Efficiency Impacts of Utilizing Soil Data in the Pricing of the Federal Crop Insurance Program

Joshua D. Woodard and Leslie J. Verteramo-Chiu

Since the Agricultural Act of 2014, the federal crop insurance program (FCIP) has been the cornerstone agricultural policy in the United States, and is the largest such program globally, with about \$100 billion in coverage annually. Given its scale and scope, the FCIP has the potential to have pervasive impacts on incentives and policy functioning if not designed and priced properly. Surprisingly, soil data are not considered by the government when establishing insurance guarantees or rates. Using soil data that could easily and feasibly be scaled nationally, we find that the pricing differentials caused by the government's failure to handle soil information leads to large errors in rating.

Key words: Federal crop insurance, soil, risk management, insurance rating.

JEL codes: Q14, Q18, G22.

The federal crop insurance program (FCIP) now serves as the cornerstone agricultural policy in the United States and is the largest direct subsidy program to domestic commercial agriculture, with around \$10 billion in expected taxpayer costs annually on \$100 billion dollars of coverage (Woodard 2016). The program is priced, regulated, and administered by the federal government via the Risk Management Agency (RMA) of the USDA, and delivered by private companies. Risk management programs such as federal crop insurance will likely only become more important in coming decades as farmers and food security are faced with increased risk from market volatility and climate change (Battisti and Naylor 2009; Lobell et al. 2011; Lobell, Schlenker, and Costa-Roberts 2011; Lobell, Sibley, and Ortiz-Monasterio 2012; Lesk, Rowhani, and Ramankutty 2016). These programs also have the potential to have pervasive impacts on conservation outcomes if not designed or priced appropriately, but may enable sustainability objectives if properly structured (Woodard et al. 2012a). Surprisingly, the government does not utilize soil data in designing products, rates, or setting guarantees.

There is a large body of literature on the modeling of crop yield distributions in the agricultural economics and actuarial literature (see, e.g., Woodard, Sherrick, and Schnitkey 2011; Woodard and Sherrick 2011), but due to data limitations, most studies tend to lack explicit consideration of soil and site-specific data on policy and insurance design, with few exceptions (e.g., Woodard 2014). While evaluating the effects of soil on crop growth is, of course, very common in crop sciences on small scales and in trial work, very little has been done on integrating these data and approaches for the purposes of large-scale insurance estimation in public policy contexts, with few exceptions (e.g., Woodard 2015a). Legally, each insurance contract (or "plan" of insurance) sold should be priced so that the amount indemnified in expectation

Joshua D. Woodard is the Zaitz Family Faculty Fellow of Agribusiness and Finance and an assistant professor in the Dyson School of Applied Economics and Management, and Leslie J. Verteramo-Chiu is a postdoctoral research assistant, both at Cornell University. This work was partially supported by the Meridian Institute. Woodard is the founder of Ag-Analytics.org, is an expert reviewer for the Federal Crop Insurance Corporation Board and United States Department of Agriculture, serves on the AGree Conservation and Crop Insurance Task Force (CCIT), and has developed crop insurance rating systems for several products currently sold in the market. Parts of this work were presented in September 2015 at a joint USDA/ERS-Farm Foundation Soil Health Workshop. Woodard conceptualized and executed the analysis, and drafted the manuscript and revisions. Verteramo-Chiu assisted with tabulating county-based model results. Woodard takes responsibility for all views and opinions expressed. Correspondence may be sent to: jdw277@cornell.edu.

Amer. J. Agr. Econ. 99(3): 757–772; doi: 10.1093/ajae/aaw099
Published online March 28, 2017

equals the premium set by the government. Historically, this has been interpreted to imply fairness to the individual producer to the extent reasonably possible given available data and intelligence. Nevertheless, attaining this type of pricing efficiency via a government agency—as many have pointed out—is likely an unrealistic task for a government agency, and is likely more suited for market determination (not only in crop insurance, but also generally; see, e.g., Priest 1996; Cummins 2006; Jaffee and Russell 2006; Michel-Kerjan and Kousky 2010).

Instead of using soil data to determine baseline insured yield levels and premium rates, the government's methodology relies on a measure of average historical yields that does not account for the number nor specific years of production reported, the weather in those years across different farmers' policies, nor even the fields being insured. Thus, the government's method does not reflect full information regarding soils, or, for example, when a producer adds or removes new land from an insured unit. The 2008 Farm Bill included provisions for the RMA and participating companies to begin collecting Common Land Unit (CLU) data for each policy and insured unit, and thus the data needed to operationalize using soil data in the rating would be feasible at reasonable cost as these data are already reported (just not used in ratings). Ignoring soil type could result in mispricing of the underlying insurance and misalignment of incentives.

The purpose of this study is to investigate the feasibility of using high resolution soil data for the modeling of crop insurance guarantees in large-scale contexts for this cornerstone agricultural program, and the implications of omitting them relative to current government methods (this is a first, to the best of our knowledge). The soil data and quality indexes employed must be nationwide and have high availability in order to be scalable and operational in practice; thus, we employ the SSURGO soil dataset from National Resources Conservation Service (NRCS), which is high resolution, nationwide soil type data (of which there are several thousand soil types). Fifty-seven different soil quality attributes such as available water storage, soil organic carbon, and other aggregate soil quality indexes are matched with the soil type data to map soil types into quantifiable indexes in estimating field-level yield guarantee models.

We conduct several analyses utilizing a uniquely constructed dataset that includes multiple yield datasets (county, farm, and psuedo-farm), as well as high-resolution weather, soil, crop cover, and modeled insurance rates according to the RMAs published methodology and parameters in order to investigate the fundamental rate bias and efficiency impacts of omitting soil information. In the first set of analyses, soil conditional expected yield models are estimated and downscaled, then calibrated back to published RMA farm-level insurance rates in order to calibrate and simulate psuedo-farm level yields on different soils, which are then used to estimate the distribution of impacts on rating errors stemming from omitting soil information. This allows for estimation of how omitting soil information in determining *baseline insurable yields* (i.e., "guarantees") flows through to affect pricing accuracy. We also perform a second set of confirmatory analyses using actual farm-level data from the Illinois Farm Business Farm Management (FBFM) dataset, which contains matched soil productivity measures, to evaluate the degree to which incorporating soil data improves guarantee determination and subsequently rating efficiency. Heterogeneity in soil-conditional risk and intra-county premium rate differentials may also, in reality, vary with soil type (see Woodard 2015a). Thus, the true premium error in existing FCIP rates from omitting this information is likely larger than that presented here. We also purposely focus on a region that has quite high soil homogeneity, as well as no designated "High Risk" land. Thus, this is the region *least* likely to find significant efficiency gains from including soil information.

