Supplementary material for "Small Area Estimation using EBLUPs under the Nested Error Regression Model"

June 23, 2023

The clp-PR and sam-RM prediction intervals

Let $\mathbf{X}_{(s)} = [\mathbf{X}_{(s)1}^T, \ldots, \mathbf{X}_{(s)g}^T]^T$, where $\mathbf{X}_{(s)i} = [\mathbf{x}_{ij}, j \in s_i]^T$, and $\mathbf{y}_{(s)} = [\mathbf{y}_{(s)1}^T, \ldots, \mathbf{y}_{(s)g}^T]^T$, where $\mathbf{y}_{(s)i} = [y_{ij}, j \in s_i]^T$, $\mathbf{Z}_{(s)} = \text{block diag}[\mathbf{1}_{s_1}, \ldots \mathbf{1}_{s_g}]$ with $\mathbf{1}_{n_i}$ the n_i-vector of ones, and $\mathbf{V}_{(s)} = \text{Var}(\mathbf{y}_{(s)}) = \text{block diag}[\dot{\sigma}_e^2 \mathbf{I}_{n_1} + \dot{\sigma}_\alpha^2 \mathbf{1}_{n_1} \mathbf{1}_{n_1}^T, \ldots, \dot{\sigma}_e^2 \mathbf{I}_{n_g} + \dot{\sigma}_\alpha^2 \mathbf{1}_{n_g} \mathbf{1}_{n_g}^T]$, with $\mathbf{I}_{n_i}$ the $n_i \times n_i$ identity matrix. Following Prasad & Rao (1990), Datta & Lahiri (2000) showed that $\text{E}(\hat{M}_i^{\text{clp}} - \dot{\eta}_i)^2$, the unconditional MSE of $\hat{M}_i^{\text{clp}}$ for predicting $\dot{\eta}_i$, can be approximated to second order by

$$\text{MSE}_{\text{PR},i} = g_{1i}(\boldsymbol{\theta}) + g_{2i}(\boldsymbol{\theta}) + g_{3i}(\boldsymbol{\theta}) + o_p(g^{-1}).$$

where

$$\begin{aligned}
g_{1i}(\boldsymbol{\theta}) &= (1-\gamma_i)\sigma_\alpha^2, \\
g_{2i}(\boldsymbol{\theta}) &= \left(\bar{\mathbf{x}}_i - \gamma_i \bar{\mathbf{x}}_{i(s)}\right)^T \left(\mathbf{X}_{(s)}^T \mathbf{V}_{(s)}^{-1} \mathbf{X}_{(s)}\right)^{-1} \left(\bar{\mathbf{x}}_i - \gamma_i \bar{\mathbf{x}}_{i(s)}\right) \\
g_{3i}(\boldsymbol{\theta}) &= (1-\gamma_i)^2 \gamma_i \sigma_e^{-4} \sigma_\alpha^{-2} M(\boldsymbol{\theta}),
\end{aligned}$$

with $M(\boldsymbol{\theta}) = \sigma_\alpha^4 \text{Var}(\hat{\sigma}_e^2) + \sigma_e^4 \text{Var}(\hat{\sigma}_\alpha^2) - 2\sigma_\alpha^2 \sigma_e^2 \text{Cov}(\hat{\sigma}_\alpha^2, \hat{\sigma}_e^2)$. The asymptotic variances and covariance of the REML estimators $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_e^2$ needed in $M(\boldsymbol{\theta})$ (i.e. under fixed or bounded small area size asymptotics) are given by

$$\mathrm{Var}(\hat{\sigma}_\alpha^2) = \frac{2}{a}\sum_{i=1}^{g}\left\{\frac{s_i - 1}{\sigma_e^4} + \frac{1}{(\sigma_e^2 + s_i\sigma_\alpha^2)^2}\right\},$$
$$\mathrm{Var}(\hat{\sigma}_e^2) = \frac{2}{a}\sum_{i=1}^{g}\frac{s_i^2}{(\sigma_e^2 + s_i\sigma_\alpha^2)^2},$$
$$\mathrm{Cov}(\hat{\sigma}_\alpha^2, \hat{\sigma}_e^2) = -\frac{2}{a}\sum_{i=1}^{g}\frac{s_i}{(\sigma_e^2 + s_i\sigma_\alpha^2)^2},$$

