Small Area Estimation Alternatives for the National Crime Victimization Survey

Robert E. Fay[1], Mamadou S. Diallo[1]
[1]Westat, Inc., 1600 Research Boulevard, Rockville, MD 20850

Abstract
The National Crime Victimization Survey (NCVS) has provided annual estimates of the number of victimizations for several types of crime since 1972, with an almost exclusive focus on national rates. Most of the programs to prevent or reduce crime are implemented locally, however. To respond to a resulting interest in subnational statistics, the Bureau of Justice Statistics (BJS) has been recently supporting research on a variety of approaches to produce subnational estimates. In this paper, we report on the potential application of model-based small area estimation methods based on the NCVS and auxiliary data, particularly the FBI's Uniform Crime Reports, using empirical best linear unbiased estimation (EBLUP). We compare a time-series model introduced by Rao and Yu to a new variant, termed here the *dynamic model*. We will also indicate how the small area approach might be integrated with other approaches that BJS is currently considering, including possible expansion of the NCVS sample size to augment the survey's capacity to produce direct estimates for some or all states.

Key Words: NCVS, Fay-Herriot, violent crime, Rao-Yu model, EBLUP

1. Introduction

Since 1972, the National Crime Survey (NCS) and its 1992 successor, the National Crime Victimization Survey (NCVS), have provided annual estimates of the frequency and consequences of crime as reported by the victims of crime. The NCVS is based on a national household sample interviewed by the Census Bureau under the protection of census confidentiality, on behalf of the Bureau of Justice Statistics. With a relatively small number of exceptions, the emphasis has been on publishing national estimates rather than subnational estimates for governmental units such as states. Microdata files available from the survey have limited geographic information; for example, they have never included state identifiers. In part, the limited NCVS sample size has reinforced this emphasis on national estimates.

Since 1930, the FBI has been the other major source of U.S. crime statistics through its Uniform Crime Report (UCR) Program (Barnett-Ryan, 2007). Law enforcement agencies report counts of crimes by type, which are then published both in disaggregated form and summarized to higher geographic levels, including state and national levels. More recently, some jurisdictions now participate in the NIBRS system, which records detailed information for each reported incident. Jurisdictions participating in NIBRS submit to the UCR by aggregating NIBRS data.

By its very nature, the UCR can only reflect crimes reported to the police, a shortcoming that the NCVS addresses. Among other limitations of the UCR are: incomplete coverage, irregular reporting by participating agencies, and possible differences in local interpretations of the UCR's categories of crimes. Although now more than 18,000 law enforcement agencies participate in the system, covering approximately 97.4% of the population in 2010 (FBI, http://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2010/crime-in-the-u.s.-2010 /aboutucrmain, downloaded 9 May 2012), undercoverage has been a serious issue until recently. Additionally, some agencies, although participating, may submit only partial reports during the year, creating additional problems of missing data (Maltz, 2007). To prepare national and state estimates, the FBI imputes for missing data, but it does not release the imputations at the level of the individual law enforcement agency. Although the overall design of the UCR includes the intent to classify crimes according to standard definitions, variation in state statutes and practice may affect classification of crimes by law enforcement personnel.

A previous paper (Li, Diallo, and Fay, 2012) outlined an initial analysis of the potential to develop small area estimates (e.g., Rao, 2003) based on the NCVS, particularly for states. The next section summarizes the findings from that paper. In short, however, we concluded that evidence from both the UCR and the NCVS favored building on the strong correlations across time in the crime rates by state. A specific approach proposed by Rao and Yu (1992, 1994) appeared to be the best candidate.

The third section of this paper describes Rao-Yu model further. Rao and Yu (1992, 1994) provided a method-of-moments estimator for their model; Rao (2003) presents this estimation approach for the model but cautions about the challenges of implementing it in practice. After encountering similar difficulties, we found that estimating the model through maximum likelihood is far more promising. The section also proposes a variant of this model, termed here the dynamic model, which offers distinct advantages over the Rao-Yu model in some situations.

The fourth section reports an empirical comparison of the Rao-Yu and dynamic models. Although we have chosen not to report on our ongoing work with the confidential NCVS data in this paper, instead we compare the fits of the Rao-Yu and dynamic models to the state-level UCR data, which is readily available on the FBI's web site. We find better results with the dynamic model, even though the improvements over the Rao-Yu model are comparatively modest, considering the large size of the underlying UCR data sets. The empirical findings do not rule out the potential usefulness of the Rao-Yu model to other applications, but they do establish the dynamic model as a promising competitor to the Rao-Yu model.

A fifth section presents our results on MSE estimation for the dynamic model, and it is followed by one presenting a small simulation study. The concluding section summarizes the findings and identifies future directions.

2. A Data-Driven Approach to the Small Area Strategy

A previous paper (Li, Diallo, and Fay, 2012) reported efforts to identify auxiliary data useful for state-level small area estimation based on the NCVS. In spite of differences between the UCR and NCVS measurement of crime, which have been extensively studied (e.g., Lynch and Addington, 2007), the considerable overlap between UCR and NCVS concepts led us to select the UCR as the most obvious initial source for auxiliary data. With the exception of simple assault, the major categories of violent crime measured by the NCVS—rape and sexual assault, robbery, and aggravated assault—have corresponding statistics reported in the UCR. Similarly, the major types of property crime in the NCVS—automobile theft, burglary, and theft—can be aligned with corresponding measures in the UCR.

At the state level, the various major UCR components of crime rates present quite different, rather than generally similar, patterns of geographic distribution (Li, Diallo, and Fay, 2012). For example, several states that are relatively high in the incidence of robbery are low on forcible rape. On the other hand, for any given type of crime, the relative ranking of the states appears quite stable over time. Fig. 1 illustrates this temporal stability.

