RéseauSPAANetwork

Structure et la Performance de l'Agriculture et de l'industrie des produits Agroalimentaires

Structure and Performance of Agriculture and Agri-products industry Network

On the Treatment of Heteroscedasticity in Area-Yield/Group-Risk Crop Insurance

Alan Ker
Aker@uoguelph.ca
Chair, department of Food, Agricultural and Resource Economics (FARE)
University of Guelph,
Guelph, Ontario, ON, N1G 2W1

Tor Tolhurst
ttolhurs@uoguelph.ca
Graduate research assistant, FARE, University of Guelph
Guelph, Ontario, ON, N1G2W1

Cahier de recherche/Working paper #2013-05

On the Treatment of Heteroscedasticity in Area-Yield/Group-Risk Crop Insurance

Résumé: Les rendements des cultures varient dans le temps et l'espace. Les programmes d'assurance récolte au Canada comme aux États-Unis dépendent de données historiques pour déterminer les primes. La modélisation des rendements joue un rôle crucial dans la spécification des contrats d'assurance. Nous exploitons des développements récents portant sur l'instabilité de la variance des rendements sur l'évaluation de probabilités de différents niveaux de rendements possibles. Nous utilisons des données sur les rendements de maïs et de soya pour des comtés de l'Illinois, l'Indiana et l'Iowa couvrant 56 années. Nous démontrons que l'estimation de coefficients d'hétéroscédasticité par secteurs stabilise la procédure d'estimation des rendements. Cette stabilité permet des économies significatives dans la détermination des primes d'assurance. Ces résultats sont particulièrement pertinents pour les programmes d'assurance récolte américains axés sur les rendements locaux et les risques collectifs, mais aussi pour les programmes canadiens. Nos résultats montrent aussi l'importance d'avoir des données de qualité au niveau des comtés.

Abstract: Crop yields vary a lot across time and space. Crop insurance programs in Canada and in the United States typically rely on historical yield data to determine premia. The econometric modeling of yields is of the utmost importance in the pricing of insurance contracts. In this manuscript we build upon recent contributions on how to account for the unstable nature of the variance of yields in predicting potential yield outcomes. With 56 years of county-level corn and soybean yield data for Illinois, Indiana and Iowa, we find that estimating heteroscedasticity coefficients using data pooled within a Crop-Reporting District greatly improves the stability of the estimation procedure. We demonstrate that this increased stability is non-trivial in that economically and statistically significant improvements in crop insurance rates are realized. These findings are particularly relevant for the U.S. area-yield or group-risk programs, but also for Canadian crop insurance programs. Our results highlight the importance of having reliable yield data at the county level.

Keywords: heteroscedasticity, yields, risk, crop insurance

Mots clés: hétéroscédasticité, rendements, risqué, assurance récolte

JEL codes/codes JEL : C53, Q18

1. Introduction

Heteroscedasticity has relatively trivial consequences for coefficient estimates in a linear regression context: estimates remain unbiased and standard errors are easily remedied by procedures such as White's heteroscedasticity-corrected standard errors. In contrast, heteroscedasticity -- by definition a non-constant error variance within the sample -- has serious consequences for estimating conditional yield densities, $\hat{f}(Y|t)$, and thus for crop insurance rates derived (implicitly or explicitly) from those estimated densities. In fact, improper treatment of heteroscedasticity will result in biased density estimates and thus biased premium rates. As we later illustrate, these biases can be non-trivial: premium rates can double or triple depending on different heteroscedasticity treatments. We consider heteroscedasticity treatments in the context of premium determination for Risk Management Agency's group-risk insurance programs which use county-level yield data to estimate rates. In 2012, group-risk insurance products in the United States accounted for a total of \$3.7 billion in liabilities and collected \$256 million in premiums.[1]

Modeling crop yield distributions have received significantly more attention than heteroscedasticity treatments in the agricultural economics literature despite the latter tending to have as much or greater impact on crop insurance rates. Recently Ke and Qiao (2012) examined the effects of yield distributions on crop insurance pricing in China, while Du, Hennessy and Yu (2012) examined the presence, determinants and economic consequences of skewness in yield distributions. Heteroscedasticity is also an important consideration when estimating yield distributions for non-crop insurance settings. For example, Cabas, Weersink and Olale (2010) and Tack, Harri and Coble (2012) both made use of estimated yield distributions to anticipate the effects of climatic change on agricultural production, while Lobell, Cassman, and Field (2009) estimated yield distributions in the context of food security. The treatment of heteroscedasticity is highly influential in both determining an appropriate crop

[1]Including the Group Risk Plan (GRP), Group Risk Income Protection (GRIP) and Group Risk Income Protection - Harvest Revenue Option (GRIP-HRO).

yield distribution and its resulting parameter estimates. Just and Weninger (1999) demonstrated that improperly assuming homoscedasticity may lead to incorrect conclusions about the best fitting distributional form.

