RESEARCH REPORT SERIES
(Statistics #2015-01)

Multivariate Seasonal Adjustment, Economic Identities, and Seasonal Taxonomy

Tucker McElroy

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: February 26, 2015

Multivariate Seasonal Adjustment, Economic Identities, and Seasonal Taxonomy

Tucker McElroy*

Abstract

The idea that economic phenomena are driven by latent components is more than a century old, and fifty years have elapsed since the assertion of Nerlove (1964) that seasonal patterns are related across time series. Although most methodological development since the 1960s has focused on univariate approaches to seasonal adjustment, a few authors have approached the problem multivariately. This paper extends these latter efforts by exploring the statistical modeling of seasonality jointly across multiple time series, using latent dynamic factor models fitted using maximum likelihood estimation. Signal extraction methods for the series then allow us to calculate a model-based seasonal adjustment. We emphasize several novel facets of our analysis: (i) we quantify the efficiency gain in multivariate signal extraction versus univariate approaches; (ii) we address the problem of the preservation of economic identities; (iii) we describe a foray into seasonal taxonomy via the device of seasonal co-integration rank. These contributions are developed through two empirical studies of aggregate U.S. retail trade series and U.S. regional housing starts. The retail series are analyzed so as to identify seasonal co-integrating rank for the sixth seasonal, allowing for a classification of the data. Our analysis of regional housing starts identifies different seasonal sub-components that are able to capture the transition from pre-recession to post-recession seasonal patterns. We also address the topic of indirect seasonal adjustment by analyzing the regional aggregate series – in this case the economic identity dictates that regional series should aggregate to the nation-wide series.

Keywords. Co-Integration, Dynamic Factor Models, Seasonality, Time Series, Trends, VAR

*Center for Statistical Research and Methodology, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, tucker.s.mcelroy@census.gov

1 Introduction

The notion that an observed time series is composed of several unobserved components has a long history in economics, going back to the mid nineteenth century in England according to Nerlove, Grether, and Carvalho (1979). The study of seasonality goes back to this epoch as well, in the field of meteorology (Buys Ballot, 1847); changing seasonality was noted by Gilbart (1852). The view of an economic time series being composed of trend, seasonal, cycle, and residual effects was exposited by Persons (1919), which substantially impacted later work in the twentieth century. The idea that the latent effects, especially the seasonal, could be common to multiple series (or at least, highly correlated) can be dated at least to Nerlove (1964, p.263): "Indeed, seasonality does not occur in isolated economic series, but seasonal and other changes in one series are related to those in another. Hence, ideally one should formulate a complete econometric model in which the causes of seasonality are incorporated directly in the equations." Bell and Hillmer (1984) provides a slightly dated, but still pertinent, overview of the literature and topics of interest in the field of seasonal adjustment. Nerlove's proposal is the subject of this paper.

This paper studies the modeling and extraction of seasonality in multiple time series, utilizing an unobserved components framework where each latent component is described via a dynamic factor model. While there is a tremendous amount of economic literature on dynamic factor models (Stock and Watson (2011) provides an overview), there is somewhat less on structural dynamic factor models wherein the identification of the factor ranks for each latent component is considerably more challenging. There are even fewer serious attempts at multivariate modeling and seasonal adjustment using an unobserved components framework. However, there exists persistent (and publicly expressed) interest among economists in multivariate seasonal adjustment, as has been highlighted recently by public concerns about the performance of univariate seasonal adjustment procedures (such as X-12-ARIMA) under the extreme economic conditions engendered by the Great Recession.

In the last twenty years or so, development of the concept of co-integration – along with its extension to component co-integrating rank – and better techniques of computation (such as state space methods) have facilitated the multivariate signal extraction project, although only a few papers seem to have addressed seasonal adjustment: Bartelsman and Cleveland (1993), Krane and Wascher (1999), and Koopman, Ooms, and Hindrayanto (2012). Herein we develop tools to model and seasonally adjust multiple economic time series, using the concept of component co-integrating rank, and explore the topics of: (i) efficiency gain in multivariate signal extraction versus univariate approaches, (ii) addressing the problem of the preservation of economic identities, and (iii) seasonal taxonomy via the device of seasonal co-integrating rank. An introduction to these three facets is provided below.

1.1 Dimension Reduction and Latent Rank

Dynamic factor models, which constitute our basic modeling framework in this paper, have been found to be quite empirically useful; see Sargent and Sims (1977), Gianone, Reichlin, and Sala (2004), Stock and Watson (2011, 2012). We next provide the statistical motivation for dynamic factor models, from the standpoint of dimension reduction and model parsimony. Consider a vector time series $\{y_t\}$, where there are m series under consideration. When m is large, it is important to consider a dimension reduction approach to the data analysis, because the number of parameters will otherwise be too large for meaningful results. Bozik and Bell (1987) appears to be one of the first papers advocating Principal Components Analysis to estimate a dynamic factor model, which can then be utilized to reduce $\{y_t\}$ to a lower-dimensional factor time series $\{x_t\}$:

$$y_t = \Lambda\, x_t + \iota_t. \tag{1}$$

Here $\{x_t\}$ is a latent vector seasonal process of dimension k and $\{\iota_t\}$ is the error process in this decomposition. Λ is an $m \times k$ matrix, called the loading matrix. The hope of such a factor analysis is that we can take k much smaller than m, while the error $\{\iota_t\}$ contributes little to the overall variation, and moreover is not serially uncorrelated. If this is the case, one can estimate Λ and the factor series using Principal Components Analysis (PCA), and proceed with modeling the estimated factors.

However, when working with monthly of quarterly raw economic time series such a decomposition often fails to result in dimension reduction, as much of the interesting trend and seasonal structure is shifted into the estimated errors, unless k is taken quite large. Notwithstanding, some authors (Stock and Watson, 2012) have found this PCA approach to be useful on seasonally adjusted data, and have drawn useful interpretations from the resulting factor series. Unfortunately, when seasonality is present in the data, it happens that the estimated factors have trend, seasonality, and cylicality all mingled together in each factor series (i.e., each component of the vector process $\{x_t\}$ will contain *all* the dynamics of the data process), and so this type of decomposition is not useful for separating out different dynamics. Another approach to estimating such a decomposition is Independent Components Analysis (ICA), but the same empirical behavior has been observed by the author.

Because the single factor model (1) is ineffective at dimension reduction in seasonal time series, and cannot cleanly separate disparate dynamics, it is natural to look for a more complex decomposition that achieves these objectives. The key is to have several factor components, each one associated with particular dynamics that are present in the series. This takes the form

$$y_t = \mu_t + \xi_t + \iota_t \tag{2}$$

This equation (2) decomposes the observed series into trend $\{\mu_t\}$, seasonal $\{\xi_t\}$, and irregular $\{\iota_t\}$ processes. The trend and seasonal in turn could be written as a dynamic factor model $\Lambda\, x_t$, where

x_t is a latent trend or seasonal process respectively, effectively generalizing (1). One could also introduce a latent cyclical process to model moderate 2-10 year swings about the long-term trend, but we avoid such devices in this paper.

The framework of (2), when appropriately extended, has proven useful in our empirical work for separating out dynamics of different types (it is also utilized in Koopman, Ooms, and Hindrayanto (2012)), while also allowing for some dimension reduction (by lowering the various k dimensions, when warranted). Other approaches, such as an initial PCA or ICA application, we have found to be unfruitful. This paper adopts the latent dynamic factor framework, and seeks to answer the following questions: how do we fit models to economic time series, such that we allow for dimension reduction? Given the models, how can we estimate and remove seasonality? What are the advantages over a univariate approach? Although there are some recent publications treating multivariate modeling and seasonal adjustment, there seems to be no systematic treatment of latent rank, and its ramifications on seasonal adjustment. In fact, our research has been partially motivated by strong public criticism of univariate seasonal adjustment procedures, such as X-12-ARIMA, and the challenges implicit in the Great Recession. See Wright (2013) for an overview of these concerns. One might hope that modeling multiple series would facilitate superior estimation of latent seasonal effects, thereby ensuring that quickly evolutive seasonality does not pollute trend and business cycle extractions, such as can occur during epochs of great change. We discuss modeling, latent rank selection, and seasonal adjustment in Section 2, and Section 3 provides applications on construction and retail series.

1.2 Maintaining Economic Identities

Although of little interest to the theoretician, economic accounting rules are extremely important to the publishers and consumers of economic data. These accounting rules may be aggregation relations across stratifications (e.g., male unemployment plus female unemployment equals total unemployment), regions (e.g., housing starts series for NorthEast, MidWest, South, and West must sum to the Total series), or epochs (e.g., three monthly figures must sum to the corresponding quarterly figure, for a flow time series). A discussion of accounting constraints is provided in Quenneville and Fortier (2012). The key challenge that arises is that data arise from diverse sources (e.g., surveys and/or administrative records) of varying quality, and are typically subject to sampling error – see Tiller (2012) for an overview – and are revised over time. Various recipes, adapted to the protocols and culture of each particular agency, are utilized to balance accounting constraints and ensure economic identities hold – these recipes include raking and utilizing controls (to a more reliable data source).

Here we are concerned with the disturbance of accounting rules resulting from seasonal adjustment. Any signal extraction, even when linear in the data, will disturb the raw data's economic identities. The naïve solution is to declare the aggregate variable's seasonal adjustment to be the appropriate aggregation of the component variables' seasonal adjustments – a procedure known as indirect adjustment. However, in many cases this results in inadequate seasonal adjustment of the aggregate, as seasonality yet remains. Where does it come from? One explanation is that the component variables are cross-correlated in their seasonal dynamics, and this is unaccounted for by a univariate seasonal adjustment methodology. Then the summation of the seasonal adjustments can still be seasonal.

Suppose that the seasonal coherency in the raw disaggregate time series is modeled and measured, and accounted for in the multivariate seasonal adjustment. Because the multivariate seasonal adjustment is produced via a filter that acts upon *all* the input series, we do not expect the cross-spectra to have seasonal peaks. This heuristic argument says that indirect seasonal adjustment is safer when utilizing a multivariate approach. We pursue this idea further in Section 3, through the analysis of regional housing starts.

1.3 Seasonal Taxonomy

The pattern of seasonality varies greatly by industry and series type (e.g., retail versus construction), but certain facets are common to batches of coherent time series. It is of interest to group and classify series by these features, in order to understand which series are driven by a common latent seasonal process. This classification, or taxonomy, can assist in detecting new patterns (e.g., departure of one series from its cluster); it can help in understanding redundancies and coherence (e.g., do some series lead or lag others with respect to their seasonal movements?); and it can provide a general portrait of the economic variable (Granger, 1966). The applications of this taxonomy are at this stage speculative, but may include the following: identification of batches of series suitable for joint multivariate analysis and adjustment (or forecasting); identification of structural changes to the economy, when series that were formerly classified as belonging to the same species no longer do so; identification of data inaccuracies, when co-integrating relations are violated at particular sample points.

Given that taxonomy is of interest, tools are needed to provide measures of clustering. Coherence is the analogue of correlation for time series, but here we focus on seasonal coherence, i.e., high spectral coherence at seasonal frequencies (those of the form $2\pi j/s$ for integer j, and s the seasonal period). The latent component model (2) implies that the spectral density (of the differenced series) evaluated at seasonal frequencies is equal to the spectrum driving the seasonal $\{\xi_t\}$, so that there is an immediate connection. As is shown in Section 2, reduced rank in the covariance matrix of the white noise driving $\{\xi_t\}$ is known by the term "reduced seasonal rank," and implies there is seasonal co-integration in the observed series; it corresponds to maximal possible coherence at the seasonal frequencies. We propose to utilize seasonal rank as a measure of taxonomic proximity. In particular, if m series have seasonal rank equal to one, they belong to the same species, all of them being driven by the same latent one-dimensional seasonal process.

2 Modeling Methodology

This section provides a discussion of a multivariate time series model involving latent components for trend, seasonality, cycle, and irregular. We discuss co-integration and latent rank, and review signal extraction methodology. For further background, see Harvey (1989), Durbin and Koopman (2001), and Koopman et al. (2012).

