



**SMALL AREA ESTIMATION WITH STATE SPACE MODELS
SUBJECT TO BENCHMARK CONSTRAINTS**

DANNY PFEFFERMANN, RICHARD TILLER

ABSTRACT

This article shows how to benchmark small area estimators, produced by fitting separate state-space models within the areas, to aggregates of the survey direct estimators within a group of areas. State-space models are used by the U.S. Bureau of Labour Statistics (BLS) for the production of the monthly Employment and Unemployment State estimates. The computation of the benchmarked estimators and their variances is accomplished by incorporating the benchmark constraints within a joint model of the direct estimators in the different areas, which requires the development of a new filtering algorithm for state-space models with correlated measurement errors. No such algorithm has been developed before. The properties and implications of the use of the benchmarked estimators are discussed and illustrated using BLS unemployment series.

The problem of Small Area Estimation is how to produce reliable estimates of area (domain) characteristics, when the sample sizes within the areas are too small to warrant the use of traditional direct survey estimates. This problem is commonly handled by borrowing strength from either neighbouring areas and/or from previous surveys, using appropriate cross-sectional/time series models. In order to protect against possible model breakdowns and for consistency in publication, it is often required to benchmark the area model dependent estimates to the direct survey estimate in a group of areas for which the survey estimate is sufficiently accurate. The latter estimate is a weighted sum of the direct estimates in the areas included in the group, so that the benchmarking process defines another way of borrowing strength across the areas.

**Southampton Statistical Sciences Research Institute
Methodology Working Paper M05/14**

Small Area Estimation with State Space Models Subject to Benchmark Constraints

Danny Pfeffermann and Richard Tiller

ABSTRACT

This article shows how to benchmark small area estimators, produced by fitting separate state-space models within the areas, to aggregates of the survey direct estimators within a group of areas. State-space models are used by the U.S. Bureau of Labor Statistics (BLS) for the production of the monthly Employment and Unemployment State estimates. The computation of the benchmarked estimators and their variances is accomplished by incorporating the benchmark constraints within a joint model of the direct estimators in the different areas, which requires the development of a new filtering algorithm for state-space models with correlated measurement errors. No such algorithm has been developed before. The properties and implications of the use of the benchmarked estimators are discussed and illustrated using BLS unemployment series.

The problem of Small Area Estimation is how to produce reliable estimates of area (domain) characteristics, when the sample sizes within the areas are too small to warrant the use of traditional direct survey estimates. This problem is commonly handled by borrowing strength from either neighboring areas and/or from previous surveys, using appropriate cross-sectional/time series models. In order to protect against possible model breakdowns and for consistency in publication, it is often required to benchmark the area model dependent estimates to the direct survey estimate in a group of areas for which the survey estimate is sufficiently accurate. The latter estimate is a weighted sum of the direct estimates in the areas included in the group, so that the benchmarking process defines another way of borrowing strength across the areas.

Key words: Autocorrelated measurement errors, Generalized least squares,
Recursive filtering, Sampling errors, Unemployment

Danny Pfeffermann is Professor of Statistics at the Hebrew University of Jerusalem, Israel, 91905 (E-mail: msdanny@huji.ac.il), and the Southampton Statistical Sciences Research Institute, University of Southampton, SO17 1BJ, United Kingdom. Richard Tiller is a Senior Mathematical Statistician at the Statistical Methods Division, U.S. Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington DC 20212 (E-mail: Tiller.Richard@bls.gov).

The authors thank the editor, an associate editor and the reviewer for constructive comments and suggestions.

1. INTRODUCTION

The problem considered in this article applies to the U.S. Labor Force estimates but as will become apparent, this problem and the proposed solution are more general with many potential applications. See, in particular, Sections 3 and 4.

The U.S. Bureau of Labor Statistics (BLS) uses state-space time series models for the production of the monthly employment and unemployment estimates in the 50 States and the District of Columbia. The models are fitted to the direct sample estimates obtained from the Current Population Survey (CPS). The use of models is necessary because the CPS State samples are too small to allow producing reliable direct estimates, which is a typical 'small area estimation' problem. The coefficient of variation (CV) of the direct estimates varies from about 8% in the large States to about 16% in the small States. The use of time series models reduces the variances of the estimators very significantly by borrowing information from past estimates, see the illustrations in Section 5. For a recent review of small area estimation methods see Pfeffermann (2002, Section 6 considers the use of time series models). The new book by Rao (2003) provides a methodological account of this topic.

The state-space models are fitted independently for each State and combine a model for the true population values with a model for the sampling errors. The direct survey estimates are the sums of these two unknown components. The published estimates are the differences between the direct estimates and the estimates of the sampling errors, as obtained under the combined model. At the end of each calendar year, the monthly model dependent estimates in any given State are modified so that their annual mean equals the corresponding mean of the direct CPS estimates. The purpose of the benchmarking is to provide protection against possible model failure. This benchmarking procedure has, however, two major disadvantages:

- 1- The mean annual CPS estimates are still unstable because the monthly estimates are highly correlated due to large sample overlaps between different months (see Section 2).
- 2- The benchmarking is retrospective, occurring at the end of each year after that the monthly model dependent estimates have already been published, and hence they provide no protection to real time estimates. (The benchmarked estimates are the input series for trend estimation.)

In this article we study a solution to the benchmarking problem that addresses the two disadvantages of the current BLS procedure. The proposed solution consists of fitting the model jointly to 'homogeneous' groups of States (States with similar 'labor force behavior', see Sections 5 and 6), and adding each month the constraint,

$$\sum_{s=1}^S w_{st} \hat{Y}_{st,model} = \sum_{s=1}^S w_{st} \hat{Y}_{st,cps}, \quad t=1,2,\dots \quad (1.1)$$

where S denotes the number of States in the group. Notice that unlike in classical benchmarking problems that use external (independent) data for the benchmarking process, like census figures or estimates from another large sample, the model dependent estimates in the left hand side of (1.1) are benchmarked to a weighted mean of the direct CPS estimates, which are the input data for the models. External data to which the monthly State estimates can be benchmarked are not available even for isolated months. See, Hillmer and Trabelsi, 1987, Doran, 1992 and Durbin and Quenneville, 1997 for benchmarking procedures to external data sources in the context of state-space modeling.

The justification for incorporating the constraints (1.1) is that the direct CPS estimators, which are unreliable in single States, can be trusted when averaged over different States. Note that the sampling errors of the direct estimates that are highly correlated within a State are independent between States. The basic idea behind the use of these constraints is that if all the direct CPS estimates in the same group jointly increase or decrease due to some sudden external effects that are not accounted for by the model, the benchmarked estimators will reflect this change much faster than the model dependent estimators obtained by fitting separate models in each of the States. This property is illustrated very strikingly in the empirical results presented in Section 5 using actual Unemployment series. Note also that by incorporating the constraints (1.1), the benchmarked estimators in a given month ‘borrow strength’ both from past data and cross-sectionally, unlike the model dependent estimators in current use that only borrow strength from the past. This property is reflected by reduction in the variance of the benchmarked estimators, which again is illustrated in Section 5.

An important question underlying the use of the constraints in (1.1) is the definition of the weights w_{st} ; $s = 1 \dots S$, $t = 1, 2, \dots$. Possible definitions include,

$$w_{1st} = 1/S ; \quad w_{2st} = N_{st} / \sum_{s=1}^S N_{st} ; \quad w_{3st} = [1/Var_{st}(cps)] / \sum_{s=1}^S [1/Var_{st}(cps)] \quad (1.2)$$

where N_{st} and $Var_{st}(cps)$ are respectively the total size of the labor force and the variance of the direct CPS estimate in State s at month t . The use of the weights $\{w_{1st}\}$ is appropriate when the direct estimates are totals; the use of $\{w_{2st}\}$ is appropriate when the direct estimates are proportions. The use of either of these two sets guarantees that the aggregated benchmarked estimate for the group of States in every month t is the same as the corresponding aggregated direct estimate, thus satisfying publication consistency requirements. The use of the weights

$\{w_{3st}\}$ minimizes the variance of the benchmark $\sum_{s=1}^S w_{st} \hat{Y}_{st,cps}$, but it does not satisfy publication consistency.

The proposed solution of combining the individual State models into a joint model with built in benchmark constrains intensifies the computations. The dimension of the state vector in the separate State models is 30 (see next section), implying that by fitting the model jointly to a group of say 12 States, the dimension of the joint state vector would be 360. Fitting state-space models of this size puts heavy demands on CPU time and memory. This creates problems in a production environment where many monthly official estimates have to be produced and published each month soon after that new direct estimates become available. A possible solution to this problem studied in the present article that allows also to compute the variances of the benchmarked estimators (see below) is to include the sampling errors as part of the error terms in the observation (measurement) equation, instead of the current practice of fitting a linear time series model for the sampling errors and including them in the state vector. Implementation of this solution reduces the dimension of each of the separate state vectors by half, because the sampling errors make up 15 elements of the state vector (see next section). As explained in Section 2, including the sampling errors in the observation equation does not change the model.

The use of this solution, however, introduces autocorrelated errors in the measurement equation of the state-space model, since as already mentioned the sampling errors are highly correlated over time. This raises the need of developing a recursive filtering algorithm with good statistical properties for state-space models with autocorrelated measurement errors. While originally motivated by computational considerations, the development of such an algorithm is of more general interest because the common practice of fitting a linear time series model for the measurement errors when they are autocorrelated is not always practical. See next section for the sampling error model approximation used by the BLS. To the best of our knowledge, no such algorithm has been studied previously in the literature. Pfeffermann and Burck (1990) likewise add constraints of the form (1.1) to a state-space model and develop an appropriate recursive filtering algorithm, but their model does not contain autocorrelated sampling errors so that the measurement errors are independent cross-sectionally and over time.