Figure 1 illustrates the economic implications of heteroscedasticity on county-level crop insurance rates derived from estimated conditional yield densities under two common heteroscedasticity treatments: one adjusted for heteroscedasticity assuming constant coefficient of variation (the dashed line) and one assuming homoscedasticity (the solid line). In this case, the raw residuals fail to reject homoscedasticity but so too do the standardized residuals after accounting for a constant coefficient of variation. That is, both the adjusted and unadjusted residuals fail to reject homoscedasticity and as such it is difficult to choose one rate in favor of the other. Furthermore, the differences in the estimated premium rates are large: at the 75% coverage level, the adjusted rate of 2.19% is 7.2 times larger than the unadjusted rate of 0.30%. At the 90% coverage level, the difference between rates is proportionally smaller but remains substantial nevertheless: an adjusted rate of 3.80% versus an unadjusted rate of 1.16%. Clearly, the differences across the adjusted and unadjusted rates are economically significant. Furthermore, this ambiguous situation is prevalent in our data: 26.6% of the county-crop combinations fail to reject homoscedasticity in both the adjusted and unadjusted residuals.[2]

Harri *et al.* (2011) highlight the fact that researchers have been limited by the use of relatively strict heteroscedasticity treatments. They note: "Our results demonstrate that no single heteroscedasticity assumption is appropriate in every case. Being correct on average or even in a majority of cases may still lead to significant problems in the design, rating, and performance of crop insurance contracts" (p. 708). In response, Harri *et al.* (2011) developed an empirical approach that both relaxed heteroscedasticity assumptions and lead to statistically significant economic improvements in the actuarial soundness of premium rates.

[2]When variance is assumed to be linearly proportional to the fitted values (as opposed to constant coefficient of variation) 56.7% of the county-crop combinations fail to reject homoscedasticity in both the adjusted and unadjusted residuals.

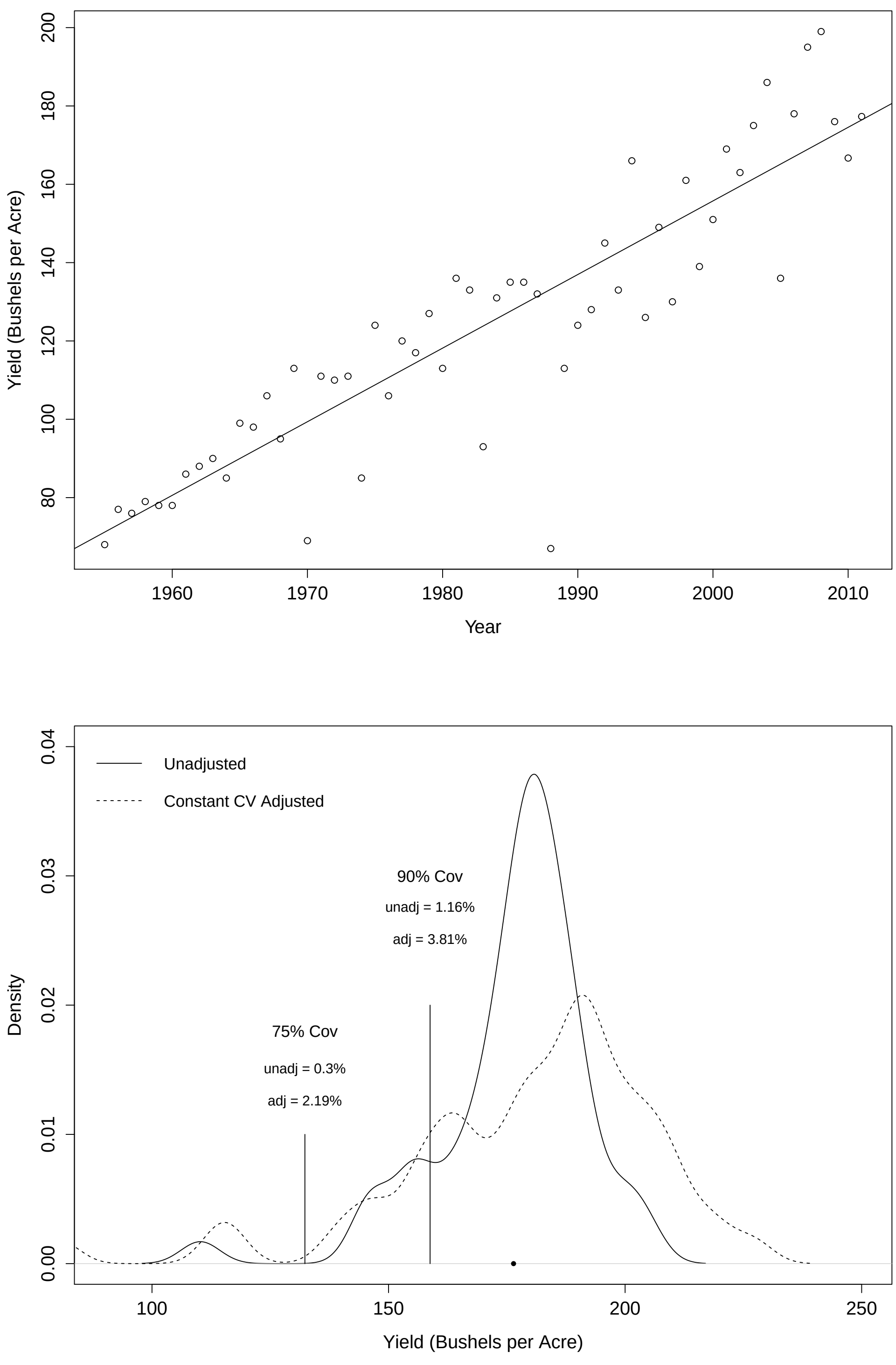

FIGURE 1. Top: historical corn yields in Bureau County, Illinois 1955-2011 with linear trend. Bottom: nonparametric estimates of Bureau County corn yield densities and estimated premium rates using different heteroscedasticity treatments.

