Constructing Farm Level Yield Densities from Aggregated Data: Analysis and Comparison of Approaches

Joseph Cooper♠ [1 2], Michael Langemeier♦, Gary Schnitkey♥, and Carl Zulauf[1]♣

April 23, 2009

Paper prepared for the Agricultural & Applied Economics Association's 2009 AAEA & ACCI Joint Annual Meeting, Milwaukee, WI, July 26-28, 2009.

♠Economic Research Service, USDA, jcooper@ers.usda.gov
♦Kansas State University, mlange@agecon.ksu.edu
♥ University of Illinois, schnitke@uiuc.edu
♣ The Ohio State University, zulauf.1@osu.edu

[1]The reviews expressed herein are those of the authors and not necessarily of their respective institutions. [2] Corresponding author.

Constructing Farm Level Yield Densities from Aggregated Data: Analysis and Comparison of Approaches

Abstract

Yield variability can be significantly higher at the farm level than at more aggregated levels, including the county. However, due to a dearth of available farm level data, much stochastic analysis involving farm yields utilizes more aggregated yield data as a proxy for the farm level. We empirically evaluate farm-level variability using longitudinal farm level data sets available from the Kansas Farm Management Association and the Illinois Farm Business and Farm Management Association. For corn, soybeans, and wheat, we compare the farm level yield variability obtained from this data to that inferred from Federal crop insurance premiums. The farm management data exhibit lower yield variability than are implied by the crop insurance premiums.

Key words

Yield variability, crop insurance, corn, wheat, soybeans

Introduction

Yield variability is significantly higher at the farm level than at more aggregated levels, including the county. Thus, using a yield density function based on aggregate data may not result in the most accurate solutions to farm decision models that incorporate uncertain decisions, including models based on expected utility. Further, certain Federal support programs, including a variety of crop insurance plans, make payments or calculate indemnities based in part on farm level yield data. Moreover, the 2008 Farm Act includes two new programs that base eligibility for payments (e.g., Average Crop Revenue Election (ACRE)) or the payments themselves (e.g., Supplemental Revenue Assistance (SURE)) on farm yields. Hence, when studying the distribution of costs to the government of these programs as well as the impacts of these programs on farm revenue and production, yield density functions estimated at the farm level will provide more accurate results than those based on county or State level data.

The dearth of farm level yield data is a barrier to this analysis. For instance, the lowest level of aggregation of the most comprehensive yield data set for the US is at the county level (U.S. Department of Agriculture (USDA), National Agricultural Statistics Service (NASS)).[1] One approach used to generate farm level yield data by previous research is to back it out of distributions implied by crop insurance premiums set by the Risk Management Agency (RMA) of the USDA (e.g., Coble and Dismukes, 2008). This approach assumes that RMA premium rates are actuarially correct, free from moral hazard and actuarial ratings problems. Even if actuarially correct, yield data generated from insurance premium rates may not be representative of the general population as

[1] NASS data is available from http://www.nass.usda.gov/Data_and_Statistics/Quick_Stats/index.asp.

adverse selection may influence participation in crop insurance. A second approach used by previous research is to assume a farm versus county or farm versus State yield relationship that appears reasonable (e.g., Goodwin, 2009).

In this paper, we use a different approach by empirically evaluating farm-level variability using longitudinal farm level data sets available from the Kansas Farm Management Association and Illinois Farm Business and Farm Management Association. We use these data sets to examine the statistical relationships between farm level yield data and yield data at the county level for corn, soybeans, and wheat. Specifically, we calculate the relationship of the standard deviations of yields using farm data and compare those standard deviations for county and state aggregations using simple ratios and through the use of econometric techniques. We also examine the distribution of the difference between farm level yield and county level yield.

Given an identified distribution for this difference in conjunction with the relationship between farm and county level standard deviations of yield, we show how to construct farm level yield densities from NASS county level yield densities. Last, the ratios of farm to county level standard deviations of yield obtained from the Kansas and Illinois farm management data are compared to the ratios of farm to county level standard deviation of yield inferred from RMA crop insurance premiums.

Data

The Illinois data used in this study came from the Illinois Farm Business Farm Management (FBFM) program. FBFM is a farmer-owned cooperative that has a working relationship with the University of Illinois at Urbana-Champaign. Farmer members maintain production and financial records for their farm. At the end of the calendar year, financial statements and production records are prepared and aggregate databases of crop and livestock production, receipts, expenses, inventories, and capital accounts are produced so as to develop benchmarks farmers can use to compare their farms. To be included in the data base, FBFM personnel must certify a farm's data as usable, meaning that it is reliable data. The data chosen for analysis cover periods over which preparation of the data and computations are consistent. The Kansas Farm Management Association (KFMA) data (Langemeier, 2005) was developed in a similar fashion.

