Journal of Official Statistics, Vol. 38, No. 3, 2022, pp. 709–732, http://dx.doi.org/10.2478/JOS-2022-0032

Hierarchical Bayesian Model with Inequality Constraints for US County Estimates

Lu Chen[1], *Balgobin Nandram*[2], *and Nathan B. Cruze*[3]

In the production of US agricultural official statistics, certain inequality and benchmarking constraints must be satisfied. For example, available administrative data provide an accurate lower bound for the county-level estimates of planted acres, produced by the U.S. Department of Agriculture's (USDA) National Agricultural statistics Services (NASS). In addition, the county-level estimates within a state need to add to the state-level estimates. A sub-area hierarchical Bayesian model with inequality constraints to produce county-level estimates that satisfy these important relationships is discussed, along with associated measures of uncertainty. This model combines the County Agricultural Production Survey (CAPS) data with administrative data. Inequality constraints add complexity to fitting the model and present a computational challenge to a full Bayesian approach. To evaluate the inclusion of these constraints, the models with and without inequality constraints were compared using 2014 corn planted acres estimates for three states. The performance of the model with inequality constraints illustrates the improvement of county-level estimates in accuracy and precision while preserving required relationships.

Key words: Administrative data; bayesian diagnostic; benchmarking; crop acreage estimates; small area estimation; sub-area models; survey data.

1. Introduction

The National Agricultural statistics Service (NASS), the primary statistical data collection agency within the U.S. Department of Agriculture (USDA), conducts the County Agricultural Production Survey (CAPS) annually. CAPS provides county-level estimates by commodity crop for the following estimands: planted acres, harvested acres, yield and production. 'Crop type by county' represents a planned domain, in the sense that the CAPS multivariate-probability-proportional-to-size design and sample selection is specifically intended to support NASS's county-level data products. However, the number of survey reports obtained for each domain can vary widely due to issues of survey nonresponse, genuine differences in planting decisions each year, and in the inherent complexity of sampling for the breadth of crops of interest nationwide. The current method of producing

[1] National Institute of Statistical Sciences, 1750 K Street, NW, Suite 1100, Washington. D.C., 20006-2306, U.S.A. Email: lchen@niss.org
[2] Worcester Polytechnic Institute and USDA National Agricultural statistics Service, Department of Mathematical Sciences, Stratton Hall, 100 Institute Road, Worcester, MA 01609. U.S.A. Email: balnan@wpi.edu
[3] USDA National Agricultural statistics Service, 1400 Independence Avenue, SW, Washington, D.C. 20250-2054. Email: nathan.cruze@gmail.com
Acknowledgments: The findings and conclusions in this article are those of the authors and should not be construed to represent any official USDA, or US Government determination or policy. This research was supported by the intramural research program of the USDA, NASS. Dr. Nandram's work was supported by a grant from the Simons Foundation (#353953, Balgobin Nandram). Dr. Cruze's contributions to this work were made during his tenure at USDA, NASS.

official county-level crop estimates is an expert assessment conducted by NASS's Agricultural statistics Board (ASB), which incorporates multiple sources of information. The information includes CAPS estimates and administrative data whenever it is available. These county-level estimates are key indicators to farmers, ranchers and a number of federal and state agencies for decision making. Two USDA agencies, the Farm Service Agency (FSA) and the Risk Management Agency (RMA), consider the estimates as part of their processes for distributing farm subsidies and insurance respectively.

Given the importance of the crops county estimates program, NASS engaged a panel of experts under the National Academies of Sciences, Engineering, and Medicine for guidance and recommendations on implementing models for integrating multiple sources of information to provide county-level crop estimates with measures of uncertainty. The panel's recommendations were issued in a publicly available report; see National Academies of Sciences, Engineering, and Medicine NASEM 2017. See also Cruze et al. (2019) for a recapitulation of some of the panel's findings. In the traditional process of setting official statistics, the ASB has relied on standard processes, multiple data sources, historical performance of these sources, and expert judgment. The ASB analyzes the survey estimates and integrates them with multiple data sources through a series of informal composite estimators. (See NASEM 2017, 27–28; Cruze et al. 2019, sec. 2.) Final estimates are checked for *coherence with external administrative totals* that are interpreted as minimum amounts of activity known to have taken place in the county, and the estimates are rounded in accordance with NASS rounding rules. In a statistical sense, the ASB results are not reproducible and measures of uncertainty have not been produced with the traditional data product.

In recent years, small area models have gained increased attention by academic researchers and government agencies. Small area estimation models can "borrow strength" from related areas across space and/or time or through auxiliary information to provide "indirect" but reliable estimates for small areas while also increasing precision. One challenge of a model-based approach is to provide reliable and coherent estimates that satisfy important relationships nested among estimates and administrative data. The NASS county-level official estimates of planted acres should be greater than or equal to the corresponding available administrative totals that represent known minimum amounts of planting activity within the county, while also satisfying benchmarking constraints so that county-level estimates add up to the state-level estimates. In this article, hierarchical Bayesian models with constraints for small area estimation are discussed and applied to NASS's planted area estimates of corn for grain with reference to the 2014 crop year. With the goal of improving transparency of processes and quantifying the uncertainty associated with each estimate, NASS implemented the described model-based approach for estimating county-level planted area totals for thirteen commodity crops nationwide beginning with the 2020 crop year.

Two major types of small area models, area-level and unit-level models, have been developed using both frequentist and Bayesian methods. Pfeffermann (2013) and Rao and Molina (2015) provide a comprehensive overview of the development, methods and application of small area estimation including various types of area-level and unit-level models. For continuous responses, the first and most common model is the Fay-Herriot (FH) model (Fay and Herriot 1979) in small area estimation. It is an area-level model based on a "normal-normal-linear" assumption. That is, the direct estimates and area-level random effects are both assumed to follow normal distribution and a linear regression function relates the true estimates of interest to covariates. The popular unit-level model, nested-error regression (NER) model, was proposed by Battese et al. (1988) when data are available on the individual sampled units. The NER model is also developed under the normality assumption.

The objective of NASS crops county estimates program is to incorporate different sources of auxiliary information with survey estimates in the model to provide coherent and reliable estimates with associated measures of uncertainty. The modeling strategies in both frequentist and Bayesian methods could operate in similar way. However, Bayesian approaches are more straightforward for obtaining estimates for any known functions of the model parameters. In addition, Bayesian methodology is well suited for inequality constrained problems as it naturally provides a framework that allows complex constraints via hierarchical models. Recent studies and papers related to the NASS crops county estimates program have shown that hierarchical Bayesian small area models can incorporate auxiliary sources of data to improve county-level survey estimation of crop totals with measures of uncertainty. Battese et al. (1988) introduced the unit-level models for small area estimation based on nested error linear regression. They combined survey indications with satellite data. Erciulescu et al. (2019) proposed and implemented a double shrinkage hierarchical Bayesian sub-area level model to provide the acreage estimates with associated measures of uncertainty. The paper discussed the results when integrating different data sources and showed that the county-level model-based acreage estimates decreased the coefficients of variation relative to the survey ones. Erciulescu et al. (2020) discussed the challenges of missing data, either survey responses or administrative data, when fitting hierarchical Bayesian sub-area level model to obtain the crops total estimates for the whole nation. In these two papers, the state-to-county benchmarking constraint is included.

Increasingly, constrained estimation problems have found application and international importance in the small area estimation literature. Sen et al. (2018) proposed the method to conduct inference for a constrained posterior and project samples to the constrained space through a minimal distance mapping. Instead of placing a prior on the constraint space and conducting posterior computation, a general formulation of projected posteriors in a Bayesian decision-theoretic framework is provided. Cruze et al. (2019) identified constraints among estimates and administrative data as a necessity and allowed for the possibility of different constraints by small area. Whereas the inequality constraint problems were not addressed in the aforementioned, NASS-authored literature, Nandram et al. (2022) addressed the inequality constraint problem and proposed several hierarchical Bayesian models for NASS crops county-level planted area estimates which have ultimately been used in practice by NASS effective with the 2020 crop year. They discussed the methodologies of fitting constrained models and provided a simulation study to show the performance of all models.

In this article, models with inequality constraints are discussed and implemented to address the needs and challenges of inequality and benchmarking constraints that NASS official statistics must satisfy. The models with inequality constraints of Nandram et al. (2022) are applied to 2014 NASS CAPS data. In Section 2, data sources and some particular needs of the NASS crops county estimates for total planted acres are presented.

In Section 3, hierarchical Bayesian models with inequality constraints are proposed to produce reliable and coherent county-level estimates and associated measures of uncertainty. External ratio benchmarking is applied to the county-level estimates so that they sum to state targets. The results are contrasted with those obtained from unconstrained models. In Section 4, a case study based on three different states shows the model-based estimation results and highlights the different performances of the constrained models and the unconstrained models. Conclusions and future work are presented in Section 5.