Harri *et al.* (2011) estimate a heteroscedasticity coefficient, β, using $\sigma^2_{\hat{e}_t} = \sigma^2 \cdot \hat{y}_t^\beta$. They constrain $\beta \in [0, 2]$ given $\beta = 0$ implies homoscedasticity and $\beta = 2$ implies constant coefficient of variation. Somewhat troubling is that we found this constraint to be binding in 57% of the cases in our data. As a result, we propose an alternative approach to Harri *et al.* (2011) that does not require *a priori* constraints on β. We estimate an unconstrained β by pooling observations across counties in a given crop reporting district and estimate one β. Although this restricts the heteroscedasticity coefficients to be equivalent across counties within a district for a given crop, empirically we can not reject this restriction. In addition, pooling leads to economically and statistically significant improvements in crop insurance rates compared to either the bounded or unbounded method of Harri *et al.* (2011).

This manuscript proceeds as follows. In the next section, we review common heteroscedasticity treatments in the crop yield distribution literature, the empirical method of Harri *et al.* (2011), and the data used for the analysis. In the third and fourth sections we present our empirical results and demonstrate the economic and statistical signficance of pooling using an out-of-sample simulation. The final section offers some concluding remarks.

2. Empirical Treatment of Heteroscedasticity

The treatment of possible heteroscedasticity in the crop insurance literature generally involves testing for homoscedasticity and if the yield data fails to reject they are assumed homoscedastic.[3] Conversely, if the yield data rejects homoscedasticity the variance is generally assumed to be proportional to the fitted values.[4] Harri *et al.* (2011) offer a third approach as a flexible alternative by estimating the degree of heteroscedasticity. They find significant improvements in program loss ratios could be realized with their approach.

Their approach proceeds as follows. Assume a function $g(t)$ exists representing the technological change of yields through time and assume that variance is a function of the fitted

[3]Examples include Nelson (1990); Coble, Heifner and Zuniga (2000); Woodard and Sherrick (2011).
[4]Examples: Gallagher (1987), Goodwin and Ker (1998); Just and Weninger (1999).

values $\hat{y}$. Then the variance term can be specified as follows:

$$\sigma^2_{\hat{e}_t} = \sigma^2 \cdot \hat{y}_t^{\,\beta} \tag{1}$$

where β is the heteroscedasticity coefficient to be estimated and $\hat{e}$ are the estimated residuals from regressing y on $g(t)$. The heteroscedasticity coefficient can be estimated from the logarithm of Equation 1:

$$\ln \hat{e}_t^2 = \alpha + \beta \ln \hat{y}_t + \epsilon_t \tag{2}$$

where α is an intercept-term and ϵ_t is a well-behaved error-term. Interpreting the heteroscedasticity coefficient is straightforward: when $\beta = 0$ the variance is homoscedastic as $\sigma^2_{\hat{e}_t} = \sigma^2$; when $\beta = 1$ the variance is increasing at a constant rate; and when $\beta = 2$ the variance follows a constant coefficient of variation $\sigma^2_{\hat{e}_t} = \sigma^2 \cdot \hat{y}^2$. Then, given $\hat{\beta}$ and a forecast[5] at time $T+2$ denoted $\hat{y}_{T+2}$, a set of yields, $\tilde{y}_1 ... \tilde{y}_t$, from the conditional yield density of interest is derived as follows:

$$\tilde{y}_t = \hat{y}_{T+2} + \hat{e}_t \cdot \frac{\hat{y}^{\hat{\beta}}_{T+2}}{\hat{y}^{\hat{\beta}}_t} \tag{3}$$

These adjusted yields are used to estimate county-level premium rates.

Surprisingly, when this approach is applied to county-level yield data the $\hat{\beta}$'s are frequently outside the range [0,2] and often by a significant magnitude. Figure 2 illustrates the unconstrained county-level heteroscedasticity coefficient estimates for corn and soybeans in Illinois, Indiana, and Iowa.[6] Although the frequency of violations is readily apparent (and summarized in Table 1), the magnitude of these violations is somewhat troubling. Indeed

[5]Forecast at time $T+2$ following Harri et al. (2011). Insurance companies only have access to yield information up to year $n-2$ to set their rate for year n.

[6]County-level data for the period 1955-2011 from the National Agricultural Statistics Service were used. Any counties with incomplete yield histories are excluded: five counties from both Illinois corn and soybeans; 13 counties from Indiana corn; 10 counties from Indiana soybeans; and one county from Iowa soybeans. This resulted in 552 crop-county combinations and represents a large portion of the area-yield programs.

TABLE 1. Frequency of heteroscedasticity coefficient estimates outside [0,2]

	Illinois		Indiana		Iowa		
	Corn	Soybean	Corn	Soybean	Corn	Soybean	**Total**
$\hat{\beta} \in (2, \infty)$	57%	56%	34%	57%	33%	54%	**49%**
$\hat{\beta} \in (-\infty, 0)$	0%	8%	5%	5%	12%	14%	**8%**
Total	**57%**	**64%**	**39%**	**62%**	**45%**	**68%**	**57%**

Harri *et al.* (2011) imposed the constraint, $\beta \in [0, 2]$, to handle this instability.[7] We propose an alternative approach to constraining: we pool $\hat{e}_t$ and $\hat{y}_t$ across counties within a crop-reporting district to estimate one heteroscedasticity coefficient per district. This has the benefit of additional, albeit correlated, observations but will nonetheless reduce the variance of the estimated coefficient. Additionally, pooling allows for coefficient estimates outside the [0,2] range if warranted by the data. This approach assumes the underlying heterscedasticity structure is the same across counties in a crop-reporting district for a given crop. As such we both test this restriction and evaluate the effects of pooling for rating crop insurance contracts using an out-of-sample simulation.