Analysis of the farm level data

Length of the farm level data set represents a consideration of two trade offs. The longer the observation period, the larger the degrees of freedom that exist in establishing yield relationships. On the other hand, farmers tend to drop in and out of membership in the Kansas and Illinois farm associations and/or do not provide verified information for all years. As such, the longer the observation period, the smaller the number of farms available for the analysis. This problem is exasperated by the fact that only counties with at least 10 sample farms are included in the analysis in order to obtain sufficient statistical power when estimating farm-to-county yield relationships.

We have a combination of data sets that track individual farmers for consecutive 10 and 17 year periods, depending on the crop and State. The benefit of the 17 year data set is more degrees of freedom in establishing the yield characteristics, while the benefit of the 10 year data sets are greater sample sizes per county. The 10 year span datasets cover 1997 to 2006. The 17 year datasets cover 1991 to 2007. The particular starting dates were chosen so that data for each state would cover the same period.

The Illinois yield data is detrended to s = 2007. Y_{it} is detrended as:

(1) $Y_{it}^{d} = E(Y_{is})(\Delta Y_{it} + 1), \forall$ i counties, t periods, $t \neq s$.

We detrend yield based on the standard practice of using a linear trend regression of $Y_t = f(t)$. The expected value of Y_t, or $E(Y_t)$, is calculated from the fitted trend equation. We specify the yield deviate ΔY_{it}^{d} as the deviation of detrended yield from expected yield in the base year. The other crop/State combinations exhibited flat trends over the time periods examined, and hence were not detrended. We denote farm, county, and State level yields as Y^F, Y^C, and Y^S, respectively.

Table 1 reports averages of the standard deviation and mean of yield at the farm, county, and State levels for the T = 10 and 17 year datasets. The statistics for each crop, the mean and standard deviation of farm level yield, are summarized as

$\overline{Y}^F = \frac{1}{N}\sum_{i=1}^{N}\overline{Y}_i^F$ and $\sigma(Y^F) = \frac{1}{N}\sum_{i=1}^{N}\sigma(Y_i^F)$, respectively, where $i = 1,\ldots,N$ farms in the sample reporting production of the crop, where Y_i^F is the $(T \times 1)$ vector of yield values for farm i, and where $\overline{Y}_i^F = \frac{1}{T}\sum_{t=1}^{T} Y_{it}^F$. The mean and standard deviation of county level yield for a crop in a state are defined as $\overline{Y}^C = \frac{1}{J}\sum_{j=1}^{J}\overline{Y}_j^C$ and $\sigma(Y^C) = \frac{1}{J}\sum_{j=1}^{C}\sigma(Y_j^C)$, respectively, where $j = 1,\ldots,J$ counties in the sample, and where the table reports the number of counties. Y_j^C is the $(T \times 1)$ vector of yield values for county j, where $Y_{jt}^C = \overline{Y}_{jt}^F$, where the latter is the mean yield of all farmers in county j in period t. State yield in t is

$Y_t^S = \frac{1}{N}\sum_{i=1}^{N} Y_{it}^F$. The ratio of the standard deviation of farm level yield to county level yield reported in the table is $\delta^{FC} = \frac{1}{N}\sum_{i=1}^{N}\frac{\sigma\left(Y_i^F\right)}{\sigma\left(Y_i^C\right)}$.

As one would expect, farm-level yields are more variable than county or state state yields. As shown in Table 1, the ratio of farm-to-county level yields ranges from 1.12 (corn in Kansas) to 1.42 (corn in Illinois), for both the 10 and 17-year datasets. Similarly, the ratio of farm-to-state yields are above 1 and range from 1.32 (soybeans in Kansas) to 1.81 (corn in Illinois). The sample of Illinois counties for the 17 year dataset still contains a relatively large number of counties but with the benefit of increased precision of the estimates. For this 17-year Illinois data, the ratio of farm-to-county level yields is 1.3 for corn and 1.4 for soybeans. The sample sizes of Kansas counties for the 17-year dataset are relatively low, and standard deviations of the ratios do not consistently decrease in moving from the 10-year to 17-year datasets, suggesting that the 10 year dataset may be more reliable for that State. For this 10-year Kansas, the ratio of farm-to-county level yields is 1.1 for corn, 1.3 for soybeans, and 1.1 for wheat.

While varying among crops and locations, the table shows that a significant portion of yield variability is not embodied even in the county level aggregation. Hence, while statistical analysis using State level yield aggregations will certainly miss a large portion of the farm level yield variability, so too can analysis using county level data.

One question is whether or not the ratio of the standard deviation of farm level yield to county level yield is constant across the observed range of county level yield variation. If so, then the researcher can feel more secure in applying the results of this analysis to other regions. Regression results can be used to generate farm level standard deviation of yield from the county level data in cases where the relationship is not constant. We do this by crop and region using a simple specification to allow easy use of the results by other researchers. Rather than regress the standard deviation of farm yield on the standard deviation of county yield, we normalize both variables by their respective mean yield values. Using the coefficient of variation as the regression variables not only helps to account for potential herosckedaticity that is a function of yield, but also generalizes the regressions by making them scale free.