where

$$a = \left\{\sum_{i=1}^{g}\frac{s_i^2}{(\sigma_e^2 + s_i\sigma_\alpha^2)}\right\}\left[\sum_{i=1}^{g}\left\{\frac{s_i - 1}{\sigma_e^4}\frac{1}{(\sigma_e^2 + s_i\sigma_\alpha^2)}\right\}\right] - \sum_{i=1}^{g}\frac{s_i}{(\sigma_e^2 + s_i\sigma_\alpha^2)^2}.$$

Rao & Molina(2015) showed that $\mathrm{E}(\hat{M}_i^{\mathrm{sam}} - \bar{y}_i)^2$, the unconditional MSE of $\hat{M}_i^{\mathrm{sam}}$ for prediction $\bar{y}_i$ can be approximated to second order by

$$\mathrm{MSE}_{\mathrm{RM},i} = k_i^2\{g_{1i}(\boldsymbol{\theta}) + \tilde{g}_{2i}(\boldsymbol{\theta}) + g_{3i}(\boldsymbol{\theta})\} + (N_i - n_i)\sigma_e^2/N_i^2,$$

where $\tilde{g}_{2i}(\boldsymbol{\theta})$ is obtained from $g_{2i}(\boldsymbol{\theta})$ by replacing $\bar{\mathbf{x}}_i$ with $\bar{\mathbf{x}}_{i(r)}$.

Using the second-order approximations, unbiased estimators of $\mathrm{MSE}_{\mathrm{PR}}$ (Datta & Lahiri 2000) and of $\mathrm{MSE}_{\mathrm{RM}}$ (Rao & Molina 2015) are respectively given by

$$\widehat{\mathrm{MSE}}_{\mathrm{PR},i} = g_{1i}(\hat{\boldsymbol{\theta}}) + g_{2i}(\hat{\boldsymbol{\theta}}) + 2g_{3i}(\hat{\boldsymbol{\theta}}).$$

and

$$\mathrm{MSE}_{\mathrm{RM},i} = k_i^2\{g_{1i}(\hat{\boldsymbol{\theta}}) + \tilde{g}_{2i}(\hat{\boldsymbol{\theta}}) + 2g_{3i}(\hat{\boldsymbol{\theta}})\} + (N_i - n_i)\hat{\sigma}_e^2/N_i^2.$$

Based on the above approximations, $100(1-\epsilon)\%$ prediction intervals for $\dot{\eta}_i$ and $\bar{y}_i$ are

$$[\hat{M}_i^{\mathrm{clp}} - \Phi^{-1}(1-\varepsilon/2)\widehat{\mathrm{MSE}}_{\mathrm{PR},i}^{1/2}, \hat{M}_i^{\mathrm{clp}} + \Phi^{-1}(1-\varepsilon/2)\widehat{\mathrm{MSE}}_{\mathrm{PR},i}^{1/2}]$$

and

$$[\hat{M}_i^{\text{sam}} - \Phi^{-1}(1-\varepsilon/2)\widehat{\text{MSE}}_{\text{RM},i}^{1/2},\ \hat{M}_i^{\text{sam}} + \Phi^{-1}(1-\varepsilon/2)\widehat{\text{MSE}}_{\text{RM},i}^{1/2}],$$

respectively.

A simulation result for the Chatterjee et al (2008) prediction interval

Table S1: Simulated coverage and length of prediction intervals when α_i has a mixture distribution and e_{ij} has a normal distribution with variances $\dot{\sigma}_\alpha^2 = 64$ and $\dot{\sigma}_e^2 = 100$, respectively.