3. An Alternative Approach

Harri *et al.* (2011) impose the restriction $\beta \in [0, 2]$ as a response to what is a high degree of instability in the heteroscedasticity estimates. While this certainly restricts $\hat{\beta}$ into an *a priori* plausible region it results in significant bunching at the endpoints of the constraint -- 57% of estimates in our county-crop combinations would be constrained. As an alternative we propose pooling the fitted yields and corresponding residuals of each crop-county combination to their respective crop-reporting district and estimate one heteroscedasticity coefficient; that is, we assume the heteroscedasticity coefficient is constant for counties within a crop-reporting district for a given crop. To this end we estimate Equation 1 using all the fitted yields and residuals for each county for a given crop pooled to the crop-reporting district.

[7] Rather than using least squares in the estimation of the heteroscedasticity coefficient, we used a variety of robust regression techniques (conditional quantile regression, trimmed least squares, and Huber M-estimation) in an attempt to stabilize $\hat{\beta}$. However, these techniques did not offer any tangible improvement.

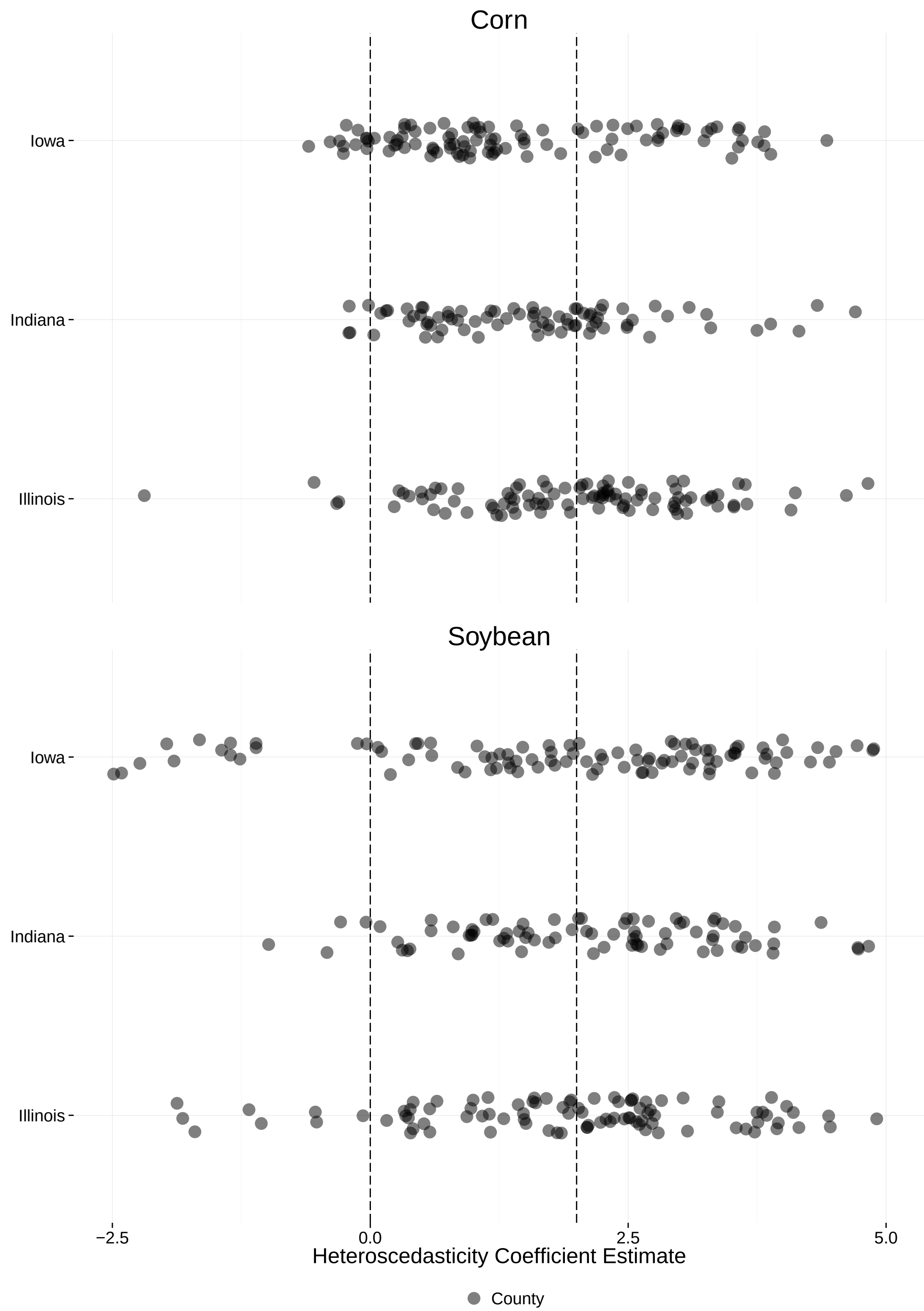

Figure 2. Unconstrained county-level heteroscedsaticity coefficient estimates for corn and soybean.