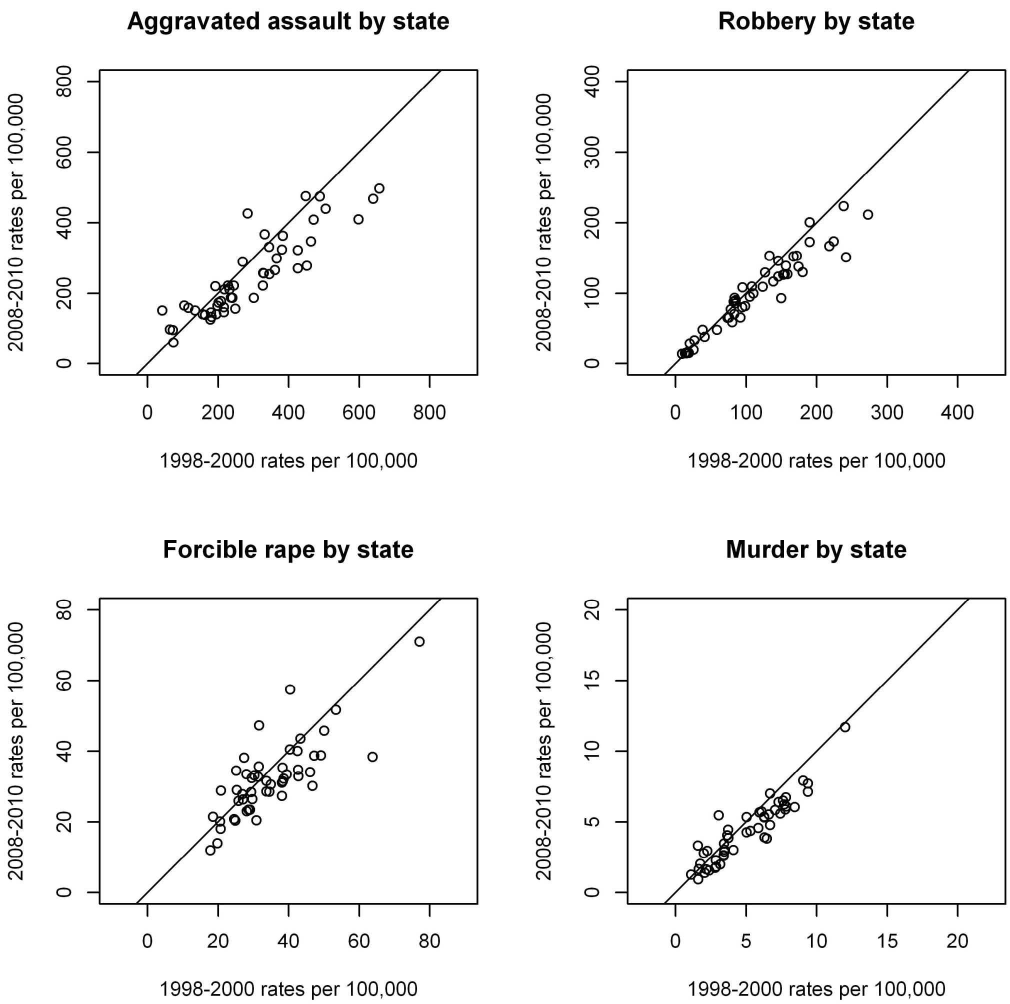

Figure. 1: Comparison of UCR state-level rates for components of violent crime for 1998-2000 and for 2008-2010. The District of Columbia is omitted from the comparisons because of high rates for robbery and murder. A line with slope 1 through the origin is shown, and the extent to which states fall generally below the line is consistent with the national drop in crime during this period. (From Li, Diallo, and Fay, 2012).

Although the UCR might be an imperfect proxy for the expected values of the state crime rates as measured by the NCVS, the strong stability across time in the UCR suggests that small area models for the NCVS should be able to incorporate time-series information. Li, Diallo, and Fay (2012) provided additional graphical evidence of temporal stability in the UCR.

The sample design of the NCVS is also a consideration in developing the modeling strategy. The NCVS has a multi-stage, rotating panel design. Sampled housing units are included for 7 interviews, spaced 6 months apart. With few exceptions, all persons age 12 and over are eligible for self-response about personal crime, including violent crime (that is, simple or aggravated assault, robbery, and rape and sexual assault). A household respondent reports for property crime (theft, motor vehicle theft, and burglary). The sample size of the NCVS has varied somewhat, going through a period of slow decline until a recent increase. In 2010, 40,974 households and 73,283 persons responded (Truman, 2011). The panel design and the use of the same first-stage sample of primary sampling units during the decade between sample redesigns produce sampling covariances between years that must be taken into account.

For our small area research, we worked with the Census Bureau's internal files to model the sampling variances and covariances for each type of crime. Although we will not use the results in the rest of the paper, we will briefly describe here the model to indicate the role of covariance across time in our small area estimation effort. We estimated design-based variance-covariance matrices for the each type of crime for 1997-2004 and 2006-2010 for self-representing (SR) and non-self-representing (NSR) areas separately. (Because of the sample redesign, the variance estimation codes changed between 2004 and 2005, leading us to consider two time periods rather than a single long one. We also excluded 2005 in computing average correlations because of additional changes in the assignment of variance estimation codes between 2005 and 2006.) We derived averaged correlation coefficients for 1-year time lags, 2-year time lags, etc., again separately for SR and NSR. The correlations between years are appreciably higher in NSR areas due to the first-stage selection of NSR counties, a selection that remains fixed between redesigns each decade. (Similar calculations can be performed for the national estimates with the public use files for the NCVS. The internal files allowed us to distinguish between SR and NSR areas.) For each state, we estimated the expected sample size and the expected proportion of the population that would fall in SR areas, using the public model of the design described by Fay and Li (2012). We then combined the estimated variances for the national rates for each year, the proportion SR, the averaged correlations for SR and NSR areas, and the expected sample size to produce a modeled variance-covariance matrix for each state.

3. Statistical Methods

In his summary of the field of small area estimation, Rao (2003) broadly distinguished between area-level models and unit-level models. In area-level models, the unit of analysis is the set of individual areas for which estimates are desired. In unit-level models, modeling occurs at the level of measurement, such as the individual, and small area estimates are produced by aggregating the available direct observations with the predicted values under the model for the unobserved units. Some researchers have developed various area-level models that feature time series aspects to incorporate survey information across time, an approach seemingly appropriate to the NCVS application. For one such approach, Rao and Yu (1992, 1994) proposed an extension to the Fay-

Herriot model (1979). Both models assume the standard relationship between a sample estimate, y_{it}, for area *i* and time *t*, and its expected population value, θ_{it},

$$y_{it} = \theta_{it} + e_{it} \tag{3.1}$$

for *i*=1,…,*m*, *t*=1,…,*T*, where the normally distributed error terms $e_i = (e_{i1},...,e_{iT})^T$ are assumed to have mean zero and covariance $\sum_i$, which is assumed known. The error terms are assumed independent between areas. Rao and Yu proposed the following model for the underlying population:

$$\theta_{it} = \mathbf{x}_{it}^T \boldsymbol{\beta} + \nu_i + u_{it} \tag{3.2}$$

with

$$u_{it} = \rho u_{i,t-1} + \varepsilon_{it} \tag{3.3}$$

where

$\mathbf{x}_{it} = (x_{it1},...,x_{itp})^T$ is the vector of auxiliary variables for area *i* at time *t*,
$\boldsymbol{\beta}$ is a vector of regression coefficients,
$\nu_i \sim N(0, \sigma_v^2)$ for *i*=1,…,*m* are iid random effects, representing time-independent differences among areas, and
$\varepsilon_{it} \sim N(0, \sigma^2)$ are iid random variables, which induce variability in the series u_{it} *t*=1,…*T*.

