Catalogue no. 12-001-X
ISSN 1492-0921

Survey Methodology

A bivariate hierarchical Bayesian model for estimating cropland cash rental rates at the county level

by Andreea Erciulescu, Emily Berg, Will Cecere and Malay Ghosh

Release date: June 27, 2019

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

Email at STATCAN.infostats-infostats.STATCAN@canada.ca

Telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following numbers:

- Statistical Information Service 1-800-263-1136
- National telecommunications device for the hearing impaired 1-800-363-7629
- Fax line 1-514-283-9350

Depository Services Program

- Inquiries line 1-800-635-7943
- Fax line 1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "Contact us" > "Standards of service to the public."

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Published by authority of the Minister responsible for Statistics Canada

An HTML version **is also available.**

Cette publication est aussi disponible en français.

Survey Methodology, June 2019
Vol. 45, No. 2, pp. 199-216
Statistics Canada, Catalogue No. 12-001-X

A bivariate hierarchical Bayesian model for estimating cropland cash rental rates at the county level

Andreea Erciulescu, Emily Berg, Will Cecere and Malay Ghosh[1]

Abstract

The National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) is responsible for estimating average cash rental rates at the county level. A cash rental rate refers to the market value of land rented on a per acre basis for cash only. Estimates of cash rental rates are useful to farmers, economists, and policy makers. NASS collects data on cash rental rates using a Cash Rent Survey. Because realized sample sizes at the county level are often too small to support reliable direct estimators, predictors based on mixed models are investigated. We specify a bivariate model to obtain predictors of 2010 cash rental rates for non-irrigated cropland using data from the 2009 Cash Rent Survey and auxiliary variables from external sources such as the 2007 Census of Agriculture. We use Bayesian methods for inference and present results for Iowa, Kansas, and Texas. Incorporating the 2009 survey data through a bivariate model leads to predictors with smaller mean squared errors than predictors based on a univariate model.

Key Words: Hierarchical Bayes; Bivariate mixed model; Benchmarking.

1 Introduction

The National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) conducts hundreds of surveys each year to obtain estimates related to diverse aspects of US agriculture. Examples of parameters that NASS estimates include total production, harvested area, and crop yield. Estimation for sub-state domains, such as counties, is difficult due to small sample sizes. Our interest is in estimation of the county-level cash rental rate, the market value of land rented on a per acre basis for cash only.

Estimates of county-level cash rental rates serve many purposes. Farmers use the estimates for guidance in determining rental agreements (Dhuyvetter and Kastens, 2009). Agronomists use the estimates to study research questions related to the interplay between cash rental rates and other economic characteristics such as commodity prices and fuel costs (Woodard, Paulson, Baylis and Woddard, 2010). NASS's published estimates of mean cash rental rates at the county level have implications for the Conservation Reserve Program, a policy that encourages agricultural landowners to conserve their land. The 2008 and 2014 Farm Bills require NASS to collect data on cash rental rates for three land use categories – non-irrigated cropland, irrigated cropland, and permanent pasture – for counties with at least 20,000 acres of cropland or pastureland.

To satisfy the requirements of the 2008 and 2014 Farm Bills, NASS conducts a Cash Rent Survey. A concern is that direct estimators of county means from the Cash Rent Surveys may be unstable due to small

1. Andreea Erciulescu, National Institute of Statistical Sciences, Washington D.C., U.S.A.; Emily Berg, Department of Statistics, Iowa State University, Ames, IA, U.S.A. E-mail: emilyb@iastate.edu; Will Cecere, Westat, Rockville, MD, U.S.A.; Malay Ghosh, Department of Statistics, University of Florida, Gainesville, FL, U.S.A.

realized sample sizes. We investigate the use of mixed models (Rao and Molina, 2015) to stabilize the estimators of average cash rental rates at the county level. NASS publishes estimates of average cash rental rates at the state level before county level estimation from the Cash Rent Survey is complete. To maintain internal consistency, the county predictors must satisfy a benchmarking restriction.

In a frequentist framework, Berg, Cecere and Ghosh (2014) use area-level models to predict county-level cash rental rates for all states and for the three land use categories of non-irrigated cropland, irrigated cropland, and permanent pasture. For each combination of land use category and state, the method of Berg et al. (2014) uses data from two years. An assumption that the variances for the two years are the same motivates the Pitman-Morgan transformation, which converts the vector of observations for the two time points into an average and a difference. After separate univariate models are applied to the average and the difference, the predictor for each time point is obtained by adding the predictor of the average to half of the predictor of the difference. The method of Berg et al. (2014) is demonstrated to provide a practical approach to obtaining reasonable predictions across a diverse range of conditions. Nonetheless, the effects of simplifying assumptions warrant additional investigation. If the variances for the two time-points differ, then, as discussed in Berg et al. (2014), the mean squared error (MSE) estimator based on the Pitman-Morgan transformtion can have a negative bias. Further, the Berg et al. (2014) method does not account for the effect of benchmarking when estimating the MSE.

This study addresses the issues of non-constant variances across time and the effect of benchmarking on efficiency in the context of the NASS Cash Rent Surveys through the use of a bivariate hierarchical Bayesian (HB) model for the unit-level data. The model is sufficiently flexible to allow the variances to differ between the two time-points. The use of Bayesian methods for inference facilitates estimation of the increase in posterior MSE due to benchmarking. Another innovation of the bivariate HB approach is that it incorporates the survey weights in the variance model. We also aim to improve the efficiency of the predictors for particular situations, relative to Berg et al. (2014), by allowing the covariates to differ across states. Datta, Day and Maiti (1998) examine HB bivariate models for the county crop acreage data of Battese, Harter and Fuller (1988). Our model extends the Datta et al. (1998) model to account for a relationship between the weight and the variance as well as an unbalanced data structure.

We focus on prediction of county level cash rental rates for non-irrigated cropland using the responses to the 2009 and 2010 Cash Rent Surveys as well as external sources of auxiliary information. In Section 2, we discuss the survey data and the auxiliary information in detail. We describe the bivariate HB model in Section 3. In Section 4, we summarize results for non-irrigated cropland in Iowa, Kansas, and Texas. In Section 5, we summarize and discuss possible future research applicable to both estimation of cropland cash rental rates and small area estimation more generally.

2 Data for modeling non-irrigated cropland cash rental rates

2.1 NASS Cash Rent Survey

NASS implemented a Cash Rent Survey in response to the 2008 Farm Bill. The specific objective of the Cash Rent Survey is to obtain county level estimates of average cash rental rates in three land use categories: non-irrigated cropland, irrigated cropland, and permanent pasture. The data for our study are from the 2009 and 2010 Cash Rent Surveys.

2.1.1 NASS Cash Rent Survey sample design

The 2009 and 2010 Cash Rent Surveys used a stratified sample design. To define the stratification, nine groups were formed on the basis of the dollars rented that an operation reported on previous surveys and censuses. The strata are the intersections of the nine groups and agricultural statistics districts. An agricultural statistics district is a group of contiguous counties within a state that are thought to have similar agricultural characteristics. The sampling fractions within strata are defined so that operations with higher dollars rented on previous surveys and censuses have greater probabilities of selection. The same sample was used for the 2009 and 2010 Cash Rent Surveys, which had a national sample size of approximately 224,000 operations. A unit may respond in only one year either because of nonresponse or because the operation only participated in a rental agreement in one of the two years.

2.1.2 Relationships between 2009 and 2010 non-irrigated cropland cash rents

A direct survey estimator for a particular land use category is a ratio of a weighted sum of the dollars rented to a weighted sum of acres rented. The weight associated with a respondent is the population size of the stratum containing the respondent divided by the number of responding units in that stratum. Berg et al. (2014) explore relationships between direct estimates for two years. For the states considered in Berg et al. (2014), the correlations between the direct estimates for the two years range from 0.20 to 0.99, where the correlation is across counties for a particular state. Because our emphasis is on unit level models, we focus on relationships over time at the unit level.