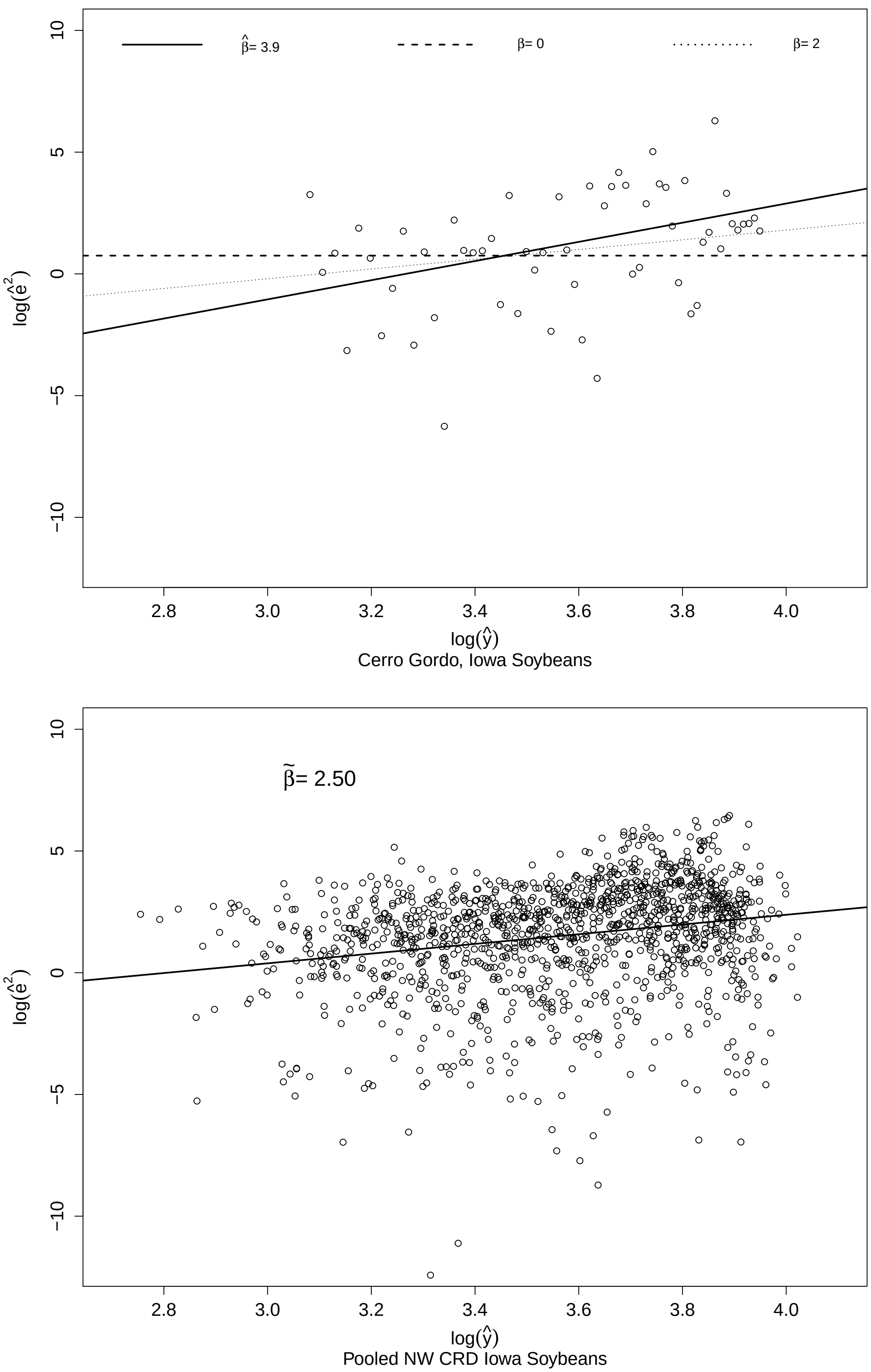

FIGURE 3. Top: regression space when $\hat{\beta}$ is estimated at the individual county level. Bottom: regression space when CRD pooled $\hat{\beta}$ is estimated.

A closer look at estimating the heteroscedasticity coefficient with yield data is revealing. Consider the Northern crop-reporting district of Iowa soybeans where the β constraint is violated in 68% of the counties. Figure 3 illustrates the heteroscedasticity regression for this district and one of its counties, Cerro Gordo county. The slopes of the lines in Figure 3 appear to fit the data equally well; one assumes homoscedasticity, another assumes constant coefficient of variation, while the third is the unconstrained estimate ($\hat{\beta} = 3.94$ which is significantly outside $[0, 2]$). The bottom panel illustrates the district-pooled heteroscedasticity procedure (with the estimated $\tilde{\beta} = 2.50$). As expected, when more observations are added the estimated heteroscedasticity coefficient has a substantially smaller variance. What is troubling is the magnitude of the differences in the premium rates resulting from the different heteroscedasticity estimates. The estimated premium rates at the 75% coverage level based on $\beta = \{0, 2, 2.5, 3.9\}$ are 0.09%, 0.58%, 1.36% and 2.91% respectively. At the 90% coverage level the rates are 0.34%, 1.15%, 1.99% and 3.54% respectively. Note that only $\beta = 0$ is rejected by the data at the 5% significance level.

A complete picture of the heteroscedasticity coefficient estimation with the pooled procedure is summarized in Figure 4. It is clear the heteroscedasticity coefficient estimate range narrows closer to the expected [0,2] range. In the case of the pooled estimates, the minimum and maximum values become more economically reasonable than in the unconstrained procedure: for the pooled $\hat{\beta}$ a minimum of -0.85 and maximum of 3.30 versus a minimum of -2.49 and maximum of 6.96 for the unconstrained county-level $\hat{\beta}$. Further, the estimates are allowed to fall out of the [0,2] range. This is important because there is no empirical evidence to suggest that the unknown but true heteroscedasticity coefficient may be outside of this [0,2] range for particular crop-county combinations. Although both pooling and constraining reduce the dispersion of heteroscedasticity coefficient estimates to more reasonable values, an outright constraint of $\beta \in [0, 2]$ may be overly restrictive. Indeed, the fact that in our sample the constraint was required in over 57% of the cases does suggest that the [0,2] range may be too restrictive. By estimating the heteroscedasticity coefficient using pooled data there is no need to determine appropriate values for restricting β estimates.

