameter $\psi_k > 0$ measures how coverage intensity responds to the farmer-paid premium among insured acres.

Estimation and counterfactual policy. I estimate (23) and (24) separately by crop. Let $\widehat{v}_{ctk}$ denote the fitted extensive-margin index evaluated at the observed subsidy $s_0 \equiv s_{ctk}$, and let $\widehat{\theta}_k$ denote the estimated coefficient on $\ln \mathrm{FP}_{ctk}(s)$. Holding fixed $(P_{ctk}, X_{ctk}, \alpha_{ck}, \tau_{tk})$, counterfactual subsidy policies replace $\ln \mathrm{FP}_{ctk}(s_0)$ with $\ln \mathrm{FP}_{ctk}(s)$:

$$v_{ctk}(s) = \widehat{v}_{ctk} - \widehat{\theta}_k \ln \mathrm{FP}_{ctk}(s_0) + \widehat{\theta}_k \ln \mathrm{FP}_{ctk}(s). \tag{25}$$

The implied counterfactual enrollment share is $\widehat{S}_{ctk}(s) = \Lambda\big(v_{ctk}(s)\big)$ and counterfactual insured acreage is $\widehat{Q}_{ctk}(s) = A_{ctk}\widehat{S}_{ctk}(s)$.

Similarly, let $\widehat{m}_{ctk}$ denote the fitted intensive-margin index at s_0 and let $\widehat{\psi}_k$ denote the estimated coefficient on $\ln \mathrm{FP}_{ctk}(s)$. Counterfactual liability per insured acre is obtained by

$$\ln \widehat{L}_{ctk}(s) = \widehat{m}_{ctk} - \widehat{\psi}_k \ln \mathrm{FP}_{ctk}(s_0) + \widehat{\psi}_k \ln \mathrm{FP}_{ctk}(s), \qquad \widehat{L}_{ctk}(s) = \exp\big(\ln \widehat{L}_{ctk}(s)\big). \tag{26}$$

The implied counterfactual total liability is

$$\widehat{\mathcal{L}}_{ctk}(s) = \widehat{Q}_{ctk}(s) \cdot \widehat{L}_{ctk}(s).$$

Consumer surplus. Under the extensive-margin logit structure, expected consumer surplus per potential acre is the log-sum term

$$CS_{ctk}(s) = \ln\big(1+\exp(v_{ctk}(s))\big). \tag{27}$$

Because discrete-choice utilities are identified only up to an additive constant, the welfare analysis focuses on changes relative to the observed baseline policy s_0, so the relevant object is $CS_{ctk}(s) - CS_{ctk}(s_0)$. In this baseline specification, the log-sum surplus is generated by the binary participation decision (buy-up versus no buy-up) and therefore values welfare changes on the extensive margin; in-

corporating welfare gains from changes in liability conditional on participation would require modeling coverage levels explicitly (e.g., a discrete choice over coverage options).

Baseline-normalized welfare change. Let the fiscal cost of subsidies in (c, t, k) under policy s be

$$C_{ctk}(s) = s \cdot P_{ctk} \cdot Q_{ctk}(s).$$

To convert budgetary subsidy outlays into social costs, I scale fiscal expenditures by $(1 + \lambda)$, where λ is the marginal excess burden (marginal cost) of public funds (Chetty 2009). Let κ_A denote a per-insured-acre real resource cost. The empirical objective function is the change in welfare relative to the observed baseline subsidy s_0:

$$\Delta W_{ctk}(s) = A_{ctk}\big[CS_{ctk}(s) - CS_{ctk}(s_0)\big] - (1+\lambda)\big[C_{ctk}(s) - C_{ctk}(s_0)\big] - \kappa_A\big[Q_{ctk}(s) - Q_{ctk}(s_0)\big]. \quad (28)$$

Equation (28) values policy-induced changes in the extensive margin through the log-sum surplus and accounts for fiscal and resource costs through induced changes in insured acreage. The joint simulation additionally tracks how subsidies affect coverage intensity through $\widehat{L}_{ctk}(s)$ and total liability $\widehat{\mathcal{L}}_{ctk}(s)$, which I report as mechanism outcomes.[14]

Optimal subsidy. Crop-specific welfare change is obtained by aggregating over counties and years:

$$\Delta W_k(s) = \sum_{c,t} \Delta W_{ctk}(s), \qquad k \in \{\text{corn}, \text{soybeans}\}.$$

The crop-specific optimal subsidy maximizes the baseline-normalized welfare objective:

$$s_k^* \in \arg\max_{s \in [0,1)} \Delta W_k(s).$$

In practice, I compute s_k^* by evaluating $\Delta W_k(s)$ on a grid of subsidy rates. I also evaluate a *common* subsidy rate applied to both crops by maximizing $\sum_k \Delta W_k(s)$.

From Model A to identification: why a control function. Equations (23) and (24) treat $\ln \text{FP}_{ctk}(s)$ as conditionally exogenous given fixed effects and controls. In practice, the farmer-paid premium can still be correlated with unobserved determinants of take-up and coverage intensity. For example, local insurance infrastructure, information frictions, and unobserved risk can jointly shape both the premium environment and insurance choices. To recover causal responses to $\ln \text{FP}_{ctk}$—the key sufficient statistic entering the welfare and counterfactual simulations—I therefore augment Model A(v5) with a control-function IV strategy described next.

Control function vs. linear instrumental variable. In the earlier (linear) specifications, I estimate the effect of insurance choices using standard instrumental variables (2SLS), which is appropriate when the structural equation is linear in the endogenous regressor. When I move to estimating the optimal insurance subsidy, the key demand object is inherently *nonlinear* (e.g., an adoption/share model with predictions constrained to $[0, 1]$), so the direct 2SLS plug-in approach is generally not well suited. Instead, I use a *control function* approach—also known as *two-stage residual inclusion*

[14] In the baseline welfare accounting, the intensive margin affects welfare only through costs and observed take-up; incorporating welfare gains from higher liability conditional on participation would require an explicit valuation of coverage intensity (e.g., a discrete-choice model over coverage levels).

(2SRI)—in which the endogenous premium (or price shifter) is first projected on the instruments and controls, and the resulting first-stage residual is included in the nonlinear second stage to purge endogeneity. This method provides an IV-style identification strategy that is compatible with nonlinear likelihood or quasi-likelihood estimation and is widely used in nonlinear models with endogenous regressors (Wooldridge 2015, sec. IV). In the fractional-share context, the nonlinear second stage follows the fractional response framework of Papke and Wooldridge (1996).

9.3 Control Function IV for Nonlinear Buy-Up Demand and Coverage Intensity in Model A (v6)

This subsection describes a control-function (two-stage residual inclusion) approach for addressing endogeneity of the farmer-paid premium in the extensive- and intensive-margin equations used in Model A(v6). The extensive-margin outcome is the buy-up insured share $S_{ctk} \in (0,1)$, and the intensive-margin outcome is liability per insured acre $L_{ctk} > 0$. The joint simulation combines these objects through

$$Q_{ctk}(s) = A_{ctk} S_{ctk}(s), \qquad \mathcal{L}_{ctk}(s) = Q_{ctk}(s) \cdot L_{ctk}(s).$$

Price shifter. Define the farmer-paid premium per acre as

$$\mathrm{FP}_{ctk}(s_{ctk}) \equiv (1 - s_{ctk}) P_{ctk}, \qquad \ln \mathrm{FP}_{ctk}(s_{ctk}) = \ln(1 - s_{ctk}) + \ln P_{ctk},$$

where P_{ctk} is the gross premium per insured acre determined by the rating methodology. The parameter of interest is the causal response of buy-up take-up and coverage intensity to $\ln \mathrm{FP}_{ctk}$.

Instruments and identification concept: premium versus generalized price. The instruments shift the *effective cost* of purchasing insurance. The farmer-paid premium is one component of this generalized price; non-pecuniary access and information frictions can also matter. I use the excluded shifters

$$Z_{ct} = \Big(\texttt{drive_km_to_usda},\ \texttt{stations_count1989},\ \texttt{L_yield},\ \texttt{tau_final}\Big),$$

which proxy (i) access costs and administrative frictions (distance to USDA service infrastructure), (ii) information quality and forecast accuracy (weather-station density), (iii) predetermined performance and rating-related inputs (lagged yield), and (iv) predetermined program parameters affecting rating- and subsidy-relevant components. In the empirical implementation, these shifters generate excluded variation in $\ln \mathrm{FP}_{ctk}$; the control-function approach then removes the component of $\ln \mathrm{FP}_{ctk}$ that is correlated with unobserved determinants of take-up or coverage intensity.

Step 1: first stage for the endogenous price shifter. I project $\ln \mathrm{FP}_{ctk}$ on the excluded shifters, controls, and fixed effects:

$$\ln \mathrm{FP}_{ctk} = Z_{ct}' \pi_1 + X_{ctk}' \pi_2 + \alpha_s + \tau_t + v_{ctk}, \tag{29}$$

where α_s and τ_t denote state and year fixed effects. I cluster standard errors at the county level to allow for arbitrary within-county serial correlation. The fitted residual $\widehat{v}_{ctk}$ summarizes the endogenous component of the farmer-paid premium.

Step 2a: extensive margin with residual inclusion. I augment the buy-up share equation with the first-stage residual:

$$S_{ctk} = \Lambda\Big(\alpha_{ck} + \tau_{tk} - \theta_k \ln \mathrm{FP}_{ctk}(s_{ctk}) + X_{ctk}' \beta_k + \rho_k^S \widehat{v}_{ctk} + \varepsilon_{ctk}^S\Big). \tag{30}$$

Step 2b: intensive margin with residual inclusion. Analogously, I estimate the intensive-margin equation with the residual included:

$$\ln L_{ctk} = \mu_{ck} + \xi_{tk} - \psi_k \ln \mathrm{FP}_{ctk}(s_{ctk}) + X'_{ctk}\gamma_k + \rho_k^L \widehat{v}_{ctk} + \varepsilon_{ctk}^L. \tag{31}$$

Endogeneity tests. The coefficients (ρ_k^S, ρ_k^L) capture endogeneity. Testing $H_0 : \rho_k^S = 0$ (or $H_0 : \rho_k^L = 0$) provides a Rivers–Vuong-style exogeneity test for the corresponding margin.

Policy counterfactuals and joint simulation. For counterfactual subsidy policies, I vary s and shift $\ln \mathrm{FP}_{ctk}(s)$ along the estimated extensive- and intensive-margin relationships while treating $\widehat{v}_{ctk}$ as fixed latent heterogeneity. Mechanically, the counterfactual changes the subsidy wedge $(1 - s)$ in the farmer-paid premium, while the control function continues to absorb the endogenous component of $\ln \mathrm{FP}_{ctk}$ that is correlated with unobserved determinants of choices. I then construct

$$\widehat{Q}_{ctk}(s) = A_{ctk}\widehat{S}_{ctk}(s), \qquad \widehat{\mathcal{L}}_{ctk}(s) = \widehat{Q}_{ctk}(s) \cdot \widehat{L}_{ctk}(s),$$

and evaluate welfare and mechanism implications under alternative subsidy rates.

9.4 Results of Model A (v6)

9.5 Discussion of Model A (v6)

Holding liability fixed, Model A estimated that the extensive margin of optimal subsidy to be 0. Holding take-up fixed, Model A estimated that the intensive margin of optimal subsidy to be xx. These results indicate that the marginal fiscal costs dominate the extensive-margin surplus gains over the relevant range. In all welfare simulations, I evaluate the objective over a policy-relevant grid of subsidy rates $s \in [0, 0.95]$ in one-percentage-point increments. When changing the parameters of kappa and gamma, the shape of response function changes, but the optimal level of subsidies stays at 0 percent.

Marginal Excess Burden of Public Funds (λ): I set the marginal excess burden of public funds to $\lambda = 0.25$, implying that each \$1 of government outlays entails \$1.25 in social cost once the distortionary cost of taxation is accounted for. This value lies within the range of estimates typically reported for the United States (approximately 0.15–0.30) and is consistent with classic general-equilibrium calculations and recent summaries of the marginal cost of public funds (Ballard, Shoven, and Whalley 1985; Browning 1987; Hendren and Sprung-Keyser 2020), as well as prior applications in the crop insurance subsidy literature (e.g., Jialing Yu (2017, ch.4)). To do: for robustness check, I test results for values of λ equal to 0 (no excess burden), 0.15, 0.30 and 0.50 (conservative).

Per Unit Real Resource (Administrative) Cost (κ): I set the per-insured-acre real resource (administrative) cost to $\kappa_A = 0$ as a baseline. Existing evidence suggests that administrative resource costs are on the order of \$1–\$5 per insured acre (Yu 2017, ch. 4). To do: For robustness check, set κ_A to 1, 3, and 5.

My results are consistent with Jialing Yu (2017, Table 4.4), who reports negative marginal welfare effects of premium subsidies at both margins. In Yu's estimates, the marginal effect is negative on the intensive margin (CARA mean -0.275, s.e. 0.085; CRRA mean -0.306, s.e. 0.258) and on the extensive margin (CARA mean -0.535, s.e. 0.175; CRRA mean -0.497, s.e. 0.427). The combined (net) marginal effect is also negative (CARA mean -2.877, s.e. 612.342; CRRA mean -2.846, s.e. 511.738), implying that, at the margin, higher premium subsidies reduce welfare.

In order to put the results in perspective, one can compare it with the optimal subsidy rate in other insurance families. A recent paper on this exercise of calculating the optimal level of insurance subsidies, is Gruber and Solomon (2026) that calculates the optimal flood insurance subsidy. The authors estimate a substantial subsidy rate of 46% as the welfare maximizing rate.