To measure the correlation between the reported 2009 and 2010 cash rental rates at the unit (farm operator) level, we compute differences between unit-level cash rental rates for non-irrigated cropland and the sample mean for a county. Only individuals that report a cash rental rate for non-irrigated cropland in both years are used to compute the differences. The difference for year t is $y_{ijt} - \overline{y}_{i.t}$, where y_{ijt} is the cash rent per acre for non-irrigated cropland reported by operator j in county i and year t, and $\overline{y}_{i.t}$ is the sample average of the y_{ijt} in county i that reported a non-irrigated cropland cash rental rate in both 2009 and 2010. The deviations between individual cash rental rates and the county means for Kansas are plotted in Figure 2.1. The deviations for 2009 and 2010 for Kansas are linearly related, and the correlation between the deviations for 2009 and the deviations for 2010 is 0.7. The extreme values in Figure 2.1 reflect the high variability among the non-irrigated cropland cash rental rates within a county in Kansas.

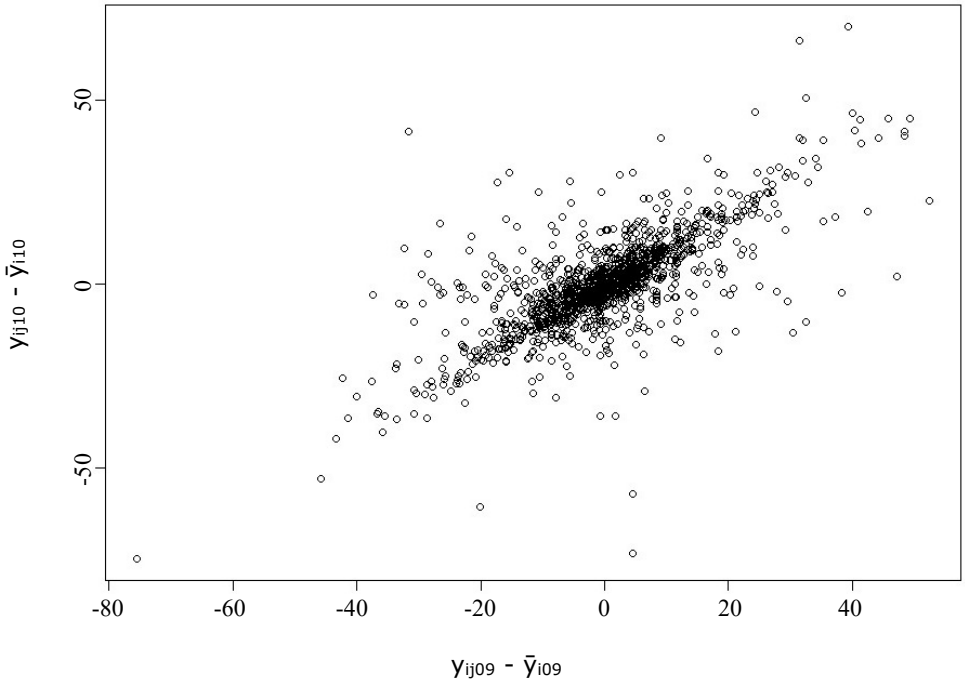

Figure 2.1 Deviations of unit-level cash rental rates from county means for 2009 (x-axis) and 2010 (y-axis) for units reporting non-irrigated cash rental rates in both years.

2.2 Auxiliary information

In an effort to improve the precision of the estimators of average cash rental rates at the county level, auxiliary variables were desired that would explain both the variability among the county means as well as the variability among units within a county. Auxiliary information for modeling cash rental rates is available from several sources external to the Cash Rent Survey. The potential covariates divide into three broad categories, depending on whether the covariate relates principally to land quality, the commodity value sold, or other farm characteristics. The list below summarizes the three categories of covariates, indicates whether each covariate is recorded at the county level or the unit level, and specifies if the covariate is only available for a particular state. Unit-level covariates are only available for units in the Cash Rent Survey sample, while area level covariates are treated as population means.

1. Land quality
    - Four National Commodity Crop Productivity Indexes (NCCPIs) are county-level covariates available for all states. Three climate-specific indexes called NCCPI-corn, NCCPI-wheat, and NCCPI-cotton reflect the quality of the soil for growing non-irrigated crops in three different climate conditions (Dobos, Sinclair and Robotham, 2012). The fourth index, Max-NCCPI, is the maximum of the three climate-specific indexes. The indexes are originally constructed at the level of a "mapunit," an area that has relatively homogeneous soil

properties. The county-level covariates are averages of the indexes across all mapunits in a county.

- An average corn yield across years 2005-2009 is available at the county level for Iowa only. All counties in Iowa have a corn yield estimate available for at least one of the years between 2005 and 2009, and years for which a yield estimate is missing for a county are excluded from the average for that county.
- Because Kansas is more agriculturally diverse than Iowa, no single crop yield is published in at least one year between 2005 and 2009 for all counties of interest. To obtain a covariate that is measured for all counties, we constructed a non-irrigated yield index for Kansas. We first averaged NASS published yields for corn, wheat, and sorghum using the method described for the Iowa corn yields. The average yields were then standardized to have mean zero and variance one. The non-irrigated yield index for a county is defined as the largest of the three standardized yields. (For Texas, availability of crop yield information was too sparse to use to define a covariate).

2. Value of the commodity sold
   - Total value of production for a county based on the 2007 Census of Agriculture is available for all states.
   - Expected sales for an operation (unit) recorded on the NASS list frame are available for all states at the unit-level.

3. Other farm characteristics
   - Farm type is a unit level categorical covariate, available for all states. Farms are partitioned into 17 farm types on the NASS list frame. To define a covariate, the farm types are aggregated into two groups: (1) grains/oilseeds, and (2) other.
   - Acres rented for non-irrigated cropland recorded on the NASS Cash Rent Survey are available at the unit level for all states.

3 Bivariate hierarchical Bayesian model

The correlation between the 2009 and 2010 cash rental rates observed in Section 2.1.1 suggests that using the information in the data from 2009 has the potential to improve the predictions for 2010. A bivariate hierarchical model for a state is specified as a way to incorporate the data for both years. Let $a_{ij,t}$ and $y_{ij,t}$ be the acres and dollars per acre, respectively, rented by operator j in county i and year $t\,(t = 09, 10)$, and let $\mathbf{x}_{ij,t}$ be the associated column vector of auxiliary variables with dimension p_t. For covariates that are constant across years and individuals, $\mathbf{x}_{ij,t} = \mathbf{x}_{i109}$. Let $w_{ij,t} = a_{ij,t} N_{g(ijt)} n_{g(ijt)}^{-1}$, where $N_{g(ijt)}$ and $n_{g(ijt)}$

are the population size and number of respondents, respectively, in year t for the stratum g that contains unit (ij).

To specify the model, we divide the respondents into three sets:

- Set 1 consists of units (ij) that report a non-irrigated cash rental rate in both 2009 and 2010.
- Set 2 consists of units (ij) that only report a non-irrigated cash rental rate in 2009.
- Set 3 consists of units (ij) that only report a non-irrigated cash rental rate in 2010.

