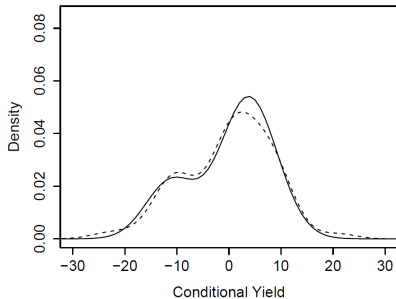


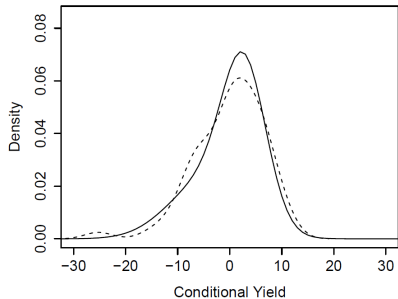
Normal Mixture Model Fitting Examples: Soybeans

Estimated Conditional Crop Yield Densities

Chatnam-Kent Soybeans



Elgin Soybeans

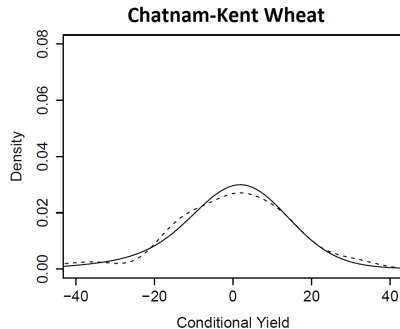
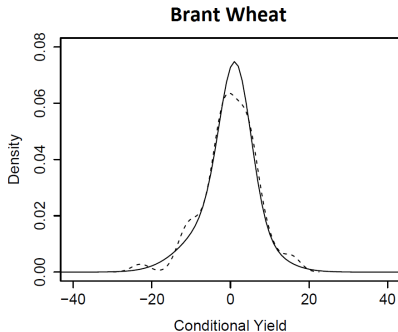


Full lines illustrate EM-estimated normal mixture model

Dashed lines illustrate nonparametric kernel density estimate for comparison

Normal Mixture Model Fitting Examples: Wheat

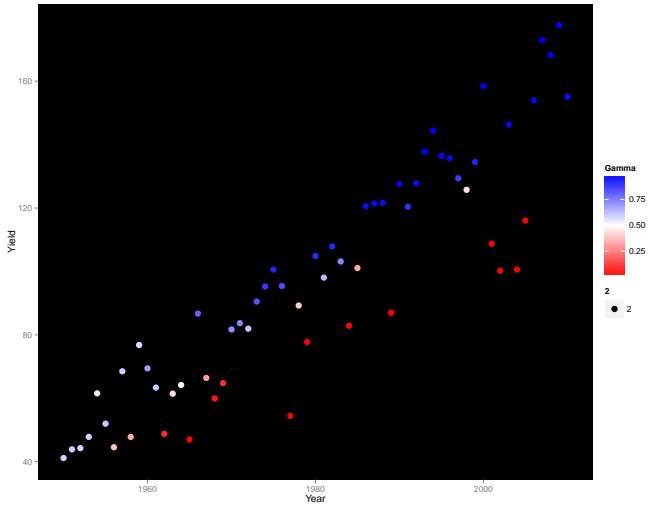
Estimated Conditional Crop Yield Densities



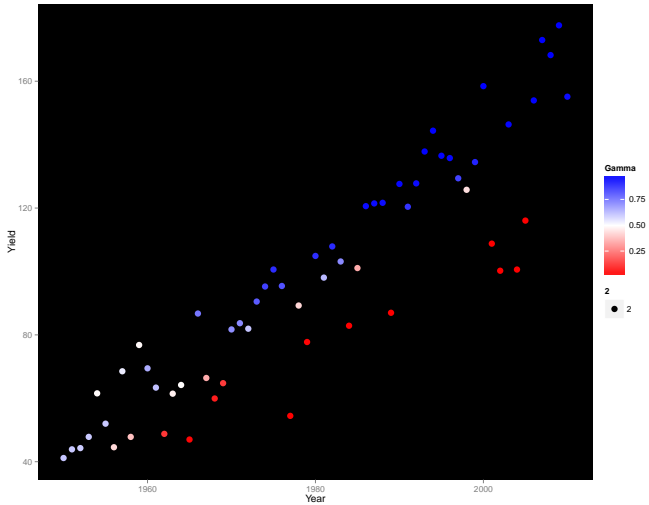
Full lines illustrate EM-estimated normal mixture model