Rao and Yu considered the case of $|\rho| < 1$ and assumed stationarity for the series in (3.3). In turn, this assumption implies for all *i* and *t*:

$$Var(u_{it}) = \sigma^2 / (1 - \rho^2) \tag{3.4}$$

When $\rho = 1$, (3.3) represents a random walk, but only by dropping the stationarity assumption (3.4). Thus, there is a discontinuity in the model at $\rho = 1$.

Assuming first that σ^2, σ_v^2, and ρ are known, the best linear unbiased predictor (BLUP) for area *i* at time *T* is

$$\tilde{\theta}_{iT} = \mathbf{x}_{iT}^T \tilde{\boldsymbol{\beta}} + (\sigma_v^2 \mathbf{1}_T + \sigma^2 \boldsymbol{\gamma}_T)^T (\boldsymbol{\Sigma}_i + \sigma^2 \boldsymbol{\Gamma} + \sigma_v^2 \mathbf{J}_T)^{-1} (\mathbf{y}_i - \mathbf{X}_i \tilde{\boldsymbol{\beta}}) \tag{3.5}$$

where

$\boldsymbol{\Gamma}$ is a $T \times T$ matrix with elements $\rho^{|i-j|} / (1 - \rho^2)$,
$\mathbf{J}_T$ is a $T \times T$ matrix with elements = 1,
$\mathbf{V}_i = \boldsymbol{\Sigma}_i + \sigma^2 \boldsymbol{\Gamma} + \sigma_v^2 \mathbf{J}_T = Cov(\mathbf{y}_i)$,
$\mathbf{V} = diag_i(\mathbf{V}_i) = Cov(\mathbf{y})$,

$\tilde{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}$ is the generalized least-squares estimator of $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}_T$ is the T th column of $\boldsymbol{\Gamma}$.

Under this approach, the small area estimates are given by the empirical best linear unbiased predictor (EBLUP), based on estimating σ_v^2, σ^2, and ρ in (3.4). Rao (2003, pp. 159) summarized the previous efforts of Rao and Yu (1994) to estimate the parameters, building on earlier work in econometrics. When ρ is known, they reported satisfactory behavior using a method of moments approach to derive closed form estimates for σ_v^2 and σ^2. They discussed a method of moments estimator for ρ from the same literature, but they found its performance to be poor—in fact extremely so in the presence of sampling error. They suggested instead trying to obtain an estimate of ρ from an external source. We similarly encountered poor performance of their suggested estimator of ρ in our initial attempts to use it.

We appeared to confront several challenges in using the Rao-Yu model for NCVS state estimates. Although our earlier paper suggested the UCR as a possible source for ρ, we preferred to be able to estimate this parameter from the NCVS data. We observed that when ρ is high in the Rao-Yu model, such as ρ > .8, separating the estimation for σ_v^2 and σ^2 in the presence of sampling error becomes increasingly challenging compared to moderate ρ, such as ρ = .4. We were also concerned about the stability of the method of moments approach when Var(e_{it}) varied widely from state to state.

These concerns motivated an alternative model, which we term the dynamic model, drawing a parallel to some models in mathematical biology (e.g., Murray, 2002). The dynamic model is based on the same sampling model (3.1), but in place of (3.2) and (3.3), we considered instead

$$\begin{aligned} \theta_{it} &= \mathbf{x}_{it}^T\boldsymbol{\beta} + \rho^{t-1} v_i^* + u_{it}^* \\ u_{i1}^* &= 0 \\ u_{it}^* &= \rho u_{i,t-1}^* + \varepsilon_{it} \qquad \text{for } t>1 \end{aligned} \tag{3.6}$$

where $v_i^* \sim N(0, \sigma_{v^*}^2)$ for i=1,…,m are iid random effects for areas at time t = 1, and ε_{it} and σ^2 are the same as in (3.3). In the new model,

$$\theta_{it} = \mathbf{x}_{it}^T\boldsymbol{\beta} + \rho\,(\theta_{i,t-1} - \mathbf{x}_{i,t-1}^T\boldsymbol{\beta}) + \varepsilon_{it} \tag{3.7}$$

which shows that the role of ρ is somewhat different in (3.6) than in (3.3). Unlike the Rao-Yu model, which assumes stationarity (3.4) for $|\rho| < 1$, the dynamic model does not assume stationarity, does not constrain ρ to be less than 1, and avoids a discontinuity at $\rho = 1$. In fact values greater than 1 can reflect systematically increasing divergence, a phenomenon called “Matthew effects” in some literatures.

When σ^2, σ^2_{v*} and ρ are known, the BLUP for this model is

$$\tilde{\theta}_{iT} = \mathbf{x}_{iT}^T \tilde{\boldsymbol{\beta}} + (\sigma^2 \boldsymbol{\gamma}_{T,u*} + \sigma^2_{v*} \boldsymbol{\gamma}_{T,v*})^T V_i^{-1} (\mathbf{y}_i - \mathbf{X}_i \tilde{\boldsymbol{\beta}}) \tag{3.8}$$

where

$\boldsymbol{\Gamma}_{u*}$ is a $T \times T$ symmetric matrix with elements:

$$\Gamma_{u*(1,j)} = 0$$

$$\Gamma_{u*(i,j)} = \rho^{(j-i)} \sum_{i'=1}^{i-1} \rho^{(2i'-2)} \qquad \text{for } 1 < i \le j$$

$\boldsymbol{\Gamma}_{v*}$ is a $T \times T$ symmetric matrix with elements $\rho^{(i+j-2)}$,

$\mathbf{V}_i = \boldsymbol{\Sigma}_i + \sigma^2 \boldsymbol{\Gamma}_{u*} + \sigma^2_{v*} \Gamma_{v*} = Cov(\mathbf{y}_i)$,

$\mathbf{V} = diag_i(\mathbf{V}_i) = Cov(\mathbf{y})$,

$\tilde{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{y}$ is the generalized least-squares estimator of $\boldsymbol{\beta}$,

$\boldsymbol{\gamma}_{T,u*}$ is the T th column of $\boldsymbol{\Gamma}_{u*}$, and

$\boldsymbol{\gamma}_{T,v*}$ is the T th column of $\boldsymbol{\Gamma}_{v*}$.