The fitting of second order polynomial functions to scatterplots of the coefficient of variation of yield at the farm level versus the coefficient of variation of yield at the farm level for each state and crop combination suggested that a linear functional form provides the best fit. Given the linearity of the relationship, Table 2 presents the regression results for a simple model of the log of the coefficient of variation of farm level yield regressed on the log of the coefficient of variation of county level yield. Logging the dependent variable prevents prediction of negative coefficient of variation. Note though that one cannot take the standard errors for the regressions seriously for use in hypothesis testing; the coefficient of variation at the farm level is not a function of the coefficient of variation at the county level given that the county data is constructed from the farm level data. However, our principal goal for these regressions is simply to fit functions to the data.

Figure 1 plots the fitted curves based on the regressions in Table 2. Growing conditions differ between Illinois and Kansas over the range of the observed data for each State and crop. Illinois is located in the middle of the corn belt in general while the Kansas is located in the more variable growing conditions of the U.S. plains.

Nonetheless, the slope of the Illinois corn function has roughly the same slope as for the Kansas crops, even if not statistically the same. The slope of the Illinois soybeans function in particular diverges from that of the other functions in the figure. In any case, based on F-tests constructed from the residual sum of squares of the regressions in Table 2, the coefficients in a joint regression including all crops and both States (or one with Illinois soybeans excluded) are significantly different at the 5 percent level from those in the individual regressions.[2] When we consider all the other possible contributions to error in generating yield densities, using the constant farm to county relationships from Table 1 is probably satisfactory for most applications.

Statistical distribution of farm versus county level yield differences

In this section, we examine for each crop the statistical distribution of $\lambda = Y_{ijt}^{F} - Y_{ijt}^{C}$, over $i = 1,\ldots,N$ farmers, $j = 1,\ldots,J$ counties, and $t = 1,\ldots,T$ years. As the next section will show, identification of the distribution of λ provides the information necessary to generate farm level data from county level yield aggregates. We use Q-Q plots to examine the departure of λ from normality. As in the previous section, only counties with at least 10 sample farms are included in the analysis.

The Q-Q plot compares the quantiles of the data against theoretical quantitles, where in our case, the latter are for the normal distributions. The closer the plot of data is to the straight line of the theoretical (normal) plot, the closer the data is to being normally distributed. The quantile plots in Figures 2 and 3 below show that the majority of yield differences conform to the normality assumption. Q-Q plots based on the uniform,

[2] In another regression, dummy variables for soybeans and wheat and for Kansas were added to the regression across both States and all crops. The soybean and Kansas dummies were significantly different from zero (and positive) at the 1 percent level but the wheat dummy was not significant at the 10% level.

exponential, logistic, and extreme value distributions were also examined, but the normal distribution provided the best overall fit to the data.

Generating the farm level yield distribution – an application

We find above that the difference between farm level yield and county level yield is distributed normally with mean zero for most farms. If we then assume that distributional choice holds generally, it becomes relatively simple to convert the detrended NASS county level yield density into a detrended farm level density.[3] This conversion is possible regardless of the distribution of yield.

Let Y_j^{Cd*} = yield for county *j*, Y_{ij}^{Fd*} = yield for farmer *i* in county *j*, and λ_{ij} = Y_{ij}^{Fd*} - Y_j^{Cd*}, $\forall i, j$ (we omit time subscript *t* for clarity), where the *d* superscript signifies detrended yields and the superscript * signifies that the data may be simulated rather than actual data. Based on this notation, $Y_j^{Cd*} + \lambda_{ij} = Y_{ij}^{Fd*}$, and it follows that var($Y_j^{Cd*} + \lambda_{ij}$) = var(Y_{ij}^{Fd*}). Since we find that cov(Y_j^{Cd*}, λ_{ij}) = 0 for the farm management datasets, the variance function simplifies to var(Y_j^{Cd*}) + var(λ_{ij}) = var(Y_{ij}^{Fd*}), or var(λ_{ij}) = var(Y_{ij}^{Fd*}) - var(Y_j^{Cd*}). Given δ_j, the ratio of the standard deviation of farm level yield to county level yield from Table 1, we define $var(Y_{ij}^{Fd*}) = \left[\delta_j \cdot \sigma\left(Y_j^{Cd*}\right)\right]^2$. Hence, the standard deviation of the farm level noise is $\sigma(\lambda_{ij}) = sqrt\left\{\left[\delta_j \cdot \sigma\left(Y_j^{Cd*}\right)\right]^2 - \mathrm{var}\left(Y_j^{Cd*}\right)\right\}$, and *z*, the farm level noise to add to the county yields to obtain farm level yield, is distributed normally, z ~$N(0, \sigma(\lambda_{ij}))$.