Implementation of the proposed benchmarking procedure, which is the primary focus of this article requires therefore solving three problems:

- 1- Develop a general recursive filtering algorithm for state-space models with correlated measurement errors,
- 2- Incorporate the benchmark constraints and compute the corresponding State benchmarked estimates (estimates of employment or unemployment measures in the present application),
- 3- Compute the variances of the benchmarked estimators.

We emphasize with regard to the third problem that the constraints in (1.1) are only imposed for the computation of the benchmarked estimators, but not when computing the variances of the benchmarked estimators, that account for all the sources of variability, including the errors in the benchmark equations. As explained below, a new filtering algorithm had to be developed in order to compute the correct variances, even if the sampling errors were left in the state vector.

Comment: An alternative simpler way of enforcing the benchmark constraints is by prorating the separate model dependent estimates each month using the equation,

$$\hat{Y}_{st,Pro} = \hat{Y}_{st,model} \left(\sum_{s=1}^S w_{st} \hat{Y}_{st,cps} / \sum_{s=1}^S w_{st} \hat{Y}_{st,model} \right) \quad (1.3)$$

Clearly, the use of (1.3) again satisfies $\sum_{s=1}^S w_{st} \hat{Y}_{st,pro} = \sum_{s=1}^S w_{st} \hat{Y}_{st,cps}$ and it does not require changing the current BLS modeling procedure. This benchmarking method is often used in small area estimation applications that employ cross-sectional models, see, e.g., Fay and Herriot (1977), Battese *et al.* (1988), and Rao (2003). However, the use of (1.3) does not lend itself to parametric estimation of the variances of the prorated estimators. Variance estimation is an essential requirement from any benchmarking procedure. For the case of a single state-space model (with no benchmarking), Pfeiffermann and Tiller (2005) developed bootstrap variance estimators that account for the estimation of the model parameters with bias of correct order, but the extension of this resampling method to the prorated predictors defined by (1.3) is not straightforward, and is in any case much more computing intensive than the solution studied in this article. Another simple benchmarking procedure is a difference adjustment of the form,

$$\hat{Y}_{st,dif} = \hat{Y}_{st,model} + \left(\sum_{s=1}^S w_{st} \hat{Y}_{st,cps} - \sum_{s=1}^S w_{st} \hat{Y}_{st,model} \right) \quad (1.4)$$

However, this procedure inflates the variances of the benchmarked estimators and has the further drawback of adding the same value to each of the model dependent estimators, irrespective of their magnitude.

The benchmarking procedure developed in this paper overcomes the problems mentioned with respect to the procedures (1.3) and (1.4) and is not more complicated once an appropriate computer program is written. Notice in this respect that any other 'built in' benchmarking procedure

would require the development of a new filtering algorithm, even if the sampling errors are left in the state vector. This is so because the benchmark errors that need to be accounted for when computing the variance of the benchmarked estimator are correlated concurrently and over time with the sampling errors. As mentioned before, available algorithms for benchmarking state-space models are not suitable for handling benchmarks of the form (1.1).

Section 2 presents the BLS State models in current use. Section 3 develops the recursive filtering algorithm for state-space models with correlated measurement errors and discusses its properties. Section 4 shows how to incorporate the benchmark constraints and compute the variances of the benchmarked estimators. The application of the proposed procedure is illustrated in Section 5 using Unemployment series in the U.S., with special attention to the year 2001 when the World Trade Center was attacked. We conclude in Section 6 by discussing some remaining problems in the application of this procedure.

2- BLS MODEL IN CURRENT USE

In this section we consider a single State and hence we drop the subscript s from the notation. The model employed by the BLS combines a model for the true State values (total Unemployment or Employment rates) with a model for the sampling errors. The model is discussed in detail, including parameter estimation and model diagnostics in Tiller (1992). Below we provide a brief description of the model for better understanding of the developments and illustrations in subsequent sections.

Let y_t denote the direct CPS estimate at time t and Y_t denote the corresponding population value, so that $e_t = (y_t - Y_t)$ is the sampling error. The model includes a covariate, X_t , see below.

2.1 Model assumed for population values

$$\begin{aligned}
 Y_t &= \beta_t X_t + L_t + S_t + I_t, \quad I_t \sim N(0, \sigma_I^2) \\
 L_t &= L_{t-1} + R_{t-1} + \eta_{L_t}, \quad \eta_{L_t} \sim N(0, \sigma_L^2); \quad R_t = R_{t-1} + \eta_{R_t}, \quad \eta_{R_t} \sim N(0, \sigma_R^2) \\
 \beta_t &= \beta_{t-1} + \eta_{\beta_t}, \quad \eta_{\beta_t} \sim N(0, \sigma_\beta^2) \\
 S_t &= \sum_{j=1}^6 S_{j,t}; \\
 S_{j,t} &= \cos \omega_j S_{j,t-1} + \sin \omega_j S_{j,t-1}^* + v_{j,t}, \quad v_{j,t} \sim N(0, \sigma_S^2) \\
 S_{j,t}^* &= -\sin \omega_j S_{j,t-1} + \cos \omega_j S_{j,t-1}^* + v_{j,t}^*, \quad v_{j,t}^* \sim N(0, \sigma_S^2); \quad \omega_j = 2\pi j/12; \quad j = 1 \dots 6
 \end{aligned} \tag{2.1}$$

The model defined by (2.1) but without the covariate X_t is known in the time series literature as the Basic Structural Model (BSM), with L_t , R_t , S_t and I_t defining respectively the trend level, slope seasonal effect and ‘irregular’ term operating at time t . The error terms $I_t, \eta_{L_t}, \eta_{R_t}, \nu_{j_t}, \nu_{j_t}^*$ are independent white noise series. The model for the trend approximates a local linear trend, whereas the model for the seasonal effect uses the classical decomposition of the seasonal component into 6 subcomponents $S_{j,t}$ that represent the contribution of the cyclical functions corresponding to the 6 frequencies (harmonics) of a monthly seasonal series. The added noise permits the seasonal effects to evolve stochastically over time but in a way that guarantees that the expectation of the sum of 12 successive seasonal effects is 0. See Harvey (1989) for development and thorough discussion of the BSM. The covariate X_t represents the ‘number of persons in the State receiving unemployment insurance benefits’ when modeling the *Unemployment* series, and the ‘ratio between the number of payroll jobs in business establishments and the State population size’ when modeling the *Employment to Population Ratio* series.

2.2 Model assumed for the sampling errors

The CPS sampling error variance varies with the level of the series. Denoting $v_t^2 = Var(e_t)$, the model assumed for the standardized residuals $e_t^* = (e_t / v_t)$ is *AR*(15), which is used as an approximation to the sum of an *MA*(15) process and an *AR*(2) process. The *MA*(15) process accounts for the autocorrelations implied by the sample overlap underlying the CPS sampling design. By this design, households selected to the sample are surveyed for 4 successive months, they are left out of the sample for the next 8 months and then they are surveyed again for 4 more months. The *AR*(2) model accounts for the autocorrelations arising from the fact that households dropped from the survey are replaced by households from the same ‘census tract’. The reduced ARMA representation of the sum of the two processes is ARMA (2,17), which is approximated by an *AR*(15) model.

The separate models holding for the population values and the sampling errors are cast into a single state-space model for the observations y_t (the CPS estimates) of the form,

$$y_t = Z_t \alpha_t^*; \alpha_t^* = T \alpha_{t-1}^* + \eta_t^*, E(\eta_t^*) = 0; E(\eta_t^* \eta_t^{*\prime}) = Q^* \quad (2.2)$$

with the first equation to the left defining the observation (measurement) equation and the second

equation defining the transition (state) equation. The state vector, α_t^* , consists of the covariate coefficient β_t , the trend level L_t , the slope R_t , the 11 seasonal coefficients $S_{j,t}, S_{k,t}^*$, $j = 1 \dots 6$, $k = 1 \dots 5$, the irregular term I_t and the concurrent and 14 lags of the sampling errors, a total of 30 elements. The sampling errors and the irregular term are included in the state vector so that there are no error terms in the observation equation. Notice in this respect that including the sampling errors in the observation equation as discussed in the introduction and pursued later instead of the present practice of including them in the state equation does not change the model holding for the observed direct estimates as long as the model fitted for the sampling errors reproduces their autocovariances.

The parameters indexing the model are estimated separately for each State in two steps; first the AR(15) coefficients and residual variance are estimated by solving the corresponding Yule-Walker equations (the sampling error variance and autocorrelations are estimated externally by the BLS), then the remaining model variances are estimated by maximizing the model likelihood, with the AR(15) model parameters held fixed at their estimated values. See Tiller (1982) for details.

The monthly Employment and Unemployment State estimates published by the BLS are obtained under the model (2.1) and the relationship $Y_t = (y_t - e_t)$ as,

$$\hat{Y}_t = (y_t - \hat{e}_t) = \hat{\beta}_t X_t + \hat{L}_t + \hat{S}_t + \hat{I}_t \quad (2.3)$$

with $\hat{\beta}_t, \hat{L}_t, \hat{S}_t, \hat{I}_t$ denoting the estimated components at time t , as obtained by application of the Kalman filter (Harvey, 1989). Notice that unlike classical Small Area models that assume common regression slopes in all the areas, different estimates $\hat{\beta}_t$ are used for different States.