9.6 Section 9.2 Conclusions

The Federal Crop Insurance Program (FCIP) insures farmers against income losses arising from yield and price risk. As the program has expanded, policy debate has increasingly emphasized potential unintended consequences—including changes in production practices, land allocation, and environmental externalities—that may offset some of the benefits from risk protection. A central question for policy design is therefore whether the welfare gains generated by premium subsidies outweigh their fiscal and real resource costs once behavioral responses are taken into account.

Model A provides a transparent welfare assessment of buy-up premium subsidies using a sufficient-statistics framework. The approach places benefits and costs in a common metric and links them

directly to estimable objects: (i) the causal response of buy-up participation to the farmer-paid premium and (ii) the fiscal cost of subsidizing premiums. I model each county–crop–year market with planted acreage as the market size and estimate a crop-specific participation equation using a fractional logit specification. Counterfactual subsidy schedules shift the farmer-paid premium while holding market fundamentals fixed, generating predicted changes in insured acreage.

I then conducted welfare comparisons relative to the observed baseline policy. Benefits are measured using the log-sum surplus implied by the extensive-margin logit participation decision, while costs are captured by subsidy outlays scaled by the marginal cost of public funds and any per-insured-acre resource cost. In addition, I simulate how subsidies affect coverage intensity—liability per insured acre—and total insured liability as complementary mechanism outcomes. Together, these components deliver a policy-relevant mapping from reduced-form demand estimates to quantitative welfare evaluations and an empirically implementable calculation of the welfare-maximizing subsidy rate for buy-up coverage.

9.7 Model A (v4): Extensive Margin Only Crop-Specific Buy-Up Demand and Optimal Subsidy

This section develops a crop-specific model of buy-up crop insurance demand and welfare grounded in the sufficient-statistics approach to social insurance (Baily 1978; Chetty and Finkelstein 2013) and a random-utility foundation for discrete-choice insurance participation (Berry and Haile 2021, sec. 3.2). The goal is to quantify the welfare effects of premium subsidies using estimable enrollment responses and fiscal parameters, without fully modeling production or consumption.

In the model, each planted acre faces a discrete choice between enrolling in buy-up coverage and remaining uninsured. Under an i.i.d. Type-I extreme value assumption on unobserved taste shocks, this acre-level random-utility model implies a logit take-up probability and therefore a logistic form for the county–year–crop buy-up share. Empirically, I estimate this share equation using a fractional logit specification (Papke and Wooldridge 1996), which is well suited to outcomes in $[0, 1]$ and accommodates fixed effects and clustered inference.

I index counties by c, years by t, and crops by $k \in \{\text{corn}, \text{soybeans}\}$. The potential market size in a county–year–crop is the planted acreage A_{ctk}. Each acre chooses between enrollment in buy-up insurance coverage ($j = 1$) and non-enrollment ($j = 0$) in a subsidized insurance contract. Let P_{ctk} be the gross actuarially fair premium per acre, and $s \in [0, 1)$ the subsidy rate, so the farmer-paid premium is $(1 - s)P_{ctk}$. Insured acreage is $Q_{ctk}(s)$, with share $S_{ctk}(s) = Q_{ctk}(s)/A_{ctk}$.

Farmer-paid premiums and empirical proxies. The farmer-paid premium per acre under subsidy rate s is $(1 - s)P_{ctk}$, where P_{ctk} denotes the gross premium per insured acre in county c, year t, and crop k. In the empirical implementation, P_{ctk} is measured by the average gross premium per insured acre,

$$P_{ctk} \approx \frac{\text{total premium payments}_{ctk}}{\text{insured acres}_{ctk}}.$$

Rather than entering the subsidy component and the gross premium separately, I summarize the relevant price faced by farmers using the log of the farmer-paid premium:

$$\ln \mathrm{FP}_{ctk}(s) \;\equiv\; \ln\big((1-s)P_{ctk}\big).$$

This formulation maps directly to the policy counterfactual: varying s shifts the farmer-paid premium through the factor $(1-s)$ while holding the gross premium component P_{ctk} fixed.

Demand. I model the buy-up enrollment using a fractional logit (Papke and Wooldridge 1996) with a logit link. Let $v_{ctk}(s)$ denote the systematic component of the logit index. The market share satisfies

$$S_{ctk}(s) = \Lambda\big(v_{ctk}(s)\big) = \frac{\exp(v_{ctk}(s))}{1+\exp(v_{ctk}(s))}, \tag{32}$$

with index

$$v_{ctk}(s) = \alpha_{ck} + \tau_{tk} - \theta_k \ln \mathrm{FP}_{ctk}(s) + X'_{ctk}\beta_k. \tag{33}$$

Here α_{ck} are *county–crop* fixed effects and τ_{tk} are *year–crop* fixed effects (implemented by estimating the model separately for each crop and including year fixed effects within crop). The vector X_{ctk} includes climate controls (degree days and precipitation), and standard errors are clustered at the county level. The parameter $\theta_k > 0$ captures the sensitivity of buy-up take-up to the farmer-paid premium.

Estimation and counterfactual policy. I estimate the buy-up share equation separately by crop. Let $\widehat{v}_{ctk}$ denote the fitted index evaluated at the observed subsidy $s_0 \equiv s_{ctk}$, and let $\widehat{\theta}_k$ denote the estimated coefficient on $\ln \mathrm{FP}_{ctk}(s)$. To evaluate counterfactual subsidy policies, I hold fixed $(P_{ctk}, X_{ctk}, \alpha_{ck}, \tau_{tk})$ and replace the baseline farmer-paid premium term $\ln \mathrm{FP}_{ctk}(s_0)$ with its counterfactual value $\ln \mathrm{FP}_{ctk}(s)$:

$$v_{ctk}(s) = \widehat{v}_{ctk} - \widehat{\theta}_k \ln \mathrm{FP}_{ctk}(s_0) + \widehat{\theta}_k \ln \mathrm{FP}_{ctk}(s). \tag{34}$$

The implied counterfactual enrollment share is $\widehat{S}_{ctk}(s) = \Lambda\big(v_{ctk}(s)\big)$, and counterfactual insured acreage is $\widehat{Q}_{ctk}(s) = A_{ctk}\widehat{S}_{ctk}(s)$.

Consumer surplus. Under the logit structure, expected consumer surplus per potential acre is given by the log-sum term

$$CS_{ctk}(s) = \ln\big(1+\exp(v_{ctk}(s))\big). \tag{35}$$

Because discrete-choice utilities are identified only up to an additive constant, I conduct the welfare analysis in changes relative to the observed baseline policy s_0. Accordingly, the relevant object is $CS_{ctk}(s) - CS_{ctk}(s_0)$.

Baseline-normalized welfare change. Let the fiscal cost of subsidies in (c,t,k) under policy s be

$$C_{ctk}(s) = s \cdot P_{ctk} \cdot Q_{ctk}(s).$$

To convert budgetary subsidy outlays into social costs, I scale fiscal expenditures by $(1+\lambda)$, where λ is the marginal excess burden (marginal cost) of public funds (Chetty 2009). Let κ_A denote a per-insured-acre real resource cost. The empirical objective function is the change in welfare relative to

the observed baseline subsidy s_0:

$$\Delta W_{ctk}(s) \equiv W_{ctk}(s) - W_{ctk}(s_0) = A_{ctk}\big[CS_{ctk}(s) - CS_{ctk}(s_0)\big] - (1+\lambda)\big[C_{ctk}(s) - C_{ctk}(s_0)\big] - \kappa_A\big[Q_{ctk}(s) - Q_{ctk}(s_0)\big]. \tag{36}$$

This normalization ensures that consumer surplus and fiscal costs are compared on the same baseline, and avoids mixing a change in surplus with spending levels.

Optimal subsidy. Crop-specific welfare change is obtained by aggregating over counties and years:

$$\Delta W_k(s) = \sum_{c,t} \Delta W_{ctk}(s), \qquad k \in \{\text{corn}, \text{soybeans}\}.$$

The crop-specific optimal subsidy maximizes the baseline-normalized welfare objective:

$$s_k^* \in \arg\max_{s \in [0,1)} \Delta W_k(s).$$

In practice, s_k^* is computed by evaluating $\Delta W_k(s)$ on a grid of subsidy rates and selecting the maximizer. In addition to crop-specific optima, the code also evaluates a *common* subsidy rate applied to both crops by maximizing $\sum_k \Delta W_k(s)$.

Mapping to the sufficient-statistics framework. Equation (36) aligns with the sufficient-statistics logic of Baily (1978) and Chetty (2006). The benefit component is summarized by the estimated enrollment response to the farmer-paid premium through the log-sum surplus change $CS_{ctk}(s) - CS_{ctk}(s_0)$, while the cost component is summarized by fiscal parameters (P_{ctk}, λ) and the induced change in insured acreage $Q_{ctk}(s) - Q_{ctk}(s_0)$. The optimal subsidy therefore depends on (i) demand responsiveness to the farmer-paid premium and (ii) the fiscal cost of subsidizing premiums, rather than on the full set of primitives governing risk preferences or production technology.

Model A treats planted acreage A_{ctk} as exogenous with respect to insurance incentives, so crops are modeled as independent markets. This abstraction is appropriate when insurance-induced substitution between corn and soybeans is empirically small, so the primary margin of adjustment is buy-up coverage intensity within each crop rather than crop choice.

Predetermined parameters from the literature. I calibrate the policy primitives using values commonly adopted in the public finance and insurance literature. I set the marginal excess burden of public funds to $\lambda = 0.25$, implying that each \$1 of government outlays entails \$1.25 in social cost once the distortionary cost of taxation is accounted for. This value lies within the range of estimates typically reported for the United States (approximately 0.15–0.30) and is consistent with classic general-equilibrium calculations and recent summaries of the marginal cost of public funds (Ballard, Shoven, and Whalley 1985; Browning 1987; Hendren and Sprung-Keyser 2020), as well as prior applications in the crop insurance subsidy literature (e.g., Jialing Yu (2017, ch.4)). To do: for robustness check, I test results for values of λ equal to 0 (no excess burden), 0.15, 0.30 and 0.50 (conservative).

I set the per-insured-acre real resource (administrative) cost to $\kappa_A = 0$ as a baseline, following common practice in optimal subsidy exercises. Existing empirical evidence suggests that administrative resource costs are on the order of \$1–\$5 per insured acre (Yu 2017, ch. 4). To do: For robustness check, set κ_A to 1, 3, and 5.

In all welfare simulations, I evaluate the objective over a policy-relevant grid of subsidy rates $s \in [0, 0.90]$ in one-percentage-point increments.

9.8 Control Function IV for Nonlinear Insurance Demand in Model A (v4)

This subsection describes how I address endogeneity of the farmer-paid premium in the nonlinear buy-up demand equation used in Model A. The extensive-margin outcome is the buy-up insured share for crop k in county c and year t, denoted $S_{ctk} \in (0,1)$. The baseline demand specification is

$$S_{ctk} = \Lambda\Big(\alpha_{ck} + \tau_{tk} - \theta_k \ln \mathrm{FP}_{ctk}(s_{ctk}) + X'_{ctk}\beta_k + u_{ctk}\Big),$$

where $\Lambda(\cdot)$ is the logistic link, α_{ck} are county–crop fixed effects, τ_{tk} are year–crop fixed effects, X_{ctk} collects exogenous controls (including climate variables), and

$$\ln \mathrm{FP}_{ctk}(s_{ctk}) \equiv \ln\big((1 - s_{ctk})P_{ctk}\big)$$

is the log farmer-paid premium per acre, with P_{ctk} denoting the gross premium per insured acre. The parameter of interest is θ_k, the causal semi-elasticity of buy-up demand with respect to the farmer-paid premium.

Motivation. The farmer-paid premium may be endogenous in the share equation. For example, unobserved county–crop risk, insurer participation, information frictions, or local program infrastructure can affect both the premium environment and insurance demand, inducing correlation between $\ln \mathrm{FP}_{ctk}(s_{ctk})$ and the unobserved demand shifter u_{ctk}. In linear models, two-stage least squares (2SLS) is a standard solution; however, in nonlinear logit or fractional logit models, 2SLS does not generally deliver consistent estimates. I therefore use a *control-function* (two-stage residual inclusion) approach; see Wooldridge (2015) read Jan 31, 2026 for formal treatments.

Step 1: first stage for the endogenous price shifter. Let Z_{ct} denote a vector of excluded instruments that shift the farmer-paid premium but are plausibly orthogonal to the structural demand shock u_{ctk}, conditional on controls and fixed effects. I estimate a reduced-form equation for the endogenous price shifter:

$$\ln \mathrm{FP}_{ctk}(s_{ctk}) = Z'_{ct}\pi_1 + X'_{ctk}\pi_2 + \alpha_{ck} + \tau_{tk} + v_{ctk},$$

where (α_{ck}, τ_{tk}) are county–crop and year–crop fixed effects. county or state? in the ols and spatial regressions, I did state fixed effects The fitted residual $\widehat{v}_{ctk}$ captures the component of the farmer-paid premium that is correlated with unobserved determinants of demand.

Step 2: structural demand with a control function. I augment the nonlinear demand equation with the first-stage residual:

$$S_{ctk} = \Lambda\Big(\alpha_{ck} + \tau_{tk} - \theta_k \ln \mathrm{FP}_{ctk}(s_{ctk}) + X'_{ctk}\beta_k + \rho_k \widehat{v}_{ctk} + \varepsilon_{ctk}\Big).$$

Under standard conditions, the control function absorbs the endogenous component of the price shifter. In particular, if the conditional expectation of the structural error satisfies $E[u_{ctk} \mid \ln \mathrm{FP}_{ctk}, Z_{ct}, X_{ctk}, \alpha_{ck}, \tau_{tk}] = h(v_{ctk})$ and $h(\cdot)$ is well approximated by a linear term, then conditioning on $\widehat{v}_{ctk}$ renders $\ln \mathrm{FP}_{ctk}$ conditionally exogenous. Estimating the augmented equation by fractional logit yields consistent estimates of $(\theta_k, \beta_k, \rho_k)$.