Rao (2003, section 6.2.4) describes how to implement maximum likelihood estimation (MLE) and restricted maximum likelihood estimation (REML) for the general area-level mixed model. His description is generally sufficient to describe an iterative Newton-Ralphson algorithm to obtain the EBLUP for the dynamic model for either estimation method. A complicating issue is that the parameters $\delta = (\sigma^2_{v*},\ \sigma^2,\ \rho)$ are subject to the constraints $\sigma^2_{v*} \ge 0$ and $\sigma^2 \ge 0$, but the Newton-Ralphson algorithm can be modified to respect these constraints. We have implemented both the MLE and REML versions in R.

The Rao-Yu model is equivalent to the dynamic model when (1) $\sigma_v^2 = 0$ in (3.2) for the Rao-Yu model, (2) $\rho < 1$, (3) the series (3.3) is stationary, and (4) $\sigma^2_{v*} = \sigma^2/(1-\rho^2)$ in the dynamic model. When $\sigma_v^2 > 0$ in (3.2) for the Rao-Yu model, however, the two models are not in general equivalent.

In fact, the Rao-Yu model has been previously fitted by another researcher through MLE (J.N.K. Rao, personal communication). The use of Newton-Ralphson in this case requires a bit more care because the parameters $\delta = (\sigma_v^2,\ \sigma^2,\ \rho)$ require a third constraint $|\rho| < 1$, but we have successfully implemented an algorithm for the MLE of the Rao-Yu model in R.

4. Comparison of the Rao-Yu and Dynamic Models

Because neither the Rao-Yu model nor the dynamic model is a special case of the other, we created a superordinate model with random effect terms from both, so that each model was a special case of the superordinate model. We applied the superordinate model to the UCR data to assess the relative performance of the Rao-Yu and dynamic models to the overall fit.

The superordinate model is of the form

$$\begin{aligned} \theta_{it} &= \mathbf{x}_{it}^T \boldsymbol{\beta} + v_i + \rho^{t-1} v_i^* + u_{it}^* \\ u_{i1}^* &= 0 \\ u_{it}^* &= \rho u_{i,t-1} + \varepsilon_{it} \qquad \text{for } t > 1 \end{aligned} \tag{4.1}$$

where $v_i \sim N(0,\sigma_v^2)$ and $v_i^* \sim N(0,\sigma_{v*}^2)$ for i=1,…,m are each mutually independent sets of iid random effects for areas. Model (4.1) holds if either (3.2) or (3.6) holds, but (4.1) allows for an optimum combination of the two sets of random effects. For high ρ, the parameters of the superordinate model are unstable, particularly because of the identifiability problem at $\rho = 1$.

Even though the UCR data are aggregate totals, we attributed Poisson variance to the observed counts. We computed the log-likelihood for the maximum-likelihood estimates for the Rao-Yu and dynamic models separately. For each of the UCR types of crime, the log-likelihood for the dynamic model consistently exceeds the log-likelihood for the Rao-Yu model.

We used the increase in the log-likelihood from either separate model to the superordinate model as a measure of the increased improvement in fit due to including the features of the other model. Large increases indicate that the initially omitted model represents an improvement, whereas no increase or small increases indicate that the initial model is adequate. We carried out the analysis for all of the states with DC included, and without DC. The results are presented in Table 1.

Table 1: Comparison of Improvements in the Log-Likelihood of the Superordinate Model Relative to the Dynamic or Rao-Yu Models, UCR 1997-2010

	Dynamic model		Rao-Yu model	
States and DC	$\hat{\rho}$	$\Delta(loglkhd)$	$\hat{\rho}$	$\Delta(loglklhd)$
Aggravated assault	.950	0.85	.990[a]	30.99
Robbery	.976	17.13	.990[a]	46.33
Forcible rape	.968	0.00	.981	1.69
Larceny	.926	0.00	.979	26.66
Burglary	.971	6.39	.987	10.12
Motor vehicle theft	.949	0.00	.976	9.07
States without DC				
Aggravated assault	.954	1.14	.989	19.89
Robbery	.944	0.99	.990[a]	40.13
Forcible rape	.969	0.00	.982	1.73
Larceny	.929	0.00	.979	23.76
Burglary	.974	4.07	.988	6.83
Motor vehicle theft	.937	0.00	.969	10.72

Note: [a] The algorithm constrained rho for the Rao-Yu model to a maximum of .990 in each iteration.

The superordinate model appreciably improves the log-likelihood for the dynamic model only for robbery and burglary when DC is included, and only for burglary when DC is omitted. In other words, almost all of the fit of the superordinate model can be attributed to the dynamic model. On the other hand, the superordinate model improves the Rao-Yu model by significant amounts. (Without the effect of the boundary constraints on the parameters, twice these differences in log-likelihoods could be interpreted as likelihood-ratio chi-square statistics on one degree of freedom, but their behavior in the presence of boundary constraints is less straight-forward.)

To put these results in perspective, however, these differences in log-likelihood are relatively modest compared to the scale of the UCR. The results establish a basis to apply the dynamic model to the NCVS, but do not suggest particularly adverse consequences if the Rao-Yu model had been employed instead. Although we have identified some potentially attractive features of the dynamic model, the practical differences with the Rao-Yu model may be more subtle than substantial. Because of their appreciable sampling variances, we do not expect the NCVS data by themselves to provide a firm basis to select one model over the other.