2. Data Sources and Requirements

2.1. Survey Data

Although NASS has been producing official county-level agricultural estimates since 1917, it was in 2011 that NASS completely implemented the large-scale probability survey, CAPS, to provide county-level official estimates for many principle small grains and row crops in several states.

The CAPS survey uses a Multivariate Probability Proportional to Size (MPPS) sample design. The target population for CAPS is all agricultural operators with cropland and/or storage capacity in any of the eligible states. The NASS list frame includes all known agricultural establishments. The list frame for CAPS consists of those NASS list frame records with positive planted acres or storage capacity of the desired commodities in the previous year (NASEM 2017, 111–117). Sample size is dependent upon the number of operations in the universe list and the variability of data among operations on a given list. Sample sizes vary widely among states, and the number of obtained reports will vary by state, commodity crop, and county, but there is some effort to treat ‘county by crop’ as a planned domain, and construct samples accordingly.

The list of crops and states in CAPS may change year to year depending on the requirement of coverage for federally mandated program crops and others. Figure 1 shows

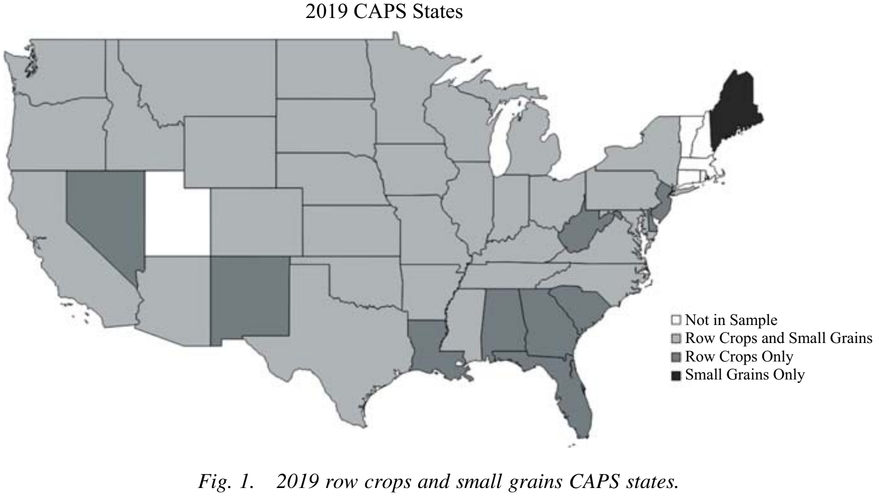

Fig. 1. 2019 row crops and small grains CAPS states.

the 2019 CAPS states. Four different colors indicate the category that each state is in for the 2019 CAPS. The state in black was for small grains CAPS only. States in dark gray were for row crops CAPS only. States in light gray were for both small grains and row crops CAPS. States in white were not included in 2019 CAPS. The row crops CAPS (e.g., corn, soybeans) was conducted in 41 states. The small grains CAPS (e.g., barley, oats) was conducted in 32 states. No other states were included in 2019 CAPS.

As discussed in the introduction, the smallest area at which CAPS produces estimates is the county. Historically, NASS has also produced estimates for an intermediate domain called the agricultural statistics district (ASD). Each ASD is comprised of contiguous counties in the state. Both county-level and ASD level survey estimates and associated variance estimates are available in CAPS summary. The state-level planted acreage estimates are published before the completion of data collection for the CAPS. Therefore, when setting the county level estimates, an external state benchmarking constraint exists. In the traditional estimation process, the ASB reviews the direct estimates based on CAPS and relevant auxiliary information to set county estimates that aggregate to those state targets.

2.2. Auxiliary Data

NASS obtains auxiliary sources of information on crop acres from FSA and RMA. Both agencies have farmer-reported administrative data on planted acres. While FSA and RMA programs are popular, they are not compulsory. The activity of some parts of the population may be absent in either record. The participation rates can vary by crop, by state, and even by locality within state. For example, the rates of enrollment in FSA programs for corn are typically higher in so-called corn-belt states, for example Illinois, than in some other states such as Ohio or Pennsylvania.

As described in NASEM (2017, 20), "FSA defines the common land unit (CLU) as an individual, contiguous farming parcel, which is the smallest unit of land that has a permanent, contiguous boundary; common land cover and land management; and a common owner or common producer association." FSAmaintains a database of these digitized, geolocated field boundaries for the entire United States. The size and location of the fields are known with accuracy. The *contents* contained within these field boundaries remain empty until planted acreages are reported by farmers to FSA each year. Farmers who opt to participate in FSA programs must certify their planting activity by prescribed due dates. The process of certification typically entails a visit by farmers to an FSA office, where the farmer is assisted in identifying the fields (CLUs) operated on maps. The farmer then provides the acreages by type of crop planted in a standard form containing all associated identifiers with the parcels operated. Deliberate misreporting is dissuaded under penalty of "loss of program benefits for noncompliance." The FSA handbook on *Acreage and Compliance Determinations* (USDA FSA 2018) details the procedures for certification, as well as quality assurance and compliance procedures at length.

As overseer of the Federal Crop Insurance Corporation, RMA receives administrative data on planted acres as farmers enroll in insurance coverage through approved insurance providers or file claims that are associated with these programs. Farmers may choose not participate in any crop insurance programs, or they may not insure all commodity crops they choose to grow. The participation rates in crop insurance can vary by state and commodity. For example, in RMA's own analysis of the 2015 federal crop insurance portfolio market penetration, defined as the percent of national planted area totals (estimates produced by NASS) that are also insured acres under these programs, it was found that 89% of corn acres, 90% of soybean acres, 73% of barley acres, and just 17% of oats acres were insured (USDA RMA 2017). Accordingly, NASS treats the RMA administrative data on planted acres as a useful lower bound on the planted acreage.

Non-probability sources like the FSA and RMA programs administrative data are not free of nonsampling errors. Foremost, neither of these collections represents a registry of the activity of all population units, and therefore, totals obtained at any geographic level from either source may subject to some degree of under coverage or provide underestimates of population totals. Good (2014) discussed this in the context of comparing national totals from NASS and from FSA. The aforementioned RMA analysis pointed to likely undercoverage nationwide. As a department, USDA, takes steps to mitigate other types of nonsampling error that could affect the quality of the reported data, through minimizing opportunities or incentives to misreport, through ongoing quality assurance procedures, and in the case of FSA, through geospatially resolving the collected data to the county, and more specifically, to the field, where the crop was planted. With these combinations of factors and the need to produce estimates that are coherent given these other USDA data sources, NASS has interpreted the administrative data as *informative lower bounds* in the construction of official county estimates produced under the traditional ASB process and seeks to retain that feature in any candidate model for planted area. While there can be significant overlap of FSA and RMA data, not all operations will participate in both. Because NASS treats both FSA and RMA data as the lower bounds of the county-level planted acreage estimates, the definition of the lower bound in the constrained models is the maximum of both sources of administrative data. That is, where FSA and RMA acreages may differ, the larger is taken as a firm lower bound NASS estimates should respect.

2.3. Important Relationships for Planted Acres

In the production of the official statistics for total acres reported by NASS, benchmarking *and inequality constraints* should be satisfied. NASS sometimes describes its procedures as 'top-down', meaning that national and state estimates are published before the sub-state ASD and county estimates, even while additional CAPS data collection may be ongoing (NASEM 2017, 24). In practical terms, it means that official county-level estimates will have an external target for benchmarking county totals to the published state total. Additionally, NASS's official estimates of planted acres should cover corresponding available administrative data: FSA and RMA planted acreage data within any given geographic boundary, such as the US, a state, and county. The differences between NASS official statistics and FSA administrative data of total planted acreage for corn, soybeans, barley and oats at US level from 2012 to 2019 are displayed in Figure 2. Each plot shows that the differences between NASS official estimates and FSA data are all positive at the US level. However, the county-level survey estimates of the planted acreage do not always satisfy the constraints.