Dashed lines illustrate nonparametric kernel density estimate for comparison

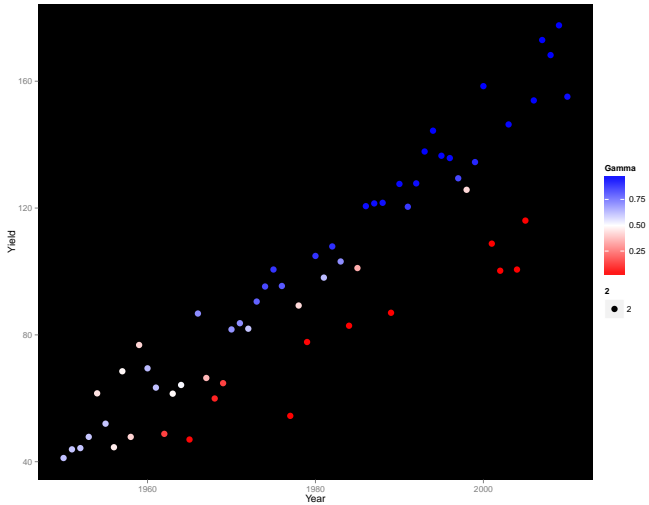
A Closer Look at the E-Step



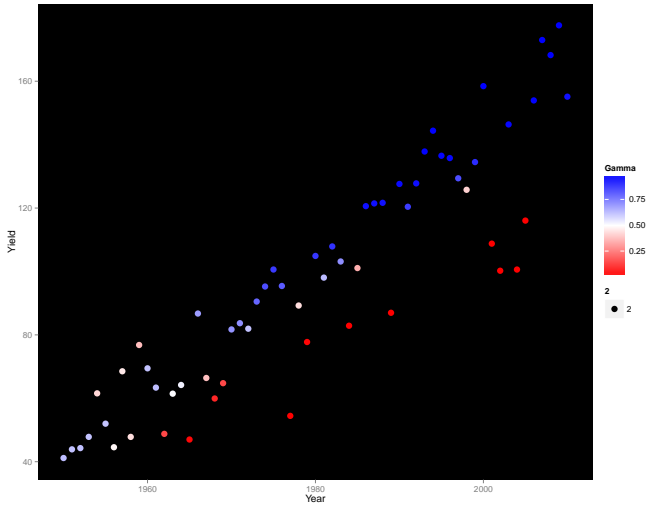
A Closer Look at the E-Step



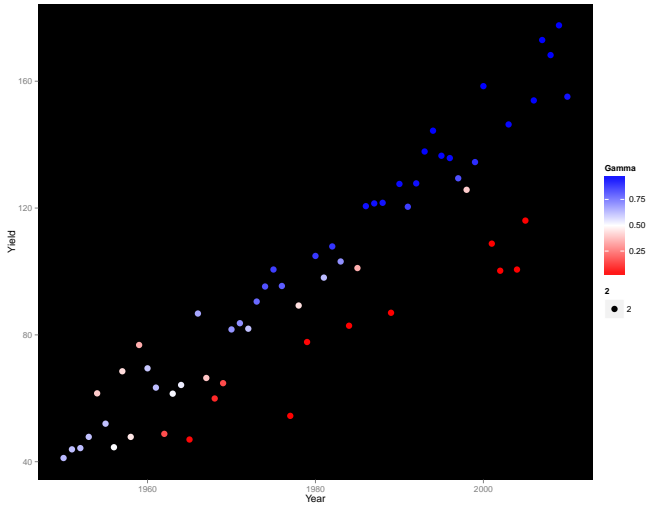
A Closer Look at the E-Step



A Closer Look at the E-Step

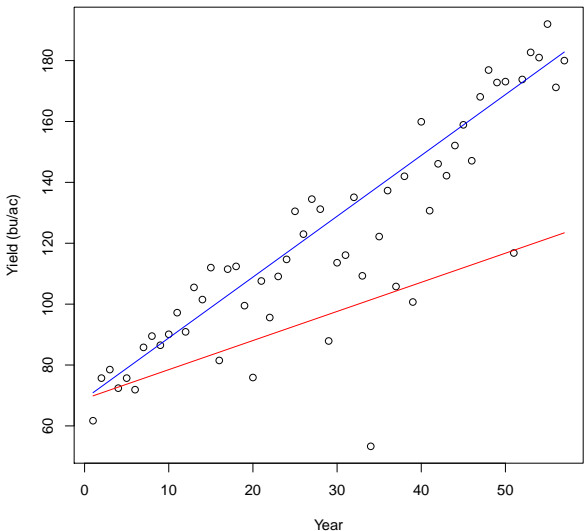


A Closer Look at the E-Step



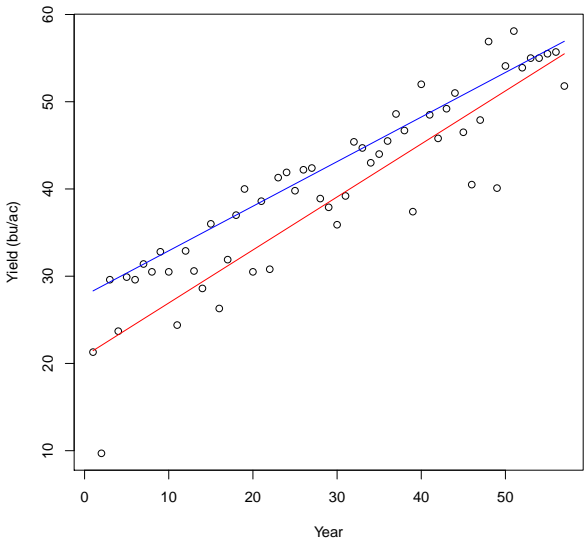
Representative Estimate

Clinton IA Corn



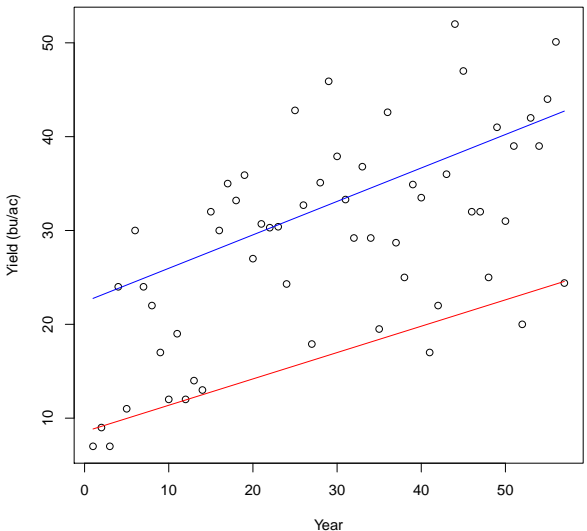
Atypical Estimate

Cherokee IA Soybean



Atypical Estimate

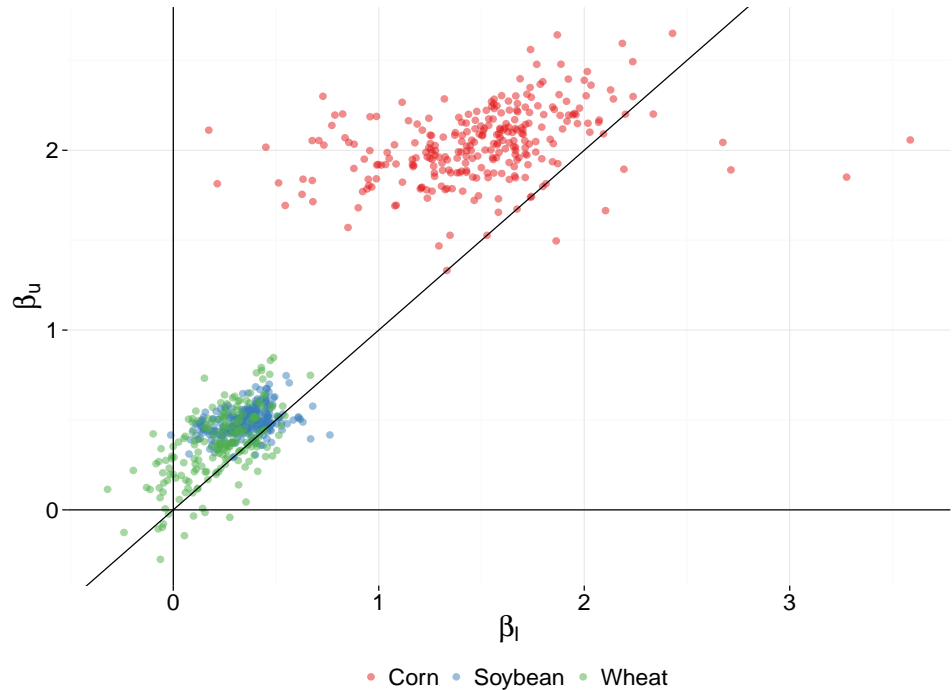
Hodgeman KS Wheat



Different Trends in Two States?

Evidence

- 8.6% have faster rate of technological change in *lower state*