3. FILTERING OF STATE-SPACE MODELS WITH AUTOCORRELATED MEASUREMENT ERRORS

In this section we develop a recursive filtering algorithm for state-space models with autocorrelated measurement errors. By a recursive filtering algorithm we mean an algorithm that updates the most recent predictor of the state vector every time that a new observation becomes available. This filter is required for implementing the benchmarking procedure discussed in the introduction (see Section 4), but as mentioned before, it is general and can be used for other applications of state-space models with autocorrelated measurement errors. In Section 3.2 we discuss the properties of the proposed filter.

3.1 Recursive Filtering algorithm

Consider the following linear state-space model for (possibly vector) times series y_t ,

$$\text{Observation equation: } y_t = Z_t \alpha_t + e_t; \quad E(e_t) = 0, \quad E(e_t e_t') = \Sigma_{tt} \quad E(e_t e_\tau') = \Sigma_{t\tau} \quad (3.1a)$$

$$\text{Transition equation: } \alpha_t = T \alpha_{t-1} + \eta_t; \quad E(\eta_t) = 0, \quad E(\eta_t \eta_t') = Q \quad E(\eta_t \eta_{t-k}') = 0, \quad k > 0 \quad (3.1b)$$

It is also assumed that $E(\eta_t e_\tau') = 0$ for all t and τ . (The restriction to time invariant matrices T and Q is for convenience. Extension of the filter to models that contain fixed effects in the observation equation is straightforward.) Clearly, what distinguishes this model from the standard linear state-space model is that the measurement errors, e_t (the sampling errors in the BLS model) are correlated over time. Notice, in particular, that unlike the BLS model representation in (2.2) where the sampling errors and the irregular term are part of the state vector so that there are no measurement errors in the observation equation, the measurement (sampling) errors feature now in the observation equation. The recursive filtering algorithm developed below takes account of the autocovariance matrices $\Sigma_{\tau t}$.

At time 1

Let $\hat{\alpha}_1 = (I - K_1 Z_1) T \hat{\alpha}_0 + K_1 y_1$ be the filtered (predicted) state estimator at time 1, where $\hat{\alpha}_0$ is an initial estimator with covariance matrix $P_0 = E[(\hat{\alpha}_0 - \alpha_0)(\hat{\alpha}_0 - \alpha_0)']$ and $K_1 = P_{10} Z_1' F_1^{-1}$ is the 'Kalman gain', with $P_{10} = T P_0 T' + Q$ and $F_1 = Z_1 P_{10} Z_1' + \Sigma_{11}$. We assume for convenience that $\hat{\alpha}_0$ is independent of the observations. The matrix P_{10} is the covariance matrix of the prediction errors $(T \hat{\alpha}_0 - \alpha_1) = (\hat{\alpha}_{10} - \alpha_1)$ and F_1 is the covariance matrix of the innovations (one step ahead prediction errors) $v_1 = (y_1 - \hat{y}_{10}) = (y_1 - Z_1 \hat{\alpha}_{10})$. Since, $y_1 = Z_1 \alpha_1 + e_1$,

$$\hat{\alpha}_1 = (I - K_1 Z_1) T \hat{\alpha}_0 + K_1 Z_1 \alpha_1 + K_1 e_1 \quad (3.2)$$

At time 2

Let $\hat{\alpha}_{21} = T \hat{\alpha}_1$ define the predictor of α_2 at time 1 with covariance matrix $P_{21} = E[(\hat{\alpha}_{21} - \alpha_2)(\hat{\alpha}_{21} - \alpha_2)']$. An unbiased predictor $\hat{\alpha}_2$ of α_2 , [i.e., $E(\hat{\alpha}_2 - \alpha_2) = 0$], based on $\hat{\alpha}_{21}$ and y_2 is the Generalized Least Square (GLS) predictor in the regression model,

$$\begin{pmatrix} T\hat{\alpha}_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} I \\ Z_2 \end{pmatrix} \alpha_2 + \begin{pmatrix} u_{21} \\ e_2 \end{pmatrix}, \quad (u_{21} = T\hat{\alpha}_1 - \alpha_2) \quad (3.3)$$

that is,

$$\hat{\alpha}_2 = \left[(I, Z_2') V_2^{-1} \begin{pmatrix} I \\ Z_2 \end{pmatrix} \right]^{-1} (I, Z_2') V_2^{-1} \begin{pmatrix} T\hat{\alpha}_1 \\ y_2 \end{pmatrix} \quad (3.4)$$

where

$$V_2 = \text{Var} \begin{pmatrix} u_{21} \\ e_2 \end{pmatrix} = \begin{bmatrix} P_{21} & C_2 \\ C_2' & \Sigma_{22} \end{bmatrix} \quad (3.5)$$

and $C_2 = \text{Cov}[u_{21}, e_2] = TK_1 \Sigma_{12}$ (follows from (3.2)). Notice that V_2 is the covariance matrix of the errors u_{21} and e_2 , and not of the predictors $T\hat{\alpha}_1$ and y_2 . As discussed below and proved in Appendix A, the GLS predictor $\hat{\alpha}_2$ is the best linear unbiased predictor (BLUP) of α_2 based on $T\hat{\alpha}_1$ and y_2 , with covariance matrix,

$$E[(\hat{\alpha}_2 - \alpha_2)(\hat{\alpha}_2 - \alpha_2)'] = \left[(I, Z_2') V_2^{-1} \begin{pmatrix} I \\ Z_2 \end{pmatrix} \right]^{-1} = P_2 \quad (3.6)$$

At time t

Let $\hat{\alpha}_{t|t-1} = T\hat{\alpha}_{t-1}$ define the predictor of α_t at time $(t-1)$ with covariance matrix $E[(\hat{\alpha}_{t|t-1} - \alpha_t)(\hat{\alpha}_{t|t-1} - \alpha_t)'] = TP_{t-1}T' + Q = P_{t|t-1}$, where $P_{t-1} = E[(\hat{\alpha}_{t-1} - \alpha_{t-1})(\hat{\alpha}_{t-1} - \alpha_{t-1})']$. Set the random coefficients regression model,

$$\begin{pmatrix} T\hat{\alpha}_{t-1} \\ y_t \end{pmatrix} = \begin{pmatrix} I \\ Z_t \end{pmatrix} \alpha_t + \begin{pmatrix} u_{t|t-1} \\ e_t \end{pmatrix}, \quad (u_{t|t-1} = T\hat{\alpha}_{t-1} - \alpha_t) \quad (3.7)$$

and define

$$V_t = \text{Var} \begin{pmatrix} u_{t|t-1} \\ e_t \end{pmatrix} = \begin{bmatrix} P_{t|t-1} & C_t \\ C_t' & \Sigma_{tt} \end{bmatrix} \quad (3.8)$$

The covariance matrix $C_t = \text{Cov}[T\hat{\alpha}_{t-1} - \alpha_t, e_t]$ is computed as follows: Let $[I, Z_j'] V_j^{-1} = [B_{j1}, B_{j2}]$ where B_{j1} contains the first q columns of $[I, Z_j'] V_j^{-1}$ and B_{j2} the remaining columns, with $q = \dim(\alpha_j)$. Define, $A_j = TP_j B_{j1}$, $\tilde{A}_j = TP_j B_{j2}$, $j = 2 \dots (t-1)$; $\tilde{A}_1 = TK_1$. Then,

$$C_t = \text{Cov}[T\hat{\alpha}_{t-1} - \alpha_t, e_t] = A_{t-1} A_{t-2} \dots A_2 \tilde{A}_1 \Sigma_{1t} + A_{t-1} A_{t-2} \dots A_3 \tilde{A}_2 \Sigma_{2t} + \dots + A_{t-1} \tilde{A}_{t-2} \Sigma_{t-2,t} + \tilde{A}_{t-1} \Sigma_{t-1,t} \quad (3.9)$$

The GLS predictor of α_t based on $T\hat{\alpha}_{t-1}$ and y_t , and the covariance matrix of the prediction errors are obtained from (3.7)-(3.8) as,

$$\hat{\alpha}_t = \left[(I, Z_t') V_t^{-1} \begin{pmatrix} I \\ Z_t \end{pmatrix} \right]^{-1} (I, Z_t') V_t^{-1} \begin{pmatrix} T\hat{\alpha}_{t-1} \\ y_t \end{pmatrix}; P_t = E[(\hat{\alpha}_t - \alpha_t)(\hat{\alpha}_t - \alpha_t)'] = \left[(I, Z_t') V_t^{-1} \begin{pmatrix} I \\ Z_t \end{pmatrix} \right]^{-1} \quad (3.10)$$

The predictor $\hat{\alpha}_t$ can be written alternatively (see Appendix C) as,

$$\hat{\alpha}_t = T\hat{\alpha}_{t-1} + (P_{t|t-1} Z_t' - C_t) [Z_t' P_{t|t-1} Z_t - Z_t' C_t - C_t' Z_t + \Sigma_{tt}]^{-1} (y_t - Z_t T\hat{\alpha}_{t-1}) \quad (3.11)$$

Written this way, the predictor of α_t at time t is seen to equal the predictor of α_t at time $(t-1)$ plus a correction factor that depends on the magnitude of the innovation (one step ahead prediction error) when predicting y_t at time $(t-1)$.

