

RESEARCH ARTICLE

Econometric identification of crop insurance participation

Francis Tsiboe  and Dylan Turner

Research Agricultural Economist, United States Department of Agriculture, Economic Research Service, Kansas City, MO, USA

Corresponding author: Francis Tsiboe, email: ftsiboe@hotmail.com

(Received 11 June 2022; revised 2 February 2023; accepted 3 March 2023; first published online 04 May 2023)

Abstract

This paper shows how econometric identification can be improved in studies making use of crop insurance participation as either an independent or dependent variable. The paper provides the reader with a succinct overview of how crop insurance contracts are priced and how to use publicly available data to derive a novel composite crop insurance design parameter that emulates existing crop insurance rating parameters using a procedure that is based on current actuarial practices. The derived design parameter performs well at predicting historic crop insurance loss-cost ratios and satisfies the requirements for an instrumental variable for a variety of empirical applications related to crop insurance. Representative empirical examples are presented where it is shown that the proposed instrument has favorable two-staged least squares diagnostic tests and is effective at eliminating endogeneity bias.

Keywords: crop insurance; endogeneity; instruments; participation; premium rate

JEL codes: Q14; Q18; G22

Introduction

The agricultural insurance sector is currently a large and rapidly expanding component of producer-oriented governmental support programs (Mahul and Stutley 2010; Smith and Glauber 2012; Glauber 2013; Belasco 2020). This has prompted a large volume of research over a diverse set of outcomes such as input and land-use decisions (Young et al. 2001; Zhong et al. 2007; Fadhliani et al. 2019; He et al. 2020), environmental and conservation outcomes (Schoengold et al. 2015; Claassen et al. 2017), agricultural marketing (Du et al. 2015; Jablonski et al. 2022), farm financing (Ifft et al. 2015; Cariappa et al. 2020), the use of other risk mitigation strategies (O'Donoghue et al. 2009; Deryugina and Konar 2017; Turner and Tsiboe 2022), and the participation in other programs that comprise the farm safety net (Möllmann et al. 2019). One challenge that researchers in this domain face is that endogeneity exists in most empirical analysis involving measures of crop insurance

© USDA, Economic Research Service, 2023. This is a work of the US Government and is not subject to copyright protection within the United States. Published by Cambridge University Press on behalf of the Northeastern Agricultural and Resource Economics Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

demand. For models making use of crop insurance demand as a dependent variable (typically measured by total premiums paid, liability, acreage insured, or coverage levels), it will typically be appropriate to include the price of insurance (usually measured by the premium rate) as an independent variable. However, by actuarial construction, premium rates increase with the quantity of insurance purchased (as measured by coverage level) which at face value would suggest demand increases with price leading to an apparent violation of the law of demand (Woodard and Yi 2020).¹ This is the result of the quantity of insurance and the premium rates being simultaneously determined.

If instead, a model is specified using a measure of crop insurance demand as an independent variable, endogeneity is also likely a concern. This is because “risk” serves as a key component in the pricing of insurance but is not easily observable to the rate setter which leads to a case of omitted variable bias if the producer’s risk is correlated with the dependent variable of interest.² For example, if a model is specified to estimate the effect of crop insurance participation on other risk-mitigating behavior such as irrigation decisions or cover crop use, this form of endogeneity would likely be present since the producer’s unique risk profile would likely influence both their crop insurance decisions and which-ever risk-mitigating decision is specified as a dependent variable.

Regardless if crop insurance demand is specified as a dependent or independent variable, researchers may consider using instrumental variables to generate the requisite exogenous variation for credible econometric identification (Roberts et al. 2006; Walters et al. 2012; Falco et al. 2014; Deryugina and Konar 2017; Connor et al. 2021). However, finding a suitable instrument is often challenging. The instrument must be highly correlated with the price of crop insurance (i.e., premium rates) yet have no plausible causal link to the measured quantity of insurance purchased (when crop insurance demand is a dependent variable) or be highly correlated with the demand for insurance yet have no plausible causal link to the outcome of interest (when crop insurance demand is an independent variable). One strategy (for either empirical setup) is to use actuarial rating parameters as instruments since they govern premium rates in a way that is plausibly exogenous to any individual farmer’s crop insurance decisions (Woodard and Yi 2020; Tsiboe and Turner 2022). In turn, these actuarial rating parameters also exogenously influence crop insurance demand meaning they can be used as instrumental variables in models using a wide variety of potential outcomes of interest (such as other risk mitigation behavior).³

Fortunately, for research focused on the Federal Crop Insurance Program (FCIP), the USDA Risk Management Agency (RMA) publicly provides the necessary data for constructing these rating parameter-based instruments. However, deriving these rating parameters can be daunting without an intimate understanding of the underlying modeling approaches used for setting crop insurance premium rates. The purpose of this paper is to provide a concise overview of the actuarial design parameters used to set crop insurance premium rates and to provide guidance on how these parameters can be estimated from existing publicly available data to provide a set of instrumental variables that can be used to address the previously mentioned sources of endogeneity.

¹Measures of the quantity of crop insurance that exclusively capture the extensive margin (like enrolled acreage) do not correlate with premium rates in the same way but would still exhibit the same correlation with respect to the total premium. Thus, there is still a positive relationship between the quantity purchased and the total cost of insurance.

²It’s worth noting that this source of endogeneity is just as likely a problem in the case of crop insurance demand being modeled as a function of the price of insurance (Yu et al. 2018).

³This identification strategy is valid for any potential outcome of interest that is uncorrelated with the actuarial rating parameters set by RMA.

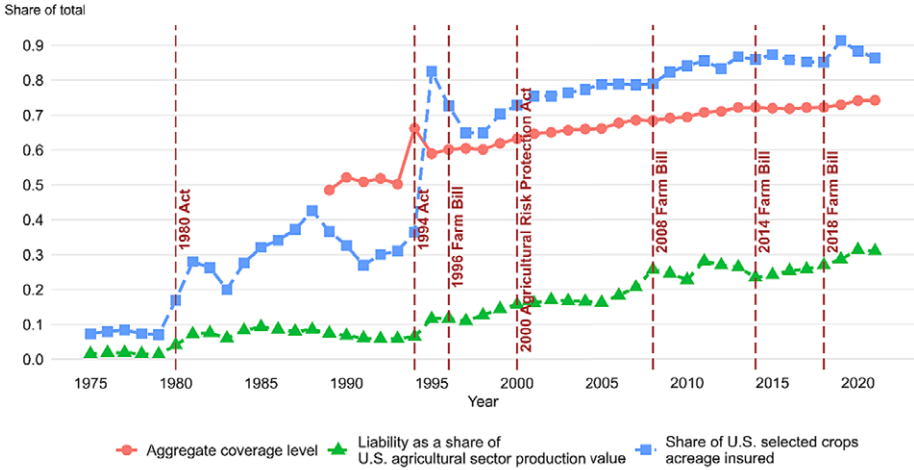
The rest of this paper is organized as follows. The next section provides an overview of major crop insurance instruments in the context of the US. Then, a conceptual overview of contracts is provided that shows how crop insurance rating parameters influence premium rates. Following the conceptual overview, a guide on how to approximate the actuarial rating parameters is presented before discussing the data sources necessary for utilizing the proposed methodology. After discussing necessary data sources, the ability of the estimated rating parameters to predict historic crop insurance outcomes is evaluated. The penultimate section provides empirical demonstrations of how the estimated rating parameters can be used as instrumental variables to mitigate endogeneity bias and improve econometric identification. The final section concludes.

Available instruments for crop insurance

Existing sources of exogenous variation for the identification of crop insurance outcomes in the context of the US fall under four broad categories. The first relies on national-level legislative events driving structural changes in FCIP (Schoengold et al. 2015; Connor and Katchova 2020; Wang et al. 2021). This first group of instruments proposed by Schoengold et al. (2015) takes on a value of unity for years within the respective policy periods and zero otherwise. The major changes in the national policy upon which these instruments are based include updates in The Crop Insurance Reform Act of 1994 (Public Law [P.L.] 103-354) (conventionally known as “1994 Act”), The Agricultural Risk Protection Act of 2000 (P.L. 106-224) (“ARPA”), The Food, Conservation, and Energy Act of 2008 (P.L. 110-246) (“2008 farm bill”), and The Agricultural Act of 2014 (P.L. 113-79) (“2014 farm bill”). Notable changes in the 1994 Act were (1) the requirement of producer enrollment in crop insurance as an eligibility criterion for the support provided by certain federal farm support programs; (2) the creation of catastrophic coverage option (CAT) which offers protection of 50% on yield and 60% of expected price all at 100% premium subsidy but with an administrative fee of 50 \$/crop/county; (3) the increment of coverage options above CAT; and (4) the authorization of revenue protection products. The ARPA expanded the geographic availability, removed restrictions on livestock products, expanded private sector product development, and authorized premium reduction plans that allow approved insurance providers (AIPs) to offer reduced premiums, thereby improving their competitive advantage. The 2008 farm bill eliminated the premium reduction plans enacted by ARPA, reduced premium subsidies for area plans, and increased CAT fees from 50 to 300 \$/crop/county. The 2014 farm bill also expanded the program by authorizing higher coverage levels with lower deductibles and introduced new products aimed at shallow loss coverage (e.g., Supplemental Coverage Option and Stacked Income Protection), providing producers the option to exclude low yields from their production history under certain circumstances, linked premium subsidy eligibility to conservation compliance, and premium subsidy reduction to planting on native sod in some states. Figure 1, retrieved from Economic Research Service [ERS] (2022), shows that these legislative changes improved crop insurance participation.