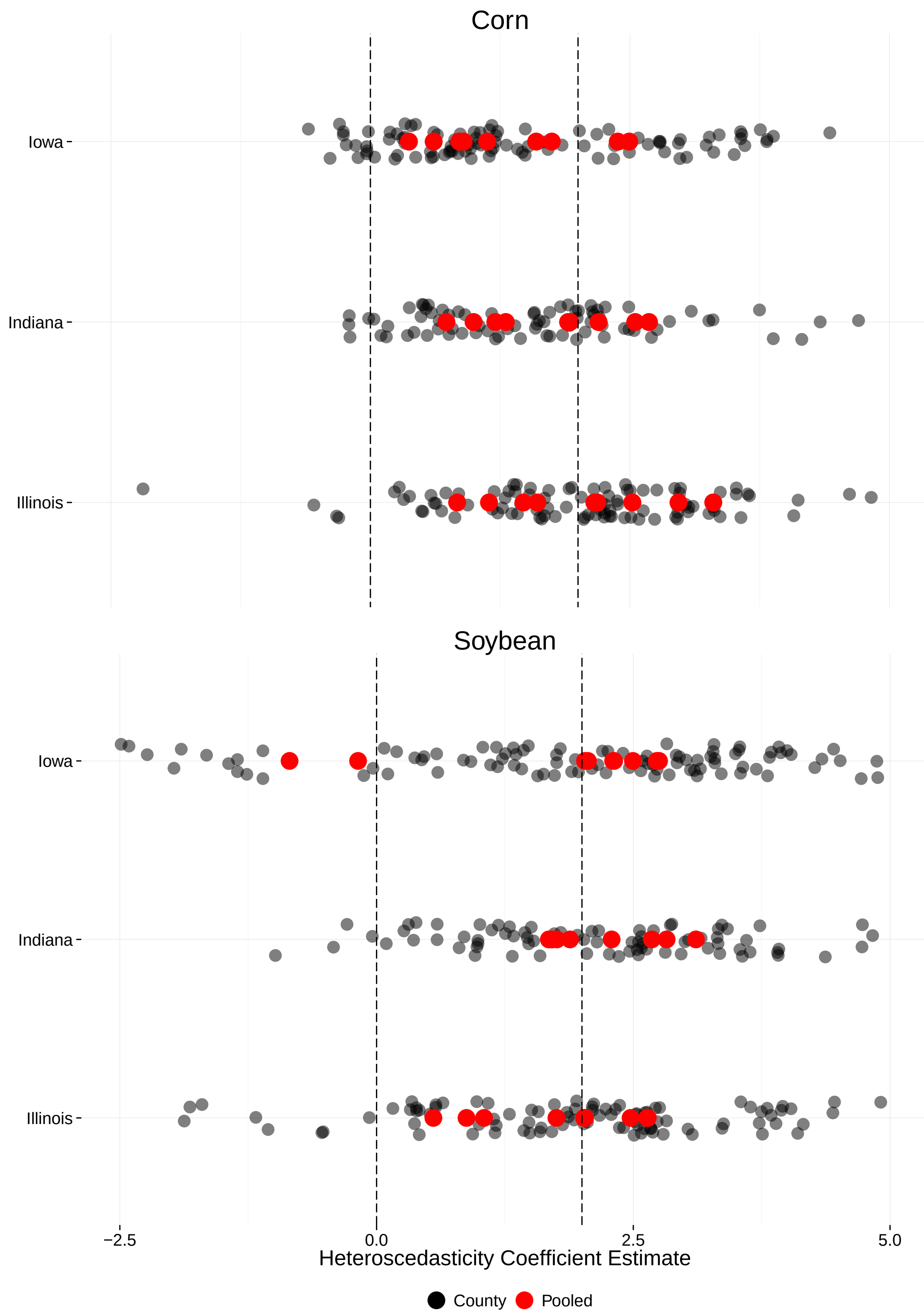

FIGURE 4. Comparison of pooled and unconstrained heteroscedasticity coefficient estimates.

Table 2. Summary of heteroscedasticity coefficient estimate t- and F-tests.

	% of Counties Failing to Reject at the 5% Significance Level			
	t-tests			F-test
State	$H_o : \hat{\beta}_i = 0$	$H_o : \hat{\beta}_i = 2$	Both	$H_o : \hat{D}_1 = \cdots = \hat{D}_i = 0$
Corn				
Illinois	38.1%	90.7%	33.0%	100.0%
Indiana	63.3%	88.6%	55.7%	100.0%
Iowa	63.6%	88.9%	52.5%	100.0%
Soybean				
Illinois	68.0%	91.8%	63.9%	100.0%
Indiana	54.9%	92.7%	53.7%	100.0%
Iowa	56.1%	84.7%	48.0%	100.0%

Table 2 presents the results of testing the heteroscedasticity coefficient from Equation 2, $\hat{\beta}$, for homoscedasticity ($\hat{\beta} = 0$) and a constant coefficient of variation ($\hat{\beta} = 2$) using a t-test with robust standard errors. Consistent with the direct heteroscedasticity tests, we fail to reject a constant coefficient of variation in a large majority of the county-crop combinations, fail to reject homoscedasticity in more than half the combinations, and fail to reject both in roughly half the combinations (assuming a 5% level of significance for all tests). We also perform a joint F-test in the crop-reporting district equation to test whether the heteroscedasticity coefficients are equivalent across the counties, using dummy variables to represent the different counties. We fail to reject the null that the heteroscedasticity coefficients are equal in almost all cases. Although this result is conssitent with pooling, it is not really surprising or informative given the low power of the test resulting from the high variance of the estimated coefficients. As such, we conduct a simulated game to consider the economic implications of using the pooling approach versus the restrictions.

