WORKING PAPER SERIES

Robust Estimation in Clustered Multiple Time Series
with Structural Change

by

Erniel Barrios
Resi O. Olivares
Ma. Shiela R. Data

Abstract

We postulate a model for clustered multiple time series where individual and cluster-specific effects were represented by random components. Structural change is also considered assuming that it occurs only in the autoregressive parameter. To induce robustness during episodes of temporary structural change, we use the forward search and bootstrap methods embedded into the backfitting algorithm to estimate the parameters. Simulation studies show that the estimated models have high predictive ability especially in long time series where temporary structural change usually occur. High predictive ability is also observed even in near-nonstationarity and in the presence of misspecification errors.

Keywords: multiple time series, clustering, structural change, forward search, bootstrap

1. Introduction

Suppose there are N credit card holders, each have generated transactions for T time points. While each card holder does their own independent transactions, company policies like credit limits, among others, drives their behavior to generate similar patterns. Loyalty incentive policies of the company may also generate dependencies among certain group of consumers to exhibit clustering. Sudden changes in price structure may further cause temporary perturbations in their transaction history. Among stock prices of certain issues, while individual stocks exhibit similar patterns, contagion, say among those in the same sectors, may lead to clustering leading towards the manifestation of temporary structural change when certain random shocks occur in that sector.

In modeling time series data, there are many stylized facts that constrain extraction of information. For example, level shifts, outliers, and structure changes that are commonly encountered could easily mislead the conventional time series analysis procedure resulting to erroneous conclusions (Tsay, 1988). These may be caused by disasters, economic crises or changes in the political leadership. Difficulties are encountered in the prediction of these time series because these perturbations in the data can influence model estimates and lead to biased or inaccurate forecasts.

Veron Cruz and Barrios (2014), developed an estimation procedure for a multiple time series by embedding the bootstrap methods, maximum likelihood estimation, and best linear unbiased prediction into the backfitting algorithm. The hybrid estimation procedure has high predictive ability, robust to the number of time series behavior, but improved estimates are realized in longer time series data. Campano and Barrios (2011) developed a procedure for estimating a time series model that exhibits robustness in the presence of temporary structural change. We modify the estimation procedure of Veron Cruz and Barrios (2014) by integrating the modified forward search methods in Campano and Barrios (2011) in aiming for robust estimates in the presence of temporary structural change.

We present some literature on multiple time series in Section 2 while we discuss in Section 3 studies related to structural change. In Section 4, we discussed details a multiple time series with temporary structural change. In Section 5, we present the postulated multiple time series with clustering and simultaneously with clustering and temporary structural change. Then, in Section 5, we present the simulation study design and the results and discussion in Section 6. We finally summarize some concluding notes.

2. Multiple Time Series

(Tsay, 2005) defined a time series as a collection of random variables through time. The simplest forms of the generating mechanisms of the time series are the common autoregressive-integrated-moving average processes (ARIMA) given as follows:

$$(1 - \phi_1 B)(1 - B)z_t = \mu + (1 - \theta_1)a_t + a_t \qquad (1)$$

Left-hand side of (1) describes the autoregressive process where observed values of z are related to its past values as well as the past and present random shocks in the right-hand side. Box-Jenkins modeling strategy for univariate time series were quite successful, (Box and Jenkins, 1976). However, there are many instances where the joint analysis of multivariate or vector time series is more meaningful. A dynamic system of k-related time series is denoted as $z_{1t}, z_{2t}, ..., z_{kt}$, and denoted by $\mathbf{Z}_t = (z_{1t}, z_{2t}, ..., z_{kt})'$, the $k \times 1$ time series vector at time t (Box et al, 2008). A general framework, some theory and application were provided by Parzen (1968) for multiple time series modeling. Sims (1980) recommended the vector autoregressive model (VAR Model) where variables are assumed to be endogenous, allowing for a dynamic analysis among variables. Variance decomposition may also be done to account for the sources of variation. An extended sample cross-correlations approach for model identification was proposed by Tiao and Tsay (1983).

Early contributions in multiple time series focused on related time series. Akman and Gooijer (1996) proposed a component-extraction method for mutually uncorrelated time series. These time series may be obtained by recording values of same variables from different locations, example, meteorological data from different locations, financial data or economic data like rice prices from different countries or regions.

Veron Cruz and Barrios (2014) proposed a backfitting strategy to combine the maximum likelihood estimation and best linear unbiased prediction to estimate a multiple time series with a common autoregressive component and different individual random effects. The method yields comparable if not better estimates to the Arellano-Bond GMM estimator of the panel model. Prado et.al (2006) proposed a multivariate time series modeling approach by mixing the paradigm of hierarchical mixture-of-experts (HME) and vector autoregressive components (VAR) presented a time-domain approach on analyzing a multivariate non-stationary time series.

Park (2011) investigated a new class of multiple time series models without the parametric assumptions and proposed a unified estimation method of minimal dimension using an Akaike Information Criterion for situations in which the dimension for multiple regressors is unknown. The model is advantageous in forecasting future values in the longitudinal data context.

3. Structural Change

Structural change affects modeling of the key economic and financial time series such as output growth, inflation, exchange rates, interest rates, stock returns, among others. Legislative, institutional or technological changes, shifts in economic policy, or large macroeconomic shocks such as the doubling of oil prices experienced over the past decades could cause these structural changes. Hence, the detection of these stylized facts became a valuable aspect in modeling.