Results indicate that pricing efficiency could be significantly improved in this program by taking into account soil data explicitly when estimating crop insurance guarantees. This could lead to savings for taxpayers, fairer premiums for lower-risk farmers, less risky underwriting exposure for companies and the taxpayers, and possibly better environmental outcomes. This framework could be applied and expanded in a fairly straightforward manner to other field-level yield databases (which, since 2009 the RMA has started to collect, but does not utilize) to operationalize in practice. Before it will be feasible to operationalize the adaptation of the FCIP to more properly accommodate emerging conservation practices, having rating systems in place with soil information at the foundation will be necessary and critical. That said, it should be recognized that the FCIP interacts with other policies (e.g., Commodity Title programs that are insurance-like) and markets, not to mention the Standard Reinsurance Agreement. Thus, one might argue that future changes to the rating system to incorporate soil data should be carefully thought out and considered in close concert with stakeholders, delivery companies, and policy-makers.

Data and Methods

Two primary analyses were conducted. In the first, a pseudo-farm yield model is constructed that allows for the explicit analysis of actual GIS-level field boundary and soil data. The second was a confirmatory analysis using farm-level yields from the FBFM dataset which also include soil information. For the first analysis, yield data were collected from the National Agricultural Statistical Service (NASS) for corn, and span thirty-nine years from 1975 to 2013 for eight states: Illinois, Indiana, Iowa, Michigan, Minnesota, Missouri, Ohio, and Wisconsin. These states were selected because they are the largest corn-producing states by total production, form a contiguous area, and have similar crop-growing seasons (C=734 counties, N=26,755 observations, unbalanced panel). Growing season weather data were obtained from the PRISM Climate Group at Oregon State University. PRISM data are gridded 4km resolution data, which we aggregate to obtain county averages. Weather values for July and August are employed as explanatory variables, as these months are critical for corn production in this region. These serve as reasonable proxies for this analysis.

For the second analysis (N=16,173), farms with *at least* twenty years of data from five Central Illinois counties for Corn were selected for analysis from the FBFM dataset, for 1972–2007 (F=612 farms). Accompanying average summer Palmer Drought Severity Z-Index (*PDSI*) data were employed as a weather proxy. The data have side-by-side soil index productivity (*SPR*) data, which are a mapping from the University of Illinois Bulletin 810 and Circular 1156 reports to NRCS SURGGO on soil type data (see Woodard 2014, and Woodard 2015a for an explanation of these indexes and their mapping to NRCS soil type data). The second dataset is similar to that used in Woodard (2014).

This study employs the last published Common Land Unit (CLU) field maps in the public domain (last released in 2008 by the Farm Service Agency). While the exact field boundaries change somewhat through time as Insured Units are combined, sold, added, or broken out, we use them here to approximate the degree of spatial heterogeneity in soils across fields within a county. The analysis could easily be replicated with updated CLU maps. Since the 2008 Farm Bill, RMA began directing participating companies (so-called Approved Insurance Providers, or AIPs) to collect CLU-specific yield and insurance policy data. Prior to that, data in the crop insurance program were only identified by the county in which the policy was located. Thus, the analysis contained herein should be fairly extensible to internal RMA field-level data currently being collected.

Soil type data layers were obtained from the National Resources Conservation Services (NRCS) SSURGO database. The NRCS publishes a gridded version of this dataset at thirty-meter resolution, along with the VALU1 table, which links soil type to fifty-seven different soil quality indexes. These soil variables are generally stated for different soil layers (depths) for different soil components, including soil organic carbon (*SOC*), average water storage (*AWS*), and thickness of soil components. There are also indexes for the National Commodity Crop Productivity Index (*NCCPI*) for various crops, Drought Vulnerable Soil Landscapes (*DROUGHTY*), Potential Wetland Soil Landscapes (*PWSL*), and root zone depth available water storage (*rootz*). We spatially aggregate over the areas of interest to obtain fifty-seven average soil quality estimates at both the county and CLU levels of aggregation for modeling. While we acknowledge that the resolutions at which the data are published (ten to thirty meters) may lend themselves to a mild degree of inaccuracy, this study evaluates at the Common Land Unit (roughly speaking, field-boundary level) which typically consists of several acres and has a reasonable scientific basis (NRCS 2015). To ensure that the soil averages for the counties were constructed only for soils upon which corn is typically grown, we filtered the

SSURGO data by the NASS Cropland Data Layers (CDL), averaging only over areas grown to corn. The CDL layers provide thirty-meter resolution estimates of crop cover type. All processed data and sources are freely available from the Ag-Analytics open-data platform (Woodard 2015b).

The *SOC* is an important determinant in soil fertility. Further, *rootz* is the volume of plant-available water that the soil can store within the root zone and is also an important determinant in yield potential. Since most variables within each group capture the same factors, they are expected to be highly correlated and may pose problems if included in the regression. We selected the variables that best explain crop productivity by running successive regressions of yields on each soil variable and evaluating the size of the average effect. We selected two variables for further analysis below with the largest average effects from each class: *SOC*, measured in g C/m^2 in standard layer 3 (twenty to fifty cm depth), and *rootz*, expressed in mm. These variables are both positively correlated with crop yields, and have a correlation of about 0.34 between them. We also generate results below using the *NCCPI* (corn and soybeans)—which is an aggregate proxy for soil quality—for comparison. Note that at the 10m resolution, the SSURGO soil data may be subject to some false precision, but the CLUs are typically several acres, which is less of a concern. Having a system in place to update soil-type data through the administration of the FCIP could be advantageous at some point, but would be a later step. As noted, this study deals with soil type, which in reality should be static. Updating these data to reflect any corrections in the soil-type maps could in the future perhaps be facilitated in the course of delivering federal crop insurance. The benefits of incorporating other measures of soil quality and their collection should be evaluated against their costs.

Several regression models of county yield on soil and weather are estimated in tables 1–3. Models are estimated with state-level fixed effects for both intercept and trend terms. Models were also estimated using a subset of the data for Illinois, Indiana, and Iowa only for robustness (tables 1–3). Several variants of the models are investigated, including those with and without weather/soil interactions, as well as with different subsets of soil variables for robustness. Explicitly, the yield model estimated is

$$(1)\quad Y = \sum_{state=1}^{8} D_{state}\delta_{state} + \sum_{state=1}^{8} TimeD_{state}\beta_{state} + W\beta_{weather} + S\beta_{soil} + \varepsilon$$

where **Y** is a vector of average county corn yields (bu./acre), D_{state} is a vector of state dummy variables, *Time* is a time trend vector, *W* is a matrix of weather variables (temperature, accumulated precipitation, and its quadratic terms, for a critical period of the crop), *S* is a matrix of soil variables, and *Xws* is a matrix of the interaction between some weather and soil variables (a subset of *W* and *S*). Further, δ_{state}, β_{state}, $\beta_{weather}$, β_{soil}, and β_{ws} are their respective estimated coefficients, and ε is a vector of error terms $\sim N(0, \sigma)$. Alternative models presented are subsamples of equation (1).

The farm-level FBFM yield models for the second analysis are similar in that they employ county-level fixed effects and county-level trends, as well as the *SPR* measure, as well as *PDSI* and squared terms. We also attempted to estimate the models using farm-level trends and intercepts, but the system was singular. This is not surprising since the soil-type data accounts for a very large and significant amount of the variation in farm performance within a county. In practice, an analyst would not have farm-level intercepts in any case, so the chosen setup is preferred for comparison to feasible operational models.