Method			sam-LW		clp-LW		clp-PR		clp-Chatterjee	
Area	N_i	n_i	Cvge	Rlen	Cvge	Rlen	Cvge	Rlen	Cvge	Rlen
1	40	20	0.956	0.03	0.948	0.03	0.974	0.23	0.739	1.65
2	86	43	0.945	0.03	0.931	0.06	0.982	0.26	0.846	2.75
3	166	42	0.962	0.06	0.966	0.05	0.972	0.15	0.765	2.38
4	105	26	0.965	0.07	0.956	0.06	0.966	0.12	0.701	1.67
5	181	45	0.967	0.03	0.966	0.03	0.977	0.13	0.767	2.42
6	190	48	0.951	0.03	0.954	0.02	0.966	0.13	0.772	2.5
7	47	20	0.95	0.03	0.947	0.01	0.966	0.17	0.724	1.52
8	124	31	0.965	0.05	0.954	0.03	0.966	0.1	0.716	1.84
9	183	46	0.953	0.04	0.954	0.03	0.973	0.14	0.782	2.47
10	128	32	0.963	0.05	0.957	0.05	0.971	0.13	0.753	1.94
11	113	28	0.961	0.04	0.956	0.02	0.962	0.09	0.703	1.69
12	193	48	0.96	0.05	0.953	0.04	0.974	0.15	0.767	2.56
13	113	28	0.966	0.08	0.963	0.06	0.966	0.13	0.715	1.78
14	148	37	0.966	0.09	0.964	0.08	0.973	0.17	0.788	2.25
15	132	33	0.954	0.02	0.948	0.01	0.97	0.08	0.721	1.87

Additional results for the consumer expenditure population

Table S2: Simulated relative design-bias and design RMSEs of the point estimators sam and clp, together with the design-averages of the LW, RM and PR estimators of their RMSEs for the consumer expenditure on fresh milk products in each state in 2002. The states are in the same order as in Table 3. * identifies states in Group 3 and † identifies states in Group 2.

STATE	ARB-sam	ARB-clp	AVE-LW	AVE-RM	AVE-PR	RMSE-sam-T	RMSE-clp-T
16	-0.0167	-0.0332	0.4412	0.3701	0.3763	0.2755	0.2250
$50^\dagger$	-0.0533	-0.1113	0.4486	0.3733	0.3746	0.5242	0.5983
31	0.0450	0.0867	0.4790	0.3858	0.3746	0.2973	0.3749
$22^\dagger$	0.0799	0.1501	0.4840	0.3876	0.3745	0.3085	0.5257
21	0.0261	0.0497	0.4888	0.3895	0.3754	0.2457	0.2650
15	-0.0139	-0.0240	0.4933	0.3910	0.3756	0.2740	0.2237
$32^\dagger$	-0.0768	-0.1400	0.4976	0.3922	0.3750	0.5596	0.7582
$37^\dagger$	0.0847	0.1472	0.5016	0.3932	0.3745	0.3679	0.5272
1	0.0025	0.0041	0.4104	0.3430	0.3618	0.1645	0.1375
45	0.0007	-0.0032	0.4143	0.3457	0.3623	0.3675	0.2918
2^*	-0.0725	-0.1447	0.3886	0.3282	0.3552	0.5809	0.8875
9^*	0.1437	0.2606	0.3788	0.3239	0.3543	0.5889	0.9045
41	-0.0082	-0.0185	0.3728	0.3168	0.3487	0.2795	0.2061
49	0.0020	0.0021	0.3562	0.3064	0.3454	0.2485	0.2101
18	0.0210	0.0403	0.3463	0.2993	0.3413	0.3341	0.2771
27^*	-0.0566	-0.1176	0.3395	0.2944	0.3379	0.5262	0.7639
$8^\dagger$	-0.0333	-0.0663	0.3268	0.2854	0.3323	0.4196	0.4474
13	0.0072	0.0114	0.3102	0.2733	0.3237	0.2245	0.1691
$24^\dagger$	0.0520	0.1032	0.3053	0.2705	0.3237	0.3205	0.4699
29	0.0145	0.0307	0.2990	0.2653	0.3191	0.1706	0.1718
53	-0.0090	-0.0213	0.2945	0.2622	0.3176	0.1861	0.1671
55	0.0647	0.0868	0.4673	0.3542	0.3526	0.3316	0.3692
51	-0.0024	-0.0033	0.4692	0.3549	0.3532	0.2490	0.2035
25	-0.0063	-0.0094	0.4519	0.3469	0.3487	0.2044	0.1659
4	-0.0097	-0.0141	0.4468	0.3442	0.3473	0.2772	0.2411
26	0.0318	0.0442	0.4327	0.3371	0.3429	0.2713	0.2599
34	0.0141	0.0187	0.4053	0.3230	0.3340	0.2577	0.2243
17	0.0271	0.0370	0.4057	0.3229	0.3337	0.2180	0.2137
39	0.0113	0.0147	0.3397	0.2868	0.3073	0.2111	0.1869
42	0.0455	0.0572	0.3181	0.2741	0.2969	0.2882	0.2954
36	0.0002	-0.0003	0.3064	0.2668	0.2909	0.2584	0.2304
12	-0.0286	-0.0389	0.3040	0.2653	0.2897	0.2679	0.2762
48	0.0234	0.0305	0.2942	0.2591	0.2840	0.2436	0.2388
6	-0.0184	-0.0254	0.2577	0.2351	0.2618	0.2309	0.2338