- 4.6% have equal rate of technological change in *upper state*
- 81.5% *upper state* trend is greater 110% of *lower state* trend
- 68.4% *upper state* trend is greater 125% of *lower state* trend
- 37.9% *upper state* trend is greater 150% of *lower state* trend
- 16.7% *upper state* trend is **double** the *lower state* trend



Different Trends in Two States?

Hypothesis Test One:

$$H_o^1 : \beta_\ell = \beta_u$$

$$H_a^1 : \beta_\ell \neq \beta_u$$

Using likelihood ratio test.

Rejection rates (5% significance level):

Corn	Soybean	Wheat	Total
84.4%	82.7%	65.5%	78.0%

Majority **reject** equivalent rate of change in two states

Stable Probability of “Belonging to” Upper State?

Hypothesis Test Two:

Where δ is the estimated slope coefficient from a linear regression of $\gamma = h(t)$

$$H_o^2 : \delta = 0$$

$$H_a^2 : \delta \neq 0$$

Using a t -test with robust standard errors.

Rejection rates (5% significance level):

Corn	Soybean	Wheat	Total
10.8%	5.3%	6.0%	7.5%

Majority **fail to reject** stable expectations vector

Corn Hypothesis Rejection Rates (5% Significance Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Ontario	32	87.50%	0.00%
Illinois	97	81.40%	16.50%
Indiana	79	84.80%	15.20%
Iowa	99	85.90%	5.10%
Sub-total	307	84.36%	10.77%
Total	854	77.99%	7.49%

Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

Soybean Hypothesis Rejection Rates (5% Sign. Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Ontario	6	100.00%	0.00%
Illinois	97	83.50%	4.10%
Indiana	82	80.50%	4.90%
Iowa	98	82.70%	7.10%
Sub-total	283	82.70%	5.28%
Total	854	77.99%	7.49%