5. Mean Square Error (MSE) Estimation for the Dynamic Model

The Rao-Yu model and the dynamic model are closely related instances of the much wider class of general mixed models. In chapter 6 of his book, Rao (2003) traces connections between the theory for the general mixed model and many small area applications. In general, MSE estimates depend on the values of parameters that typically must be estimated. For example, the parameters of the dynamic model are $\delta = (\sigma_{v*}^2,\ \sigma^2,\ \rho)$. Rao's development of the MSE progresses from (1) the MSE of the best linear unbiased estimator (BLUP) when the true values of the parameters are known but the fixed and random effects are estimated, to (2) the MSE of the empirical BLUP (EBLUP) when the parameters are estimated (e.g., by REML) but assumed known, to (3) estimation of the MSE for the EBLUP without prior knowledge of the parameters. In this section, we follow this general outline for the dynamic model.

Following Rao's notation, a general linear mixed model for the sample data may be written

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \mathbf{e}$$

where $\mathbf{X}$ and $\mathbf{Z}$ are known $n \times p$ and $n \times h$ matrices of full rank, $\mathbf{v}$ and $\mathbf{e}$ are independently distributed with means $\mathbf{0}$ and covariance matrices $\mathbf{G}$ and $\mathbf{R}$ depending on variance parameters $\boldsymbol{\delta}$. The variance of $\mathbf{y}$ is thus $\mathrm{Var}(\mathrm{y}) = \mathrm{V} = \mathrm{V}(\delta) = \mathbf{R} + \mathbf{Z}\mathbf{G}\mathbf{Z}^{\boldsymbol{T}}$**.** The theory for the general linear mixed model is developed for any linear estimate of the form

$$\mu = \mathbf{l}^{\boldsymbol{T}}\boldsymbol{\beta} + \mathbf{m}^{\boldsymbol{T}}\mathbf{v}.$$

For the dynamic model, $n = T \times m$, where T is the number of years and m is the number of areas. Because the covariance matrices $\mathbf{G}$ and $\mathbf{R}$ each have block diagonal form when arranged by area, most expressions can be stated for each area i separately. The random effects for each area can be expressed as $\mathbf{v}_i = (u_{i1}^* + v_{i1}^*, u_{i2}^* + v_{i2}^*, \ldots, u_{iT}^* + v_{iT}^*)^T$, with $\boldsymbol{Z}_i = \mathbf{I}_{\boldsymbol{T}}$, and

$$\mathbf{G}_i = \mathbf{G}_i(\delta) = \sigma_{v*}^2 \boldsymbol{\Gamma}_{v*}(\rho) + \sigma^2 \boldsymbol{\Gamma}_{u*}(\rho).$$

The general expression for the BLUP is

$$\tilde{\mu}^H = \mathbf{l}^T\widetilde{\boldsymbol{\beta}} + \mathbf{m}^T\tilde{\mathbf{v}} = \mathbf{l}^T\widetilde{\boldsymbol{\beta}} + \mathbf{m}^T\mathbf{G}\mathbf{Z}^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\widetilde{\boldsymbol{\beta}}).$$

The BLUP from (3.8) can be written in this form as:

$$\tilde{\theta}_{iT} = \mathbf{l}_i^T\widetilde{\boldsymbol{\beta}} + \mathbf{m}_i^T\mathbf{v}_i = \mathbf{x}_{iT}^T\widetilde{\boldsymbol{\beta}} + \mathbf{m}_i^T\mathbf{G}_i\mathbf{Z}_i^T\mathbf{V}_i^{-1}(\mathbf{y}_i - \mathbf{X}_i\widetilde{\boldsymbol{\beta}})$$

where $\boldsymbol{m}_i^T = (0,\ldots,0,1)$ with 1 in the Tth position. The other quantities are defined as in equation (3.8).

5.1 MSE of the BLUP

Assuming that the variance components, $\boldsymbol{\delta} = (\sigma_{v*}^2, \sigma^2, \rho)$, are known, the MSE of the BLUP is:

$$MSE\big(\tilde{\theta}_{iT}(\boldsymbol{\delta})\big) = g_{1iT}(\boldsymbol{\delta}) + g_{2iT}(\boldsymbol{\delta}) \qquad (5.1)$$

where

$$\begin{aligned} g_{1iT}(\boldsymbol{\delta}) = \boldsymbol{m}_i^T\big(\boldsymbol{G}_i - \boldsymbol{G}_i\boldsymbol{Z}_i^T\boldsymbol{V}_i^{-1}\boldsymbol{Z}_i\boldsymbol{G}_i\big)\boldsymbol{m}_i \\ = \rho^{2(T-1)}\sigma_{v*}^2 + \sum_{t=1}^{T-1}\rho^{2(t-1)}\,\sigma^2 \\ - \big(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\big)^{\boldsymbol{T}}\boldsymbol{V}_i^{-1}\big(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\big) \end{aligned}$$

with $\boldsymbol{\gamma}_{v*}$ and $\boldsymbol{\gamma}_{u*}$ being respectively the Tth column of the matrices $\boldsymbol{\Gamma}_{v*}$ and $\boldsymbol{\Gamma}_{u*}$, and

$$g_{2iT}(\delta) = \boldsymbol{d}_i^T\left(\sum_{i=1}^{m}\boldsymbol{X}_i^T\boldsymbol{V}_i^{-1}\boldsymbol{X}_i\right)^{-1}\boldsymbol{d}_i$$

where

$$\boldsymbol{d}_i^T = \mathbf{x}_{iT}^T - \boldsymbol{b}_i^T\boldsymbol{X}_i^T$$

$$\boldsymbol{b}_i^T = \boldsymbol{m}_i^T\boldsymbol{G}_i\boldsymbol{Z}_i^T\boldsymbol{V}_i^{-1} = \big(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\big)^{\boldsymbol{T}}\boldsymbol{V}_i^{-1}$$

Thus,

$$\begin{aligned} g_{2iT}(\boldsymbol{\delta}) = \mathbf{x}_{iT}^T\left(\sum_{i=1}^{m}\boldsymbol{X}_i^T\boldsymbol{V}_i^{-1}\boldsymbol{X}_i\right)^{-1}\mathbf{x}_{iT} \\ + \big(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\big)^{\boldsymbol{T}}\boldsymbol{V}_i^{-1}\boldsymbol{X}_i^T\left(\sum_{i=1}^{m}\boldsymbol{X}_i^T\boldsymbol{V}_i^{-1}\boldsymbol{X}_i\right)^{-1}\boldsymbol{X}_i\boldsymbol{V}_i^{-1}\big(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} \\ + \sigma^2\boldsymbol{\gamma}_{T,u*}\big) - \mathbf{2}\big(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\big)^{\boldsymbol{T}}\boldsymbol{V}_i^{-1}\boldsymbol{X}_i^T\left(\sum_{i=1}^{m}\boldsymbol{X}_i^T\boldsymbol{V}_i^{-1}\boldsymbol{X}_i\right)^{-1}\mathbf{x}_{iT} \end{aligned}$$

The term $g_{2iT}(\delta)$ accounts for the variability in the estimation of $\boldsymbol{\beta}$.

