Benchmarking Options for Model-Based County-Level Estimation of Agricultural Cash Rental Rates

Michael E. Bellow[1] , Andreea L. Erciulescu[1,2] , Nathan Cruze[1]

[1]*National Agricultural Statistics Service*
[2]*National Institute of Statistical Sciences*

FCSM Research and Policy Conference
Washington, DC
March 7, 2018

Cash Rents Estimation

- *Cash rent* refers to cropland or pastureland rented for cash on a per acre basis (impacted by crop yield, value of production, market conditions)
- NASS publishes estimates of cash rental rates at different levels for three land use practices (non-irrigated, irrigated and pasture)
- NASS only publishes rental rates (not total rented acres or total rent series)
- Applications include farm program administration, farm rental agreements and agronomics research
- Main data source is Cash Rent Survey (CRS) (conducted by NASS on an annual or biennial basis since 2009)
- Model-based approach developed by Berg, Cecere and Ghosh (2014) has been integrated into the operational estimation procedure

Berg-Cecere-Ghosh (BCG) Method

- Model-based approach involving separate univariate area-level models for average and difference of county-level cash rental rates over two survey years (*Fay-Herriot* type formulation)
- Prior year data incorporated as random variable
- *Covariate index* combines information from the National Commodity Crop Productivity Indexes (NCCPI) , NASS crop yields, total value of production (TVP)
- Closed form estimators (fast computation time)
- Employs version of Ghosh-Steorts difference benchmarking method

Motivation for Benchmarking

- Aggregation of model-based county-level estimates of cash rents to states and *agricultural statistics districts* can be widely different from corresponding *direct* (survey-based) estimates for the larger areas
- NASS requires consistency of county-level estimates with state-level estimates (previously published) and district-level estimates before figures can be released to the public
- Possible protection against model misspecification (Pfefferman, 2013)
- Reduction of bias that may result from modification of outliers (Gershunskaya and Lahiri, 2010)

Agricultural Statistics Board (ASB)

- Panel of NASS statisticians and commodity experts charged with examining multiple sources of information and setting official estimates of various agricultural items at different levels (national, state, district, county)
- CRS is main basis for estimation of cash rental rates but ASB considers other indications as well

Ratio vs. Difference Benchmarking

- Ratio Benchmarking
  - model-based estimates adjusted via multiplication by state-level and/or district-level ratios between weighted sums of rented acreage values
  - benchmarked estimates guaranteed to have same sign as corresponding unadjusted estimates
- Difference Benchmarking (Ghosh-Steorts Method)
  - model-based estimates adjusted via addition of an estimated parameter involving CRS-based weights (or alternatively, ASB-based weights)
  - change in estimate due to benchmarking same for all counties in a district
  - benchmarked county-level estimates could be negative even if corresponding model-based estimates are positive

Research Questions

- How does the Ghosh-Steorts difference benchmarking method compare with ratio-based procedures with respect to different efficiency criteria?
- Does the choice of weights (CRS or ASB-based) matter?

<u>Note</u> – Benchmarking is not a 'cure' for a poor model fit

Notation

$\tilde{\theta}_i$ = unadjusted model-based estimate of cash rental rate in county *i*

D_j = agricultural statistics district *j*

$\tilde{\lambda}_{j,w}$ = unadjusted estimate of rental rate in district *j* (weighted sum)

$\hat{\lambda}_j^{(DB)}$ = adjusted (difference benchmarked) estimate of rental rate in district *j*

z_i = estimate of acres rented in county *i* *(*from CRS or ASB data)

a_s = target state-level value of rental rate

Difference Benchmarking for Cash Rents

Objective – minimize conditional expected loss -

$$L = \sum_{j \in s} \sum_{i \in D_j} w_i (\hat{\theta}_i^{(DB)} - \theta_i)^2 + \sum_{j \in s} \eta_j (\hat{\lambda}_j^{(DB)} - \lambda_j)^2$$

subject to:

$$\sum_{i \in D_j} w_i \hat{\theta}_i^{(DB)} = \hat{\lambda}_j^{(DB)}, \quad \sum_{j \in s} \eta_j \hat{\lambda}_j^{(DB)} = a_s$$

where:

$w_i = z_i / \sum_{k \in D_{j(i)}} z_k$ (first stage weights)

$D_{j(i)}$ = district containing county *i*

$\eta_j = (\sum_{k \in D_j} z_k) / (\sum_{k \in s} \sum_{i \in D_k} z_i)$ (second stage weights)

Difference Benchmarked Estimates

District Level -

$$\hat{\lambda}_j^{(DB)} = \tilde{\lambda}_{j,w} + (a_s - \hat{\varphi}_w)\eta_j(1+\eta_j)^{-1} / \sum_{k\in s}\eta_k{}^2(1+\eta_k)^{-1}$$

where:

$$\hat{\varphi}_w = \sum_{j\in s}\eta_j\,\tilde{\lambda}_{j,w}$$

County Level -

$$\hat{\theta}_i^{(DB)} = \tilde{\theta}_i + \hat{\lambda}_{j(i)}^{(DB)} - \tilde{\lambda}_{j(i),w}$$

Two Versions of Ratio Benchmarking

- *RB1* (traditional approach)
  – common adjustment factor applied to all model-based estimates in state (district level bypassed)

$$\hat{\theta}_i^{(RB1)} = (a_s \,/\, \sum_{k \in s} w_k \, \tilde{\theta}_k) \, \tilde{\theta}_i$$

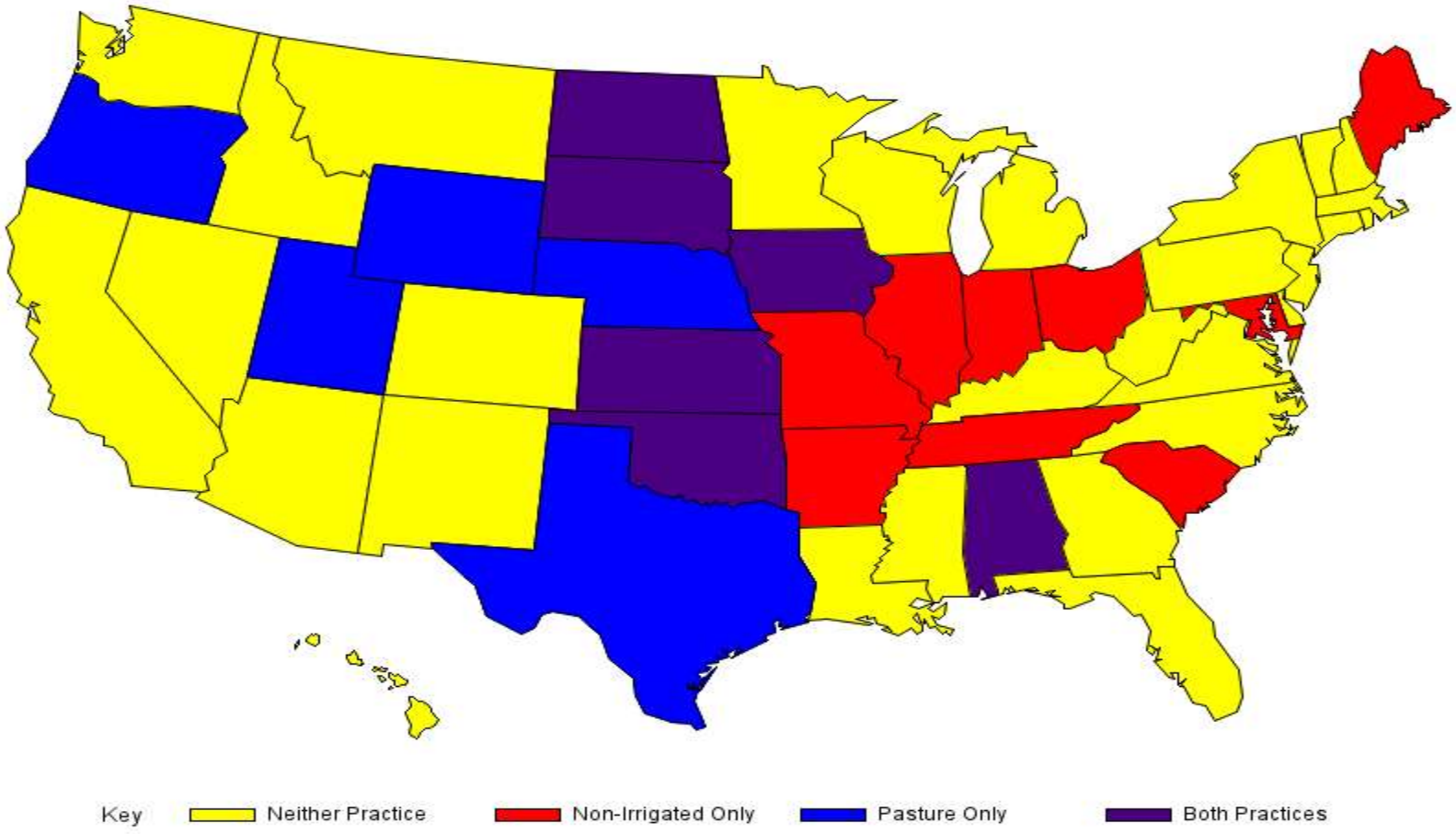
- *RB2* (alternative approach)
  - separate state-to-district and district-to-county adjustment factors