Insurance Analysis and Simulation Design

Pseudo-Yield and Ratemaking Analysis

First we outline the pseudo-farm yield and rate analysis. Using the estimated yield model that employs the NCCPI above, field-level expected yields for McLean County were estimated for all CLUs by downscaling the model estimated with NASS county data. Since this model is used to estimate *expected* yield conditional on soil, these models, which are conditional on soil, can be downscaled to the field level without bias for a linear model (this would not necessarily be the case with

Table 1. Yield Regression Results, State-Level Fixed Effects, Soil, and Weather Controls

Variable	Coefficient	T-value	Coefficient	T-value	Coefficient	T-value
IL	−456.883	−16.26	−436.836	−16.29	−434.083	−16.18
IN	−463.182	−16.47	−439.837	−16.39	−433.746	−16.15
IA	−471.315	−16.77	−455.120	−16.97	−446.69	−16.65
TimeIL	1.815	67.81	1.813	71.02	1.802	70.53
TimeIN	1.626	57.49	1.631	60.45	1.625	60.17
TimeIA	2.054	76.58	2.052	80.23	2.048	80.00
Temp	50.999	21.12	46.629	20.22	44.919	19.43
Prec	0.732	42.09	0.748	45.05	0.747	44.94
Temp2	−1.252	−24.16	−1.163	−23.50	−1.129	−22.75
Prec2	−0.0027	−38.86	−0.0027	−41.65	−0.0027	−41.92
SOC	0.00198	19.92	0.00153	16.02		
Rootz			0.1503	33.50		
NCCPI					80.1144	39.12
Adj.R^2	0.659		0.690		0.689	
SSR	345.66		314.41		315.08	
AIC	66,035		64,967		64,989	
DW	1.46		1.59		1.57	
N = 11,293						

Table 2. Yield Regressions: Three States' Fixed Effects with Soil and Weather Interactions

Variable	Coefficient	T-value	Coefficient	T-value	Coefficient	T-value
IL	−500.4923	−13.64	−442.8347	−11.51	21.759	10.33
IN	−503.9415	−13.75	−442.9726	−11.52	31.091	16.31
IA	−518.2541	−14.12	−455.2866	−11.83	15.025	6.78
TimeIL	1.8208	71.78	1.7979	70.47	1.621	48.00
TimeIN	1.6572	61.60	1.6405	60.69	1.500	41.87
TimeIA	2.0272	79.50	2.0394	79.70	1.984	58.57
Temp	47.9908	16.78	43.5799	16.06		
Prec	1.0487	35.42	0.9909	26.11		
Temp2	−1.1617	−19.88	−1.1008	−21.57		
Prec2	−0.0026	−39.39	−0.0027	−40.40		
*Temp*Soc*	−0.0001	−1.37				
*Prec*Soc*	0.0000	−4.25				
*Temp*rootz*	−0.0044	−1.52				
*Prec*rootz*	−0.0012	−10.38				
SOC	0.0044	2.94				
Rootz	0.3704	5.13				
*Temp*NCCPI*			0.0533	0.04		
*Prec*NCCPI*			−0.3533	−7.11		
NCCPI			113.5153	3.75	102.144	37.87
Adj.R^2	0.6947		0.6909		0.4469	
SSR	309.79		313.63		561.19	
AIC	64,808		64,941		71,499	
DW	1.58		1.57		1.76	
N	11,293		11,293		11,293	

conditional variance). Note that the models above use the published NCCPI and other soil characteristic variables as published in the VALU1 table by NRCS, which link to soil *type*. These are indicative of soil quality so far as soil type goes, but soil type does not change through time. The maps may be periodically updated, but these measures are linked to soil type, and are not dynamic measures of any other soil quality

Table 3. Yield Regressions: Eight States with Soil and Weather Variables Interactions

	Model 7		Model 8	
Variable	Coefficient	T-value	Coefficient	T-value
IL	−524.8002	−44.46	−457.663	−41.07
IN	−527.5552	−44.58	−458.311	−40.91
IA	−542.0921	−46.01	−468.923	−42.05
MI	−547.8125	−46.95	−473.706	−42.79
MN	−564.3944	−48.05	−482.413	−43.33
MO	−536.5075	−45.61	−466.259	−41.86
OH	−531.3276	−44.83	−468.267	−41.77
WI	−542.5202	−46.32	−469.348	−42.22
TimeIL	1.813	66.37	1.7913	66.40
TimeIN	1.641	56.68	1.6184	56.75
TimeIA	2.0455	74.43	2.0347	75.18
TimeMI	1.6838	50.77	1.5679	47.93
TimeMN	2.3631	76.92	2.3312	77.03
TimeMO	1.6033	58.01	1.6349	60.07
TimeOH	1.6355	54.96	1.6393	55.93
TimeWI	1.7391	51.48	1.6693	50.06
Temp	50.4064	49.95	39.198	39.25
Prec	0.9192	47.55	0.773	41.54
Temp2	−1.1914	−51.88	−0.8181	−34.76
Prec2	−0.0026	−48.74	−0.0026	−48.80
*Temp*SOC*	−0.0004	−10.67		
*Prec*SOC*	0.00000	−3.92		
*Temp*Rootz*	0.0035	2.52		
*Prec*Rootz*	−0.0007	−8.53		
SOC	0.0119	13.72		
Rootz	0.1103	3.31		
Temp NCCPI*			−11.7458	−25.8865
Prec NCCPI*			−0.0771	−3.0224
NCCPI			338.1252	32.1762
Adj. R2	0.6756		0.6847	
SSR	363.82		353.60	
AIC	157,820		157,050	
DW	1.37		1.3754	
N	26,755		26,755	

Table 4. Rate Error Simulation Results (Percentiles) by Field, McLean County Illinois

	Percentile						
	5th	10th	25th	50th	75th	90th	95th
Own-Field, Relative	0.417	0.526	0.745	1.047	1.426	1.851	2.153
Mixed-Field, Relative	0.352	0.528	0.958	1.696	2.927	4.880	6.661
Own-Field, Nominal	−0.020	−0.015	−0.007	−0.001	0.004	0.008	0.010
Cross-Field, Nominal	−0.105	−0.070	−0.034	−0.012	0.001	0.008	0.011

components. Next, actual RMA rates were collected for each field based on the expected yield to perform a simulated rating analysis. The RMA rating methodology assigns a premium rate within a county depending on the Actual Production History (APH) level (which is a proxy of expected yield).