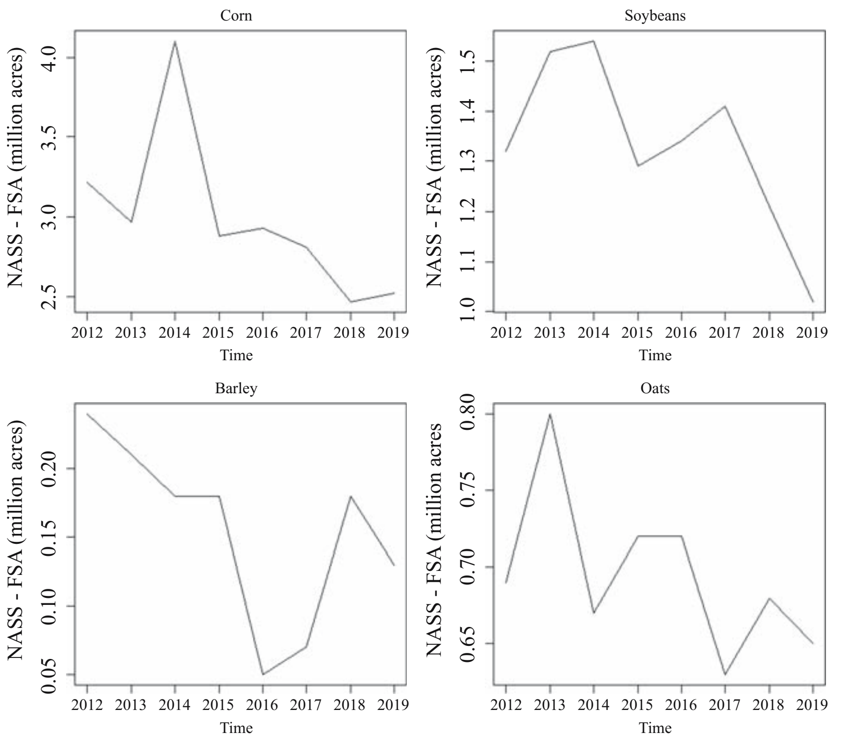

Fig. 2. The differences of US-level planted acreage estimates of several commodities between NASS and FSA.

Figure 3 indicates that the points in the plot of the survey estimates in log10 scale versus the FSA data in log10 scale are scattered around the 45 degree line. Some of the survey estimates are one or two standard deviations below the corresponding FSA or RMA data. This introduces difficulties for models without constraints to preserve the relationships. However, inequality constraints must be incorporated into the model so that all known relationships are satisfied at all levels before NASS can rely on model-based estimates as the foundation for the final official estimates.

3. Models

Bayesian area-level and sub-area level models are popular in small area estimation. They are excellent reproducible tools that combine survey data and auxiliary data to produce reliable estimates for areas. In this paper, models with constraints are considered based on Nandram et al. (2022). For comparison, the model without constraints are from Erciulescu et al. (2020).

Two model assumptions are made for both constrained and unconstrained model. First, it is assumed that the sampling variances are known and valid estimates from the survey summary in both area-level and subarea level sampling models. The modeling strategies are developed to deal with the crop county estimates including different commodities in all states covered by CAPS. Whereas Erciulescu et al. (2019) developed and compared models for direct estimates scaled by the sample sizes with a hierarchy for sampling variances, here

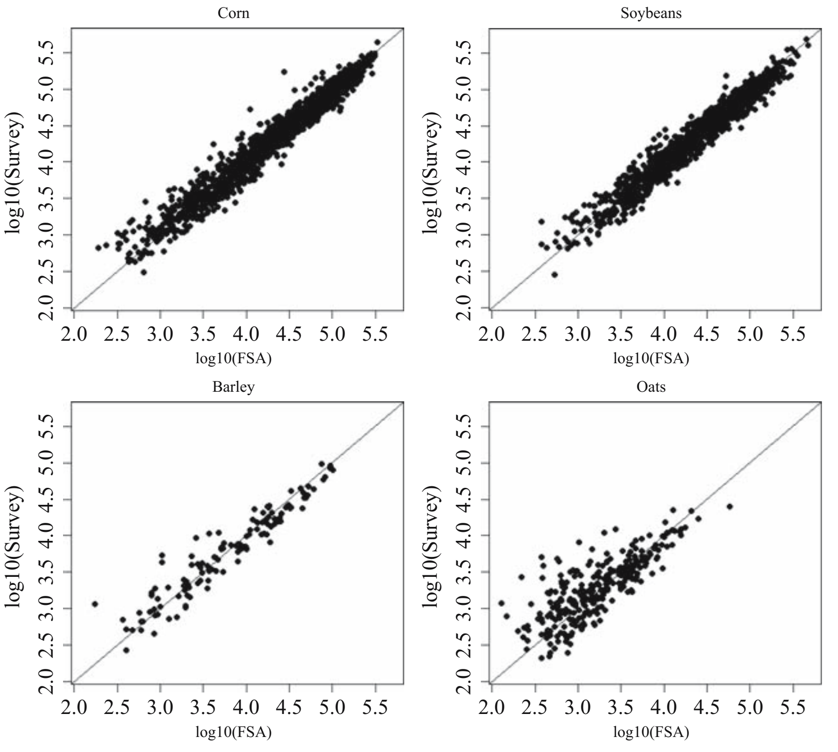

Fig. 3. The county-level planted acreage estimates (log10 scale) of several commodities for CAPS and FSA in all eligible counties.

we adopted the models for the direct estimates described in Erciulescu et al. (2020). Adding an extra model assumption for the sampling variance is not feasible, and so we assume that the sampling variances are fixed to avoid computational difficulties. Second, assuming normality of direct estimates is a practical method with good performance that provides estimates for counties with sample sizes as small as one and zero. This is impossible under the model specification in Erciulescu et al. (2019) because the sample sizes are denominators in the models. On the other hand, in our case, each county has its own unique inequality constraint and the sum needs to satisfy another benchmarking constraint.

In this section, models, with and without constraints, are presented and applied in a case study of 2014 corn data. They are illustrated for one state and one commodity, that is, all parameters are state and commodity-specific. The area-level model without inequality constraints was first introduced by Fay and Herriot (1979), where an area represents a county. The sub-area level models without inequality constraints were discussed by Fuller and Goyeneche (1998) and Torabi and Rao (2014) as an extension of FH model. Nandram et al. (2022) propose and discuss both area and sub-area level models to address the inequality constraints into the models.

3.1. Models Without Constraints

Erciulescu et al. (2020) discussed and applied a hierarchical Bayesian sub-area model to estimate the number of planted and harvested acres. In their paper, the state benchmarking constraint is handled by ratio benchmarking in the output analysis, but inequality constraints are not addressed either in the model or in the output analysis. In this article, this model without inequality constraints is referred to the unconstrained model and several comparisons between this type of model and models with inequality constraints (constrained models) will be presented in Section 4.

In the sub-area level models, an area is an ASD and a subarea is a county. Let $i = 1,\ldots,m$ be an index for m ASDs in the state and $j = 1,\ldots,n_i$ be an index for the county in the i^{th} ASD. The survey estimate of planted acreage in county j in district i is denoted by $\hat{\theta}_{ij}$ and the associated survey variance is $\hat{\sigma}^2_{ij}$. The total number of counties in a state is $\sum_{i=1}^{m} n_i$. The auxiliary data used in the models are x_{ij}, including an intercept.

The sub-area hierarchical Bayesian model is

$$\hat{\theta}_{ij}|\theta_{ij}, \hat{\sigma}^2_{ij} \overset{ind}{\sim} N(\theta_{ij}, \hat{\sigma}^2_{ij}), i = 1, \ldots, m,$$

$$\theta_{ij}|\beta, \sigma^2_{\mu} \overset{ind}{\sim} N(x'_{ij}\beta + v_i, \sigma^2_u), j = 1, \ldots, n_i, \quad (1)$$

$$v_j|\sigma^2_v \overset{ind}{\sim} N(0, \sigma^2_v),$$

where $(\beta, \sigma^2_{\mu}, \sigma^2_v)$ is a set of nuisance parameters. The county-level FSA and RMA planted acreage data are highly correlated. To avoid the multicollinearity problem, we choose to use the maximum of these two data sources. Thus, the vector of regressors for the county j with in the district i consists of $x_{ij} = (1, \max(\text{FSA}_{ij}, \text{RMA}_{ij}))'$.

Note that the above sub-area level model without area level (ASD) effects, v_i, reduces to the basic area-level FH model without constraints, that is,

$$\hat{\theta}_{ij}|\theta_{ij}, \hat{\sigma}^2_{ij} \overset{ind}{\sim} N(\theta_{ij}, \hat{\sigma}^2_{ij}), i = 1, \ldots, m, \quad (2)$$

$$\theta_{ij}|\beta, \sigma^2_{\mu} \overset{ind}{\sim} N(x'_{ij}\beta, \sigma^2_{\mu}), j = 1, \ldots, n_i.$$

A diffuse prior is adopted to the coefficients β, that is, a bivariate normal prior distribution with fixed and known mean and variance and covariance matrix $\beta \sim MN(\hat{\beta}, 1000 \hat{\sum}_{\hat{\beta}})$. Here, $\hat{\beta}$ are the least squares estimates of β obtained from fitting a simple linear regression model of the county-level survey estimates on the auxiliary data x_{ij} and $\hat{\sum}_{\hat{\beta}}$ is the estimated covariance matrix of $\hat{\beta}$. The proper diffused prior on β is used because it provides a degree of computational stability. The prior distributions for σ^2_{μ} and σ^2_v are Uniform $(0,10^{10})$ and Uniform $(0,10^{10})$. The discussion in Browne and Draper (2006) motivates the use of an uniform prior distribution for the random-effect variance components.