Prediction intervals treating the random effects as fixed

To treat the random effects as fixed, we rewrite the model (2.1) and (2.2) as a regression model

$$y_{ij} = \mathbf{z}_{ij}^{*T}\boldsymbol{\chi} + e_{ij}, \qquad \text{for } j = 1,\dots,N_i,\ i = 1,\dots,g,$$

where $\mathbf{z}_{ij}^{*} = [\mathbf{u}_i^T, \mathbf{x}_{ij}^{(w)T}, \mathbf{v}_i^T]^T$ with $\mathbf{v}_i$ a g-vector of zeros with a one in position i, and $\boldsymbol{\chi} = [\boldsymbol{\xi}^T, \boldsymbol{\beta}_2^T, \alpha_1, \dots, \alpha_g]^T$ with $\sum_{i=1}^{g}\alpha_i = 0$. The normal maximum likelihood estimator $\hat{\boldsymbol{\chi}}$ of $\boldsymbol{\chi}$ is the (constrained) ordinary least squares estimator which can be computed using `lm` in R with the sum to zero constraint on the α_i. The optimal model-based predictor of $\bar{y}_i$ is the composite estimator

$$\bar{y}_i^{\text{com-fixed}} = (1-k_i)\,\bar{y}_{i(s)} + k_i\bar{\mathbf{z}}_{i(r)}^{*T}\hat{\boldsymbol{\chi}},$$

where $\bar{\mathbf{z}}_{i(r)}^{*} = [\mathbf{u}_i^T, \bar{\mathbf{x}}_{i(r)}^{(w)T}, \boldsymbol{v}_i^T]^T$ and an approximate $100(1-\epsilon)\%$ prediction interval for $\bar{y}_i$ is

$$\left[\bar{y}_i^{\text{com-fixed}} - \Phi^{-1}(1-\epsilon/2)\,k_i\left\{\frac{\hat{\sigma}_e^2}{N_i - n_i} + \bar{\mathbf{z}}_{i(r)}^{*T}\hat{\mathbf{V}}_{\boldsymbol{\chi}}\bar{\mathbf{z}}_{i(r)}^{*}\right\}^{1/2},\ \bar{y}_i^{\text{com-fixed}} + \Phi^{-1}(1-\epsilon/2)\,k_i\left\{\frac{\hat{\sigma}_e^2}{N_i - n_i} + \bar{\mathbf{z}}_{i(r)}^{*T}\hat{\mathbf{V}}_{\boldsymbol{\chi}}\bar{\mathbf{z}}_{i(r)}^{*}\right\}^{1/2}\right],$$

where $\hat{\sigma}_e^2$ estimates σ_e^2, and $\hat{\mathbf{V}}_{\boldsymbol{\chi}}$ estimates the variance of $\hat{\boldsymbol{\chi}}$. Similarly, a synthetic estimator of $\bar{y}_i$ is $\bar{y}_i^{\text{syn-fixed}} = \bar{\mathbf{z}}_i^{*T}\boldsymbol{\chi}$, where $\bar{\mathbf{z}}_i^{*} = [\mathbf{u}_i^T, \bar{\mathbf{x}}_i^{(w)T}, \boldsymbol{v}_i^T]^T$, and a second approximate $100(1-\epsilon)\%$ prediction interval for $\bar{y}_i$ is

$$\left[\bar{y}_i^{\text{syn-fixed}} - \Phi^{-1}(1-\epsilon/2)(\bar{\mathbf{z}}_i^{*T}\hat{\mathbf{V}}_{\boldsymbol{\chi}}\bar{\mathbf{z}}_i)^{1/2},\ \bar{y}_i^{\text{syn-fixed}} + \Phi^{-1}(1-\epsilon/2)(\bar{\mathbf{z}}_i^{*T}\hat{\mathbf{V}}_{\boldsymbol{\chi}}\bar{\mathbf{z}}_i)^{1/2}\right].$$