As it regards APH, to the extent that the expected yield is different conditional on soil type, then when planted area and CLUs

Table 5. Illinois FBFM Farm Yield Regression Results

Variable	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
SPR	1.081	36.956	1.076	40.381	4.470	8.716	4.411	9.446
SPR^2					−0.020	−6.618	−0.020	−7.153
PDSI	7.920	59.573	9.402	75.976			9.398	76.060
PDSI^2			−2.809	−57.721	7.916	59.623	−2.808	−57.796
INT1	−3,875.850	−43.743	−3,684.195	−45.625	−4,011.285	−44.166	−3,817.506	−46.133
INT2	−3,444.345	−42.082	−3,238.909	−43.409	−3,577.795	−42.496	−3,370.270	−43.925
INT3	−3,619.601	−46.909	−3,335.856	−47.363	−3,748.730	−47.159	−3,462.986	−47.743
INT4	−4,197.899	−38.852	−3,944.480	−40.053	−4,338.730	−39.449	−4,083.116	−40.741
INT5	−3,882.385	−39.234	−3,620.300	−40.128	−4,032.696	−39.770	−3,768.265	−40.772
TREND1	1.969	44.286	1.876	46.312	1.966	44.275	1.873	46.310
TREND2	1.751	42.636	1.651	44.100	1.748	42.595	1.648	44.064
TREND3	1.842	47.550	1.703	48.164	1.836	47.438	1.697	48.053
TREND4	2.134	39.347	2.010	40.662	2.133	39.387	2.009	40.714
TREND5	1.973	39.723	1.845	40.743	1.977	39.864	1.849	40.902
R-Sq	0.462		0.554		0.463		0.555	
Sigma-Sq	637.964		528.946		636.279		527.308	
DW	1.194		1.173		1.196		1.176	
N	16,172		16,172		16,172		16,172	

Note: The variables *TREND1-5* and *INT1-5* are county-level trend and fixed effects for La Salle, Livingston, McLean, Marshall, and Woodford counties, Illinois; *SPR* is the soil productivity rating, *PDSI* is the Palmer Drought Severity Z-index.

Table 6. Simulation Results: Loss Cost Ratio Error, Conventional vs. Soil Conditioned using FBFM Corn Data, 85% Coverage

	MSE		Relative Error		95th Percentile Error		
County	APH	Soil APH	APH	Soil APH	APH	Soil APH	Avg. LCR
La Salle	0.80%	0.33%	29.64%	12.29%	195.41%	133.32%	2.69%
Livingston	0.71%	0.26%	38.25%	14.10%	215.15%	137.87%	1.85%
McLean	0.79%	0.32%	32.56%	13.03%	197.98%	135.58%	2.43%
Marshall	0.88%	0.35%	31.20%	12.43%	190.09%	133.31%	2.81%
Woodford	0.94%	0.35%	34.57%	12.94%	202.74%	135.55%	2.72%
All 5 Counties	0.81%	0.32%	33.27%	13.03%	200.27%	135.12%	2.44%

Note: Results by county of Mean Square Error (MSE) for conventional loss cost ratio (LCR) and soil conditioned LCR. The third column is the average expected LCR for each county and the full sample, measured in percentage of liability for corn farms (F=612) with at least twenty years of data from 1972–2007 from the FBFM dataset. The 500 bootstrapped sample trials are executed for each farm, and then the average LCR calculated under the sampled APH and compared to the true LCR as approximated by the burn rate LCR for each farm. Six years of data are drawn for the APH calculation. Relative MSE is also reported as MSE of the LCR divided by the expected LCR. The 95th Percentile Error is the 95th percentile of the ratio of the simulated expected LCR under the respective APH method to the true LCR for the farm, across all bootstrapped samples and farms.

within a given Insured Unit through time do not include some measure of soil, then current methods will lead to biased APHs. This will lead to biased rates through at least two avenues.[1] The first is via the granting of artificially high or low guarantees, which effectively lead to higher or lower effective coverage levels versus nominal coverage levels, upon which rates are based; that is, the liability will be biased.

Since the rate within a county is then benchmarked against APH (we simply refer to this published schedule relationship between rate charged and APH as the "rate curve"—which is understood in this case), there is also additional rate bias admitted through this route (the rate "looked up" on the schedule conditional on the APH will be biased). Additionally, if fields/CLUs constituting an Insured Unit are mixed through time, or if APHs are generated in small samples, then additionally both the "rate curve" schedules

[1] It is also well-known that yield trends create bias in guarantees (Skees and Reed 1986), and the Trend APH Endorsement was introduced in recent years to address this source of bias.

and the APHs generated could be made more efficient by incorporating soil data. Thus, current methods excluding soil data are both biased and inefficient, with respect to both determination of premium rates and determining liability/effective coverage. The first analysis proposed below using the downscaled soil conditional models (estimating appropriately using county data) addresses only the errors in rates admitted through the APH determination, and thus are conservative.

The mean RMA quoted rate (which we take as true for the purposes of the counterfactual analysis) was 0.0173 for this analysis, with a standard deviation of 0.00089. Using those rates, Weibull distribution parameters were backed out for each field by estimating the parameter values that would result in the premium rate quoted by RMA. In general, the actuarially fair premium rate is set equal to the expected loss cost ratio, *E(LRC)*:

$$(2)\quad E(LCR) = AF\ Premium\ Rate = \int_0^{E(Y)\cdot Cov} Max(0, APH\cdot Cov - y)\cdot f(y)dy/(APH\cdot Cov)$$

where *Cov* is the coverage level election (or one minus the deductible percent), $f(y)$ is the yield distribution, *y* is the yield outcome, and *APH* is expected yield; in this context, the product of $APH \cdot Cov$ is the insurance guarantee. Simply put, the actuarially fair premium rate, or expected loss cost, is the expectation of indemnities (or losses) weighted over all possible yield outcomes as a fraction of the amount of coverage (or guarantee). This analysis focuses on results for the 85% coverage level (the highest election available), although the same contextual results hold for any coverage level.

Farm-Level Rating Analysis

We also conduct a confirmatory analysis using the farm-level yield data and model from the FBFM dataset. The rating efficiency analysis proceeds as follows. We conduct a simulation whereby the effective rate error—as caused by inefficient guarantee determination from omitting soil information or not—is inferred by comparing the following: 1) no soil data, where 500 bootstrap samples from each farm are conducted from the original farm yields (detrended to reflect the Trend Yield endorsement, approximately) in order to calculate APH, and then expected (empirical) loss costs under that APH are calculated for each trial and farm; 2) including soil data, whereby using the estimated soil conditioned farm-yield models, a soil-adjusted APH is calculated for each farm for 500 samples. We sample out of the farms error distribution when constructing this soil conditioned APH in order to reflect the normal sampling error that would still be present in reality if soil-conditioned APH models were employed. These are again compared to the empirical loss rates for each farm to evaluate the distribution of pricing errors.[2]

Yields for 2015 are calculated at the mean of the PDSI value (fundamentally equal to zero for the Z-index). In both cases, we sample six years of detrended yields (or equivalently, six years of residuals and construct detrended yields) to serve as the basis for calculating the alternative APH measures (traditional or soil-conditioned), which are then compared to what expected loss costs ratios (LCRs) would have been generated in both cases. The LCR is the expected indemnity divided by the liability, as indicated in equation 2.

Results

Pseudo-Yield and Ratemaking Analysis

Figure 1 displays the average National Commodity Crop Productivity Index (NCCPI) for corn and soybeans by county. Soil quality is highly spatially correlated across counties and varies widely across the United States, from 0.15 (5th percentile of corn-producing counties) in marginal counties regions to 0.75 (95th percentile) in intensive production regions in the heart of the Midwest. The standard deviation across counties is about 0.187. Not surprisingly, soil quality tends to match the density of planted acres closely. Figure 2 displays the NCCPI for McLean County, Illinois (the largest county by acreage in the state in one of the most productive areas) by Common Land Unit (CLU) field boundary. Within the

[2] We detrended the yields to net out/remove any bias from yield trends. This handicaps the RMA rating system favorably relative to not making this adjustment. Most farms in this region take the APH Trend Yield endorsement, so this is appropriate.