We assume that observations in set 1 satisfy the bivariate model

$$\begin{pmatrix} y_{ij,09} \\ y_{ij,10} \end{pmatrix} = \begin{pmatrix} \mathbf{x}'_{ij,09}\boldsymbol{\beta}_{09} + \nu_{i,09} + e_{ij,09} \\ \mathbf{x}'_{ij,10}\boldsymbol{\beta}_{10} + \nu_{i,10} + e_{ij,10} \end{pmatrix}, \tag{3.1}$$

where

$$\begin{pmatrix} e_{ij,09} \\ e_{ij,10} \end{pmatrix} \sim \mathrm{N}\left(\mathbf{0},\ \mathbf{D}_{wij}^{-0.5}\boldsymbol{\Sigma}_{ee}\mathbf{D}_{wij}^{-0.5}\right), \tag{3.2}$$

$\mathbf{D}_{wij} = \mathrm{diag}\left(w_{ij,09},\ w_{ij,10}\right)$, and

$$\begin{pmatrix} \nu_{i,09} \\ \nu_{i10} \end{pmatrix} \sim \mathrm{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\nu\nu}\right). \tag{3.3}$$

We denote the diagonal elements of $\boldsymbol{\Sigma}_{ee}$ corresponding to 2009 and 2010 by σ_{ee09} and σ_{ee10}, respectively. For units (ij) in set 2 or 3, we assume

$$y_{ij,t} = \mathbf{x}'_{ij,t}\boldsymbol{\beta}_t + \nu_{i,t} + e^*_{ij,t}, \tag{3.4}$$

where $e^*_{ij,t} \sim \mathrm{N}\left(0,\ w_{ij,t}^{-1}\,\tau^2_{e,t}\right)$, $t = 09$ for set 2, and $t = 10$ for set 3. The model not only allows the variances for the unit-level errors to differ across time points but also allows the variances of unit-level errors for units that respond in both time points to differ from the variances for units that only respond in one time-point. The quantity to predict for 2010 is

$$\theta_{i,10} = \bar{\mathbf{x}}'_{N_i,10}\boldsymbol{\beta}_{10} + \nu_{i,10}, \tag{3.5}$$

where $\bar{\mathbf{x}}_{N_i,10}$ is the population mean of the covariates for county i.

The variances of the unit-level errors, $e_{ij,t}$ and $e^*_{ij,t}$, are assumed to be inversely proportional to the weight, $w_{ij,t}$, for two reasons. First, incorporating the weights in the model aims to reduce bias that could arise if the design is informative for the model. As explained in Section 2, the weights depend on the dollar value of the land rented from the previous year. Therefore, the possibility that the sample design may be informative for a model without the weights is plausible. If $\boldsymbol{\Sigma}_{ee}$ and $\boldsymbol{\Sigma}_{\nu\nu}$ are diagonal, and if $\tau^2_{e,t} = \sigma_{eet}$, then in a frequentist framework, an empirical best linear unbiased predictor for the county i mean in year

t is the design-consistent pseudo-eblup of You and Rao (2002). The second reason to incorporate the weights is that the variances of residuals from preliminary analyses decrease as the acres increase.

Diffuse, proper priors are specified for the unknown regression coefficients and variances. Specifically, $\boldsymbol{\beta}_t \sim \mathrm{N}(\mathbf{0}, 10^6\mathbf{I})$, and $\tau_{e,t}^2 \sim \text{inverse} - \text{gamma}(0.001, 0.001)$. The covariance matrices, $\boldsymbol{\Sigma}_{ee}$ and $\boldsymbol{\Sigma}_{vv}$ have inverse-Wishart prior distributions with shape parameter 0.01 and a diagonal scale matrix with diagonal elements 0.001. The parameterizations for the inverse-gamma and inverse-Wishart distributions are from Gelman, Carlin, Stern and Rubin (2009). We choose priors with conjugate forms for computational simplicity. The choices of the hyperparameters are selected to be un-informative relative to the data for the Cash Rents Survey application.

3.1 Gibbs sampling and posteriors

We use Gibbs sampling to obtain a Monte Carlo approximation to the posterior distribution. An analysis of BGR statistics (Gelman et al., 2009) based on three MCMC chains, each with 20,000 iterations, indicated that 1,000 iterations is sufficient for burn-in. The analyses in Section 4 are based on one chain of length 20,000 for each of the three states, Iowa, Kansas and Texas, where the first 1,000 iterations are discarded for burn-in. By the choices of the likelihood and the priors, the full conditional distributions are known distributions. See Appendix A.

3.2 Prediction and MSE estimation

If $\overline{\mathbf{x}}_{N_i,10}$ is known, the Bayes predictor of $\theta_{i,10}$ for squared error loss is

$$\tilde{\theta}_{i,10}^{B} = E\left[\theta_{i,10} \,\middle|\, (\mathbf{y}, \mathbf{x}), \overline{\mathbf{x}}_{N_i,10}\right] = \overline{\mathbf{x}}'_{N_i,10}\hat{\boldsymbol{\beta}}_{10} + E\left[v_{i,10} \,\middle|\, (\mathbf{y}, \mathbf{x})\right], \tag{3.6}$$

where $\hat{\boldsymbol{\beta}}_{10} = E[\boldsymbol{\beta}_{10} \,|\, (\mathbf{y}, \mathbf{x})], (\mathbf{y}, \mathbf{x})$ denotes the observed cash rental rates and covariates for the two years, and the second equality in (3.6) follows from (3.5) and linearity of expectation. The posterior mean squared error of $\tilde{\theta}_{i,10}^{B}$ is

$$E\left[\left(\tilde{\theta}_{i,10}^{B} - \theta_{i,10}\right)^2 \middle| (\mathbf{y}, \mathbf{x}), \overline{\mathbf{x}}_{N_i,10}\right] = V\left\{\theta_{i,10} \,\middle|\, (\mathbf{y}, \mathbf{x}), \overline{\mathbf{x}}_{N_i,10}\right\}. \tag{3.7}$$

As discussed in Section 2, the population mean of the covariates, $\overline{\mathbf{x}}_{N_i,10}$, is not available for unit-level covariates in the Cash Rent Survey application. To define a predictor, we add a model for the covariate mean. See Lohr and Prasad (2003) for an approach that begins with a model specification for the unit level covariates. Partition $\mathbf{x}_{ij,10}$ into two sub-vectors, $\mathbf{x}_{ij,10}^{(1)}$ and $\mathbf{x}_{ij,10}^{(2)}$, where $\mathbf{x}_{ij,10}^{(1)}$ contains county-level covariates, and $\mathbf{x}_{ij,10}^{(2)}$ contains unit-level covariates. Assume $\overline{\mathbf{x}}_{wi10} \,|\, \overline{\mathbf{x}}_{N_i,10} \sim \mathrm{N}\left(\overline{\mathbf{x}}_{N_i,10}, \mathbf{V}_{xxi,10}\right)$, where $\overline{\mathbf{x}}_{wi10} = \left(\sum_{j=1}^{n_{i10}} w_{ij,10}\right)^{-1}\left(\sum_{j=1}^{n_{i10}} w_{ij,10}\mathbf{x}_{ij,10}\right)$, n_{i10} is the sum of the number of units in set 1 and in set 3, and $\mathbf{V}_{xxi,10}$ is known. The elements of $\mathbf{V}_{xxi,10}$ corresponding to $\mathbf{x}_{ij,10}^{(1)}$ are 0, and we explain how we obtain the elements of $\mathbf{V}_{xxi,10}$ corresponding to unit-level covariates in Appendix B. The Central Limit Theorem supports the assumption of normality for $\overline{\mathbf{x}}_{wi10}$ even if the distribution of the unit-level covariate values is not normal

(Kim, Park and Lee, 2017). Assuming $\overline{\mathbf{x}}_{N_i,10}$ has a flat prior, $\overline{\mathbf{x}}_{N_i,10} \left| \overline{\mathbf{x}}_{wi10} \right. \sim \mathrm{N}\left(\overline{\mathbf{x}}_{wi10}, \mathbf{V}_{xxi,10}\right)$. The Bayes predictor of $\theta_{i,10}$ for squared error loss under the extended model in which the population mean of the covariates is unknown is

$$\hat{\theta}_{i,10}^{B} = \overline{\mathbf{x}}'_{wi10}\hat{\boldsymbol{\beta}}_{10} + E\left[v_{i,10} \left| (\mathbf{y}, \mathbf{x}) \right.\right]. \tag{3.8}$$