3.2 Properties of the filtering algorithm

Assuming known model parameters, the recursive GLS filter defined by (3.10) or (3.11) has the following properties:

1- At every time point t , the filter produces the 'best linear unbiased predictor' (BLUP) of α_t based on the predictor $\hat{\alpha}_{t|t-1} = T\hat{\alpha}_{t-1}$ from time $(t-1)$ and the new observation y_t . The BLUP property means that $E(\hat{\alpha}_t - \alpha_t) = 0$ and $Var[d'(\hat{\alpha}_t - \alpha_t)] \leq Var[d'(\hat{\alpha}_t^L - \alpha_t)]$ for every vector coefficient d' and any other linear unbiased predictor of the form $\hat{\alpha}_t^L = L_1 \hat{\alpha}_{t|t-1} + L_2 y_t + l$, with general matrices L_1, L_2 and vector l . See Appendix A for proof of this property.

2- When the measurement errors are independent, the GLS filter algorithm coincides with the familiar Kalman filter; see Appendix B for proof. Thus, the recursive GLS filter can be viewed as an extension of the recursive Kalman filter for the case of correlated measurement errors.

3- Unlike the Kalman filter that assumes independent measurement errors and yields therefore the BLUP of α_t based on all the individual observations $y_{(t)} = (y_1 \dots y_t)'$, i.e., the BLUP out of all the linear unbiased predictors of the form $\sum_{i=1}^t C_{it} y_i + c_t$ for general matrices C_{it} and vector c_t , the filter (3.11) yields the BLUP of α_t based on $\hat{\alpha}_{t|t-1}$ and the new observation y_t . (The predictor

$\hat{\alpha}_{t|t-1}$ is itself a linear combination of all the observations until time $(t-1)$, but the values of the matrix coefficients are fixed by the previous steps of the filter, see also the comment below).

Computation of the BLUP of α_t under correlated measurement errors out of all possible unbiased predictors that are linear combinations of the individual observations $y_{(t)}$, (or the minimum mean square error predictor, $E(\alpha_t | y_t, y_{t-1}, \dots, y_1; \lambda)$, under normality of the model error terms, where λ defines the model parameters), requires joint modeling of $y_{(t)}$ for every time t . For long series the computations become very heavy and generally not practical in a production environment that requires routine runs of many series with high dimensional state vectors in a short time. Empirical evidence so far suggests that the loss in efficiency from using the GLS algorithm instead of the BLUP based on all the individual observations is mild. See Section 5 for empirical illustrations.

Comment: For general covariance matrices $\Sigma_{\tau t}$ between the measurement errors, it is impossible to construct a recursive filtering algorithm that is a linear combination of the predictor from the previous time point and the new observation and is BLUP out of all unbiased predictors of the form $\sum_{i=1}^t C_{it} y_i + c_t$. To see this, consider a simple case of 3 observations y_1, y_2, y_3 with common mean μ and variance σ^2 . If the three observations are independent, the BLUP of μ based on the first 2 observations is $\bar{y}_{(2)} = (y_1 + y_2)/2$, and the BLUP based on the three observations is $\bar{y}_{(3)} = (y_1 + y_2 + y_3)/3 = (2/3)\bar{y}_{(2)} + (1/3)y_3$. The BLUP $\bar{y}_{(3)}$ is the Kalman filter predictor for time 3. Suppose, however, that $Cov(y_1, y_2) = Cov(y_2, y_3) = \sigma^2 \rho_1$ and $Cov(y_1, y_3) = \sigma^2 \rho_2 \neq \sigma^2 \rho_1$. The BLUP of μ based on the first 2 observations is again $\bar{y}_{(2)} = (y_1 + y_2)/2$, but the BLUP of μ based on the 3 observations is in this case $\bar{y}_{(3)}^{opt} = ay_1 + by_2 + cy_3$ where $a = (1 - \rho_1)/(3 - 4\rho_1 + \rho_2)$ and $b = (1 - 2\rho_1 + \rho_2)/(3 - 4\rho_1 + \rho_2)$. Clearly, since $a \neq b$, the predictor $\bar{y}_{(3)}^{opt}$ cannot be written as a linear combination of $\bar{y}_{(2)}$ and y_3 . For example, if $\rho_1 = 0.5, \rho_2 = 0.25 \Rightarrow \bar{y}_{(3)}^{opt} = 0.4y_1 + 0.2y_2 + 0.4y_3$.

4. INCORPORATING THE BENCHMARK CONSTRAINTS

4.1 Joint modeling of several concurrent estimates and their weighted mean

In this section we model jointly the concurrent observations (estimates) for S series (States) and their weighted mean. In Section 4.2 we show how to incorporate the benchmark constraints and compute the variances of the benchmarked predictors. We follow the BLS modeling practice and assume that the state vectors and the measurement errors are independent between the series. Section 6 considers extensions of the joint model that allow for cross-sectional correlations between corresponding components of the state vectors operating in different series.

Suppose that the models underlying the S series are as in (3.1) with the (correlated) measurement errors included in the observation equation. Below we add the subscript s to all the model components in order to distinguish between the series. The observations y_{st} (the CPS estimates in the BLS models) and the measurement errors e_{st} (the sampling errors in the BLS model) are scalars and Z_{st} is a row vector (denoted hereafter, z_{st}). Let

$\tilde{y}_t = (y_{1t} \dots y_{St}, \sum_{s=1}^S w_{st} y_{st})'$ define the concurrent observations in the S series and their weighted mean (the right hand side of the benchmark equation 1.1). The corresponding vector of measurement errors is, $\tilde{e}_t = (e_{1t} \dots e_{St}, \sum_{s=1}^S w_{st} e_{st})'$. Let $Z_t^* = I_S \otimes z_{st}$ (a block diagonal matrix with z_{st} as the s^{th} block), $\tilde{T}_t = I_S \otimes T$, $\tilde{Z}_t = \begin{bmatrix} Z_t^* \\ w_{1t} z_{1t}, \dots, w_{St} z_{St} \end{bmatrix}$, $\tilde{\alpha}_t = (\alpha_{1t}, \dots, \alpha_{St})'$ and $\tilde{\eta}_t = (\eta_{1t}, \dots, \eta_{St})'$.

By (3.1) and the independence of the state vectors and measurement errors between the series, the joint model holding for \tilde{y}_t is,

$$\tilde{y}_t = \tilde{Z}_t \tilde{\alpha}_t + \tilde{e}_t ; E(\tilde{e}_t) = 0 , E(\tilde{e}_t \tilde{e}_t') = \tilde{\Sigma}_{\tau t} = \begin{bmatrix} \Sigma_{\tau t} & h_{\tau t} \\ h_{\tau t}' & v_{\tau t} \end{bmatrix} \quad (4.1a)$$

$$\tilde{\alpha}_t = \tilde{T} \tilde{\alpha}_{t-1} + \tilde{\eta}_t ; E(\tilde{\eta}_t) = 0 , E(\tilde{\eta}_t \tilde{\eta}_t') = I_S \otimes Q_{st} = \tilde{Q} , E(\tilde{\eta}_t \tilde{\eta}_{t-k}') = 0 , k > 0 \quad (4.1b)$$

$$\Sigma_{\tau t} = \text{Diag}[\sigma_{1\tau t} \dots \sigma_{S\tau t}] ; \sigma_{s\tau t} = \text{Cov}[e_{s\tau}, e_{st}] , v_{\tau t} = \sum_{s=1}^S w_{s\tau} w_{st} \sigma_{s\tau t} = \text{Cov}[\sum_{s=1}^S w_{s\tau} e_{s\tau}, \sum_{s=1}^S w_{st} e_{st}]$$

$$h_{\tau t} = (h_{1\tau t} \dots h_{S\tau t})' ; h_{s\tau t} = w_{st} \sigma_{s\tau t} = \text{Cov}[e_{s\tau}, \sum_{s=1}^S w_{st} e_{st}]$$

Comment: The model (4.1) is the same as the separate models defined by (3.1) for the different States. Adding the model for $\sum_{s=1}^S w_{st} y_{st}$ to the observation equation provides no new information.

4.2 Recursive filtering with correlated measurements and benchmark constraints

In order to compute the benchmarked predictors we apply the recursive GLS filter (3.11) to the joint model defined by (4.1), setting the variance of the benchmarked errors, $\sum_{s=1}^S w_{st} e_{st}$ to zero. The idea behind this procedure is as follows. By (4.1), the model dependent predictor of the signal (true population value) of series s at time t takes the form, $\hat{Y}_{st, \text{model}} = z_{st}' \hat{\alpha}_{st}$. (See 2.3 for the predictor under the BLS model.) Thus, the benchmark constraints (1.1) can be written as,

$$\sum_{s=1}^S w_{st} z_{st}' \hat{\alpha}_{st} = \sum_{s=1}^S w_{st} y_{st} \quad , \quad t=1,2,\dots \quad (4.2)$$

Since, in fact, $\sum_{s=1}^S w_{st} y_{st} = \sum_{s=1}^S w_{st} z_{st}' \alpha_{st} + \sum_{s=1}^S w_{st} e_{st}$, a simple way of satisfying the benchmark constraints is by imposing $\sum_{s=1}^S w_{st} y_{st} = \sum_{s=1}^S w_{st} z_{st}' \alpha_{st}$, or equivalently, by setting

$$\text{Var}[\sum_{s=1}^S w_{st} e_{st}] = \text{Cov}[e_{st}, \sum_{s=1}^S w_{st} e_{st}] = 0; \quad s=1\dots S, \quad t=1,2,\dots \quad (4.3)$$

Notice that the use of (4.3) is just a convenient technical way of forcing the constraints and hence computing the benchmarked predictors. In Appendix D we show how to compute the variances of the benchmarked predictors, accounting for the errors of the constraints (and no longer imposing 4.3). The imposition of (4.3) in the GLS filter (3.11) is implemented by replacing the covariance matrix $\tilde{\Sigma}_t$ of the observation equation (4.1a) by the matrix $\tilde{\Sigma}_t^* = \begin{bmatrix} \Sigma_{tt} & , & 0_{(S)} \\ 0_{(S)} & , & 0 \end{bmatrix}$, where $0_{(S)}$ is the null vector, and setting the last column of the covariance matrix $C_t^{bmk} = \text{Cov}[\tilde{T} \tilde{\alpha}_{t-1}^{bmk} - \alpha_t, \tilde{e}_t]$ to zero.