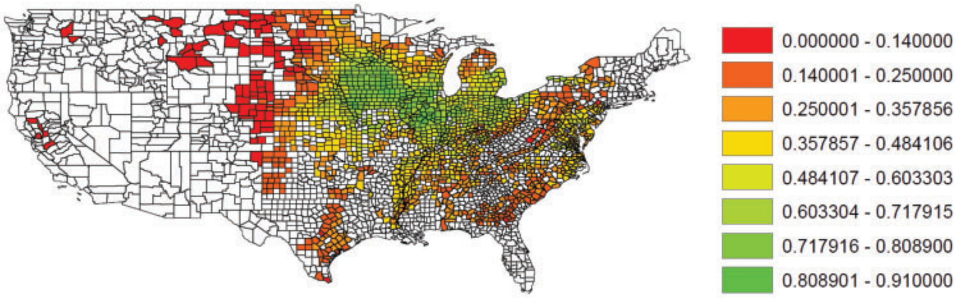

Figure 1. National commodity crop productivity index, corn, United States, by county

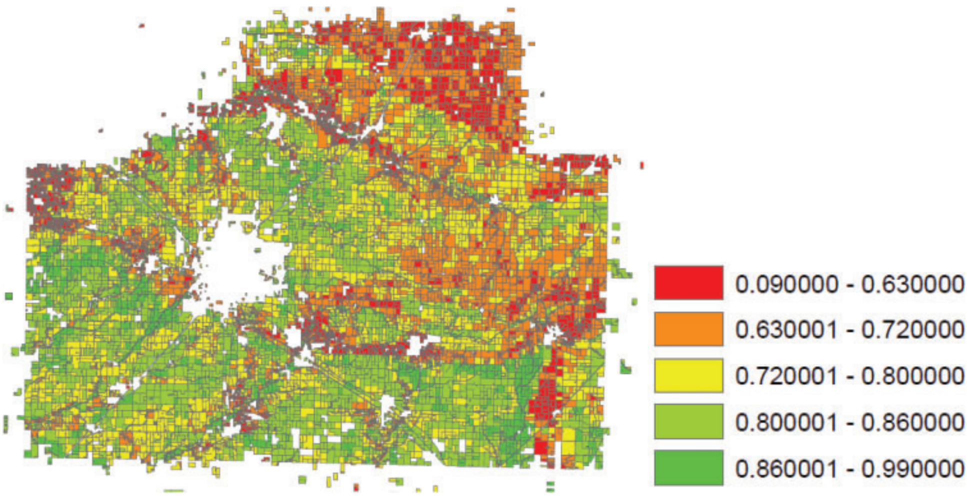

Figure 2. National commodity crop productivity index by field, McLean County, Illinois

county, substantial variation in soil quality is still very apparent; the average NCCPI value across all fields in McLean is 0.73, and the standard deviation is 0.139 (0.443 and 0.896 at the 5th and 95th percentiles, respectively).

The crop insurance program does not track individual field location in determining guarantees, but rather uses an average of between four and ten years of producer data (depending on how much is available); this average does not take into account which years are reported, and so two fields that have very similar soil could get much different guarantees simply by virtue of which years are reported. Likewise, it is not unusual for lower-quality fields to obtain higher guarantees than higher-quality fields simply because of which years were reported. How overall quality of soil within a policy changes from adding or removing land within a unit is also not factored into the guarantee.

To evaluate how this intra-county soil variation translates into expected errors in determining guarantees, and resultant impact on rates, we estimate soil conditional expected yield models using nationwide yield data. Regression results for a model using data from the Central Corn Belt states are displayed in table 1. Tables 2 and 3 display supplementary regression results to evaluate robustness of the yield models, including models that incorporate soil and weather interactions for the three Central Corn Belt states of Illinois, Indiana, and Iowa, models for the larger eight-state region (state fixed effects, including soil and weather controls), and the eight-state region also including weather and soil interactions. The results across all models were fairly consistent, and using any of the below models did not affect the qualitative results of the main rate analysis.

The soil parameters are economically and statistically significant. The first model uses *Soil Organic Carbon*, and the second model includes *Root Zone Available Water Storage.* These variables represent soil quality proxies. The third model uses the NCCPI. The measure of *Soil Organic Carbon* was taken for the layer between twenty and fifty cm, and corresponds to units of grams of carbon per square meter (NRCS 2015). This layer provided the best response to corn yields, although NCCPI reveals similar (stronger) results. Weather controls and state-level fixed effects were also taken into account and found to be significant.

Simulation experiments were conducted using the estimated yield models—downscaling to the field level, which can be conducted without bias in the *mean* yield—for two different cases (*own-field* and *mixed-field*, defined below). After downscaling to estimate the expected yield, we used the RMA rate to "back out" what the farm-yield distribution would have been that would have corresponded to that published rate, as well as expected yield, to develop a reasonable set of farm-level distributions as implied by RMA average rates. A Weibull distribution is assumed (Woodard, Sherrick, and Schnitkey 2011). We solve for each field what the parameters of the corresponding Weibull distribution would have been, subject to the mean being equal to that which results from the soil models, and with a level of risk such that the expected rate equates to the actual unsubsidized RMA rate. We then simulate out of these distributions for each field.

In the first experiment (***own-field***), between four and ten yields were simulated for each field from each field's own estimated yield distribution. The guarantee (so called, APH) and expected loss rate was then calculated by integrating over the yield distribution. This design allows for the evaluation of the likely inefficiency that results from the small-sample nature of determining guarantees and ignoring soil information. If soil information were incorporated, then presumably this would significantly improve the estimation of expected yield (which APH is meant to proxy). In the second experiment (***mixed-field***, to account for the fact that in reality not only is soil taken into account, but the APH is not necessarily pegged to an individual field), we first select at random a field in the county, then randomly draw between four and ten years of yield data to determine APH; we then impose that APH on another field at random, and calculate the expected loss rate under that field's yield distribution. Each experiment was run for 3,524,500 iterations (500 iterations per field for 7,049 fields).

Note that the *own field* results basically show what the pricing inefficiency is that results simply from sampling APH for small samples *and* ignoring soil information—even if soils did not vary at all within the county—while the *mixed field* results are indicative of additional inefficiency that is brought about by the fact that soils vary within the county and that the Insured Unit is comprised of multiple CLUs that are mixing and changing in composition through time. The issue under consideration here is whether yields vary enough within a county based on soil and other sampling noise to substantiate an effort that includes soil. A prerequisite for that is that the inefficiency in rate determination based on APH variability be of a magnitude fundamentally greater to warrant such intervention. If soils do not vary much in a county, then the sampling error impact on APH (and in turn, pricing inefficiency) will be small, then *mixed field* results will be suitably similar. If not, then they can be much different. The difference is indicative of what types of efficiency gains might be expected if soil information was employed.

