Current Issues in Modeling Yield and Price Risk: Implications for the Design and Rating of Crop Insurance Contracts

Barry K. Goodwin
(With Sujit Ghosh, Ying Zhu, and Ying-Erh Chen)

North Carolina State University

January 12, 2011

2011 workshop of the ERCA Research Network on the Structure and Performance of Agriculture and Agri-Products Industry (SPAA)
Introduction

- Publicly subsidized crop insurance continues to increase in prominence as a means of providing assistance to agriculture
  - Large programs in US, Canada
  - Many new programs being developed around the world
  - See recent World Bank book that summarizes crop insurance development around the world
- In US, 2011 program will exceed $90 billion in total insured liability and $10 billion in premium
- Crop insurance presents the applied economist/statistician with an excellent opportunity to work with data, frontier methods, and real world policy issues
Objectives of This Presentation

- Provide a quick overview of the US crop insurance program
- Highlight three current research topics central to modeling yield, price, and revenue risks
  - Time-varying distributions
  - Modeling multivariate (non-independent) sources of risk
  - Designing multi-year insurance contracts
Some Perspective in the US: The Current Farm Bill (Shields, Monke, and Schnepf (CRS))

Figure 1. Farm Safety Net Programs Authorized Under the 2008 Farm Bill and Other Legislation
(average annual projected outlays by the Congressional Budget Office)

Risk Management
- Yield-based
- Revenue-based
- Whole-farm
- Non-insured disaster assistance (NAP), $0.092 bil.

Commodity Programs
- Field crops, $6 bil.
- Counter-cyclical payments (CCP), $0.559 bil.
- Average Crop Revenue Election (ACRE), $0.311 bil.
- Marketing Assistance Loan Program (MAL), $0.225 bil.

Direct payments (DP), $4.9 bil.

Farm Safety Net
- $15 bil.
- Supplemental Revenue Assistance Payments Program (SURE)
- Ad hoc disaster payments
- Emergency Assistance for Livestock, Honey Bees, and Farm-Raised Fish Program (ELAP)
- Emergency (FM) disaster loans
- Livestock Indemnity Program (LIP)
- Livestock Forage Disaster Program (LFP)
- Tree Assistance Program (TAP)
- Loan deficiency payments (LDP)
- Marketing loan gains (MLG)

Source: Congressional Research Service (CRS)
Crop Insurance

- Becoming pervasive support instrument
- Almost always heavily subsidized and uptake limited without large subsidies
- In US, current program:
  - > 1.1 million policies
  - ≈ 275 million acres insured
  - ≈ $90 billion in liability
  - ≈ $10 billion in premium with ≈ $6 billion in premium subsidy
  - Most participation in individual revenue insurance plans at high coverage levels on aggregated units
  - Over long-run, farmers receive ≈ $2 in indemnities for every $1 paid in premium

Goodwin (NCSU): January 12, 2011
Federal Crop Insurance

- Private insurance available in U.S. since 1879 (usually single-peril)
- Earlier origins in Scotland (1780) and Germany (1797)
- Federal crop insurance program introduced in 1938
- Editorials of the day:
  - *Christian Science Monitor*: “Will the program become, in effect, an underwriting of high-risk farming areas which, in fact, ought to be retired from farming... instead of burdening steadier farms with cutthroat competition in good years and a demand on them for assistance in bad years?”
  - *Barron’s*: “[don’t let it become]...a subsidy to the politically important agricultural industry.”
Why is the Government Involved?

- Multiple peril crop insurance is nearly *always* subsidized
- Is there a persuasive case for market failure?
  - Systemic risks—agricultural risks are not sufficiently diversifiable and private reinsurance cannot cover this magnitude of systemic risk.
  - Transactions costs—the government can do it cheaper.
  - The multiple-peril nature of crop insurance leads to insurmountable problems of moral hazard and adverse selection—private plans will fail.
- Social and political objectives:
  - In developing economies, insurance is seen as a way to improve access to credit.
  - Often a mechanism for transferring welfare from taxpayers to farmers ($1.80 paid for $1 premium)
Market Failure Considerations