3.2. Models With Constraints

Because of the advantage of shrinkage estimation in small area models without constraints discussed in Subsection 3.1, smaller survey estimates are likely to be pulled upwards. This will help to meet the bounds, but it does not solve the problem. As discussed in Subsection 2.2, the county-level estimates must be larger than the corresponding FSA and RMA planted acres data. If the model does not incorporate inequality constraints, the final estimates do not necessarily cover the lower bounds in all cases. The inequality constraints need to be incorporated in the models. In this section, the hierarchical Bayesian models with inequality constraints by Nandram et al. (2022) are discussed.

First, inequality constraints between the true parameter θ_{ij} of interest and administrative values need to be included in the model, that is,

$$\theta_{ij} \geq c_{ij}, i = 1, \ldots, m; j = 1, \ldots, n_i, \tag{3}$$

where the c_{ij} is fixed known quantity.

In our application on planted acres, $c_{ij} = \max(\mathrm{FSA}_{ij}, \mathrm{RMA}_{ij})$ is the maximum value between FSA and RMA corresponding values in the same county. Notice that in Figure 3, some of the survey estimates are one or two standard deviations below their corresponding c_{ij}, thereby creating some difficulties for the model estimates to do the same. The benchmarking constraint creates an additional challenge because the state target may be only slightly larger than the state total from administrative data, $c = \sum_{i=1}^{m}\sum_{j=1}^{n_i} c_{ij}$. This may be a tight condition, as discussed in Cruze et al. (2019).

In addition, under NASS's top-down approach, the benchmarking constraint needs to be considered as well. In this article, we fit Bayesian models using Markov chain Monte Carlo (MCMC) simulation. After model fitting, a series of MCMC samples are obtained to construct the posterior summaries of interest. We perform ratio benchmarking in each iteration of the MCMC samples. Erciulescu et al. (2020) discussed and applied the ratio benchmarking adjustment method at the (MCMC) iteration level in the output analysis to address the county-state benchmarking constraint. It provides a suitable benchmarking adjustment to ensure consistency of county-level estimates with the state target efficiently.

Let $\tilde{\theta}_{ij}^{B}$ be the adjusted model estimate for county j in district i. Let $\theta_{ij,k}$ denote the draw of θ_{ij} and $\theta_{ij,k}^{B}$ denote the adjusted (after benchmarking) draw, where k denotes the draw from the posterior distribution and $k = 1, \ldots, K$. Let a be the benchmarking state target.

The arithmetic mean of the MCMC samples is used to construct the point estimates of interest. After the ratio benchmarking adjustment,

$$\tilde{\theta}_{ij}^{B} = \frac{1}{K}\sum_{k=1}^{K}\theta_{ij,k}^{B} = \frac{1}{K}\sum_{k=1}^{K} r_k\,\theta_{ij,k}, \tag{4}$$

where r_k is the adjusted ratio at iterate level and the ratio r_k is

$$r_k = a \times \left(\sum_{i=1}^{m}\sum_{j=1}^{n_i}\theta_{ij,k}\right)^{-1}. \tag{5}$$

Therefore, the following relationship holds for state benchmarking constraint,

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}\theta_{ij}^{B} = a. \tag{6}$$

However, we need to make sure the adjusted final estimate $\tilde{\theta}_{ij}^{B}$ can satisfy inequality constraint as well. Given Equation (3), the inequality constraint can be preserved for $\theta_{ij,k}$ in each k^{th} iteration. If $r_k \geq 1$ for each k, the following relationship follows from combining

Equations (3) and (4):

$$\tilde{\theta}_{ij}^{B} = \frac{1}{K}\sum_{k=1}^{K} r_k \theta_{ij,k} \geq \frac{1}{K}\sum_{k=1}^{K} \theta_{ij,k} \geq \frac{1}{K}\sum_{k=1}^{K} c_{ij} \geq c_{ij}. \tag{7}$$

Therefore, when $r_k \geq 1$ $\left(\sum_{i=1}^{m}\sum_{j=1}^{n_i}\theta_{ij,k} \leq a\right)$ for each iteration k, it follows that that all model estimates are raked up, and no individual county's inequality constraint will be violated.

Based on the discussion above, θ_{ij} should be drawn subject to the constraints

$$\theta_{ij} \geq c_{ij}, i = 1, \ldots, m; j = 1, \ldots, n_i, \sum_{i=1}^{m}\sum_{j=1}^{n_i}\theta_{ij} \leq a \tag{8}$$

to address both inequality and benchmarking constraints in the models.

According to the constraints Equation (8),

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i} c_{ij} \leq \sum_{i=1}^{m}\sum_{j=1}^{n_i}\theta_{ij} \leq a. \tag{9}$$

Therefore, the support of θ_{ij} given $\underset{\sim}{\theta}_{(ij)}$ is

$$\mathcal{T} = \left\{\theta_{ij} : \max\left(c_{ij,} \sum_{i=1}^{m}\sum_{j=1}^{n_i} c_{ij} - \sum_{i'=1, i'\neq i}^{m}\sum_{j'=1, j'\neq j}^{n_i}\theta_{i'j'}\right) \leq \theta_{ij} \leq a - \sum_{i'=1, i'\neq i}^{m}\sum_{j'=1, j'\neq j}^{n_i}\theta_{i'j'}\right\}, \tag{10}$$

where the lower bound c_{ij} is known and fixed and $i = 1, \ldots, m;\ j = 1, \ldots, n_i$.

To preserve the relationships, the constraint Equation (10) is added to the FH model and the sub-area model in the priors to get the joint posterior density of θ_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n_i$. This problem falls under the general heading of constraint problems in statistics (Nandram et al. 1997).

Therefore, the sub-area hierarchical Bayesian model with constraints is proposed as

$$\begin{aligned} \hat{\theta}_{ij}|\theta_{ij}, \hat{\sigma}_{ij}^2 &\overset{ind}{\sim} N(\theta_{ij}, \hat{\sigma}_{ij}^2), j = 1, \ldots, n_i, \\ \theta_{ij}|\beta, \sigma_{\mu}^2 &\overset{ind}{\sim} N(x_{ij}'\beta + v_i, \sigma_{\mu}^2),\ \theta_{ij} \in \mathcal{T}, \\ v_i|\sigma_v^2 &\overset{ind}{\sim} N(0, \sigma_v^2), i = 1, \ldots, m, \end{aligned} \tag{11}$$

where $\mathcal{T}$ denotes the support in Equation (10) of θ_{ij} such that both the benchmarking constraint and inequality constraints are satisfied. Here, $(\beta, \sigma_{\mu}^2, \sigma_v^2)$ is a set of nuisance parameters and $x_{ij} = (1, x_{ij1}, \ldots, x_{ijp})$ is the vector of covariates and the intercept. In the constrained model, the vector of regressors for the county j with in the district i are the same with those in the unconstrained model, that is, $x_{ij} = (1, \max(\mathrm{FSA}_{ij}, \mathrm{RMA}_{ij}))'$. Note that the above sub-area level model without sub-area level (ASD) effects, v_i, reduces to the area-level FH model with constraints, that is,

$$\hat{\theta}_{ij}|\theta_{ij}, \hat{\sigma}_{ij}^2 \overset{ind}{\sim} N(\theta_{ij}, \hat{\sigma}_{ij}^2), j = 1, \ldots, n_i, \tag{12}$$

$$\theta_{ij}|\beta, \sigma_{\mu}^2 \overset{ind}{\sim} N(x_{ij}'\beta, \sigma_{\mu}^2),\ \theta_{ij} \in \mathcal{T}.$$

A diffuse prior is adopted to the coefficients β, the same as the prior mentioned in Subsection 3.1. The prior distributions in subarea-level model for σ_{μ}^2 and σ_{ν}^2 are Uniform $(0, 10^{10})$ and Uniform $(0, 10^{10})$, respectively and the prior distribution in area level model for σ_{μ}^2 is Uniform $(0, 10^{10})$. Notice that without benchmarking constraint based on ratio benchmarking, the θ_{ij} s are not correlated a priori, but they are correlated a posteriori because of the common parameters over areas. With benchmarking constraint, they are correlated because they must add up to the state target a.