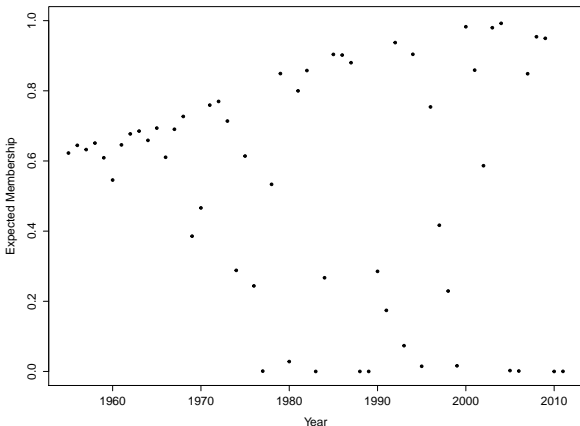
Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

Wheat Hypothesis Rejection Rates (5% Significance Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Ontario	25	80.00%	20.00%
Kansas	93	79.60%	1.10%
Nebraska	50	56.00%	8.00%
Texas	96	53.10%	6.20%
Sub-total	264	65.53%	6.05%
Total	854	77.99%	7.49%

Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

Stable Probability of “Belonging to” Upper State?



- Possible determinants of δ ? (technology, weather, etc.)
- Might be (a) no effect or (b) no *net* effect

An Application: Crop Insurance Rates

Out-of-Sample Simulated Game

- Common technique in crop insurance literature for comparing two rate-setting techniques (Ker & McGowan 2000; Ker & Coble 2003; Racine & Ker 2006; Harri *et al.* 2011.)
- Out-of-sample: mimics real life
- If new method has no advantage, will not perform better

Number of “winning” states:

Corn	Soybean	Wheat	Total
4/6 (1)	4/6 (0)	5/6 (4)	13/18 (5)

Statistically significant wins at the 5% level in brackets

An Application: Crop Insurance Rates

Some Minor Caveats

- Comparison at lower-end of the density tail only
- Not quite an apples-to-apples comparison
 - Two component trend method does not exhibit heteroscedasticity
- Two trend method seems to performs better
- Rate improvement in 13 of 18 crop-state combinations
- Statistically significant improvement in 5 of 18 states
- No statistically significant wins in opposite direction
- Higher bar to exceed than earlier in-sample likelihood-ratio test - power of the test is weaker

Out-of-Sample Rating Game Results: Corn

Set	Coverage	Retained	Psuedo Loss Ratio		p -value	% Payoff
	Level	by Private	Private	Government		
IL	75%	85.9%	0.092	0.026	0.787	3.0%
	90%	87.8%	0.287	0.465	0.012	18.6%
IN	75%	81.1%	0.164	0.134	0.604	3.7%
	90%	82.4%	0.395	0.434	0.373	19.8%
IA	75%	83.9%	0.357	0.409	0.361	6.6%
	90%	86.8%	0.406	0.444	0.291	12.2%

Note: **Winner** and $p < 0.05$ or $p < 0.95$

Out-of-Sample Rating Game Results: Wheat

Set	Coverage Level	Retained by Private	Pseudo Loss Ratio		p -value	% Payoff
			Private	Government		
KS	75%	52.6%	1.297	1.921	0.023	18.5%
	90%	41.9%	1.268	1.399	0.217	33.0%
NE	75%	40.8%	0.309	0.747	0.070	5.3%
	90%	45.3%	0.652	0.611	0.587	19.8%
TX	75%	73.5%	0.957	2.096	0.001	19.5%
	90%	69.1%	1.112	1.593	0.001	43.8%

Note: **Winner** and $p < 0.05$ or $p < 0.95$

Summary

- Crop yields are important measuring stick of agricultural productivity
- Crop science literature suggests yields have distinct subsets of the yield distribution—are the rates of technological change equal across these subsets?
- We propose an empirical model that explicitly allows for unique trends in two states