The second set of instruments first introduced by Yu et al. (2018) is based on policy changes to the suite of subsidy rates. Yu et al. (2018) argue that changes in legislation create a structural break that shifts the suite of subsidy rates exogenously in a way that is not driven by endogenous factors related to crop production. These instruments used by Yu et al. (2018) are generated by dividing the annual subsidies for yield and revenue protection policies for a given coverage level and dividing this by the sum of their associated total premiums. Yu et al. (2018) recommended the 65% and 75% coverage levels as the point of aggregation since they are the most patronized and observations exist for all years;

U.S. Federal Crop Insurance Program participation, 1975–2021



GDP = gross domestic product.

Note: The aggregate coverage level is calculated as total annual Federal Crop Insurance Program insured liability, divided by potential liability. U.S. agricultural sector production value is taken from USDA, ERS Farm Income and Wealth Statistics data product and is recorded on a calendar year basis, while Federal Crop Insurance outcomes are on a crop-year basis. Selected crop acreage is the sum of planted acres for barley, corn, cotton, flaxseed, oat, peanut, potato, rice, rye, sorghum, soybean, sugarbeet, sunflower, sweet potato, and wheat; and harvested acres for sugarcane, and tobacco. Public Law (PL) numbers for each piece of referenced legislation are as follows: 1980 Act (PL 96–365), 1994 Act (PL 103–354), 1996 Farm Bill (PL 104–127), 2000 Agricultural Risk Protection Act (PL 106–224), 2008 Farm Bill (PL 110–246), 2014 Farm Bill (PL 113–79), 2018 Farm Bill (PL 115–334). FCIP data are based on the Risk Management Agency Summary of Business as of March 2, 2023.

Source: USDA, Economic Research Service (ERS), using data from USDA, Risk Management Agency, USDA, National Agricultural Statistics Service, and USDA, ERS Farm Income and Wealth Statistics.

Figure 1. US Federal Crop Insurance program participation improved with legislation.

thus, two sets of variables are developed and used as the instruments. Even though instruments by Yu et al. (2018) are also based on one-time changes in legislation, unlike the first they are scalar. Thus, they can capture exogenous changes in legislation, and hence crop insurance outcomes, beyond the initial year of passage of the associated law.

The first and second sets of instruments all rely on crop insurance legislation set at the national level. Thus, even though they can account for within-county variation in crop insurance outcomes over time, they do not provide the spatial (between-county) variation necessary for proper instrumentation in a cross-sectional setting. The third set of instruments overcomes this by using a measure of the degree to which crop insurance legislation affects a county. First introduced by DeLay (2019), these instruments are based on a measure of preferences for different types of insurance before the enactment of legislative changes discussed above. Particularly, these instruments are generated by interacting a policy exposure measure with dummy variables representing each of the four policy change periods (essentially the first set of instruments). In their application, DeLay (2019) used total insured acres divided by total cropland as the policy exposure measure. Total acres for each period were taken as reported in the associated latest Census of Agriculture survey. According to DeLay (2019), the instrument captures how much a county sees its effective subsidy rates increase and hence can be used to estimate the variation in insured acreage due only to changes in government policy and not to contemporaneous land-use decisions. Connor et al. (2021) adopted a version of this instrument that used a policy exposure measure constructed as the proportion of insured acres for coverage levels.

The final set of instruments, upon which this study is based, relies on the current rating framework used by RMA (i.e., continuous rating). The main idea is that FCIP premium rates

are determined via exogenously set rating parameters, and hence, the parameters if known can be used to generate exogenous variation in premium rates that producers face when making crop insurance decisions. To the best of our knowledge, Woodard and Yi (2020) and Tsiboe and Turner (2022) are the only studies relying on this strategy. According to Tsiboe and Turner (2022), even though information on producers that participated in the FCIP in the past is used in the determination of the policy variables, it is unlikely that the information from a single producer will influence their outcome in such a way that will be advantageous to them in the current season. Particularly, previous yields reported by producers contribute to a producer's historical loss experience but are fixed and not under the control of producers once reported. Accordingly, FCIP premium rates can be thought of as sequentially exogenous to producer insurance decisions. This assumption is plausible because the rating parameters are finalized a season or two in advance. Consequently, the exogenous portion of FCIP premium rates is driven by the rating parameter and the endogenous portion by the insured's choice of insurance unit, coverage level, and their historic production experience.⁴

Although Woodard and Yi (2020) and Tsiboe and Turner (2022) use a similar methodology of approximating rating parameters, a few nuanced differences persist. Given the form of the continuous rating formula, Woodard and Yi (2020) estimated the rating parameters from a latent rate curve using observed loss experience data via a regression framework.⁵ According to Tsiboe and Turner (2022), estimating the rating parameters from the loss experience data including the year for which they were set does not make them entirely void of endogeneity problems. Alternatively, Tsiboe and Turner (2022) essentially take the rating parameters as given (i.e., the same parameters farmers faced when they made the decisions leading to the loss experience data), under the presumption that this approach makes their instrument less likely to be afflicted by the same source of previously described simultaneity bias.

In this study, we aim to improve the identification of crop insurance for both temporal and spatial settings. Currently, an instrument based on structural dummy variables (Schoengold et al. 2015) and those based on national-level subsidy rates (Yu et al. 2018) lack the needed variation for applications in a cross-sectional setting. On the other hand, those (Woodard and Yi 2020; DeLay 2019) using instruments based on loss experience outcomes including the year of the study introduce new sources of bias that likely makes these alternative instruments not entirely void of the same endogeneity problems they hope to address. Neither Woodard and Yi (2020) and Tsiboe and Turner (2022) guide how to arrive at premium rating-based instruments for years before 2000. While Tsiboe and Turner (2022)'s approach most closely reflects reality, it is only relevant for years after 2000 when the rating method upon which their instruments are based was in effect.⁶

⁴Historic production experience is influenced by past producer production experiences which are under the control of the producer, thus adding an endogenous component to historic production. Production practices are likely to be relatively constant meaning historic production experiences are mostly driven by stochastic elements (ex: weather).

⁵Woodard and Yi (2020) estimate an instrument that mimics the relationship between RMA's rate differential factors and coverage level to capture curvature in the premium schedule. However, the rate differential factors exhibit relatively little variation across the many different insurance pools. In other words, much of the variation in crop insurance premium rates is exhibited in the form of changes in intercept shifts, rather than variation in the change in premium rate across coverage levels. Here, we propose a method that uses the observed demand data to estimate an instrument that captures the level of premiums, which is both easier to estimate and provides a means to estimate instruments for the entire history of the FCIP.

⁶Before the continuous rating formula, RMA published Base Premium Rates for APH yield-based crops using a fixed rate for a span of yields, which was typically nine rate spans (R-span) per crop. The continuous

Currently, the rating parameters from 2011 to 2023 are available in easily accessible files via RMA's Actuarial Data Master (ADM) data set, but those from 2001 to 2010 are not easily accessible.⁷ Additionally, since the continuous rating was not in existence pre-2001 there are no existing values for rating parameters for most of the FCIP historical time series data. However, as will be shown, some of the continuous rating parameters (specifically the base rate reflecting the county reference rate and catastrophic fixed loading factor) can be approximated using historical data.

Premium rating system

The system for rating crop insurance contracts within the FCIP, which is based on a principle known as loss-cost ratio (LCR) rate-making, is understandably very complex (see for example K. H. Coble et al. (2010) and K. Coble et al. (2020)). This section attempts to provide a simplified overview of FCIP contract rating as it relates to this study. Focus is given to individual-level yield protection policies, but the discussion that follows can be applied to other policy types. By extension, output price is assumed to be equal to unity without loss of generality.

Individual-level yield insurance contracts have eight main components: rate yield (\bar{y}), approved yield (\check{y}), coverage level (θ), indemnity ($I(y)$), premium rate (τ), premium (P), and subsidy ($S(\theta, u)$). Rate and approved yields are essentially the averages of actual production history (APH) reported by farmers with the main difference being that the latter is adjusted upward via several aspects of RMA's actuarial process (e.g., yield exclusion, yield substitution, and trend adjustments).⁸ For a given contract, the producer can choose to insure their approved yield at a coverage level of θ with a respective yield guarantee of $\theta\check{y}$. Thus, the per-acre indemnity for a given yield outcome, y , is given by $I(y) = \max[0, \theta\check{y} - y]$.