The methodology for creating the state targets guarantees that a state target a is greater than or equal to the administrative state total c. That is, $a = \sum_{i=1}^{m}\sum_{j=1}^{n_i}\tilde{\theta}_{ij}^{B} \geq \sum_{i=1}^{m}\sum_{j=1}^{n_i}c_{ij} = c$. Therefore, there are feasible solutions to the inequality constraint problem in Equation (8), and a feasible solution clearly depends on the state target and the FSA and RMA data. As discussed in Subsection 2.2, most of the survey estimates are within two standard deviations of the bounds, but many of the smaller ones are much further below the bounds. If the model does not incorporate inequality constraints, the final model estimates do not necessarily cover the lower bounds in all cases. Therefore, inequality constraints need to be incorporated in the models to provide not only reliable but also coherent estimates.

4. Case Study

Three states, Illinois (IL), Ohio (OH), and Pennsylvania (PA), are considered in the case study. The four models discussed in Section 3 are compared: the sub-area level model with inequality constraint, the area-level model with inequality constraint, the sub-area level model without inequality constraint and the area-level model without inequality constraint. All models are fit using the administrative data sources described in Subsection 2.2.

All models produce 2014 CAPS estimates of planted acres for corn in IL, OH, and PA. FSA and RMA administrative data in IL usually have very high coverage rates of the planted acres for corn in each county. But in some specific counties in OH, both sets of administrative data have relatively low coverage rates for planted acres. In PA, administrative data in many counties have low coverage rates for planted acreage estimates. Therefore, these three states have different features. The model performance is evaluated for all different scenarios.

As mentioned in Subsection 2.2, the county-level survey estimates did not automatically cover all FSA and RMA administrative data. The relationship between survey estimates and the corresponding lower bounds based on administrative data (the maximum of FSA and RMA data) is displayed in Figure 4. The plotted pairs of survey estimates and administrative data are scattered around the 45 degree line. Around 31% of the county-level survey estimates cover FSA and RMA for IL. About 56% of the survey estimates cover FSA and RMA for OH. About 83% of the survey estimates cover FSA and RMA for PA.

In Subsection 4.1, a summary of the model fitting process is provided. Subsection 4.2 includes the internal checks for all four models. Several diagnostic tools are explored to

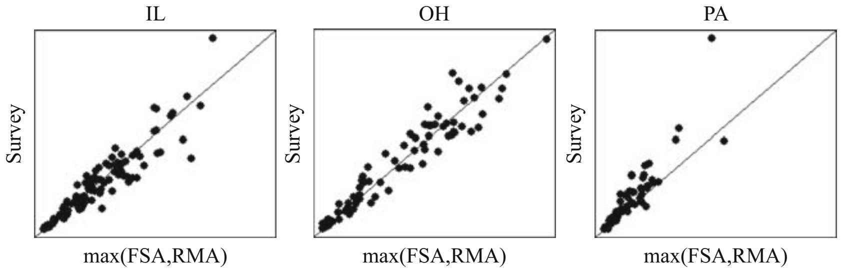

Fig. 4. The county-level planted acreage estimates of corn for CAPS and the lower bounds in IL, OH, and PA.

check the adequacy of the models. External checks between model estimates, survey estimates and official statistics from NASS are presented for all models in Subsection 4.3.

4.1. Model Estimation

All four models are applied to all counties with positive data in one state for which $\left(\hat{\theta}_{ij}, \hat{\sigma}^2_{ij}, x_{ij}\right)$ are available. In IL, there are 102 counties and 9 ASDs in the CAPS samples for planted acreage. In OH, there are 88 counties and 9 ASDs. In PA, there are 65 counties and 9 ASDs.

MCMC simulation method is used to fit all four hierarchical Bayesian models using R and JAGS (Plummer 2003). In each model, three chains are run for our MCMC simulation. Each chain contains 50,000 Monte Carlo samples, and the first 15,000 iterates are discarded as a burn-in to improve the mixing of each chain. After that point, 35,000 further iterations were produced for each of the three chains. In order to eliminate the correlations among neighboring iterations, those iterations are thinned by taking a systematic sample of 1 in every 35 samples. Finally 1,000 MCMC samples in each chain are obtained for constructing the posterior distributions of all the parameters, the nuisance parameters and the parameters for the planted acres.

Markov Chain Monte Carlo (MCMC) methods have been used to approximate the posterior marginals in Bayesian Hierarchical models and are computationally intensive if models are complicated and intractable. Computation time is one key factor when candidate models are evaluated for production especially for crops county estimates project involving multiple commodities for all related counties in US. As mentioned before, all models are fit by MCMC simulation using RJAGS. The computation time in reaching convergence for the different parameters in the unconstrained models is one to two minutes for each state and each commodity depends on the sizes of the data. But the computational time for the constrained models are five to six minutes since two inequality constraints nested with parameters are incorporated in the models. Their posterior distributions are more complicated than the unconstrained models, involving truncated normal distributions. Therefore, it takes more times to fit the constrained models than unconstrained ones. However, the computation time to produce county-level estimates with associated uncertainties is acceptable in current production procedure.

Convergence diagnostics are conducted. The convergence is monitored using trace plots, the multiple potential scale reduction factors ($\hat{R}$ close to 1) and the Geweke test of stationarity for each chain (Gelman and Rubin 1992; Geweke 1992). Also, once the simulated chains have mixed, the effective number of independent simulation draws to monitor simulation accuracy is determined. Effective sample sizes and the $\hat{R}$ are shown in Tables 1 and 2, resulting in good convergence for all four models: area and sub-area models with inequality constraint (C) and without constraint (NC) for IL, OH, and PA. The values of $\hat{R}$ of most coefficient parameters are close to 1. The effective sample sizes of coefficient parameters in sub-area level models are 3,000 and those in area-level models are around 2,000 for IL. The effective sample sizes vary from 1,100 to 3,000 for OH. The effective sample sizes vary from 1,900 to 3,000 for PA.

4.2. Internal Check

Several diagnostic tools are available to check the adequacy of all four models to the observed data considered in this article. First, the fit of the models to the data is assessed using Bayesian predictive checks. If a model fit is adequate to all observations $\hat{\theta}$, replicated values θ_{rep} that generated data from the model would be similar to observations. We

Table 1. Sub-area level models: Effective sample sizes (ESS) and $\hat{R}$ for 2014 IL, OH and PA corn.

State	Parameters	ESS		$\hat{R}$	
		C Sub-area	NC Sub-area	C Sub-area	NC Sub-area
	β_0	3000	3000	1.001	1.001
IL	β_1	3000	3000	1.001	1.002
	σ_μ^2	1900	2100	1.004	1.012
	σ_v^2	3000	3000	1.001	1.003
	β_0	1800	2800	1.006	1.006
OH	β_1	3000	1400	1.006	1.002
	σ_μ^2	2000	1200	1.007	1.007
	σ_v^2	2300	1800	1.003	1.003
	β_0	2400	2800	1.004	1.003
PA	β_1	3000	3000	1.001	1.001
	σ_μ^2	1900	2000	1.019	1.007
	σ_v^2	2500	3000	1.011	1.009

Table 2. Area level models: Effective sample sizes (ESS) and $\hat{R}$ for 2014 IL, OH, and PA corn.

State	Parameters	ESS		$\hat{R}$	
		C Area	NC Area	C Area	NC Area
	β_0	1500	1700	1.010	1.002
IL	β_1	2000	1900	1.002	1.002
	σ_μ^2	2100	2500	1.008	1.004
	β_0	2300	3000	1.007	1.001
OH	β_1	1700	1100	1.009	1.017
	σ_μ^2	1900	1200	1.008	1.007
	β_0	2400	2400	1.009	1.011
PA	β_1	2600	2700	1.011	1.005
	σ_μ^2	2800	3000	1.001	1.003

calculate the Bayesian predictive p-value (BPP) to measure the adequacy of all models to the data from Rubin (1984), Meng (1994), and Gelman et al. (2013). The Bayesian posterior predictive p-value (BPP) is defined as

$$p = Pr\left(T\left(\theta^{rep}, \Omega\right) > T\left(\hat{\theta}, \Omega\right) | \hat{\theta}\right), \tag{13}$$

where discrepancy function, $T(\theta, \Omega)$, is selected as $T(\theta, \Omega) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \frac{\left(\theta_{ij} - E(\theta_{ij}|\hat{\theta})\right)^2}{Var\left(\theta_{ij}|\Omega\right)}$ and Ω are the nuisance parameters in each model. The p-value is the probability of the sum of square residuals based on replicated estimates larger than the one from observed data. If the value is extreme, smaller than 0.05 or larger than 0.95, it indicates a discrepancy between the model and the data, meaning the model is not adequate. The BPP for each model is presented in Table 3. For IL, the BPPs in the area-level and subarea level models with constraints are 0.663 and 0.504, respectively, which are not close to 0 or 1. The models without constraints have high BPP, 0.903 and 0.947, respectively. Those BPP are close to 0.95. Similar results show for OH in Table 5. Noticed that the sub-area level model without constraints for OH is 0.967, which is a borderline case. It indicated that the model is not adequate when comparing with survey estimates. The model's predictions are "biased" to be too high. For PA, all BPPs are smaller than 0.5 but they are not close to 0. However, models cannot be ranked based on BPPs.