2.1 Latent Component Models

Let us consider (2) in more detail; we will further decompose the seasonal into its atomic (referred to as "trigonometric" by Harvey (1989)) components corresponding to the complete factorization of the seasonal unit root differencing polynomial. The latent processes are related to the observed m-dimensional time series $\{y_t\}$ via (2), and each latent process in turn is a difference stationary vector time series driven by potentially collinear white noise. When the white noise is collinear, the latent process is said to be "common," and there is a reduced dimension representation. When the white noise is not collinear, the latent process is said to be "related", and there is no dimension reduction possible; if the white noise covariance matrix is also diagonal, then the latent process is said to be "unrelated."

We assume that the latent trend process $\{\mu_t\}$ is differenced to stationarity by application of $(1-B)^d$ for $d = 0, 1, 2$, where this scalar polynomial is applied to each component of the process, i.e.,

$$(1-B)^d \mu_t = \epsilon_t^\mu$$

for an m-variate white noise process $\{\epsilon_t^\mu\}$ of covariance matrix Σ^μ. Likewise, the latent seasonal process $\{\xi_t\}$ is reduced to stationarity by application of $U(B) = 1 + B + \cdots + B^{s-1}$ where s is the seasonal period. This operator is factorized into

$$U(B) = \prod_{j=1}^{s/2-1} (1 - 2\cos(2\pi j/s)B + B^2)\,(1+B)$$

when s is even, and when s is odd we obtain

$$U(B) = \prod_{j=1}^{(s-1)/2} (1 - 2\cos(2\pi j/s)B + B^2).$$

We focus on the s even case below, as modifications for s odd are trivial to make. We next suppose the existence of latent *atomic* seasonal processes $\{\xi_t^{(j)}\}$ for $1 \leq j \leq s/2$, which are defined such that a single factor of $U(B)$ reduces each atomic seasonal to stationarity. By definition, the jth atomic seasonal $\{\xi_t^{(j)}\}$ for $1 \leq j \leq s/2$ satisfies

$$
\begin{aligned}
(1 - 2\cos(2\pi j/s)B + B^2)\xi_t^{(j)} &= \epsilon_t^{(j)} \qquad 1 \leq j \leq s/2 - 1 \\
(1 + B)\xi_t^{(s/2)} &= \epsilon_t^{(s/2)},
\end{aligned}
$$

where each $\{\epsilon_t\}$ is m-variate white noise with covariance matrix Σ (i.e., $\Sigma^{(j)}$ for the jth atomic seasonal, and Σ^{ι} for the irregular).

The latent seasonal process $\{\xi_t\}$ is related to the atomic seasonal components via aggregation: $\xi_t = \sum_{j=1}^{s/2} \xi_t^{(j)}$. It is easy to check that its minimal differencing polynomial is $U(B)$ when all the atomic seasonal processes are non-zero. The difference polynomial for the aggregate $\{y_t\}$ is $(1-B)^d U(B)$, and by its application we obtain

$$
(1 - B)^d U(B) y_t = g_\mu(B)\, \epsilon_t^{\mu} + \sum_{j=1}^{s/2} g_j(B)\, \epsilon_t^{(j)} + g_\iota(B)\, \iota_t \tag{3}
$$

$$
\begin{aligned}
g_\mu(B) &= U(B) \\
g_j(B) &= (1 - B)^d \prod_{k \neq j} \delta^{(k)}(B) \\
g_\iota(B) &= (1 - B)^d U(B) \\
\delta^{(j)}(B) &= 1 - 2\cos(2\pi j/s)B + B^2 \qquad 1 \leq j \leq s/2 - 1 \\
\delta^{(s/2)}(B) &= 1 + B.
\end{aligned}
$$

The differenced observed process on the left hand side of (3) will then be denoted by $\{\partial y_t\}$. For all of our applications (model fitting and signal extraction) it is necessary to compute the autocovariance function of each summand process in (3), so we now discuss how these functions can be easily computed. First note that the spectral density of $\{\partial y_t\}$ is real-valued and is given by

$$
f(\lambda) = |g_\mu(z)|^2\, \Sigma^{\mu} + \sum_{j=1}^{s/2} |g_j(z)|^2\, \Sigma^{(j)} + |g_\iota(z)|^2\, \Sigma^{\iota} \tag{4}
$$

with $z = e^{-i\lambda}$, under the assumption that all the latent process' white noises are uncorrelated with one another, and are uncorrelated with $\{\iota_t\}$. Each summand of (4) is a known scalar function times a covariance matrix Σ, and hence corresponds to the spectral density of a simple vector moving average – the autocovariance is then extremely easy to compute, and we can simply sum up these autocovariances to obtain the sequence for $\{\partial y_t\}$. These autocovariances, together with the multivariate Durbin-Levinson algorithm (Brockwell and Davis, 1991), provide a stable and efficient method for computing the Gaussian likelihood.

There are models implied for each of the individual series, which of course can differ quite a bit from a univariate model fitted to the particular series. Due to the extremely simple structure of the unobserved component models, these implied models are simple to derive. Let e_ℓ denote the ℓth unit vector of $\mathbb{R}^m$; then the ℓth series has trend innovation variance $e_\ell' \Sigma^\mu e_\ell = \Sigma^\mu_{\ell,\ell}$, and so forth. If we filter the ℓth series $\{y_t^{(\ell)}\}$ with the univariate signal extraction filter (for details, see below) corresponding to this implied univariate model, we obtain $\mathbb{E}[s_t^{(\ell)} | \{y_t^{(\ell)}\}]$, which can be quite different from $\mathbb{E}[s_t^{(\ell)} | \{y_t\}]$. Moreover, for a Gaussian process the MSE that is generated from the same methodology will correspond to $\mathrm{Var}[s^{(\ell)} | \{y_t^{(\ell)}\}]$ rather than the smaller $\mathrm{Var}[s^{(\ell)} | \{y_t\}]$.

2.2 Collinearity, Orthogonality, and Co-Integration

The case of collinear latent innovations can now be discussed. If any latent process' white noise covariance matrix is not full rank, i.e., has a zero eigenvalue, then collinearity results. The covariance matrix Σ has a unit lower triangular Cholesky decomposition:

$$\Sigma = L\, D\, L',$$

where L is unit lower triangular and D diagonal with non-negative entries. In such a decomposition, the diagonal entries of D are the successive Schur complements of the matrix, and are interpretable as partial variances (see below). If the rank is $k \leq m$, then $m - k$ of these partial variances will be zero; let J denote the sub-indices of $\{1, 2, \cdots, m\}$ such that $d_j > 0$ for $j \in J$. Then with $L_{\cdot j}$ denoting the jth column of L, we can write

$$\Sigma = \sum_{j \in J} d_j\, L_{\cdot j}\, L_{\cdot j}'. \tag{5}$$

Note that the partial variances need not be ordered, so that zero values of the diagonals can occur at any index (however, a value of $d_1 = 0$ will typically not occur in practice, as it means that the first variable of that latent component is deterministic). If estimating Σ through a parametric model – say via Maximum Likelihood Estimation (MLE) – we can proceed as described in Pinheiro and Bates (1996): all lower triangular values of L are unconstrained real numbers, whereas the non-zero values of D can be described as the exponentials of real numbers. Clearly, the number and format of such parameters depends on knowing the rank k, and each choice of k constitutes a different model, requiring separate estimation.

Each choice of restrictions on the rank of D constitutes a nested model within the nesting model, which is the fully unconstrained case wherein all covariance matrices have full rank. In order to obtain a more parsimonious model, it is of interest to determine whether collinearity of the innovations is valid. Because co-integration tests in the econometric literature are focused on the case of common trends (see Nyblom and Harvey (2000, 2001)), we take a fresh approach to the problem of seasonal co-integration that is based off the MLEs. Our method here lacks a distribution theory, although given the asymptotic distribution of the MLEs it seems plausible that such a theory could be developed. We are looking for any conditional variances that are suitably close to zero – this is further developed in the next subsection. (It is important *not* to use a bounding box in the nonlinear optimization producing the MLEs.)

Given one or more small values for the conditional variances, we consider the nested model given by setting these Schur complements to zero, thereby obtaining the index set J of size k. The corresponding columns of L are also eliminated – these are the $L_{\cdot j}$ for $j \notin J$ in (5). Labeling the resulting rectangular lower triangular matrix by Λ and the diagonal matrix of corresponding non-zero Schur complements by Δ, it is seen that this Λ exactly corresponds to the factor loading matrix described in (1). The corresponding factor time series corresponds to a (nonstationary) latent stochastic process of dimension k, which has no cross-correlation; indeed, the innovations driving the factor latent stochastic process (whether trend or seasonal) will have covariance matrix Δ. This is the latent dynamic factor model interpretation of the model described herein.

Determining which Schur complements to replace with zero requires some care (next subsection), but note that any such zeroes result in a model that is nested on the boundary of the parameter space (since each $d_j \geq 0$). The likelihoods for the nested and nesting models can then be directly compared. Because the distribution – under the null hypothesis that the nested model is correct – of the log likelihood ratio is not χ^2, and the true mixture distribution is unknown, we recommend instead that one do an AIC comparison. An important caution is that such zeroes should not be placed in Σ^ι, because then the spectral density f will be non-invertible. Actually, f is allowed to be non-invertible at a finite number of frequencies (see the discussion in McElroy and Trimbur (2012)). For example, it is permissible to have collinear trend innovations and/or collinear seasonal innovations, so long as the irregular has full rank.

There are other parameter restrictions of the matrix L that may be of interest. If all the off-diagonal entries are constrained to be zero, then the corresponding Σ will be diagonal, and there will be no cross-correlation between components of the corresponding random vector. This restriction is tantamount to fitting a univariate model to each series, with model fit determined in an aggregate sense across all m series. We refer to this particular sub-model as the orthogonality restriction; in a sense it is the opposite of the collinear innovations case, and also involves a reduction in the number of parameters, from $\binom{m+1}{2}$ down to m.

Individual lower triangular entries of L correspond to partial covariances, and each entry may vary independently of the others. As established in McElroy (2014), the ijth entry of L equals d_j^{-1} (when $d_j > 0$) times the partial covariance between the ith and jth variables, conditional on variables one through $j-1$. The variables here refer to the various components of each differenced latent component ϵ_t (for any t). When $i > j$, we obtain the lower triangular entries of L, which as a covariance can be any real number. When $i = j$, the partial covariance becomes the partial variance of the jth variable, and just equals d_j. When $i < j$, the partial covariance is zero because the ith variable is in the conditioning set, and this corresponds to the upper triangular entries of L.

The decision to replace a small entry of L with a zero can be made on the basis of the statistical uncertainty of the parameter. Because we utilize a Gaussian likelihood, the inverse of the Hessian of the objective function should provide an estimate of the parameter error covariance matrix, due to the efficiency of MLEs. The resulting nested model can be checked against the nesting competitor via the GLR test, utilizing χ^2 quantiles, because the parameter restriction of zero does not lie on the boundary of the parameter space. Because collinearity can eliminate entire columns of the L matrix, one should determine collinearity first, and then pursue orthogonality.

We next discuss the relationship to co-integration, which is also discussed in McElroy and Trimbur (2012). Generalizing the basic concepts presented in Engle and Granger (1987) and Stock and Watson (1988), we say that when an m-vector α exists such that $\{\alpha' y_t\}$ has reduced non-stationarity, then α is a co-integrating vector. By reduced non-stationarity, we mean that the minimal differencing polynomial required to reduce $\{\alpha' y_t\}$ to stationarity (up to fixed effects) has lower degree than the polynomial required for the original $\{y_t\}$. If $\{\alpha' y_t\}$ has only trend nonstationarity, α is said to be a seasonal co-integrating vector, whereas if there only seasonal nonstationarity remains, α is said to be a trend co-integrating vector.