In the U.S., FCIP premium rates are required to be set such that the resulting policies are actuarially fair; meaning premiums are set equal to expected per-acre indemnities. Standardizing expected per-acre indemnities by insured liability can be characterized by the following equation.

$$\tau(\theta) = \frac{E[I(y)]}{\theta\check{y}} = \frac{1}{\theta\check{y}} \int_0^{\theta\check{y}} (\theta\check{y} - y)f(y)dy \quad (1)$$

where $f(y)$ is the probability density function of y . Since indemnities are stochastic and not known at the time the policy is written, crop insurance policies are typically generated under the assumption that $f(y)$ is conditional on an adjustment mechanism that relies on the underlying risk profile of the producer seeking insurance. This risk profile is not observable and is approximated by the productivity of the producer seeking insurance relative to their peers. The magnitude of this adjustment is determined using RMA's

rating formula develops a rate for each yield rather than for ranges of yields as the R-span does. Recent work shows that the main component of the continuous rating formula that induces rating heterogeneity among farmers does a relatively good job at incorporating soil information into rates when a long yield history is used to approximate the farmers' yield (Tsiboe and Tack 2022).

⁷Obtaining the rating parameters for 2001–2010 is possible but requires merging and aggregating many different data sources and then reconstructing the rating parameters from scratch using actuarial relationships. Doing so can be difficult without an intimate understanding of each data source.

⁸Higher approved yields result in higher yield guarantees leading to higher indemnities for a given yield outcome which improves producer welfare (Adhikari et al. 2013).

“continuous rating formula” (Milliman and Robertson 2000; Risk Management Agency [RMA] 2000, 2009).

The default form of the continuous rating formula is specified according to the following equation

$$\tau_{ijt} = \left[\alpha_{jt} \left(\bar{y}_{ijt} / \bar{y}_{jt} \right)^{\beta_{jt}} + \delta_{cjt} \right] \vartheta(\theta_{ijt}, u_{ijt}) \quad (2)$$

where the subscript ijt represents an insured (i) seeking protection defined by an insurance pool (j) for crop year t .⁹ The parameters α_{jt} and δ_{jt} , respectively, represent the reference rate and catastrophic fixed loading factor for a common coverage level (conventionally at the 65% level) for the insurance pool. In what follows, the continuous rating formula is used to adjust the common coverage level rate for the insurance pool based on the specific risk and choices of the producer seeking insurance. These mechanisms are analogous to how other property and casualty insurance rate factors are developed from a combination of experience and differential exposure information (Sherrick et al. 2014).

The first step in calculating rates for a given producer is to adjust the reference rate based on the assumption that historical yield and risk are correlated. This adjustment is represented in equation (3) by $\left(\bar{y}_{ijt} / \bar{y}_{jt} \right)^{\beta_{jt}}$ which is the ratio of a given producer's historic mean yield, \bar{y}_{ijt} , divided by a reference yield (\bar{y}_{jt}) of the insurance pool in which they are rated, raised to a power, β_{jt} , known as the continuous rating exponent. The continuous rating exponent takes on a negative value which assumes that higher production levels correlate with less risk. This means that when a producer's historical mean yield is greater than the county reference yield (i.e., $\bar{y}_{ijt} / \bar{y}_{jt} > 1$), the county reference rate is adjusted downward for that producer.¹⁰ The next step in determining a producer's premium rate is to adjust their initial rate, $\alpha_{jt} \left(\bar{y}_{ijt} / \bar{y}_{jt} \right)^{\beta_{jt}} + \delta_{jt}$, via a scaling function, $\vartheta(\cdot)$, which is based on the producer's choice of coverage level (θ_{ijt}) and insurance unit election (u_{ijt}).¹¹

The total price of the insurance contract, P , is set equal to the product of the premium rate, $\tau(\cdot)$, and the yield guarantee, $\theta\bar{y}$, such that $P = \tau(\cdot)\theta\bar{y}$. The final price paid by the insured is subsidized at a rate $S(\theta, u)$ that is tied to coverage level and insurance unit and not to location or the crop. RMA internally sets the values for the parameters (α , δ , β , and \bar{y}_c) in the continuous rating formula using historical loss data from the FCIP coupled with empirical approaches originating in the early 1980s (Sherrick et al. 2014). Thus, because coverage in the 1980s was limited to only 65% for yield protection, the rating system harmonizes all historical contracts to a 65% coverage on yield. Given the harmonized data, the reference rate (α) for an insurance pool is determined as the annual mean of the pool's historical LCR which is capped to limit the influence of outliers from anomalous catastrophic loss events (K. H. Coble et al. 2010). The excess exposure resulting from the capped LCR is then used to determine the catastrophic fixed loading factor (δ). According to the literature, the pool's reference yield (\bar{y}_c) is determined as the ten-year average of the

⁹RMA defines an insurance pool as a group of producers formed by the unique combination of county, crop, crop type, and production practices.

¹⁰(Milliman and Robertson 2000) indicated that this assumption is based on research that demonstrated an inverse relationship between yield and risk. (Botts and Boles 1958; Skees and Reed 1986) corroborate this finding. However, Goodwin (1994) has questioned the robustness of this relationship.

¹¹For a producer seeking revenue protection, there are additional adjustments that compare the covariance between yield and price to determine the additional cost of revenue insurance based on volatility factors. These volatility factors attempt to replicate the potential price spread reflected in the implied volatility of actual options traded on the underlying futures price against which the policy is priced and adjusted for the insurance period.

yield of the insurance pool (Rejesus et al. 2010), and the continuous rating exponent (β) is estimated by regressing LCR on $\bar{y}_{ijt}/\bar{y}_{jt}$ using farm-level data (Tsioboe and Tack 2022).

From the foregoing discussion, the exogenous portion of FCIP premium rates ($\tau(\theta, u, \bar{y}; \omega)$) is driven by the policy parameter space $\omega \in \{\alpha, \beta, \delta, \bar{y}_c\}$ and the endogenous portion by the insured's choice of insurance unit (u), coverage level (θ), and their historic production experience \bar{y} . In what follows, we outline how a target rate reflecting the county reference rate (α) and catastrophic fixed loading factor (δ) can be approximated using historical data.

Deriving crop insurance rating parameters

It is worth stressing here that our goal is not to provide alternative rating procedures but to illustrate how key parameters in current individual policy rating could be approximated for instrumental variable purposes. Several third-party FCIP reports on crop insurance rating methodology are of interest to this study which is used to draw insights to generate target rates that are comparable to county reference rates and catastrophic fixed loading factors under the continuous rating framework. We primarily draw from the RMA-commissioned review of the APH and COMBO rating methods (K. H. Coble et al. 2010). For this study, the FCIP has publicly available historic loss experience data from 1948 to 2021 (i.e., 74 years of data) of which data from 1961 to 2020 (i.e., 60 years of data) are relevant for this study.

This study first defines an aggregation unit characterized by a cross-sectional index i (e.g., county-crop, county, state, etc.) and time index t (e.g., year) for which an approximation of a target rate (τ_{it}) is desired. The stochastic historic empirical LCR is denoted as $\{LCR_{it-s}\}_{s=2}^T$, where $LCR_{it} = \min(1, \max[0, \text{indemnity}_{it} \div \text{liability}_{it}])$.¹² For each aggregation unit, the LCR reflects the proportion of insured liability that was paid to producers each year, that is, the return on insurance. In practice, RMA first adjusts both historic liabilities and indemnities to a common coverage level (conventionally at the 65% level) and harmonizes all contracts to yield-based protection before taking this ratio. However, since the objective is to estimate target rates over the entire history of the FCIP, this study forgoes these adjustments because historic coverage level choice data before 1989, and the price data needed to convert all contracts to yield-based protection are currently not available in the public domain.

The timing of actuarial releases by RMA suggests that rating parameters are derived two years ahead of a given crop year; thus, the historic empirical LCR series for a given cross-sectional index end at $t - 2$. Given $\{LCR_{it-s}\}_{s=2}^T$, an empirical burn rate ($\bar{\tau}_{it,BR}$) is estimated as the mean of the series. Next, the burn rate is loaded (by dividing by 0.88) to obtain the raw rate ($\bar{\tau}_{it,RR}$). The underlying data for rating crop insurance (i.e., LCRs) are often highly variable. Thus, the RMA rate-making process includes several steps intended to smooth some of these fluctuations to prevent outsized influence from a singular data point. One such process is a spatial smoothing procedure that takes a weighted average of the raw rate and surrounding raw rates (i.e., county group rate) to determine a final rate. This study applies a similar strategy where a county group rate ($\bar{\tau}_{it,CR}$) is defined as the weighted average (by the sum of the liability for periods less than $t - 2$) of the raw rates of all adjoining counties. The final target rate which can be used as an instrument is then taken as

¹²The derivation of the observed loss cost ratio is analogous to the case of index-based insurance where the positive difference between the trigger grid index, less the final grid index, is replaced with the indemnity and the trigger grid index is replaced with the liability.

$\tau_{it} = 0.6\tilde{\tau}_{it,RR} + 0.4\tilde{\tau}_{it,CR}$, where the weighting values of 0.6 and 0.4 are based on past RMA practices (K. H. Coble et al. 2010).¹³

The resulting target rate, τ_{it} , is slightly different than any existing RMA premium rating parameters but performs a similar function of explaining variation in lost cost ratios (this is validated empirically in a later section). Thus, τ_{it} is analogous to the exogenous parameters from equation (2) in the sense that τ_{it} is exogenous to any producer insurance decisions and is also correlated with final premium rates by being derived with the purpose of empirically fitting historical lost cost ratio data. However, τ_{it} has the benefit of being derivable for any spatiotemporal setting from publicly available data, making it a more accessible instrument relative to other potential rating parameter-based instruments. More specifically, τ_{it} approximates RMA's initial rate, (defined in equation (3) as $(\alpha_{jt} (\bar{y}_{ijt}/\bar{y}_{jt})^{\beta_{jt}} + \delta_{jt})$ for the special case where the producer's mean yield, \bar{y}_{ijt} , is exactly equal to the reference yield, \bar{y}_{jt} . In other words, τ_{it} reduces to an approximation of $\alpha_{jt} + \delta_{jt}$ which tracks variation in premium rates without the influence of the endogenous portions of the continuous rating formula.