Endogeneity test. The coefficient ρ_k is an endogeneity parameter: testing $H_0 : \rho_k = 0$ provides a Rivers–Vuong-style exogeneity test. If ρ_k is small, the simpler single-equation fractional logit may

be adequate; if $\rho_k \neq 0$, the control-function specification recovers the causal effect of the farmer-paid premium on buy-up take-up while preserving the nonlinear probability structure used for welfare calculations.

Policy counterfactuals. For counterfactual subsidy policies that vary s while holding instruments fixed, I shift $\ln \mathrm{FP}_{ctk}(s)$ along the estimated demand curve, while treating $\widehat{v}_{ctk}$ as fixed latent heterogeneity. Intuitively, the policy experiment changes the mechanical subsidy component $(1 - s)$ in the farmer-paid premium, while the control function continues to net out the endogenous premium component correlated with demand shocks.

9.9 Results of Model A (v4)

Optimal buy-up subsidy. Figure 42 plots the baseline-normalized welfare change $\Delta W(s)$ implied by Model A over a grid of subsidy rates. Panel A reports crop-specific welfare, $\Delta W_k(s) = \sum_{c,t} \Delta W_{ctk}(s)$, separately for corn and soybeans. Panel B reports total welfare under a common subsidy rate applied to both crops, $\sum_k \Delta W_k(s)$. Throughout, welfare is reported per planted acre to facilitate interpretation and comparability across crops and years.

The welfare-maximizing subsidy rate is obtained by grid search:

$$s_k^* \in \arg\max_{s \in [0,1)} \Delta W_k(s), \qquad s_{\text{common}}^* \in \arg\max_{s \in [0,1)} \sum_k \Delta W_k(s).$$

In the baseline calibration, the welfare objective is minimized at an interior subsidy rate and maximized at a boundary of the policy grid, implying a corner solution for s^*.[15]

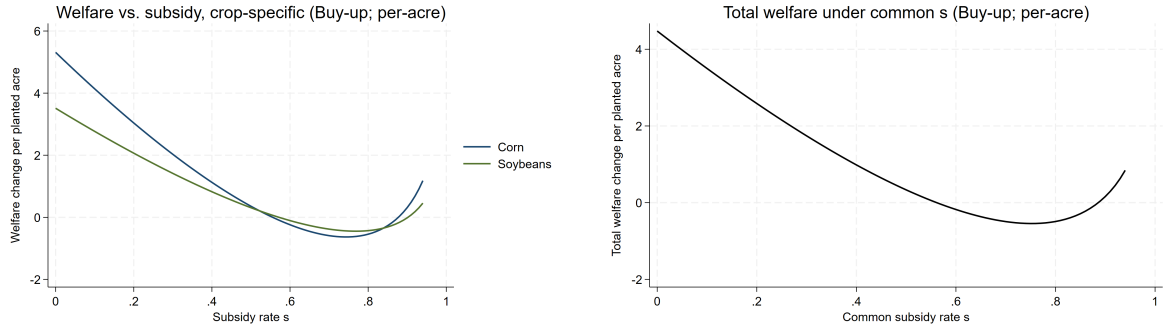
Table 26: Optimal buy-up subsidy rates in Model A

	Corn	Soybeans	Common s
Optimal subsidy s^*			
Welfare at s^* (per acre)			

Notes: Entries are computed by grid search over $s \in [0, 0.95]$ in one-percentage-point increments. Welfare is baseline-normalized relative to the observed subsidy s_0.

[15] This result is conditional on holding gross premiums P_{ctk} fixed when varying s, and on the assumed values of the marginal excess burden of public funds λ and per-insured-acre resource cost κ_A. Appendix **??** reports sensitivity to alternative calibrations.

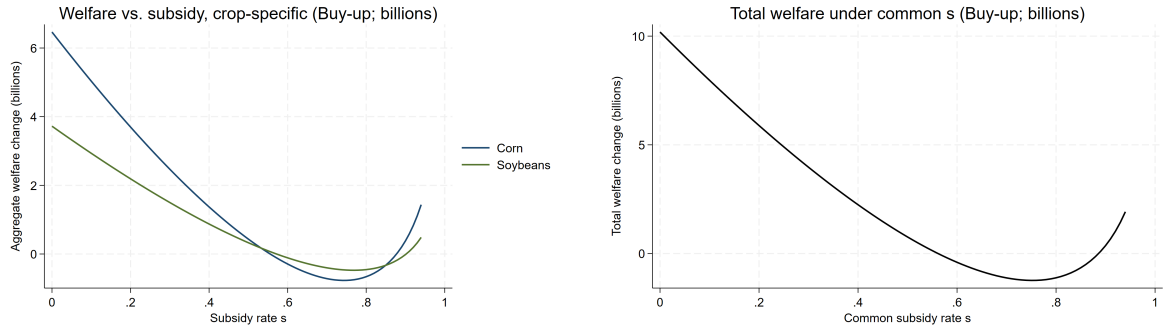
Figure 42: Model A: Welfare versus subsidy rate for buy-up coverage (per planted acre)



(a) Crop-specific welfare $\Delta W_k(s)$ (b) Total welfare under a common subsidy rate

Notes: Each curve reports baseline-normalized welfare relative to the observed policy s_0, aggregated over county–year markets. The optimal subsidy is the maximizer of the curve on the evaluation grid. Welfare per acre equals aggregate welfare divided by total planted acres.

Figure 43: Model A: Welfare versus subsidy rate for buy-up coverage (aggregate billions)



(a) Crop-specific welfare $\Delta W_k(s)$ (b) Total welfare under a common subsidy rate

Notes: Same as Figure 42, but aggregated welfare is reported in billions of dollars.

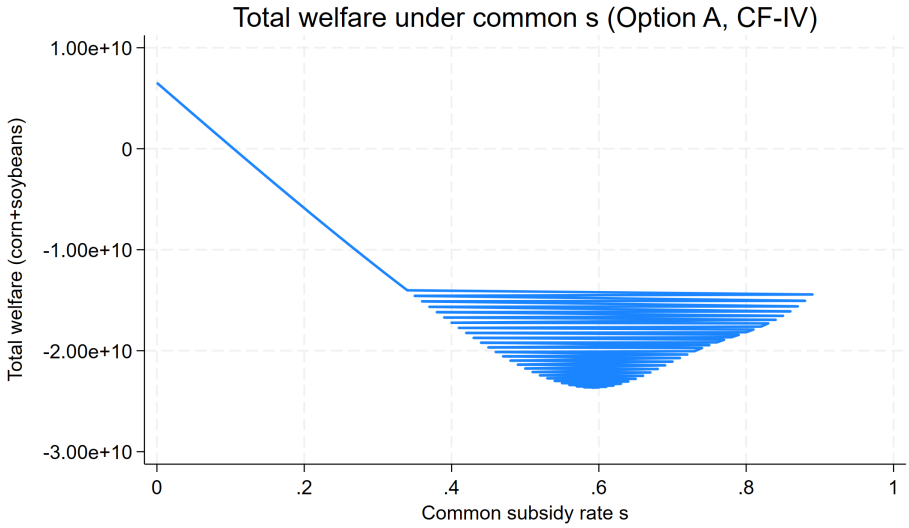
9.10 Discussion of Model A (v4)

Holding liability fixed, Model A estimated that the extensive margin of optimal subsidy to be 0. Holding take-up fixed, Model A estimated that the intensive margin of optimal subsidy to be xx. This model used sufficient statistics to skip the grid search for finding the optimal value.

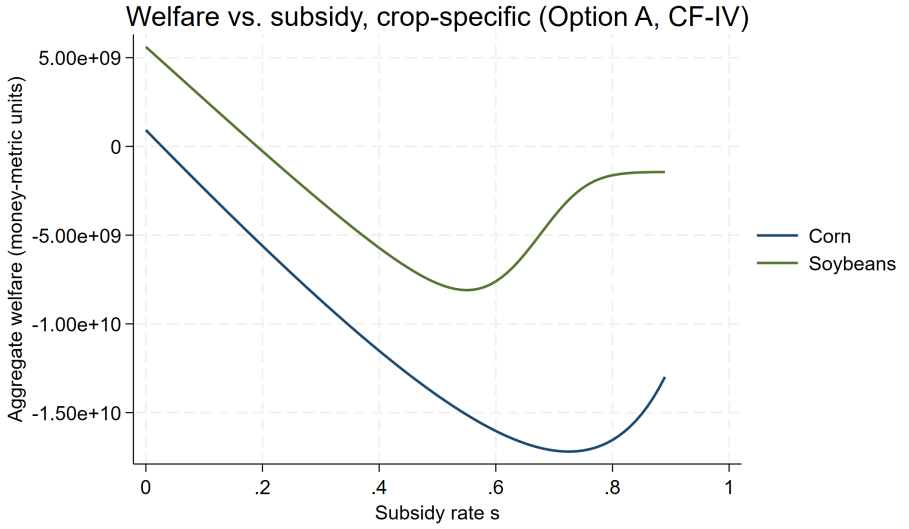
9.11 Results of Model A with IV

Figure 44 reports results for the sufficient statistics exercise. The first panel plots crop-specific welfare as a function of the subsidy rate s using the control function IV (CF-IV) estimates of insurance demand. Because the CF-IV procedure corrects for the endogeneity of the net price $\ln(1-s)$, the implied demand elasticities differ markedly from the baseline no–IV specification. For both corn and soybeans, welfare initially declines as subsidies increase from zero, reflecting rising fiscal costs and only moderate consumer–surplus gains. However, for soybeans the welfare schedule reaches an interior minimum around $s \approx 0.55$ and then begins to rise at higher subsidy levels. This non–monotonicity indicates that, once endogeneity is corrected, the model implies a relatively strong insurance demand response at low net prices, generating substantial consumer–surplus improvements that partially offset the government cost at high subsidy rates. The resulting curvature underscores how optimal–subsidy calculations can change when the price endogeneity in insurance participation is properly addressed.

The second panel aggregates welfare across both crops under a common subsidy rate imposed on the entire market. Unlike the crop–specific case, the combined schedule is dominated by the much larger corn sector and therefore exhibits a more monotonic pattern. Total welfare falls sharply from $s = 0$ to approximately $s = 0.40$, where it reaches its minimum, and then flattens without displaying the strong rebound observed for soybeans alone. At high subsidy levels, the continued welfare losses in corn outweigh the modest gains implied by the soybean demand estimates, producing a relatively smooth and predominantly downward–sloping welfare profile. These results highlight that welfare conclusions can differ substantially between crop–specific and uniform subsidies, and that the CF–IV correction affects the two cases in distinct ways due to differences in market size and estimated demand elasticities.



(a) Two crops combined



(b) Crop-specific

Figure 44: The change with welfare measured with consumer surplus with change in crop insurance subsidy. The county-level data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).

9.12 Model B(v4): Acreage Substitution and Spatial Equilibrium with Crop Insurance

This section extends Model A by allowing farmers to reallocate acreage across crops in response to insurance incentives and closes the model with a competitive land market. The resulting framework delivers a spatial equilibrium in which crop insurance subsidies affect insurance take-up, crop allocation, and land rents within counties.

Environment. Consider counties $c \in \mathcal{C}$, crops $k \in \mathcal{K}$, and years t. Each county has a fixed agricultural land endowment $\bar{A}_{ct}$ that can be allocated across crops:

$$\sum_{k \in \mathcal{K}} A_{ckt} = \bar{A}_{ct}.$$

Output prices p_{kt} are taken as given. Yield per acre for crop k in county c is stochastic and given by

$$y_{ckt} = \mu_{ckt} + \eta_{ckt},$$

where η_{ckt} captures weather-related production risk.

Crop insurance. For each (c, k, t), farmers may purchase insurance coverage indexed by $d \in \mathcal{D}$, where $d = 0$ denotes no insurance. Let $I_{ckt}(d, y)$ denote indemnity payments and $\pi^{AF}_{ckt}(d)$ the actuarially fair premium. A subsidy schedule $s_{ckt}(d)$ reduces the farmer-paid premium to

$$P_{ckt}(d) = (1 - s_{ckt}(d))\, \pi^{AF}_{ckt}(d).$$

In addition, farmers face a non-price insurance access cost τ_{ict} that varies across counties due to administrative frictions and insurance infrastructure.

Per-acre profit. Let r_{ct} denote the competitive land rental rate in county c. Per-acre profit from planting crop k and choosing insurance coverage d is

$$\pi_{ickt}(d) = p_{kt} y_{ckt} - g_{ckt} - r_{ct} + I_{ckt}(d, y_{ckt}) - P_{ckt}(d) - \tau_{ict} \mathbf{1}\{d \neq 0\},$$

where g_{ckt} denotes non-land production costs.