Given our particular latent factor model, α is a jth atomic seasonal co-integrating vector if and only if α is a left null-vector of $\Sigma^{(j)}$, whereas α is a trend co-integrating vector if and only if α is a left null-vector of Σ^{μ}. This follows from the form of our model – see equation (4). A basis for the co-integrating vectors' space can be computed from the rows of L^{-1}, utilizing the rows that correspond to zero d_j values. For example, a rank of one implies that there exists a basis of $m-1$ co-integrating vectors. From the standpoint of taxonomy, we say that all such time series belong to the same latent species, where the type of species is defined by the particular latent frequency. For example, if $\Sigma^{(j)}$ has rank one we say that all the series belong to the same jth-atomic seasonal species.

Observe that Σ (excepting the case of the irregular covariance matrix) is the value of the spectral density f of the differenced time series at the latent process' corresponding frequency, i.e., $\Sigma^{\mu} = f(0)$ and $\Sigma^{(j)} = f(\pi j/s)$. This discussion leads to the following definition.

Definition 1 *Two time series following a latent dynamic factor model are j-equivalent if and only if the bivariate spectrum evaluated at the jth seasonal frequency ($0 \leq j \leq s/2$) has rank one. We denote this equivalency with the notation $\sim_j$.*

This is well-defined, in the sense that permuting the series' order does not change the rank. Moreover, we have the following result:

Proposition 1 *$\sim_j$ is an equivalence relation, and therefore partitions the set of difference stationary time series.*

We make a few comments about this result. First, when estimating the covariance matrix of multiple time series, we can only make probabilistic assertions about the rank, and therefore statistical errors can arise; also, changing samples can alter the species classification. Second, there is a classification pertaining to each frequency $2\pi j/s$, so that up to $s/2$ different partitions exist for these types of time series. Two series may be equivalent according to frequency zero, but not according to frequency $j = 1$. This concept is applied in Section 3.1, with $j = 6$.

2.3 Taxonomic Identification

Having proposed a theoretical definition for taxonomy, it is crucial to have an empirical procedure for its application on data. Formal tests for whether particular d_js are zero is the ideal target for the statistician, but using the Hessian – as with the partial covariances in L – will lead to incorrect inference because the parameter space for the Schur complements has boundary. Our recommendation at present is to examine the d_j parameter estimates in the context of other entries in Σ, and thereby obtain nested models – by reducing the rank of Σ – to fit, and finally to use AIC comparisons to evaluate the models. To that end, we must study how the partial variance and partial covariance parameters determine singularity and near singularity.

First consider the $m = 2$ case, which contains the main features of the general case as well. The partial correlation between the first and second variables (so it is the same as the unconditional correlation) is

$$\kappa_{21} = \frac{\text{sign}(L_{21})}{\sqrt{1 + d_2/(L_{21}^2\, d_1)}}$$

when $d_1 > 0$ (this case rarely comes up empirically, because it corresponds to the first component of the latent process being completely deterministic). This formula shows that the absolute correlation approaches unity as $L_{21} \to \pm\infty$, or if $d_2/d_1 \to 0$. Therefore, although small values of d_2 indicate singularity, it is also sufficient merely that d_2 be small relative to d_1. Defining $\tau_{21} = \log(d_2) - \log(L_{21}^2\, d_1 + d_2)$, we have $\kappa_{21} = 1 - e^{\tau_{21}}$, and values of τ_{21} tending to $-\infty$ correspond to the rank of Σ going from two to one. Observe that the determinant of Σ is equal to the product of the diagonal entries times $1 - \kappa_{21}^2$, which represents a scale-free quantity. We can define a condition number by computing the scale-free determinant of Σ in log-scale, i.e.,

$$\log(1 - \kappa_{21}^2) = \tau_{21}.$$

For example, if $|\kappa_{21}| = .9$ then $\tau_{21} = -1.66$, whereas $|\kappa_{21}| = .99$ corresponds to $\tau_{21} = -3.92$ and $|\kappa_{21}| = .999$ implies $\tau_{21} = -6.22$. The advantage of examining τ_{21} and κ_{21} to determine approximate singularity is that we've removed the scale of the series from the analysis.

Now let us generalize to $m > 2$. Let ϵ denote the random m-vector of covariance matrix Σ; as noted above,

$$d_j = \text{Var}_{\epsilon_1,\cdots,\epsilon_{j-1}}(\epsilon_j) \qquad L_{ij} = d_j^{-1}\, \text{Cov}_{\epsilon_1,\cdots,\epsilon_{j-1}}(\epsilon_i, \epsilon_j).$$

When $d_j = 0$, the jth column of L will be eliminated in the formula for Σ, so that L_{ij} need not be defined. Defining $\kappa_{ij} = \mathrm{Corr}_{\epsilon_1,\cdots,\epsilon_{j-1}}(\epsilon_i, \epsilon_j)$, we wish to express it in terms of the L_{ij} and d_j parameters; values of the absolute partial correlations close to unity will indicate near singularity, and will also allow us to identify exactly which d_j are chiefly responsible (which condition numbers cannot do). These relations are given in the Appendix, and it can be shown that $\det \Sigma_m$ is equal to the product of the diagonal entries times $\prod_{i>j}(1-\kappa_{ij}^2)$. With $\tau_{ij} = \log(1-\kappa_{ij}^2)$, we define our condition number to be

$$\sum_{i>j} \log(1-\kappa_{ij}^2) = \sum_{i>j} \log \tau_{ij}.$$

This generalizes the $m = 2$ case. Any values of κ_{ij}^2 close to unity (or equivalently, large negative values of τ_{ij}) indicate that the ith variable can potentially be eliminated. If $\kappa_{ij}^2 \approx 1$ for several j for a given i, no additional action needs to be taken. For example, we might propose eliminating variables until the condition number $\sum_{i>j} \tau_{ij}$ exceeds $\log(.95)$, so that 95% of the scale-free determinant is explained. We apply these measures to a trivariate analysis in Section 3.1 below.

The goal of such elimination of variables is parsimony. Conceptually, we may imagine an essential dimension for a given model and time series, which corresponds to that subset of Euclidean space wherein most of the likelihood has its probability mass. This essential dimension is the number of parameters, and its ratio to the full sample size should be roughly constant across sub-spans of the data. As an *ad hoc* rule of thumb, we might venture that the essential dimension should be no more than the number of series (m) times the square root of sample length n, so that the ratio of data to essential dimension should be less than $\sqrt{n}$.

2.4 Multivariate Filtering

Here we consider the main application of the preceding modeling methodology to signal extraction by describing the minimum MSE linear filters corresponding to the fitted structural model. We rely upon the formulas derived in Theorem 2 of McElroy and Trimbur (2012). In order to implement the smoothing formula, we write the data vector as collected by time and listed over vector components. Then the covariance matrices for the differenced latent components can be computed quite easily. Let s and n denote signal and noise, where s consists of the sum of any components given in (2) that are of interest, and n consists of the remaining components. For example, n could consist of the sum all $s/2$ atomic seasonals, and s consist of the sum of trend and irregular; then the signal extraction corresponds to seasonal adjustment.

Identification of the signal components of interest in turn implies a signal differencing operator $\delta^s(B)$, and a spectrum f_s for the differenced signal; similarly, we will have a noise differencing operator $\delta^n(B)$ and noise spectrum f_n. The signal and noise spectra will actually correspond to various summands of (4), in the following sense: the squared gain of the noise differencing operator will multiply the signal spectrum in (4), whereas the squared gain of the signal differencing operator will multiply the noise spectrum. In the seasonal adjustment example, the noise differencing operator is $U(B)$ and the signal spectrum is

$$f_s(\lambda) = \Sigma^\mu + |1-z|^{2d}\, \Sigma^\iota.$$

On the other hand, the signal differencing operator is $(1-B)^d$ and the noise spectrum is

$$f_n(\lambda) = \sum_{j=1}^{s/2} \Big| \prod_{k\neq j} \delta^{(k)}(z) \Big|^2 \, \Sigma^{(j)}.$$

This is just one example; we might be interested in various atomic seasonals as signals, or combinations of such, and in each case f_s and f_n can be defined. The signal extraction filter for a bi-infinite sample has frf given by

$$f_s(\lambda)\, f^{-1}(\lambda)\, |\delta^n(z)|^2,$$

as proved in McElroy and Trimbur (2012) for cases including co-integration. For samples of finite length we instead use a matrix filter F, whose formula is also computed from the signal and noise spectra, as well as the differencing operators. For a sample of size n, there are two alternative ways of stacking the data into a matrix. First we have

$$Y = [y_1 \, y_2 \cdots y_n],$$

which is $m \times n$ dimensional. We call this *series-by-time*, and is conventional in many textbooks. The other representation is

$$Y' = [y^{(1)} \, y^{(2)} \cdots y^{(m)}],$$

where $y^{(j)}$ is an n-vector consisting of all observations for the jth series. Thus Y' is $n \times m$ dimensional, and is referred to as the *time-by-series* representation. The description of F in McElroy and Trimbur (2012) presumes the time-by-series representation, so that $F \,\mathrm{vec}[Y']$ yields vecced signal extraction estimates written in time-by-series format.

Now the application of F to the vectorization of the time-by-series data matrix Y' is appropriate when the mean is zero. When the mean of the differenced series is nonzero, say given by an m-vector $\mathbf{m}$, then instead we apply F to the mean-corrected time-by-series data, where the mean correction involves subtracting $\mathbf{m} \otimes \tau$ from $\mathrm{vec}Y'$, where τ is a column vector $\tau = [1^d, 2^d, \cdots, n^d]'$ for $d \geq 0$. For the seasonal, cycle, and irregular components we compute $F\,[\mathrm{vec}Y' - \mathbf{m} \otimes \tau]$, but for the trend we compute $\mathbf{m} \otimes \tau + F\,[\mathrm{vec}Y' - \mathbf{m} \otimes \tau]$. This procedure is justified in the Appendix (in the Supplement).

Moreover, the error covariance matrix – whose diagonal entries are the signal extraction MSEs, or conditional variances – is given by by a matrix V, which is expressed (McElroy and Trimbur, 2012) as the difference of two positive definite matrices. Essentially, the first matrix corresponds to univariate signal extraction error, and the second matrix brings cross-series information into play, in order to increase precision when warranted. Section 3.2 further explores the precision increases due to multivariate signal extraction. As discussed in Section 2.1, we propose to measure the ratios

$$\frac{\mathrm{Var}[s_t^{(j)} | \{y_t\}]}{\mathrm{Var}[s_t^{(j)} | \{y_t^{(j)}\}]}$$

for each $1 \leq j \leq m$; here the numerator is given by the appropriate entries of the matrix V, and the denominator is computed from V under the assumption that all series are uncorrelated with one another. This ratio of MSEs will give an idea of how much reduction in MSE is attributable to the multivariate filtering.

2.5 Indirect Seasonal Adjustment

The problem of indirect seasonal adjustment is that the total of the seasonal adjustments of several disaggregate series (e.g., corresponding to regions) might not equal the seasonal adjustment of the total, if this latter adjustment is done separately. This is actually a repercussion of conditional expectation calculations, and has nothing to do *per se* with non-linearity of filtering. The direct seasonal adjustment of the aggregate would be $\mathbb{E}[\sum_j n_t^{(j)} | \{\sum_j y_t^{(j)}\}]$, where $n_t^{(j)}$ is the jth series' non-seasonal component; indirect multivariate adjustment proceeds by summing the disaggregate multivariate adjustments, namely $\sum_j \mathbb{E}[n_t^{(j)} | \{y_t\}]$. Note that the latter information set includes the information set of the direct case, so we should favor indirect adjustment. We should also prefer this indirect multivariate adjustment to univariate adjustment, which proceeds by summing the univariate disaggregate adjustments: $\sum_j \mathbb{E}[n_t^{(j)} | \{y_t^{(j)}\}]$.

We propose adopting the indirect multivariate seasonal adjustment, as its expectation conditions on the most amount of information. The economic identity – which for example states that the total shall equal the sum of the disaggregate series – is preserved automatically, and both the total seasonal adjustment and the individual seasonal adjustments are coherent, as they are computed from the same information set. Part of the problem with direct seasonal adjustment is that it proceeds from the information set $\{\sum_j y_t^{(j)}\}$, which can be quite different from the univariate information sets.