Data and construction of relevant variables

Primary FCIP information for the construction of the target rate which can be used as an instrument was retrieved from RMA's summary of business and cause of loss files. RMA provides three different versions of the summary of business which are aggregated by:

- (1) State/County/Crop,
- (2) State/County/Crop/Coverage, and
- (3) State/County/Crop/Coverage/Type/Practice/Unit Structure.

For the cause of loss, RMA provides three different versions:

- (1) Summary of Business with Month of Loss,
- (2) Indemnities Only, and
- (3) Indemnities with Month of Loss.

We use information from all six sources to construct a county-crop panel data set 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 procedure used to aggregate the information from all six sources first sums up the variables of interest for each county-crop-year combination, separately for each of the six data sources. Then for each county-crop-year combination, the maximum among the six data sources is used as the representative value for each variable. The rationale for doing this is that RMA periodically updates the summary of business and cause of loss files as new information, which is continuously provided by AIPs, becomes available. This means any given data set for a particular year continues to receive minor changes for some time after that particular year has ended. Thus, there will sometimes be differences in the same metric across different file sources, but presumably, the metrics tracked in these sources, such as policies sold, are

¹³Since 2006, RMA replaced the historical weights (0.6 and 0.4) with a procedure which credibly weights the target county and adjoining counties based on a function of acres. We do not do this here in the spirit of presenting a method that is easily accessible to a wider audience.

monotonically increasing over time meaning the maximum value from all available data sources should be the most reflective of reality.

In addition to this data, the study also retrieved actuarial information (i.e., actual values for the county base rates) from RMA's ADM spanning from 2011 to 2020.¹⁴ The specific data retrieved from the ADM include the county reference rate and catastrophic fixed loading factor. For each county-crop-year, the RMA-set county reference rate and catastrophic fixed loading factor were taken as the mean of all the respective values retrieved from RMA's ADM that could be associated with that county-crop-year combination. Finally, a representative target rate comparable to those generated from this study was taken as the sum of the county reference rate and catastrophic fixed loading factor.

Table 1 shows the descriptive statistics of the data used for the approximation of target rates. The sample is composed of 564,340 observations with 26,078 unique county-crop FCIP programs. On average, an FCIP county-crop program in the data had an annual insured liability of \$3.5 Million on about 15 thousand acres. Total collected premiums for a county-crop program averaged \$309,000 of which \$91,000 was paid for by government subsidies. The annual mean LCR for FCIP county-crop program from 1948 to 2020 is estimated at 0.12 indicating that for every dollar of liability insured, about 12 cents was paid to farmers to indemnify their losses.

Lost cost ratio predictive performance

The RMA's rate-making procedures are intended to allow for the estimation of the expected loss component of rates, represented by the LCR (K. H. Coble et al. 2010). Thus, the approximation of the target rate proposed by this study should be able to predict LCRs to a similar degree of accuracy compared to the status quo methods used by RMA. In that regard, this study evaluates the LCR predictive performance of all the various rates estimated in the approximation procedure by using both out-sample (via five-fold-cross-validation) and in-sample prediction performance. Both measures are operationalized on an annual basis from 2011 to 2020 and are evaluated in terms of the LCR prediction error (mean squared error) generated in this study relative to the mean error when actual observed rates set by RMA are used to predict the LCR. Thus, values below one are preferred with values significantly larger than 1 indicating worse performance over LCR predictability.

Figure 2 shows that the final target rate estimated in the approximation procedure reflects expected losses better than the burn rates, raw rate, and county group rate. Model diagnostics (shown in Figure S1) also indicate that the final target rate performs slightly better in terms of R-squared, AIC, and BIC. Figure 3 shows box plots of the final target rate and the official rates produced by RMA. As can be seen in the figure, the approximation proposed in this study yields target rate distributions that are quite close to the target rate distributions produced by RMA for the years 2011–2020.¹⁵

¹⁴ADM files for each insurance year are available at: https://ftp.rma.usda.gov/pub/References/actuarial_data_master/. The aggregation of ADM information is based on initial work by Tsiboe and Tack (2021) using Beocat, a High-Performance Computing (HPC) cluster at Kansas State University (<https://beocat.ksu.edu/>).

¹⁵On a pairwise basis, Figure S2 shows that the final target rates are systematically higher than the mean of actual observed rates set by RMA. The source of this apparent difference can be gleaned from Figure S3 where this is due to revenue-based contracts. Figure S3 shows that while the yield-protection final target rates follow the observed rates closely in both magnitude and annual dynamics, the revenue-protection rates are significantly higher even though they also follow the underlying annual dynamics.

Table 1. Means and standard deviations of selected variables on US Federal Crop Insurance county-crop programs (1948–2020)

Variables	Mean (Standard deviation)
Net insured area (1,000 acres)	14.736 (48.668)
Total liability insured (\$ 1,000) [A]	3,449.423 (11,946.263)
Total premium (\$ 1,000)	309.396 (1,042.076)
Total premium subsidy (\$ 1,000)	90.752 (312.426)
Total Indemnity paid [B]	264.576 (1,536.565)
Loss cost ratio [(B/A)]	0.120 (0.193)
Number of county-crop combinations	26,078
Number of observations	564,340

Note: The data was constructed by the authors using primary data from the Risk Management Agency’s summary of business and cause of loss files.

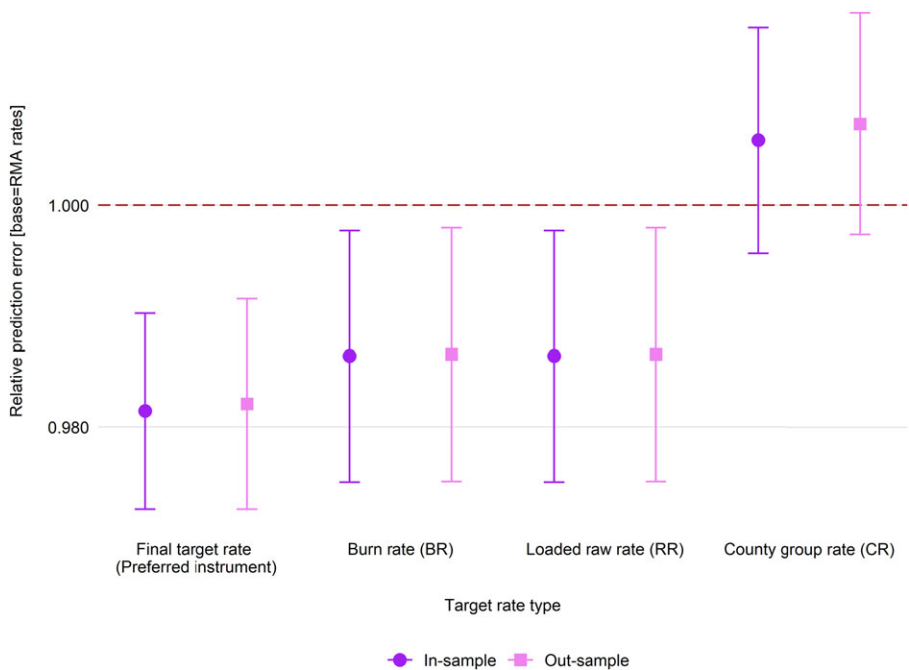


Figure 2. Loss cost ratio predictive performance of different rate approximation methods.

Notes: Graph shows the in- and out-sample loss-cost ratio predictive performance of various approximations of USDA Risk Management Agency (RMA) crop insurance target rates (i.e., the sum of county reference rate and catastrophic fixed loading factor). The values are evaluated in terms of relative performance to the mean of RMA target rates for 2011–2020.

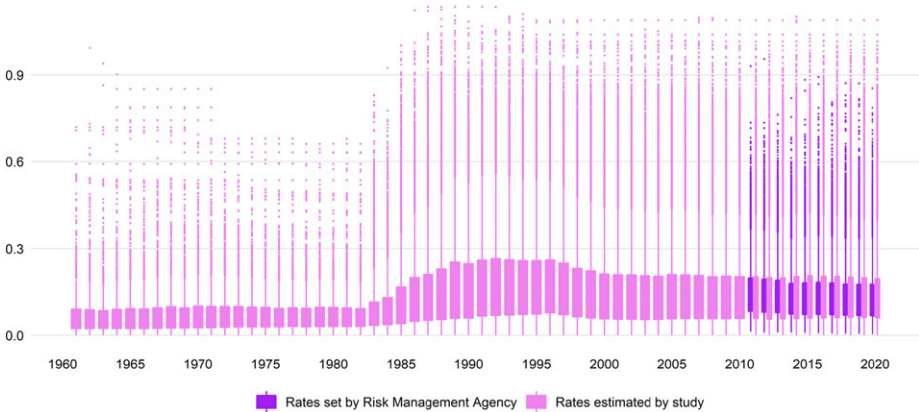


Figure 3. County-crop approximated target rate distribution by crop year.