$$\hat{\theta}_i^{(RB2)} = [(a_s \,/\, \sum_{k \in s} \eta_k \, \tilde{\lambda}_k)(\, \tilde{\lambda}_{j(i)} \,/\, \sum_{k \in D_{j(i)}} w_k \, \tilde{\theta}_k)] \, \tilde{\theta}_i$$

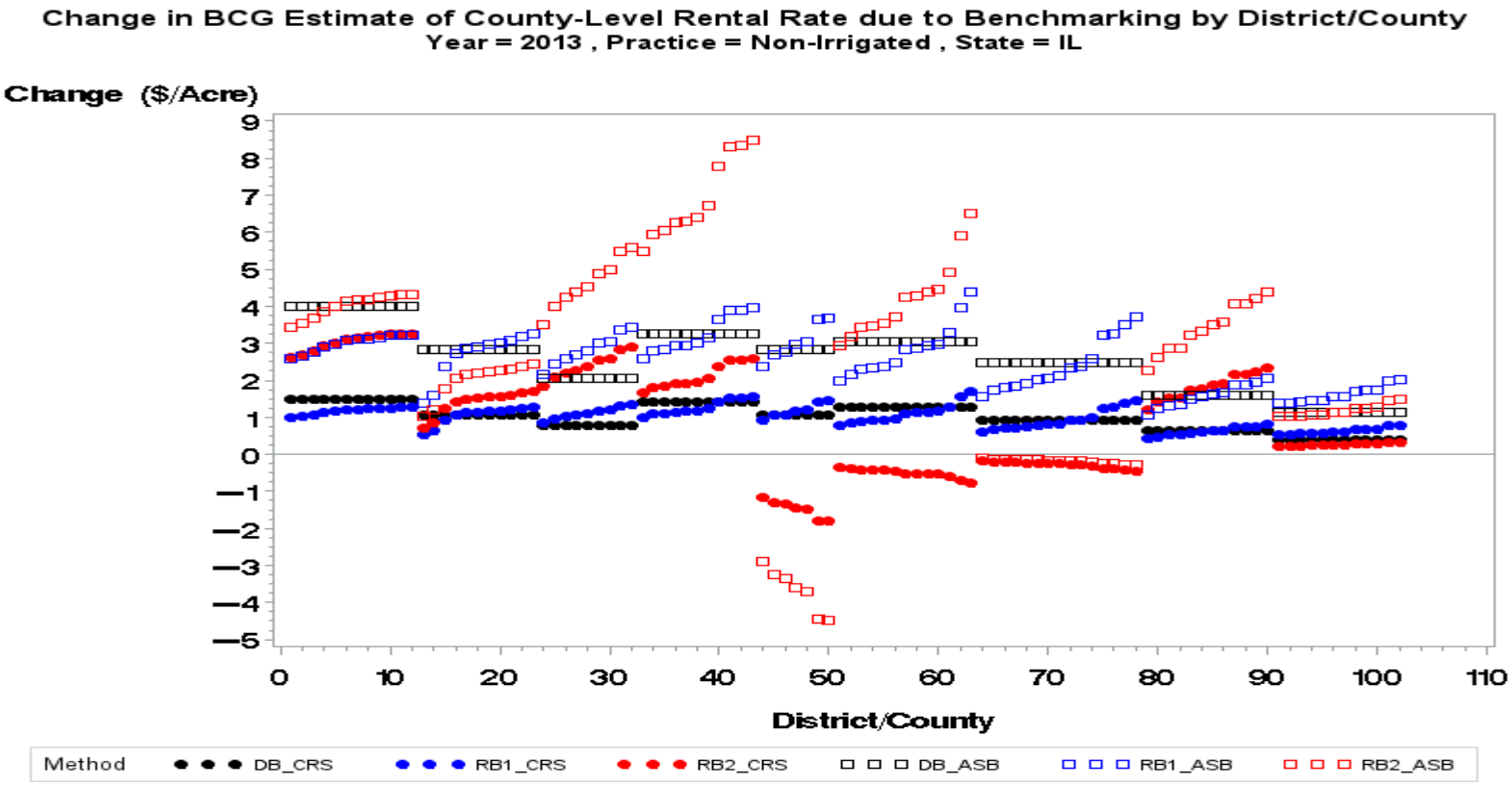
Empirical Study Using Data from 2013-14

- Only states with adequate CRS data, ASB data and model-based rental rate estimates for all counties included (handling of 'incomplete' states is an issue for future research)
  - irrigated practice not considered due to only two states satisfying inclusion criteria
- Six benchmarking options
  - Three methods (*DB*, *RB1*, *RB2*) evaluated using both CRS and ASB-based weights (so referred to as *DB_CRS*, *DB_ASB*, *RB1_CRS*, *RB1_ASB*, *RB2_CRS* and RB2_ASB) for non-irrigated (*NIR*) and pasture (*PAS*)
- Criteria for comparison
  - change in estimate due to benchmarking vs. 1) county, 2) CRS sample size
  - estimation accuracy metrics (using official county-level rental rates as "gold standard")