Another goodness-of-fit measure for models is the deviance information criterion (DIC) (Spiegelhalter et al. 2002) shown in Table 3. It is not well suited to make the model selection based on DICs between constrained and unconstrained models. In particular, we consider the following type of comparison based on DIC only: between sub-area models and area models within either constrained models and unconstrained models. Table 3 shows the DICs from sub-area models are slightly smaller than those in area-level models. They indicate that the sub-area level models are better than the area-level models because sub-area models can borrow information from both area and sub-area levels.

Therefore, based on DIC diagnostics, sub-area level models are better than the area level models. To check model performance between sub-area level constrained and unconstrained models, external comparisons are discussed in the next section.

4.3. *External Check*

Internal checks show that sub-area level models have slightly smaller DICs than area-level models. Comparisons between area level and sub-area level unconstrained models and

Table 3. DICs and BPPs for constrained and unconstrained models.

Type	Model	DIC		BPP	
		C	NC	C	NC
IL	Sub-area	2334.6	2285.3	0.504	0.947
	Area	2335.7	2285.2	0.633	0.903
OH	Sub-area	1881.1	1766.7	0.331	0.967
	Area	1884.7	1776.4	0.248	0.908
PA	Sub-area	1613.8	1313.7	0.551	0.178
	Area	1618.4	1281.1	0.151	0.111

comparisons between area level and sub-area level constrained models are fine. However, the internal checks considered are not appropriate in terms of the model comparison of both sub-area level constrained and unconstrained models. Therefore, in this section, several external checks are conducted. In the guidance provided in NASEM (2017), the recommendation related the external comparisons is to use published estimates in assessing the quality and reasonableness of the model-based county-level estimates, especially at the research stage. In addition, before NASS can adopt a model-based approach to producing crops county estimates, the model must incorporate all known relationships. The inequality constraints check is another important factor in our evaluation.

First, the inequality check between the final model estimates of planted acres and the corresponding FSA and RMA administrative data is conducted for each model. Figure 5 shows the results in the unconstrained models for IL and OH. Counties in white indicate that the corresponding model estimates are smaller than FSA and RMA data. Counties in gray mean that their estimates are larger than the maximum of both FSA and RMA administrative data. In unconstrained model, 34 out of 102 counties in IL, 8 out of 88 counties in OH, and 3 out of 65 counties in PA do not satisfy with the constraints. However, based on the constrained model setting, all counties in both states satisfy the constraints after ratio benchmarking.

For the unconstrained model, the coverage rate on administrative data depends on the relationship between survey estimates and the administrative data. In PA, when many administrative data are smaller than survey estimates, only a few model-based estimates of planted acres are smaller than the administrative data.

Figure 6 shows comparison of constrained model estimates (denoted C) and unconstrained model estimates (denoted NC) relative to the FSA administrative totals in each county. In each panel, counties have been sorted on the horizontal axis in ascending order by number of CAPS survey reports collected and assigned a corresponding index value, for example, ranging from 1 to 102 for all counties in IL. Within each state, all modeled county estimates are benchmarked to the same fixed state total. On the vertical axis, values greater than one indicate that the estimated county acreage covers the corresponding FSA administrative data. Eliciting an acceptable tolerance below the FSA or RMA acreages has been difficult; at 640 acres (approximately 259 hectares) to the square mile, even some of the apparently modest differences below one become points of

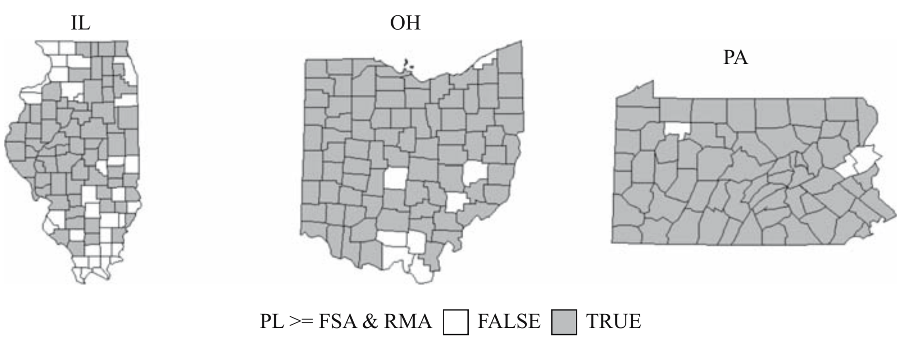

Fig. 5. Inequality check for unconstrained models for IL, OH, and PA.

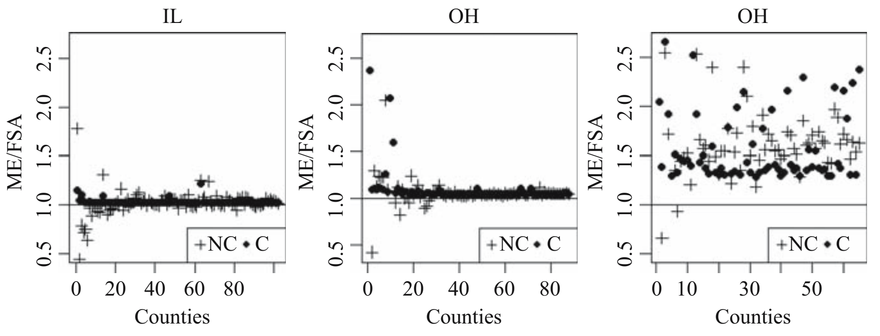

Fig. 6. Ratio of constrained model estimates (C) and unconstrained model estimates (NC) to FSA administrative acreage for IL, OH, and PA.

concern as the unconstrained model estimate begins to disagree with minimum amounts of planting activity on record in FSA's accurately geolocated database. The rightmost panel for PA points to some of the differences by state and commodity in NASS's estimation program. It speaks to the importance of quantifying the uncertainty of estimates when official statistics based on a blend of data may have properties more like administrative data in some scenarios, and more like survey or unconstrained model estimates in others.

In addition, both model-based estimates and survey estimates are compared to the published estimates. Let $\tilde{\theta}^{NC}_{MERB}$ be the unconstrained (NC) model estimates after ratio benchmarking and $\tilde{\theta}^{C}_{MERB}$ be the constrained (C) model estimates after ratio benchmarking. Let $\tilde{\theta}^{DE}_{MERB}$ be the survey (DE) estimates. The absolute relative differences between those estimates and published estimates,

$$\text{ARD} = \frac{|\tilde{\theta}^{t}_{MERB} - \text{Published}|}{\text{Published}}, \tag{14}$$

are calculated and presented, where $t =$ NC, C, DE. A small ARD is one key check on the performance of model-based point estimates. It is true that ARD will not be useful in the current year because the published estimates will not be available. However, it is a good check in a previous year when the official estimates are already decided and published. Note that we are using 2014 data and corresponding official estimates were published. Indeed, ARD is a check on models for future applications in the research stage recommended by NASEM (2017).

The posterior coefficients of variation (CV),

$$\text{CV} = \frac{\text{PSD}^{t}}{\tilde{\theta}^{t}_{MERB}}, \tag{15}$$

are calculated, where $t =$ NC, C, DE and the posterior standard deviation (PSD) is the corresponding posterior standard deviation of $\tilde{\theta}^{t}_{MERB}$, $t =$ NC, C, DE from different models and survey (see Table 4 and 5).

The sample sizes for planted acres in CAPS varies with county in each state. Many counties in these three states have relatively large sample sizes. However, many counties have small sample sizes as well. Small area models tend to improve the accuracy of estimates comparing to the accuracy of survey estimates, especially in areas with small

Table 4. 2014 IL corn planted acres: comparisons of ARDs and CVs among survey, sub-area unconstrained model and constrained model.

Sample size	Statistics	ARD (%)			CV (%)		
		DE	NC	C	DE	NC	C
Overall	Min	0.259	0.007	0.003	10.501	1.899	0.144
	Median	14.914	0.948	0.194	19.210	5.199	0.272
	Max	82.973	51.346	34.908	92.283	125.905	12.705
[0,30)	Min	0.259	0.622	0.273	25.315	20.544	1.466
	Median	16.585	13.530	0.978	42.421	34.905	2.187
	Max	66.174	51.346	34.908	92.283	125.905	12.705
[30,60)	Min	0.575	0.007	0.007	10.501	2.459	0.185
	Median	9.721	1.204	0.176	19.885	5.812	0.278
	Max	39.620	17.036	1.940	33.961	21.985	2.336
≥ 60	Min	7.474	0.096	0.003	9.108	1.899	0.144
	Median	33.990	0.646	0.196	15.731	3.151	0.214
	Max	82.973	2.032	1.199	53.570	5.522	1.740

Table 5. 2014 OH corn planted acres: comparisons of ARDs and CVs among survey, sub-area unconstrained model and constrained model.