Notes: Graph shows the distribution of USDA Risk Management Agency (RMA) crop insurance target rates (i.e., the sum of county reference rate and catastrophic fixed loading factor) and its preferred approximation from the stud.

Empirical application

As was discussed in the introduction, several sources of endogeneity are of concern in empirical studies making use of crop insurance demand as either an independent or dependent variable. What follows are two sets of empirical examples of how to address these endogeneity concerns by using derived rating parameters as instrumental variables. The primary goal of the empirical examples is not to provide a detailed analysis of the underlying issues but to illustrate how the target rates approximated in this study could be used to deal with endogeneity as it relates to the context of this paper. The examples can easily be generalized to other applications and research settings that make use of direct crop insurance outcomes like participation and premiums as independent variables.

Empirical application: crop insurance demand as a function of premiums

This first set of empirical applications exemplifies the use of the derived instrument to address the situation where the total premium for crop insurance is an independent variable of interest. Past studies try to circumvent the endogeneity problem in modeling crop insurance demand by using Heckman-type models of participation (Goodwin and Kastens 1993; Richards 2000; Shaik et al. 2008). Unfortunately, the majority of existing studies do not consider this endogeneity problem (Gardner and Kramer 1986; Hojjati and Bockstael 1988; Barnett et al. 1990; Calvin 1990; Goodwin 1993; Goodwin and Kastens 1993; K. H. Coble et al. 1996; Yi et al. 2020; Bulut and Hennessy 2021). Thus, empirical approaches seeking to econometrically identify FCIP demand elasticities, which do not explicitly account for this source of endogeneity, are suspect from a causal perspective. Here, county-level corn production data are used to estimate the responsiveness of crop insurance demand to premium rates at both the intensive and extensive margins.

Crop insurance data for corn from 1975 to 2020 used in this application are obtained from the same data sources used in approximating the target rates. Specifically, the

variables of interest include the total potential liability, total insured liability, total premium, and total subsidy for corn in county i for crop year t .¹⁶ Data on planted acres of corn were also retrieved from NASS Quick Stats, and where missing, the planted acres were filled in with harvested acres which were also obtained from the same source. Using this data, several variables of interest were calculated; the aggregate coverage level (θ_{it}) as total insured liability divided by total potential liability; premium per dollar of liability (r_{it}) as the total premium divided by total insured liability; and subsidy per dollar of premium (s_{it}) as the total premium subsidy paid divided by the total premium. Finally, this study used the same corn price discovery method employed by RMA in estimating projected prices to approximate state-level expected prices for corn from 1975 to 2020. Table S1 shows the descriptive statistics of all variables used in this empirical application.

The empirical model is formulated as

$$\ln \theta_{it} = \beta_0 + \tau \ln r_{it} + \beta_s \ln s_{it} + \beta_w w_{it} + v_{it} + \varepsilon_{it} \quad (4)$$

where the variable θ_{it} is a measure of crop insurance uptake at both the intensive margin and extensive margin (A second version of this model is also specified to explain demand exclusively at the extensive margin by setting θ_{it} equal to the share of eligible acres enrolled in crop insurance as the dependent variable).¹⁷ The vector w_{it} contains the log expected price, log planted acres, and state-specific time trends. The terms v_{it} and ε_{it} are county fixed effects (FE) and error terms, respectively. The error terms are assumed to have an expected value of 0 but can be heteroskedastic so standard errors are clustered by the county to allow for ε_{it} to be spatially correlated within each crop year (Petersen 2009; Cameron et al. 2011; Thompson 2011).

The premium per dollar of liability (r_{it}) in equation (4) is the weighted average of farm-level FCIP premium rates represented by equation (2). As noted in previous sections, FCIP premium rates are driven by both exogenous and endogenous factors, which makes them potentially endogenous to crop insurance demand. For example, a corn farmer with a high-risk profile will end up paying a higher premium compared to a lower-risk farmer. At the same time, farmers tend to allocate fewer corn acres in a county characterized by a greater risk profile, *ceteris paribus* (Hojjati and Bockstael 1988). Alternatively, corn growers in counties that have a high-risk profile for corn production are likely to increase their crop insurance coverage. In these cases, estimating the empirical model via ordinary least squares (OLS) will result in biased coefficient estimates since “risk” is an omitted independent variable that partially determines crop insurance coverage and is correlated with the premium per dollar of liability. The identification strategy in this empirical application is to utilize the exogenous RMA-set variations in continuous rating formula policy parameters to generate exogenous variation in premium per dollar of liability via instrumental variables estimation. However, since these parameters are not publicly available for the entire period under review (1975–2020), the empirical application uses a comparable approximation of the sum of county reference rate and catastrophic fixed loading factor as

¹⁶Potential liability for a given producer is calculated by taking the producer’s total insured liability and dividing by the producer’s chosen coverage level which represents the total insured liability if that producer was to insure at a 100% coverage level. Summing this metric across all participants in a county for a given year provides a metric for the total liability that could conceivably be insured in that county for that year given the current participants and enrolled acreage.

¹⁷By construction, θ_{it} increases when individual coverage levels are increased for a fixed amount of enrolled acreage (an increase at the intensive margin), but also increases if individual coverage levels remain static with additional acreage enrolled (an increase at the extensive margin).

outlined in previous sections. In this setting, the aggregation unit to generate the instrument (i.e., the final target rate) is characterized by a cross-sectional index defined as the unique combination of county crops in the underlying data for the analysis.

Table 2 reports the estimated results for Equation (4) using the combined measure of demand at the intensive and extensive margins. Columns (1) and (2) report the estimation results based on two-staged least squares (2SLS) regression with and without FE, respectively. Similarly, columns (3) and (4) report OLS results with and without FE, respectively. The Anderson-Rubin chi-sq test for exogeneity of premium per dollar of liability (r_{it}) is rejected ($p < 0.01$) for both 2SLS models, suggesting that the previously described concerns over endogeneity are valid and that an instrumental variables approach is appropriate. The question that remains is if the approximated target rates are strong enough to aid in identifying the causal effect of premium rates on crop insurance demand. First-stage F-statistics are well above the thresholds suggested by Stock and Yogo (2005) indicating that the final target rate is not a weak instruments.

Generally, Table 2 shows that the preferred model [Model 1: FE-2SLS] has estimated coefficients that are varied in direction and larger in magnitude than the model without instrumental variables (IVs) [Model 3 and 4: OLS and FE-OLS] or without FE [Model 2: 2SLS] suggesting that either crop insurance choices or time-varying unobserved individual risk profiles bias the responsiveness of crop insurance demand.

Table 3 reports estimated results for Equation (4) as specified to represent demand exclusively at the extensive margin and yields qualitatively similar results when compared to Table 2. The Anderson-Rubin chi-sq test is rejected ($p < 0.01$) for the 2SLS model making use of fixed effects (although this is not the case if fixed effects are excluded). Weak instruments are again not a problem in this specification with both 2SLS models exhibiting first-stage F-stats that are well above conventionally set thresholds. Again, the preferred model [Model 1: FE-2SLS] yields a coefficient on the premium per dollar of liability that is opposite in sign and larger in magnitude compared to the OLS models in Table 3.

Concerning the effect that subsidies have on demand for insurance, Table 2 suggests a negative effect while Table 3 suggests a positive relationship between increasing subsidies per dollar of premium and demand at the extensive margin. This could be interpreted as evidence that increasing subsidy rates induce a trade-off between demand at the intensive margin in favor of increased demand at the extensive margin. However, as noted previously, this analysis is not intended to be a comprehensive analysis of demand for crop insurance, but instead serve as a demonstration of the validity of the proposed instrumental variable.

Empirical application: farm debt as a function of crop insurance participation

The second set of empirical applications demonstrates how to instrument for measures of crop insurance demand that appear as independent variables of interest but remain endogenous with respect to the outcome of interest. This example is based on existing work assessing the role that crop insurance has on farm debt use (Ifft et al. 2015) and demonstrates that the proposed instrument also has favorable diagnostics when demand for insurance appears as an endogenous independent variable.