Table 4 presents results of the rate simulation analysis. Figure 4 reports kernel densities of the simulated relative pricing error multiples for experiments 1 (*own-field*) and 2 (*mixed-field*). Relative and nominal rate errors are displayed by percentile for all iterations. Nominal rate error was calculated as the quoted rate minus the true expected loss rate once inefficiency in the APH measure from omitting soil is accounted for; a value of less than zero indicates underpricing, while a value greater than zero indicates overpricing. Relative rate error was calculated as the expected loss rate divided by the quoted rate; a value of less than one indicates overpricing, while a value greater than one indicates underpricing. The standard deviation of *own-field* and *mixed-field* nominal rate errors was 0.0095 and 0.0203, respectively, while the coefficient of variation of *own-field* and *mixed-field* nominal rate errors was 54.7% and 113.7%, respectively. The results show the potential for substantial mispricing to stem from inefficient guarantee

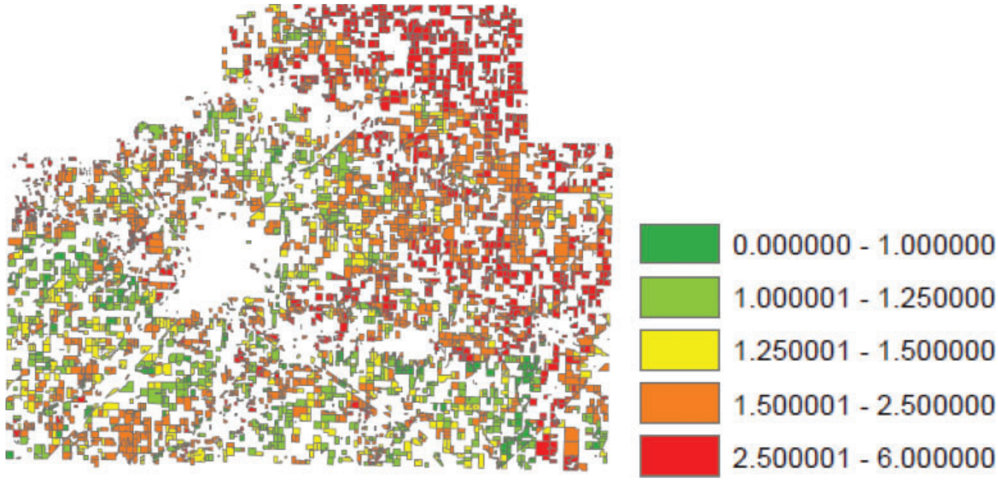

Figure 3. Average pricing error multiple (*Mixed Field*), by field, McLean County, Illinois

determination. In 25% of all simulated cases (assuming the APH is not mixed across fields), the expected loss rate to the charged premium rate would be less than 0.745, or the equivalent of an overcharge in premium of approximately 34%. The lowest-risk cases (10% of cases) are approximately double-charged. On the other hand, for the riskiest 10% of cases, the policies would pay out 1.851 times the premium charged, meaning that the premium would need to be 185.1% of the current value charged to be actuarially fair.

Figure 3 displays the average relative pricing error *by field*. Only fields that the 2013 Cropland Data Layer maps indicate as having a majority of corn were considered for the analysis (although results were consistent regardless). Note that values of less than one indicate that on average, the field will be overpriced and vice-versa for values greater than one. Comparing to the soil quality map in figure 2, it is clear that higher-quality fields will be overpriced substantially, and lower-quality fields will be underpriced. The rate error multiples are also large. For example, the lowest one-third of fields by soil quality had estimated rate error multiples of 2.349 or greater, meaning that the expected loss on the field under the RMA's rating methodology would result in indemnities about 2.3 times greater than premium. On the other hand, the highest 10% quality fields would, on average, be overpriced. Note that this analysis is done under the conservative assumption that RMA rates are, for a given true expected yield level, in fact correct. To the extent that the conditional loss rate is more responsive across APH levels than indicated by RMA's methodology, the rate error levels could be even higher.

Farm-Level Rating Analysis

Results for the confirmatory farm-level analysis are presented in tables 5–7. Table 5 presents results for the farm-level yield regressions. The models cover La Salle, Livingston, McLean, Marshall, and Woodford Counties in Illinois. Soil productivity rating is denoted as *SPR*, while *PDSI* is the Palmer Drought Severity Z-index. All soil and weather variables had the expected and plausible signs, and are consistent with earlier research that used different functional forms (Woodard 2014). The rating analysis results are presented in table 6, and present the Mean Square Error (MSE) of the expected loss ratio by farm for both the conventional loss cost ratio (LCR) and soil conditioned LCR. The third column is the average expected LCR for each county and the full sample, measured in percentage of liability for corn farms. A total of 500 boostrapped sample trials were executed for each farm, and then the average LCR calculated under the sampled APH and compared to the "true" LCR as approximated by the burn rate LCR for each farm. Relative MSE is also reported, which equals MSE of the LCR divided by the expected LCR. The 95th Percentile Error is the 95th percentile of the ratio of the simulated expected LCR under the respective APH method to the "true" LCR for the farm, across all bootstrapped samples and farms. Overall, the

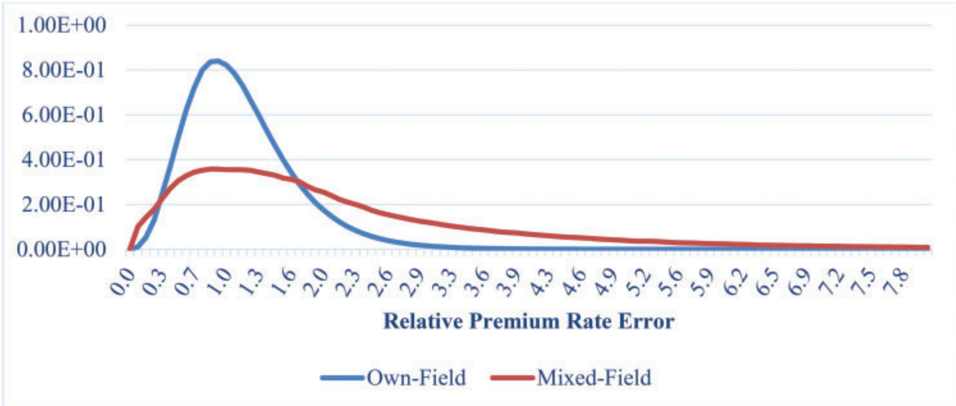

Figure 4. Kernel density, simulated relative pricing error multiples, McLean County, Illinois

results are fairly consistent in magnitude to the pseudo-yield analysis. The MSE rate for the conventional APH method is about 2–3 times that of the soil conditioned rate. For example, for Marshall County the MSE of the conventional method versus the soil-conditioned method is 0.88% vs. 0.35%. This is a fairly significant amount of normal error, given that the expected premium rate in the sample was 2.4%. Thus, the relative error, on average, is about 33% under current RMA methods, but would only be about 13% if soil were considered. We also evaluated the percentiles of rate error and compared across the two. About 5% of the cases under the conventional method result in expected loss costs that are double their RMA rate (>100% of the premium); on the other hand, when considering soil data, the worst 5% of cases only experience about a 35% error in rates. Compared with the earlier results, they are fairly consistent in that when ignoring soil data, it is not uncommon to have cases in which there is a doubling or more in expected loss costs over RMA quoted rates.