- Few persuasive arguments for market failure.
- As for systemic risk, consider the Credit Default Swap market—
  - Current outstanding notional debt is $2.75 trillion
  - (the late) Lehman Brothers held $72 billion alone
- Unlikely that government bureaucracy will have efficiency advantages.
- Theory predicts risk-averse agents will fully insure at actuarially-fair rates
  - The fact that large subsidies have been needed to achieve even modest levels of participation raises questions
  - Worldwide experience with all-risk insurance has realized low rates of participation
  - Prior to 1994 CIRA, participation around 10-30%
- The “Samaritan Effect”—Since 1985, nearly $30 billion in ad-hoc disaster assistance
Externalities

- In some cases, transactions costs may preclude private market solutions
- A good example involves plant and animal diseases and deliberate threats to food supplies
  - Agents may be unwilling to report disease or take preventative measures without government involvement (through coercion or persuasion—i.e., fines or subsidies)
  - In some cases, disease may spread rapidly (e.g., Soybean Rust and Citrus Canker)
  - Private contracts to address disease risks too costly to develop, monitor, and enforce
- We have considered two case studies—
  - Soybean rust
  - Citrus canker
### Federal Crop Insurance Corp
#### Summary of Business Report for 2008 thru 2011
As of 01-10-2011

*(Net Acre and Dollars in Thousands)*

<table>
<thead>
<tr>
<th></th>
<th>2008 Crop Year To Date</th>
<th>2009 1 Year Ago To Date</th>
<th>2009 Crop Year To Date</th>
<th>2010 1 Year Ago To Date</th>
<th>2010 Crop Year Prev Week</th>
<th>2010 Crop Year To Date</th>
<th>2011 Crop Year Prev Week</th>
<th>2011 Crop Year To Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policies with Premium</td>
<td>1,149,291</td>
<td>1,171,354</td>
<td>1,171,972</td>
<td>180,071</td>
<td>1,138,782</td>
<td>1,139,085</td>
<td>119,183</td>
<td>138,269</td>
</tr>
<tr>
<td>Units with Premium</td>
<td>3,023,185</td>
<td>2,728,875</td>
<td>2,729,675</td>
<td>455,290</td>
<td>2,569,436</td>
<td>2,570,233</td>
<td>281,719</td>
<td>344,185</td>
</tr>
<tr>
<td>Net Acres Insured</td>
<td>272,272</td>
<td>264,584</td>
<td>264,764</td>
<td>57,666</td>
<td>255,827</td>
<td>255,900</td>
<td>26,915</td>
<td>39,912</td>
</tr>
<tr>
<td>Liability</td>
<td>89,897,337</td>
<td>79,495,875</td>
<td>79,573,760</td>
<td>9,816,144</td>
<td>77,897,512</td>
<td>77,932,871</td>
<td>7,600,174</td>
<td>8,421,296</td>
</tr>
<tr>
<td>Total Premium</td>
<td>9,851,159</td>
<td>8,940,192</td>
<td>8,949,428</td>
<td>961,892</td>
<td>7,572,919</td>
<td>7,576,764</td>
<td>670,678</td>
<td>819,365</td>
</tr>
<tr>
<td>Subsidy</td>
<td>5,690,844</td>
<td>5,419,600</td>
<td>5,426,178</td>
<td>605,323</td>
<td>4,697,390</td>
<td>4,699,715</td>
<td>426,481</td>
<td>517,780</td>
</tr>
<tr>
<td>Indemnity</td>
<td>8,680,382</td>
<td>4,032,077</td>
<td>5,218,289</td>
<td>1,772</td>
<td>3,184,180</td>
<td>3,374,147</td>
<td>127</td>
<td>135</td>
</tr>
<tr>
<td>Loss Ratio</td>
<td>0.88</td>
<td>0.45</td>
<td>0.58</td>
<td>0.00</td>
<td>0.42</td>
<td>0.45</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
US Crop Insurance

Goodwin (NCSU): January 12, 2011
Modeling Yield and Price Risk
US Crop Insurance—Transfers to Farmers

Loss-Experience: 1981-2007 Ratio of Indemnity Payments to Farmer-Paid Premiums Weighted Across Years by Annual Insured Acreage
An Aside: Yield Guarantee Drag

McClean County, IL (Corn Rotation)
Key parameters of any insurance contract—the guarantee and the premium rate

Both are intimately related to the probability density describing the variable or event being insured

Consider a standard yield plan that guarantees a certain proportion $\lambda$ of expected yield $\mu$.