[3] Future analysis could use nonparametric approaches to generate the distribution of the yield difference, and compare the sensitivity of policy simulation results to the assumed distribution of the yield difference.

Farm level yield densities inferred from Federal crop insurance premiums

In our application of the Coble and Dismukes approach to backing-out the farm level standard deviation of yield from crop insurance premiums from RMA data, we assume that our representative farmers purchase revenue assurance (RA) with the base price option and 70 percent coverage (RMA, 2009). In our notation, the RA indemnity payment per acre for producer *i* of crop *j* in period *t* is

$$RA_{ijt} = \max\{0, (0.7 \cdot E(P_{jt}) \cdot Y_{ijt}^{APH} - P_{jt} \cdot Y_{ijt}^{Fd*})\},$$

where $E(P_{jt})$ and P_{jt} are expected and harvest time futures prices, respectively, and Y_{ijt}^{APH} is the actual production history for the farm.

Cooper (2009) discusses the simulation of the P_{jt} in a manner that preserves their inverse correlation with national level yield. A (1×*S*) vector of prices is simulated for each g = 1,…,*G* values of a simulated national yield vector, where *S* = *G* = 1000. The latter is drawn from a kernel density function estimated from NASS national level data over 1975 to 2008. County level yield data is also drawn from a kernel density function estimated from NASS county level data over the same period. We generate the county level yield density functions from NASS data using a kernel approach described in Cooper (2009). The actual Pearson's correlation between national and county yield over the period is imposed on the simulated yields (*ibid.*).

In our simulation context, the insurance premium, $PREM_{ijt}$, is actuarially correct if it is set equal to $E(RA_{ijt})$, where $E(RA_{ijt})$ is the mean of all outcomes of the equation above given our ($S \times G$) matrix of prices and ($G \times 1$) vector of farm yields. Using a quasi-

Newton technique with the BFGS algorithm, we find the value of $\sigma(\lambda_{ij})$ that minimizes $abs(PREM_{ijt} - E(RA_{ijt}))$, where $PREM_{ijt}$ is the full premium including the farmer paid portion and the portion subsidized by the government. The simulation is conducted for a representative farmer in four counties, including one outside the States covered by the farm management data sets. The farmer paid premium is downloaded from the RMA website (RMA, 2009) using the price and APH yield values from Table 3, and divided by 0.41 to generate $PREM_{ijt}$. The 0.41 inflates the premium up to its full cost, given that the government subsidizes on average 59 percent of the insurance premium.[4]

Generally, we find that $\sigma(\lambda_{ij})$ is higher when inferred from the RMA premiums than when calculated from the farm management data.[5] Table 4 shows the RMA crop insurance premium, the ratio $\sigma(Y_{ijt}^{Fd*})/\sigma(Y_{jt}^{Cd*})$ inferred from it, and also what the full (unsubsized) premium would be using the $\sigma(Y_{ijt}^{Fd*})/\sigma(Y_{jt}^{Cd*})$ ratios from the farm management data sets. For our farmer/crop combinations for Illinois and Kansas in columns 4 and 5 of Table 4, the ratio of $\sigma(Y_{ijt}^{Fd*})/\sigma(Y_{jt}^{Cd*})$ is on average 1.94 times higher based on the RMA data. However, given the standard errors of the farm management based ratios reported in Table 1, the differences between these ratios and the RMA ratios are not necessarily statistically different. Of course, the smaller the farm management

[4] In Table 3, the Revenue Assurance base price for the 2009 calendar year harvest of winter wheat was established by RMA using August and September 2008 values of the 2009 KCBT July hard red winter wheat futures contract. This 2008 period exhibited significant commodity price spikes and is unlikely to reasonably reflect updated expected prices for the 2009 crop. Hence, we recalibrate the $E(P_{2009})$ for winter wheat to \$5.85, which is the average of end-of-week closing prices in February 2009 for 2009 KCBT July hard red winter wheat (Cooper, 2009).

[5] A future research avenue could be to explore the feasibility and utility of backing out from RMA premiums the loads (i.e., prevented planting, replant) that it puts on rates.

ratio of yields relative to the RMA inferred values, the lower the actuarially correct simulated premium in the last column.

Conclusions

As one would expect, farm yields are more variable than either state or county yields. Our research on the Kansas and Illinois farm management datasets suggests that there are practical ways of modeling the variety of farm yields, even if farm level yields are not available or are unavailable in sufficient sample size to use in a flexible specification for a yield density function. Specifically, results suggest modeling farm yields as normal deviations from NASS-based aggregate yield values holds promise. Still, our results cannot be generalized without some limitations. One is the ratio of farm yields to county yields standard deviations varies across crops and counties. Hence, one specific factor relating farm to county yields likely does not exist. However, the relationships that we estimate between the coefficient of variation of farm level yield and county level yield do tend to be within the general ballpark of one another, both by crop and by States, suggesting that out-of-sample application of our results would still be an improvement over ignoring farm level yield variability in economic analysis.