5.2 MSE of the EBLUP

The EBLUP obtained from (3.8) after estimating the variance components has a second-order approximation to the MSE equal to:

$$MSE\big(\tilde{\theta}_{iT}(\widehat{\boldsymbol{\delta}})\big) = g_{1iT}(\boldsymbol{\delta}) + g_{2iT}(\boldsymbol{\delta}) + g_{3iT}(\boldsymbol{\delta}) \qquad (5.2)$$

Where $g_{1iT}(\boldsymbol{\delta})\ and\ g_{2iT}(\boldsymbol{\delta})$ are defined above and

$$g_{3iT}(\boldsymbol{\delta}) = tr[(\frac{\partial \boldsymbol{b}_{iT}^T}{\partial \boldsymbol{\delta}})V_i \left(\frac{\partial \boldsymbol{b}_{iT}^T}{\partial \boldsymbol{\delta}}\right)^T \bar{V}(\widehat{\boldsymbol{\delta}})]$$

With

$$\bar{V}(\widehat{\boldsymbol{\delta}}) = \mathcal{I}^{-1}(\delta)$$

The information matrix $\mathcal{I}(\delta)$, which is a 3x3 matrix, is defined as:

$$\mathcal{I}_{jk}(\boldsymbol{\delta}) = \frac{1}{2} tr\left(\boldsymbol{V}^{-1}\boldsymbol{V}_{(j)}\boldsymbol{V}^{-1}\boldsymbol{V}_{(k)}\right) \text{ where } \boldsymbol{V}_{(r)} = \frac{\partial \boldsymbol{V}}{\partial \delta_j} \tag{5.3}$$

Moreover, for the dynamic model, the elements of the information matrix are

$$\mathcal{I}_{vv}(\boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} tr(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \sigma_{v*}^2})^2 = \frac{1}{2}\sum_{i=1}^{m} tr(\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{v*})^2$$

$$\mathcal{I}_{uu}(\boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} tr(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \sigma^2})^2 = \frac{1}{2}\sum_{i=1}^{m} tr(\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{u*})^2$$

$$\mathcal{I}_{\rho\rho}(\boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} tr(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \rho})^2 = \frac{1}{2}\sum_{i=1}^{m} tr(\boldsymbol{V}_i^{-1}(\sigma_{v*}^2\boldsymbol{\Gamma}'_{v*} + \sigma^2\boldsymbol{\Gamma}'_{u*}))^2$$

with

$$\boldsymbol{\Gamma}'_{v*} = \frac{\partial \boldsymbol{\Gamma}_{v*}}{\partial \sigma_{v*}^2} \quad and \quad \boldsymbol{\Gamma}'_{u*} = \frac{\partial \boldsymbol{\Gamma}_{u*}}{\partial \sigma^2}$$

$$\mathcal{I}_{vu}(\boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} tr[\left(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \sigma_{v*}^2}\right)\left(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \sigma^2}\right)] = \frac{1}{2}\sum_{i=1}^{m} tr[(\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{v*})(\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{u*})]$$

$$\mathcal{I}_{v\rho}(\boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} tr[\left(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \sigma_{v*}^2}\right)\left(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \rho}\right)] = \frac{1}{2}\sum_{i=1}^{m} tr[(\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{v*})(\sigma_{v*}^2\boldsymbol{\Gamma}'_{v*} + \sigma^2\boldsymbol{\Gamma}'_{u*})]$$

$$\mathcal{I}_{u\rho}(\boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} tr[\left(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \sigma^2}\right)\left(\boldsymbol{V}_i^{-1}\frac{\partial \boldsymbol{V}_i}{\partial \rho}\right)] = \frac{1}{2}\sum_{i=1}^{m} tr[(\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{u*})(\sigma_{v*}^2\boldsymbol{\Gamma}'_{v*} + \sigma^2\boldsymbol{\Gamma}'_{u*})]$$

and

$$\frac{\partial \boldsymbol{b}_{iT}^T}{\partial \sigma_{v*}^2} = \boldsymbol{\gamma}_{T,v*}^T\boldsymbol{V}_i^{-1} - \left(\sigma_v^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\right)^T\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{v*}\boldsymbol{V}_i^{-1}$$

$$\frac{\partial \boldsymbol{b}_{iT}^T}{\partial \sigma^2} = \boldsymbol{\gamma}_{T,u*}^T V_i^{-1} - \left(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\right)^T\boldsymbol{V}_i^{-1}\boldsymbol{\Gamma}_{u*}\boldsymbol{V}_i^{-1}$$

$$\frac{\partial \boldsymbol{b}_{iT}^T}{\partial \rho} = \left(\sigma_{v*}^2\boldsymbol{\gamma}'_{T,v*} + \sigma^2\boldsymbol{\gamma}'_{T,u*}\right)^T V_i^{-1} - \left(\sigma_{v*}^2\boldsymbol{\gamma}_{T,v*} + \sigma^2\boldsymbol{\gamma}_{T,u*}\right)^T\boldsymbol{V}_i^{-1}(\sigma_{v*}^2\boldsymbol{\Gamma}'_{v*} + \sigma^2\boldsymbol{\Gamma}'_{u*})\boldsymbol{V}_i^{-1}$$

with $\boldsymbol{\gamma}'_{T,v*} = \frac{\partial \boldsymbol{\gamma}_{T,v*}}{\partial \sigma_{v*}^2}$ *and* $\boldsymbol{\gamma}'_{T,u*} = \frac{\partial \boldsymbol{\gamma}_{T,u*}}{\partial \sigma^2}$.

5.3 Estimating the MSE of the EBLUP

The MSE in (5.2) assumes that all the variance components and the autocorrelation coefficient are known, in reality those quantities are unknown and need to be estimated.