The actuarial premium is set to equal expected losses:

$$E(\text{Losses}) = Pr(y < \lambda \mu)(\lambda \mu - E(y|y < \lambda \mu)),$$ \hspace{1cm} (1)

which, for a pdf $f(y)$, is given by:

$$E(\text{Losses}) = \int_{0}^{\lambda \mu} f(y)dy \left[ \lambda \mu - \frac{\int_{0}^{\lambda \mu} yf(y)dy}{\int_{0}^{\lambda \mu} f(y)dy} \right],$$ \hspace{1cm} (2)

In rate terms, $Rate = \frac{E(\text{Losses})}{\lambda \mu}$
Issues in Modeling $f(y)$

- What are the conditioning factors?
- What is the appropriate representation of $f(y)$?
  - Parametric (if so, which family?) Beta is overwhelming choice but not always appropriate.
  - Nonparametric (Goodwin and Ker, 1998) and semi-parametric densities
- Does the nature of the data permit assumptions of iid distributions?
  - If not, what is the nature of the correlation?
  - Is correlation state-dependent (key to systemic risk)?
  - How do we characterize spatio-temporal relationships?
  - Multivariate risks—how do we characterize dependence (much research on copula methods)?
  - Revenue ($P \times Q$), multi-year contracts, multi-crop contracts
- Structural changes over time (big issue in modeling crop yields).
- Latent catastrophic risks (e.g., the 1 in 100 year flood event) and loading
A (Quick) Overview of Research on Time-Varying Distributions

- Work with Ying Zhu and Sujit Ghosh of NCSU
- It is very common to work with yields observed over time (e.g., NASS county averages)
- Problem is that technology has changed considerably over time
- Common (nearly universal) approach is two-step:
  - Detrend data (e.g., linear or quadratic regression)—
    \[ y_t = \alpha + \beta T + e_t \]
  - Ignore step 1 and proceed to work with detrended data—
    - Additive deviations:
      \[ \tilde{y}_t = \hat{y}_{2009} + \hat{e}_t \]
    - Proportional deviations:
      \[ \tilde{y}_t = \hat{y}_{2009}(1 + \frac{\hat{e}_t}{\hat{y}_t}) \]
- Analogous to generated regressor problem—treats estimated yields as though directly observed
Scatter Plot of Corn Yield over Time---Adair County, Iowa
An Alternative

- Allow shape, scale, and location parameters of the density of interest to vary in a systematic way over time.
- Very standard to adopt a Beta distribution for crop yields:

\[
f(y) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)(\gamma - \sigma)^{\alpha+\beta-1}} (y - \sigma)^{\alpha-1}(\gamma - y)^{\beta-1}
\]

where \((\alpha, \beta)\) are shape parameters and \((\gamma, \sigma)\) scale and location parameters.

- The idea is to let these parameters vary as a function of time (we use linear and quadratic):
  \[\alpha_t = \exp(g(t, \theta))\] ensures positive values.
- We use MLE to estimate the parameters and generate a density that evolves over time.
Time-Varying Beta

Empirical Models of Yield Risk

Goodwin (NCSU): April 17, 2009
Summary of Results

- Applied to county NASS yields for corn, cotton, soybeans, and wheat
- Outperforms more conventional (2-Step) models in terms of
  - In-sample predictive power
  - Out-of-sample predictive power
  - Estimation of premium rates in simulation model
We are often interested in modeling multiple (correlated) sources of risk

Copula means “join, link, tie, bond, etc.”