Atkinson, et.al (1997) studied the effects on the parameter estimates of various kinds of intervention in time series analysis. The study made use of scored statistics and simulation envelopes to form a scaled distance to assess the importance of the interventions. Wang and Zivot (2000) developed a Bayesian approach for analyzing a dynamic time series model with multiple structural changes in level, trend, and error variance based on the Gibbs sampler. The Monte Carlo study demonstrated that for a fixed number of structural changes, the procedure produces sharp parameter estimates.

The forward search (FS) algorithm is a powerful method for detecting multiple masked outliers, in determining their effects on models fitted to data, and for detecting systematic model inadequacy (Riani, 2002). Atkinson, et.al (2008) used forward search algorithm for multivariate outlier detection. They proposed a new calibration method for obtaining reliable cut-off points of distances derived from the MCD estimator of a scatter. Campano and Barrios (2011) estimated a time series model with structural change using a nonparametric bootstrap and FS algorithm. Nonparametric bootstrap (block bootstrap or AR sieve) was used to generate a series of estimates via a modified forward search (FS) algorithm. The estimation procedure yields robust estimates of ARIMA models with temporary structural changes generated from AR (1), MA (1), and ARMA (1, 1) processes. The estimates from FS algorithm is more robust and better captures the overall structure of data than the estimates from standard methods like conditional least squares (CLS) for relatively long, stationary, and invertible time series. However, the procedure performed poorly on short time series.

The forward search method makes use of a small data set and then tracks the effect of adding more observation to the initial data until all observations were included. This powerful method avoids the masking effect of multiple outliers (Riani and Atkinson, 2000). The forward search algorithm is also used with other statistical analysis such as regression diagnostics for binomial data (Atkinson and Riani, 2001), exploring multivariate data (Atkinson et al, 2004) or clustering multivariate data (Atkinson and Riani, 2006).

Time series data are also prone to outliers. This may result from typographic error or real life crisis such us strikes or oil crisis (Tsay, 1986). An iterative procedure was proposed by Chang et al (1988) to detect innovative and additive outliers and estimate time series parameters of autoregressive-integrated-moving average models in the presence of outliers.

Riani (2004), proposed and illustrated the effectiveness of extending the forward search procedure to time series data using the diffuse Kalman filter with missing observations. Econometric applications of the forward search was also illustrated by Atkinson (2009). The modified forward search procedures using segments of time series with overlapping and independent blocks was proposed by Campano and Barrios (2011).

The success of the backfitting procedure as used by Veron Cruz and Barrios (2014) in estimating a mixed model of independent multiple time series and the robustness of the estimates produced by the modified forward search procedure using segments of time series with overlapping and independent blocks as proposed by Campano and Barrios (2011) makes robust estimation of multiple time series viable. Furthermore, a more "general" approach may be developed accounting for a cluster parameter as suggested by Mouchart and Rombouts (2005). They proposed a strategy that considers a clustering approach into the usual panel model specification which allows for nowcasting on short time series and many missing data.

4. Multiple Time Series with Temporary Structural Change

Suppose the N units are uncorrelated but assume similar pattern-specifications, i.e., the N time series shares a common autocorrelation parameter ϕ. Veron Cruz and Barrios (2014), postulated a multiple time series (no structural change) in Equation (2).

$$Y_{it} = \phi Y_{it-1} + \lambda_i + \epsilon_{i.t} \quad i = 1, 2, \dots, N, \ t = 1, 2, \dots, T \qquad (2)$$
$$\lambda_i \sim N\left(\mu_i, \sigma_{\lambda_i}^2\right), \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

For example, consider rice production per capita as multiple independent time series where: Y_{it} is rice production per capita of the ith province at time t, ϕ is the common component for all the provinces attributed to the national agricultural policies or programs governing the production, λ_i is the provincial specific-production endownment, ϵ_{it} is the error component.

When disturbances in the implementation of national agricultural policies and projects happen (e.g. calamities, food crisis, deficit of funds, etc), it greatly influences the rice production at the provincial level. These disturbances can be reflected as structural change in model (2). Hence, Equation (2) is modified to account for time series that exhibit structural change in Equation (3).

$$Y_{it} = \phi_{it}^* Y_{it-1} + \lambda_i + \epsilon_{it} \quad i = 1, 2, \dots, N, \ t = 1, 2, \dots, T \qquad (3)$$
$$\lambda_i \sim N\left(\mu_i, \sigma_{\lambda_i}^2\right), \epsilon_i \sim N(0, \sigma_\epsilon^2)$$

Where $\phi_{it}^* = \phi + \psi_{it}$ and $\psi_{it} = 0$ during ordinary times and $\psi_{it} \neq 0$ during periods of temporary structural change.

Estimation Procedure

We estimate Model (3) following the algorithm of Veron Cruz and Barrios (2014) imbedded with Campano and Barrios (2011) in aiming for robustness of estimates typically generated from the forward search algorithm. The forward search algorithm for independent blocks of similar lengths was used to estimate the parameters that mitigates the influence of the presence of structural change. We further assume that Individual units are uncorrelated but shares common autoregressive pattern and individual effects λ_i are independent and normally distributed.