The posterior mean squared error of $\hat{\theta}_{i,10}^{B}$ is

$$\begin{aligned}
E\left[\left(\hat{\theta}_{i,10}^{B} - \theta_{i,10}\right)^2 \left| (\mathbf{y}, \mathbf{x}) \right.\right] &= E\left\{\left(\hat{\theta}_{i,10}^{B} - \tilde{\theta}_{i,10}^{B} + \tilde{\theta}_{i,10}^{B} - \theta_{i,10}\right)^2 \left| (\mathbf{y}, \mathbf{x}) \right.\right\} \\
&= E\left\{\left(\hat{\theta}_{i,10}^{B} - \tilde{\theta}_{i,10}^{B}\right)^2 \left| (\mathbf{y}, \mathbf{x}) \right.\right\} \\
&\quad + 2E\left\{E\left[\left(\hat{\theta}_{i,10}^{B} - \tilde{\theta}_{i,10}^{B}\right)\left(\tilde{\theta}_{i,10}^{B} - \theta_{i,10}\right) \left| (\mathbf{y}, \mathbf{x}), \overline{\mathbf{x}}_{N_i,10} \right.\right] \left| (\mathbf{y}, \mathbf{x}) \right.\right\} \\
&\quad + V\left\{\theta_{i,10} \left| (\mathbf{y}, \mathbf{x}) \right.\right\} \\
&= \hat{\boldsymbol{\beta}}'_{10} V\left\{\overline{\mathbf{x}}_{N_i,10} \left| \overline{\mathbf{x}}_{wi10} \right.\right\} \hat{\boldsymbol{\beta}}_{10} + V\left\{\theta_{i,10} \left| (\mathbf{y}, \mathbf{x}) \right.\right\} \\
&= \hat{\boldsymbol{\beta}}'_{10} V\left\{\overline{\mathbf{x}}_{N_i,10} \left| \overline{\mathbf{x}}_{wi10} \right.\right\} \hat{\boldsymbol{\beta}}_{10} \\
&\quad + V\left\{\overline{\mathbf{x}}'_{wi10}\boldsymbol{\beta}_{10} + v_{i,10} + \left(\overline{\mathbf{x}}_{N_i,10} - \overline{\mathbf{x}}_{wi10}\right)' \boldsymbol{\beta}_{10} \left| (\mathbf{y}, \mathbf{x}) \right.\right\} \\
&\approx \hat{\boldsymbol{\beta}}'_{10} V\left\{\overline{\mathbf{x}}_{N_i,10} \left| \overline{\mathbf{x}}_{wi10} \right.\right\} \hat{\boldsymbol{\beta}}_{10} + V\left\{\overline{\mathbf{x}}'_{wi10}\boldsymbol{\beta}_{10} + v_{i,10} \left| (\mathbf{y}, \mathbf{x}) \right.\right\},
\end{aligned} \tag{3.9}$$

where the final approximation assumes that the $\mathbf{Cov}\left\{\overline{\mathbf{x}}'_{wi10}\boldsymbol{\beta}_{10} + v_{i,10}, \left(\overline{\mathbf{x}}_{N_i,10} - \overline{\mathbf{x}}_{wi10}\right)' \boldsymbol{\beta}_{10} \left| (\mathbf{y}, \mathbf{x}) \right.\right\}$ is negligible. A comparison of (3.7) and (3.9) shows that the term $\hat{\boldsymbol{\beta}}'_{10} V\left\{\overline{\mathbf{x}}_{N_i,10} \left| \overline{\mathbf{x}}_{wi10} \right.\right\} \hat{\boldsymbol{\beta}}_{10}$ accounts for the increase in posterior MSE due to replacing $\overline{\mathbf{x}}_{N_i,10}$ in (3.6) with $\overline{\mathbf{x}}_{wi10}$ in (3.8). To quantify the posterior MSE of $\hat{\theta}_{i,10}^{B}$, we use

$$\mathrm{MSE}\left(\hat{\theta}_{i,10}^{B}\right) = \widehat{\mathrm{MSE}}_{1i} + \widehat{\mathrm{MSE}}_{2i}, \tag{3.10}$$

where $\widehat{\mathrm{MSE}}_{1i} = V\left\{\overline{\mathbf{x}}'_{wi10}\boldsymbol{\beta}_{10} + v_{i,10} \left| (\mathbf{y}, \mathbf{x}) \right.\right\}$, and $\widehat{\mathrm{MSE}}_{2i} = \hat{\boldsymbol{\beta}}'_{10}\mathbf{V}_{xxi,10}\hat{\boldsymbol{\beta}}_{10}$. In the application of Section 4, we evaluate the effect of including the term $\widehat{\mathrm{MSE}}_{2i}$, which accounts for the increase in posterior MSE due to use of the sample mean of the covariate instead of the population mean, on the posterior MSE of the predictor.

3.3 Two-stage benchmarking

NASS obtains estimates of cash rental rates at the state level using data from a national survey conducted in June (the June Area Survey) in addition to the Cash Rent Survey. The state estimates are published before the county-level data from the Cash Rent Survey are fully processed. NASS also establishes estimates of cash rental rates for agricultural statistics districts. To retain internal consistency, appropriately weighted sums of county estimates must equal the district estimates and appropriately weighted sums of district estimates must equal the previously published state estimate. Letting $\hat{\theta}_{i10}$ be the benchmarked predictor for 2010, the benchmarking restrictions for a single time-point are defined by

$$\sum_{i \in d_k} w_{i10} \hat{\theta}_{i10} = \hat{\lambda}_{k10}, \tag{3.11}$$

and

$$\sum_{k=1}^{K} \eta_{k10} \hat{\lambda}_{k10} = \theta_{\text{pub10}}, \tag{3.12}$$

where $k = 1, \ldots, K$ index the districts, $w_{i10} = \left(\sum_{i \in d_k} z_{i10} \right)^{-1} z_{i10}$,

$$\eta_{k10} = \left(\sum_{k=1}^{K} \sum_{i \in d_k} z_{i10} \right)^{-1} \sum_{i \in d_k} z_{i10},$$

z_{i10} is the direct estimator of the acres rented in county i in year 2010, d_k is the index set for the counties in district k, $\hat{\lambda}_{k10}$ is the final estimate of the average cash rental rate for district k, and θ_{pub10} is the published estimate of the state-level cash rent per acre. We consider estimates for the year 2010 in (3.11) and (3.12) because we focus on estimation for 2010 in the analysis of Section 4.