If we are able to accurately characterize the parametric form of the relevant joint distribution, the problem is easy

However, it may be difficult to analytically derive an adequate joint distribution

For example, when:

- Marginals are from different parametric families
- Components of each marginal are taken from different sources (more on this below)
- One or more marginals is nonparametric
The Basic Idea

- A copula is a function that joins the marginal distribution functions to form the multivariate distribution function.
- Copulas make use of the fact that if \( x \sim F \) then \( F(x) \sim U(0, 1) \).
- To define a copula, consider \( m \) uniform (on the unit interval) random variables \( u_1, u_2, ..., u_m \). The joint cdf of \( m \) uniform random variables is:

\[
C(u_1, u_2, ..., u_m) = \text{Prob}(U_1 \leq u_1, U_2 \leq u_2, ... U_m \leq u_m) \quad (1)
\]

where \( C \) is a “copula,” which is unique for continuous cdf’s.
- For an \( m \)-variate function \( F \), the copula associated with \( F \) is a distribution function \( C : [0, 1]^m \rightarrow [0, 1] \) that satisfies

\[
F(y_1, ..., y_m) = C(F_1(y_1), ..., F_m(y_m); \theta), \quad (2)
\]

where \( \theta \) is a set of parameters that measures dependence.
Development of the Literature:

- Relatively new area of research
- Many financial and actuarial mechanisms make applied use of copulas
- Seminal work that introduced concept is by Sklar (1959)
- Famous paper that stimulated interest is by Geneset and MacKay (1986)
- Schweizer and Wolff (1981) linked copulas to models of dependence
- Great book by Nelsen (1999), *An Introduction to Copulas*
Applications:

- In pricing annuities, actuaries have noted the “broken heart syndrome,” where the death of a mate substantially increases the probability that their spouse will die in a given amount of time.

- “Systemic Risk,” which we hear so much about in the current financial situation, is an example of a situation where correlation may differ according to the state of individual random variables.

- When the housing market went bad, everything else went bad as well.

- In agriculture:
  - Crop revenue insurance (revenue = price * yield)
  - Spatial correlation of crop yields
  - Whole farm insurance
  - Pricing portfolios
The formula that felled Wall St
By Sam Jones
Published: April 24 2009 22:04 | Last updated: April 24 2009 22:04

Johnny Cash and June Carter met backstage at the Grand Ole Opry. It was a little like a country song: he was married, she recently divorced, and an affair ensued; both singers had young children, and Cash would have three more with his first wife before she left him in 1966, citing his drinking and carousing. Two years later, he proposed to Carter on stage and, despite having turned him down numerous times before, she accepted. They’d each found a life match.

It ended like a country song, too. In 2003, Carter died in Nashville of complications from heart surgery, and Cash followed her to the grave four months later. The heart complication for him, it seemed, was that it was broken: “It hurts so bad,” he told the audience at the last concert he would give. The pain, he said, tuning his guitar, close to tears, was “the big one. It’s the biggest.”
Different Correlation Scenarios:

Modeling Dependence Using Copulas
State Dependence in Spatial Correlation (Goodwin 2001):

Figure 2. Pearson correlation coefficients vs. distance: normal yield years

Figure 3. Pearson correlation coefficients vs. distance: extreme yield years
Practical Issues:

- To model a copula, I need:
  - Specification of the marginals (and estimates of the relevant parameters)
  - Specification of the copula function, which includes
  - Some measure of dependence (Pearson Correlation, Kendall’s Tau, etc.)
- Different copula functions imply different forms of dependence
- Two common approaches to estimation:
  - Standard MLE applied to the copula and marginals jointly
  - A two-step approach—
    - First estimate the marginals
    - Then estimate the copula conditional on estimates of marginal parameters (called “inference functions for margins” or IFM)
- R has several nice packages to work with copulas
- SAS also has implemented copulas (e.g., in proc model
- Higher ordered joint distributions become complicated, difficult, and subject to the curse of dimensionality
- “Archimedean” copulas are one class that have relatively
Examples of Copula Functions:

A Practical Example: Pricing Revenue Insurance

- Two approaches to pricing revenue coverage:
  - Collect a series of revenue realizations and estimate a univariate marginal
  - Measure yield risk from one source (history) and price risk from another (options) and combine the information using a copula or a measure of dependence

- Step 2 has advantages if market volatilities and price expectations are time-varying

- Revenue = Price * Yield

- Of course, if price and yield are correlated, $E_{Revenue} \neq E_{Price} \times E_{Yield}$

- The “natural hedge” suggests inverse correlation

- Common to assume constant linear correlation, but what if it is not?