Phase I: Initialization

Following Campano and Barrios (2011), we partition each time series into k blocks of similar length. For each of the k blocks, fit Model (2) to generate initial parameter estimates.

Step 1: Using the modified forward search (FS) algorithm with independent blocks of same length, ignore the term, ϕY_{it-1}, to estimate the random effects λ_{ik} as $\hat{\lambda}_{ik}$ using the best linear unbiased predictors (BLUP) method.

Step 2: Compute the residuals: $r_{it}^{ar} = Y_{it} - \hat{\lambda}_{ik}$.

Step 3: In each block for each time series, estimate ϕ through the autoregression of the residuals in Step 2 $\Rightarrow \hat{\phi}_{ik}$.

Phase I yield $k \ x \ N$ estimates of λ and ϕ, from each block of individual time series.

Phase II: Iteration

1. Given $\widehat{\phi}_{ik}$ from Phase I, compute new residuals $r_{it}^{re} = Y_{it} - \hat{\phi}_{ik}Y_{it-1}$ for each unit i and each block k.
2. For each unit i and block k, estimate λ_{it} as the BLUP from $r_{it}^{re} = \lambda_{it} + \epsilon'_{it}$, where $\epsilon'_{it} \sim N(0, \sigma^2) \Rightarrow \hat{\lambda}_{it}$.
3. Compute new residuals $r_{i,t}^{ar} = y_{it} - \hat{\lambda}_{it}$.
4. For each unit i and each block k, estimate AR(p) from the residuals in 3, $\Rightarrow \hat{\phi}_{ik}$.
5. Iterate, i.e., repeat Steps 1-4 until convergence.
6. For each block, use bootstrap procedure on the estimates $\hat{\lambda}_{ik}$ within each time series to obtain $\hat{\lambda}_i^{BS}$
7. For each unit, use bootstrap method to obtain $\hat{\phi}_k$ for each block k. Use bootstrap method again on $\hat{\phi}_k$ to estimate $\hat{\phi}^{BS}$. Compute the bootstrap confidence interval (BCI) using the adjusted bootstrap percentile method (BC_a) as discussed by Davison and Hinkley (1997).

Steps 1-5 of Phase II resulted to converging k estimates of λ and ϕ. Convergence is desirable to assure stability of the estimates. Step 6-7 uses bootstrap on N estimates of λ and common autoregressive parameter ϕ. Estimates generated from bootstrap method is endowed with monte carlo measures of accuracy.

Phase III. Detection of Structural Change, Adjustment of Estimates from Robust Estimation.

Phase III illustrates the procedure of identifying the presence or absence of structural change within segment. Then estimates within the block is adjusted based on identified structural change.

1. Using the BCI from Algorithm II, compare the individual estimates in each block to the BCI to detect the segments that exhibits perturbation.
2. After identifying the segment where the estimate is outside the BCI, compute the difference. $\omega_{im} = BCI_l - \hat{\phi}_{im}$, where m = the segment where $\hat{\phi}_{im} \notin$ BCI for ϕ_i^{BS} BCI_l is the BCI limit nearest to $\hat{\phi}_i^{BS}$.

5. Clustered Multiple Time Series with Structural Change

Suppose N time series are grouped into C clusters, each cluster contains N_k time series, $k=1,...,C$, each of length T, $\sum N_k = N$. Let $Y_{i_k t}$ be the ith time series that belong to the kth cluster and have a common higher level parameter ψ_k shared with other time series in the same cluster. Time series in different clusters are assumed to be independent. Further assume that ϕ, the common autoregressive parameter is shared by all N time series. Let λ_{i_k} be the component unique to the ith individual time series in the kth cluster. Model (3) is generalized into the following clustered time series with random components model:

$$Y_{i_k t} = \phi Y_{i_k t-1} + \lambda_{i_k} + \psi_k + \epsilon_i, \ i_k = 1, 2, ..., N_k, \ t = 1, 2, ..., T, k = 1, 2, ..., C \qquad (4)$$
$$\lambda_{i_k} \sim N\left(\mu_{\lambda_{i_k}}, \sigma_\lambda^2\right), \qquad \psi_k \sim N\left(\mu_{\psi_k}, \sigma_\psi^2\right), \qquad \epsilon_{i_k} \sim N(0, \sigma_\epsilon^2)$$

Each λ_{i_k} of the individual time series are assumed to be normally distributed with unique means but common variance. The C cluster parameters ψ_k are also normally distributed with each component having its own mean but common variance. The error term of the model is also assumed to be normally distributed with zero mean and has variance σ_ϵ^2. It is also assumed that σ_ϵ^2 is much less than σ_λ^2 and σ_ϵ^2 is much less than σ_ψ^2. This model differs slightly from the model used by Veron Cruz and Barrios (2014) as this model considers the cluster-specific effects to account for inherent grouping (eg. spatial aggregation) for some time series.

The N time series are grouped into C clusters of length T, ψ_k represents the unique contribution of the kth cluster while the λ_{i_k} reflects the individual peculiarities among the time series within each cluster. As an illustration, consider the monthly prices of rice in each province of a country. In the Philippines for instance, each province is administered by a provincial agriculturist, and is part of a regional grouping which is also governed by a higher agricultural officer. $Y_{i_k t}$ is the price of rice in the ith province in the kth cluster on the tth month. The parameter λ_{i_k} represents the unique price level in the ith province of the kth region and ψ_k accounts for the higher level effect of the uniqueness of the geographical, meteorological and political nature of the region.