Application of the algorithm (3.11) to the model (4.1), with the benchmark constraints imposed by (4.3), yields the benchmarked predictor,

$$\tilde{\alpha}_t^{bmk} = \tilde{T} \tilde{\alpha}_{t-1}^{bmk} + (P_{t|t-1}^{bmk} \tilde{Z}_t' - C_{t,0}^{bmk}) \left[\tilde{Z}_t P_{t|t-1}^{bmk} \tilde{Z}_t' - \tilde{Z}_t C_{t,0}^{bmk} - C_{t,0}^{bmk} \tilde{Z}_t' + \tilde{\Sigma}_t^* \right]^{-1} (\tilde{y}_t - \tilde{Z}_t \tilde{T} \tilde{\alpha}_{t-1}^{bmk}) \quad (4.4)$$

where, $P_{t|t-1}^{bmk} = \tilde{T}P_{t-1}^{bmk}\tilde{T}' + \tilde{Q}$; $P_{t-1}^{bmk} = E[(\tilde{\alpha}_{t-1} - \tilde{\alpha}_{t-1}^{bmk})(\tilde{\alpha}_{t-1} - \tilde{\alpha}_{t-1}^{bmk})']$, and by defining $\tilde{e}_{t,0} = (e_{1t} \dots e_{St}, 0)'$, $C_{t,0}^{bmk} = Cov[\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \alpha_t, \tilde{e}_{t,0}]$.

Comment: The matrix $P_{t|t-1}^{bmk} = E[(\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t)(\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t)']$ is the true prediction error covariance matrix. See Appendix D for the computation of $P_t^{bmk} = E[(\tilde{\alpha}_t - \tilde{\alpha}_t^{bmk})(\tilde{\alpha}_t - \tilde{\alpha}_t^{bmk})']$ and $C_t^{bmk} = Cov[\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \alpha_t, \tilde{e}_t]$ under the model (4.1) without the constraints. The matrix P_t^{bmk} accounts for the variability of the state vector components and the variances and covariances of the sampling errors (defining the matrices $\tilde{\Sigma}_t$ and C_t^{bmk}).

It should be emphasized again that the benchmarking filter developed in the present article, and in particular the covariance matrix P_t^{bmk} is different from the state-space benchmarking filters developed in other articles, where the series observations are benchmarked to external (independent) data sources. For example, Doran (1992) considers benchmark constraints (possibly a set) of the form $R_t\tilde{\alpha}_t = r_t$, where the r_t 's are constants. In our case the benchmarks $\sum_{s=1}^S w_{st} y_{st}$ are random and depend heavily on the sum $\sum_{s=1}^S w_{st} z_{st} \hat{\alpha}_{st}$. Another difference between the present filter and the other filters developed in the context of state-space modeling is the accounting for correlated measurement errors in the present filter.

5. EMPIRICAL ILLUSTRATIONS

We start by illustrating that the proposed benchmarking algorithm with correlated measurement errors performs properly. For this, we simulated 10,000 series of length 45 for each of 3 models and computed the empirical variances of the benchmark prediction errors $(\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t)$, which we compare to the theoretical variances under the model, $P_t^{bmk} = E[(\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t)(\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t)']$, defined by (D.3) (Appendix D). We also compare the empirical covariances, $Cov[\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \alpha_t, \tilde{e}_t]$ with the theoretical covariances C_t^{bmk} under the model as defined by (D.4). The models used for generating the three sets of series have the general form,

$$\begin{aligned} y_{st} &= Y_{st} + e_{st}; Y_{st} = Y_{s,t-1} + \eta_{st} \\ e_{st} &= \varepsilon_{st} + 0.55\varepsilon_{s,t-1} + 0.30\varepsilon_{s,t-2} + 0.10\varepsilon_{s,t-3}; s = 1, 2, 3 \end{aligned} \quad (5.1)$$

where $\{\eta_{st}\}$ and $\{\varepsilon_{st}\}$ are independent white noise series. The random walk variances for the three models are, $Var(\eta_{st}) = (0.01, 0.88, 1.2)$ respectively. The corresponding measurement error variances are, $Var(e_{st}) = (0.30, 0.08, 1.21)$, with autocorrelations $Corr(e_t, e_{t-1}) = 0.53$, $Corr(e_t, e_{t-2}) = 0.25$, $Corr(e_t, e_{t-3}) = 0.07$ (same autocorrelations for the three models). The benchmark constraints are,

$$\hat{Y}_{1t} + \hat{Y}_{2t} + \hat{Y}_{3t} = y_{1t} + y_{2t} + y_{3t}, t = 1 \dots 45. \quad (5.2)$$

Table 1 shows the theoretical and empirical variances and covariances (over the 10,000 replications) for the last time point ($t=45$).

Table 1 about here

The results presented in Table 1 show a close fit even for the third model where both the population value variance and the measurement error variance are relatively high. These results can be viewed as validation of the theoretical expressions.

The remaining empirical results presented below refer to the monthly estimates of total unemployment in the 9 Census divisions of the U.S.A. for the period January 1998-December 2003. Very similar results and conclusions are obtained when analyzing the corresponding employment estimates. As described in the Introduction, the direct CPS estimates are fitted by the BLS using the model (2.2), yielding the model dependent predictors defined by (2.3). The year 2001 is of special interest for illustration since it is affected by the start of a recession in March and the attack on the World Trade Center in September. These two events provide a good test for the performance of the proposed benchmarking procedure.

It is mentioned in Section 3 that the loss in efficiency from using the recursive GLS filter (3.11) instead of the optimal predictors based on all the individual observations is mild. Table 2 shows for each division the means and standard deviations (STD) of the monthly ratios between the STD of the GLS and the STD of the optimal predictor when predicting the total unemployment Y_t and the Trend levels L_t (Equation 2.1). As can be seen, the largest loss in efficiency, when measured by the means is 3%.

Table 2 about here

Next we combined the individual Division models into the joint model (4.1). The benchmark constraints are as defined in (1.1) with $w_{st} = 1$, such that the model dependent predictors of the

Divisions' total unemployment are benchmarked to the total national unemployment. The CV of the CPS estimator of total national unemployment is 2%, which is viewed as sufficiently precise.

Figure 1 compares the sum of the model dependent predictors over the 9 Divisions without benchmarking, with the CPS national unemployment estimates. In the first part of the observation period the sums of the model predictors are close to the corresponding CPS estimates. In 2001, however, there is evidence of systematic model underestimation which, as explained above, results from by the start of a recession in March and the attack on the World Trade Center in September. The bias of the model dependent predictors is further highlighted in Figure 2, which plots the differences between the two sets of estimators. As can be seen, for a period of about one year starting in March 2001, all the differences are negative and in some months the absolute difference is larger than twice the standard deviation of the CPS estimator.

Figures 3-5 show the direct CPS estimators, the unbenchmarked predictors and the benchmarked predictors for three out of the 9 census divisions. We restrict the graph to the period of 1.2000-1.2003, so as to better illuminate how benchmarking corrects in real time the underestimation of the unbenchmarked predictors in the year 2001. Similar corrections are observed for the other divisions except for New England (not shown), where the benchmarked predictors have a slightly larger positive bias than the unbenchmarked predictors. This is explained by the fact that unlike in the other eight divisions, in this division the unbenchmarked predictors in 2001 are actually higher than the CPS estimators, which serves as an excellent example of the need to apply the benchmarking in 'homogeneous groups' (see the concluding remarks in Section 6).

Another important conclusion reached from Figures 3-5 is that imposing the benchmark constraints in regular periods when they are not really needed affects very mildly the predictors. This is expected since in regular periods the benchmark constraints are approximately satisfied under the correct model even without imposing them directly. In fact, the degree to which the unbenchmarked predictors satisfy the constraints can be viewed as a model diagnostic tool. In order to illustrate this point further, we show in Table 3 for each of the 9 census divisions the means and STDs of the monthly ratios between the benchmarked predictor and the unbenchmarked predictor, separately for 1998-2003 excluding 2001, and for 2001. As can be seen, in 2001 some of the mean ratios are about 4% but in the other years the means never exceed 1%, showing that in normal times the effect of benchmarking on the separate model dependent predictors is indeed very mild.