The signal extraction MSE for the total can be determined from the error covariance matrices for the individual series; if V is the error covariance matrix for the multivariate seasonal adjustment, then

$$[[1, 1, \cdots, 1] \otimes 1_n] \;\; V \;\; [[1, 1, \cdots, 1] \otimes 1_n]'$$

is the error covariance matrix for the aggregate, whose diagonal entries provide the time-varying MSEs. This concept is further explored in Section 3.3.

This approach becomes problematic if the data has been log transformed. The application of a log transformation implies a multiplicative decomposition in the original scale of the data, so that $y_t = s_t \cdot n_t$ for signal and noise vectors s_t and n_t, and $\cdot$ denoting Hadamard product. The total is defined as $z_t = [1, 1, \cdots, 1]' y_t$, and it is unclear how to define the corresponding signal and noise decomposition for z_t, since it equals

$$z_t = \sum_{j=1}^{m} s_t^{(j)} n_t^{(j)}.$$

If s_t is the seasonal, and n_t the nonseasonal component, then one possibility is to arbitrarily define $\sum_{j=1}^{m} n_t^{(j)}$ as the nonseasonal component of the aggregate z_t, and their quotient as the seasonal. In order to apply classical signal extraction methodology, a log transformation is applied to each component of y_t, which transforms the multiplicative decomposition into an additive decomposition. After applying the signal extraction methodology in the log domain, one can exponentiate all the estimates in order to translate results into the original scale. Unfortunately, these transformations will interfere with the indirect multivariate adjustment method described above[1].

We have no satisfactory resolution of this problem, and instead recommend shortening series' length to a degree such that a log transformation is no longer necessary[2]. For monthly data, ten years of data is typically sufficient to fit common univariate models, and longer spans actually seem to warrant more complicated models (e.g., time-varying coefficient ARIMA models, or regime-switching models). When modeling multivariately, we have additional sample cross-sectionally, and hence there is less danger in restricting the series length. We henceforth proceed to work with such shortened spans, finding that our simple models work quite well without a log transformation; this enables the method of indirect adjustment described above.

3 Empirical Illustrations

We first model three retail series, and give an application of taxonomy; then, we consider multivariate seasonal adjustment of four regional construction series, with application to indirect adjustment.

3.1 Retail Series and Seasonal Taxonomy

In the course of modeling retail and construction series, we found that inclusion of a business cycle component gives little improvement to the overall models, and moreover had an obnoxious impact on seasonal adjustment: either one allows the cycle period to freely vary – in which case it can become coincident with seasonal frequencies and lead to misidentification – or one constrains the cycle period arbitrarily to some band. This latter choice produced period estimates on the

[1]Some statistical agencies (e.g., the Bureau of Economic Analysis) prefer using a log transform, so that the final seasonal adjustment results can be interpreted as percentage adjustments, and utilize benchmarking algorithms to enforce accounting rules. This approach is alien to the methodology of this paper, which employs model-based methods so as to carefully quantify signal extraction uncertainty.

[2]Many series manifest their seasonal amplitude as proportional to the trend level, and thus warrant a multiplicative decomposition; by shortening the series, the impact of trend growth becomes linearized.

boundary, and moreover there was little evidence in the estimated spectra to indicate a cycle's existence in the first place. For these reasons, we forego inclusion of the cycle.

Thus we employ the model (2) without modification. For the purpose of seasonal adjustment, the sum of all six seasonal processes constitutes the seasonal. The parameters present in the six covariance matrices allows for nuanced modeling of the seasonal structure, and is actually highly necessary to capture the transition of seasonal behavior from pre- to post- Great Recession. Essentially, the additional parameters due to use of atomic seasonals allows each of the six spectral peaks in the spectral density to have parameters controlling their width and height independently of the other peaks. If instead only one parameter (or matrix) controls all the spectral peaks' features, then some peaks may be modeled with heights and/or widths that are inappropriate. This can gravely impact seasonal adjustment: if a peak is modeled too narrowly, then the resulting seasonal adjustment filter's frequency response function will have seasonal troughs that are too narrow, and seasonality at that particular seasonal frequency may well remain. This problem is of lesser concern in times of economic regularity, but when transitioning between economic regimes (e.g., from pre-GR to post-GR) in reality the spectral seasonal peaks grow wider, reflecting the more highly evolutive nature of seasonality.

These comments result from our own experience modeling these series with the common seasonal model, with data spans encompassing the GR. Spectral peak width does indeed vary between seasonal frequencies, reflecting the change *to trend as well as seasonal patterns* (see figures below). Our presentation focuses on the final, best results possible, but it is worth enunciating on methods that will surely fail. Our custom R code allows one to include/exclude any of these components, and impose any degree of collinearity on any component, and also to enforce parameter constraints (such as independence).

For our first example, we study three series from the Advance Monthly Sales for Retail and Food Services data (representing a preliminary estimate of each series, featuring the largest retailers), which are published each month. We consider the following fairly highly aggregated series: 448 (Clothing and Clothing Accessories Stores); 451 (Sporting Goods, Hobby, Book, and Music Stores); 452 (General Merchandise Stores). The sampling period was 1992 through 2012, and values pertain to the entire U.S. geography. Each series was first cleaned of fixed effects (via regression ARIMA modeling of the log-transformed data), and then modeled with an unrestricted model including $I(2)$ trend, irregular, and six seasonals.

There are $n = 252$ observations for each series, for a total of 756 data points, resulting in relative convergence; residual acf plots indicated that no serial correlation remained. This full model involves six parameters for each latent component, for a total of 48, plus an additional three parameters for the mean of the differenced data. Hence the ratio of data to parameters is 14.82, which is less than $\sqrt{252}$ (our *ad hoc* upper bound), so we may wish to find a more parsimonious model if possible. We will employ the co-integration identification techniques discussed above.

The initial trivariate model is modified to allow for co-integration at the sixth frequency, obtained by enforcing that $d_3 = 0$ for that component. Whereas the nesting model has AIC of -1006.64, the nested model with 50 parameters has AIC of -1007.052, being marginally better. Signal extraction estimates for trend as well as the six seasonals (and their aggregate) for the three series are given in Figure 1. In this case, the signal extraction uncertainty was quite low, but there is a subtle shading around each estimate corresponding to a two standard error width confidence interval.

Next, we considered the three bivariate analyses from the pairings (448,451), (451,452), and (448,452). The first two analyses were run with unconstrained models, and no co-integration was identified. The last analysis – as expected given the structure of the trivariate covariance matrix – yields a high degree of correlation between the two series, arguing that we can reduce to a rank one nested model. Refitting, the AIC drops from -370.76 to -370.81 with the loss of one parameter; barely an improvement.

To obtain a better understanding of the correlation patterns, we display the partial correlation matrices L for the various models. In the unconstrained trivariate model, we have

$$L = \begin{bmatrix} 1 & 0 & 0 \\ .027 & 1 & 0 \\ .949 & .058 & 1 \end{bmatrix}.$$

Also the Schur complements are $d_1 = e^{-8.04}$, $d_2 = e^{-11.45}$, and $d_3 = e^{-27.90}$. As a result, the partial correlations are $\kappa_{21} = .14707$, $\kappa_{31} = .99994$, and $\kappa_{32} = .99999$. The first two are actually interpretable as straight correlations, whereas the third (κ_{32}) is the correlation for series 451 and 452 conditional on series 448. We conclude that 448 and 451 have little relationship, but 448 and 452 are highly linked – also 451 appears to be highly linked to 452, but this is only "through" the linkage of 448 to 452. The co-integration measures are $\tau_{21} = -.02$, $\tau_{31} = -9.03$, and $\tau_{32} = -10.82$; the latter two values indicate that d_3 should be set to zero to get a nested model. The values of the parameters in L change slightly when fitting the nested model – only in the third decimal places (while the Schur complements d_1 and d_2 change minorly as well). Next, the L matrices for the three bivariate unconstrained models, in order, are:

$$L = \begin{bmatrix} 1 & 0 \\ .012 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1.132 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ .947 & 1 \end{bmatrix}.$$

More importantly, the Schur complements are $d_1 = e^{-8.03}$ and $d_2 = e^{-11.40}$ for the first model, $d_1 = e^{-11.43}$ and $d_2 = e^{-8.21}$ for the second model, and for the third model $d_1 = e^{-8.05}$ and $d_2 = e^{-28.92}$. The (partial) correlations between the two variables are .221, .301, and 1 (to numerical precision) respectively. This confirms that series 448 and 452 are strongly correlated, and this is responsible for the rank reduction in the trivariate model. The co-integration measures τ_{21} are

2.973, 2.303, and -20.766 respectively. So we have evidence to pursue a nested model for the third bivariate analysis, and the resulting parameter in L only changes in the third decimal place, again.

In terms of taxonomy, we have 448 $\sim_6$ 452, whereas 448 $\not\sim_6$ 451 and 451 $\not\sim_6$ 452. The co-integrating vector β for the three series, i.e., that vector such that its application reduces the nonstationarity by the factor $1+B$, is given by taking the nesting trivariate model and computing the bottom row of L^{-1}, or in other words

$$\beta' = [L_{32}\, L_{21} - L_{31}, -L_{32}, 1] = [-.947, -.058, 1].$$

If we apply β' to the three series, we should obtain a series that has reduced order of integration. That is, excepting possible deterministic terms in the null space of $1+B$ (i.e., sequencies proportional to $(-1)^t$), the application of $(1-B)^2(1+B^2+\cdots+B^{10})$ (observe that $U(B)$ divided by $1+B$ equals $1+B^2+\cdots+B^{10}$). To the extent that the signal extraction estimates share the co-integration properties of the underlying signals, we can expect that application of the co-integrating vector to the 6th seasonal extraction will be a stationary time series, plus a deterministic function of period 2. This is exactly the case: we computed the application of β to the 6th seasonal extraction, and find the result to be purely deterministic sine wave of period 2; similarly, the co-integrating vector for the third bivariate analysis is $\beta' = [-.947, 1]$, and again its application to the 6th seasonal extraction is purely deterministic with frequency π.

This shows that 448 and 452 are in the same species (according to $\sim_6$), and as a result their 6th seasonal extractions are approximately – up to a deterministic function – scale multiples of one another, across all time points. The applications of taxonomy suggested in the introduction do not apply here, because this form of coherency carries across the GR. A full taxonomic classification of the retail database would provide insight into how different variables are related, and also indicate which batches of series would be amenable to multivariate seasonal adjustment. This could have possible ramifications to missing data problems, changing sampling frequency, and anticipation of data revisions – these speculations are left unto future work for refinement, but are only mentioned here to provoke interest.

3.2 Modeling Housing Starts

For a second illustration we consider housing starts data that is published by the U.S. Census Bureau on a monthly basis, for the regions corresponding to South, NorthEast (NE), West, and MidWest (MW). We study "New Residential Construction, 1964-2012, Housing Units Started, Single Family Units" from the Survey of Construction of the U.S. Census Bureau, available at `http://www.census.gov/construction/nrc/how_the_data_are_collected/soc.html`. It is also of interest to seasonally adjust the Totals, defined to be the straight sum of the four regional series. We proceed to illustrate the indirect multivariate seasonal adjustment procedure via the housing starts data. We also use the co-integration techniques described in the previous subsection rather freely, and will report only the ultimate refined models.