Insurance choice. Farmers are risk averse with Bernoulli utility $u(\cdot)$. Conditional on planting crop k in county c, the farmer chooses insurance coverage to maximize expected utility. Let

$$\bar{V}_{ckt}(d) = \mathbb{E}\left[u\left(\pi_{ickt}(d)\right)\right]$$

denote the systematic component of expected utility. Assuming i.i.d. Type-I extreme value preference shocks over insurance choices, the probability of purchasing insurance for crop k in county c is

$$\sigma_{ckt} = 1 - \frac{\exp\left(\bar{V}_{ckt}(0)\right)}{\sum_{d \in \mathcal{D}} \exp\left(\bar{V}_{ckt}(d)\right)}.$$

Crop choice and acreage allocation. Let the expected indirect utility from planting crop k in county c incorporate the optimal insurance decision:

$$W_{ckt} = \mathbb{E}\left[\max_{d \in \mathcal{D}} u\left(\pi_{ickt}(d)\right)\right].$$

Assuming random utility over crop choices, acreage shares follow a multinomial logit structure:

$$\frac{A_{ckt}}{\bar{A}_{ct}} = \frac{\exp\left(\alpha_k + \lambda W_{ckt}\right)}{\sum_{m \in \mathcal{K}} \exp\left(\alpha_m + \lambda W_{cmt}\right)},$$

where α_k captures crop-specific amenities and $\lambda > 0$ governs the responsiveness of acreage allocation to expected returns.

Land market clearing and spatial equilibrium. Land is allocated competitively within each county. The land rental rate r_{ct} adjusts such that total acreage demand equals available land:

$$\sum_{k \in \mathcal{K}} A_{ckt} = \bar{A}_{ct}.$$

A spatial equilibrium is a set $\{A_{ckt}, \sigma_{ckt}, r_{ct}\}$ such that insurance choices are optimal, acreage allocations satisfy crop choice probabilities, and land markets clear in every county.

Policy implications. An increase in insurance subsidies $s_{ckt}(d)$ raises insurance take-up by increasing the expected utility of insured states. Through the crop choice margin, higher insurance generosity increases the relative attractiveness of insured crops, shifting acreage allocation and bidding up local land rents. Consequently, insurance subsidies generate both production responses and capitalization effects in equilibrium.

9.13 Model B(v3): Acreage Substitution and the Nested Logit Framework

To extend Model A and account for potential crop-level reallocation in response to variation in insurance subsidies, I adopt a nested logit framework. In this structure, farmers make two sequential choices: first, they allocate total planted acreage across crops based on the relative attractiveness of insurance-inclusive bundles; second, they decide whether to insure each acre within the chosen crop. This model captures both the extensive margin (crop choice) and the intensive margin (insurance enrollment), offering a unified framework for analyzing policy incentives.

Are random coefficient model and LATE substitutes? Which one should I use? In random coefficients model, I assume stationarity assumption for the coeffcients. In random coefficient models, the partial effects of the endogenous explanatory variables can vary across individual units in unobserved ways. Estimating such models has fallen somewhat out of favor in empirical research yet they (implicitly) play a role in the recent program evaluation literature, where treatment effects are assumed to be heterogeneous. In the last decade or so, the focus in the treatment effects literature has been on quantities that are identified using standard IV methods under general assumptions, with the "local average treatment effect" (LATE) introduced in Imbens and Angrist (1994) being the most popular (Wooldridge 2015, p. 421).

While Model A treats planted acreage as fixed and focuses exclusively on the insurance decision, such a restriction may understate the full behavioral response to differential subsidies. Insurance generosity varies modestly across crops and locations, and if farmers perceive higher net returns to one crop due to more favorable insurance terms, they may shift acreage accordingly. Model B accommodates this possibility.

Lower Nest: Insurance Enrollment Conditional on Crop Choice For each crop $k \in \{\text{corn}, \text{soybeans}\}$, I retain the acre-level logit from Model A to model insurance demand conditional on crop planted:

$$S^{\text{ins}}_{ctk}(s) = \frac{\exp\left(\nu^{\text{ins}}_{ctk}(s)\right)}{1 + \exp\left(\nu^{\text{ins}}_{ctk}(s)\right)},$$

where

$$\nu^{\text{ins}}_{ctk}(s) = \alpha_{ck} + \tau_{tk} - \theta_k \ln(1-s) - \phi_k \ln P_{ctk} + X'_{ctk}\beta_k.$$

This stage remains unchanged from Model A. It captures how insured share within crop responds to price, premium, and controls. The inclusive value for insurance in money-metric utility units is defined as:

$$IV^{\text{ins}}_{ctk} = \ln\left(1 + \exp\left\{\nu^{\text{ins}}_{ctk}(s)\right\}\right).$$

Upper Nest: Crop Choice Based on Insurance-Inclusive Utility Let total planted acreage $A_{ct} = \sum_k A_{ctk}$ be fixed within a county-year. I model the share of land allocated to crop k using a logit response to inclusive utility:

$$\frac{A_{ctk}}{A_{ct}} = \frac{\exp\left(\gamma IV^{\text{ins}}_{ctk} + \alpha_c + \delta_t + Z'_{ctk}\eta\right)}{\sum_j \exp\left(\gamma IV^{\text{ins}}_{ctj} + \alpha_c + \delta_t + Z'_{ctj}\eta\right)},$$

where γ captures the degree to which insurance attractiveness affects crop allocation. When $\gamma = 0$, crop shares are unaffected by insurance and Model B collapses back to Model A. The vector Z_{ctk}

includes agronomic or economic shifters that influence crop profitability independently of insurance.

Welfare Insured acres are given by:

$$Q_{ctk}(s) = A_{ctk} \cdot S^{\text{ins}}_{ctk}(s).$$

Welfare from insurance provision is then:

$$W_{ctk}(s) = A_{ctk} \cdot \left[CS^{\text{ins}}_{ctk}(s) - CS^{\text{ins}}_{ctk}(s_0)\right] - (1+\lambda)sP_{ctk}Q_{ctk}(s) - \kappa_A Q_{ctk}(s),$$

where

$$CS^{\text{ins}}_{ctk}(s) = \ln\left(1 + \exp\left\{\nu^{\text{ins}}_{ctk}(s)\right\}\right).$$

Total welfare aggregates across counties, crops, and years. I numerically solve for the optimal subsidy s^* that maximizes this total welfare, accounting for both crop allocation and insurance decisions.

Interpretation Model B offers a flexible extension of the sufficient-statistics approach in Model A. It allows me to evaluate not only whether insurance subsidies increase take-up, but also whether they distort land use by inducing acreage shifts across crops. A test of $H_0 : \gamma = 0$ provides direct evidence on whether insurance-inclusive utility affects planting decisions. If the estimated γ is small and counterfactual acreage shifts are minimal, Model A may suffice. Otherwise, Model B provides a more comprehensive basis for policy evaluation.

9.14 Model B(v2-adapted to Model A)- Acreage Substitution and the Nested Logit Model

To allow for crop-level reallocation in response to insurance subsidy variation, I extend the crop-specific demand model in Model A into a nested logit structure. In this model, farmers jointly choose which crop to plant and whether to insure it, recognizing that crop choice may respond to relative insurance attractiveness across crops. This formulation permits insurance-induced changes in crop acreage shares, capturing intensive and extensive margin responses within a unified framework.

Model A treats planted acreage as exogenous with respect to insurance subsidies. While this assumption is standard in the crop insurance literature, insurance subsidies vary slightly across crops and counties, raising the possibility that relative insurance attractiveness may affect crop allocation. To allow for insurance-induced acreage substitution between corn and soybeans, I extend Model A to a nested logit framework in which farmers jointly determine crop allocation and insurance enrollment.

Lower Nest: Insurance Enrollment Conditional on Crop Choice. For each crop $k \in \{\text{corn}, \text{soybeans}\}$, county c, and year t, I model the conditional insurance enrollment share among planted acres using the same acre-level logit specification as in Model A:

$$S^{\text{ins}}_{ctk}(s) = \frac{\exp(\nu^{\text{ins}}_{ctk}(s))}{1 + \exp(\nu^{\text{ins}}_{ctk}(s))}, \tag{37}$$

where

$$\nu^{\text{ins}}_{ctk}(s) = \alpha_{ck} + \tau_{tk} - \theta_k \ln(1-s) - \phi_k \ln P_{ctk} + X'_{ctk}\beta_k. \tag{38}$$

The parameters retain the same interpretation as in Model A. As before, identification of θ_k relies on exogenous variation in the net price $\ln(1-s)$, instrumented using policy-driven variation in subsidy schedules.

From this index, I compute the insurance inclusive value,

$$IV^{\text{ins}}_{ctk} = \ln\big(1 + \exp\{\nu^{\text{ins}}_{ctk}(s)\}\big), \tag{39}$$

which represents the money-metric expected utility of insurance conditional on crop choice.

Upper Nest: Crop Acreage Allocation. Let total planted acreage in a county-year be fixed. Farmers allocate land across crops based on the relative attractiveness of each crop's insurance-inclusive value. I model crop acreage shares using a fractional logit:

$$s^{A}_{ctk} = \frac{A_{ctk}}{\sum_j A_{ctj}} = g\big(\gamma\, IV^{\text{ins}}_{ctk} + \alpha_c + \delta_t + Z'_{ctk}\eta\big), \tag{40}$$

where $g(\cdot)$ is the logit link, α_c and δ_t are county and year fixed effects, and Z_{ctk} includes crop-specific agronomic and economic controls. The coefficient γ captures the extent to which insurance attractiveness affects crop allocation. When $\gamma = 0$, acreage allocation is independent of insurance incentives and Model B collapses to Model A.

Welfare. Let insured acres be $Q_{ctk}(s) = S^{\text{ins}}_{ctk}(s) \cdot A_{ctk}$. County–crop–year welfare is

$$W_{ctk}(s) = A_{ctk}\Big[CS^{\text{ins}}_{ctk}(s) - CS^{\text{ins}}_{ctk}(s_0)\Big] - (1+\lambda)\, s\, P_{ctk}\, Q_{ctk}(s) - \kappa_A\, Q_{ctk}(s), \tag{41}$$

where

$$CS_{ctk}^{\text{ins}}(s) = \ln\bigl(1 + \exp\{\nu_{ctk}^{\text{ins}}(s)\}\bigr).$$

Aggregate welfare is obtained by summing over counties, years, and crops. I compute optimal subsidy rates by maximizing total welfare with respect to s, allowing acreage to adjust endogenously through the upper nest.

Interpretation. Model B generalizes Model A by allowing insurance subsidies to affect both the intensive margin of insurance coverage within crops and the extensive margin of crop allocation. The nested structure provides a transparent test of whether insurance subsidies distort land use: a Wald test of $H_0 : \gamma = 0$ evaluates whether insurance incentives induce meaningful acreage substitution between corn and soybeans.

9.15 Model B (v1)- Acreage Substitution and the Nested Logit Model

The crop-specific insurance demand model described above (Model A) treats planted acreage as exogenous with respect to insurance subsidy policy. This assumption is standard in the crop insurance literature, reflecting the view that agronomic constraints, crop rotations, and relative output prices are the primary determinants of acreage allocation between corn and soybeans in the U.S. Midwest. Nonetheless, because insurance subsidies differ slightly across crops and across counties, it is possible that a more generous subsidy for one crop could alter relative expected returns and induce a reallocation of land across crops. To allow for this possibility, we estimate an augmented nested model (Model B) that embeds the plan-level insurance demand structure from Model A within an upper-level acreage allocation equation linking planted area to the "inclusive value" of insurance. to add: graph of comparing corn and soybeans subsidies in a specific county in one state. Like Adams in Illinois from the RIRS data.

Formally, the lower nest recovers, for each crop k, county c, and year t, the conditional insurance demand index x^{ins}_{ctk} from a fractional logit specification. This index captures how the insurance participation share responds to the net premium and other covariates holding acreage fixed. From this index, I compute an insurance inclusive value,

$$IV^{\text{ins}}_{ctk} = \log\left(1 + e^{\,x^{\text{ins}}_{ctk}}\right),$$

which is proportional to the money-metric expected utility of insurance conditional on crop choice. In the upper nest, we model crop acreage shares $s^{A}_{ctk} = \text{acres}_{ctk} / \sum_j \text{acres}_{ctj}$ using a fractional logit equation of the form

$$s^{A}_{ctk} = g\left(\gamma\, IV^{\text{ins}}_{ctk} + \alpha_c + \delta_t + Z'_{ctk}\eta\right),$$

where $g(\cdot)$ is the logit link, α_c and δ_t denote county and year fixed effects, and Z_{ctk} includes optional agronomic or economic controls. The key coefficient is γ, which captures whether counties allocate more acreage to crops for which the insurance bundle is more attractive. When $\gamma = 0$, the upper nest is independent of insurance incentives, and Model B collapses to Model A; when $\gamma \neq 0$, acreage responds endogenously to subsidy-induced changes in insurance value.

Model B therefore nests Model A and provides a straightforward framework for testing whether insurance subsidies distort land allocation. I estimate both the unrestricted specification containing IV^{ins}_{ctk} and the restricted specification imposing $\gamma = 0$.

Update this paragraph with the new updated results:A Wald test of the null hypothesis $H_0 : \gamma = 0$ yields a statistically insignificant result ($\chi^2(1) = 0.99$, $p = 0.32$), and information criteria (AIC and BIC) are virtually identical across the two models (see Section 28). These results indicate that the insurance inclusive value has no detectable effect on acreage shares. In other words, there is no evidence that county-level crop mix responds to insurance subsidy incentives once fixed effects and agronomic controls are accounted for. Consequently, Model A—which treats acreage as predetermined with respect to insurance generosity—is empirically supported and is used as our preferred model for welfare analysis, while Model B serves as a robustness check confirming that subsidy-induced acreage substitution is negligible in our data.