Design-based simulation

Tables S3 and S4 show the empirical design-coverage and the design-average relative length of the intervals for the settings with variances $\dot{\sigma}_\alpha^2 = 4$ and $\dot{\sigma}_e^2 = 100$ when e_{ij} has a normal distribution and α_i has either a normal distribution or a mixture distribution; settings with non-normal e_{ij} are included in the supplementary material. The areas are again presented and labeled in order of increasing size. The Monte Carlo standard errors for the design-coverage probabilities are approximately less than 0.01.

In all settings, the model holds and the within area variances are constant. Nonetheless, when σ_e^2/σ_a^2 is large (as in Tables S3 and S4), the kind of results we saw in Table 3 occur. In Table S3, there are 5 Group 3 areas, numbers 6 ($\hat{\alpha}_6/\hat{\sigma}_\alpha = -1.278,\ n_6 = 20$), 7 ($\hat{\alpha}_7/\hat{\sigma}_\alpha = -1.115,\ n_7 = 27$), 21 ($\hat{\alpha}_{21}/\hat{\sigma}_\alpha = -1.359,\ n_{21} = 41$), 24 ($\hat{\alpha}_{24}/\hat{\sigma}_\alpha = 0.991,\ n_{24} = 46$) and 28 ($\hat{\alpha}_{28}/\hat{\sigma}_\alpha = 1.195,\ n_{28} = 52$), and 6 Group 2 areas, numbers 2 ($\hat{\alpha}_2/\hat{\sigma}_\alpha = -0.290,\ n_2 = 5$), 4 ($\hat{\alpha}_4/\hat{\sigma}_\alpha = 0.860,\ n_4 = 17$), 10 ($\hat{\alpha}_{10}/\hat{\sigma}_\alpha = 0.948,\ n_{10} = 32$), 15 ($\hat{\alpha}_{15}/\hat{\sigma}_\alpha = 0.797,\ n_{15} = 36$), 22 ($\hat{\alpha}_{22}/\hat{\sigma}_\alpha = -0.837,\ n_{22} = 42$) and 25 ($\hat{\alpha}_{25}/\hat{\sigma}_\alpha = 0.810,\ n_{25} = 48$). In Table S4, there are 4 Group 3 areas, numbers 11 ($\hat{\alpha}_{11}/\hat{\sigma}_\alpha = 1.073,\ n_{11} = 31$), 16 ($\hat{\alpha}_{16}/\hat{\sigma}_\alpha = -0.807,\ n_{16} = 34$), 25 ($\hat{\alpha}_{25}/\hat{\sigma}_\alpha = -1.941,\ n_{25} = 44$) and 28 ($\hat{\alpha}_{28}/\hat{\sigma}_\alpha = 1.257,\ n_{28} = 53$), and 2 Group 2 areas, numbers 2 ($\hat{\alpha}_2/\hat{\sigma}_\alpha = -0.613,\ n_2 = 7$) and 10 ($\hat{\alpha}_{10}/\hat{\sigma}_\alpha = 0.789,\ n_{10} = 31$). On the other hand, when $\dot{\sigma}_e^2/\dot{\sigma}_\alpha^2$ is not large, there are no Group 2 or 3 areas. Overall, we see that when $\dot{\sigma}_e^2/\dot{\sigma}_\alpha^2$ is large, areas with extreme EBLUPs and small to moderate sample sizes are more difficult than other areas to estimate well in the design-based framework.

Table S3: Simulated design-coverage and design-average length of confidence intervals when α_i and e_{ij} have normal distributions with variances $\dot{\sigma}_\alpha^2 = 4$ and $\dot{\sigma}_e^2 = 100$, respectively. * identifies states in Group 3 and † identifies states in Group 2.