Estimation Procedure

Since there is a shared higher level parameter within clusters and the model has additive parameters, we modify the backfitting algorithm proposed by Veron Cruz and Barrios (2014) to allow estimation of the clustering parameter. Furthermore, to produce robust estimates in the presence of structural change, we incorporate the modified forward search algorithm using independent blocks of similar lengths of contiguous time points and adjust using the bootstrap confidence interval as proposed by Campano and Barrios (2011).

Phase I. Initialization

1. Using the modified forward search (FS) algorithm with independent blocks of uniform lengths, initial estimates $\hat{\lambda}_{i_k j}$ and $\hat{\psi}_{kj}$ of the random effects $\lambda_{i_k j}$ and ψ_{ij} for the jth block of the ith individual on the kth cluster were obtained ignoring the $\phi Y_{i_k t-1}$ component, using the best linear unbiased prediction (BLUP).

2. Using the initial estimates $\hat{\lambda}_{i_k j}$ and $\hat{\psi}_{kj}$ for each blocks of each time series, compute the residuals $r_{i_k,t}^{ar} = Y_{i_k,t} - \hat{\lambda}_{i_k j} - \hat{\psi}_{kj}$.

3. For each block on each individual time series, obtain the estimates $\hat{\phi}_{i_k j}$ of $\phi_{i_k j}$ of the jth block of the ith time series on the kth cluster as parameters of AR(p) models.

After Phase I, there will be $b \times N$ $\lambda_{i_k j}$ and $\phi_{i_k j}$ estimates. Each estimate corresponds to the estimate of the parameters of the model within each block of each individual time series. Also, there will be $b \times C$ ψ_{kj} estimates with one estimate per block for each cluster. These estimates are the initial values of the backfitting algorithm.

Phase II. Iteration

1. Using the estimates $\hat{\phi}_{i_k j}$, compute the residuals $r_{i_k t} = Y_{i_k t} - \hat{\phi}_{i_k j} Y_{i_k t-1}$.

2. Use the modified forward search (FS) algorithm with independent blocks of uniform length, estimates $\hat{\lambda}_{i_k j}$ and $\hat{\psi}_{kj}$ of the random effects $\lambda_{i_k j}$ and ψ_{kj} for the jth block of the

i^{th} individual on the kth cluster are obtained using best linear unbiased prediction (BLUP).

3. Using the new estimates $\hat{\lambda}_{i_k j}$ and $\hat{\psi}_{kj}$ for each blocks of each time series, compute the residuals $r_{i_k t}^{ar} = Y_{i_k t} - \hat{\lambda}_{i_k j} - \hat{\psi}_{kj}$.

4. For each block on each individual time series, obtain the new estimates $\hat{\phi}_{i_k j}$ of $\phi_{i_k j}$ of the jth block of the ith time series using the ARIMA estimation procedure.

5. Iterate, i.e., repeat Steps 1-4 until convergence.

Steps $1 - 5$ allows convergence of the estimates for λ_{i_k}, ψ_k and ϕ. These estimates represent the structure of the time series within each block.

6. Perform bootstrap method on the estimates $\hat{\lambda}_{i_k j}$ within each time series to obtain $\hat{\lambda}_{i_k}^{BS}$ for each individual time series.

7. Perform bootstrap method also on the estimates $\hat{\psi}_{kj}$ within each cluster to obtain $\hat{\psi}_k^{BS}$, an estimate of ψ_k for each cluster.

8. Perform bootstrap method also on the estimates $\hat{\phi}_{i_k j}$ within each block to obtain $\hat{\phi}_j^{BS}$, the estimate of ϕ within the jth block.

9. Perform bootstrap method on the estimates $\hat{\phi}_j^{BS}$ among blocks to obtain $\hat{\phi}^{BS}$. Compute the bootstrap confidence interval (BCI) using the adjusted bootstrap percentile method (BC_a).

Phase III. Detection and Adjustment of Estimates for Robust Estimation

1. Using the BCI obtained in step 9 of Phase II, identify the blocks within each time series that contain structural change. A block is identified to have structural change if the estimate $\hat{\phi}_{ij}$ is outside the BCI.

2. Adjust the estimate for the block with structural change using $\hat{\phi}_{i_k j}^{adj} = \hat{\phi}^{BS} + \omega_{i_k,t} I(t_j)$, where $\omega_{i_k,t} = BCI_L - \hat{\phi}_{i_k j}$ and $BCI_L = \min(|\hat{\phi}_{ij} - BCI_{Lower}|, |\hat{\phi}_{ij} - BCI_{Upper}|)$.

This adjustment based on the BCI helps provide better estimates by incorporating the information that there is structural change within the time segment.