Moreover, the dataset for farm yields with the largest geographic scope is that which can be generated from crop insurance premiums from the Risk Management Agency. The yields inferred from this data tend to not show a strong relationship to farm yields from the farm management databases. The irony of the farm management versus RMA datasets is that the former may have a tendency to draw farmers with lower than average yield variability – say due to better than average management skills – given that membership in these associations may itself be an indication of a strong interest in farm management.[6] The RMA data may be more influenced by the opposite group, particularly in areas with lower levels of participation in crop insurance. Baring the development of a longitudinal and random sample of farm yield, judging which approach gives more accurate indications of farm level yield variability is difficult, if not impossible. It remains to be seen in what general direction yield volatilities implied by RMA premiums will move over time.

[6] Another difference between the RMA and farm management data is that RMA data is based on farm unit data that tends to be less aggregated than farm-level. For example, basic and optional units used by RMA are smaller than "farms" as defined in the farm management data sets.

References

Coble, K., and R. Dismukes. "Distributional and Risk Reduction Effects of Commodity Revenue Program Design", *Review of Agricultural Economics* v.30, n. 3 (Fall 2008): 543-553.

Cooper, J. "Payments under the Average Crop Revenue Election Program - Implications for Government Costs and Producer Preferences," accepted for presentation as an AAEA Selected Paper at the 2009 AAEA & ACCI Joint Annual Meeting in Milwaukee, Wisconsin, July 26-July 28.

Goodwin, B. "Payment Limitations and Acreage Decisions under Risk Aversion: A Simulation Approach," American Journal of Agricultural Economics, V. 91, n. 1 (February 2009): 19-41.

Langemeier, M. "Comparison of 2002 Census and KFMA Farms," Staff Paper No. 06-01, Department of Agricultural Economics, Kansas State University, December 2005.

Risk Management Agency. "RMA Premium Calculation," USDA, accessed April 2009. http://www3.rma.usda.gov/apps/premcalc/

Schnitkey, Gary D., Bruce J. Sherrick, and Scott Irwin. "Evaluation of Risk Reductions Associated with Multi-Peril Crop Insurance Products." *Agricultural Finance Review*, Spring, 2003, pp. 1-21.

Table 1. Standard deviation and mean of yield (bu./acre) for the Kansas and Illinois data at different levels of aggregation

Crop (State)	Standard deviation (bu/acre) Farm	Standard deviation (bu/acre) County	Standard deviation (bu/acre) State	Mean (bu/acre) Farm	Mean (bu/acre) County	Mean (bu/acre) State	Counties with 10 or more sample farms	Number of farms	Ratio of farm to county standard deviation of yield: Mean	Ratio of farm to county standard deviation of yield: Standard deviation	Ratio of farm to State standard deviation of yield: Mean	Ratio of farm to State standard deviation of yield: Standard deviation
I. 10 year data sets												
Corn (IL)	21.49	16.76	11.58	169.27	164.00	168.77	72	2674	1.42	0.85	1.86	1.16
Soybean (IL)	6.80	5.25	4.14	50.06	48.94	49.94	70	2673	1.35	0.39	1.64	0.47
Corn (KS)	31.72	28.37	21.58	106.94	105.79	100.39	7	82	1.12	0.22	1.46	0.31
Wheat (KS)	10.23	8.41	5.65	43.75	43.93	42.84	17	291	1.31	0.38	1.81	0.58
Soybean (KS)	10.49	9.39	7.94	28.02	28.14	27.76	20	273	1.13	0.22	1.32	0.30
II. 17 year datasets												
Corn (IL)	26.22	21.54	16.93	174.92	173.03	172.53	40	823	1.26	0.26	1.55	0.33
Soybean (IL)	6.70	5.10	3.60	50.56	50.22	49.99	40	768	1.38	0.31	1.86	0.43
Wheat (KS)	13.22	10.71	8.41	38.70	38.42	38.67	8	110	1.17	0.19	1.48	0.24
Soybean (KS)	9.64	7.97	7.27	27.28	27.38	28.88	4	52	1.21	0.27	1.33	0.27