We use the two-stage benchmarking procedure proposed by Ghosh and Steorts (2013) to define benchmarked estimates. The benchmarked estimates minimize the quadratic form

$$g(\mathbf{c}, \mathbf{b}) = \sum_{k=1}^{K} \sum_{i \in d_k} \xi_i \left(\hat{\theta}_{i10}^{B} - c_i \right)^2 + \sum_{k=1}^{K} \rho_k \left(\hat{\theta}_{k10,w}^{B} - b_k \right)^2 \tag{3.13}$$

subject to the constraints in (3.11) and (3.12), where $\mathbf{c} = (c_1, \ldots, c_D)$, D denotes the total number of counties, $\mathbf{b} = (b_1, \ldots, b_K)$, $\hat{\theta}_{k10,w}^{B} = \sum_{i \in d_k} w_{i10} \hat{\theta}_{i10}^{B}$, and (ρ_k, ξ_i) are constants selected by the analyst. We set $\xi_i = w_{i10}$ and $\rho_k = \eta_{k10}$, which gives the benchmarked estimates

$$\hat{\theta}_{i10} = \hat{\theta}_{i10}^{B} + \hat{\lambda}_{k(i)10} - \hat{\theta}_{k(i)10,w}^{B}, \tag{3.14}$$

with

$$\hat{\lambda}_{k(i)10} = \hat{\theta}_{k(i)10,w}^{B} + \frac{\left(\theta_{\text{pub10}} - \hat{\theta}_{w10}^{B} \right) \eta_{k(i)10} \left(1 + \eta_{k(i)10} \right)^{-1}}{\sum_{i \in d_{k(i)10}} \eta_{k(i)10}^2 \left(1 + \eta_{k(i)10} \right)^{-1}}, \tag{3.15}$$

for county i and district $k(i)$, respectively, where $k(i)$ is the district containing county i. In (3.15), $\hat{\theta}_{w10}^{B} = \sum_{k=1}^{K} \eta_{k10} \hat{\theta}_{k10,w}^{B}$. Each of the benchmarked estimates in (3.14) and (3.15) is a sum of the hierarchical Bayes predictor and an adjustment term. If the hierarchical Bayes predictor for the state is larger (smaller) than the previously published state total, then the adjustment is negative (positive), and the benchmarked county and district estimates are smaller (larger) than the hierarchical Bayes predictors. The posterior mean squared error of the benchmarked predictor for year t is

$$\text{MSE}_{i10}^{B\text{Bench}} = \text{MSE}\left(\hat{\theta}_{i10}^{B} \right) + \left(\hat{\theta}_{i10}^{B} - \hat{\theta}_{i10} \right)^2, \tag{3.16}$$

where $\text{MSE}\left(\hat{\theta}_{i10}^{B}\right)$ is defined in (3.10). See (You, Rao and Dick, 2004) for a derivation of the posterior MSE of a benchmarked predictor.

4 Results for non-irrigated cropland in Iowa, Kansas, and Texas

The model of Section 3 was fit to the non-irrigated cropland cash rental rates reported on the 2009 and 2010 Cash Rent Surveys for Iowa, Kansas, and Texas. These three states were chosen to reflect a range of situations. All counties in Iowa have estimates for corn yields, and cash renting is a relatively common way to rent non-irrigated cropland. Kansas is more agriculturally diverse than Iowa. According to agricultural specialists at NASS, share-renting is a more common way to rent land than cash renting in many parts of Texas, which may explain why realized sample sizes for some Texas counties are as small as zero or one report.

4.1 Covariate selection

The potential covariates for Iowa, Kansas, and Texas are listed in Section 2.2. For each state, the covariates include four variables related to the NCCPI, the total value of production for a county based on the 2007 Census of Agriculture, the expected sales for an operation recorded on the NASS list frame, the farm type recorded on the NASS list frame, and the acres rented for non-irrigated cropland recorded on the NASS Cash Rent Survey. For Iowa, an additional covariate is the corn yield for the county. For Kansas, an additional covariate is the non-irrigated yield index.

The covariates for each state were selected according to the following procedure. First, univariate models were fit to the data for 2009 and 2010 separately using maximum likelihood estimation. The univariate model used for covariate selection is of the form

$$y_{ijt} = \mathbf{x}'_{ijt}\boldsymbol{\alpha}_t + \nu_{it} + \epsilon_{ijt}, \tag{4.1}$$

where $\epsilon_{ijt} \sim \text{N}\left(0, \sigma^2_{\epsilon,t}\right)$, and $\nu_{it} \sim \text{N}\left(0, \sigma^2_{\nu,t}\right)$. The data for each farm operator who reported a non-irrigated cropland cash rental rate in year t were used to fit the univariate model for year t, regardless of whether or not the unit also reported a cash rental rate in year $s\,(s \neq t)$. The `R` function `lmer` in the package `nlme` is used for maximum likelihood estimation. For each year, step-wise selection using the `R` function `stepAIC` is performed using the BIC measure. The selected covariates are the variables that are in the minimum BIC models for both the 2009 and 2010 univariate models. We acknowledge that the minimum BIC model is a local minimum identified by the `stepAIC` procedure rather than a global minimum. The selected covariates for Iowa, Kansas, and Texas are as follows:

- Iowa: corn yield, expected sales, non-irrigated acres rented for cash.
- Kansas: non-irrigated yield index, expected sales, farm type.
- Texas: max-NCCPI, expected sales, farm type.

4.2 Estimates of correlation parameters

The exploratory analysis of Section 2.1 suggests a substantial correlation between the non-irrigated cropland cash rental rates for 2009 and 2010. Table 4.1 contains summaries of the posterior distributions of the correlations in the bivariate HB model defined in Section 3.1. The columns labeled "Median" are the posterior medians of the correlations, and lower and upper endpoints of the 95% credible intervals are the 2.5 and 97.5 percentiles of the posterior distributions of the correlations. Even though the variances of e_{ij09} and e_{ij10} are proportional to the inverses of the weights, the correlation is a constant because the weights cancel in the definition of the correlation.

Table 4.1
Posterior distributions of correlations between 2009 and 2010

	$\text{Cor}\{v_{i09}, v_{i10}\}$		$\text{Cor}\{e_{ij09}, e_{ij10}\}$	
State	**Median**	**95% Credible Interval**	**Median**	**95% Credible Interval**
Iowa	0.746	[0.611, 0.839]	0.570	[0.548, 0.592]
Kansas	0.919	[0.870, 0.950]	0.727	[0.701, 0.751]
Texas	0.884	[0.831, 0.921]	0.691	[0.667, 0.714]

The posterior medians of the county-level and unit-level correlations exceed 0.74 and 0.57, respectively. The lower endpoints of the 95% credible intervals exceed 0.61 and 0.54 for the county-level and unit-level correlations, respectively. For each state, the correlations at the level of the county are larger than the correlations for individual units. The significant correlations suggest the potential for an efficiency gain for the predictors relative to a univariate model.

4.3 Comparison of 2010 predictors for bivariate and univariate models

To demonstrate the gain in efficiency due to the use of the bivariate model relative to a univariate model, we compare the posterior mean squared errors of the predictors from the bivariate model to the posterior mean squared errors of the predictors from a corresponding univariate model. The assumptions of the univariate models are the same as the assumptions of the bivariate models except that the covariance parameters in $\mathbf{\Sigma}_{ee}$ and $\mathbf{\Sigma}_{vv}$ are assumed to equal zero. To fit the univariate models, we use inverse-gamma prior distributions for σ_{eet} and σ_{vvt} $(t = 09, 10)$.

To compare the bivariate and univariate models, we define the relative posterior MSE (RelMSE) for county i by

$$\text{RelMSE}_{i,10} = \frac{\text{MSE}_{i10}^{B\text{Bench}}}{\text{MSE}_{i10}^{\text{UNIBench}}}, \tag{4.2}$$

where $\text{MSE}_{i10}^{B\text{Bench}}$ is defined in (3.16) and $\text{MSE}_{i10}^{\text{UNIBench}}$ is the posterior MSE based on the corresponding univariate model. The average relative MSEs for Iowa, Kansas, and Texas are 88.71%, 97.27%, and 88.65%, respectively, where the average relative mean squared error for a state is $D^{-1}\sum_{i=1}^{D}\text{RelMSE}_{i,10}$. Note that the effects of both estimating the covariate mean and benchmarking are incorporated in the forms for the

posterior MSE for both the bivariate and univariate models. Because of the significant correlations in the model errors for the two time points, the posterior MSE from a bivariate model is smaller than the posterior MSE from the corresponding univariate model, and the average relative efficiencies are less than one.