The four series were first cleaned of additive regression effects, such as trading day and holiday effects, and then multivariate structural models involving $I(2)$ trend, irregular, and six seasonals. As mentioned before, results involving a cycle and/or an aggregate seasonal gave poor results, and were abandoned. We have run analyses for the entire span of data, but here focus our discussion on the span 2004-2012, nine years that include both pre- and post- GR time periods. The sample size is $n = 108$ with $m = 4$ series for a total of 432 observations; the unconstrained nesting model has AIC of 1009.101, with 84 parameters (in each latent component, there are 6 parameters in L and 4 parameters in D, plus four mean parameters), whereas the restricted nested model has AIC of 951.85 with 54 parameters. This substantial increase in parsimony is achieved with the following choices of J (the subset of $\{1, 2, 3, 4\}$ corresponding to nonzero Schur complements) for the trend and six seasonals, respectively:

$$J_\mu = \{1\}, \quad J_{\xi^{(1)}} = \{1, 2\}, \quad J_{\xi^{(2)}} = \{1, 2\}, \quad J_{\xi^{(3)}} = \{1\}$$
$$J_{\xi^{(4)}} = \{1, 2\}, \quad J_{\xi^{(5)}} = \{1, 2\}, \quad J_{\xi^{(6)}} = \{1\}.$$

As usual, the irregular is enforced to have full rank. The restrictions were first identified using the condition numbers described in Section 2.3. The final residuals from both this nested model, as well as the nesting model, are both adequate, neither set of plots indicating any substantive residual serial correlation. The final ratio of data to parameters is 8 (additional parameter reduction can be achieved by setting some small values in the L matrices to zero – these seems possible for 9 of the parameters), which is a bit less than $\sqrt{108}$. Signal extraction results, including the aggregate seasonal and the seasonal adjustment along with shaded uncertainty, are given in Figure 2.

We make several observations about the series and the results. The data are initially rising, but by the end of 2006 the decline has begun. Although the low frequency behavior of the series is of substantial interest, we draw attention to the rapidly evolving seasonal pattern. We enunciate here a stylized fact that is well-known to the seasonal adjustment community, but is yet to be absorbed by the broader enclave of economists – the seasonal pattern can change rapidly, making the antiquated use of seasonal regressors a dire mistake. The seasonal pattern, as subsequent analysis will show, is fairly stable when focus in retained on the pre-GR years, but the transition to mid- and post-GR behavior involves a gradual and yet substantial change to the seasonal pattern. The change is on both amplitude and yearly pattern. Standard checks – spectral plots of seasonal extractions and seasonal adjustments, as well as autocorrelation plots – indicate that the resulting adjustment is adequate. (Prior results, not shown here, that utilized the aggregate seasonal failed to capture the post-GR shift in seasonal pattern, essentially passing the pre-GR seasonal pattern forward to the GR years, resulting in residual seasonal swings – visible to the naked eye – in the seasonal adjustment in years 2010, 2011, and 2012.)

The series belong to the same species according to the third and sixth seasonal frequencies, and they also belong to the same trend species. The resulting trend extractions are (up to linear shifts) scalar multiples of one another; however, note that the trend extractions are not displayed in Figure 2, the focus being on the seasonal adjustment (trend plus irregular) instead. Another feature is the higher signal extraction uncertainty, as compared to the retail analysis, which may be due to the smaller sample size of nine years. These confidence intervals contain the true signal with probability approximately 95%, the width being equal to four standard errors. In some sense, the multivariate signal extraction has increased precision over univariate approaches, due to the information we can glean from other related series.

We can explore this facet by comparing the multivariate MSE to that obtained from the implied univariate models, as discussed in Section 2.1. The ratio of multivariate MSE to implied univariate model's MSE is displayed in Figures 1 and 2 of the Supplement, for the case of the trend estimate and seasonal estimate, respectively (the MSE for the seasonal and the seasonal adjustment are the same, since each signal is the other component's noise). We expect all these ratios to be less than unity, indicating some degree of efficiency gain from the multivariate model. This ratio is not uniform across time points, there being a notable boundary effect implicitly due to differences in forecasting performance. The case of series with little efficiency gain in the multivariate modeling can arise when the series are independent of one another, although lack of efficiency gain can occur for other reasons. Efficiency gain can be dramatic, although in practice are somewhat overstated – we are comparing to an implied univariate model, rather than refitting that univariate model to each individual series. We've chosen to display the results this way, because it isolates the impact of parameter choice (the parameters being the same for both specifications), and focuses purely on the impact of multivariate signal extraction. (We also studied precision comparisons with a cycle in the model, and with the aggregate seasonal, but the story is similar.)

Leaving our discussion of univariate extractions, we now further explore the quality of the multivariate extractions. A potential criticism of the results in Figure 2 is that the resulting new seasonal is much more dynamic and swiftly changing then many seasonal adjusters would be comfortable with. However, this directly follows from the model that we utilized together with the highly evolutive period of the GR. This model accommodates regime change in seasonal patterns, and thus offers a gradual and gentle alternative to modeling the GR with ramp regressors or other intervention effects, as some statisticians have pursued. There are some drawbacks to the use of intervention effects: their identification in real-time is not immediate, but requires at least a year to realistically utilize; once they are inserted, they are likely to produce a substantial effect (else why use them?) as compared to previous publications that omitted these regressors, resulting in large revisions. Another criticism is laid on aesthetic grounds: one must ask "when does it stop?" Ever more ramps and level shifts can be added, until no time series analysis remains to be done! We prefer the gradual changes produced by a filter based upon stochastic models, and recommend restricting the use of regressors to known effects such as reclassification or trading day or economic shocks. Although the GR is certainly a verifiable economic condition, identification of its start and end is subject to debate, and cannot be precisely defined in the way that trading day or reclassification level shifts can be.

We proceed to analyze seasonal adjustment sensitivity to model span. We look at several prior spans of the construction series, corresponding to 2000-2008, 2001-2009, 2002- 2010, and 2003-2011. We fitted the same overall model to each span, and produced nested models by co-integration restrictions in each case (these varied by span, so ultimately the model is not fixed). The resulting extractions are displayed in Figures 2, 3, 4, and 5 of the Supplement. Even the first span contains some of the GR period, but we notice that the seasonal extraction is fairly stable, and this carries through to the next span. Then in the latter two spans, the GR effect is now causing quite a bit of change to the seasonals. In each case, the seasonal adjustments and seasonal extractions have the requisite properties. For the span 2000-2008, the AIC is 1228.129 for the full model, and 1169.299 (54 parameters) for the restricted model; for the span 2001-2009, the AIC is 1198.619 for the nesting model, and 1144.619 (57 parameters) for the nested model. Next, for the span 2002-2010 the AIC is 1175.882 for the full model and 1133.439 (62 parameters) for the restriction. Finally, the span 2003-2011 has AIC 1101.773 for the nesting model versus 1055.265 (54 parameters) for the nested model.

Models and signal extraction results are fairly similar across these spans. Although the co-integrating orders for various components may differ, the overall number of parameters (between 54 and 62) is similar, indicating that the essential dimension for this sample size is close to the ratio eight, of data to parameters.

3.3 Regional Aggregation of Housing Starts

Here we focus on the topic of direct and indirect adjustment, and display results for the total housing starts using both definitions of indirect seasonal adjustment. Recall that the economic identity in play here dictates that the sum of the four regional series should equal the total, and we wish this to be true of the signal extraction estimates as well. The indirect approach computes the total's signal extraction as the appropriate linear combination (i.e., straight summation) of the individual series' signal extractions, so that economic accounting is automatic. The concern is whether this indirect extraction has the right statistical properties, i.e., is the indirect seasonal adjustment adequate? The direct approach, in contrast, models the totals with a univariate model and constructs signal extractions, which need not equal the sum of the individual series' extractions. We begin with fitting univariate models (these consist of $I(2)$ trend, irregular, and six seasonals, for nine parameters total) to each of the four series, and display the signal extraction results in Figure 6 of the Supplement. Comparing with Figure 2, there is quite a discrepancy between signal extraction results, although ultimately both seasonal adjustments – multivariate and univariate – are adequate, both having the ability to adapt their extraction filters to the GR regime change. We may then very well ask: what are the advantages of the multivariate method?

Note that the practical danger of utilizing a univariate indirect approach is that potentially some residual seasonal effects – indistinguishable according to an individual series' seasonal adjustment diagnostics – can (in aggregate) generate some actual seasonality. This danger, while not provably eliminated, is lessened in principle by the multivariate approach because each individual series' adjustment is obtained from a filter acting on *all* the regional series. The presence of correlation between two series at seasonal frequencies will be ignored by univariate seasonal adjustment methods, potentially resulting in residual seasonality in the aggregate series; the multivariate approach to seasonal adjustment has an opportunity to model and capture this correlation, and perfectly account for it in signal extraction.

One can simply add the univariate extractions to get an indirect extraction for totals, but there will be no quantification of uncertainty – one must either model the regional series jointly, or model the totals directly in order to get signal extraction uncertainty. This type of "univariate" indirect extraction yields results that are ultimately similar, in this case, to the multivariate extraction for totals shown in the left panel of Figure 3. (Uncertainty for the total is *not* obtained by summing the regional uncertainties, as discussed above.) Alternatively, one can model totals with a univariate model and produce a direct adjustment, with signal extraction error quantified, but there is no longer any guarantee that aggregation constraints are respected (they are not); this estimate is displayed in the right panel of Figure 3. Comparing both panels of this figure, we see that the multivariate indirect extractions and the univariate direct extractions are broadly similar, and both seasonal adjustments are indeed adequate, but the MSE for the former (left panel) is less than for the latter (right panel). To summarize, the multivariate indirect method seems to be superior to both the univariate indirect method and the univariate direct method: against the former methodology, the advantage is quantification of uncertainty; against the latter methodology, the advantage is respecting accounting rules. This discussion has omitted the possibility of raking or other *ad hoc* reconciliation measures, because these nonparametric techniques destroy all possibility of quantifying signal extraction uncertainty (Quenneville and Fortier, 2012).

We mention that some modelers may prefer to take a log transformation of the data, although this makes the indirect method of adjustment unworkable. Model fitting results are not reported here, but were similar or slightly simpler than the results given above for no transformation. The signal extraction estimates are extremely similar, although the uncertainty is quite a bit lower – this benefit can be weighed against the algebraic awkwardness associated with maintaing economic identities.

4 Conclusion

This paper attempts to address an important and long-standing question in seasonal adjustment and signal extraction, namely *is there a benefit to multivariate techniques*? Our proposals herein rely on available tools, such as multivariate structural tools (encoded in R). We have attempted to motivate these models as latent dynamic factor models that expand the basic dynamic factor model (1) in a manner that takes account of time series structure, associating additional latent dynamic factors with frequencies of interest in the process' spectral density. The factor loadings of each latent factor are then naturally associated with the lower Cholesky factors of the respective innovations' covariance matrix. Each entry of these lower Cholesky factors is interpretable as a scaled partial covariance, and therefore gives some information about how the respective time series are related to one another at trend or seasonal frequencies.

A pleasing facet of these models is their ready interpretability. The reduced rank in a latent process' innovation covariance matrix, corresponding to collinear innovations, is easily modeled, and moreover can be interpreted as frequencies of non-invertibility for the differenced process' spectral density. This in turn implies a co-integration intepretation for the undifferenced series, in a generalized sense; the basis of the co-integrating spaces are obtained at once by taking the appropriate rows of the inverse of the lower Cholesky factors. It has been shown that application of these co-integrating vectors reduces the order of non-stationarity of the original process, by exactly eliminating the need for the differencing operator corresponding to that particular latent process. We laboriously walk through this interpretation with a trivariate retail analysis.

Having identified the ranks of each latent component, we can then contemplate taxonomy of economic data, because co-integration is the same things as full spectral coherency among series at the respective frequency. We propose that fully reduced rank, of unity, be utilized as the definition of species, and establish some preliminary results for taxonomic classification. A key empirical facet is being able to determine the actual rank of each latent process' innovation covariance matrix, and we describe some tools involving partial correlations to tease out potential reduced rank models. We illustrate this procedure on the three retail series, showing how these partial correlation measures do indeed indicate which series are redundant for a particular latent component (the sixth atomic seasonal in our illustration). Once the full and restricted models have been fitted, an AIC comparison can be used to decide between competitors; other parameters (the entries of the lower Cholesky factors, or factor loading matrices) can also be zeroed out if warranted by a likelihood ratio test.

A key facet of this methodology is the ability to compute likelihoods and signal extraction results with relative ease. Our own implementation utilizes the innovations algorithm to evaluate the likelihood, which in our experience is faster than the Kalman filter; also, this method is more general, allowing treatment of processes not amenable to state space representations. Likewise, our signal extraction algorithms proceed from direct formulas – rather than state space smoothing algorithms – yielding the full error covariance matrix, which is needed to compute signal extraction mean squared errors for aggregates of the given variables.