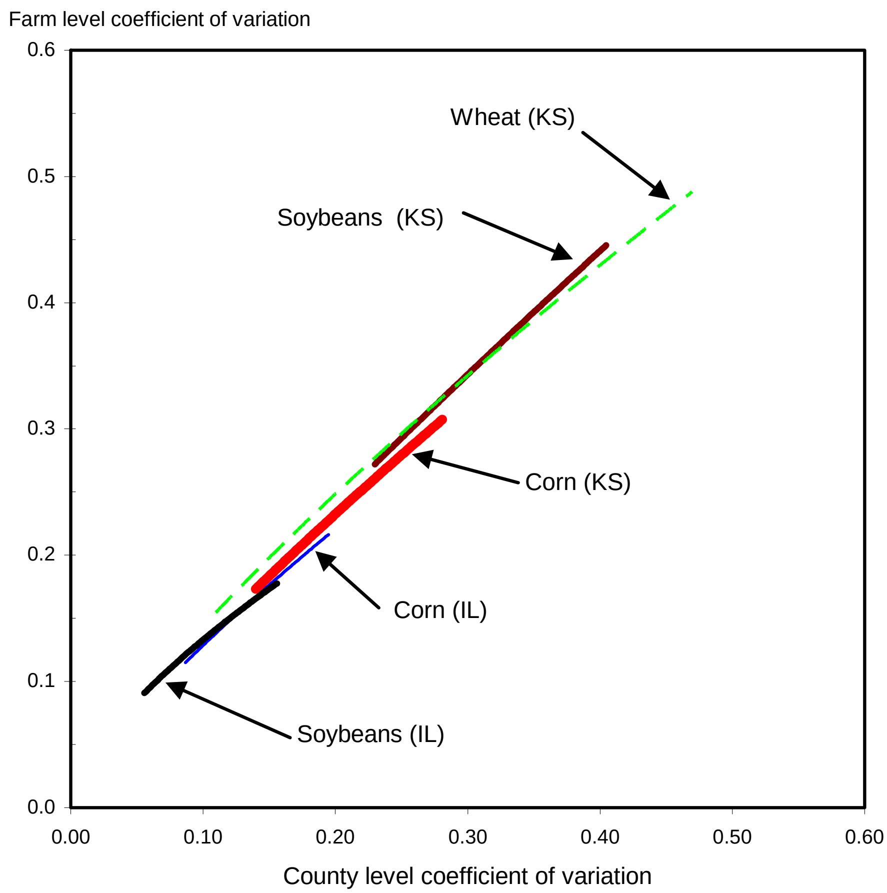
Notes: Data only includes observations in counties with at least 10 sample farms. Data in section I covers 1997 to 2006. Data in section II covers 1991 to 2007. For the 17 year data set for Kansas corn, no counties had at least 10 sample farms. The Illinois yield data is detrended to 2007.

Table 2. OLS Regressions for the coefficient of variation of farm level yield regressed on the coefficient of variation of county level yield[a]

Coefficient	Corn Illinois	Corn Kansas	Soybeans Illinois	Soybeans Kansas	Wheat Kansas	All crops and States	All crops and States except IL soybeans
Intercept	-0.253	-0.134	-0.515	-0.020	-0.120	-0.076	-0.014
	(2.72)	(0.52)	(5.16)	(0.23)	(1.40)	(3.39)	(0.54)
Ln(coefficient of variation of county level yield)	0.781	0.823	0.652	0.871	0.790	0.846	0.885
	(17.94)	(4.26)	(15.39)	(11.08)	(16.17)	(77.35)	64.82
Adjusted R^2	0.531	0.185	0.235	0.309	0.473	0.728	0.741
Error sum of squares	40.79	5.62	42.60	13.66	23.64	130.33	85.84
F value	322	18	237	123	262	5982	4202
Sample size	823	82	768	273	291	2238	1470

Notes:
Data sources are the same as for the 10-year datasets in Table 1. The numbers in parenthesis are the regression coefficients divided by their standard error.
a The dependent variable is the log of the coefficient of variation of farm level yield.

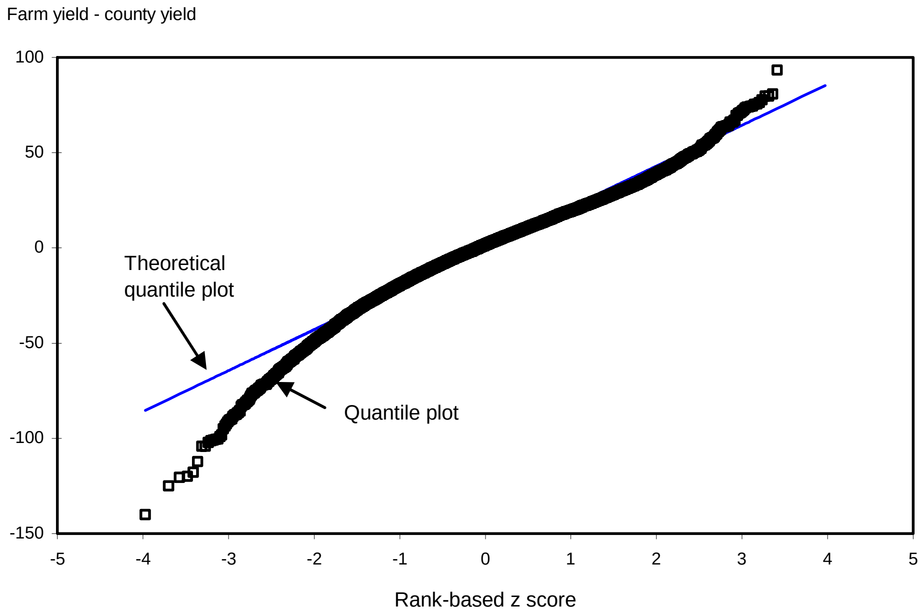
Figure 1. Estimated Relationship between Farm Level and County Level Coefficient of Variation of Yield, Illinois and Kansas