Both theoretical and empirical research on risk balancing has shown that producers could increase financial risk in response to a decline in business risk. By extension, this also means that farm policy programs like the FCIP that decrease business risk could have causal impacts on financial risk levels. However, producers' decisions on coverage through crop insurance and farm debt are likely influenced by the same unobservables, and thus endogenous. Ifft et al. (2015) account for this simultaneity bias in their estimation of the

Table 2. Crop insurance demand for corn production in the US for 1989–2020 at The intensive margin is inelastic to changes in premiums rates

Variables	(1) FE-2SLS ^{††}	(2) 2SLS	(3) OLS	(4) FE-OLS
Natural logarithm of premium per dollar of liability	−0.106*** (0.017)	−0.054*** (0.005)	0.061*** (0.003)	−0.001 (0.003)
Natural logarithm of subsidy per dollar of premium	−0.355*** (0.016)	−0.353*** (0.012)	−0.435*** (0.011)	−0.392*** (0.010)
Natural logarithm of planted acres	0.000 (0.000)	0.000* (0.000)	0.001*** (0.000)	0.002*** (0.000)
Natural logarithm of real expected corn price	−0.002 (0.002)	0.003* (0.002)	0.006*** (0.002)	0.004** (0.002)
Number of observations	33,314	33,314	33,314	33,314
Number of counties	1,159	1,159	1,159	1,159
Endogeneity: Anderson-Rubin chi-sq test	60.684***	131.613***	–	–
Endogeneity: MM distance test statistic	153.965***	285.326***	–	–
Weak identification test (Kleibergen-Paap rk Wald F statistic)	216.532***	3376.583***	–	–
Model F statistic	276.588***	345.868***	513.385***	426.181***

Notes: The outcome variable is county-level corn crop insurance demand at the intensive margin is measured as the total liability divided by the total potential liability implied by those who purchased corn crop insurance in that county. The data used was constructed by the authors using primary data from (1) Risk Management Agency’s summary of business and cause of loss files. The preferred model is ††. Significance levels – * $p < 0.1$ ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by county.

Table 3. Crop insurance demand for corn production in the US for 1948–2020 at The extensive margin is inelastic to changes in premiums rates

Variables	(1) FE-2SLS ^{††}	(2) 2SLS	(3) OLS	(4) FE-OLS
Natural logarithm of premium per dollar of liability	1.094*** (0.200)	0.038 (0.038)	−0.095*** (0.028)	−0.068*** (0.022)
Natural logarithm of subsidy per dollar of premium	0.113*** (0.022)	−0.048** (0.022)	0.060*** (0.019)	−0.044* (0.022)
Natural logarithm of planted acres	−0.015*** (0.002)	−0.025*** (0.003)	−0.020*** (0.002)	−0.029*** (0.002)
Natural logarithm of real expected corn price	−0.002 (0.002)	0.003* (0.002)	0.006*** (0.002)	0.004** (0.002)
Number of observations	41,230	41,230	41,230	41,230
Number of counties	1,159	1,159	1,159	1,159
Endogeneity: Anderson-Rubin chi-sq test	38.964***	1.027	–	–
Endogeneity: MM distance test statistic	43.325***	14.364***	–	–
Weak identification test (Kleibergen-Paap rk Wald F statistic)	231.791***	2653.871***	–	–
Model F statistic	260.755***	291.366***	388.119***	291.697***

Notes: The outcome variable is county-level corn crop insurance demand at the extensive margin measured as the total insured acres for corn divided by the total planted of corn in that county. The data used was constructed by the authors using primary data from (1) Risk Management Agency’s summary of business and cause of loss files, and (2) NASS Quick Stats. The preferred model is ††. Significance levels – * $p < 0.1$ ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses are clustered by county.

difference in debt use attributable to FCIP via two methods: (1) propensity score matching (to approximate a controlled experiment) and (2) seemingly unrelated regression (to explicitly model joint determination of farm debt structure and crop insurance participation). In this example, we provide another way to use the final target rate in this paper under an instrumental variable regression framework.

Like Ifft et al. (2015), we use data from the nationally representative Agricultural Resource Management Survey (ARMS), with the key exception that the sample is pooled for the ARMS fielded from 2011 to 2020.¹⁸ We follow Ifft et al. (2015) by limiting the sample to farm businesses as defined officially by USDA, which includes (1) farms where the principal operator's primary occupation is farming, (2) farms with sales over \$250,000, and/or (3) non-family farms.

The empirical model is formulated as

$$\operatorname{asinh} D_i = \beta_0 + \beta_\tau \operatorname{asinh} E_i + \beta_z z_i + \mu_i \quad (5)$$

where the variable D_i is a measure of debt use which, following Ifft et al. (2015), is defined as one of (1) total farm financial debt; (2) non-real estate financial debt; (3) real estate financial debt; (4) short-term financial debt (outstanding); (5) short-term financial debt (all); (6) farm business debt to asset ratio; (7) share of short-term financial debt (all) to total operating expenses; and (8) current ratio, all in per-acre terms. The key independent variable is crop insurance participation measured by the total per-acre premium paid by the farm (E_i). This is the only reasonable scalar measure of FCIP participation in the ARMS data as it has been shown that the key function that determines the premium paid by insured is increasingly convex in coverage (Woodard and Yi 2020). The vector z_i contains control variables and is defined, again based on (Ifft et al. 2015), to include operators' age and education; total operating acres; percent of cropland; unit structure; sales class as defined by USDA; ERS Farm Resource Region; and an annual trend variable. Because some of the debt-use decision outcome variables contain zero and non-positive values, we used the Inverse hyperbolic sine transformation (asinh). Likewise, the total premium paid by farmers also has zero values for non-participating farms. After the estimation of Equation (5), we used procedures outlined in the literature to approximate effects in elasticity terms (Bellemare and Wichman 2020). ARMS is based on a stratified sample design as such it requires weighted estimation of sample statistics; thus, USDA recommended delete-a-group (out of 30 groups) jackknife procedure was used to estimate standard errors. In this setting, the aggregation unit to generate the instrument is characterized by a cross-sectional index defined as the county.

Table 4 provides results for the empirical application focused on estimating farm debt as a function of total crop insurance expenditures. The null is rejected for the Anderson-Rubin test for 6 out of 8 considered debt outcomes suggesting that, in general, the instrumentation is an appropriate choice for this application. For all specifications in Table 4, weak instruments do not appear to be a problem (first-stage F-stats are over 100 for all specifications). Across all models, instrumental variables generally produce large coefficients on crop insurance expenditures relative to OLS results, but with similar or decreased statistical significance. This is potentially a result of the decreased efficiency of two-staged least squares. However, given the favorable diagnostic tests and theoretical basis for endogeneity, two-staged least squares estimates are still preferable to OLS in this empirical application.

¹⁸The Agricultural Resource Management Survey (ARMS) is the U.S. Department of Agriculture's primary source of information on the production practices, resource use, and economic well-being of America's farms and ranches (Kuethe and Morehart 2012).

Table 4. Federal Crop Insurance led to higher farm debt use in the US Farm economy from 2011 to 2020

Outcome	Responsiveness to Federal Crop Insurance expenditure		Number of observations	Endogeneity: Anderson-Rubin chi-sq test	Endogeneity: MM distance test statistic	Weak identification test (Kleibergen-Paap rk Wald F statistic)	2SLS Model F statistic
	(1) OLS	(2) 2SLS ^{††}					
Total farm financial debt	0.065*** (0.006)	0.608*** (0.165)	114,868	32.489***	26.261***	104.864***	27.746***
Non-real estate financial debt	0.104*** (0.010)	0.222 (0.203)	120,938	2.041	0.581	104.352***	2.063
Real estate financial debt	0.100*** (0.012)	0.399 (0.289)	120,938	5.064**	2.848*	104.352***	4.847**
Short-term financial debt (outstanding)	0.116*** (0.010)	0.361 (0.232)	120,938	4.450**	2.044	104.352***	4.402**
Short-term financial debt (all)	0.134*** (0.012)	0.424 (0.270)	120,938	4.305**	2.015	104.352***	4.249**
Farm business debt to asset ratio	0.064*** (0.006)	0.703*** (0.170)	120,938	35.798***	30.136***	104.352***	29.983***
Share of short-term financial debt (all) to total operating expenses	0.004 (0.004)	-0.046 (0.080)	120,936	0.909	1.077	104.355***	0.903
Current ratio	0.069*** (0.006)	0.739*** (0.179)	120,938	18.343***	15.249***	104.352***	17.380***

Notes: The farm-level data is from the 2011 to 2020 Agricultural Resource Management Survey (ARMS) Phase 3 data. The preferred model is ††. Significance levels - * $p < 0.1$ ** $p < 0.05$, *** $p < 0.01$. Standard errors are calculated using the delete-a-group jackknife.

Conclusion

Employing existing publicly available data, this paper provides an overview of how to derive a crop insurance rating parameter that is based on current actuarial practices employed by the USDA RMA. This derived rating parameter is exogenous to individual farmers' decisions by construction. Because of this, this parameter can be used to generate exogenous variation in both crop insurance premium rates and crop insurance participation via instrumental variables estimation to eliminate endogeneity bias in empirical models that use measures of crop insurance demand as either a dependent or independent variable.

After validating the derived rating parameter in a prediction exercise, it is employed in a couple of representative empirical applications to demonstrate how it can be used to improve econometric identification. One important caveat is that although the methods presented here have been validated by showing that the derived parameters predict LCRs with sufficient accuracy for the years 2011–2020, lack of data availability prevents explicit validation outside of this range which should be considered when using this methodology on historical data. The empirical applications presented here are based on both aggregate-level (county) and farm-level data. However, conceptually, these methods could be applied to other levels of aggregation as well. While we only considered the two-stage least squares estimator, there is no reason the approach presented here could not be applied to other estimators (e.g., matching estimators, selection models, three-stage least squares, etc.). As a final note, although the proposed instrument performs well in the presented empirical applications and is likely to perform well in other related settings, it is still up to the researcher to assess the appropriateness of the instrument in their specific research application from a theoretical point of view and to heed to suggestions of the relevant diagnostic tests.