One benefit of the multivariate signal extraction methodology is increased precision, as demonstrated through the precision comparisons on the housing starts data. Another benefit is the improvement of the indirect method of seasonal adjustment for preservation of economic identities. The direct method – running a univariate methodology on the totals – fails to preserve aggregation relations, while univariate indirect methods (summing the individual adjustments) will not take into account cross-series correlation, and will not allow for quantification of the aggregate series' signal extraction uncertainty. The multivariate indirect method addresses both of these latter issues, while preserving economic identities.

Given the benefits in terms of interpretability, taxonomic classification, and preservation of economic identities, what are the demerits of the multivariate methodology? We explored signal extraction revisions for construction series, demonstrating that the models are able to adapt to pre- and post-GR phenomena, indeed having the flexibility to accommodate rapidly changing seasonal patterns. This accommodation resulted in substantial revisions, which is no surprise given that the new information radically altered prior understandings of trend and seasonal patterns. Overall, the revisions behavior seemed satisfactory, although we noted that the actual models identified (the particular co-integrating ranks) can change dramatically as the data span is altered.

In our own opinion, the chief criticism is in the additional time required of the analyst to perform the modeling task; secondly, and related, is the huge number of parameters involved when $m = 4$ or higher. Thirdly, the use of log transformations interferes with our proposed method of handling the preservation of economic identities. Regarding the first two points, to achieve parsimony and a feasible computation time, one is naturally led to seeking co-integrating relationships and other reductions of the parameter space, and these modeling efforts take a substantial amount of investigative time. Future research must focus on regularization techniques, or other methods to enforce sparsity on the parameter space.

References

[1] Bartelsman, E. and Cleveland, W. (1993) Joint seasonal adjustment of economic time series. Finance and Eocnomics Discussion Series, Federal Reserve Board, Washington, D.C.

[2] Bell, W. and Hillmer, S. (1984) Issues involved with the seasonal adjustment of economic time series. *Journal of Business and Economics Statistics* **2**, 291–320.

[3] Bozik, J. and Bell, W. (1987) "Forecasting age specific fertility using principal components." Bureau of the Census Research Report RR-87/19.

[4] Brockwell, P. and Davis, R. (1991) *Time Series: Theory and Methods.* New York: Springer.

[5] Buys Ballot, C. (1847) *Les Changements Pérodiques de Tempèrature*, Utrecht: Kemink et Fils.

[6] Durbin, J. and Koopman, S. (2001) *Time Series Analysis by State Space Methods.* Oxford University Press.

[7] Engle, R. and Granger, C. (1987) Cointegration and error correction: representation, estimation, and testing. *Econometrica* **55**, 251–276.

[8] Giannone, D., Reichlin, L., and Sala, L. (2004) Monetary policy in real time. *NBER Macroeconomics Annual*, 161-200.

[9] Gilbart, J. (1852) On the laws of the currency of Ireland, as exemplified in the changes that have taken place in the amount of bank notes in circulation in Ireland, since the passing of the acts of 1845, *Journal of the Statistical Society of London* **15**, 307–325.

[10] Granger, C. (1966) The typical spectral shape of an economic variable. *Econometrica* **34**, 150–161.

[11] Koopman, S., Ooms, M., and Hindrayanto, I. (2012) A multivariate periodic unobserved components time series analysis for sectoral U.S. employment. In W. Bell, S. Holan, and T. McElroy (Eds.), *Economic Time Series: Modeling and Seasonality.* New York: Chapman and Hall.

[12] Krane, S. and Wascher, W. (1999) The cyclical sensitivity of seasonality in U.S. unemployment. *Journal of Monetary Economics* **44**, 523–553.

[13] McElroy, T. (2014) Recursive computation for block nested covariance matrices. *Mimeo*

[14] McElroy, T. and Trimbur, T. (2014) Signal extraction for nonstationary multivariate time series with illustrations for trend inflation. Forthcoming, *Journal of Time Series Analysis.* Also: Finance and Economics Discussion Series, Federal Reserve Board. 2012-45

[15] Nerlove, M. (1964) Spectral analysis of seasonal adjustment procedures. *Econometrica* **32**, 241–286.

[16] Nerlove, M., Grether, D., and Carvalho, J. (1979) *Analysis of Economic Time Series: A Synthesis*, New York: Academic Press.

[17] Nyblom, J. and Harvey, A. (2000) Tests of common stochastic trends. *Econometric Theory* **16**, 176–199.

[18] Nyblom, J. and Harvey, A. (2001) Testing against smooth stochastic trends. *Journal of Applied Econometrics* **16**, 415–429.

[19] Persons, W. (1919) Indices of business conditions, *Review of Economics and Statistics* **1**, 5–107.

[20] Pinheiro, J. and Bates, D. (1996) Unconstrained parametrizations for variance-covariance matrices. *Statistics and Computing* **6**, 289–296.

[21] Quenneville, B. and Fortier, S. (2012) Restoring accounting constraints in time series – methods and software for a statistical agency. In W. Bell, S. Holan, and T. McElroy (Eds.), *Economic Time Series: Modeling and Seasonality.* New York: Chapman and Hall.

[22] Sargent, T. and Sims, C. (1977) "Business cycle modeling without pretending to have too much a-priori economic theory," in *New Methods in Business Cycle Research*, ed. by C. Sims et al., Minneapolis: Federal Reserve Bank of Minneapolis.

[23] Stock, J. and Watson, M. (1988) Testing for common trends. *Journal of the American Statistical Association* **83**, 1097–1107.

[24] Stock, J. and Watson, M (2011) "Dynamic Factor Models," Ch. 2 in M.J. Clements and D.F. Hendry (eds.), *Oxford Handbook on Economic Forecasting.* Oxford: Oxford University Press.

[25] Stock, J. and Watson, M. (2012) Disentangling the channels of the 2007-2009 recession. Prepared for the Brookings Panel on Economic Activity, March 22-23, 2012.

[26] Tiller, R. (2012) Frequency domain analysis of seasonal adjustment filters applied to periodic labor force survey series. In W. Bell, S. Holan, and T. McElroy (Eds.), *Economic Time Series: Modeling and Seasonality.* New York: Chapman and Hall.

[27] Wright, J. (2013) "Unseasonal seasonals?" Fall 2013 Brookings Panel on Economic Activity, Septermber 19-20, 2013.

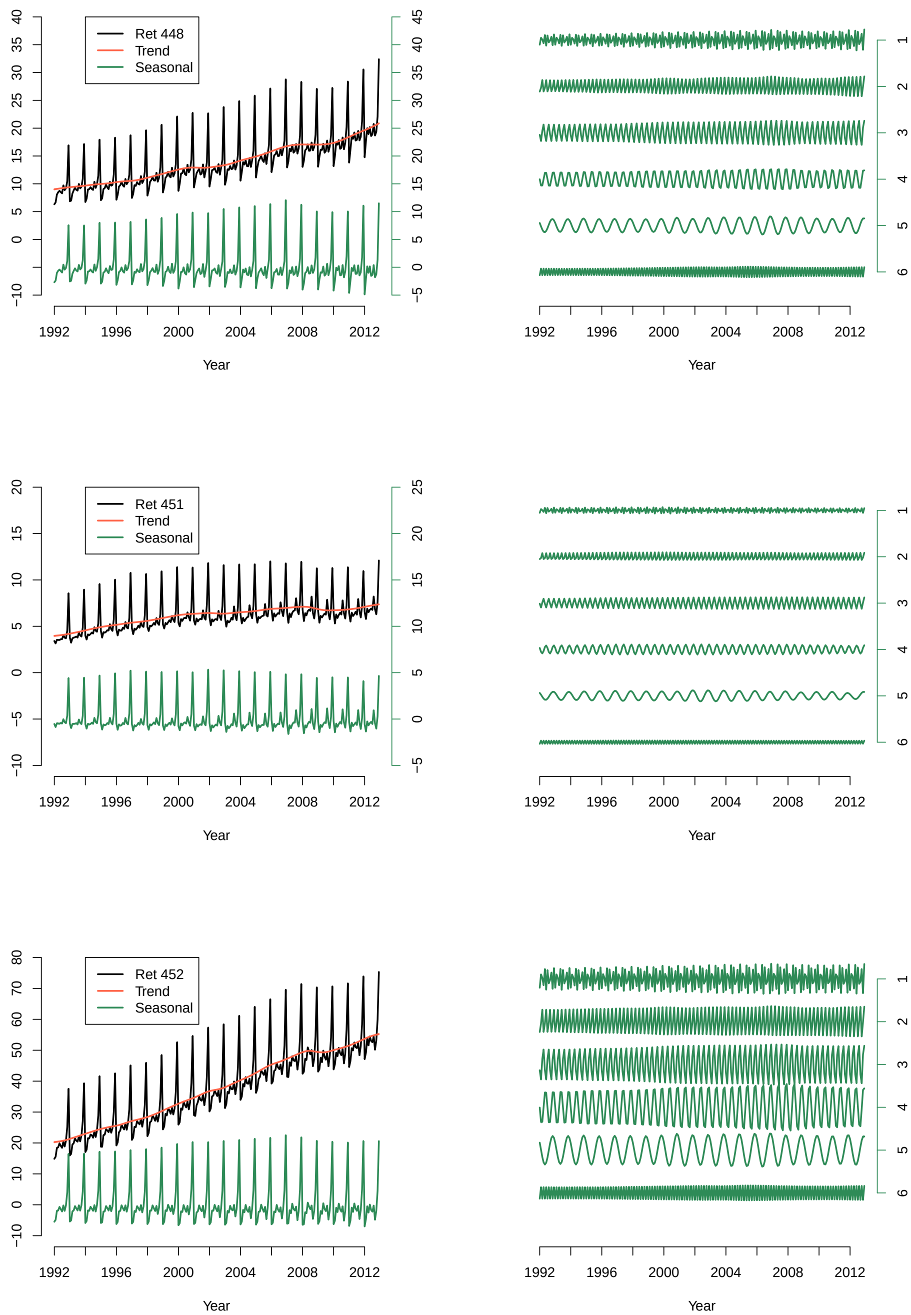

Figure 1: Left panel: Retail series (black) for 1992-2012 period, for 448, 451, and 452, with trend (red) and seasonal (green) estimates based on the best fitted trivariate model. Right panel: seasonal extractions for first through sixth seasonal components. Shaded bands corresponds to confidence intervals of width given by two standard errors.

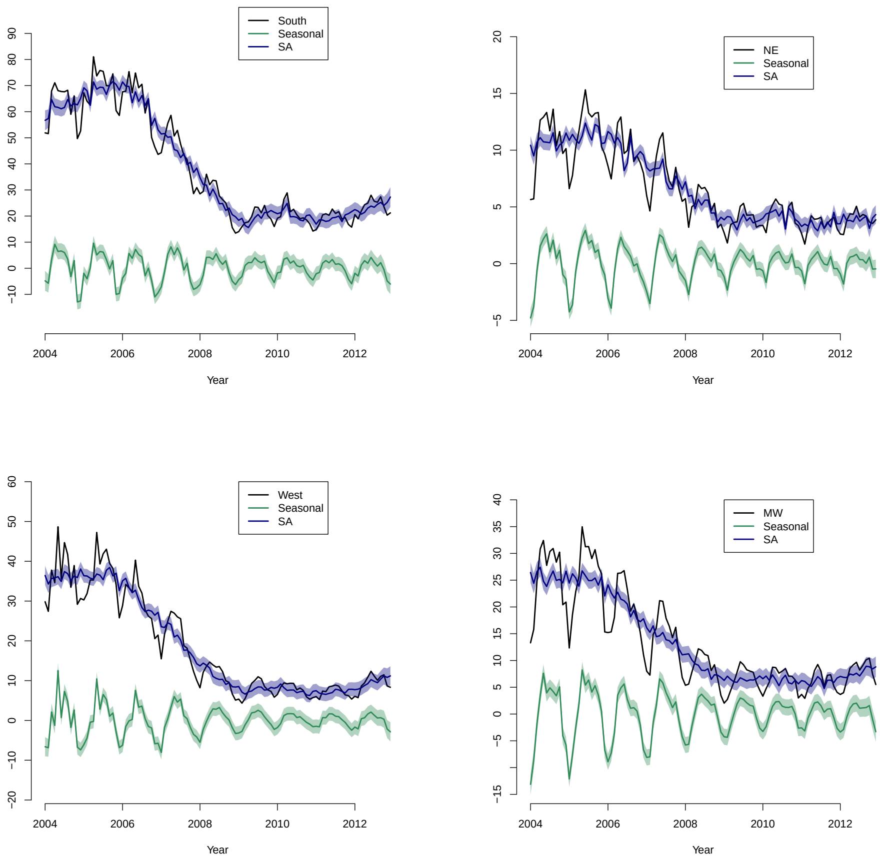

Figure 2: Housing Starts (black) for 2004-2012 period, for four regions of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates based on the best fitted multivariate model. Shaded bands corresponds to confidence intervals of width given by two standard errors.