4. Economic Implications

In this section we compare the pooled heteroscedasticity estimates with both the unpooled and restricted county-level estimates in an out-of-sample simulation of rating insurance contracts. This approach has been used extensively in the literature (see Ker and McGowan

2000; Ker and Coble 2003; Racine and Ker 2006; and Harri *et al.* 2011) to compare alternative rating methodologies. The simulation involves two agents: (1) the Risk Management Agency (RMA) which uses the base case insurance valuation technique, and (2) a private insurance company, which uses the proposed technique. In this case, the government insurance company uses the heteroscedasticity coefficient estimation technique of Harri *et al.* (2011), including the [0,2] bounds on heteroscedasticity coefficient estimates. The private insurance company uses the proposed pooled approach. Both agents use identical detrending procedures so that differences in the performance are exclusively due to differences in heteroscedasticity treatments.

The design of the simulation imitates the decision rules of the Standard Reinsurance Agreement. Under the Standard Reinsurance Agreement, private insurance companies may effectively retain or cede insurance contracts of their choice *ex ante*.[8] Let $\hat{\pi}_{tk}^{p}$ be the estimated premium rate of the private insurance company for county k in year t based on yield data from 1955 to $t-2$. Also let $\hat{\pi}_{tk}^{g}$ be RMAs estimated premium rate for county k in year t again based on yield data from 1955 to $t-2$. The private insurance company will retain policies it expects to be profitable; that is, policies with rates lower than the government rates ($\hat{\pi}_{tk}^{p} < \hat{\pi}_{tk}^{g}$). Loss ratios are calculated for the set of retained policies and the set of ceded policies using actual yield realizations. The simulation is performed on a by crop and state basis and counties are weighted by acreage. Consistent with the literature (Ker and McGowan 2000; Ker and Coble 2003; Racine and Ker 2006; and Harri *et al.* 2011) we use 15 years for the out-of-sample component and p-values are calculated using randomization methods (1000 randomizations).

Table 3 reports the results of the out-of-sample simulation for all state-crop combinations at the 75% and 90% coverage levels. Both the constrained and unconstrained government heteroscedasticity processes are reported in the table. In most cases the loss ratios are quite low but correspond to actual loss ratios for the program over this period.

[8]The SRA contains multiple funds and is more complicated but essentially a private insurance company can significantly reduce their exposure to unwanted policies.

TABLE 3. Out-of-sample simulation results for the 75% and 90% coverage levels.

Crop	State	Private % of policies	Psuedo Loss Ratio Private	Psuedo Loss Ratio Government	p-value	Underpayment to program	Overcharge to farmers
Government $\beta \in [0, 2]$, 75% Coverage Level							
Corn	IA	26.0%	0.000	0.043	0.000	-1,283,504	1,251,888
	IL	27.8%	0.040	0.049	0.033	-18,396,166	2,071,608
	IN	15.8%	0.070	0.145	0.053	-3,154,020	1,187,909
Soybean	IA	21.6%	1.068	0.429	1.000	-1,110,723	3,288
	IL	36.9%	0.178	0.578	0.001	-1,213,266	349,334
	IN	18.5%	0.072	0.245	0.007	2,775,644	150,125
Government $\beta \in [0, 2]$, 90% Coverage Level							
Corn	IA	26.0%	0.032	0.114	0.000	-2,085,988	1,847,554
	IL	27.8%	0.206	0.227	0.025	-25,276,303	3,201,127
	IN	15.8%	0.250	0.330	0.013	-5,573,420	1,816,622
Soybean	IA	24.0%	1.237	0.489	1.000	-1,603,765	11,279
	IL	37.5%	0.434	0.685	0.007	-4,741,512	284,890
	IN	18.8%	0.486	0.609	0.026	-2,156,519	769,971
Government $\beta \in (-\infty, \infty)$, 75% Coverage Level							
Corn	IA	55.2%	0.005	0.027	0.000	-1,099,013	8,559,892
	IL	58.1%	0.009	0.030	0.000	-8,468,842	17,272,550
	IN	48.9%	0.020	0.103	0.000	-2,045,060	7,538,648
Soybean	IA	59.1%	0.071	0.475	0.062	-392,103	2,588,397
	IL	65.2%	0.053	0.739	0.000	-420,548	5,741,677
	IN	54.5%	0.018	0.281	0.000	-1,977,644	9,473,880
Government $\beta \in (-\infty, \infty)$, 90% Coverage Level							
Corn	IA	55.2%	0.021	0.107	0.000	-1,877,533	10,255,324
	IL	58.1%	0.063	0.221	0.000	-13,531,817	19,958,888
	IN	48.9%	0.098	0.285	0.000	-4,112,465	9,222,908
Soybean	IA	60.3%	0.163	0.823	0.000	-555,301	2,971,117
	IL	65.5%	0.173	0.825	0.000	-985,400	7,440,351
	IN	54.7%	0.134	0.792	0.000	-3,521,869	11,242,732

Note: All counties weighted by share of total acreage. Underpayment and overcharge are in dollars.

For the pooled estimation procedure versus the Harri *et al.* (2011) procedure with $\beta \in [0, 2]$, the pooled estimation procedure tends to perform better: in ten of the twelve cases loss ratios are lower and statistically significant with the exception of Iowa soybean at both coverage levels. With respect to Iowa soybean the heteroscedasticity coefficient estimates are highly volatile both across the counties and within a county over the simulation period. As a result, even the pooled estimate is somewhat volatile. Comparing the pooled approach to the unconstrained methodology, pooling outperforms all state-crop combinations and the private insurance company retains a far greater proportion of the contracts compared to the bounded scenario. This is not surprising since leaving the heteroscedasticity coefficient estimates unbounded at the county-level results in wildly-fluctuating premium rates. Summarizing, in twenty-two of the twenty-four cases the pooled approach outperforms the non-pooled approaches.