Method				Direct Estimator		sam-LW		sam-RM		clp-LW		clp-PR	
Area	N_i	n_i	$\hat{\alpha}_i/\hat{\sigma}_\alpha^2$	Cvge	Alen	Cvge	Alen	Cvge	Alen	Cvge	Alen	Cvge	Alen
1	123	4	0.079	0.999	7.418	1.000	4.915	0.998	2.243	1.000	4.915	0.997	2.127
$2^\dagger$	147	5	-0.290	0.994	7.151	1.000	4.393	0.941	2.162	1.000	4.393	0.920	2.072
3	114	7	0.079	1.000	5.468	1.000	3.660	0.997	2.090	1.000	3.660	1.000	2.004
$4^\dagger$	40	17	0.860	0.994	2.322	0.964	1.838	0.898	1.554	0.926	1.838	0.828	1.717
5	49	17	-0.073	1.000	3.166	0.997	1.959	0.989	1.610	1.000	1.959	1.000	1.717
6^*	46	20	-1.278	0.981	2.192	0.916	1.680	0.821	1.450	0.613	1.680	0.586	1.651
7^*	54	27	-1.115	0.995	2.121	0.937	1.360	0.903	1.227	0.798	1.360	0.860	1.521
8	57	28	0.637	1.000	2.085	0.987	1.347	0.967	1.217	0.985	1.347	0.991	1.505
9	107	32	-0.193	1.000	2.144	0.987	1.479	0.976	1.297	0.995	1.479	0.996	1.445
$10^\dagger$	108	32	0.948	0.998	2.272	0.966	1.482	0.919	1.298	0.952	1.482	0.925	1.445
11	111	33	-0.426	0.999	2.224	0.985	1.458	0.963	1.283	0.988	1.458	0.988	1.432
12	115	34	-0.013	1.000	2.202	0.994	1.439	0.977	1.269	0.998	1.439	0.999	1.418
13	112	34	0.504	0.998	2.170	0.974	1.430	0.950	1.264	0.977	1.430	0.983	1.418
14	113	34	-0.229	1.000	2.223	0.992	1.433	0.985	1.265	0.994	1.433	0.994	1.417
$15^\dagger$	120	36	0.797	0.998	2.094	0.955	1.394	0.923	1.239	0.953	1.394	0.939	1.392
16	122	37	0.015	1.000	2.061	0.992	1.371	0.984	1.222	0.997	1.371	0.998	1.378
17	74	37	0.060	0.999	1.812	0.970	1.162	0.953	1.073	0.991	1.162	0.998	1.378
18	128	38	-0.402	0.998	1.994	0.984	1.360	0.967	1.213	0.990	1.360	0.994	1.365
19	81	40	-0.174	1.000	1.555	0.995	1.124	0.988	1.042	1.000	1.124	1.000	1.342
20	136	41	-0.096	0.999	1.892	0.981	1.305	0.968	1.174	0.995	1.305	0.997	1.331
21^*	82	41	-1.359	0.998	1.667	0.920	1.104	0.893	1.026	0.748	1.104	0.864	1.331
$22^\dagger$	141	42	-0.837	1.000	2.299	0.957	1.292	0.933	1.165	0.956	1.292	0.956	1.320
23	153	46	-0.379	1.000	1.958	0.979	1.232	0.966	1.121	0.984	1.232	0.988	1.279
24^*	152	46	0.991	0.999	2.000	0.942	1.231	0.894	1.119	0.915	1.231	0.919	1.277
$25^\dagger$	97	48	0.810	0.999	1.579	0.962	1.025	0.949	0.962	0.943	1.025	0.984	1.258
26	171	51	0.156	0.999	1.819	0.984	1.172	0.974	1.075	0.991	1.172	0.993	1.230
27	172	52	-0.453	1.000	1.720	0.980	1.158	0.967	1.064	0.990	1.158	0.992	1.221
28^*	174	52	1.195	0.996	1.899	0.931	1.161	0.898	1.066	0.895	1.161	0.910	1.221
29	183	55	-0.208	0.998	1.827	0.980	1.127	0.965	1.041	0.987	1.127	0.987	1.197
30	185	56	0.392	1.000	1.796	0.974	1.115	0.964	1.033	0.978	1.115	0.990	1.191