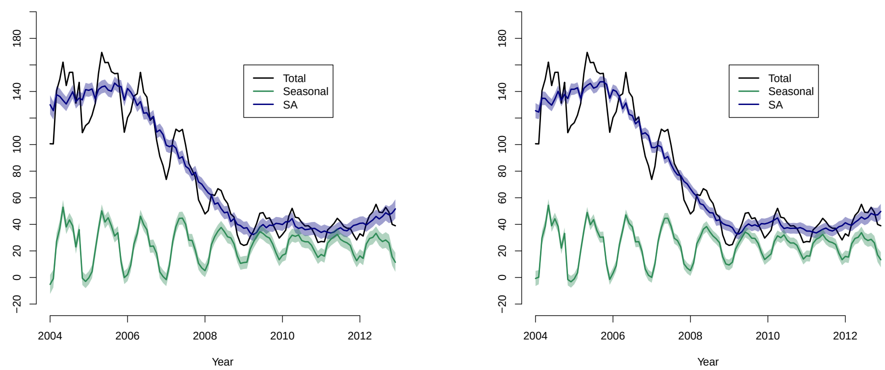

Figure 3: Housing Starts (black) for 2004-2012 period, for total starts of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates. Shaded bands corresponds to confidence intervals of width given by two standard errors. The indirect adjustment (left panel) is constructed from the multivariate model, whereas the direct adjustment (right panel) is constructed from the univariate model fitted to totals.

Supplement to “Multivariate Seasonal Adjustment, Economic Identities, and Seasonal Taxonomy”

Tucker McElroy*

Appendix

Proof of Proposition 1. Consider three time series a, b, c that are difference stationary. Then the spectral density of a with a evaluated at frequency $2\pi j/s$ is a covariance matrix with correlation ± 1, and hence has rank at most one. Hence $a \sim_j a$. Likewise, given that $a \sim_j b$, then $b \sim_j a$ because the correlation is ± 1 in the covariance matrix, in either case. Finally, suppose that $a \sim_j b$ and $b \sim_j c$; we wish to show that $a \sim_j c$. Consider the covariance matrix $f(2\pi j/s)$ where f is the spectral density for all three series (appropriately differenced). From the Cholesky decomposition discussed in Section 2, we have

$$\Sigma_3 = \left[\begin{array}{lll} d_1 & L_{21}\, d_1 & L_{31}\, d_1 \\ L_{21}\, d_1 & L_{21}^2\, d_1 + d_2 & L_{21}\, L_{31}\, d_1 + L_{32}\, d_2 \\ L_{31}\, d_1 & L_{21}\, L_{31}\, d_1 + L_{32}\, d_2 & L_{31}^2\, d_1 + L_{32}^2\, d_2 + d_3 \end{array} \right].$$

Let us suppose that the series are ordered (a, b, c) in this matrix – similar arguments apply to other orderings. Then the covariance matrix for (a, b) is given by

$$\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right] \Sigma_3 \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{array} \right] = \left[\begin{array}{ll} d_1 & L_{21}\, d_1 \\ L_{21}\, d_1 & L_{21}^2\, d_1 + d_2 \end{array} \right].$$

By assumption this has rank at most one, so that either $d_1 = 0$ or $d_2 = 0$, or both. Similarly, the covariance matrix for (b, c) is

$$\left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right] \Sigma_3 \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{array} \right] = \left[\begin{array}{ll} L_{21}^2\, d_1 + d_2 & L_{21}\, L_{31}\, d_1 + L_{32}\, d_2 \\ L_{21}\, L_{31}\, d_1 + L_{32}\, d_2 & L_{31}^2\, d_1 + L_{32}^2\, d_2 + d_3 \end{array} \right].$$

*Center for Statistical Research and Methodology, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, tucker.s.mcelroy@census.gov

Because this matrix has rank at most one by assumption, if $d_1 = 0$ we find that we must have $d_2 = 0$ or $d_3 = 0$ (or both). Likewise, if $d_2 = 0$, then we must have $d_1 = 0$ or $d_3 = 0$ or both. Hence, at most one of d_1, d_2, d_3 is nonzero, which shows that Σ_3 has rank at most one. It also follows that the covariance matrix for (a, c) has rank at most one; this yields the transitive property, and hence $\sim_j$ is an equivalence relation. $\square$

Derivation of Partial Correlations. We first compute another conditional variance. Denote the ijth entry of the m-dimensional covariance matrix by $\Sigma_{i,j}(m)$. Then for $i > j$ we have

$$\mathrm{Var}_{\epsilon_1,\cdots,\epsilon_{j-1}}(\epsilon_i) = \Sigma_{i,i}(m) - \Sigma_{i,1:j-1}(m)\, \Sigma(j-1)^{-1}\, \Sigma_{1:j-1,i}(m) = \sum_{p=j}^{i} L_{ip}^2\, d_p.$$

As a result,

$$\kappa_{ij} = \frac{\mathrm{sign}(L_{ij})}{\sqrt{1 + [d_i + \sum_{p=j+1}^{i-1} L_{ip}^2\, d_p]/(L_{ij}^2\, d_j)}}.$$

It is also possible to express the entries of $\Sigma(m)$ in terms of the various κ_{ij}, from which the stated formula for the determinant follows at once. The condition number is defined in terms of the τ_{ij}, which can be written in terms of the LDL' parametrization via

$$\tau_{ij} = \log(\sum_{p=j+1}^{i} L_{ip}^2 d_p) - \log(\sum_{p=j}^{i} L_{ip}^2 d_p).$$

Derivation of Trend Centering. The fixed mean function is $\mathbf{m}_t = \mathbf{m}_t^{\mu} + \mathbf{m}_t^{\xi}$ for trend and seasonal effects, such that $U(B)\mathbf{m}_t^{\xi} = 0$ and $\mathbf{m}_t^{\mu} = \sum_{j=0}^{d} \nu_j t^j$. This places any nonzero mean of the seasonal effect with the trend polynomial. The mean $\mathbf{m}$ of ∂y_t is

$$\mathbf{m} = \mathbb{E}[\partial y_t] = (1-B)^d U(B)[\mathbf{m}_t^{\mu} + \mathbf{m}_t^{\xi}] = d!\nu_d U(1).$$

Therefore, $\nu_d = \mathbf{m}/(d!U(1))$. The action of signal extraction matrices for ξ, ψ, and ι all involve trend differencing, while the action of signal extraction matrices for μ, ψ, and ι all involve seasonal differencing. Write

$$\mathbf{m}^{\mu} = K \begin{bmatrix} \mathbf{m}_1^{\mu} \\ \mathbf{m}_2^{\mu} \\ \vdots \\ \mathbf{m}_n^{\mu} \end{bmatrix}, \qquad \mathbf{m}^{\xi} = K \begin{bmatrix} \mathbf{m}_1^{\xi} \\ \mathbf{m}_2^{\xi} \\ \vdots \\ \mathbf{m}_n^{\xi} \end{bmatrix},$$

with K the transposition matrix. Then the various signal extraction matrices satisfy

$$F^{\mu}\mathbf{m}^{\xi} = 0, \qquad F^{\xi}\mathbf{m}^{\xi} = \mathbf{m}^{\xi}, \qquad F^{\psi}\mathbf{m}^{\xi} = 0, \qquad F^{\iota}\mathbf{m}^{\xi} = 0.$$

However, the action of the trend-differencing reduces $\mathbf{m}^\mu$ to $d!\nu_d \otimes (1,1,\cdots,1)'$, which we denote by β. Thus

$$\begin{aligned} F^\mu \mathbf{m}^\mu &= \mathbf{m}^\mu + \left(\overline{F^\xi} + \overline{F^\psi} + \overline{F^\iota} \right) \beta \\ F^\xi \mathbf{m}^\mu &= \overline{F^\xi} \beta \\ F^\psi \mathbf{m}^\mu &= \overline{F^\psi} \beta \\ F^\iota \mathbf{m}^\mu &= \overline{F^\iota} \beta, \end{aligned}$$

where in each case $\overline{F}$ is defined via $F = \overline{F}[1_n \otimes \Delta_\mu]$. The trend estimate, which follows from conditional expectation calculations, is

$$\begin{aligned} & \mathbf{m}^\mu + F^\mu [\text{vec} Y' - (\mathbf{m}^\mu + \mathbf{m}^\xi)] \\ & = F^\mu \text{vec} Y' - \left(\overline{F^\xi} + \overline{F^\psi} + \overline{F^\iota} \right) \beta \\ & = \widetilde{\mathbf{m}}^\mu + F^\mu [\text{vec} Y' - \widetilde{\mathbf{m}}^\mu], \end{aligned}$$

where $\widetilde{\mathbf{m}}^\mu$ is obtained from $\mathbf{m}^\mu$ by setting each ν_j equal to zero except for ν_d (so that only the leading coefficient matters). Similarly, the seasonal, cycle, and irregular estimates are

$$\begin{aligned} \mathbf{m}^\xi + F^\xi [\text{vec} Y' - (\mathbf{m}^\mu + \mathbf{m}^\xi)] &= F^\xi [\text{vec} Y' - \widetilde{\mathbf{m}}^\mu] \\ F^\psi [\text{vec} Y' - (\mathbf{m}^\mu + \mathbf{m}^\xi)] &= F^\psi [\text{vec} Y' - \widetilde{\mathbf{m}}^\mu] \\ F^\iota [\text{vec} Y' - (\mathbf{m}^\mu + \mathbf{m}^\xi)] &= F^\iota [\text{vec} Y' - \widetilde{\mathbf{m}}^\mu] \end{aligned}$$

respectively. So in each case, we can subtract $\widetilde{\mathbf{m}}^\mu = \mathbf{m} \otimes \tau$ from $\text{vec} Y'$, apply the filter, and then in the case of a trend estimate, add $\widetilde{\mathbf{m}}^\mu$ back.