As the empirical quality standards for research in the fields of agricultural economics and agribusiness rise, so does the demand for high-quality data to augment the research process. As was previously discussed, purging endogeneity from studies focused on the demand for crop insurance is one burgeoning use for the methods described in this study. However, as was demonstrated with the second empirical application in this study, the methods utilized here are also potentially valuable in other research applications where controlling for variation in crop insurance participation or crop insurance premiums may be necessary. Focus on tangentially related outcomes of interest such as farm production decisions, land-use allocation, or other risk mitigation behavior may benefit by incorporating metrics tracking crop insurance demand. In such a case, incorporating the methods outlined in this paper could prove to be valuable.

Supplementary material. For supplementary material accompanying this paper visit <https://doi.org/10.1017/age.2023.13>

Data availability statement. Replication materials are available from the authors upon request.

Acknowledgments. The findings and conclusions in this publication are those of the author(s) and should not be construed to represent any official USDA or U.S. Government determination or policy. This research was supported by the U.S. Department of Agriculture Economic Research Service.

Author contributions. Francis Tsiboe conceptualized the study, sourced, processed, and analyzed the data, and partly wrote the abstract, introduction, methods, results, and discussion sections. Dylan Turner partly wrote the abstract, introduction, methods, results, and discussion sections. All authors reviewed and approved the final version of the manuscript for publication.

Funding statement. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interests. The authors declare that they have no competing interests.

Ethical standards. This study does not contain results involving human participants performed by any of the authors.

References

- Adhikari, S., T.O. Knight, and E.J. Belasco.** 2013. "Yield Guarantees and the Producer Welfare Benefits of Crop Insurance." *Journal of Agricultural and Resource Economics* 38(1): 78–92. Available at: <https://www.jstor.org/stable/23496740>.
- Barnett, B.J., J.R. Skees, and J.D. Hourigan.** 1990. "Explaining Participation in Federal Crop Insurance." Annual Meeting, August 5–8, Vancouver, Canada 270875, American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association). 1990.
- Belasco, E.J.** 2020. "WAEA Presidential Address: Moving Agricultural Policy Forward: Or, There and Back Again." *Journal of Agricultural and Resource Economics* 45(3): 397–409. doi: [10.22004/ag.econ.304764](https://doi.org/10.22004/ag.econ.304764).
- Bellemare, M.F., and C.J. Wichman.** 2020. "Elasticities and the Inverse Hyperbolic Sine Transformation." *Oxford Bulletin of Economics and Statistics* 82(1): 50–61. doi: [10.1111/obes.12325](https://doi.org/10.1111/obes.12325).
- Botts, R.R., and J.N. Boles.** 1958. "Use of Normal-Curve Theory in Crop Insurance Ratemaking." *Journal of Farm Economics* 40(3): 733. doi: [10.2307/1235383](https://doi.org/10.2307/1235383).
- Bulut, H., and D.A. Hennessy.** 2021. "Estimating Elasticities via Market Share Impacts for Crop Insurance Coverage Options." 2021. Available at: <https://ageconsearch.umn.edu/record/312738>.
- Calvin, L.** 1990. *Participation in the U.S. Federal Crop Insurance Program*: US Department of Agriculture, Economic Research Service. Available at: <https://naldc.nal.usda.gov/download/CAT10879095/PDF>.
- Cameron, A., C.J.B. Gelbach, and D.L. Miller.** 2011. "Robust Inference with Multiway Clustering." *Journal of Business & Economic Statistics* 29(2): 238–249. doi: [10.1198/jbes.2010.07136](https://doi.org/10.1198/jbes.2010.07136).
- Cariappa, A.G.A., D.P. Mahida, P. Lal, and B.S. Chandel.** 2020. "Correlates and Impact of Crop Insurance in India: Evidence from a Nationally Representative Survey." *Agricultural Finance Review* 81(2): 204–221. doi: [10.1108/AFR-03-2020-0034](https://doi.org/10.1108/AFR-03-2020-0034).
- Claassen, R., C. Langpap, and J.J. Wu.** 2017. "Impacts of Federal Crop Insurance on Land Use and Environmental Quality." *American Journal of Agricultural Economics* 99(3): 592–613. doi: [10.1093/ajae/aa075](https://doi.org/10.1093/ajae/aa075).
- Coble, K., B. Goodwin, M.F. Miller, R. Rejesus, A. Harri, and D. Linton.** 2020. "Review of the Pasture, Rangeland, Forage Rainfall Index Crop Insurance Program Indexing and Rating Methodology Final Report." Report to RMA by Sigma Agricultural Risk and Actuarial Services, LLC, 2020. Available at: https://www.rma.usda.gov/-/media/RMA/Publications/Risk-Management-Publications/prf_contractor_report.ashx?la=en.
- Coble, K.H., T.O. Knight, B.K. Goodwin, M.F. Miller, R.M. Rejesus, and G. Duffield.** 2010. "A Comprehensive Review of the RMA APH and COMBO Rating Methodology Final Report." Available at: <https://legacy.rma.usda.gov/pubs/2009/comprehensivereview.pdf>.
- Coble, K.H., T.O. Knight, R.D. Pope, and J.R. Williams.** 1996. "Modeling Farm-Level Crop Insurance Demand with Panel Data." *American Journal of Agricultural Economics* 78(2): 439–447. doi: [10.2307/1243715](https://doi.org/10.2307/1243715).
- Connor, L., and A.L. Katchova.** 2020. "Crop Insurance Participation Rates and Asymmetric Effects on U.S. Corn and Soybean Yield Risk." *Journal of Agricultural and Resource Economics*. doi: [10.22004/ag.econ.298431](https://doi.org/10.22004/ag.econ.298431).
- Connor, L., R.M. Rejesus, and M. Yasar.** 2021. "Crop Insurance Participation and Cover Crop Use: Evidence from Indiana County-level Data." *Applied Economic Perspectives and Policy*. doi: [10.1002/aep.13206](https://doi.org/10.1002/aep.13206).
- DeLay, N.** 2019. "The Impact of Federal Crop Insurance on the Conservation Reserve Program." *Agricultural and Resource Economics Review* 48(02): 297–327. doi: [10.1017/age.2019.9](https://doi.org/10.1017/age.2019.9).
- Deryugina, T., and M. Konar.** 2017. "Impacts of Crop Insurance on Water Withdrawals for Irrigation." *Advances in Water Resources* 110(December): 437–444. doi: [10.1016/j.advwatres.2017.03.013](https://doi.org/10.1016/j.advwatres.2017.03.013).