9.16 Results of Model B

Figure 45 shows the welfare simulations based on Model B—the specification that allows insurance subsidies to affect acreage allocation across crops—yield results that are remarkably similar to those from the baseline crop-specific model. Panel (a) of Figure **??** shows that total welfare declines monotonically as the common subsidy rate s increases, with no interior optimum; welfare losses become increasingly large for higher subsidy levels because government outlays scale directly with insured acreage while consumer-surplus gains remain small. Panel (b) illustrates the crop-specific welfare curves for corn and soybeans. Both crops exhibit strongly negative welfare slopes, although the magnitude is larger for corn due to its larger acreage base and higher baseline premium volume. Importantly, the welfare curves are nearly linear in s, and the difference between crops is driven almost entirely by the scale of the insured area rather than by any substitution response across crops. This pattern is consistent with the estimation results from the upper nest, which indicated that the insurance inclusive value IV_{ctk}^{ins} has no significant effect on acreage shares. As a consequence, the welfare implications under Model B closely mirror those from the simpler Model A: insurance subsidies generate substantial fiscal costs with limited offsetting consumer-surplus benefits, and allowing for acreage substitution does not materially change this conclusion.

9.17 Test for Whether Nested Logit (Model B) is Needed After Estimating Model A

In order to assess whether crop insurance subsidies induce acreage substitution between corn and soybeans, we estimated a nested fractional–logit model in which the lower nest governs insurance take-up within crop and the upper nest allocates acreage shares across crops as a function of the inclusive value of insurance, IV_{ctk}^{ins}. In the upper nest we test the null hypothesis $H_0 : \gamma_{IV} = 0$, where γ_{IV} is the coefficient on IV_{ctk}^{ins}. The Wald test reported in Table 27 fails to reject this restriction ($\chi^2(1) = 0.99$, $p = 0.32$), implying that acreage shares do not respond systematically to the insurance inclusive value. Consistently, the information–criterion comparison in Table 28 shows that the unrestricted nested model (Model B) delivers virtually identical log-likelihood, AIC, and BIC values as the restricted specification (Model A) that excludes IV_{ctk}^{ins}. Taken together, these results provide no empirical evidence of subsidy-induced acreage substitution between corn and soybeans; in the rest of the analysis we therefore use the simpler crop-specific model without acreage substitution as our benchmark, and treat the nested specification as a robustness check.

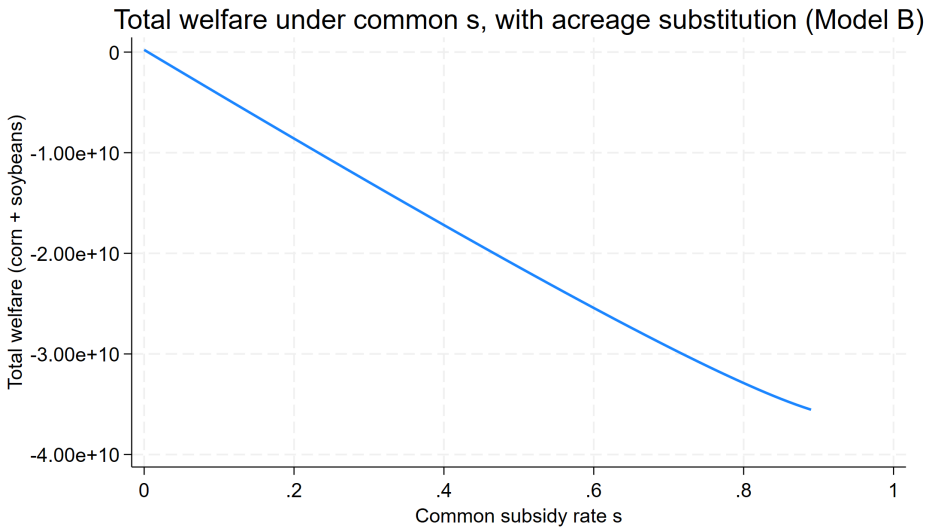
Table 27: Nested test of acreage response to insurance inclusive value

Regressor	$\chi^2(1)$	p–value
IV_{ctk}^{ins}	0.99	0.3205

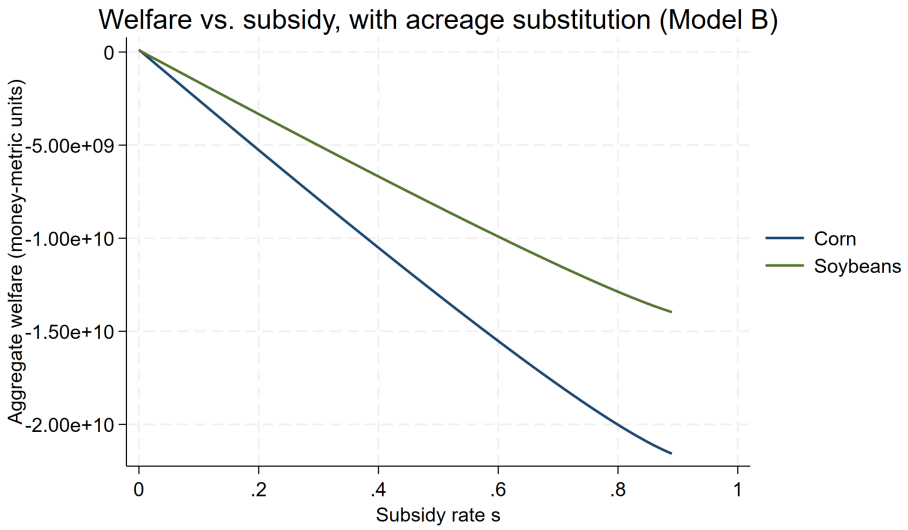
Table 28: Information–criterion comparison of acreage models

Model	N	ℓ(null)	ℓ(model)	df	AIC	BIC
NEST_B	5	−20,760.47	−20,550.02	676	42,452.04	42,188.02
RESTRICTED_A	5	−20,760.47	−20,550.69	676	42,453.39	42,189.37

This code implements a diagnostic test that evaluates whether the *upper nest* in a nested logit model is empirically needed. The logic is as follows:



(a) Two crops combined



(b) Crop-specific

Figure 45: The change with welfare measured with consumer surplus with change in crop insurance subsidy. The county-level data is from U.S. Department of Agriculture, Risk Management Agency (RMA) (2025).

- Model A assumes that crop insurance demand for corn and soy is independent across crops (no correlated unobservables across corn and soy).
- Model B adds an "upper nest" that allows correlation in unobserved utility through the inclusive value (IV) parameters.
- If the inclusive value terms for corn and soy are jointly insignificant, then Model A is sufficient; if they matter, Model B is required.

The Stata test (shown below) proceeds as follows:

1. Use a dataset where crop shares sum to one across {corn, soybeans}:

$$\texttt{share_corn} + \texttt{share_soy} = 1.$$

2. Estimate a fractional multinomial logit model with inclusive value terms:

$$\texttt{share_corn} = f(\texttt{c.IV_corn},\ \texttt{c.IV_soybeans},\ \text{county FE}, \text{year FE}),$$
$$\texttt{share_soy} = f(\texttt{c.IV_corn},\ \texttt{c.IV_soybeans},\ \text{county FE}, \text{year FE}).$$

3. Test whether the inclusive values matter jointly:

$$H_0 : \alpha_{\text{IV,corn}} = \alpha_{\text{IV,soybeans}} = 0.$$

If the p, value of the joint test is above 0.10 and the coefficients are economically tiny, then the inclusive values do not explain crop mix changes, and the simpler Model A is likely sufficient.
The Stata code:

```
* This code tests whether nested logit is needed after Model A.
* If the inclusive value parameters (upper nest) do not matter,
* then Model A is sufficient and Model B adds little.

* Crop shares summed to 1 across {corn, soy}
use crop_shares_with_IV.dta, clear
* Variables: county_id year share_corn share_soy IV_corn IV_soy

* Fractional multinomial logit with inclusive values
ssc install fmlogit, replace
fmlogit share_corn share_soy c.IV_corn c.IV_soy ///
    i.county_id i.year, basevar(share_soy) vce(cluster county_id)

* Test whether inclusive-value parameters matter jointly
test IV_corn IV_soy

* Interpretation:
* If p > 0.10 and effects are tiny, Model A is enough.
* If coefficients are large or significant, Model B (nested logit) is required.
```

```
* Back-of-envelope counterfactual:
* Raise s by +0.05 in the lower nest, recompute IVs,
* Plug into the model, predict new crop shares,
* Report share changes $\Delta$share_corn, $\Delta$share_soy.
```

Parameters not necessary for Model A (but needed for a structural model) The following parameters are unnecessary for Model A (which uses a revealed-preference, inclusive-value welfare formula), but are used in structural risk-utility models:

- Risk aversion (CARA/CRRA parameters), γ,
- Variance terms (c_v), profit gap Δ,
- Correlation parameters (ρ_1, ρ_2),
- Adverse-selection terms (inverse Mills ratio λ_L, marginal indemnity τ_b),
- Moral-hazard elasticities ($\eta_{R,b}$, $\eta_{\sigma,b}$),
- Adverse-selection elasticities ($\eta_{I,\tau}$),
- Subsidy-response elasticities ($\eta_{\tau,s}$).

These are required only if you build a full CARA-Normal or sufficient-statistics structural welfare model. They are not needed for the baseline Model A.

9.18 Model A+

In this section, I use the parameters of Table **??** from Jialing Yu (2017) for an alternative modeling.

Model A treats the welfare gain from insurance as the change in consumer surplus (CS) revealed by the estimated take-up response, minus the social cost of public funds and real resource costs. Three blocks from your parameter table map directly into the model:

Deadweight loss of taxation The Deadweight loss of Table **??** φ (Uniform$(0.2, 0.5)$) corresponds to the public-finance marginal excess burden of public funds. In Model A this is the parameter λ. A canonical baseline is:

$$\lambda = 0.25 \qquad \text{(robustness: } \lambda \in [0.15, 0.30]\text{)}.$$

Government cost therefore enters as:

$$(1 + \lambda) \cdot s \cdot P \cdot Q.$$

Administrative and loading costs Let P denote the gross premium per acre and μ the expected indemnity per acre. Define the total loading:

$$L = P - \mu.$$

Split loading into a real-resource share $\omega \in [0.1, 0.3]$ (A&O + admin) and a redistribution share $(1 - \omega)$:

$$RC = \kappa_A + \omega(P - \mu),$$

where RC is the resource cost per acre. This replaces the simple $\kappa_A Q$ term in Model A. Thus welfare per (c, t, k) becomes:

$$W_{ctk}(s) = A_{ctk}\,\Delta CS_{ctk}(s) - (1+\lambda)\,s\,P_{ctk}\,Q_{ctk}(s) - \big(\kappa_{A,k} + \omega_k(P_{ctk} - \mu_{ctk})\big)\,Q_{ctk}(s).$$

The following parameters from Table ?? can be used for the estimation:

- A&O cost loads (Normal, mean ≈ 0.17),
- CAT loads,
- a general loading factor ι.

Calibration baselines

$$\lambda \in [0.15, 0.30], \qquad \omega \approx 0.15\text{–}0.25, \qquad \kappa_A \approx \text{A\&O dollars per insured acre.}$$

These parameters allow you to separate *fiscal* cost from *real resource* cost in a transparent way.

9.19 Structural Model

Zhong et al. (2023) develops two theoretical models using the game theory framework to analyze the full-cost insurance and the income insurance, and how should governments design insurance subsidies. Their results show that the government's optimal insurance subsidies under different insurance formats are closely related to the yield uncertainty and the production efficiency under the risk of extreme weather. When the production efficiency of the farmers is low, regardless of the subsidy level, the sufficient supply of the agricultural products can be guaranteed under full-cost insurance. However, when production efficiency is high, a lower subsidy level should be set under full-cost insurance, while a higher subsidy level is required under income insurance. Zhong et al. (2023) also finds the optimal levels of the government insurance subsidies under different insurance formats may not always align with the interests of farmers. They also show that the government can maximize social welfare by ensuring the stable supply of agricultural products.

9.20 Discussion of Sufficient Statistics Modeling

Earlier work over roughly the same time period (1989-2012) using county-level data had shows that while increases in subsidies can induce farmers to enroll more land, they primarily encourage them to adopt higher levels of coverage on land already enrolled (E. O'Donoghue 2014).

According to E. O'Donoghue (2014), based on total premiums (the sum of the premium the producer pays for a policy plus associated premium subsidies), the average response for corn producers in the Midwest (IL, IN, IA, and OH) was to increase the demand for crop insurance when the price of crop insurance fell. That is, a 1-percent increase in subsidies per acre aof roughly 0.96 percent in total premium purchases. The average response of corn producers in the Lake States (MI, MN, and WI) and the Northern Plains (KS, NE, and SD) exhibited smaller changes, with total premium purchases increasing by 0.83 percent and 0.80 percent, respectively, following a 1-percent increase in subsidies.

According to E. O'Donoghue (2014), the average responses of soybean producers were similar but smaller than those of corn producers (not surprising, as corn producers often grow soybeans as well). For a 1-percent increase in subsidies, the average response of soybean producers was to increase total

premium purchases by 0.84 percent in the Midwest, 0.59 percent in the Lake States, and 0.47 percent in the Northern Plains.

According to E. O'Donoghue (2014), the effects of subsidy increases are also revealed in other measures of demand for crop insurance during the period. Premiums per acre (total premiums divided by acres insured) and liabilities per acre (total liability divided by acres insured) also tended to rise, with premiums increasing the most. The change in demand measured as acres enrolled (both total and buy-up policies—those policies above the minimal, fully subsidized catastrophic coverage available to growers) typically showed very small, and statistically insignificant, movements in response to subsidy changes. Further, these results were often negative, suggesting the potential for cross-product substitution effects. Together, these findings suggest that subsidies may not (during this time frame) have drawn many new acres into crop insurance, but they appear to have induced higher levels of coverage on enrolled acres.