To assess the effect of estimating the covariate population mean on the MSE of the predictor, we calculate the average of the ratios $\widehat{\text{MSE}}_{2i}\widehat{\text{MSE}}_{1i}^{-1}$ for $i = 1, \ldots, D$, where $\widehat{\text{MSE}}_{2i}$ and $\widehat{\text{MSE}}_{1i}$ are defined following (3.10). The ratios are 18.21%, 28.20%, and 21.07% for Iowa, Kansas, and Texas, respectively. Compared to Iowa and Texas, the contribution to the prediction MSE due to using the sample covariate mean instead of the population covariate mean is higher in Kansas, and this makes sense since Kansas is more agriculturally diverse. The relatively large average relative MSE for Kansas (97.27%) reflects the relatively large increase in posterior MSE due to estimating the covariate mean.

4.4 Model assessment

To assess model fit, we use the posterior predictive $\text{p}-$value, which measures departures between the observed data and the model. The posterior predictive $\text{p}-$value compares the posterior predictive distribution of selected summary statistics to the corresponding values obtained using the original sample. For the analysis below, we use only the elements observed in both 2009 and 2010 (set 1).

We consider two summary statistics: the mean for each year and the multivariate skewness. The mean for year t is the mean of the observations in set 1 for year t and is defined

$$\bar{y}_t = \left(\sum_{i=1}^{D} |A_i|\right)^{-1} \sum_{i=1}^{D} \sum_{j \in A_i} y_{ijt},$$

where A_i denotes the elements in set 1 for county i. The multivariate skewness is defined by

$$\hat{\gamma}_{1,p} = \left(\sum_{i=1}^{D} |A_i|\right)^{-1} \sum_{i=1}^{D} \sum_{k=1}^{D} \sum_{j \in A_i} \sum_{\ell \in A_i} m_{ijk\ell}^3,$$

where $m_{ijk\ell} = (\mathbf{y}_{ij} - \bar{\mathbf{y}})'\mathbf{S}^{-1}(\mathbf{y}_{k\ell} - \bar{\mathbf{y}})$, $\mathbf{y}_{ij} = (y_{ij,09}, y_{ij,10})'$, $\bar{\mathbf{y}} = (\bar{y}_{09}, \bar{y}_{10})'$, and $\mathbf{S} = \left(\sum_{i=1}^{D} |A_i| - 1\right)^{-1} \sum_{i=1}^{D} \sum_{j \in A_i} (\mathbf{y}_{ij} - \bar{\mathbf{y}})(\mathbf{y}_{ij} - \bar{\mathbf{y}})'$.

The posterior predictive $\text{p}-$value is defined as the proportion of summary statistics calculated with samples generated from the posterior predictive distribution that exceed the corresponding value based on the original sample. To be specific, let $T(\mathbf{y}^{(r)})$ be the summary statistic based on the r^{th} data set generated from the posterior predictive distribution, where the procedure to generate data from the posterior predictive distribution is defined in Appendix C. Let $T(\mathbf{y})$ be the corresponding statistic based on the original sample. The posterior predictive $\text{p}-$value is $R^{-1}\sum_{r=1}^{R} I[T(\mathbf{y}^{(r)}) > T(\mathbf{y})]$. A $\text{p}-$value close to 0.5 indicates that the model provides a reasonable fit to the sample data.

Table 4.2 contains the posterior predictive $\text{p}-$values for Iowa, Kansas, and Texas. For Kansas, the posterior predictive values indicate that the model is a good fit to the data. For Iowa and Texas, the posterior

predictive $p-$values indicate lack of fit. A further analysis of residuals suggests that the lack of fit may result from outliers. The posterior predictive $p-$values far from 0.5 may also arise because we only use the observations sampled in both 2009 and 2010 to calculate the posterior predictive $p-$values, while we use the full data set to fit the model.

Table 4.2
Posterior predictive P – values

State	Statistic	P – value
IA	Mean $t = 09$	1.000
	Mean $t = 10$	1.000
	Skewness	0.931
KS	Mean $t = 09$	0.291
	Mean $t = 10$	0.507
	Skewness	0.371
TX	Mean $t = 09$	0.025
	Mean $t = 10$	0.039
	Skewness	0.004

5 Conclusions and future work

We use a bivariate HB model to obtain predictors of county-level cash rental rates for non-irrigated cropland in Iowa, Kansas, and Texas. The model incorporates auxiliary information related to land quality, commodity values, and farm characteristics. Significant correlations exist between the 2009 and 2010 model random effects at both the unit and county levels. As a consequence, using the information in the 2009 cash rent estimates reduces the posterior MSE relative to a univariate model. The analysis of the bivariate HB model provides support that a more refined approach than that of Berg et al. (2014) is possible. To incorporate unit-level covariates with unknown population means, we add a level to the hierarchical model that justifies adding a term to the posterior mean squared error to account for uncertainty in the unknown population means of the unit-level covariates. Unlike Berg et al. (2014), the proposed bivariate HB model allows variability to change over time and accounts for effects of benchmarking on the MSE.

The analysis of the residuals and the posterior predictive $p-$values suggests that accounting for outliers may be an important way to substantially improve the model fit. One option is to consider a heavy-tailed distribution, such as a $t-$distribution or a mixture of normal distributions, that may represent the observed responses more appropriately than the assumed normal distribution. An extension of Gershunskaya (2010) to bivariate framework and Bayesian estimation is one possible way to approach the issue of outliers.

Acknowledgements

The National Agricultural Statistics Service (NASS) of the United States Department of Agriculture (USDA) supported this work. The authors are grateful to Wendy Barboza, Dan Beckler, Angie Considine,

Mark Harris, Sharyn Lavender, Joe Parsons, Scot Rumberg, Scott Shimmin, Curt Stock, and Linda Young from the National Agriculture Statistics Service. Further, the authors thank Bob Dobos from the National Resource Conservation Service and Rich Iovanna from the Farm Service Agency for their assistance in acquiring the National Commodity Crop Productivity Index. Without the generous assistance from these individuals, this research would have been impossible. The views expressed in this paper are those of the authors and do not necessarily represent the views of NASS or the USDA.

Appendix A

To specify the full conditional distributions for Gibbs sampling, we introduce notation. Let Θ_γ be the set of parameters except for the parameter denoted by γ. Let $\mathbf{X}_{ij} = \left(\mathbf{z}_{ij,09}, \mathbf{z}_{ij,10}\right)'$, where $\mathbf{z}_{ij,09} = \left(\mathbf{x}'_{ij,09}, \mathbf{0}'_{p_{10}}\right)'$, and $\mathbf{z}_{ij,10} = \left(\mathbf{0}'_{p_{09}}, \mathbf{x}'_{ij,10}\right)'$. Let $\mathbf{y}'_{ij} = \left(y_{ij,09}, y_{ij,10}\right)$. Let A_i be the set of units (farm operators) in county i that are in set 1, $B_{i,09}$ be the set of units in county i that are in set 2, and $B_{i,10}$ be the set of units in county i that are in set 3, where set 1, set 2, and set 3 are defined in Section 3. Full conditionals are as follows.