- As we observed, it may be that correlation is “state-dependent”

- If so, assuming constant correlation may result in pricing errors
As an alternative to using functional copulas, one approach is to reorder the data to achieve a desired correlation structure. Iman and Conover suggested a method for doing so. The method uses a scoring function usually based on a standard normal function. The fact is, however, this is not an alternative to adopting a copula but rather is entirely analogous. Use of a normal score function in implementing the IC method is entirely equivalent to assuming a Gaussian copula with constant linear correlation. Other score functions correspond to different copulas and can have very important implications for the results.
Examples of Coupla Functions:

(Mildenhall, CAS Report of the Research Working Party on Correlations and Dependencies Among All Risk Sources, 2005.)

Figure 4.1: Bivariate distributions with normal, uniform and exponential scores.
A Postscript:

Systemic Risk and the Financial Meltdown
The Discussion:

60 Minutes looked at one of the selling documents of such a security with Frank Partnoy, a former derivatives broker and corporate securities attorney, who now teaches law at the University of San Diego.

“It’s hundreds and hundreds of pages of very small print, a lot of detail here,” Partnoy explains. Asked if he thinks anyone ever reads all this fine-print, Partnoy says, “I doubt many people read it.”

These complex financial instruments were actually designed by mathematicians and physicists, who used algorithms and computer models to reconstitute the unreliable loans in a way that was supposed to eliminate most of the risk.

“Obviously they turned out to be wrong,” Partnoy says. Asked why, he says, “Because you can’t model human behavior with math.”
A Multiyear Crop Insurance Plan

Multiyear Crop Insurance Plans

• Lower actuarially fair premium rate

• A hybrid of single year and multiyear coverage

  It guarantees $\gamma \mu$ and $\beta \cdot 2\mu$

  $\mu$: expected yield

  $\gamma$: single year coverage (partial payment)

  $\beta$: two-year coverage (total indemnity)

• Partial Payment in each insured period
  a) help producers to repay operation loans
  b) assure yield loss will not bankrupt producers
Simulation Result

Relationship between the Actuarially Fair Premium Rate and Correlation of Yield Across Years

Note: The rate for a single year plan is 0.165, which is the same as rates for multiyear plans when \( \rho = 1 \)
Empirical Analysis (County Data)

Summary

- **Pearson’s correlation coefficient:**
  - Scott (Iowa), Logan (Illinois) have negative correlation
  - Union (Indiana), most counties in Ohio have positive correlation
- **P-value:**
  - Carroll (Indiana) is significantly (negatively) correlated when significant level is 0.05 (p-value=0.04)
  - No p-value is significant when significant level is 0.025 or 0.01
- **Yields across years are not significantly correlated**
Empirical Analysis- Farm Data

Data

- **State and Crop:**
  - Corn and soybean (Iowa, Illinois, Indiana, Ohio)
  - Wheat (Montana, Kansas, Oklahoma, Texas)
  - Barley (Montana, North Dakota)
  - Cotton (Texas)
  - Grain (Kansas, Nebraska)

- **Periods:**

- **Data Information Includes:**
  - yearly yield, insured acres, county and state location
Summary

• Aggregate to state level (farm-level data):
  Only Texas wheat (p-value=0.025) is significantly negatively correlated at 0.05 significant level