Conclusions

- 1 Trajectory of technological change
 - different trajectories in different states
 - statistically significant difference in 78% of cases
 - > 85% slower rate of technological change in *lower state*
- 2 Inconclusive results w.r.t. stable probability of the *upper state*
- 3 Mixture model opens the door for a lot of interesting questions. . .
- 4 Results consistent with plant science research expenditures (corn versus beans, racehorse versus suboptimal)
- 5 Results consistent with stylized facts regarding heteroscedasticity
- 6 Results suggest little effect of climate change.
- 7 Results inconsistent with what you would find looking at moments from single technology estimation
- 8 Results are fairly robust within a crop

Going Forward

- Can we isolate the underlying cause(s) of different trends in the two states?
- How different are the results at the farm-level?
 - Trial farm project in initial stages
- What are the determinants of yields “belonging to” the *upper state*?
 - What climatic/weather conditions can we isolate as important?

Thank You

Questions?

Funding for this research was generously provided by the Ontario Ministry of Food and Agriculture (OMAF) and the Institute for the Advanced Study of Food and Agricultural Policy

Operationalizing the Model

Maximum likelihood approach to incomplete data
(Dempster, Laird & Rubin, 1977)

- We want to know the parameters of each state
- Assume there exists an identity variable vector \mathbb{Z} that identifies the states

Operationalizing the Model

Maximum likelihood approach to incomplete data
(Dempster, Laird & Rubin, 1977)

- We want to know the parameters of each state
- Assume there exists an identity variable vector \mathbb{Z} that identifies the states
- If we knew \mathbb{Z} it would be easy to estimate θ_j
- However \mathbb{Z} is latent and we only observe yields y_t

Operationalizing the Model

Maximum likelihood approach to incomplete data
(Dempster, Laird & Rubin, 1977)

- We want to know the parameters of each state
- Assume there exists an identity variable vector \mathbb{Z} that identifies the states
- If we knew \mathbb{Z} it would be easy to estimate θ_j
- However \mathbb{Z} is latent and we only observe yields y_t
- Let $\gamma \in [0, 1]$ be a **weakly-assigned** estimate of \mathbb{Z} called the “expectations vector”
- We can use y_t to estimate γ and the respective parameters of each state in an iterative algorithm (the EM algorithm)

Approaches from the Literature

Technological Trend

Some examples:

- Simple linear trend
- Piecewise linear splines (Skees & Reed 1986)
- Stochastic Kalman filter (Kaylen & Koroma 1991)
- ARIMA (Goodwin & Ker 1998)
- Polynomial trend (Just & Weninger 1999)
- Spatio-temporal approach (Ozaki & Silva 2009)

(But all these approaches estimate the trend at the mean)

Approaches from the Literature

Conditional Yield Density Models

Some examples:

- Normal (Botts & Boles 1958; Just & Weninger 1999)
- Lognormal (Day 1965)
- Gamma (Gallagher 1987)
- Beta (Nelson & Preckel 1989)
- **Mixture of two normals (Ker 1996)**
- Nonparametric kernel densities (Goodwin & Ker 1998)
- Semiparametric (Ker and Coble 2003)
- Logistic (Atwood, Shaik & Watts 2003)
- Weibull (Sherrick et al 2004)

(But all these approaches assume distribution constant over time)

Hypothesis Rejection Rates (5% Significance Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Corn	307	84.36%	10.77%
Soybean	283	82.70%	5.28%
Wheat	264	65.53%	6.05%
Total	854	77.99%	7.49%

Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

Corn Hypothesis Rejection Rates (5% Significance Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Ontario	32	87.50%	0.00%
Illinois	97	81.40%	16.50%
Indiana	79	84.80%	15.20%
Iowa	99	85.90%	5.10%
Sub-total	307	84.36%	10.77%
Total	854	77.99%	7.49%

Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

Soybean Hypothesis Rejection Rates (5% Sign. Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Ontario	6	100.00%	0.00%
Illinois	97	83.50%	4.10%
Indiana	82	80.50%	4.90%
Iowa	98	82.70%	7.10%
Sub-total	283	82.70%	5.28%
Total	854	77.99%	7.49%

Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

Wheat Hypothesis Rejection Rates (5% Significance Level)