The estimator of $MSE\left(\tilde{\theta}_{iT}(\widehat{\boldsymbol{\delta}})\right)$ depends on the method used to estimate the parameters. If REML is used for estimating the parameters, then the estimator of $MSE\left(\tilde{\theta}_{iT}(\widehat{\boldsymbol{\delta}})\right)$ is equal to:

$$mse\left(\tilde{\theta}_{iT}(\widehat{\boldsymbol{\delta}})\right) = g_{1iT}\left(\widehat{\boldsymbol{\delta}}\right) + g_{2iT}\left(\widehat{\boldsymbol{\delta}}\right) + 2g_{3iT}\left(\widehat{\boldsymbol{\delta}}\right). \tag{5.4}$$

For calculating the information matrix involved in the calculation of $g_{3iT}\left(\widehat{\boldsymbol{\delta}}\right)$, one will use the REML version of (5.3) which is equal to:

$$\mathcal{I}_{jk}^{REML}(\delta) = \frac{1}{2} tr\left(\boldsymbol{P}\boldsymbol{V}_{(j)}\boldsymbol{P}\boldsymbol{V}_{(k)}\right) \tag{5.5}$$

where $\boldsymbol{P} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-\mathbf{1}}\boldsymbol{X})\boldsymbol{X}'\boldsymbol{V}^{-\mathbf{1}}$.

The REML method estimates the variance components and the autocorrelation coefficient without being affected by the fixed effects. This means that the REML estimates are invariant to the values of the fixed effects. Also, REML implicitly takes into account the degrees of freedom for the fixed effects. In general, the degrees of freedom can play an important role in the estimation of the parameters when the rank of the matrix X is large compared to the data. The MLE method does not have those two properties (e.g., McCulloch et al., 2008).

When MLE is used to estimate the parameters of the model, then the estimator of the mean square error is equal to (5.4) where the term $\boldsymbol{b}_{\hat{\delta}}^{T}(\widehat{\boldsymbol{\delta}})\nabla g_{1iT}(\widehat{\boldsymbol{\delta}})$ is subtracted. This gives:

$$mse_{MLE}\left(\tilde{\theta}_{iT}(\widehat{\boldsymbol{\delta}})\right) = g_{1iT}\left(\widehat{\boldsymbol{\delta}}\right) + g_{2iT}\left(\widehat{\boldsymbol{\delta}}\right) + 2g_{3iT}\left(\widehat{\boldsymbol{\delta}}\right) - \boldsymbol{b}_{\hat{\delta}}^{T}\left(\widehat{\boldsymbol{\delta}}\right)\nabla g_{1iT}\left(\widehat{\boldsymbol{\delta}}\right). \tag{5.6}$$

In the expression (5.6), the information matrix is the same as in (5.3),

$\nabla g_{1iT}(\boldsymbol{\delta}) = \left(\frac{\partial g_{1iT}(\delta)}{\partial \sigma_{v*}^{2}}, \frac{\partial g_{1iT}(\delta)}{\partial \sigma^{2}}, \frac{\partial g_{1iT}(\delta)}{\partial \rho}\right)^{T}$ and $\boldsymbol{b}_{\hat{\delta}}^{T}(\boldsymbol{\delta}) = \frac{1}{2m}(\mathcal{I}^{-1}(\delta) col_{1\le j\le 3} tr[\mathcal{I}_{\beta}^{-1}(\delta)\frac{\partial \mathcal{I}_{\beta}(\delta)}{\partial \delta}])$

where $\mathcal{I}_{\beta}(\boldsymbol{\delta})$ is the information matrix associated with the fixed effects $\boldsymbol{\beta}$. Given that $\mathcal{I}_{\beta}(\boldsymbol{\delta}) = \sum_{i=1}^{m} \boldsymbol{X}_{i}^{\boldsymbol{T}}\boldsymbol{V}_{i}^{-\mathbf{1}}\boldsymbol{X}_{i}$, we get:

$$\boldsymbol{b}_{\hat{\delta}}(\boldsymbol{\delta}) = -\frac{1}{2m}\left(\mathcal{I}^{-1}(\boldsymbol{\delta}) col_{1\le j\le 3} tr\left[\left(\sum\nolimits_{i=1}^{m} \boldsymbol{X}_{i}^{\boldsymbol{T}}\boldsymbol{V}_{i}^{-\mathbf{1}}\boldsymbol{X}_{i}\right)^{-1}\left(\sum\nolimits_{i=1}^{m} \boldsymbol{X}_{i}^{\boldsymbol{T}}\boldsymbol{V}_{i}^{-\mathbf{1}}\frac{\partial \boldsymbol{V}_{i}}{\partial \delta_{j}}\boldsymbol{V}_{i}^{-\mathbf{1}}\boldsymbol{X}_{i}\right)\right]\right).$$

6. Simulations using the Dynamic Model

In this section, we conducted simulation study to further assess the dynamic model. We used annual rape and robbery rates from UCR as auxiliary variables. We used $\sigma^2 = 0.001^2$, $\sigma_{v*}^{2} = (3\sigma)^2$, and $\rho = 0.90$. The sampling covariance matrix was diagonal and defined as $\Sigma_i = \mathrm{Adj_i} * \sigma_{v*}^{2} * \mathrm{I}_T$, where Adj_i varies from 0.1 (for the largest state California) to 10 (for the smallest state Wyoming) by an equal increment of 0.198. In total 5,000 populations were simulated and for each population two dependent variables corresponding to T=7 and T=14 were generated using model (3.6).

Table 2 shows the average of the relative bias for estimating the parameters of the model. The estimation of ρ is very precise with relative bias smaller than 1% in absolute value. Also the estimation of the variance components is more precise when using 14 years of data rather than 7 years.

Table 2: Relative bias (%) of the parameters using the dynamic model with REML

Length of the series	$\boldsymbol{\rho}$	$\boldsymbol{\sigma_{v*}^2}$	$\boldsymbol{\sigma^2}$
T=7	0.03	-4.68	3.22
T=14	-0.51	1.25	2.75

In table 3, the states are ordered by population size according the 2010 U.S. census. Table 3 gives the gain in efficiency associated with the small area estimates. The efficiency is defined as the ratio of the 2010 variance estimate to the mse estimate for a given state. As expected, the smaller states have more gain in efficiency than the larger states and also using more data improves the efficient of the small area estimates.