Table 7 presents the APH results for both the soil-conditioned and naive methods. The MSE of conventional APH is around 10 bu./acre in this region (mean yields of approx. 180 bu./acre), while for the soil-conditioned APH, the typical error is half of that. We also saw from the previous table that this translates into substantial rate bias.

Table 7. Simulation Results - APH Error, Conventional vs. Soil Conditioned using FBFM Data

MSE		
APH	Soil APH	Avg. APH
9.74	3.94	188.31
11.96	4.40	176.41
10.54	4.19	185.97
9.89	3.96	190.05
10.88	4.10	189.06
10.72	4.15	185.18

Note: Results by county of Mean Square Error (MSE) for conventional APH measure, and soil-conditioned APH measure. The third column is the average APH for each county and the full sample, measured in bu./acre for corn farms (F=612) with at least twenty years of data from 1972–2007. The 500 bootstrapped sample trials are executed for each farm. Six years of data are drawn for the APH calculation.

Discussion and Conclusion

We have shown that widely-available high resolution soil data can feasibly be integrated into crop insurance guarantee determination, and that a large boost in rating efficiency is likely relative to currently employed rating methods used by the government, which ignore explicit field location and soil data. While the data exist, RMA does not use this information when estimating premium rates. This study serves as a proof of concept and provides a strong motivation for modifying the program to incorporate these data. Of course, additional analyses and modeling efforts will be necessary to build the same into RMAs' existing rating methodology, but this work provides a sound basis and motivation for doing so by quantifying the actuarial efficiency impacts from ignoring information content in soils.

The availability of insurance affects investment and production decisions. Past research has shown that poorly-designed insurance can also lead to adverse incentives regarding practice choice (Woodard et al. 2012a). Approaches to integrating soil information explicitly into yield risk and insurance models is a necessary precursor to later accommodating and quantifying impacts from different soil sustainability practices in dynamic frameworks. For example, the amount of soil organic carbon (SOC) can be modified though managerial practices. These practices include the adoption of no-till farming, cover crops, and reforestation. They are commonly referred to as sustainable agricultural practices as they increase SOC, while also increasing soil fertility/productivity and sequestering carbon. This increase of SOC, however, is not immediate, but contributes to the long-term productivity of soils. Increased soil organic matter may also contribute to water retention in the long run (Arriaga and Lowery 2003). That being said, in certain conditions sustainable agricultural practices can lead to higher risk, and thus should be considered carefully.

For example, some practices may have complex impacts that also vary by soil type or climatology. Increasing the use of cover crops can potentially result in higher soil water retention in some soils (Teasdale and Mohler 1993). However, in arid and semi-arid regions, cover crops may have negative effects on soil water retention since they compete with the commodity crop for available water (Unger and Vigil 1998; Snapp et al. 2005). Usually these types of factors are not adequately accounted for in the pricing and administration of the program by the RMA.

Building a soils-based pricing foundation is a first step towards creating a crop insurance system that can accommodate future program modifications related to sustainability. This would open the door for improving conservation outcomes by appropriately incentivizing (or at least not disincentivizing) adoption via insurance that is appropriately designed and rated. Failure to properly determine guarantees under this program may lead producers to not adopt otherwise potentially optimal conservation practices such as cover crop use, skip-row (Woodard et al. 2012a), adaptive nitrogen management (van Es et al. 2007), or others. Nevertheless, the current structure of how insurance rules are determined, and how insurance is priced (i.e., via an agency, as opposed to market discovery of rates), leaves the system unable to respond flexibly. We would note that there is no guarantee that private firms would necessarily integrate soil into rates if left to private market pricing; however, the advent of "big data" technologies and the fact that a variety of firms have emerged in the farm management sector that utilize just as technologically-demanding data and models for related problems suggests a high likelihood that in today's market, most companies would make use of these data in a competitive market pricing situation.

That being said, insuring different units may create fraud opportunities such as those mentioned in Atwood, Robison-Cox, and Shaik (2006) if not properly tracked. In practice, the CLUs making up the unit would be combined to develop a unit-level soil quality measure (and there would be no changes in any case to the adjustment procedures or unit structures) so this is perhaps of less concern. Baselining APHs and rates against objective soil models could in fact be a superior alternative to current reporting procedures, but are outside of the scope of this paper. Optimally, a mix of producer experience data and objective soil data should be captured and weighed appropriately based on proper credibility models. Research exploring explicit credibility-adjusted approaches to including both soil and producer history data for determining APHs and rates should be pursued.

A lack of understanding surrounding the shortcomings of the current actual production history (APH) methods employed—as well as program rigidities—may hamper the adoption of these newer and more appropriate approaches. Since 2009, the government began collecting insurance policy records and yields tied to CLUs, but to date it simply does not make much use of the data.[3] In order to scale up the analysis to the entire country for the purposes of operationalizing modifications to the program to explicitly take into account soil quality, the government would likely need to release proper data to allow for further research and development by companies and the research community. To date, RMA has refused to make these data available to the research community, citing privacy concerns and interpretations of Section 1619 of the 2008 Farm Bill (see, e.g., Woodard 2016a and Woodard 2016b for a discussion of these issues). We urge the government to make these data available to researchers in a suitable form. Future data-intensive research applications will take on an increasingly larger role in shaping policy

[3] Specifically, the common land units (CLUs, or roughly speaking, fields) within each insured unit.

and sustainability science going forward (Woodard 2016b). The fact that these data exist, but that viable researchers may not access them, provides a strong motivation for the creation of a secure data warehousing facility to foster data-intensive analytics that otherwise cannot be conducted without the large troves of uncurated data whose access is restricted (sometimes reasonably, and sometimes not) by various government agencies. Given the privacy provisions in Section 1619 of the 2008 Farm Bill, it is likely that without repealing Section 1619 the only mechanism through which such gaps could be bridged would be through the development of such data facilities for researchers (and indeed there are precedents). These important program modifications could have overarching conservation and sustainability impacts.

There is fairly broad agreement and much empirical evidence (in crop insurance and other fields) that government agencies are likely not in a natural position to be jointly administering, regulating, and pricing these types of insurance programs, and indeed the track record in pricing efficiency is not favorable. Agencies tend to be reactive rather than proactive, and often lack sufficient bandwidth to suitably replicate pricing mechanisms that markets would yield (see e.g., Priest 1996; Cummins 2006; Jaffee 2006; Michel-Kerjan and Kousky 2010; Woodard et al. 2012b). This makes a very strong case for replacing the current government-controlled rating system with a more realistic, accurate, and flexible set of pricing mechanisms where companies and the market can have some hand in informing premium rates. This would allow the market to flexibly integrate information and data that a government agency cannot adequately handle, but which have a foundational importance in the classification of the risks to be underwritten (e.g., the influence of different soil types on crop growth). The FCIP as a policy has made many inroads into helping farmers manage risk, and it is well accepted that such markets would likely not exist in the absence of government intervention due to systemic risk. Much like deposit insurance, terrorism insurance, or even health insurance, government intervention and subsidization likely has a legitimate role to play in economically optimal outcomes. However, the government arguably should not be in charge of actually setting prices/rates for the program.