Corn data

We examine the corn data presented in Battese et al (1988). This data set comprises 37 observations on the area of corn (hectares) per segment spread across 12 counties, as shown in Table S5. The objective was to predict the average area of corn (HECTARECORN) per segment in these counties. Satellite data on the number of pixels representing corn (PIXELCORN) and the number of pixels representing soybeans (PIXELSOYBEAN) in each segment were used as auxiliary variables. Additionally, we had access to the area population means for the number of pixels corresponding to corn and soybeans in each county.

We pre-processed the auxiliary variables by centering them around their respective population means means, distinguishing these new variables by appending *cent* to the variable names. Further, we incorporated the population means, distinguished by appending *avg* to the variable names, as variables distinguishing between counties. This provided us with $p_b = 2$ between county variables and $p_w = 2$ within county variables.

We then fitted the model (0.1)

$$\begin{aligned} HECTARECORN_{ij} =& \beta_0 + \beta_1 PIXELCORNavg_i + \beta_2 PIEXELSOYavg_i \\ &+ \beta_3 PIXELCORNcent_{ij} + \beta_4 PIEXELSOYcent_{ij} + \alpha_i + e_{ij}. \end{aligned} \tag{0.1}$$

We used the `lmer` function to estimate the parameters in the model and calculated 3 point estimates, the county sample mean, small area mean (sam) estimator and conditional linear predictor (clp), as well as the design-based variance of the sample mean (DE), the mean squared error of our proposed approximation (LW), the Prasad-Rao (PR) and Rao-Molina (RM) mean squared error estimates. The results are presented in in Table S5.

We see in Table S5, that the sam and clp predictors are very similar. The LW estimator of the MSE proposed in this study is larger than both the RM and PR estimators when the sample size is very small, specifically in the range from 1 to 4. Nevertheless, the difference between the proposed LW estimator and both RM and PR estimators tends to decrease with increasing sample size. Notably, when the sample size reaches 6, the proposed LW estimator is smaller than both the RM and PR estimators. There is no simple pattern in the relationship between the sample mean and the sam/clp predictors or between the variance of the direct estimator, DE, and the other MSE estimators.

Table S4: Simulated design-coverage and design-average length of confidence intervals when α_i has a mixture distribution and e_{ij} has a normal distribution with variances $\dot{\sigma}^2_\alpha = 4$ and $\dot{\sigma}^2_e = 100$, respectively. * identifies states in Group 3 and † identifies states in Group 2.