Table 3: 2010 Efficiency obtained from using the dynamic model with REML

State	*Efficiency (T=7)*	*Efficiency (T=14)*	*State*	*Efficiency (T=7)*	*Efficiency (T=14)*
Wyoming	17.08	19.95	Louisiana	10.23	11.59
District of Columbia	9.03	12.80	South Carolina	9.93	11.24
Vermont	16.82	19.51	Alabama	9.65	10.89
North Dakota	16.02	18.77	Colorado	9.06	10.33
Alaska	14.13	17.27	Minnesota	8.97	10.14
South Dakota	15.05	17.81	Wisconsin	8.79	9.85
Delaware	15.35	17.92	Maryland	8.39	9.43
Montana	15.11	17.60	Missouri	8.19	9.13
Rhode Island	15.19	17.58	Tennessee	7.83	8.74
New Hampshire	14.74	17.11	Arizona	7.51	8.35
Maine	14.48	16.78	Indiana	7.24	8.02
Hawaii	14.41	16.61	Massachusetts	6.93	7.65
Idaho	13.81	16.03	Washington	6.56	7.24
Nebraska	13.51	15.65	Virginia	6.31	6.90
West Virginia	13.66	15.67	New Jersey	5.92	6.48
West Mexico	12.88	14.97	North Carolina	5.63	6.11
Nevada	12.84	14.84	Georgia	5.26	5.70
Utah	12.53	14.48	Michigan	4.81	5.23
Kansas	12.20	14.08	Ohio	4.50	4.85
Arkansas	11.82	13.64	Pennsylvania	4.11	4.41
Mississippi	11.88	13.60	Illinois	3.67	3.94
Iowa	11.55	13.21	Florida	3.21	3.45
Connecticut	11.40	12.98	New York	2.69	2.92
Oklahoma	10.93	12.51	Texas	2.09	2.32
Oregon	10.70	12.20	California	1.17	1.55
Kentucky	10.46	11.88			

7. Discussion

We have introduced the dynamic model as an alternative to the Rao-Yu model for applications in which the new model appears to summarize the data more successfully than either an assumption of stationarity or an underlying random walk. But we do not intend to displace the Rao-Yu model by this new approach, and we can report that maximum likelihood estimation substantially improves upon the estimation approach described by Rao (2003) for the Rao-Yu model. In practice, the dynamic and Rao-Yu models may often produce quite similar small area estimates.

We noted the general arguments to prefer REML estimation over MLE. In the NCVS application, involving relatively few fixed effects, the differences between the two may be small. Furthermore, the performance of either approach is affected by the boundary conditions on the parameters. A limited simulation suggested a slight advantage to REML in this application, but more such work is required before stating a firm conclusion. We also plan to examine further whether the estimated MSEs can be used to choose between the two estimation methods; in particular, if the estimated MSEs for MLE are consistently better than for REML (which we have observed in our application), does this support the conclusion that MLE is preferable? A more extensive set of simulation experiments may be the most effective method to investigate this question.

Our NCVS work thus far has focused on estimating each type of crime separately. However, totals for violent crime and property crime are of considerable interest. In our application, we have observed that differences between modeling the totals and the sum of their modeled components are large enough to be of substantive importance. Rao (2003) reviewed a number of small area applications involving multivariate formulations. The general theory for BLUP covers all linear combinations of the fixed and random effects, so that producing EBLUPs for the components of violent or property crime modeled simultaneously leads to the EBLUP for their sums. We plan to pursue such an extension.

Acknowledgements

The findings and views expressed in this paper are solely those of the authors. They do not necessarily reflect those of the Bureau of Justice Statistics, of the Census Bureau, or of Westat.

We want to thank Dr. Michael Planty of BJS for his consultation and significant contributions at different stages of this research. We also thank the Bureau of Justice Statistics for support of this work.

References

Barnett-Ryan, C. (2007), “Introduction to the Uniform Crime Reporting Program,” in Lynch, J.P. and Addington, L.A. (eds.) (2007), *Understanding Crime Statistics: Revisiting the Divergence of the NCVS and UCR*, pp. 55-89.

Datta, G.S. and Lahiri, P. (2000), *A Unified Measure of Uncertainty of Estimated Best Linear Predictors in Small Area Estimation Problems*, Statistica Sinica, 10, 613-627.

Fay, R.E., and Herriot, R.A. (1979), “Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data,” Journal of the American Statistical Association, *74*, 269-277.

Fay, R.E. and Li, J. (2012) “Rethinking the NCVS: Subnational Goals through Direct Estimation,” presented at the 2012 Federal Committee on Statistical Methodology Conference, Washington, DC, Jan. 10-12, 2012, available at http://www.fcsm.gov/12papers/Fay_2012FCSM_I-B.pdf.

Li, J., Diallo, M.S., and Fay, R.E., (2012) "Rethinking the NCVS: Small Area Approaches to Estimating Crime," presented at the 2012 Federal Committee on Statistical Methodology Conference, Washington, DC, Jan. 10-12, 2012, available at http://www.fcsm.gov/12papers/Li_2012FCSM_I-B.pdf.

Lynch, J.P. and Addington, L.A. (eds) (2007a), *Understanding Crime Statistics: Revisiting the Divergence of the NCVS and UCR*, Cambridge University Press, New York, NY.

Maltz, M.D. (2007), "Missing UCR Data and Divergence of the NCVS and UCR Trends," in Lynch, J.P. and Addington, L.A. (eds.) (2007), *Understanding Crime Statistics: Revisiting the Divergence of the NCVS and UCR*, pp. 269-294.

McCulloch, C.E, Searle, S. R., and Neuhaus, J.M. (2008), *Generalized, Linear, and Mixed Models*, John Wiley & Sons, Hoboken, NJ.

Murray, J.D. (2002), *Mathematical Biology: I. An Introduction*, Springer-Verlag, New York.

Prasad, N.G.N., and Rao, J.N.K. (1990), *The Estimation of Mean Squared Error of Small-Area Estimators*, Journal of the American Statistical Association, 85, 163-171.

Rao, J.N.K. (2003), *Small Area Estimation*, John Wiley & Sons, Hoboken, NJ.

Rao, J.N.K. and Yu, M. (1992), "Small Area Estimation Combining Time Series and Cross-Sectional Data," Proceedings of the Survey Research Methods Section, American Statistical Association, pp. 1-9.

______ (1994), "Small Area Estimation by Combining Time Series and Cross-Sectional Data," *Canadian Journal of Statistics*, 22, 511-528.

Truman, J.L. (2011), "Criminal Victimization, 2010," Report NCJ 235508, U.S. Bureau of Justice Statistics, September, 2011.