References

Arriaga, F.J., and B. Lowery. 2003. Soil Physical Properties and Crop Productivity of an Eroded Soil Amended with Cattle Manure. *Soil Science* 168 (12): 888–99.

Atwood, J.A., J.F. Robison-Cox, and S. Shaik. 2006. Estimating the Prevalence and Cost of Yield-switching Fraud in the Federal Crop Insurance Program. *American Journal of Agricultural Economics* 88 (2): 365–81.

Battisti, D.S., and R.L. Naylor. 2009. Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. *Science* 323 (5911): 240–4.

Coumou, D., and S. Rahmstorf. 2012. A Decade of Weather Extremes. *Nature Climate Change* 2 (7): 491–6.

Cummins, J.D. 2006. Should the Government Provide Insurance for Catastrophes? *Federal Reserve Bank of St. Louis Review* 88 (4): 337–79.

Hammer G.L., G. McLean, C. Messina, M.J. Roberts, and W. Schlenker. 2013. The Critical Role of Extreme Heat for Maize Production in the United States. *Nature Climate Change* 3 (5): 497–501.

Jaffee, D.M., and T. Russell. 2006. Should Governments Provide Catastrophe Insurance? *Economists' Voice* 3 (5): 1–8. Available online at http://faculty.haas.berkeley.edu/jaffee/papers/095BEPress06.pdf.

Lesk, C., P. Rowhani, and N. Ramankutty. 2016. Influence of Extreme Weather Disasters on Global Crop Production. *Nature* 529 (7584): 84–7.

Lobell, D.B., A. Sibley, and J. Ivan Ortiz-Monasterio. 2012. Extreme Heat Effects on Wheat Senescence in India. *Nature Climate Change* 2: 186–9.

Lobell, D.B., M. Bänziger, C. Magorokosho, and B. Vivek. 2011. Nonlinear Heat Effects on African Maize as Evidenced by Historical Yield Trials. *Nature Climate Change* 1: 42–5.

Lobell, D.B., W. Schlenker, and J. Costa-Roberts. 2011. Climate Trends and Global Crop Production since 1980. *Science* 333 (6042): 616–20.

Michel-Kerjan, E.O., and C. Kousky. 2010. Come Rain or Shine: Evidence on Flood Insurance Purchases in Florida. *Journal of Risk and Insurance* 77 (2): 369–97.

Natural Resources Conservation Service, Soil Survey Staff. 2015. National Value Added Look Up (valu) Table Database for the Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America and the Territories, Commonwealths, and Island Nations served by the USDA-NRCS. United States Department of Agriculture. Available online at https://gdg.sc.egov.usda.gov/.

Priest, G.L. 1996. The Government, the Market, and the Problem of Catastrophic Loss. *Journal of Risk and Uncertainty* 12 (2): 219–37.

Roberts, M.J., W. Schlenker, N. Braun, B.B. Little, R.M. Rejesus, and G.L. Hammer. 2014. Greater Sensitivity to Drought Accompanies Maize Yield Increase in the U.S. Midwest. *Science* 344 (6183): 516–9.

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): 653–59.

Snapp, S.S., S.M. Swinton, R. Labarta, D. Mutch, J.R. Black, R. Leep, J. Nyiraneza, and K. O'Neil. 2005. Evaluating Cover Crops for Benefits, Costs and Performance within Cropping System Niches. *Agronomy Journal* 97 (1): 322–32.

Soil Survey Staff. 2015. National Value Added Look Up (valu) Table Database for the Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America and the Territories, Commonwealths, and Island Nations served by the USDA-NRCS. United States Department of Agriculture, Natural Resources Conservation Service. Available online at https://gdg.sc.egov.usda.gov/.

Teasdale, J.R., and C.L. Mohler. 1993. Light Transmittance, Soil Temperature, and Soil Moisture under Residue of Hairy Vetch and Rye. *Agronomy Journal* 85 (3): 673–680

Unger, P.W., and M.F. Vigil. Cover Crop Effects on Soil Water Relationships. *Journal of Soil and Water Conservation* 53 (3): 200–7.

U.S. Department of Agriculture, Natural Resources Conservation Service, Soil Survey Staff. 2014. Web Soil Survey. Washington DC. Available at: http://websoilsurvey.nrcs.usda.gov/.

van Es, H.M., B.D. Kay, J.J. Melkonian, and J.M. Sogbedji. 2007. Nitrogen Management under Maize in Humid Regions: Case for a Dynamic Approach. In *Managing Crop Nutrition for Weather*, ed. T. Bruulsema, 6–13. Norcross, GA: International Plant Nutrition Institute Publication.

Woodard, J.D. 2014. A Conditional Distribution Approach to Assessing the Impact of Weather Sample Heterogeneity on Yield Risk Estimation and Crop Insurance Ratemaking. *North American Actuarial Journal* 18 (2): 279–93.

Woodard, J.D., B.J. Sherrick, and G.D. Schnitkey. 2011. Actuarial Impacts of Loss Cost Ratio Ratemaking in U.S. Crop Insurance Programs. *Journal of Agricultural and Resource Economics* 36 (1): 211–28.

Woodard, J.D., and B.J. Sherrick. 2011. Estimation of Mixture Models using Cross-Validation Optimization: Implications for Crop Yield Distribution Modeling. *American Journal of Agricultural Economics* 93 (4): 968–82.

Woodard, J.D., A.D. Pavlista, G.D. Schnitkey, K. Ward, and P. Burgener. 2012a. Government Insurance Program Design, Incentive Effects, and Technology Adoption: The Case of Skip-Row Crop Insurance. *American Journal of Agricultural Economics* 94 (4): 823–37.

Woodard, J.D., G.D. Schnitkey, B.J. Sherrick, N. Lozano-Gracia, and L. Anselin. 2012b. A Spatial Econometric Analysis of Loss Experience in the U.S. Crop Insurance Program. *Journal of Risk and Insurance* 79 (1): 261–85.

Woodard, J.D. 2013. Theme Overview: Current Issues in Risk Management and U.S. Agricultural Policy. *Choices* 28 (3): 1–2.

———. 2015a. Integrating High Resolution Soil Data into Federal Crop Insurance Policy: Implications for Policy and Conservation. *Environmental Science & Policy* 66: 93–100.

———. 2015b. Ag-Analytics.org Data Warehouse. Available at: ag-analytics.org.

———. 2016. Data Science and Management for Large Scale Empirical Applications in Agricultural and Applied Economics Research. *Applied Economic Perspectives and Policy* 38 (3): 373–88.

———. 2016. Big Data and Ag-Analytics: An Open Source, Open Data Platform for Agricultural & Environmental Finance, Insurance, and Risk. *Agricultural Finance Review* 76 (1): 15–26.

———. 2016. Crop Insurance Demand More Elastic than Previously Thought. *Choices*. 31 (3): 1–11. Available at http://www.choicesmagazine.org/choices-magazine/th-5 eme-articles/crop-insurance-in-the-20182019-farm-bill/crop-insurance-demand-more-elastic-than-previously-thought.