- Du, X., J. Ifft, L. Lu, and D. Zilberman. 2015. "Marketing Contracts and Crop Insurance." *American Journal of Agricultural Economics* 97(5): 1360–1370. doi: [10.1093/ajae/aav024](https://doi.org/10.1093/ajae/aav024).
- Economic Research Service [ERS]. 2022. "Crop Insurance at a Glance." Available at: <https://www.ers.usda.gov/topics/farm-practices-management/risk-management/crop-insurance-at-a-glance/>.
- Fadhliani, Z., J. Luckstead, and E.J. Wailes. 2019. "The Impacts of Multiperil Crop Insurance on Indonesian Rice Farmers and Production." *Agricultural Economics* 50(1): 15–26. doi: [10.1111/agec.12462](https://doi.org/10.1111/agec.12462).
- Falco, S.D., F. Adinolfi, M. Bozzola, and F. Capitanio. 2014. "Crop Insurance as a Strategy for Adapting to Climate Change." *Journal of Agricultural Economics* 65(2): 485–504. doi: [10.1111/1477-9552.12053](https://doi.org/10.1111/1477-9552.12053).
- Gardner, B.L., and R.A. Kramer. 1986. "Experience with Crop Insurance Programs in the United States." In P. Hazell, C. Pomareda, and A. Valdes (eds). *Crop Insurance for Agricultural Development: Issues and Experience* (pp. 195–222). Baltimore and London: The Johns Hopkins University Press. <https://ebrary.ifpri.org/digital/collection/p15738coll2/id/126103>.
- Glauber, J.W. 2013. "The Growth of the Federal Crop Insurance Program, 1990–2011." *American Journal of Agricultural Economics* 95(2): 482–488. doi: [10.1093/ajae/aas091](https://doi.org/10.1093/ajae/aas091).
- Goodwin, B.K. 1993. "An Empirical Analysis of the Demand for Multiple Peril Crop Insurance." *American Journal of Agricultural Economics* 75(2): 425–434. doi: [10.2307/1242927](https://doi.org/10.2307/1242927).
- Goodwin, B.K. 1994. "Premium Rate Determination In The Federal Crop Insurance Program: What Do Averages Have To Say About Risk?" *Journal of Agricultural and Resource Economics* 19(2): 382–395. <https://ideas.repec.org/a/ags/jlaare/30747.html>.
- Goodwin, B.K., and T.L. Kastens. 1993. "Adverse Selection, Disaster Relief, and the Demand for Multiple Peril Crop Insurance." Contract Report for the Federal Crop Insurance Corporation.
- He, J., X. Zheng, R. Rejesus, and J. Yorobe. 2020. "Input Use under Cost-of-production Crop Insurance: Theory and Evidence." *Agricultural Economics* 51(3): 343–357. doi: [10.1111/agec.12558](https://doi.org/10.1111/agec.12558).
- Hojjati, B., and N.E. Bockstael. 1988. "Modelling the De-Mand for Crop Insurance." In P. Harry (ed). *A Collection of Empirical Studies*, Mapp Southern Cooperative Series Bulletin No. 334: Oklahoma State University. Available at: https://pdf.usaid.gov/pdf_docs/pnabg857.pdf.
- Ifft, J.E., T. Kuethe, and M. Morehart. 2015. "Does Federal Crop Insurance Lead to Higher Farm Debt Use? Evidence from the Agricultural Resource Management Survey (ARMS)." *Agricultural Finance Review* 75(3): 349–367. doi: [10.1108/AFR-06-2014-0017](https://doi.org/10.1108/AFR-06-2014-0017).
- Jablonski, B.B.R., J. Hadrich, and A. Bauman. 2022. "The Role of Federal Crop Insurance for Farms and Ranches That Sell through Local Food Markets." *Agricultural Finance Review* 82(1): 113–132. doi: [10.1108/AFR-12-2020-0178](https://doi.org/10.1108/AFR-12-2020-0178).
- Kuethe, T., and M. Morehart. 2012. "The Agricultural Resource Management Survey." Edited by Charles B. Moss. *Agricultural Finance Review* 72(2): 191–200. doi: [10.1108/00021461211250429](https://doi.org/10.1108/00021461211250429).
- Mahul, O., and C.J. Stutley. 2010. *Government Support to Agricultural Insurance. Government Support to Agricultural Insurance: The World Bank*. doi: [10.1596/978-0-8213-8217-2](https://doi.org/10.1596/978-0-8213-8217-2).
- Milliman & Robertson. 2000. *Actuarial Documentation of Multiple Peril Crop Insurance Ratemaking Procedures: RMA Actuarial Methodology Publications*. Available at: https://legacy.rma.usda.gov/pubs/2000/mpci_ratemaking.pdf.
- Möllmann, J., M. Michels, and O. Musshoff. 2019. "German Farmers' Acceptance of Subsidized Insurance Associated with Reduced Direct Payments." *Agricultural Finance Review* 79(3): 408–424. doi: [10.1108/AFR-09-2018-0071](https://doi.org/10.1108/AFR-09-2018-0071).
- O'Donoghue, E.J., M.J. Roberts, and N. Key. 2009. "Did the Federal Crop Insurance Reform Act Alter Farm Enterprise Diversification?" *Journal of Agricultural Economics* 60(1): 80–104. doi: [10.1111/j.1477-9552.2008.00166.x](https://doi.org/10.1111/j.1477-9552.2008.00166.x).
- Petersen, M.A. 2009. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies* 22(1): 435–480. doi: [10.1093/rfs/hhn053](https://doi.org/10.1093/rfs/hhn053).
- Rejesus, R.M., B.K. Goodwin, K.H. Coble, and T.O. Knight. 2010. "Evaluation of the Reference Yield Calculation Method in Crop Insurance." *Agricultural Finance Review* 70(3): 427–445. doi: [10.1108/00021461011088530](https://doi.org/10.1108/00021461011088530).
- Richards, T.J. 2000. "A Two-Stage Model of the Demand for Specialty Crop Insurance." *Journal of Agricultural and Resource Economics* 25(1): 177–194.
- Risk Management Agency [RMA]. 2000. Premium Rate Calculations for the Continuous Rating Model. Available at: http://www.rma.usda.gov/ftp/Continuous_Rating/cr_instr.pdf.

- Risk Management Agency [RMA].** 2009. Rate Methodology Handbook Actual Production History (APH). Available at: <https://www.rma.usda.gov/-/media/RMA/Publications/ratemethodology.ashx?la=en>.
- Roberts, M.J., N. Key, and E. O'Donoghue.** 2006. "Estimating the Extent of Moral Hazard in Crop Insurance Using Administrative Data." *Review of Agricultural Economics*. doi: [10.1111/j.1467-9353.2006.00303.x](https://doi.org/10.1111/j.1467-9353.2006.00303.x).
- Schoengold, K., Y. Ding, and R. Headlee.** 2015. "The Impact of AD HOC Disaster and Crop Insurance Programs on the Use of Risk-Reducing Conservation Tillage Practices." *American Journal of Agricultural Economics* 97(3): 897–919. doi: [10.1093/ajae/aau073](https://doi.org/10.1093/ajae/aau073).
- Shaik, S., K.H. Coble, T.O. Knight, A.E. Baquet, and G.F. Patrick.** 2008. "Crop Revenue and Yield Insurance Demand: A Subjective Probability Approach." *Journal of Agricultural and Applied Economics* 40(3): 757–766. doi: [10.1017/S1074070800002303](https://doi.org/10.1017/S1074070800002303).
- Sherrick, J., B.G.D. Schnitkey, and J.D. Woodward.** 2014. "Crop Insurance Loss Experience, Ratings Changes, and Impacts on Participants." *Agricultural Finance Review* 74(4): 443–463. doi: [10.1108/AFR-09-2014-0027](https://doi.org/10.1108/AFR-09-2014-0027).
- Skees, J.R., and M.R. Reed.** 1986. "Rate Making for Farm-Level Crop Insurance: Implications for Adverse Selection." *American Journal of Agricultural Economics* 68(3): 653. doi: [10.2307/1241549](https://doi.org/10.2307/1241549).
- Smith, V.H., and J.W. Glauber.** 2012. "Agricultural Insurance in Developed Countries: Where Have We Been and Where Are We Going?" *Applied Economic Perspectives and Policy* 34(3): 363–390. doi: [10.1093/aep/pps029](https://doi.org/10.1093/aep/pps029).
- Stock, J.H., and M. Yogo.** 2005. "Asymptotic Distributions of Instrumental Variables Statistics with Many Instruments." *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. doi: [10.1017/CBO9780511614491.007](https://doi.org/10.1017/CBO9780511614491.007).
- Thompson, S.B.** 2011. "Simple Formulas for Standard Errors That Cluster by Both Firm and Time." *Journal of Financial Economics* 99(1): 1–10. doi: [10.1016/j.jfineco.2010.08.016](https://doi.org/10.1016/j.jfineco.2010.08.016).
- Tsiboe, F., and J. Tack.** 2022. "Utilizing Topographic and Soil Features to Improve Rating for Farm-Level Insurance Products." *American Journal of Agricultural Economics* 104(1): 52–69. doi: [10.1111/ajae.12218](https://doi.org/10.1111/ajae.12218).
- Tsiboe, F., and D. Turner.** 2022. *The Causal Dependence of Crop Insurance Demand Response on Premium Subsidies*. U.S. Department of Agriculture, Economic Research Service. https://www.dylanturner.org/img/fcip_demand.pdf.
- Turner, D., and F. Tsiboe.** 2022. "The Crop Insurance Demand Response to the Wildfire and Hurricane Indemnity Program Plus." *Applied Economic Perspectives and Policy*. doi: [10.1002/aep.13314](https://doi.org/10.1002/aep.13314).
- Walters, C.G., C.R. Shumway, H.H. Chouinard, and P.R. Wandschneider.** 2012. "Crop Insurance, Land Allocation, and the Environment." *Journal of Agricultural and Resource Economics* 37(2): 301–320.
- Wang, R., R.M. Rejesus, and S. Aglasan.** 2021. "Warming Temperatures, Yield Risk and Crop Insurance Participation." *European Review of Agricultural Economics* 48(5): 1109–1131. doi: [10.1093/erae/jbab034](https://doi.org/10.1093/erae/jbab034).
- Woodard, J.D., and J. Yi.** 2020. "Estimation of Insurance Deductible Demand under Endogenous Premium Rates." *Journal of Risk and Insurance* 87(2): 477–500. doi: [10.1111/jori.12260](https://doi.org/10.1111/jori.12260).
- Yi, J., H.L. Bryant, and J.W. Richardson.** 2020. "How Do Premium Subsidies Affect Crop Insurance Demand at Different Coverage Levels: The Case of Corn." *Geneva Papers on Risk and Insurance: Issues and Practice*. doi: [10.1057/s41288-019-00144-8](https://doi.org/10.1057/s41288-019-00144-8).
- Young, C.E., M.L. Vandever, and R.D. Schnepf.** 2001. "Production and Price Impacts of U.S. Crop Insurance Programs." *American Journal of Agricultural Economics* 83(5): 1196–1203. doi: [10.1111/0002-9092.00267](https://doi.org/10.1111/0002-9092.00267).
- Yu, J., A. Smith, and D.A. Sumner.** 2018. "Effects of Crop Insurance Premium Subsidies on Crop Acreage." *American Journal of Agricultural Economics* 100(1): 91–114. doi: [10.1093/ajae/aax058](https://doi.org/10.1093/ajae/aax058).
- Zhong, F., M. Ning, and L. Xing.** 2007. "Does Crop Insurance Influence Agrochemical Uses under Current Chinese Situations? A Case Study in the Manasi Watershed, Xinjiang." *Agricultural Economics* 36(1): 103–112. doi: [10.1111/j.1574-0862.2007.00180.x](https://doi.org/10.1111/j.1574-0862.2007.00180.x).