To give an idea of the scope of rate differences, the two columns on the right hand side of Table 3 sum the dollar difference in premiums generated by different heteroscedasticity treatments.[9] These calculations assume the private insurance company's premium rates are correct such that if $\pi^p > \pi^g$ the difference between the rates is the underpayment to the program and if $\pi^p < \pi^g$ is the overcharge to farmers for one year (2011). The underpayment and overcharge are only for one year of the program—in perpetuity, these values represent significant transfers of wealth between insurers and policyholders. Notably for corn, all of the underpayment and overcharge values exceed \$1 million. The underpayment and overcharge values for soybeans are smaller but nevertheless roughly half exceed \$1 million. Not surprisingly, the differences between the private and government insurance rates are largest at higher coverage levels and when the government heteroscedasticity coefficient is left unbounded. To put these values into perspective, for the bounded scenario at the 75% (90%) coverage level, the underpayment and overcharge represent 61.5% (44.8%), 56.1% (37.5%), and 55.7% (38.6%) of the estimated total premiums for 2011 in Illinois, Indiana and Iowa respectively. Similarly for soybean, the underpayment and overcharge represent 53.0% (35.7%), 64.1% (39.9%) and 62.8% (44.4%) of total estimated 2011 premiums. In the unbounded scenario the proportion of underpayment and overcharge are even higher,

[9]Premiums for a given country are calculated as the product of the county's estimated premium rate (bu./ac.), the RMA projected price (\$/bu.) and insured acreage (ac.), all for 2011.

often exceeding 100% of the total estimated premiums. The underpayment and overcharge values demonstrate the economically significant differences in crop insurance premium rates generated under different heteroscedasticity treatment scenarios.

5. Conclusions

This purpose of this manuscript was to recover heteroscedasticity estimates that aligned with *a priori* expectations without imposing restrictive bounds. We estimated a single heteroscedasticity coefficient per crop-reporting district combination rather than per county-crop combination by pooling the data across counties within a district. Not only did the pooling lead to more *a priori* plausable estimates, but our out-of-sample simulations found that it offers economically and statistically signficant improvements in crop insurance rating over Harri *et al.* (2011) in twenty-two of the twenty-four cases considered. This finding is noteable as the Harri *et al.* (2011) approach is used for rating group-risk insurance contracts (a program with liability in excess of $3.7 billion in 2012) as well as the proposed shallow loss area programs proposed under the new farm bill.

References

Cabas, J., A. Weersink, and E. Olale. 2010. "Crop yield response to economic, site and climatic variables." *Climatic Change* 101(3):599–626.

Coble, K.H., R. Heifner, and M. Zuniga. 2000. "Implications for crop yield and revenue insurnace for produces hedging." *Journal of Agricultural and Resource Economics* 23:539–551.

Day, R.H. 1965. "Probability distributions of field crop yields." *Journal of Farm Economics* 47(3):713–741.

Du, X., D.A. Hennessy, and C.L. Yu. 2012. "Testing Day's conjecture that more nitrogen decreases crop yield skewness." *American Journal of Agricultural Economics* 94(1):225–237.

Gallagher, P. 1987. "US soybean yields: Estimation and forecasting with nonsymmetric disturbances." *American Journal of Agricultural Economics* 69(4):796–802.

Goodwin, B.K., and A.P. Ker. 1998. "Nonparametric estimation of crop yield distributions: Implications for rating group-risk crop insurance contracts." *American Journal of Agricultural Economics* 80:139–153.

Harri, A., K.H. Coble, A.P. Ker, and B.J. Goodwin. 2011. "Relaxing heteroscedasticity assumptions in area-yield crop insurance rating." *American Journal of Agricultural Economics* 93(3):707–717.

Just, R.E., and Q. Weninger. 1999. "Are crop yields normally distributed?" *American Journal of Agricultural Economics* 81(2):287–304.

Ke, W., and Z. Qiao. 2012. "Influence of flexible crop yield distributions on crop inurance premium rate - A case study on cotton insurance in three areas of Xinjiang." *Asian Agricultural Research* 4(5):7–12.

Ker, A.P., and K. Coble. 2003. "Modeling conditional yield densities." *American Journal of Agricultural Economics* 85(2):219–304.

Ker, A.P., and P. McGowan. 2000. "Weather-based adverse selection and the U.S. crop insurance program: The private insurance company perspective." *Journal of Agricultural and Resource Economics* 25(2):386–410.

Lobell, D.B., K.G. Cassman, and C.B. Field. 2009. "Crop yield gaps: their importance, magnitudes and causes." Working paper, NCESR Publications and Research. Paper 3.

Nelson, C.H. 1990. "The influence of distributional assumptions on the calculation of crop insurance premia." *North Central Journal of Agricultural Economics* 12(1):71–78.

Park, R.E. 1966. "Estimation with Heteroscedastic Error Terms." *Econometrica* 34(4):888.

Racine, J., and A.P. Ker. 2006. "Rating crop insurance policies with efficient nonparametric estimators that admit mixed data types." *Journal of Agricultural and Resource Economics* 31(1):27–39.

Tack, J., A. Harri, and K. Coble. 2012. "More than mean effects: Modeling the effect of climate on the higher order moments of crop yields." *American Journal of Agricultural Economics* 94(5):1037–1054.

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):968–982.