According to E. O'Donoghue (2014), across all regions, for a 1-percent increase in subsidies, demand for crop insurance rose about 0.86 percent for corn, 0.74 percent for soybeans, and 0.64 percent for wheat in terms of total premiums. For liabilities, increases in demand ranged from roughly 0.2 percent for soybeans and wheat to 0.3 percent for corn. For acres enrolled, the rise in subsidies typically resulted in statistically insignificant changes in demand with magnitudes below 0.2 percent.

I used county-level data in this research. This means I could not control for the idiosyncratic behavior of individuals. So care must be taken in interpreting results, as what holds at the aggregate (county) level may mask considerable variation at the individual level. Furthermore, the analysis was undertaken one crop at a time and only with corn and soybeans, assuming that each enterprise on a farm is operated independently of the next. If cross-product substitution effects exist that are not taken into account here, results may change.

10 Economic Model

Environment and choice. I consider a farmer indexed by county c, crop i, and year t, who chooses insurance coverage intensity D_{ict}, which may be interpreted as the fraction of planted area insured, the coverage level, or any monotone index of insurance coverage.

Let random profit without insurance be

$$\pi^0_{ict} = p_{ict} y_{ict} - k_{ict},$$

where output price is p_{ict}, production cost is k_{ict}, and yield y_{ict} depends on weather shocks ω_{ict}, such as degree days and precipitation realizations.

With insurance coverage D_{ict}, profit becomes

$$\pi_{ict}(D) = \pi^0_{ict} + I_{ict}(D, \omega_{ict}) - P_{ict}(D),$$

where $I_{ict}(\cdot)$ denotes state-contingent indemnity payments and $P_{ict}(D)$ is the insurance premium.

Participation frictions and instruments. In addition to the premium, I assume that farmers incur non-price participation frictions associated with insurance use. These frictions take the form

$$\kappa_{ict}(D; Z_c) = D \cdot \tau(Z_c),$$

where Z_c is a vector of predetermined county-level shifters. Specifically,

$$Z_{1c} = \text{Distance}_c \quad \text{and} \quad Z_{2c} = \text{Stations}_c,$$

corresponding to distance to the nearest USDA service office and the number of weather stations in the county in 1989, respectively.

Distance to USDA offices captures administrative, informational, and transaction costs associated with enrollment and servicing, while historical weather-station density proxies for local information infrastructure and administrative capacity. The sign of $\tau(Z_c)$ is ultimately empirical.

Net profit can therefore be written as

$$\tilde{\pi}_{ict}(D) = \pi^0_{ict} + I_{ict}(D, \omega_{ict}) - P_{ict}(D) - D\tau(Z_c).$$

Expected-utility maximization. Farmers choose insurance coverage to maximize expected utility,

$$\max_{D \in [0,1]} \mathbb{E}[U(\tilde{\pi}_{ict}(D)) \mid \mathcal{I}_{ict}],$$

where $\mathcal{I}_{ict}$ denotes the information set available at the insurance sign-up date, including predetermined climate risk, historical weather variability, and prices.

First-order condition. For an interior solution, the first-order condition is

$$\mathbb{E}\left[U'(\tilde{\pi}_{ict}(D)) \left(\frac{\partial I_{ict}}{\partial D} - \frac{\partial P_{ict}}{\partial D} - \tau(Z_c)\right)\right] = 0.$$

Define the risk-adjusted marginal benefit and marginal cost of insurance as

$$MB_{ict}(D) \equiv \mathbb{E}\left[U'(\tilde{\pi}_{ict}(D)) \frac{\partial I_{ict}}{\partial D}\right], \qquad MC_{ict}(D) \equiv \mathbb{E}\left[U'(\tilde{\pi}_{ict}(D)) \frac{\partial P_{ict}}{\partial D}\right],$$

and define the marginal participation friction as

$$FC_{ict}(D; Z_c) \equiv \mathbb{E}[U'(\tilde{\pi}_{ict}(D))]\, \tau(Z_c).$$

The optimal insurance condition can then be written compactly as

$$MB_{ict}(D) - MC_{ict}(D) - FC_{ict}(D; Z_c) = 0.$$

This expression highlights that optimal coverage equates the marginal, risk-adjusted benefits of insurance to the sum of marginal premium costs and marginal participation frictions.

Linear approximation and empirical first stage. To connect the theoretical model to the empirical specification, I take a first-order approximation of the optimality condition around a typical coverage level D^*. I assume that (i) marginal benefits and costs depend approximately linearly on observable risk and price shifters, and (ii) participation frictions enter linearly through $\tau(Z_c)$.

This yields the reduced-form approximation

$$D_{ict} = \pi_1 Z_{1c} + \pi_2 Z_{2c} + \gamma' W_{ict} + \alpha_s + \delta_t + u_{ict},$$

where W_{ict} includes growing-degree-day measures, precipitation moments, and output prices, while α_s and δ_t denote state and year fixed effects.

Economic interpretation. The coefficient π_1 captures the semi-elastic response of insurance coverage to administrative and transaction costs associated with distance to USDA offices. The coefficient π_2 reflects the effect of local information and administrative capacity, as proxied by historical weather-station density. The vector γ captures how baseline climate risk and price incentives shape insurance demand. This formulation provides the economic foundation for the first-stage regressions used in the empirical analysis.

11 Discussion

The results of both the sublia and fracInsured specifications highlight a consistent trade-off in the design of the Federal Crop Insurance Program (FCIP). Subsidies and insurance coverage reduce yield losses under moderate stress of heat and normal rainfall, but increase vulnerability under extreme conditions, particularly heat. This asymmetric effect reflects the weather trends in our study period: as shown in the stylized facts, the Midwest has experienced a rising number of moderate days that are relatively more conducive to soybean production, but also more frequent extreme heat events that disproportionately affect corn yields. In this context, our findings suggest that insurance subsidies amplify short-term protection under moderate conditions, while leaving farmers more exposed to yield losses when weather crosses damaging thresholds.

The crop-specific results reinforce this pattern. Corn, which has higher yields but also greater yield variability (Figure 6), benefits more in the short run from insurance support under favorable weather but is especially vulnerable to extreme heat, consistent with its physiological sensitivity. Soybeans, by contrast, face relatively fewer extreme hot days and have lower variance, meaning that insurance dampens incentives for adaptation but in a more muted way. These dynamics mirror farmers' planting choices: despite relative price signals that should have favored soybeans (Figure 4), acreage decisions did not fully adjust, likely because insurance reduced the need to reallocate risk through crop switching.

Taking together, these results contribute to a growing literature (Annan and Schlenker (2015); Jisang Yu, A. Smith, and Sumner (2018); Chemeris, Liu, and Alan P. Ker (2022)) showing that while crop insurance stabilizes income and smooths consumption in the short run, it may undermine long-run resilience to climate change. The policy implication is that subsidy design must better account for heterogeneous responses across the climate distribution and across crops: without adjustment, crop insurance may inadvertently amplify the weather risks that it was designed to mitigate.

12 Limitation

In this paper, we did not consider the spatial effect of subsidies on acreage. As Barry K. Goodwin and Vincent H. Smith (2013) has shown, the geographic distributions of participation and relative returns to insurance are varied across the geography in the US. To the extent that the benefits of subsidized crop insurance are heterogeneous across regions and crops, distortions in production patterns may follow. Participation (in terms of total liability) is particularly high in the Corn Belt, the Mississippi Delta, and in the Upper Great Plains. At the same time, the relative returns to insurance

as represented by the ratio of indemnities to subsidy-adjusted premiums, are lower in the Corn Belt but appear to be particularly high in Western Texas, the mid-south, and in certain portions of the Northern Plains. This suggests the potential for production distortions as certain crops and growing areas receive significantly higher subsidies. Next work could address this issue.

On a separate note, this paper took a "positive" approach to study the impact of subsidies on crop acreage, taking the existing policies as given. An alternative "normative" approach would be needed to study whether the presence of an insurance subsidy is an optimal welfare-improving policy, and if yes, what would the optimal level of subsidy be. Section 3 in the Handbook chapter of Chetty and Finkelstein (2013) could be a good starting point for answering this question.

In calculating the optimal level of insurance subsidy, I did not model the choice of farmers over percentage coverage of insurance, i.e. yield or revenue coverage that can take values of 55%, 60%, ..., 90%. A follow up research can take that dimension into account. Adjesy the three limitations below for this paper. These limitations are from the 2006 paper of Raj Chetty, "A general formula for optimal level of social insurance".
Fiscal externalities. I assumed that the only distortionary tax/subsidy in the economy is the UI benefit. In practice, many behaviors are taxed, and these taxes create bfiscal externalitiesQ that could change the formula for the optimal benefit rate. For example, higher UI benefits will in general lower private savings, which could in turn reduce tax collections from capital gains or dividends. This tax revenue effect is not internalized by the agent and therefore affects the optimal benefit rate directly. It would be useful to determine the magnitude of such fiscal externalities too assess whether they affect the calculation of the optimal benefit rate significantly.

General equilibrium effects: In the models analyzed here, all behavioral responses to UI were solely determined by the agent. This assumption was important because the envelope conditions used to obtain the formula for optimal benefits relied on the idea that all endogenous variables in the model are chosen to maximize the agent's utility. Obtaining a reduced-form formula that takes equilibrium responses by firms into account would be useful.

Myopic agents. The envelope arguments above rely on the assumption that agents are optimizing. If agents experience large consumption drops during unemployment because they are myopic and do not save enough, the formula for the optimal benefit level may be different.

I used this Annan and Schlenker (2015) data set as the basis for building the main data set of analysis, and only focused on counties, crops and years covered in (Annan and Schlenker 2015). With this choice, I eliminated the step for calculating the number of moderate degree days, number of heating degree days and temperature for other counties, crops and years. A follow-up research can expand on the scope of work on either of these three dimensions.

13 Conclusion

I have shown that aggregated annual rainfall and exposure to heat measures mask intra-seasonal heterogeneity in rainfall/heat exposure effects and attenuate estimated impacts of rainfall shocks on crop yield First line of conclusion of Farris et al. (2023).

This study investigated how federal crop insurance subsidies influence the sensitivity of US crop

yields to weather extremes. Across specifications, two consistent patterns emerge. First, insurance subsidies reduce yields on average, a result robust in the baseline models where subsidies per dollar of liability (sublia) and the fraction of acreage insured (fracInsured) are treated as endogenous. Focusing on the number of heating degree days as the most important dimension of extreme weather, we observed that the effect of sublia on yield is negative across the distribution of hot degree days. The impact was small and mostly insignificant at lower quantiles, but it becomes larger and statistically significant at higher heat levels, reaching yield losses of about 7–9 percent per additional percentage point of adoption. At the 99th, 95th, 90th, 75th percentiles and average of hot degree days, a one percentage point increase in sublia decrease yield by 8.6%, 7.8%, 7.4%, 7% and 6.8%. This negative relationship supports the hypothesis that subsidies can discourage adaptive investment and shift cultivation toward more marginal land, thereby lowering resilience.

Crop-specific analyses reinforce these findings. For corn, the evidence is especially strong: subsidies magnify vulnerability to heat and rainfall extremes, with robust first-stage instruments supporting the reliability of these results. For soybeans, the effects are weaker and less systematic, with subsidy–weather interactions often statistically insignificant. This divergence highlights crop-specific agronomic differences in sensitivity to weather stress and the role of insurance. Taken together, the results indicate that while crop insurance reduces income volatility in the short run, it may inadvertently increase sensitivity of yield to weather extremes, particularly for corn, by altering both adaptive incentives and land-use decisions. From a policy perspective, these findings underscore the need to balance the stabilizing function of crop insurance with incentives for climate adaptation.

Designing subsidies that encourage resilience-enhancing practices such as adoption of drought-tolerant varieties, conservation tillage, and improved water management, could mitigate the disincentive effects documented here and better align federal crop insurance with the realities of a changing climate.

14 Acknowledgement

The author reports no conflicts of interest in this research. She received funding from Indiana University Bloomington for presenting an earlier version of this paper at the 89th *Midwest Economics Association Conference* in 2025.

15 Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

16 References

References

Abrigo, Michael RM and Inessa Love (2016). "Estimation of panel vector autoregression in Stata". In: *The Stata Journal* 16.3, pp. 778–804.

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, Donald WK and Biao Lu (2001). "Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models". In: *Journal of econometrics* 101.1, pp. 123–164.

Angel, Jonathan et al. (2018). "Midwest". In: *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II*. Ed. by David R. Reidmiller et al. Washington, DC, USA: U.S. Global Change Research Program, pp. 872–940.

Angrist, Joshua D and Jörn-Steffen Pischke (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

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.

Anselin, Luc (2024a). *Applied Spatial Regression Analysis - Lecture: Robust Tests*. YouTube, viewed *minute 00:00*. URL: https://youtu.be/nWk2eBPEsR4?si=scT1to8JxGaElejQ.

— (2024b). *Applied Spatial Regression Analysis - Lecture: Robust Tests*. YouTube, viewed *minute 05:20*. URL: https://youtu.be/nWk2eBPEsR4?si=E8e-BJWACJlci0NG.

— (2024c). *Applied Spatial Regression Analysis - Lecture: Spatial Regression Specifications*. YouTube, viewed *minute 2:20*. URL: https://www.youtube.com/watch?v=yMwQ4MljEWc&list=PLzREt6r1NenlbfMWgHZ9aRpI

— (2024d). *Applied Spatial Regression Analysis - Lecture: Spatial Spillovers*. YouTube, viewed *minute 14:30*. URL: https://youtu.be/C-sBZ9XWC9g?si=iwqPTRCNpBQYddtg&t=876.