Table 3 about here

Finally, we mentioned in the introduction that by imposing the benchmark constraints the predictor in any given area “borrows strength” from other areas. This can be illustrated by comparing the STDs of the benchmarked predictors with the STDs of the unbenchmarked predictors. Figures 6-8 show the two sets of STDs, along with the STDs of the direct CPS estimators, for the same three divisions as in Figures 3-5. Notice that the STDs of the CPS estimators are with respect to the distribution of the sampling errors over repeated sampling, whereas the STDs of the two other predictors account also for the variability of the state vector components. As can be seen, the STDs of the benchmarked predictors are somewhat lower than the STDs of the unbenchmarked predictors for all the months, and both sets of STDs are significantly lower than the STDs of the corresponding CPS estimators. This pattern repeats itself in all the divisions. The STDs of the benchmarked and unbenchmarked predictors are further compared in Table 4. The mean ratios in Table 4 show gains in efficiency of up to 15% in some of the divisions by use of benchmarking. The ratios are very stable throughout the years despite the fact that the STDs of the two sets of model predictors change between months due to changes in the STDs of the sampling errors. See also Figures 6-8.

Table 4 about here

6. CONCLUDING REMARKS, OUTLINE OF FUTURE RESEARCH

Agreement of small area model dependent estimators with the direct sample estimate in a ‘large area’ is a common requirement by statistical agencies producing official statistics. See, for example, Battese *et al.* (1989), Pfeiffermann and Barnard (1991) and Rao (2003) for proposed modifications to meet this requirement. This article shows how this requirement can be implemented with the use of state-space models. As emphasized in the article, benchmarking constraints of the form (1.1) cannot be incorporated by use of the standard Kalman filter, requiring instead the development of a filter that produces the correct variances of the benchmarked estimators under the model. The filter developed in this article has the additional property of addressing situations where the measurement errors are correlated over time. The GLS filter developed for the case of correlated measurement errors (without incorporating the benchmark constraints) produces the BLUP of the state vector at time t out of all the predictors that are linear combinations of the state predictor from time $(t-1)$ and the new observation at time t , but generally not the BLUP among all the predictors that are linear combinations of past and present observations. Nonetheless, empirical evidence presented in Section 5 illustrates that the loss of efficiency from using the GLS filter is mild. When the measurement errors are independent in time,

the GLS filter is the same as the Kalman filter. Another important property illustrated in Section 5 is that by joint modeling a large number of areas and incorporating the benchmark constraints, the benchmarked predictor in any given area borrows strength from other areas, resulting in reduced variance, which is not possible when fitting the model independently in the various areas.

An important condition for the success of the benchmarking procedure is that the small areas (States in the present application) are 'homogeneous' with respect of the behavior of the true (estimated) quantities of interest (the true Employment or Unemployment values in the population in the present application). The need to satisfy this condition is illuminated in the empirical illustrations in Section 5 where the benchmarking of the census division predictors to the direct CPS national estimate induced a small positive bias in the division of New England. This happened because unlike in all the other divisions, the model dependent predictors in New England were already higher than the corresponding CPS estimators. Since benchmarking of the employment and unemployment estimates in the U.S.A. will be implemented by the BLS for the State estimates, our next immediate task is to classify the 50 States and the District of Columbia into homogeneous groups. Several factors need to be taken into account when defining the groups. Geographic proximity accounting for weather conditions and other environmental effects, and breakdown of the Labor Force into the major categories of employment (percentages employed in manufacturing, services, farming etc.) are important factors that will be considered. The classification of the States will reflect also the behavior of past estimates and their components, like the trend and seasonal effects. Accounting for all the factors mentioned above for the grouping process may end up in small groups, but it should be emphasized that the groups must be sufficiently large to justify the benchmarking to the corresponding aggregate CPS estimate in the group. Thus, the sensitivity of the benchmarking to the definition of the groups needs to be investigated.

Another topic that is currently under final stages of investigation is the development of a smoothing algorithm that accounts for correlated measurement errors and incorporates the benchmarking constraints. Clearly, as new data accumulate it is desirable to modify past predictors, which is particularly important for trend estimation. An appropriate 'fixed point' smoother has already been designed and is presently tested. This smoother augments the state vector at every time $t \geq d$ by the state vector corresponding to month d for which the smoothed predictor has to be computed. See Harvey (1989) for the computation of the fixed point smoother under a state-space model with time independent measurement errors.

Finally, the present BLS models assume independence between the state vectors operating in different States. It seems plausible that changes in the trend or seasonal effects are correlated between homogeneous States and accounting for these correlations might improve further the efficiency of the predictors. In fact, the existence of such correlations underlies implicitly the use of the proposed benchmarking procedure. Accounting explicitly for the existing correlations is simple within the joint model defined by (4.1) and may reduce (but not eliminate) the need for the benchmark process.

REFERENCES

- Battese, G. E., Harter, R.M., and Fuller, W. A. (1988), "An Error Component Model for Prediction of County Crop Areas using Survey and Satellite Data," *Journal of the American Statistical Association*, 83, 28-36.
- Doran, H. E. (1992), "Constraining Kalman Filter and Smoothing Estimates to Satisfy Time-Varying Restrictions," *The Review of Economics and Statistics*, 74, 568-572.
- Durbin, J., and Quenneville, B. (1997), "Benchmarking by State Space Models," *International Statistical Review*, 65, 23-48.
- Fay, R. E. and Herriot, R. (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.
- Harvey, A. C. (1989), *Forecasting Structural Time Series with the Kalman Filter*. Cambridge: Cambridge University Press.
- Hillmer, S. C., and Trabelsi, A. (1987), "Benchmarking of Economic Time Series," *Journal of the American Statistical Association*, 82, 1064-1071.
- Pfeffermann, D. (1984), "On Extensions of the Gauss-Markov Theorem to the Case of Stochastic Regression Coefficients. *Journal of the Royal Statistical Society*, Ser. B, 46, 139-148.
- Pfeffermann, D., and Burck, L. (1990), "Robust Small Area Estimation Combining Time Series and Cross-Sectional Data," *Survey Methodology*, 16, 217-237.
- Pfeffermann, D., and Barnard, C.H. (1991), "Some New Estimators for Small Area Means with Application to the Assessment of Farmland Values," *Journal of Business & Economic Statistics*, 9, 73-83.
- Pfeffermann, D. (2002), "Small Area Estimation- New Developments and Directions," *International Statistical Review*, 70, 125-143.
- Pfeffermann, D., and Tiller, R. B. (2005), "Bootstrap Approximation to Prediction MSE for State-Space Models with Estimated Parameters," *Journal of Time series Analysis* (Forthcoming).
- Rao, J. N. K. (2003), *Small Area Estimation*. New York: Wiley.
- Tiller, R. B. (1992), "Time Series Modeling of Sample Survey Data from the U.S. Current Population Survey," *Journal of Official Statistics*, 8, 149-166.

APPENDIX A. Proof of BLUP property of the GLS predictor $\hat{\alpha}_t$ in (3.10)

The model holding for $Y_t = (\hat{\alpha}_{t|t-1}, y_t)'$ is,

$$Y_t = \begin{bmatrix} \hat{\alpha}_{t|t-1} \\ y_t \end{bmatrix} = \begin{bmatrix} I \\ z_t \end{bmatrix} \alpha_t + \begin{bmatrix} u_{t|t-1} \\ e_t \end{bmatrix}; u_{t|t-1} = \hat{\alpha}_{t|t-1} - \alpha_t$$

where $\hat{\alpha}_{t|t-1}$ is the predictor of α_t from time $t-1$. Denote $u_t = (u_{t|t-1}, e_t)'$, $X_t = [I, z_t]'$. In what follows all the expectations are over the joint distribution of Y_t and α_t , so that $E(u_t) = 0$; $E(u_t u_t') = V_t$ (Equation (3.8)). A predictor α_t^* is unbiased for α_t if $E(\alpha_t^* - \alpha_t) = 0$.

Theorem: The predictor $\hat{\alpha}_t = (X_t' V_t^{-1} X_t)^{-1} X_t' V_t^{-1} Y_t = Q_t^{-1} X_t' V_t^{-1} Y_t$ is the Best Linear Unbiased Predictor (BLUP) of α_t in the sense of minimizing the prediction error variance out of all the unbiased predictors that are lineal combinations of $\hat{\alpha}_{t|t-1}$ and y_t .

Proof: $E(\hat{\alpha}_t - \alpha_t) = E[Q_t^{-1} X_t' V_t^{-1} (X_t \alpha_t + u_t) - \alpha_t] = E[Q_t^{-1} X_t' V_t^{-1} u_t] = 0$ so that $\hat{\alpha}_t$ is unbiased for α_t and $Var(\hat{\alpha}_t - \alpha_t) = E[(\hat{\alpha}_t - \alpha_t)(\hat{\alpha}_t - \alpha_t)'] = Q_t^{-1}$. Clearly, $\hat{\alpha}_t$ is linear in Y_t .