6. Simulation Studies

To evaluate the performance of the proposed estimation procedure, data were generated using various settings. Various scenarios included settings for coefficient of variation of λ and ϕ, magnitude of structural change, length of time series, and location of structural change. Simulation settings and implications to real life data are below:

Table 1. Simulation Settings for Multiple Time Series

Variable	Setting	Implications
T	60, 1560	60 represents a short time series and 1560 represents a long time series
ϕ	0.5, 0.75, 0.95	0.5 and 0.75 represents stationary time series while 0.95 represents near nonstationary series

Coefficient of Variation of λ	1, 5	1 represents low and 5 represents high variations in individual component
Misspecification Error	0.05, 0.15	0.05 represents low (absence) and 5 represents high misspecification error

The study is limited to 150 individual units, 100 replicates, and 6% of time points exhibiting structural change. Additionally, both λ and ϕ are assumed to be normally distributed. Note that the period where structural change occurs is divided by the number of locations where the structural change was inserted, i.e., if only have one location (e.g. start), the length of time series with structural change is 6%. But if we have two locations (e.g. middle and end), 6% will be divided into 2, making the middle and end to have 3% length each.

Table 2. Simulation Settings for Multiple Time Series with Structural Change

Variable	Values	Implications
Magnitude of Structural Change	0.04, -0,04, -0.1, -0.20	0.04 will make series to be more nonstationary (0.04) while -0.1, -0.2, -0.04 will make it more stationary
Location of structural change	Start; middle; end; start and middle; start and end; middle and end; start, middle, and end.	Different locations of structural change in the series

For the settings on structural change, we add 0.04 to ϕ for the time series to approach closer nonstationarity, in fact, adding 0.04 to ϕ =0.95 would make the series almost nonstationary (ϕ =0.99). Table 3 summarizes the values of AR coefficient after adding different magnitude of structural change.

Table 3. ϕ after Adding Values Reflective of Structural Change.

Magnitude of Structural Change	Values of ϕ		
	0.50	0.75	0.95
0.04	0.54	0.79	0.99
-0.04	0.46	0.71	0.91
-0.1	0.40	0.65	0.85
-0.2	0.30	0.55	0.75

We also assessed if different locations of structural changes will affect the estimates. Embedding of structural change was done by replacing the selected segments (start, middle, end, start and middle, start and end, middle and end) of the uncontaminated series with observations from a model with structural change.

For clustered data, we generated varying C, N_k, T, coefficient of variation of λ and ψ, ϕ, magnitude of structural change, length of time series with structural change and location of structural change. Structural change was also introduced to the data through the autoregressive parameter ϕ. The ϕ values for the time points where structural change occur were modified by adding or subtracting differing magnitudes based on the specification for the setting. The different values taken by each parameter are listed below.

Table 4. Simulation Settings for Clustered Data (without Structural Change)

Variable	Values	Implications
T	60 and 1560	Short vs. Long time series
ϕ	0.9 and 0.6	near non-stationary vs stationary time series
Coefficient of Variation of λ and ψ	0.05 and 0.15	Different variations in the values of the individual and cluster components
Cluster N (C)	5 and 10	Few big clusters vs many smaller clusters
Cluster Size (N_k)	5 and 20	

Table 5. Simulation Settings for Clustered Time Series with Structural Change

Variable	Values	Implications
Magnitude of Structural Change	0.09, ±0.2 and -0.09	More stationary (-0.09) vs more nonstationary (+0.2 and +0.09)
Location of structural change	Start; middle;end; start and middle; start and end; middle and end; start, middle, and end.	Will structural change in different parts of the time series cause problems?
Length of time points with SC	6% and 12%	Shorter vs longer structural change

When the time series is stationary($\phi = 0.6$), we are interested on the effect of a structural change that drives ϕ to go up (nearly nonstationary) or down (more stationary or independence) in some parts of the data. On the other hand, if the time series is nearly nonstationary ($\phi = 0.9$), we are interested with the implications if it becomes more stationary (ϕ goes down by 0.09 or 0.2) in some time segments. Much more interesting is what happens if we make the time series more nearly nonstationary ($\phi = 0.99$). By varying the locations of the structural change, we determine if estimates are still reliable if the structural change were divided or moved to different time segments.

7. Results and Discussions

We evaluate performance of the proposed estimation procedure on multiple time series (including clustering) with or without structural change. Assessment of predictive ability was done across the individual and combination of different scenarios for short and long time series, for stationary and near-nonstationary autoregressive models, for low and high error variance, for location and magnitude of structural change, and for varying coefficient of variation of random components.

7.1 Presence of Structural Change

Length of Time Series

Short time series consisting of 60 time points long time series with 1560 time points were considered. Without structural change, predictive ability (measured in terms of mean absolute prediction error-MAPE) is better for longer time series. The autoregressive parameter ϕ is also slightly better (smaller bias) in shorter series while relative bias of λ does not change prominently both different time series lengths. Similar is true even in the presence of structural change. See Table 6 for details.

Table 6. MAPE and Percentage of Correctly Identified Structural Change by Time Series Length

	MAPE		% Correctly Identified	
Time Series Length	No SC	With SC	No SC	With SC
Short (T=60)	14.38	15.07	50.40	40.61
Long (T=1560)	4.13	4.26	51.73	25.72

Stationarity

The hybrid estimation procedure yield estimated models with high predictive ability whether or not a structural change exists. In Table 7, predictive ability of the estimated model increases (lower MAPE) as the autoregressive approaches a near nonstationary level. Increasing proportion of structural changes can also be observed as the autoregressive parameter increases.