— (2024e). *Applied Spatial Regression Analysis - Lecture: Specification Tests*. YouTube, viewed *minute 35:46*. URL: https://youtu.be/XyIokAi01Ko?si=RCYHrXi_zI3ouBkS&t=2146.

Anselin, Luc and Anil K Bera (1998). "Spatial dependence in linear regression models with an introduction to spatial econometrics". In: *Statistics textbooks and monographs* 155, pp. 237–290.

Arellano, Manuel (1987). "Computing robust standard errors for within-groups estimators." In: *Oxford Bulletin of Economics & Statistics* 49.4.

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

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.

Balli, Hatice Ozer and Bent E. Sørensen (2013). "Interaction Effects in Econometrics". In: *Empirical Economics* 45.1. Earlier draft (2010) available at https://www.uh.edu/~bsorense/Ozer-Balli_Sorensen.pdf, pp. 583–603. DOI: 10.1007/s00181-012-0604-2.

Baquie, Sandra et al. (2025). *Industrial Policies: Handle with Care*. International Monetary Fund.

Bellemare, Marc F and Casey J Wichman (2020). "Elasticities and the inverse hyperbolic sine transformation". In: *Oxford Bulletin of Economics and Statistics* 82.1, pp. 50–61.

Bera, Anil K, Osman Doğan, et al. (2020). "Specification tests for spatial panel data models". In: *Journal of Spatial Econometrics* 1.1, p. 3.

Bera, Anil K and Mann J Yoon (1993). "Specification testing with locally misspecified alternatives". In: *Econometric theory* 9.4, pp. 649–658.

Berry, Steven T and Philip A Haile (2021). "Foundations of demand estimation". In: *Handbook of industrial organization*. Vol. 4. 1. Elsevier, pp. 1–62.

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.

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 (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#eyJhcHBpZCI6MTksInN0ZXBzIjpbMSwyLDMsM10sImRhdGEiOltbImNhdGVnb3JpZXMiLCJTdXJ2ZXkiX
=.

Byrnes, James P, David C Miller, and William D Schafer (1999). "Gender differences in risk taking: A meta-analysis." In: *Psychological bulletin* 125.3, p. 367.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2011). "Robust inference with multiway clustering". In: *Journal of Business & Economic Statistics* 29.2, pp. 238–249.

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.

— (2009). "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods". In: *Annu. Rev. Econ.* 1.1, pp. 451–488.

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.

Chung, Wonho (2014). "Analysis of social welfare effects of crop revenue insurance". In: *Journal of Rural Development/Nongchon-Gyeongje* 37.4, pp. 75–95.

Claassen, Roger, Christian Langpap, and JunJie Wu (2017). "Impacts of federal crop insurance on land use and environmental quality". In: *American Journal of Agricultural Economics* 99.3, pp. 592–613.

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.

Deschênes, Olivier and Michael Greenstone (2007). "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather". In: *American economic review* 97.1, pp. 354–385.

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).

Diekmann, Florian et al. (2012). "Judging farmers' willingness to trade distance and taxes for extension services". In: *Applied economic perspectives and policy* 34.3, pp. 454–471.

Dijkstra, Edsger W. (1959). "A note on two problems in connexion with graphs". In: *Numerische Mathematik* 1, pp. 269–271. DOI: 10.1007/BF01386390.

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

Falco, Salvatore Di et al. (2014). "Crop insurance as a strategy for adapting to climate change". In: *Journal of Agricultural Economics* 65.2, pp. 485–504.

Farris, Jarrad et al. (2023). "Growing pains: Timing of in Utero rainfall shocks and child growth in rural Rwanda". In: *Economic Development and Cultural Change* 71.2, pp. 793–818.

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.

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.

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 and Alan P Ker (1998). "Nonparametric estimation of crop yield distributions: implications for rating group-risk crop insurance contracts". In: *American journal of agricultural economics* 80.1, pp. 139–153.

Goodwin, Barry K. and Vincent H. Smith (2013). "What Harm Is Done by Subsidizing Crop Insurance?" In: *American Journal of Agricultural Economics* 95.2, pp. 489–497.

Granger, Clive W. J. and Paul Newbold (1974). "Spurious Regressions in Econometrics". In: *Journal of Econometrics* 2.2, pp. 111–120.

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).

Halek, Martin and Joseph G Eisenhauer (2001). "Demography of risk aversion". In: *Journal of Risk and Insurance*, pp. 1–24.

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

Hansen, Lars Peter (1982). "Large sample properties of generalized method of moments estimators". In: *Econometrica: Journal of the econometric society*, pp. 1029–1054.

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.

Jefferson, Thomas (Aug. 23, 1785). *Letter to John Jay*. The Avalon Project. URL: https://avalon.law.yale.edu/18th_century/let32.asp (visited on 12/15/2025).

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.

Kleven, Henrik J (2021). "Sufficient statistics revisited". In: *Annual Review of Economics* 13.1, pp. 515–538.

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

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).

Lashkaripour, Ahmad (2021). "The cost of a global tariff war: A sufficient statistics approach". In: *Journal of International Economics* 131, p. 103419.

Lashkaripour, Ahmad and Volodymyr Lugovskyy (2023). "Profits, scale economies, and the gains from trade and industrial policy". In: *American Economic Review* 113.10, pp. 2759–2808.

LeSage, James and Robert Kelley Pace (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

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).

Licht, Mark and Sotirios Archontoulis (July 21, 2017). *Influence of Drought on Corn and Soybean*. Integrated Crop Management News, Iowa State University Extension and Outreach. URL: https://crops.extension.iastate.edu/cropnews/2017/07/influence-drought-corn-and-soybean.

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

Maisashvili, Aleksandre, Henry L Bryant, and Jason PH Jones (2020). "Implications of alternative crop insurance subsidies". In: *Journal of Agricultural and Applied Economics* 52.2, pp. 240–263.

Maisashvili, Aleksandre, Bart Fischer, and Henry Bryant (2023). "Crop insurance implications of permanently authorizing the emergency relief program". In: *Journal of Agricultural and Applied Economics* 55.1, pp. 171–193.

Manski, Charles F (1993). "Identification of endogenous social effects: The reflection problem". In: *The review of economic studies* 60.3, pp. 531–542.

Meena, A et al. (2024). "Spatial Econometrics in Agriculture: Modelling Spatial Dependencies in Data". In: *Archives of Current Research International* 24.6, pp. 488–506.

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.

Moran, Patrick AP (1950). "Notes on continuous stochastic phenomena". In: *Biometrika* 37.1/2, pp. 17–23.

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.

Nickell, Stephen (1981). "Biases in dynamic models with fixed effects". In: *Econometrica: Journal of the econometric society*, pp. 1417–1426.

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.

— (2026). *nClimDiv County Palmer Drought Severity Index (PDSI), Version 1.0.0*. https://www.ncei.noaa.gov/pub/data/cirs/climdiv/climdiv-pdsicy-v1.0.0-20260305. County-level monthly PDSI file from the nClimDiv bulk archive; accessed 2026-03-17.

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].

Papke, Leslie E and Jeffrey M Wooldridge (1996). "Econometric methods for fractional response variables with an application to 401 (k) plan participation rates". In: *Journal of applied econometrics* 11.6, pp. 619–632.

Paredes, Alexandre et al. (2022). "Does R&D tax credit impact firm behaviour? Micro evidence for Portugal". In: *Research Evaluation* 31.2, pp. 226–235.

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).

Pawlowski, Boguslaw, Rajinder Atwal, and Robin IM Dunbar (2008). "Sex differences in everyday risk-taking behavior in humans". In: *Evolutionary Psychology* 6.1, p. 147470490800600104.

Quiggin, John (1994). "The optimal design of crop insurance". In: *Economics of Agricultural Crop Insurance: Theory and Evidence*. Springer, pp. 115–134.

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, Erik J O'Donoghue, and Nigel D Key (2007). "Does Crop Insurance Affect Crop Yields?" In.

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.

Sadoulet, Elisabeth, Alain De Janvry, and Benjamin Davis (2001). "Cash transfer programs with income multipliers: PROCAMPO in Mexico". In: *World development* 29.6, pp. 1043–1056.

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.

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.

Staiger, Douglas and James H Stock (1997). "Instrumental Variables Regression with Weak Instruments". In: *Econometrica* 65.3, pp. 557–586.

Stock, James H and Motohiro Yogo (2002). *Testing for weak instruments in linear IV regression*.

Tibshirani, Robert (1996). "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 58.1, pp. 267–288.

Tsiboe, Francis and Dylan Turner (2023a). "Econometric identification of crop insurance participation". In: *Agricultural and Resource Economics Review* 52.3, pp. 476–497.

— (2023b). "The Crop Insurance Demand Response to Premium Subsidies: Evidence from US Agriculture". In: *Food Policy*.

U.S. Department of Agriculture, Economic Research Service (2026a). *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, 2026b). *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.

— (2026c). *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.

— (2026d). *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, Risk Management Agency (2025). *RMA Information Reporting System (RIRS)*. Accessed January 8, 2026. U.S. Department of Agriculture. URL: https://webapp.rma.usda.gov/apps/RIRS/.

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. Department of Labor, Bureau of Labor Statistics (2016). *Producer Price Index*. U.S. Department of Labor, Bureau of Labor Statistics. Accessed: September 21, 2025. URL: https://www.bls.gov/ppi/detailed-report/.

Ugarte-Ruiz, Alfonso (July 2025a). *Locproj & Lpgraph: Stata Commands to Estimate Local Projections*. Working Paper WP 25-09. Accessed October 28, 2025. BBVA Research. URL: https://www.bbvaresearch.com/en/publications/locproj-lpgraph-stata-commands-to-estimate-local-projections/.

— (2025b). "LOCPROJ: Stata module to estimate local projections". In.

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.

— (June 2025). *RMA Volatility Factor Calculation Methodology*. Tech. rep. Accessed: 2025-11-19. Risk Management Agency (USDA). URL: https://www.rma.usda.gov/sites/default/files/2025-06/volatilitymethodology.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.

Walters, Cory and Richard Preston (2018). "Net income risk, crop insurance and hedging". In: *Agricultural Finance Review* 78.1, pp. 135–151.

Williams, Richard (2012). "Using the margins command to estimate and interpret adjusted predictions and marginal effects". In: *The Stata Journal* 12.2, pp. 308–331.

Woodard, Joshua D and Jing Yi (2020). "Estimation of insurance deductible demand under endogenous premium rates". In: *Journal of Risk and Insurance* 87.2, pp. 477–500.

Wooldridge, Jeffrey M (2015). "Control function methods in applied econometrics". In: *Journal of Human Resources* 50.2, pp. 420–445.

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.

Yi, Fujin, Wuyi Lu, and Yingheng Zhou (2016). "Cash transfers and multiplier effect: Lessons from the grain subsidy program in China". In: *China Agricultural Economic Review* 8.1, pp. 81–99.

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.

Zhong, Ling et al. (2023). "Optimal design of agricultural insurance subsidies under the risk of extreme weather". In: *International Journal of Production Economics* 263, p. 108920.

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, we often referred to technical expressions in the context of agricultural insurance. To be on the same page, we provide an explanation for the major terms from a Handbook of the United States Department of Agriculture, Risk Management Agency (2008)[16].

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.

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

- 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:

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

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

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 Causality

We are interested in the causal relationship between enrollment of farmers in FCIC and their crop yield. For simplicity, let's assume enrollment in insurance is a dummy variable, and we denote it with C_i (see also Angrist and Pischke 2009, sec. 3.2). To address this question, we imagine two potential yield outcomes:

[16] See page 50.

$$\text{Potential Outcome} = \begin{cases} Y_{1i} & \text{if } C_i = 1, \\ Y_{0i} & \text{if } C_i = 0 \end{cases}$$

In this case, $Y_1 i$ and $Y_0 i$ are the farmer i's yield when insured and not insured, respectively. We would like to know the difference between $Y_1 i$ and $Y_0 i$; which is the causal effect of insurance enrollment for farmer i. The observed outcome, Y_i, can be written in terms of potential outcomes as

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})C_i$$

We aim to measure $Y_{1i} - Y_{0i}$ for those who purchased insurance. But the challenge is that we get to see one of Y_{1i} and Y_{0i}, but never both. To simplify the analysis, we assume that the conditional independence assumption (CIA) holds (see also Angrist and Pischke 2009, p. 53). Let us call all other covariates X_i. The CIA would assert that conditional on observed characteristics, X_i, selection bias for buying insurance disappears. Formally, this means

$$\{Y_{0i}, Y_{1i}\} \perp\!\!\!\perp C_i \mid X_i$$

Or alternatively, it is equivalent to

$$E[Y_i \mid X_i, C_i = 1] - E[Y_i \mid X_i, C_i = 0] = E[Y_{1i} - Y_{0i} \mid X_i]$$

Next, let us consider the insurance coverage as a continuous variable. The causal relationship between the level of coverage and the crop yield is likely to be different for each farmer (or county). We therefore use individual-specific functional notation,

$$Y_{\theta i} \equiv f_i(\theta),$$

to denote the potential yield that farmer i would obtain at coverage level θ. $f_i(\theta)$ answers the causal "what if" questions. The CIA in this more general setup becomes

$$\{Y_{\theta i}\} \perp\!\!\!\perp \theta_i \mid X_i \qquad \text{for all } \theta$$