	Number of Counties	Null Hypothesis	
		One $\hat{\beta}_P = \hat{\beta}_R$	Two $\delta = 0$
Ontario	25	80.00%	20.00%
Kansas	93	79.60%	1.10%
Nebraska	50	56.00%	8.00%
Texas	96	53.10%	6.20%
Sub-total	264	65.53%	6.05%
Total	854	77.99%	7.49%

Note: H_0^1 evaluated using a likelihood ratio test. H_0^2 evaluated using a t -test with robust standard errors where δ is the estimated slope coefficient of a linear regression $\hat{\gamma}_i = f(t)$.

An Application: Crop Insurance Rates

Out-of-Sample Simulated Game

- Common technique in crop insurance literature for comparing two rate-setting techniques (Ker & McGowan 2000; Ker & Coble 2003; Racine & Ker 2006; Harri *et al.* 2011.)
- Out-of-sample: mimics real life
- If new method has no advantage, will not perform better

Number of “winning” states:

Corn	Soybean	Wheat	Total
4/6 (1)	4/6 (0)	5/6 (4)	13/18 (5)

Statistically significant wins at the 5% level in brackets

An Application: Crop Insurance Rates

Some Minor Caveats

- Comparison at lower-end of the density tail
- Does not reflect ability of the model to fit all the data
- *And* not quite an apples-to-apples comparison
 - Two component trend method more constrained in its heteroscedasticity treatment (minor disadvantage)

An Application: Crop Insurance Rates

Some Minor Caveats

- Comparison at lower-end of the density tail
- Does not reflect ability of the model to fit all the data
- *And* not quite an apples-to-apples comparison
 - Two component trend method more constrained in its heteroscedasticity treatment (minor disadvantage)
- Two trend method performs better (despite minor disadvantage)
- Rate improvement in 13 of 18 crop-state combinations
- Statistically significant improvement in 5 of 18 states
- No statistically significant wins in opposite direction
- Higher bar to exceed than earlier in-sample likelihood-ratio test

Out-of-Sample Rating Game Results: Corn

Set	Coverage Level	Retained by Private	Pseudo Loss Ratio		p -value	% Payoff
			Private	Government		
IL	75%	85.9%	0.092	0.026	0.787	3.0%
	90%	87.8%	0.287	0.465	0.012	18.6%
IN	75%	81.1%	0.164	0.134	0.604	3.7%
	90%	82.4%	0.395	0.434	0.373	19.8%
IA	75%	83.9%	0.357	0.409	0.361	6.6%
	90%	86.8%	0.406	0.444	0.291	12.2%

Note: **Winner** and $p < 0.05$ or $p < 0.95$

Out-of-Sample Rating Game Results: Soybean

Set	Coverage Level	Retained by Private	Pseudo Loss Ratio		p -value	% Payoff
			Private	Government		
IL	75%	77.7%	0.366	0.374	0.431	3.0%
	90%	67.4%	0.540	0.474	0.695	12.9%
IN	75%	86.9%	0.257	0.258	0.479	3.4%
	90%	78.4%	0.611	0.746	0.201	19.6%
IA	75%	80.6%	0.809	0.467	0.757	6.1%
	90%	78.7%	0.751	0.911	0.076	16.1%

Note: Winner and $p < 0.05$ or $p < 0.95$

Out-of-Sample Rating Game Results: Wheat

Set	Coverage Level	Retained by Private	Pseudo Loss Ratio		p -value	% Payoff
			Private	Government		
KS	75%	52.6%	1.297	1.921	0.023	18.5%
	90%	41.9%	1.268	1.399	0.217	33.0%
NE	75%	40.8%	0.309	0.747	0.070	5.3%
	90%	45.3%	0.652	0.611	0.587	19.8%
TX	75%	73.5%	0.957	2.096	0.001	19.5%
	90%	69.1%	1.112	1.593	0.001	43.8%

Note: Winner and $p < 0.05$ or $p < 0.95$