1. $\boldsymbol{\beta}\,|\left(\Theta_{\boldsymbol{\beta}}, \mathbf{y}\right) \sim \mathrm{N}\left(\boldsymbol{\Sigma}_{\beta\beta}\mathbf{r}_\beta, \boldsymbol{\Sigma}_{\beta\beta}\right)$, where

$$\begin{aligned}
\boldsymbol{\Sigma}_{\beta\beta} &= \left[\sum_{i=1}^{D}\sum_{j\in A_i}\mathbf{X}'_{ij}\mathbf{D}_{wij}^{0.5}\,\boldsymbol{\Sigma}_{ee}^{-1}\mathbf{D}_{wij}^{0.5}\mathbf{X}_{ij} + 10^{-6}\,\mathbf{I}_{p_{09}+p_{10}} + \boldsymbol{\Omega}\right]^{-1} \\
\boldsymbol{\Omega} &= \text{block-diag}\left(\tau_{e,09}^{-2}\sum_{i=1}^{D}\sum_{j\in B_{i,09}} w_{ij,09}\mathbf{x}_{ij,09}\mathbf{x}'_{ij,09}, \tau_{e,10}^{-2}\sum_{i=1}^{D}\sum_{j\in B_{i,10}} w_{ij,10}\mathbf{x}_{ij,10}\mathbf{x}'_{ij,10}\right) \\
\mathbf{r}_\beta &= \sum_{i=1}^{D}\sum_{j\in A_i}\mathbf{X}'_{ij}\mathbf{D}_{wij}^{0.5}\boldsymbol{\Sigma}_{ee}^{-1}\mathbf{D}_{wij}^{0.5}\left(\mathbf{y}_{ij} - \mathbf{v}_i + \mathbf{r}_{\beta 2}\right),
\end{aligned} \tag{A.1}$$

and

$$\mathbf{r}_{\beta 2} = \begin{pmatrix} \sum_{i=1}^{D}\sum_{j\in B_{i,09}} \tau_{e,09}^{-2} w_{ij,09}\mathbf{x}_{ij,09}\left(y_{ij,09} - \nu_{i,09}\right) \\ \sum_{i=1}^{D}\sum_{j\in B_{i,10}} \tau_{e,10}^{-2} w_{ij,10}\mathbf{x}_{ij,10}\left(y_{ij,10} - \nu_{i,10}\right) \end{pmatrix}. \tag{A.2}$$

2. $\boldsymbol{\Sigma}_{ee}\,|\left(\boldsymbol{\Theta}_{\boldsymbol{\Sigma}_{ee}}, \mathbf{y}\right) \sim \text{Inverse-Wishart}\left(\mathbf{A}_e, d_e\right)$, where $d_e = \sum_{i=1}^{D}|A_i| + 0.001$, and

$$\mathbf{A}_e = \sum_{i=1}^{D}\sum_{j\in A_i}\mathbf{D}_{w_{ij}}^{0.5}\left(\mathbf{y}_{ij} - \mathbf{v}_i - \mathbf{X}_{ij}\boldsymbol{\beta}\right)\left(\mathbf{y}_{ij} - \mathbf{v}_i - \mathbf{X}_{ij}\boldsymbol{\beta}\right)'\mathbf{D}_{w_{ij}}^{0.5}. \tag{A.3}$$

3. $\boldsymbol{\Sigma}_{vv}\,|\left(\boldsymbol{\Theta}\mathrm{Q}_{\boldsymbol{\Sigma}_{vv}}, \mathbf{y}\right) \sim \text{Inverse-Wishart}\left(\mathbf{A}_v, d_v\right)$, where

$$d_v = D + 0.001, \tag{A.4}$$

and

$$\mathbf{A}_{\nu} = \sum_{i=1}^{D} \mathbf{v}_i \mathbf{v}_i'. \tag{A.5}$$

$\tau_{e,t}^2 \left| \left(\boldsymbol{\Theta}_{\tau_{e,t}^2}, \mathbf{y} \right) \right. \sim$ Inverse-Gamma (a_{et}, d_{et}), where

$$d_{et} = \sum_{i=1}^{D} \left| B_{i,t} \right| + 0.001, \tag{A.6}$$

and

$$a_{et} = \sum_{i=1}^{D} \sum_{j \in B_{it}} \mathbf{D}_{wi_j} \left(y_{ij,t} - \nu_{i,t} - \mathbf{x}_{ij,t} \boldsymbol{\beta}_t \right)^2. \tag{A.7}$$

4. $\mathbf{v}_i \left| \left(\boldsymbol{\Theta}_{\nu_i}, \mathbf{y} \right) \right. \sim \mathrm{N}\left(\boldsymbol{\mu}_{\nu\nu}, \mathbf{M}_i^{-1} \right)$, where

$$\mathbf{M}_i = \left(\boldsymbol{\Sigma}_{\nu\nu}^{-1} + \boldsymbol{\Sigma}_{ee,wi}^{-1} + \boldsymbol{\Omega}_{ee,wi}^{-1} \right)^{-1}, \tag{A.8}$$

$$\boldsymbol{\mu}_{\nu\nu} = \mathbf{M}_i^{-1} \left(\mathbf{r}_{i_1} + \mathbf{r}_{i_2} \right), \boldsymbol{\Sigma}_{ee,wi} = \sum\nolimits_{j \in A_i} \mathbf{D}_{w_{ij}}^{0.5} \boldsymbol{\Sigma}_{ee}^{-1} \mathbf{D}_{w_{ij}}^{0.5},$$

$$\begin{aligned} \mathbb{W}_{ee,wi} &= \mathrm{diag}\left(\tau_{e,09}^{-2} \sum_{j \in B_{i09}} w_{ij,09}, \tau_{e,10}^{-2} \sum_{j \in B_{i10}} w_{ij,10} \right), \\ \mathbf{r}_{i_1} &= \sum_{j \in A_i} \mathbf{D}_{w_{ij}}^{0.5} \boldsymbol{\Sigma}_{ee}^{-1} \mathbf{D}_{w_{ij}}^{0.5} \left(\mathbf{y}_{ij} - \mathbf{X}_{ij}' \boldsymbol{\beta} \right), \end{aligned} \tag{A.9}$$

and

$$\mathbf{r}_{i_2} = \begin{pmatrix} \sum_{j \in B_{i,09}} w_{ij,09} \left(y_{ij,09} - \mathbf{x}_{ij,09}' \boldsymbol{\beta}_{09} \right) \tau_{e,09}^{-2} \\ \sum_{j \in B_{i,10}} w_{ij,10} \left(y_{ij,10} - \mathbf{x}_{ij,10}' \boldsymbol{\beta}_{10} \right) \tau_{e,10}^{-2} \end{pmatrix}. \tag{A.10}$$

Appendix B

We define an estimator of the diagonal elements of $V\left\{ \overline{\mathbf{x}}_{wi10} \left| \overline{\mathbf{x}}_{N_i,10} \right. \right\} := \mathbf{V}_{xxi,10}$ corresponding to unit-level covariates, $\mathbf{x}_{ijk10}$ for $k = 1, \ldots, p_{10}$. The variance estimator is based on a working assumption that a probability proportional to size with replacement (PPSWR) sample is a reasonable approximation for the cash rent survey design. As discussed in Cochran (1977), use of a PPSWR approximation is often reasonable if the sampling fraction is less than 10%. Suppose the draw probability for element j in area i for the PPSWR design is $p_{ij} = n_{i10}^{-1} w_{ij,10}^{-1}$. Because $n_{i10} \leq 1$ for some counties, we define the estimator of the diagonal elements of $\mathbf{V}_{xxi10}$ corresponding to unit level covariates as a convex combination of a direct estimator of the within-area variance and a variance estimator that pools information across all counties in

a state. For area i with $n_{i10} > 1$, the estimate of the within-area variance of x_{ijk10} under the assumed PPSWR design (Särndal, Swensson and Wretman, 1992) is given by

$$S_{ik10}^2 = \frac{n_{i10}^2}{\left(\sum_{j=1}^{n_{i10}} w_{ij,10}\right)^2 (n_{i10} - 1)} \sum_{j=1}^{n_{i10}} w_{ij,10}^2 \left(x_{ijk10} - \overline{x}_{wik10}\right)^2,$$

where $\overline{x}_{wik10}$ is the k^{th} element of $\overline{\mathbf{x}}_{wi10}$. The pooled estimator of the variance is defined by