Notes:
Data sources are the same as for the 10 year datasets (Kansas) and 17 year datasets (Illinois) in Table 1.
Values fitted over the observed data ranges using the log-log regression estimates from Table 2.

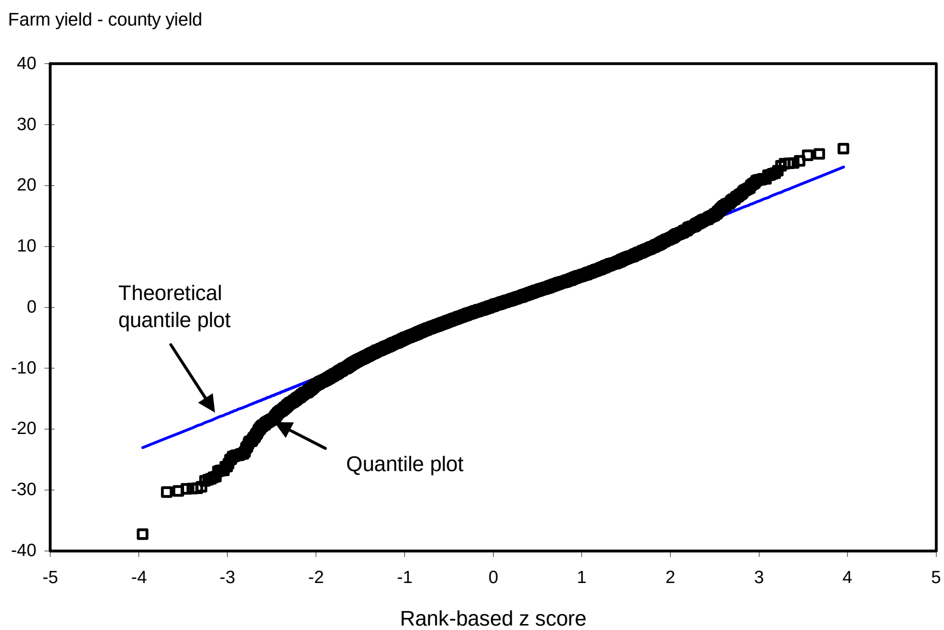
Figure 2. Normal Quantile Plot for Farm Level Minus County Level Yield (Illinois)

a) Corn



Note: Data covers 1991 to 2007 and only includes observations for counties with 10 or more sample farms. Mean and standard deviation of the yield difference are 0 and 21.46, respectively.