Sample size	Statistics	ARD (%)			CV (%)		
		DE	NC	C	DE	NC	C
Overall	Min	0.002	0.103	0.093	8.754	1.043	0.473
	Median	12.942	2.394	2.575	22.292	3.670	0.797
	Max	114.123	95.376	49.858	100.000	104.411	89.816
[0,30)	Min	0.002	0.103	0.671	17.169	3.266	0.533
	Median	24.898	9.791	4.650	35.280	22.292	5.044
	Max	95.687	95.376	49.858	100.000	104.411	89.816
[30,60)	Min	1.574	0.136	0.093	10.224	1.206	0.473
	Median	12.699	2.266	2.191	19.468	2.546	0.660
	Max	114.123	10.968	14.864	33.072	29.994	10.548
≥ 60	Min	6.172	0.322	0.216	8.755	1.043	0.499
	Median	11.982	0.876	1.241	14.699	1.507	0.765
	Max	18.915	5.001	6.785	19.384	5.231	4.136

sample sizes. In order to examine the effect of sample size among our models, we split counties of IL, OH, and PA, respectively, into three groups according to their number of reports in CAPS: small sizes (less than 30); median sizes (between 30 and 60); large sizes (larger than 60). All statistics are shown in Table 4, 5, and 6 as well.

Among all counties in IL, the median ARD value between survey estimates and published estimates in IL is 14.914%. Substantial improvement can be noticed from both the constrained model and the unconstrained model. Again compared to published estimates, the median ARD value based on the constrained model is 194%, less than the median ARD value based on the unconstrained model, 0.948%. Moreover, the range of ARD values from the constrained model (0.003%, 34.908%) are much narrower than the range based on survey estimates (0.259%, 82.973%) and also less than those from the

Table 6. 2014 PA corn planted acres: comparisons of ARDs and CVs among survey, sub-area unconstrained model and constrained model.

Sample size	Statistics	ARD (%)			CV (%)		
		DE	NC	C	DE	NC	C
Overall	Min	0.128	0	0	9.685	2.874	2.874
	Median	12.537	12.272	11.198	22.644	14.237	11.584
	Max	73.300	733.127	33.318	70.941	75.132	44.113
[0,30)	Min	0.128	0	0	19.421	9.057	2.874
	Median	20.000	16.343	14.284	32.611	31.736	14.532
	Max	73.300	733.127	33.318	70.941	75.132	44.113
[30,60)	Min	2.520	1.71	0	14.165	6.778	2.795
	Median	11.200	15.943	10.910	21.143	12.924	11.484
	Max	54.600	41.624	19.247	38.203	27.496	21.941
≥ 60	Min	0.303	7.510	1.513	9.685	4.521	3.294
	Median	10.654	15.933	9.343	12.943	8.943	8.223
	Max	27.382	41.015	26.477	19.257	20.140	17.858

unconstrained model (0.007%, 51.349%). Therefore, for IL, the sub-area level model with constraints performs the best among the unconstrained model and survey estimates as measured by the ARD. In addition, Table 4 shows the ARD values based on the sample sizes of counties in IL. The ranges of ARD values based on both models are large for counties with small number of reports. ARD values from the constrained model are within 2% for median size counties but those from the unconstrained model are from 0.007% to 17.036 %. For large counties, the relative differences from all models are the smallest among all three types of counties. They are within 2% difference for constrained models and 3% from unconstrained model. As expected, all estimates are closer to the published estimates with increasing sample size. Overall, the comparisons of ARD values show that the constrained model increases the accuracy of the estimates significantly.

The CVs of the IL model and survey estimates are shown in Table 4. The sub-area level model can borrow information from both covariates and other counties in the district (sub-area) level. Therefore, the posterior CVs would have a greater reduction compared with the CVs of the survey estimates. The median CVs among all counties in IL are in decreasing order: survey, the unconstrained model and the constrained model. In the unconstrained model, the CVs of small size counties are the largest (20.544%, 125.905%).

The maximum estimated CVs exceeds the maximum of CVs from survey. The CVs of the constrained model are much smaller than those from survey and the unconstrained model. As expected, the CVs are smaller when sample sizes increase. In the model with inequality constraints, the maximum CVs is in the small size counties as well.

Table 5 shows the comparisons for OH. The median of ARDs between survey estimates and published estimates is 12.942%. Substantial improvement can also be noticed from both constrained and unconstrained models. The median ARD value between model-based estimates and the published estimates is around 2%. The smallest median of the relative differences is 2.394% in the unconstrained model. However, the range of ARD values from the constrained model is (0.093%, 49.858%), which is narrower than the one from the unconstrained model, (0.103%, 95.376%). Notice that the ranges of ARDs in OH are larger than those in IL. The administrative data for OH are not stronger comparing with those in IL. In several counties, FSA and RMA administrative data have the undercoverage issue.

To examine the effect of sample sizes, OH is split into three groups and all statistics are presented in Table 5. The ranges of the ARD values based on models and the survey are relatively large in small size counties. Both model estimates are much closer to the published estimates. The estimates of the constrained model in small size counties are closest to the published estimates based on the range of the ARD values. However, the median ARD value from the constrained model is 1.241% for large size counties, which is larger than the one from the unconstrained model, 0.876%. The maximum ARD value is similar as well.

For the median size counties, constrained model tends to provide larger estimates compared with those from unconstrained model.

The CVs are compared among models and the survey estimates for OH as well. Similar to IL, the posterior CVs based on the models are small compared with the CVs from survey. The median CV in the unconstrained model is 3.67%, larger than the one in the constrained model. The maximum CV in the unconstrained model is the highest among models and survey. As expected, the CVs are smaller when sample sizes increase. The maximum of CVs is in small size counties as well. The CVs based on constrained model are much smaller than those of constrained model and survey. For OH, the range of CVs in model with inequality constraints are wider than those for IL.

Table 6 shows the comparisons for PA. The median of ARDs between survey estimates and published estimates is 12.272%. Slight improvement can be noticed from both constrained and unconstrained models in terms of median ARD value but big improvement from the constrained model when comparing with the maximum of ARD value. The median ARD value between model-based estimates and the published estimates is around 12.218% and 11.198% for unconstrained and constrained model respectively. However, the range of ARD values from the unconstrained model is (0%, 733.127%). That biggest ARD value, 733.127%, is in county with small sample sizes, far from the published estimate comparing with survey. the range of ARD values from the unconstrained model is (0%, 33.318%), which is narrower than both from survey and unconstrained model. Notice that the median of ARDs in PA is larger than those in IL and OH.

The administrative data for PA are not stronger comparing with those in IL and OH.

PA is also split into three groups and all statistics are presented in Table 6. The ranges of the ARD values based on models and the survey are relatively large in small size counties. Both model estimates are closer to the published estimates in small size counties. Models have better performance than the survey estimates when sample sizes are small. The estimates of the constrained model in small size counties are closest to the published estimates based on the range of the ARD values. However, the median ARD values from the constrained model in both medium and large size counties are only slightly smaller than the one in the survey. For unconstrained model, those are larger than the median ARD value in survey. As stated before, the administrative data in PA are not strong compared with those in IL and OH. If there was no inequality constraint, the model estimates would be affected by undercoverage from the administrative data when borrowing information from them.

The posterior CVs, based on the models, are smaller than those from the survey for PA. However, the reductions are much smaller than those in IL and OH. The loose lower bounds based on administrative data allow estimates to have more room to move. The median CV in the constrained model is 11.584%, smaller than the survey of 22.264% and the one in the unconstrained model, 14.237%. As expected, the CVs are smaller when sample sizes increase. The maximum of CVs is in small size counties as well.

5. Conclusion

NASS puts extensive research efforts on crops county estimate program aimed primarily to improve the precision of the estimates at county level while preserving the underlying relationships among the estimates and administrative data. Different small area estimation models are implemented to integrate multiple sources of auxiliary information with CAPS data. In this paper, models with inequality constraints are discussed and implemented to address the needs and challenges of the inequality and benchmarking constraints that NASS official statistics need to satisfy. That is, the county-level estimates of planted acreage should be greater than or equal to the corresponding administrative data while the total acreage of all available county-level estimates are equal to the state target.

We apply both sub-area and area-level models with inequality constraints to construct reliable and coherent county-level planted acreage estimates. In the case study of 2014 corn based on IL, OH, and PA, we show model diagnostics and provide internal checks between area-level models and sub-area level models. DICs indicate that the sub-area level models are slightly better than the area-level model. However, the residual-type internal checks are not very suitable for comparing the constrained and unconstrained models since our focus is to provide coherent estimates close to the official estimates but not to the survey estimates. For the model with inequality constraint, one would need to check it against external constraints.