In an observational study like ours, the CIA means that θ_i is "as good as randomly assigned", conditional on X_i. Conditional on X_i, the average causal effect of one extra percentage point of insurance coverage is $E[f_i(\theta) - f_i(\theta - 1) \mid X_i]$, while the average causal effect of 5 percentage (a typical coverage incremental step in the federal crop insurance program) increase in insurance coverage is $E[f_i(\theta) - f_i(\theta - 5) \mid X_i]$. The data reveals only $Y_i \equiv f_i(\theta_i)$, that is, $f_i(\theta)$ for $\theta = \theta_i$. But given the CIA, conditional on X_i, comparisons of average yield with different levels of insurance coverage have a causal interpretation. In other words, for any value of θ,

$$E[Y_i \mid X_i, \theta_i = \theta] - E[Y_i \mid X_i, \theta_i = \theta - 1] = E[f_i(\theta) - f_i(\theta - 1) \mid X_i] \tag{B.1}$$

Using the law of iterated expectations, the unconditional (overall) average causal effect would be

$$\begin{aligned} E\Big[\, E(Y_i \mid X_i, \theta_i = \theta) - E(Y_i \mid X_i, \theta_i = \theta - 1) \,\Big] &= E\Big[\, E\big(f_i(\theta) - f_i(\theta - 1) \,\big|\, X_i\big) \Big] \\ &= E[f_i(\theta) - f_i(\theta - 1)] \end{aligned} \tag{B.2}$$

C Omitted Variable Bias

A riskier crop or a riskier county has a higher premium rate for a given crop insurance choice (Coble and Barnett 2013). As a result, farmers tend to plant less risky crops and plant less in riskier counties, ceteris paribus. Riskiness is positively correlated with Γ_ijt and negatively correlated with Y_{ijt}. So, omitting the "riskiness" variable that has a positive correlation with Γ_ijt, would cause a downward bias in OLS estimate for coefficient of Γ_ijt. To verify this, let's define the long and short regressions as discussed by Angrist and Pischke (2009, pp. 59–60) as follows:

$$\textit{Long regression:} \quad \ln Y_{ijt} = \alpha_0 + \alpha_1 \Gamma_{ijt} + \alpha_2 \Omega_{ijt} + \epsilon_{ijt}$$

In which Ω denotes "riskiness". Absent the riskiness, the regression would change to

$$\textit{Short regression:} \quad \ln Y_{ijt} = \tilde{\alpha}_0 + \tilde{\alpha}_1 \Gamma_{ijt} + \eta_{ijt}$$

Then the omitted variable bias formula Angrist and Pischke (2009, p. 59) would specify the relationship between α_1, $\tilde{\alpha}_1$ and α_2 as

$$\tilde{\alpha}_1 = \alpha_1 + \alpha_2 \, \delta_{\Gamma,\text{riskiness}} = \alpha_1 + \alpha_2 \underbrace{Cov(\Gamma_{ijt}, \Omega_{ijt})}_{\textit{omitted variable bias}}$$

α_2 is the effect of the omitted control (i.e. riskiness) on the outcome, and $\delta_{\Gamma,\Omega}$ is the correlation of those omitted controls with the variables of interest. Having $Cov(\Gamma_{ijt}, \Omega_{ijt}) > 0$ and $\alpha_2 < 0$, would yield

$$\text{Omitted variable bias} = \alpha_2 \, Cov(\Gamma_{ijt}, \Omega_{ijt}) < 0$$

Therefore, the omitted variable bias is downward. This completes the proof. [0.05in]

A second source of endogeneity is the dependence of Γ_ijt on the choice of coverage level. Although the sign of relationship is not determined. For the sake of brevity, let's focus on the sublia as a measure of intensity of insurance enrollment. Following Jisang Yu, A. Smith, and Sumner (2018), we write the premium subsidy per dollar of insured liability as $\Gamma_j(\theta_j) = s(\theta_j)p_j(\theta_j)$, where $s(\theta_j)$ is the subsidy rate and $p_j(\theta_j)$ is the premium. Then we will have:

- An increase in the chosen coverage level for a crop may cause the subsidy rate to decrease because the subsidy rate is lower for the higher coverage levels ($s'(\theta_j) < 0$),
- The premium rate increases with the coverage level ($p'_j(\theta_j) > 0$).

So, the sign of $\Gamma'_j(\theta_j)$ is unknown. To determine is unknown. To determine the sign of omitted variable bias from absence of coverage level in the model, we set up the long and short regressions as follows:

$$\textit{Long regression:} \quad \ln Y_{ijt} = \alpha_0 + \alpha_1 \Gamma_{ijt} + \beta_\theta \theta_j + \epsilon$$

Absent the riskiness, the regression would change to

$$\textit{Short regression:} \quad \ln Y_{ijt} = \tilde{\alpha}_0 + \tilde{\alpha}_1 \Gamma_{ijt} + u$$

Then the omitted variable bias formula would specify the relationship among α_1, $\tilde{\alpha}_1$ and β_θ as

$$\tilde{\alpha}_1 = \alpha_1 + \beta_2 \underbrace{Cov(\Gamma_{ijt}, \theta_j)}_{omitted\ variable\ bias}$$

We know $\beta_\theta > 0$. Then we would have

$$if \begin{cases} Cov(\Gamma_{ijt}, \theta_j) > 0 & \rightarrow \quad OVB > 0, \\ Cov(\Gamma_{ijt}, \theta_j) < 0 & \rightarrow \quad OVB < 0 \end{cases}$$

So, if Γ and θ are positively (negatively) correlated, and θ has a positive impact on crop acreage, the OLS would underestimate (overestimate) the impact of Γ on acreage.

D Marginal Effects Tables

Tables D1-D8 are the bases for Figures 6-13, respectively.

Table D1 – Marginal change in sublia at different values of number of heating degree days, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	−6.350 519	3.495 379	−1.82	0.069	−13.201 34	0.500 298 7
2	−6.369 202	3.496 005	−1.82	0.068	−13.221 25	0.482 842 8
3	−6.392 519	3.496 811	−1.83	0.068	−13.246 14	0.461 105 1
4	−6.478 521	3.500 023	−1.85	0.064	−13.338 44	0.381 398 2
5	−6.829 677	3.517 018	−1.94	0.052	−13.722 91	0.063 551 7
6	−7.011 729	3.528 262	−1.99	0.047	−13.926 99	−0.096 462 6
7	−7.484 423	3.565 075	−2.10	0.036	−14.471 84	−0.497 004 8
8	−7.847 696	3.600 672	−2.18	0.029	−14.904 88	−0.790 508 5
9	−8.636 591	3.698 866	−2.33	0.020	−15.886 24	−1.386 947

Table D2 – Marginal change in sublia at different values of number of moderate degree days, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	−0.835 372	2.507 928	−0.33	0.739	−5.750 82	4.080 076
2	−0.094 631 3	2.429 246	−0.04	0.969	−4.855 867	4.666 604
3	0.374 620 7	2.388 565	0.16	0.875	−4.306 881	5.056 122
4	1.178 171	2.336 635	0.50	0.614	−3.401 55	5.757 893
5	2.181 952	2.305 083	0.95	0.344	−2.335 929	6.699 832
6	3.176 285	2.311 63	1.37	0.169	−1.354 425	7.706 996
7	4.065 262	2.349 121	1.73	0.084	−0.538 930 6	8.669 455
8	4.583 149	2.384 209	1.92	0.055	−0.089 814 4	9.256 113
9	5.369 331	2.455 002	2.19	0.029	0.557 615 9	10.181 05

Table D3 – Marginal change in sublia at different values of precipitation, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	−12.346 94	2.612 372	−4.73	0.000	−17.467 10	−7.226 789
2	−13.348 20	2.545 514	−5.24	0.000	−18.337 32	−8.359 090
3	−13.936 16	2.520 123	−5.53	0.000	−18.875 51	−8.996 813
4	−15.150 67	2.501 683	−6.06	0.000	−20.053 88	−10.247 460
5	−16.891 88	2.555 890	−6.61	0.000	−21.901 34	−11.882 430
6	−18.317 22	2.667 181	−6.87	0.000	−23.544 80	−13.089 640
7	−20.038 92	2.871 306	−6.98	0.000	−25.666 58	−14.411 260
8	−21.309 33	3.062 539	−6.96	0.000	−27.311 80	−15.306 860
9	−24.912 59	3.739 803	−6.66	0.000	−32.242 47	−17.582 710

Table D4 – Marginal change in sublia at different values of precipitation squared, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	−4.204 523	3.788 029	−1.11	0.267	−11.628 92	3.219 877
2	−3.433 429	3.900 081	−0.88	0.379	−11.077 45	4.210 59
3	−2.920 682	3.976 266	−0.73	0.463	−10.714 02	4.872 657
4	−1.739 376	4.156 425	−0.42	0.676	−9.885 82	6.407 068
5	0.631 242 3	4.534 779	0.14	0.889	−8.256 762	9.519 246
6	2.161 364	4.788 885	0.45	0.652	−7.224 678	11.547 41
7	4.784 996	5.238 917	0.91	0.361	−5.483 093	15.053 08
8	6.938 978	5.619 31	1.23	0.217	−4.074 666	17.952 62
9	14.142 65	6.940 934	2.04	0.042	0.538 665 8	27.746 63

Table D5 – Marginal change in fracInsured at different values of number of heating degree days, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	0.081 928 4	0.062 291 3	1.32	0.188	−0.040 160 2	0.204 017
2	0.082 351	0.062 430 7	1.32	0.187	−0.040 010 8	0.204 712 8
3	0.082 878 4	0.062 605 4	1.32	0.186	−0.039 825 9	0.205 582 8
4	0.084 823 8	0.063 257 4	1.34	0.180	−0.039 158 5	0.208 806
5	0.092 766 9	0.066 033 5	1.40	0.160	−0.036 656 4	0.222 190 1
6	0.096 884 8	0.067 539 5	1.43	0.151	−0.035 490 2	0.229 259 9
7	0.107 577 1	0.071 638 1	1.50	0.133	−0.032 831	0.247 985 2
8	0.115 794 3	0.074 950 7	1.54	0.122	−0.031 106 4	0.262 694 9
9	0.133 638 9	0.082 543 3	1.62	0.105	−0.028 143	0.295 420 9

Table D6 – Marginal change in fracInsured at different values of number of moderate degree days, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	−0.021 019 6	0.039 574 9	−0.53	0.595	−0.098 584 9	0.056 545 7
2	−0.034 843 2	0.041 758 0	−0.83	0.404	−0.116 687 4	0.047 001 1
3	−0.043 600 3	0.043 772 9	−1.00	0.319	−0.129 393 5	0.042 193 0
4	−0.058 596 0	0.048 155 6	−1.22	0.224	−0.152 979 3	0.035 787 3
5	−0.077 328 3	0.054 895 1	−1.41	0.159	−0.184 920 8	0.030 264 1
6	−0.095 884 4	0.062 536 8	−1.53	0.125	−0.218 454 4	0.026 685 5
7	−0.112 474 3	0.069 925 8	−1.61	0.108	−0.249 526 4	0.024 577 7
8	−0.122 139 1	0.074 408 4	−1.64	0.101	−0.267 976 8	0.023 698 7
9	−0.136 810 6	0.081 405 3	−1.68	0.093	−0.296 362 1	0.022 740 9

Table D7 – Marginal change in fracInsured at different values of precipitation, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	0.133 148 4	0.067 510 5	1.97	0.049	0.000 830 3	0.265 466 5
2	0.141 715 9	0.070 562 6	2.01	0.045	0.003 415 7	0.280 016 2
3	0.146 746 9	0.072 577 7	2.02	0.043	0.004 497 4	0.288 996 5
4	0.157 139 2	0.077 196 3	2.04	0.042	0.005 837 2	0.308 441 1
5	0.172 038 3	0.084 715 1	2.03	0.042	0.005 999 8	0.338 076 8
6	0.184 234 6	0.091 493 3	2.01	0.044	0.004 910 9	0.363 558 2
7	0.198 966 7	0.100 256 1	1.98	0.047	0.002 468 4	0.395 465 1
8	0.209 837 3	0.107 040 9	1.96	0.050	0.000 041 0	0.419 633 6
9	0.240 669 5	0.127 345 0	1.89	0.059	−0.008 922 2	0.490 261 2

Table D8 – Marginal change in fracInsured at different values of precipitation squared, taking all other variables in Equation 13 at their average levels.

Percentiles are as follows: 1: 1%, 2: 5%, 3: 10%, 4: 25%, 5: mean, 6: 75%, 7: 90%, 8: 95% and 9: 99%.

	Margin	std. err.	z	P>|z|	[95% conf. interval]	
_at					lower	upper
1	0.067 222 6	0.064 941 4	1.04	0.301	−0.060 060 1	0.194 505 4
2	0.061 974 9	0.066 199 8	0.94	0.349	−0.067 774 3	0.191 724 2
3	0.058 485 4	0.067 111 4	0.87	0.384	−0.073 055 6	0.190 026 4
4	0.050 446 0	0.069 440 4	0.73	0.468	−0.085 654 8	0.186 546 7
5	0.034 312 7	0.074 931 5	0.46	0.647	−0.112 550 4	0.181 175 7
6	0.023 899 4	0.078 971 2	0.30	0.762	−0.130 881 4	0.178 680 2
7	0.006 044 2	0.086 622 3	0.07	0.944	−0.163 732 4	0.175 820 7
8	−0.008 614 8	0.093 452 2	−0.09	0.927	−0.191 777 8	0.174 548 1
9	−0.057 639 6	0.118 698 3	−0.49	0.627	−0.290 284 1	0.175 004 8

E Crop Insurance Subsidies in the World

sectionCrop 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.