Table 7. MAPE and Percentage of Correctly Identified Structural Change by State of Stationarity

	MAPE		% Correctly Identified	
Autoregressive Parameter	**No SC**	**With SC**	**No SC**	**With SC**
$\phi = 0.5$	10.39	10.66	46.83	41.54
$\phi = 0.75$	8.96	9.36	43.00	36.13
$\phi = 0.95$	8.41	8.99	63.35	21.98

Misspecification Error

The estimation procedure is fairly robust to misspecification errors are exhibited in Table 8. Furthermore, predictive ability remains comparable whether structural change is present or not. Similar of true in the correct identification of absence/presence of structural change. The method is sensitive enough in correctly identifying presence of structural change.

Table 8. MAPE and Percentage of Correctly Identified Absence and Presence of Structural Change by Varying Values of σ_ϵ^2

	MAPE		% Correctly Identified	
Error Variance	**No SC**	**With SC**	**No SC**	**With SC**
1	9.22	9.65	51.73	33.10
5	9.28	9.68	50.39	33.23

Coefficient of Variation of λ_i

Predictive ability slightly declines as coefficient of variation of λ_i increases. Note that MAPE with or without structural change are also comparable. The method is again exhibiting sensitivity in the identification of structural change.

Table 9. MAPE and Percentage of Correctly Identified Absence and Presence of Structural Change by Varying Coefficient of Variation of λ_i

	MAPE		% Correctly Identified	
CV of λ_i	**No SC**	**With SC**	**No SC**	**With SC**
0.05	8.60	9.00	52.83	30.94

0.15	9.90	10.33	49.30	35.39

Structural Change Locations

Structural change was embedded at the following locations in the series: start; middle; end; start and middle; start and end; middle and end; start, middle and end. A span of 96 time points of structural perturbation was embedded in the long series while a span of four (4) time points was used in the short series. Table 10 show that predicting ability without structural change and with structural change at different locations are comparable. This implies robustness of the estimation procedure to various locations of structural change. Structural changes that occur at the early part of the time series are easy to identify while those changes that occur in the more recent part of the time series are more difficult to identify.

Table 10. MAPE and Percentage of Correctly Identified f Structural Change by Location of Structural Change

Location	MAPE	%Correctly Identified SC
No Structural Change	9.25	
Start	9.79	38.43
Middle	9.66	30.74
End	9.49	33.81
Start and Middle	9.72	32.85
Start and End	9.61	33.21
Middle and End	9.71	28.89
Start, Middle, and End	9.68	34.28

Magnitude of Structural Change

The method is also robust to the magnitude of the change in parameter values. Varying magnitude of structural change yield comparable predictive ability. Expectedly, larger magnitude of changes in the parameters are easier to identify that those parameters with small magnitude of changes, see Table 11 for details.

Table 11. MAPE and Percentage of Correctly Identified Structural Change by Magnitude of Structural Change

Magnitude	MAPE	% Correctly Identified SC
-0.20	10.62	23.42
-0.10	9.43	29.23
-0.04	9.23	40.79
0.04 (Near nonstationary)	9.34	39.33

7.2 Clustering in the Presence of Temporary Structural Change

Convergence Rate

In the backfitting estimation procedure, convergence is achieved if difference of estimates from two iterations do not change by at most 1%. For each setting, number of converging estimates were counted. In Table 12, non-converging replicates happened whenever the structural change forces the time series towards near nonstationarity. Nonconverging replicates also occur when the structural change happens in the beginning of a

long time series. However, even with non-converging settings, predictive ability is still good MAPE is in the range of 3-4%.

Table 12. Percentage of Converging Replicates by length of series (T), coefficient of variation (CV), ϕ and magnitude of structural change

T	CV	ϕ	Structural Change -0.09	0.09	0.2	All
60	0.05	0.6	100%	100%	100%	100%
		0.9	100%	100%	100%	100%
	0.15	0.6	100%	100%	100%	100%
		0.9	100%	100%	100%	100%
1560	0.05	0.6	100%	100%	100%	100%
		0.9	100%	79%	100%	93%
	0.15	0.6	100%	100%	100%	100%
		0.9	100%	82%	100%	94%
All			100%	95%	100%	98%

Predictive Ability

The estimated model generally yields a MAPE of about 10%, generally higher for short time series (16%), and much lower for longer time series (4%). The big difference in predictive ability is explained by the division of the time series into shorter blocks, often highlighting localized nonstationarity in shorter time series. Cluster size did not influence predictive ability of the estimates in a multiple time series. Increasing magnitude of structural change also leads to the deterioration of predictive ability.