Method				Direct Estimator		sam-LW		sam-RM		clp-LW		clp-PR	
Area	N_i	n_i	$\hat{\alpha}_i/\hat{\sigma}^2_\alpha$	Cvge	Alen	Cvge	Alen	Cvge	Alen	Cvge	Alen	Cvge	Alen [/hline]
1	146	4	0.117	0.998	6.302	1.000	4.867	0.996	1.870	1.000	4.867	0.977	1.730
$2^\dagger$	113	7	-0.613	0.967	5.017	0.998	3.613	0.213	1.809	0.999	3.613	0.121	1.667
3	77	7	0.098	1.000	4.783	1.000	3.557	0.997	1.863	1.000	3.557	0.998	1.668
4	40	16	-0.010	1.000	3.251	0.994	1.911	0.970	1.518	1.000	1.911	0.999	1.523
5	47	16	-0.035	1.000	2.748	1.000	2.004	0.996	1.542	1.000	2.004	1.000	1.518
6	42	16	0.367	1.000	3.438	0.998	1.941	0.984	1.526	1.000	1.941	0.985	1.518
7	49	17	-0.245	1.000	3.372	0.999	1.934	0.995	1.507	1.000	1.934	0.999	1.503
8	60	30	0.506	1.000	2.257	0.981	1.274	0.964	1.125	0.990	1.274	0.992	1.342
9	104	31	-0.004	1.000	2.123	0.996	1.485	0.990	1.237	0.999	1.485	0.999	1.329
$10^\dagger$	103	31	0.789	0.999	2.355	0.961	1.482	0.891	1.236	0.964	1.482	0.877	1.329
11^*	102	31	1.073	0.996	2.056	0.933	1.479	0.822	1.235	0.842	1.479	0.726	1.329
12	109	33	0.084	1.000	2.220	0.991	1.435	0.976	1.208	0.997	1.435	0.996	1.310
13	112	34	-0.115	1.000	2.261	0.999	1.413	0.995	1.194	1.000	1.413	1.000	1.299
14	114	34	0.006	1.000	2.105	0.995	1.418	0.984	1.198	0.999	1.418	0.999	1.301
15	113	34	-0.031	0.999	2.314	0.996	1.415	0.991	1.196	0.998	1.415	0.997	1.299
16^*	114	34	-0.807	1.000	2.452	0.929	1.418	0.870	1.198	0.919	1.418	0.856	1.301
17	74	37	0.143	1.000	1.512	0.997	1.147	0.993	1.031	1.000	1.147	1.000	1.270
18	75	38	-0.375	0.999	1.869	0.960	1.125	0.938	1.015	0.992	1.125	0.995	1.263
19	125	38	-0.197	0.998	2.068	0.980	1.336	0.958	1.161	0.984	1.336	0.980	1.288
20	127	38	-0.049	0.999	1.945	0.992	1.340	0.988	1.149	0.995	1.340	0.994	1.261
21	134	40	-0.299	1.000	2.169	0.981	1.307	0.961	1.128	0.994	1.307	0.992	1.244
22	82	41	0.487	0.998	1.752	0.969	1.090	0.953	0.988	0.982	1.090	0.994	1.234
23	86	43	0.171	0.999	1.667	0.965	1.064	0.945	0.969	0.990	1.064	0.998	1.219
24	147	44	-0.265	0.997	1.831	0.987	1.246	0.976	1.088	0.993	1.246	0.993	1.210
25^*	148	44	-1.941	0.961	1.962	0.704	1.247	0.608	1.089	0.451	1.247	0.437	1.209
26	88	44	0.267	1.000	1.609	0.977	1.052	0.962	0.962	0.990	1.052	0.999	1.217
27	91	46	-0.189	1.000	1.613	0.993	1.023	0.983	0.937	0.997	1.023	1.000	1.193
28^*	176	53	1.257	0.998	1.772	0.924	1.133	0.854	1.012	0.832	1.133	0.817	1.142
29	181	54	-0.127	0.999	1.930	0.984	1.125	0.970	1.008	0.991	1.125	0.993	1.138
30	200	60	-0.063	0.999	1.604	0.991	1.066	0.987	0.965	0.995	1.066	0.997	1.097

Table S5: Small area estimation for reported hectares of corn in 12 Iowa Counties

County	Area size	sample size	$\hat{\alpha}_i/\hat{\sigma}^2_\alpha$	sample mean	DE	sam	clp	LW	RM	PR
Cerro Gordo	545	1	0.199	165.760	—	104.929	104.817	305.690	100.367	100.580
Hamilton	566	1	-0.159	96.320	—	144.880	144.966	305.711	100.512	100.716
Worth	394	1	-0.147	76.080	—	120.949	121.063	305.475	99.523	99.796
Humboldt	424	2	0.329	150.875	1180.862	100.694	100.456	152.404	93.178	93.113
Frankin	564	3	0.212	158.623	10.786	137.060	136.945	101.541	77.220	76.941
Pocahontas	570	3	-0.086	102.523	624.721	111.231	111.277	101.547	78.461	78.175
Winnebago	402	3	-0.042	112.773	308.713	116.989	117.021	101.322	77.190	76.800
Wright	567	3	0.242	144.297	966.822	119.754	119.623	101.544	78.962	78.690
Webster	687	4	0.075	117.595	112.743	111.870	111.836	76.117	67.921	67.573
Hancock	569	5	-0.205	109.382	48.621	121.837	121.947	60.712	59.873	59.387
Kossuth	965	5	-0.033	110.252	29.218	112.267	112.278	60.933	58.861	58.571
Hardin	556	6	-0.251	114.810	205.886	127.496	127.635	50.491	54.477	53.918

RM is intended to be used with sam; PR is intended to be used with clp.