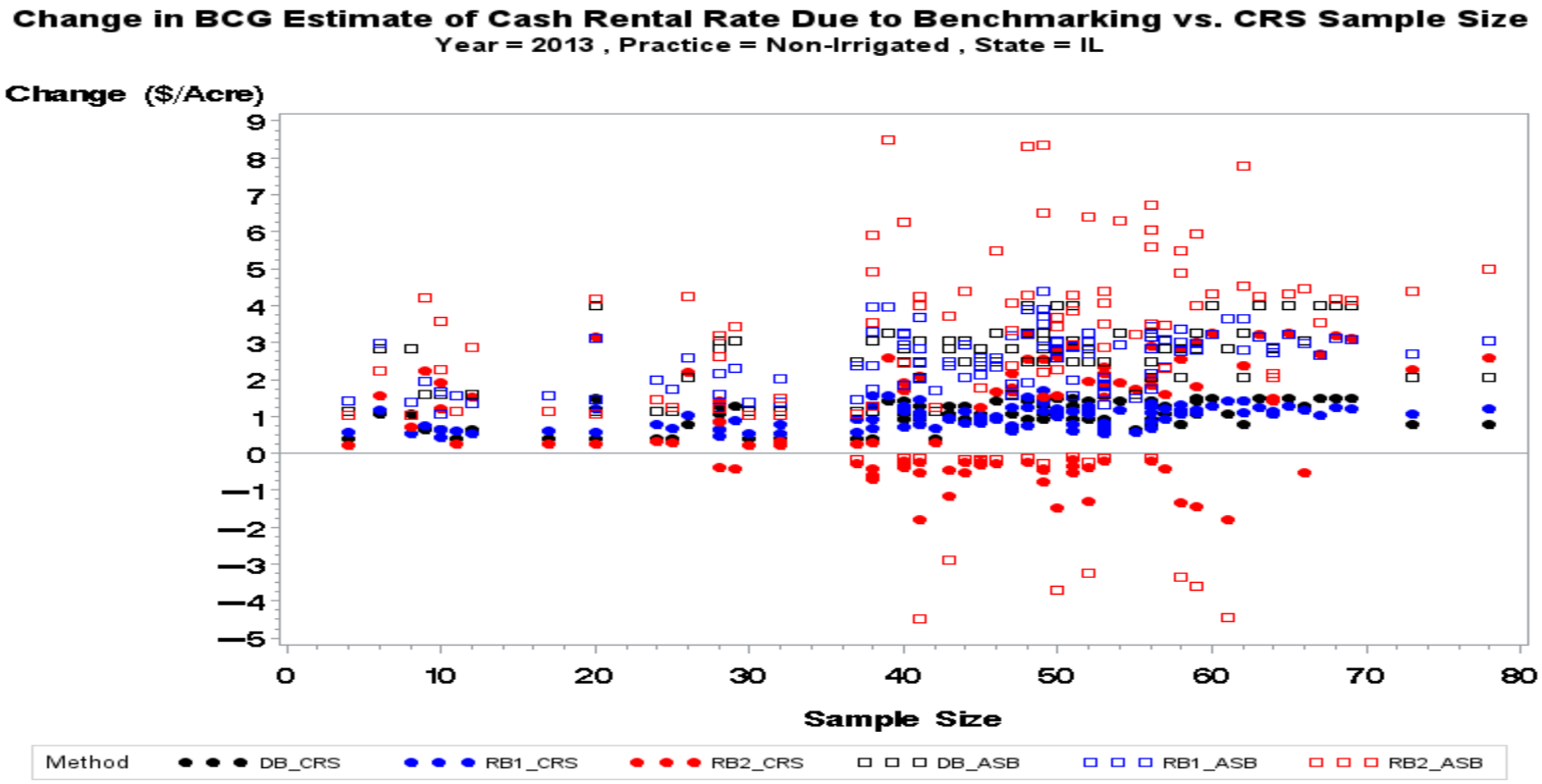
States Included in Study Areas by Practice



Illustrative Example - Non-Irrigated for Illinois (102 Counties) in 2013



Illustrative Example (cont.)



Metrics for Comparing Benchmarking Methods (Computed at State Level)

Mean Absolute Deviation (MAD) – mean of absolute values of county-level deviations (of benchmarked cash rental rate estimates) from official estimates

Mean Absolute Relative Deviation (MARD) – mean of ratios between absolute deviations and official estimates

Root Mean-Squared Deviation (RMSD) – square root of mean of squared deviations from official estimates

Comparison Based on Rankings of Benchmarking Options by Metric

- For each year/practice combination, three metrics (MAD, MARD and RMSD) computed by state for all six benchmarking options
- Study area included 14 states for *2013/NIR*, 12 for *2014/NIR*, 9 for *2013/PAS*, 10 for *2014/PAS*
- Options ranked from 1 (best) to 6 (worst) for each year/practice/state/metric combination (so 78 rankings for *NIR*, 57 for *PAS*)
- Ranks averaged over states by metric for each practice

Example – MAD for 2013/Pasture

(Ranks in Parentheses)