Table 13a. Summary of MAPE (No Structural Change)

T	Number of Clusters	Cluster Size 5	20
60 (Short)	None	14.27	14.24
	5	14.26	14.21
	10	14.27	14.27
1560 (Long)	None	3.45	3.44
	5	3.44	3.44
	10	3.45	3.44

Table 13b. Summary of MAPE (Magnitude of Structural Change=-0.09)

T	Number of Clusters	Cluster Size 5	20
60 (Short)	None	14.51	14.50
	5	14.50	14.55
	10	14.51	14.46
1560 (Long)	None	3.64	3.64
	5	3.64	3.64
	10	3.64	3.64

Table 13c. Summary of MAPE (Magnitude of Structural Change=0.09)

T	Number of Clusters	Cluster Size 5	20
60 (Short)	None	14.67	14.67
	5	14.68	14.67
	10	14.66	14.67
1560 (Long)	None	3.82	3.82
	5	3.82	3.82
	10	3.92	3.82

Table 13d. Summary of MAPE (Magnitude of Structural Change=0.2)

T	Number of Clusters	Cluster Size 5	20
60 (Short)	None	18.25	
	5	18.26	18.25
	10	18.24	18.28
1560 (Long)	None	3.86	18.23
	5	3.86	3.87
	10	3.86	3.87

Predictive ability improves as the autoregressive parameter increases. However, predictive ability generally worsens as the magnitude of structural change increases. Furthermore, predictive ability deteriorates as the length of the time series manifesting structural change increases.

Table 14: MAPE by Magnitude of Autoregression Coefficient and Magnitude of Structural Change

ϕ	Magnitude of Structural Change			
	None	-0.09	0.09	0.2
0.6	9.28	9.30	9.51	10.43
0.9	8.41	8.85	8.98	11.69

Table 15. MAPE by Magnitude and Length of Structural Change

Magnitude of Structural Change	Length of Structural Change		
	None	6%	12%
0	8.85		
-0.09		9.01	9.14
0.09		9.16	9.33
0.2		10.53	11.58

Location of structural change have minimal effect on predictive ability of the estimated models. Increasing coefficient of variation of λ_i, leads to higher MAPE (inferior predictive ability). There is no interaction between location of the structural change and the coefficient of variation of λ_i.

Table 16. MAPE by Location and Length of Structural Change

Magnitude of Structural Change	Length of Structural Change		
	None	6%	12%
0	8.85		
Start		9.58	9.93
Middle		9.73	10.03
End		9.40	9.54
Start and Middle		9.67	10.39
Start and End		9.42	10.10
Middle and End		9.60	10.22
Start, Middle, End		9.60	9.91

Table 17. MAPE by Location of Structural Change and CV of λ_i

Magnitude of Structural Change	CV of λ_i	
	0.05	0.15
None	8.47	9.23
Start	9.35	10.16
Middle	9.46	10.29
End	9.07	9.87
Start and Middle	9.60	10.46
Start and End	9.35	10.17
Middle and End	9.49	10.32
Start, Middle, End	9.34	10.16

8. Application to Palay Production Per Capita

We used the estimation algorithm estimating a model for rice production per capita in Philippine provinces. Mean absolute percentage error was computed for each of the provinces. The average MAPE for all the provinces is relatively small at about 10.56%. Top 10 Provinces with lowest MAPE are summarize in Table 18, with values ranging from 0.109 to 2.816.

Table 18. Top 10 Provinces with Least MAPE

Province	MAPE
Davao del Sur	0.411
Zamboanga City	0.441
Rizal	0.924
Sarangani	1.036
Cagayan	1.107
La Union	1.529
Pampanga	1.586
Catanduanes	2.052
North Cotabato	2.062
Bohol	2.067

Top 10 Provinces with lowest mean rice production per capita from the actual data and based on the estimates from the fitted model are shown at Table 19. The estimated model captured the same provinces at same ranks as the raw data on rice production.

Table 19. Lowest 10 Provincial Rice Production Per Capita

Provinces	Raw		Provinces	Estimated
Cebu	1.506		Cebu	2.080
Basilan	3.821		Basilan	3.558
Rizal	3.943		Rizal	3.907
Cavite	5.026		Cavite	4.542
Davao City	5.507		Davao City	5.088
Siquijor	6.712		Siquijor	6.270
Camiguin	6.844		Camiguin	6.328
Batangas	6.936		Batangas	6.438
Misamis Oriental	7.885		Misamis Oriental	6.741
Benguet	9.006		Benguet	8.368

We present in Table 20 provinces with highest mean rice production per capita generated using the actual data and from the predictions from the fitted model. The top ten provinces with highest estimated rice production per capita are the same although not on same ranks with the raw data.

Table 20. Highest 10 Provinces with Highest Mean Rice Productions Per Capita

Highest			Highest	
Provinces	Actual		Provinces	Estimated
Isabela	180.052		Nueva Ecija	166.459
Kalinga	166.459		Kalinga	157.504
Apayao	163.484		Occidental Mindoro	163.484
Nueva Ecija	159.144		Apayao	180.052
Occidental Mindoro	157.504		Isabela	134.457
Cagayan	134.457		Cagayan	133.423
Sultan Kudarat	133.423		Sultan Kudarat	109.596
Nueva Vizcaya	126.495		Iloilo	126.495
Iloilo	109.596		Nueva Vizcaya	108.457
Ilocos Norte	108.457		Ilocos Norte	96.464

Effect of provincial component (e.g. soil composition, terrain, availability of irrigation, farm to market road, proximity to technologies) can be estimated as a random effect of individual times series. Table 21 shows the top 10 lowest and highest provincial effect. Nueva Ecija is known as the top producer of agricultural products in the Philippines, had the highest provincial component. On the other hand, Cebu, a corn-producing province, had the least component individual effect.