• Aggregate to county level data (farm-level data):
  ➢ Correlation (color map): match some weather patterns (precipitation or temperature) or geographical pattern (elevation)
  ➢ P-value (histogram): most of the correlation is not significant
<table>
<thead>
<tr>
<th><strong>Three-Year Insurance Contract</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman’s rank correlation</td>
<td>$0.013(1&amp;2), 0.015(1&amp;3), 0.014(2&amp;3)$</td>
</tr>
<tr>
<td>Parameter of Beta Distribution</td>
<td>$(\alpha_1, \beta_1) = (9.14, 5.56)$</td>
</tr>
<tr>
<td></td>
<td>$(\alpha_2, \beta_2) = (9.07, 5.53)$</td>
</tr>
<tr>
<td></td>
<td>$(\alpha_3, \beta_3) = (9.00, 5.48)$</td>
</tr>
<tr>
<td>Actuarially Fair Premium Rate</td>
<td>$0.0084$</td>
</tr>
<tr>
<td></td>
<td>$(0.009, 0.022, 0.061$ for 80%, 90%, 100%)</td>
</tr>
<tr>
<td>Expected Yield</td>
<td>155</td>
</tr>
<tr>
<td>Coverage Rate of Guaranteed Yield</td>
<td>70%</td>
</tr>
<tr>
<td>Guaranteed Yield</td>
<td>$325.5 (=3 \times 70% \times 155)$ for total indemnity</td>
</tr>
<tr>
<td>Single Year Coverage</td>
<td>67%</td>
</tr>
<tr>
<td>Guaranteed yield in each year</td>
<td>$103.85 (=67% \times 155)$ for partial payment</td>
</tr>
<tr>
<td>Realized Yield in Year 1, 2 &amp; 3</td>
<td>90, 135, 100</td>
</tr>
<tr>
<td>Partial Payment</td>
<td>$8379.25$</td>
</tr>
<tr>
<td></td>
<td>$(=2.5 \times (103.85 - 90) \times 242)$ in year 1</td>
</tr>
<tr>
<td></td>
<td>$2329.25$</td>
</tr>
<tr>
<td></td>
<td>$(=2.5 \times (103.85 - 100) \times 242)$ in year 3</td>
</tr>
<tr>
<td>Total Indemnity</td>
<td>$0 (325.5 &lt; 103.85 + 135 + 103.85)$</td>
</tr>
</tbody>
</table>
Two- and Three-Year Insurance Plan
(80% Two-Year Coverage and 40% Single Year Coverage)

<table>
<thead>
<tr>
<th>Three-Year Insurance Contract</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Actuarially Fair Premium Rate</td>
<td>0.010</td>
</tr>
<tr>
<td>Expected Yield</td>
<td>155</td>
</tr>
<tr>
<td>Coverage Rate of Guaranteed Yield</td>
<td>80%</td>
</tr>
<tr>
<td>Guaranteed Yield</td>
<td>$248 (=2*80%*155)$ for total indemnity</td>
</tr>
<tr>
<td>Single Year Coverage</td>
<td>40%</td>
</tr>
<tr>
<td>Guaranteed yield in each year</td>
<td>$62 (=40%*155)$ for partial payment</td>
</tr>
<tr>
<td>Realized Yield in Year 1, 2 &amp; 3</td>
<td>90,135</td>
</tr>
<tr>
<td>Partial Payment</td>
<td>0</td>
</tr>
<tr>
<td>Total Indemnity</td>
<td>$13915=2.5(248-90-135)242$</td>
</tr>
</tbody>
</table>
Two- and Three-Year Insurance Plan
(80% Two-Year Coverage and 40% Single Year Coverage)

<table>
<thead>
<tr>
<th>Three-Year Insurance Contract</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Actuarially Fair Premium Rate</td>
<td>0.004</td>
</tr>
<tr>
<td>Expected Yield</td>
<td>155</td>
</tr>
<tr>
<td>Coverage Rate of Guaranteed Yield</td>
<td>80%</td>
</tr>
<tr>
<td>Guaranteed Yield</td>
<td>372 (=3*80%*155) for total indemnity</td>
</tr>
<tr>
<td>Single Year Coverage</td>
<td>40%</td>
</tr>
<tr>
<td>Guaranteed yield in each year</td>
<td>62 (=40%*155) for partial payment</td>
</tr>
<tr>
<td>Realized Yield in Year 1, 2 &amp; 3</td>
<td>90,135</td>
</tr>
<tr>
<td>Partial Payment</td>
<td>0</td>
</tr>
<tr>
<td>Total Indemnity</td>
<td>28435 = 2.5(372-90-135-100)242</td>
</tr>
</tbody>
</table>