State	Option						
	No. Counties	DB_CRS	DB_ASB	RB1_CRS	RB1_ASB	RB2_CRS	RB2_ASB
Alabama	67	1.55 (4)	1.56 (5)	1.53 (3)	1.51 (2)	1.56 (6)	1.41 (1)
Kansas	105	2.46 (3)	2.75 (6)	2.31 (1)	2.61 (4)	2.41 (2)	2.64 (5)
Nebraska	93	6.57 (2)	6.14 (1)	12.74 (6)	12.26 (3)	12.66 (5)	12.27 (4)
(6 others)	…	…	…	…	…	…	…
Average Rank		1.89	2.44	4.33	3.89	4.56	3.89

Average Ranks for Non-Irrigated

Year	No. States	Metric	Average Rank					
			DB-CRS	DB-ASB	RB1-CRS	RB1-ASB	RB2-CRS	RB2-ASB
2013	14	*MAD*	3.5	2.93	3.79	2.86*	4.14	3.79
		MARD	3.14*	3.21	3.5	3.29	3.64	4.21
		RMSD	4.0	2.79	4.36	2.57*	4.0	3.29
2014	12	*MAD*	4.0	3.75	3.33	3.0*	3.42	3.5
		MARD	4.0	3.75	2.58*	2.83	3.58	4.25
		RMSD	3.92	3.75	3.75	3.58	3.17	2.83*

* - lowest value for year/metric

Average Ranks for Pasture

Year	No. States	Metric	Average Rank					
			DB-CRS	DB-ASB	RB1-CRS	RB1-ASB	RB2-CRS	RB2-ASB
2013	9	*MAD*	1.89*	2.44	4.33	3.89	4.56	3.89
		MARD	2.33*	2.78	4.11	3.89	4.22	3.67
		RMSD	2.11*	2.44	4.33	4.0	4.33	3.78
2014	10	*MAD*	3.0	2.8*	3.8	4.1	3.7	3.6
		MARD	3.2	2.9*	3.5	3.4	4.6	3.5
		RMSD	3.3*	3.5	3.5	3.7	3.7	3.5

* - lowest value for year/metric

Concluding Remarks

- Exploratory analysis based on 2013-14 cash rents data suggests that relative effectiveness of benchmarking methods varies with land use (*RB1* appeared to be best for non-irrigated, *DB* for pasture)
- Adjustments can have strong correlation with CRS sample size (undesirable property); *DB* least likely to cause excessive changes
- Choice of weights (CRS or ASB) did not appear to be a major factor
- **Future Research**
  - further evaluation of accuracy properties (e.g., via simulation)
  - variance properties
  - handling of states having counties with incomplete data
  - additional benchmarking methods

References

1. Bellow, M., Cruze, N. and Erciulescu, A. (2017). "Developments in Model-Based County-Level Estimation of Agricultural Cash Rental Rates," In *Proceedings of the Section on Survey Research Methods*, American Statistical Association.
2. Berg, E., Cecere, W. and Ghosh, M. (2014). "Small Area Estimation for County-Level Farmland Cash Rental Rates," Journal of Survey Statistics and Methodology, 2, 1-37.
3. Cruze, N. (2017). "Assessing the Suitability of the Berg-Cecere-Ghosh Model for Official County Estimates of Cash Rental Rates", In *Panel on Methods for Integrating Multiple Data Sources to Improve Crop Estimates,* Presentation Book, Fourth Meeting, National Academies of Sciences, Engineering and Medicine.
4. Gershunskaya, J. and Lahiri, P. (2010). "Robust Small Area Estimation Using a Mixture Model," In *Proceedings of the Section on Survey Research Methods,* American Statistical Association.
5. Ghosh, M. and Steorts, R. (2013). "Two-Stage Benchmarking as Applied to Small Area Estimation"*, TEST,* 22*,* 670-87.
6. Pfefferman, D. (2013). "New Important Developments in Small Area Estimation," *Statistical Science*, 28, 1, 40-68.

Contact Information – Mike.Bellow@nass.usda.gov