Table 21. Provinces with Highest 10 Provincial Components

Top 10 Maximum			Top 10 Minimum	
Provinces	Lambda		Provinces	Lambda

Nueva Ecija	261.0396323	Cebu	3.0431357
Kalinga	259.8882175	Basilan	3.3852745
Occidental Mindoro	212.1639205	Rizal	3.7860184
Apayao	185.7854483	Davao City	3.9349124
Cagayan	135.4461482	Cavite	4.1717525
Iloilo	128.5659914	Misamis Oriental	5.0390390
Isabela	123.6402183	Nueva Vizcaya	5.2117274
Tarlac	123.5821666	Camiguin	5.3036906
Ilocos Norte	121.0493553	Albay	5.4141803
Sultan Kudarat	118.3524787	Siquijor	5.9234651

9. Conclusions

In the presence of structural change, the forward search method and BLUP estimation embedded into the backfitting procedure yield robust estimated model with better predictive ability specially when the time series is longer with perturbations located towards the end of the time series. The procedure is also better on series with low coefficient of variation of λ and less misspecification error. The proposed method is also working well in identification of of presence of structural changes in a multiple time series.

The combination of the backfitting algorithm and forward search algorithm coupled with the BCI to estimate a clustered multiple time series with structural change is a robust method that was able to predict the time series well. The procedure has robust estimates for the autoregressive parameters especially for those with large values of the autoregressive parameter that is still a stationary time series. The estimation procedure has an excellent predictive ability for long time series where structural changes frequently occur.

References:

Akman, I. and Gooijer, J., 1996, Component extraction analysis of multivariate time series. *Computational Statistics and Data Analysis*, 21:487-499.

Atkinson, A. and Riani, M., 2001, Regression Diagnostics for Binomial Data from the Forward Search, *Journal of the Royal Statistical Society*, Series D, 50(1):63-78.

Atkinson, A. and Riani, M., 2006, Exploratory tools for clustering multivariate data, *Computational Statistics & Data Analysis*, 52:272–285.

Atkinson, A., 2009, Econometric Applications of the Forward Search in Regression:Robustness, Diagnostics, and Graphics, *Econometric Reviews*, 28(1):21-39.

Atkinson, A., Riani, M., Cerioli, A., 2004, Exploring Multivariate Data with the Forward Search. *New York: Springer–Verlag*.

Atkinson, A., Koopman, S., Shephard, N., 1997, Detecting shocks: Outliers and breaks in time series, *Journal of Econometrics,* 80:387-422.

Box, G. and Jenkins, G., 1976, Time Series Analysis: Forecasting and Control. *San Francisco: Holden-Day*.

Box, G., Jenkins, G. and Reinsel, G., 2008, Time Series Analysis:Forecasting and Control, Fourth Edition, *New York: John Wiley & Sons*

Campano, W. and Barrios, E., 2011, Robust estimation of a time series model with structural change. *Journal of Statistical Computation and Simulation*, 81(7):909-927.

Chang, I., Tiao, G. and Chen, C., 1988, Estimation of Time Series Parameters in the Presence of Outliers, *Technometrics*, 30(2): 193-204.

Davison, A. and Hinkley, D., 1997, Bootstrap Methods and Their Application, *Cambridge: Cambridge University Press*.

Mouchart, M. and Rombouts, J., 2005, Clustered panel data models: An efficient approach fornowcasting from poor data, *International Journal of Forecasting,* 21:577– 594.

Park, J., 2011, Multiple-index approach to multiple autoregressive time series model, *New York: Springer Science.*

Parzen, E., 1968, Multiple Time Series Modeling, Department of Statistics, Stanford University, Technical Report No. 22

Prado, R., Molina, F., and Huerta, G., 2006, Multivariate time series modeling and classification via hierarchical VAR mixtures, Computational Statistics & Data Analysis, 51(3): 1445-1462.

Riani, M. and Atkinson, A., 2000, Robust Diagnostic Data Analysis: Transformations in Regression, *Technometrics*, 42(4):384-394.

Riani, M., 2004, Extensions of the Forward Search to Time Series, *Studies in Nonlinear Dynamics & Econometrics*, Volume 8, Issue 2, Article 2

Sims, C., 1980, Macroeconomics and Reality, *Econometrica*, 48(1):1-48.

Tiao, G. and Tsay, R., 1983, Multiple time series modeling and extended sample cross-correlation. *Journal of Business and Economic Statistics*, 1(1):43-56.

Tiao, G. and Box, G., 1981, Modeling Time Series with Applications, *Journal of the American Statistical Association*, 76(376):802-816.

Tsay, R. (1988), Outliers, Level Shifts, and Variance Changes in Time Series, Journal of Forecasting, 7(1): 1-20

Tsay, R., 1986, Time Series Model Specification in the Presence of Outliers, *Journal of the American Statistical Association*, 81(393):132-141.

Tsay, R., 2005, Analysis of Financial Time Series, Second Edition, *New York: John Wiley & Sons*

Veron Cruz, R. and Barrios, E., 2014, Estimation Procedure for a Multiple Time Series Model, *Communications in Statistics - Simulation and Computation*, 43(10:2415-2431.

Wang, J., and Zivot, E., 2000, A Bayesian Time Series Model of Multiple Structural Changes in Level, Trend, and Variance. *Journal of Business and Economic Statistics*, 18(3):374–386.