Let $\hat{\alpha}_t^L = L_1 \hat{\alpha}_{t|t-1} + L_2 y_t + l = L Y_t + l$ be any other linear unbiased predictor of α_t and define,

$D_t = L - Q_t^{-1} X_t' V_t^{-1}$, such that $L = D_t + Q_t^{-1} X_t' V_t^{-1}$. Since $\hat{\alpha}_t^L$ is unbiased for α_t ,

$$\begin{aligned} E(\hat{\alpha}_t^L - \alpha_t) &= E[(D_t + Q_t^{-1} X_t' V_t^{-1})(X_t \alpha_t + u_t) + l - \alpha_t] \\ &= E[D_t X_t \alpha_t + D_t u_t + Q_t^{-1} X_t' V_t^{-1} u_t + l] = E[D_t X_t \alpha_t + l] = 0. \end{aligned}$$

Thus, for $\hat{\alpha}_t^L$ to be unbiased for α_t irrespective of the distribution of α_t , we need that

$D_t X_t = 0$ and $l = 0$, which implies,

$$Var(\hat{\alpha}_t^L - \alpha_t) = Var[(D_t + Q_t^{-1} X_t' V_t^{-1})(X_t \alpha_t + u_t) - \alpha_t] = Var[D_t u_t + Q_t^{-1} X_t' V_t^{-1} u_t].$$

Hence,

$$\begin{aligned} Var(\hat{\alpha}_t^L - \alpha_t) &= E[(D_t u_t + Q_t^{-1} X_t' V_t^{-1} u_t)(u_t' D_t' + u_t' V_t^{-1} X_t' Q_t^{-1})] \\ &= D_t V_t D_t' + Q_t^{-1} X_t' V_t^{-1} V_t D_t' + D_t V_t V_t^{-1} X_t' Q_t^{-1} + Q_t^{-1} \\ &= Q_t^{-1} + D_t V_t D_t' = Var(\hat{\alpha}_t - \alpha_t) + D_t V_t D_t', \text{ since } D_t X_t = 0. \text{ QED} \end{aligned}$$

APPENDIX B. *Equality of the GLS and the Kalman Filter predictors when the measurement errors are uncorrelated over time*

Consider the model, $y_t = Z_t \alpha_t + e_t$; $\alpha_t = T \alpha_{t-1} + \eta_t$; where e_t and η_t are independent white noise series with $Var(e_t) = \Sigma_t$ and $Var(\eta_t) = Q_t$. Let $\hat{\alpha}_{t|t-1} = T \hat{\alpha}_{t-1}$ define the predictor of α_t at time (t-1), with prediction error covariance matrix $P_{t|t-1}$, assumed to be positive definite.

GLS set-up at time t:
$$\begin{pmatrix} T \hat{\alpha}_{t-1} \\ y_t \end{pmatrix} = \begin{pmatrix} I \\ Z_t \end{pmatrix} \alpha_t + \begin{pmatrix} u_{t|t-1} \\ e_t \end{pmatrix} \quad (\text{Equation 3.7})$$

where now $V_t = Var \begin{pmatrix} u_{t|t-1} \\ e_t \end{pmatrix} = \begin{bmatrix} P_{t|t-1} & 0 \\ 0 & \Sigma_t \end{bmatrix}$, (same as in 3.8 except that $C_t = 0$).

The GLS predictor at time t is,

$$\begin{aligned} \hat{\alpha}_t &= \left[(I, Z_t') V_t^{-1} \begin{pmatrix} I \\ Z_t \end{pmatrix} \right]^{-1} (I, Z_t') V_t^{-1} \begin{pmatrix} \hat{\alpha}_{t|t-1} \\ y_t \end{pmatrix} = [P_{t|t-1}^{-1} + Z_t' \Sigma_t^{-1} Z_t]^{-1} [P_{t|t-1}^{-1} \hat{\alpha}_{t|t-1} + Z_t' \Sigma_t^{-1} y_t] \\ &= [P_{t|t-1}^{-1} - P_{t|t-1}^{-1} Z_t' F_t^{-1} Z_t P_{t|t-1}^{-1}] [P_{t|t-1}^{-1} \hat{\alpha}_{t|t-1} + Z_t' \Sigma_t^{-1} y_t]; \quad (F_t = [Z_t P_{t|t-1} Z_t' + \Sigma_t]) \\ &= [I - P_{t|t-1}^{-1} Z_t' F_t^{-1} Z_t] \hat{\alpha}_{t|t-1} + P_{t|t-1}^{-1} Z_t' \Sigma_t^{-1} y_t - P_{t|t-1}^{-1} Z_t' \Sigma_t^{-1} y_t + P_{t|t-1}^{-1} Z_t' F_t^{-1} y_t \\ &= \hat{\alpha}_{t|t-1} + P_{t|t-1}^{-1} Z_t' F_t^{-1} (y_t - Z_t \hat{\alpha}_{t|t-1}), \end{aligned}$$

which is the same as the Kalman Filter predictor (Harvey, 1989, Equation 3.2.3a). *QED*

APPENDIX C. *Computation of $\hat{\alpha}_t$ (Equation 3.11)*

Consider the GLS predictor (3.10), $\hat{\alpha}_t = \left[(I, Z_t') V_t^{-1} \begin{pmatrix} I \\ Z_t \end{pmatrix} \right]^{-1} (I, Z_t') V_t^{-1} \begin{pmatrix} T \hat{\alpha}_{t-1} \\ y_t \end{pmatrix}$, with $V_t = \begin{bmatrix} P_{t|t-1} & C_t \\ C_t' & \Sigma_t \end{bmatrix}$. A familiar result on the inverse of a partitioned matrix applied to the matrix V_t yields the expressions,

$$(I, Z_t') V_t^{-1} = [(P_{t|t-1}^{-1} + P_{t|t-1}^{-1} C_t H_t C_t' P_{t|t-1}^{-1} - Z_t' H_t C_t' P_{t|t-1}^{-1}), (-P_{t|t-1}^{-1} C_t H_t + Z_t' H_t)] \quad (\text{C.1})$$

$$P_t = \left[(I, Z_t') V_t^{-1} \begin{pmatrix} I \\ Z_t \end{pmatrix} \right]^{-1} = [P_{t|t-1}^{-1} + (Z_t' - P_{t|t-1}^{-1} C_t) H_t (Z_t - C_t' P_{t|t-1}^{-1})]^{-1} \quad (\text{C.2})$$

where $H_t = [\Sigma_t - C_t' P_{t|t-1}^{-1} C_t]^{-1}$. It follows from (C.1) and (C.2) that,

$$\hat{\alpha}_t = P_t(I, Z_t')V_t^{-1} \begin{pmatrix} T\hat{\alpha}_{t-1} \\ y_t \end{pmatrix} = T\hat{\alpha}_{t-1} + P_t(Z_t' - P_{t|t-1}^{-1}C_t)H_t(y_t - Z_tT\hat{\alpha}_{t-1}) \quad (C.3)$$

Computing the matrix P_t in the right hand side of (C2) by use of a standard matrix inversion lemma (Harvey, 1989, Page 108), and substituting in the right hand side of (C.3) yields after some algebra the equation (3.11). *QED*

APPENDIX D. *Computation of $P_t^{bmk} = Var(\tilde{\alpha}_t^{bmk} - \alpha_t)$ and $C_t^{bmk} = Cov[\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t, \tilde{e}_t]$*

The benchmarked predictor defined by (4.4) can be written as,

$$\tilde{\alpha}_t^{bmk} = [I - (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}Z_t]\tilde{T}\tilde{\alpha}_{t-1}^{bmk} + (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}\tilde{y}_t \quad (D.1)$$

where $R_t = [\tilde{Z}_tP_{t|t-1}^{bmk}\tilde{Z}_t' - \tilde{Z}_tC_{t,0}^{bmk} - C_{t,0}^{bmk}, \tilde{Z}_t', \tilde{\Sigma}_{tt}^*]$. Substituting $\tilde{y}_t = \tilde{Z}_t\tilde{\alpha}_t + \tilde{e}_t$ (Equation 4.1a) in

(D.1) and decomposing $\tilde{\alpha}_t = [I - (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}Z_t]\tilde{\alpha}_t + (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}Z_t\tilde{\alpha}_t$ yields,

$$\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t = [I - (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}Z_t](\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t) + (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}\tilde{e}_t \quad (D.2)$$

Denote, $G_t = [I - (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}Z_t]$ and $K_t = (P_{t|t-1}^{bmk}\tilde{Z}_t' - C_{t,0}^{bmk})R_t^{-1}$. Then, by (D.2), the variance-covariance matrix of the prediction error $(\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t)$ under the model (4.1) is,

$$P_t^{bmk} = E[(\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t)(\tilde{\alpha}_t^{bmk} - \tilde{\alpha}_t)'] = G_tP_{t|t-1}^{bmk}G_t' + K_t\tilde{\Sigma}_{tt}K_t' + G_tC_{t,0}^{bmk}K_t' + K_tC_{t,0}^{bmk}'G_t' \quad (D.3)$$

where $P_{t|t-1}^{bmk} = E[(T\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t)(T\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t)'] = \tilde{T}P_{t-1}^{bmk}\tilde{T}' + \tilde{Q}$, $\tilde{\Sigma}_{tt} = E(\tilde{e}_t\tilde{e}_t')$ (Equation 4.1a) and

$C_t^{bmk} = Cov[\tilde{T}\tilde{\alpha}_{t-1}^{bmk} - \tilde{\alpha}_t, \tilde{e}_t]$. The matrix C_t^{bmk} is computed similarly to (3.9) using the chain,

$$C_t^{bmk} = A_{t-1}^{bmk}A_{t-2}^{bmk}\dots A_2^{bmk}\tilde{A}_1^{bmk}\tilde{\Sigma}_{1t} + A_{t-1}^{bmk}A_{t-2}^{bmk}\dots A_3^{bmk}\tilde{A}_2^{bmk}\tilde{\Sigma}_{2t} + \dots + A_{t-1}^{bmk}\tilde{A}_{t-2}^{bmk}\tilde{\Sigma}_{t-2,t} + \tilde{A}_{t-1}^{bmk}\tilde{\Sigma}_{t-1,t} \quad (D.4)$$

where $A_j^{bmk} = \tilde{T}G_j$, $\tilde{A}_j^{bmk} = \tilde{T}K_j$, with G_j, K_j defined as above.