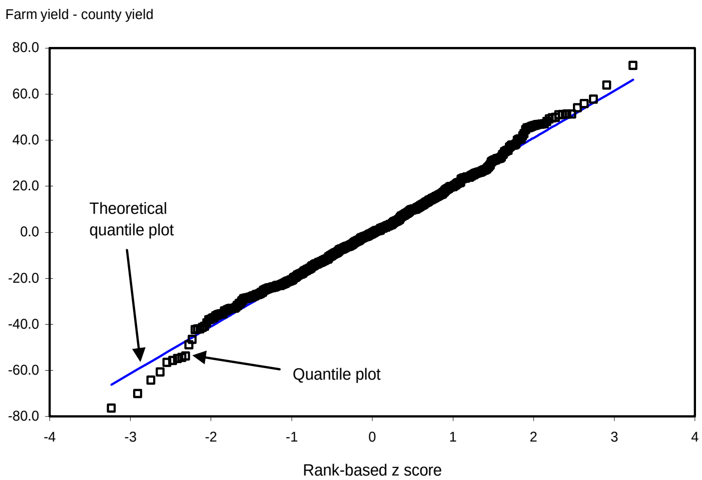
b) Soybeans



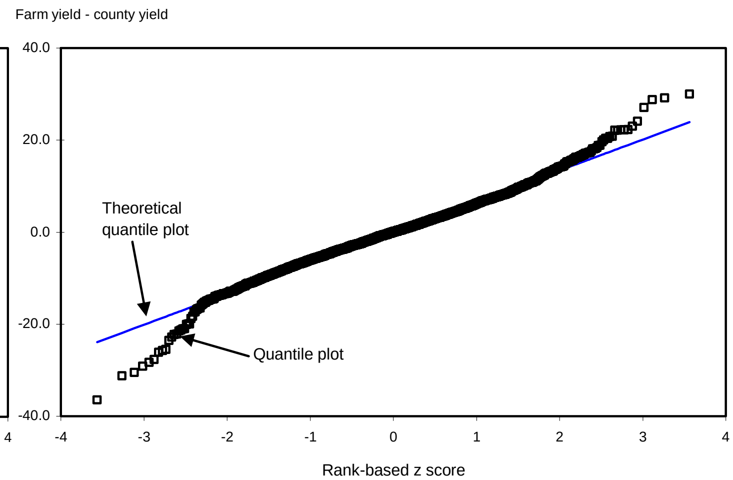
Note: Data covers 1991 to 2007 and only includes observations for counties with 10 or more sample farms. Mean and standard deviation of the yield difference are 0 and 5.82, respectively.

Figure 3. Normal Quantile Plots for Farm Level Minus County Level Yield (Kansas)

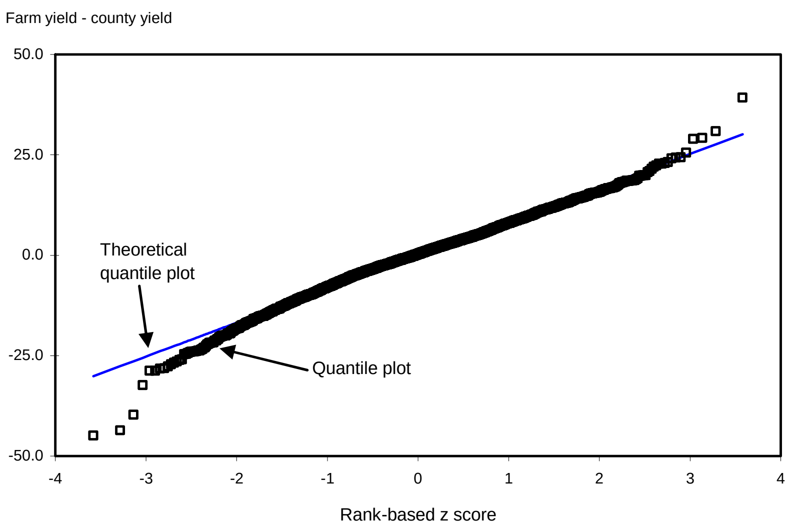
A. Corn



B. Soybeans



C. Wheat



Notes. Data covers 1997 to 2006 and only includes observations for counties with 10 or more sample farms. For corn, the mean and standard deviation of the yield difference is 0 and 20.5, respectively. For soybeans, the mean and standard deviation of the yield difference is 0 and 6.70, respectively. For wheat, the mean and standard deviation of the yield difference is 0 and 8.41, respectively.

Table 3. Parameters used in the insurance simulations

	Corn	Soybeans	Wheat
Revenue Assurance base price, 2009 ($/bu.)	$4.05	$8.80	$8.77($6.20)[a]
Expected price for simulation ($/bu.)	$4.05	$8.80	$5.85
Farm's APH yield (bu./acre):			
Logan County, Illinois	180	51	--
Butler County, Kansas	100	34	34
Finney County, Kansas	160	50	37
Barnes County, North Dakota	113	31	45[b]

Notes:
[a] Winter wheat (Spring wheat) $/bu .
[b] Spring wheat.
Sources: USDA's National Agricultural Statistics Service and Risk management Agency; futures markets.

Table 4. RMA Federal crop insurance premiums and their inferred ratios of farm to county standard deviation of yield versus results based on the farm management data (2009 crop year)

Farm location	Crop	RMA full premium ($/acre)	RMA implied ratio of farm to county standard deviation of yield	Farm management ratio of farm to county standard deviation of yield (State average)	Full premium based on farm management data ($/acre)
Logan,	corn	23.76	1.91	1.26	7.65
IL	soybeans	14.56	3.39	1.38	0.77
Butler,	corn	46.73	1.40	1.12	32.49
KS	soybeans	30.02	1.39	1.31	26.52
	wheat	32.98	1.64	1.13	22.88
Finney,	corn	53.20	4.87	1.12	0.86
KS	soybeans[b]	39.12	3.16	1.31	3.91
	wheat[b]	33.76	1.17	1.13	32.63
Barnes,	corn	60.59	1.54	1.26[c]	40.49
ND	soybeans	24.32	2.45	1.38[c]	5.46
	wheat	18.27	1.77	1.13[d]	7.71

Notes

[a] Revenue assurance with base price option, 70% coverage (source, RMA/USDA). These are the full premiums unsubsidized by the Federal government, i.e., (1-0.41)*farmer premium.

[b] Premiums for irrigated acres.

[c] Illinois corn and soybean standard deviation of yield ratios are used for the North Dakota ratios.

[d] Kansas wheat's standard deviation of yield ratio is used for the North Dakota ratio.