$$S_{pk10}^2 = \frac{1}{w_{..10}^2 \left(\tilde{n}_{10} - \tilde{D}_{10}\right)} \sum_{i=1}^{D} \left(n_{i10}^2 \sum_{j=1}^{n_{i10}} w_{ij,10}^2 \left(x_{ijk10} - \overline{x}_{wik10}\right)^2 \right) I\left[n_{i10} > 1\right],$$

where $w_{..10} = \sum_{i=1}^{D} \left(\sum_{j=1}^{n_i} w_{ij,10}\right) I\left[n_i > 1\right]$, $\tilde{n}_{10} = \sum_{i=1}^{D} n_{i10} I\left[n_{i10} > 1\right]$, and $\tilde{D}_{10} = \sum_{i=1}^{D} I\left[n_{i10} > 1\right]$. The element of the diagonal covariance matrix $\mathbf{V}_{xxi,10}$ corresponding to the k^{th} unit level covariate is then given by

$$\hat{V}\left\{\overline{x}_{wik10}\right\} = n_{i10}^{-1} \hat{S}_{ik10}^2 I\left[n_{i10} \neq 1\right] + n_{i10}^{-1} S_{pk10}^2 I\left[n_{i10} = 1\right], \tag{B.1}$$

where

$$\hat{S}_{ik10}^2 = \frac{n_{i10}}{n_{i10} + 1} S_{ik10}^2 + \frac{1}{n_{i10} + 1} S_{pk10}^2. \tag{B.2}$$

We provide a heuristic justification for the combination in (B.2), which is related to Haff (1980). Let $S^2 = n^{-1} \sum_{i=1}^{n} X_i^2$, where $X_i \sim \mathrm{N}(0, \sigma^2)$. Assume $\sigma^2 \sim$ Inverse-Gamma (α, β), where $E\left[\sigma^2\right] := v = \beta (\alpha - 1)^{-1}$. Then,

$$E\left[\sigma^2 \mid S^2\right] = \frac{2(\alpha - 1) v}{n + 2(\alpha - 1)} + \frac{n S^2}{n + 2(\alpha - 1)}.$$

In application to estimation of county-level cash rental rates, S_{ik10}^2 plays the role of S^2 and S_{pk10}^2 plays the role of v. Taking $\alpha = 1.5$ gives the desired multiplier.

Appendix C

Data simulation from the posterior distributions

Consider the posterior samples for $\boldsymbol{\beta}_{09}$, $\boldsymbol{\beta}_{10}$, $\boldsymbol{\Sigma}_{vv}$ and $\boldsymbol{\Sigma}_{ee}$, denoted by $\boldsymbol{\beta}_{09}^s$, $\boldsymbol{\beta}_{10}^s$, $\boldsymbol{\Sigma}_{vv}^s$ and $\boldsymbol{\Sigma}_{ee}^s$, respectively, for $s = 1, \ldots, S$. Define

$$\boldsymbol{\Sigma}_{eeij}^s := \mathbf{D}_{wij}^{-0.5} \boldsymbol{\Sigma}_{ee}^s \mathbf{D}_{wij}^{-0.5},$$

for $s = 1, \ldots, S$. Draw replicates ν^r_{i09}, ν^r_{i10}, y^r_{ij09} and y^r_{ij10}, for $r = 1, \ldots, R$, following model (1-3) and properties of the multivariate conditional normal distribution as follows:

$$
\begin{aligned}
\nu^r_{i09} &\sim N\left(\mathbf{0},\, \boldsymbol{\Sigma}^r_{\nu\nu,(11)}\right) \\
\nu^r_{i10} &\sim N\left(\left(\boldsymbol{\Sigma}^r_{\nu\nu,(11)}\right)^{-1} \boldsymbol{\Sigma}^r_{\nu\nu,(12)} \nu^r_{i09},\, \left(\boldsymbol{\Sigma}^r_{\nu\nu,(11)}\right)^{-1} \boldsymbol{\Sigma}^r_{\nu\nu,(11)}\, \boldsymbol{\Sigma}^r_{\nu\nu,(22)} - \left(\boldsymbol{\Sigma}^r_{\nu\nu,(12)}\right)^2\right), \\
\mu^r_{i09} &= \mathbf{x}'_{ij09} \boldsymbol{\beta}^r_{09}, \\
y^r_{ij09} &\sim N\left(\mu^r_{i09} + \nu^r_{i09},\, \boldsymbol{\Sigma}^r_{eeij,(11)}\right) \\
\mu^r_{i10} &= \mathbf{x}'_{ij10} \boldsymbol{\beta}^r_{10}, \\
y^r_{ij10} &\sim N\left(\mu^r_{i10} + \nu^r_{i10} + \left(\boldsymbol{\Sigma}^r_{eeij,(11)}\right)^{-1} \boldsymbol{\Sigma}^r_{eeij,(12)} \left(y^r_{ij09} - \mu^r_{i09} - \nu^r_{i09}\right),\right. \\
&\qquad \left. \left(\boldsymbol{\Sigma}^r_{eeij,(11)}\right)^{-1} \left(\boldsymbol{\Sigma}^r_{eeij,(11)} \boldsymbol{\Sigma}^r_{eeij,(22)} - \left(\boldsymbol{\Sigma}^r_{eeij,(12)}\right)^2\right)\right).
\end{aligned}
$$

Although the number of posterior samples is $S = 20{,}000$, we construct $R = 1{,}901$ replicates, where r is selected from the sequence 1,000 to T by skipping every 10 samples.

References

Battese, G.E., Harter, R.M. and Fuller, W.A. (1988). An error-components model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*, 83, 28-36.

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.

Cochran, W.G. (1977). *Sampling Techniques*. *3rd Edition*, New York: John Wiley & Sons, Inc.

Datta, G.S., Day, B. and Maiti, T. (1998). Multivariate Bayesian small area estimation: An application to survey and satellite data. *Sankhyā: The Indian Journal of Statistics*, 60, 344-362.

Dhuyvetter, D., and Kastens, T. (2009). *Kansas Land Values and Cash Rents at the County Level*.

Dobos, R.R., Sinclair, H.R. and Robotham, M.P. (2012). *User Guide for the National Commodity Crop Productivity Index (NCCPI), Version 2.0*, NRCS/USDA publication.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2009). *Bayesian Data Analysis: Second Edition*, CRC Press.

Gershunskaya, J. (2010). Robust small area estimation using a mixture model. *Proceedings of the Survey Research Methods Section,* American Statistical Association.

Ghosh, M., and Steorts, R. (2013). Two-stage Bayesian benchmarking as applied to small area estimation. *TEST*, 22, 670-687.

Haff, L.R. (1980). Empirical Bayes estimation of the multivariate normal covariance matrix. *Annals of Statistics*, 586-597.

Kim, J.K., Park, S. and Lee, Y. (2017). Statistical inference using generalized linear mixed models under informative cluster sampling. *Canadian Journal of Statistics*, DOI: 10.1002/cjs.11339.

Lohr, S.L., and Prasad, N.G.N. (2003). Small area estimation with auxiliary survey data. *Canadian Journal of Statistics*, 31, 383-396.

Rao, J.N.K., and Molina, I. (2015). *Small Area Estimation*, New York: John Wiley & Sons, Inc.

Särndal, C.-E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*, New York: Springer-Verlag.

Woodard, S., Paulson, N., Baylis, K. and Woddard, J. (2010). Spatial analysis of Illinois agricultural cash rents. *The Selected Works of Kathy Baylis*. http://works.bepress.com/kathy_baylis/29.

You, Y., and Rao, J.N.K. (2002). A pseudo-empirical best linear unbiased prediction approach to small area estimation using survey weights. *Canadian Journal of Statistics*, 30, 431-439.

You, Y., Rao, J.N.K. and Dick, P. (2004). Benchmarking hierarchicial Bayes small area estimators with applications in census undercoverage estimation. *Statistics in Transition*, 6, 631-640.