TABLES AND FIGURES

Table 1. Theoretical and Empirical Variances and Covariances Under the Model (5.1) for Last Time Point (t=45). 10,000 Replications.

Series	P_{45}^{bmk} (Model)			P_{45}^{bmk} (Empirical)			C_{45}^{bmk} (Model)			C_{45}^{bmk} (Empirical)		
1	.274			.276			.039			.041		
2		1.122			1.119			.615			.614	
3			.337			.344			.063			.068

Table 2. Means and STDs (in parentheses) of Monthly Ratios Between STD of GLS Predictor and STD of Optimal Predictor, 1998-2003.

Division	Prediction of Unemployment	Prediction of Trend
New England	1.03 (.002)	1.02 (.002)
Middle Atlantic	1.02 (.002)	1.02 (.002)
East North Central	1.00 (.001)	1.00 (.001)
West North Central	1.02 (.002)	1.02 (.002)
South Atlantic	1.02 (.001)	1.02 (.001)
East South Central	1.00 (.001)	1.00 (.001)
West South Central	1.02 (.001)	1.01 (.001)
Mountain	1.03 (.002)	1.03 (.002)
Pacific	1.02 (.001)	1.02 (.001)

Table 3. Means and STDs (in parentheses) of Ratios Between Benchmarked and Unbenchmarked Predictor of Total Unemployment in Census Divisions.

Division	1998-2000, 2002-2003	2001
<i>New England</i>	1.01 (.016)	1.03 (.017)
Middle Atlantic	1.00 (.014)	1.03 (.019)
East North Central	1.00 (.013)	1.03 (.013)
West North Central	1.01 (.014)	1.03 (.011)
South Atlantic	1.00 (.016)	1.04 (.016)
East South Central	1.01 (.016)	1.03 (.014)
West South Central	1.00 (.014)	1.04 (.022)
Mountain	1.00 (.011)	1.02 (.011)
Pacific	1.00 (.017)	1.04 (.020)

Table 4. Means and STDs (in parentheses) of Ratios between STD of Benchmarked and Unbenchmarked Predictor of Total Unemployment in Census Divisions, 1998-2003.

Division	Means (STDs)
New England	0.85 (.013)
Middle Atlantic	0.94 (.005)
East North Central	0.96 (.004)
West North Central	0.89 (.008)
South Atlantic	0.96 (.004)
East South Central	0.86 (.007)
West South Central	0.92 (.006)
Mountain	0.88 (.009)
Pacific	0.96 (.004)

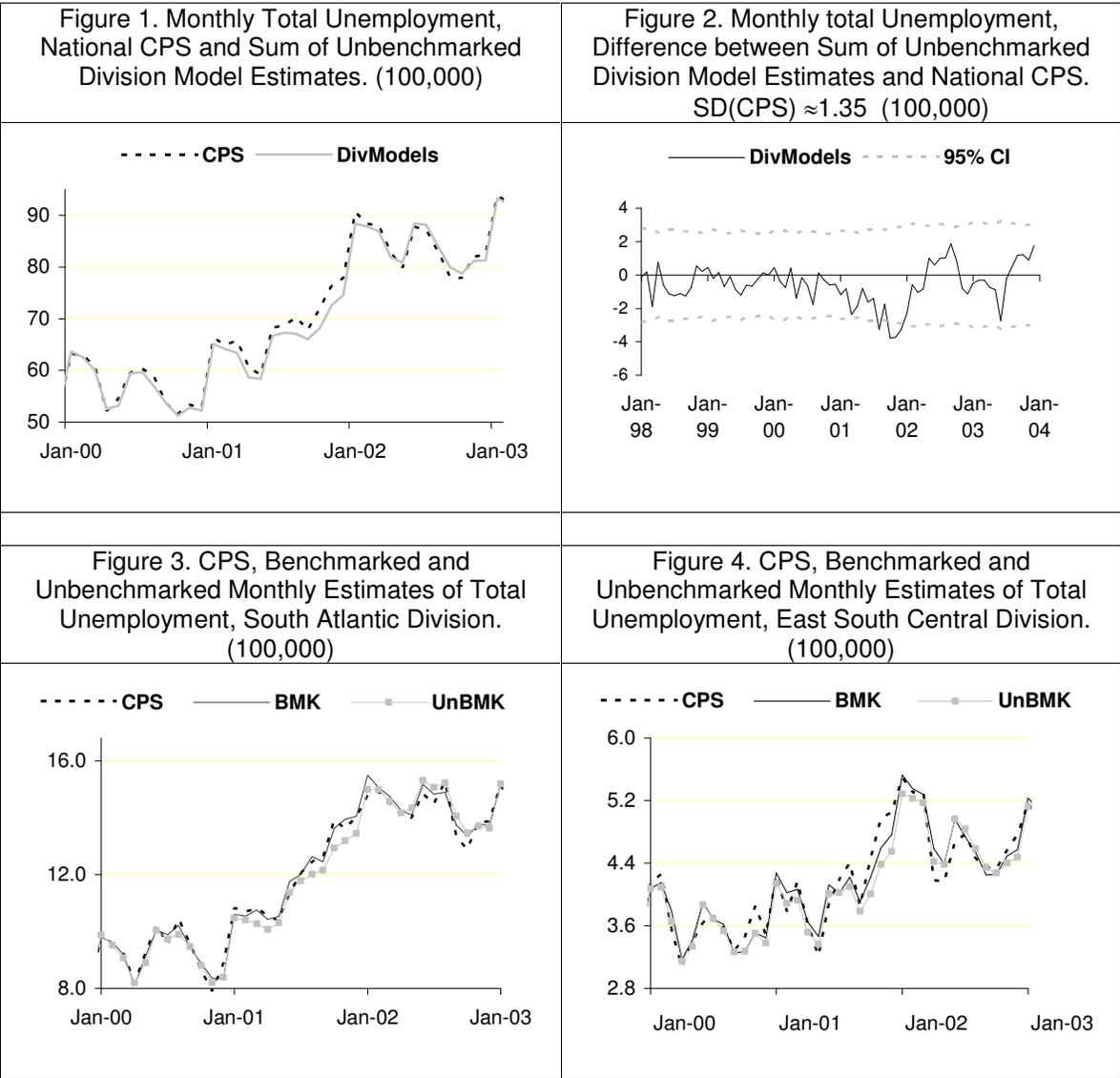


Figure 5. CPS, Benchmarked and Unbenchmarked Monthly Estimates of Total Unemployment, Pacific Division. (100,000)

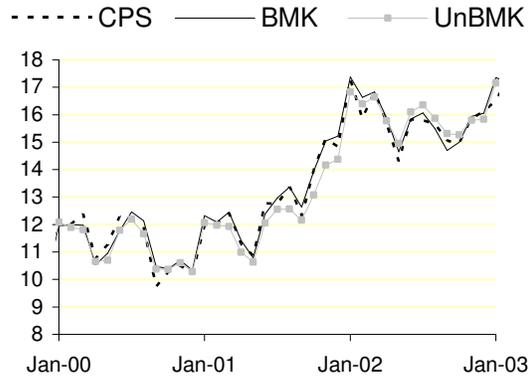


Figure 6: STD of CPS, Benchmarked and Unbenchmarked Estimates of Total Monthly Unemployment, South Atlantic Division. (10,000)

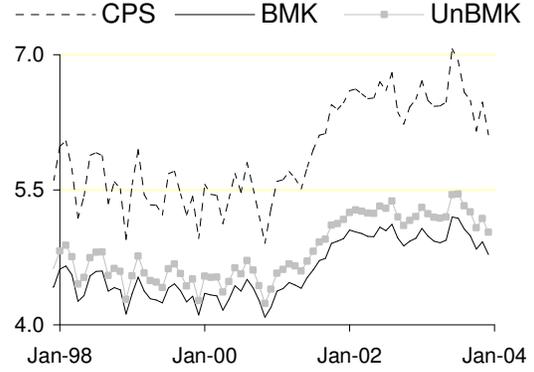


Figure 7. STD of CPS, Benchmarked and Unbenchmarked Estimates of Total Monthly Unemployment, East South Central Division. (10,000)

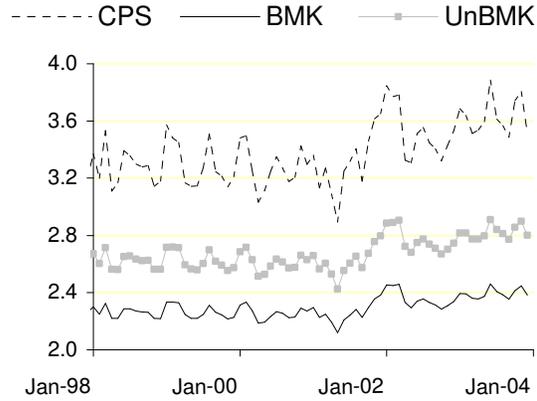


Figure 8. STDs of CPS, Benchmarked and Unbenchmarked Estimates of total Monthly Unemployment, Pacific Division. (10,000)