Now more comparisons among both sub-area level model estimates and survey estimates are made. We pick three different states because their administrative data have different coverage rates. The results show that the performances of the constrained model are different among these three states. When many survey estimates are larger than the administrative data as shown in PA, the improvements are not that significant when comparing with IL and OH. However, the constrained model is still better than the survey and unconstrained model in terms of the external check. Inequality checks show that constrained model can preserve the relationships among estimates and administrative data. But this is not necessarily the case for the unconstrained model. It is true that including inequality constraints in some areas is unnecessary. But if we relax the inequality constraints for those counties that meet the constraint, they may not be satisfied in the model estimates. Not putting the constraint on the areas that are much higher than the lower bound is incoherent from a Bayesian view. Therefore, we have to put the constraints on all counties.

In addition, the statistics of ARD values show that the constrained model provides estimates closer to the published values than those from the unconstrained model as well as those from the survey, especially for IL. FSA and RMA are very significant covariates for the estimates of planted acres. Moreover, the associated measures of uncertainty (CVs) from models are significantly smaller than the CVs of the survey estimates. The basic sub-area models can reduce the CVs while borrowing strength from auxiliary information and all counties in one district and all districts in one state. In addition, for the constrained model, the prior information based on the lower bound information from FSA and RMA data and the upper bound related to the state target reduce the CVs of the model-based estimates since estimates can be drawn only in the restricted support. Therefore, the performance of the sub-area level model with inequality constraints illustrates significant improvement of county-level estimates of planted acres in accuracy and precision.

Major ongoing and future research related to sub-area level constrained model involves the investigation of different auxiliary information. The auxiliary information considered here is the key data sources of planted acres (the combination of FSA and RMA administrative data). Future efforts will be on searching and applying other useful data sources to strengthen the model. Remote sensing data, NASS cropland data layer (see, Boryan et al., 2011), and weekly weather data are available at the county level. Variable selections should be investigated for different states and commodities because weather conditions influence the planting progress and the planted acres within different time periods based on different states and commodities.

In addition, missing data problems are another challenge for the application of the constrained model. In this article, case studies related to IL, OH, and PA, which do not have missing data in 2014 corn, are provided. However, it is not always the case for other states or other commodities. As mentioned in Subsection 2.1, CAPS is conducted for different commodities among all eligible states. In some cases, the survey may not indicate any planted area with respect to a particular commodity, but administrative data might represent some positive acres or vice versa. Erciulescu et al. (2020) used the nearest neighbor methods to impute missing data for either survey or covariates. This approach of imputing and borrowing information from previous year or the average of several years estimates are being explored. How to deal with missing data and provide reliable and coherent predictions are ongoing research.

6. References

Battese, G.E., R.M. Harter, and W.A. Fuller. 1988. "An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data." *Journal of the American Statistical Association* 83(401): 28–36. DOI: https://doi.org/10.2307/2288915.

Boryan, C., Z. Yang, R. Mueller, and M. Craig. 2011. "Monitoring US agriculture: the US Depart ment of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program." *Geocarto International* 26(5): 341–358. DOI: https://doi.org/10.1080/10106049.2011.562309.

Browne, W.J., and D. Draper. 2006. "A Comparison of Bayesian and Likelihood-Based Methods for Fitting Multilevel Models." *Bayesian Analysis* 1(3): 473–514. DOI: https://doi.org/10.1214/06-BA117.

Cruze, N.B., A.L. Erciulescu, B. Nandram, W.J. Barboza, and L.J. Young. 2019. "Producing Official County-Level Agricultural Estimates in the United States: Needs and Challenges." *Statistical Science* 34(2): 301–316. DOI: https://doi.org/10.1214/18-STS687.

Erciulescu, A.L., N.B. Cruze, and B. Nandram. 2019. "Model-Based County Level Crop Estimates Incorporating Auxiliary Sources of Information." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 182(1): 283–303. DOI: https://doi.org/10.1111/rssa.12390.

Erciulescu, A.L., N.B. Cruze, and B. Nandram. 2020. "Statistical Challenges in Combining Survey and Auxiliary Data to Produce Official Statistics." *Journal of Official Statistics* 36(1): 63–88. DOI: https://doi.org/10.2478/jos-2020-0004.

Fay, R.E., and R.A. Herriot. 1979. "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data." *Journal of the American Statistical Association* 74(366a): 269–277. DOI: https://doi.org/10.2307/2286322.

Fuller, W.A., and J.J. Goyeneche. 1998. "Estimation of The State Variance Component." Unpublished manuscript.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 2013. *Bayesian Data Analysis.* CRC press. DOI: https://doi.org/10.1201/b16018.

Gelman, A., and D.B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical science* 7(4): 457–472. DOI: https://doi.org/10.1214/ss/1177011136.

Geweke, J. 1992. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In *Bayesian Statistics* 4: 169–193. DOI: https://doi.org/10.21034/sr.148.

Good, D. 2014. "Comparing NASS and FSA Planted Acreage Data." *Farmdoc Daily* 4(9). Available at: https://farmdocdaily.illinois.edu/2014/01/comparing-nass-fsa-planted-acreage-data.html (accessed January 2014).

Meng, X.-L. 1994. "Posterior Predictive *p*-Values." *The Annals of Statistics* 22(3): 1142–1160. DOI: https://doi.org/10.1214/aos/1176325622.

Nandram, B., N.B. Cruze, and A.L. Erciulescu. 2022. Bayesian Small Area Models under Inequality Constraints with Benchmarking and Double Shrinkage. *Research Report RDD-22-02, National Agricultural Statistics Service*, USDA. Available at: https://www.nass.usda.gov/Education_and_Outreach/Reports,_Presentations_and_Conferences/reports/ResearchReport_constraintmodel.pdf /(ACCESSEDE jYKY 2022).

Nandram, B., J. Sedransk, and S.J. Smith. 1997. "Order-Restricted Bayesian Estimation of The Age Composition of A Population of Atlantic Cod." *Journal of the American Statistical Association* 92(437): 33–40. DOI: https://doi.org/10.2307/2291447.

NASEM (National Academies of Sciences, Engineering, and Medicine). 2017. *Improving Crop Estimates by Integrating Multiple Data Sources.* National Academies Press. DOI: https://doi.org/10.17226/24892.

Pfeffermann, D. 2013. "New Important Developments in Small Area Estimation." *Statistical Science* 28(1): 40–68. DOI: https://doi.org/10.1214/12-STS395.

Plummer, M. 2003. "JAGS: A Program for Analysis of Bayesian Graphical Models using Gibbs Sampling." 3rd International Workshop on Distributed Statistical Computing (DSC 2003); March 21, Vienna, Austria: 124. Avaiable at: https://www.r-project.org/conferences/ DSC-2003/Proceedings/Plummer.pdf.

Rao, J.N.K., and I. Molina. 2015. *Small Area Estimation.* 2015 John Wiley/Sons, Inc. DOI: https://doi.org/10.1002/9781118735855.

Rubin, D.B. 1984. "Bayesianly Justifiable and Relevant Frequency Calculations for the Applies Statistician." *The Annals of Statistics*: 1151–1172. DOI: https://doi.org/10.1214/aos/1176346 785.

Sen, D., S. Patra, and D. Dunson. 2018. "Constrained Inference Through Posterior Projections". arXiv preprint. DOI: https://doi.org/10.48550/arxiv.1812.05741.

Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. van der Linde. 2002. "Bayesian Measures of Model Complexity and Fit." *Journal of the Royal Statistical Society: Series B* 64(4): 583–639. DOI: https://doi.org/10.1111/1467-9868.00353.

Torabi, M., and J.N.K. Rao. 2014. "On Small Area Estimation under A Sub-Area Level Model." *Journal of Multivariate Analysis* 127: 36–55. DOI: https://doi.org/10.1016/j.jmva.2014.02.

USDA FSA (U.S. Department of Agriculture Farm Service Agency). 2018. *FSA Handbook: Acreage and Compliance Determinations.* 2-CP, Revision 16. Available at: https://www.fsa.usda.gov/Internet/FSA_File/2cp16-a1.pdf (accessed April 2018).

USDA RMA (U.S. Department of Agriculture Risk Management Agency). 2017 The Risk Management Safety Net: Market Penetration and Market Potential-Analysis of the Federal Crop Insurance Portfolio. Available at: https://www.rma.usda.gov/-/media/RMA/Publications/Market-Penetration-and-Market-Potential-2017.ashx?la = en (accessed September 2017).

Received January 2021
Revised May 2021
Accepted January 2022