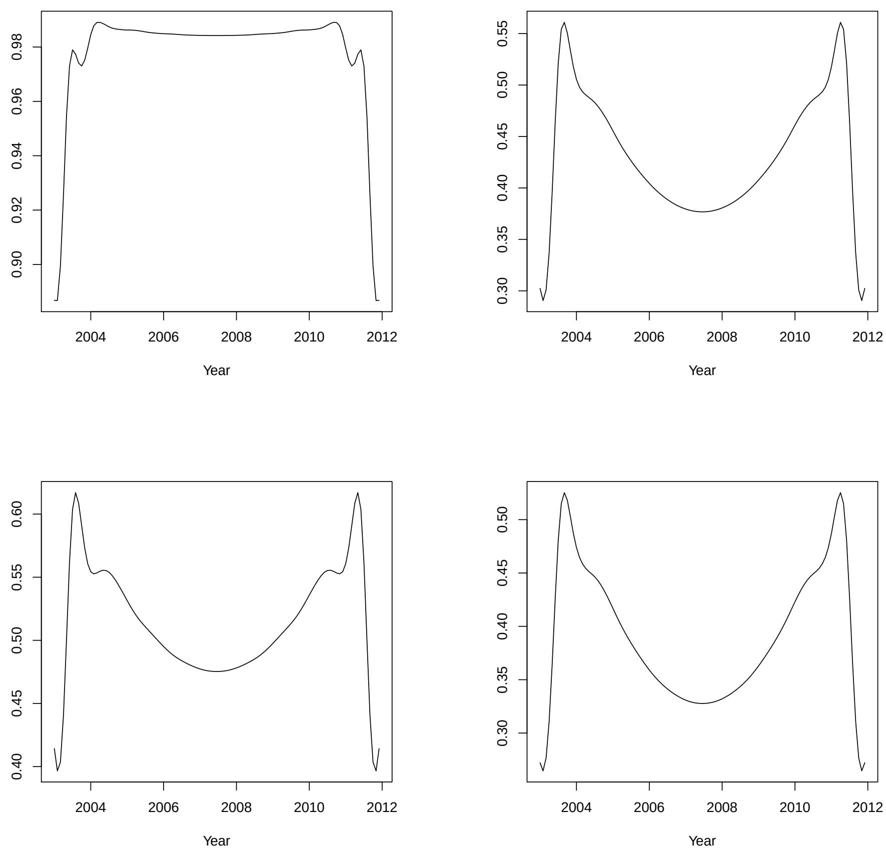

Figure 1: Trend extraction MSE ratios for Housing Starts data (2004-2012) for four regions of the U.S. (in thousands of housing units), comparing MSE of the multivariate model to MSE of the implied univariate model. These correspond to South (upper left), NorthEast (upper right), West (lower left), and MidWest (lower right) respectively.

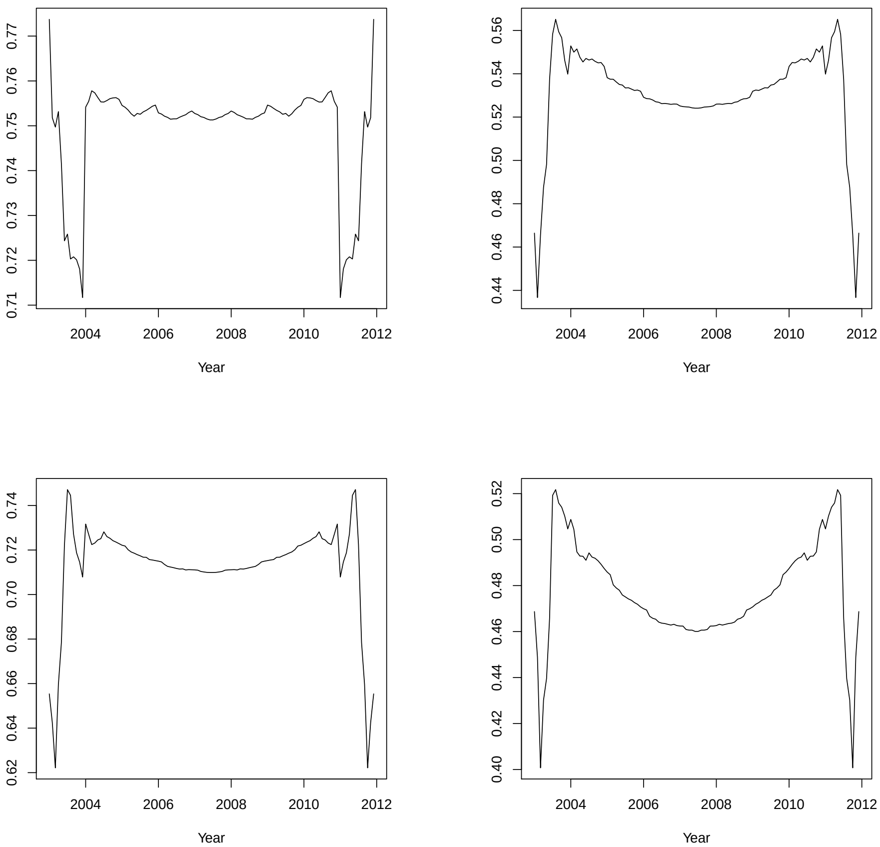

Figure 2: Seasonal extraction MSE ratios for Housing Starts data (2004-2012) for four regions of the U.S. (in thousands of housing units), comparing MSE of the multivariate model to MSE of the implied univariate model. These correspond to South (upper left), NorthEast (upper right), West (lower left), and MidWest (lower right) respectively.

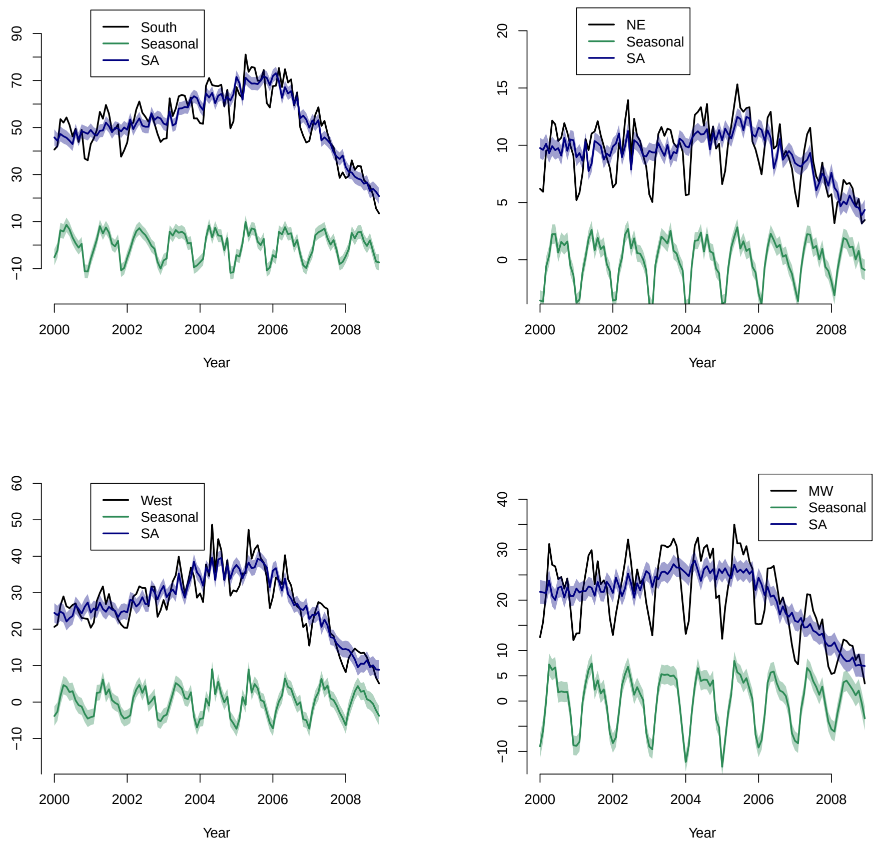

Figure 3: Housing Starts (black) for 2000-2008 period, for four regions of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates based on the best fitted multivariate model. Shaded bands corresponds to confidence intervals of width given by two standard errors.

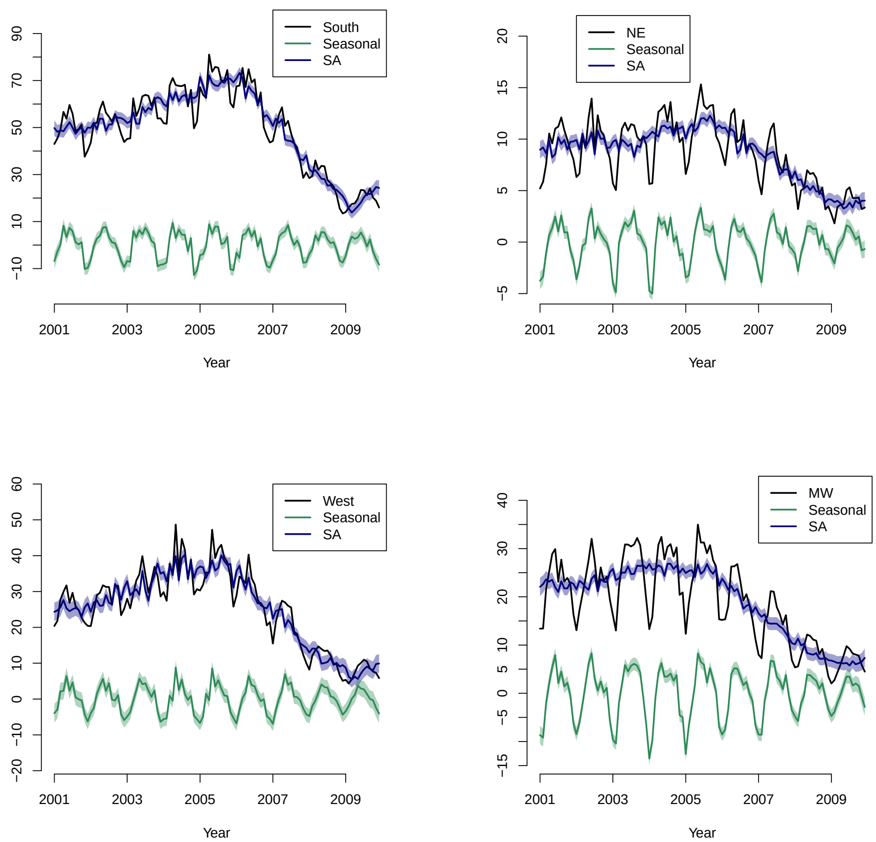

Figure 4: Housing Starts (black) for 2001-2009 period, for four regions of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates based on the best fitted multivariate model. Shaded bands corresponds to confidence intervals of width given by two standard errors.

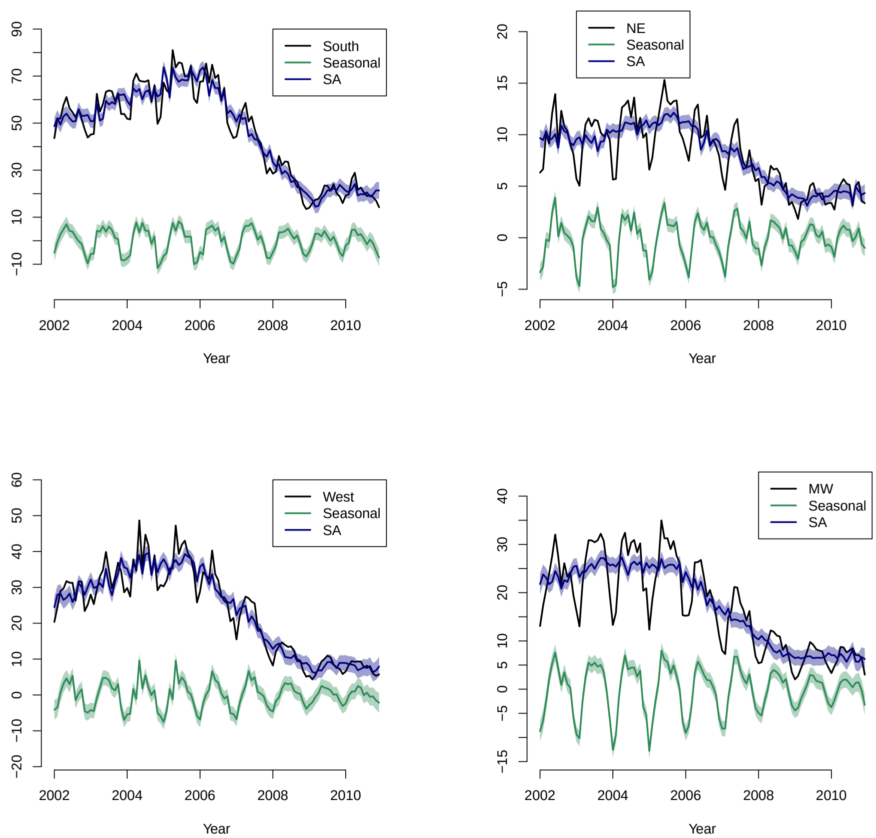

Figure 5: Housing Starts (black) for 2002-2010 period, for four regions of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates based on the best fitted multivariate model. Shaded bands corresponds to confidence intervals of width given by two standard errors.

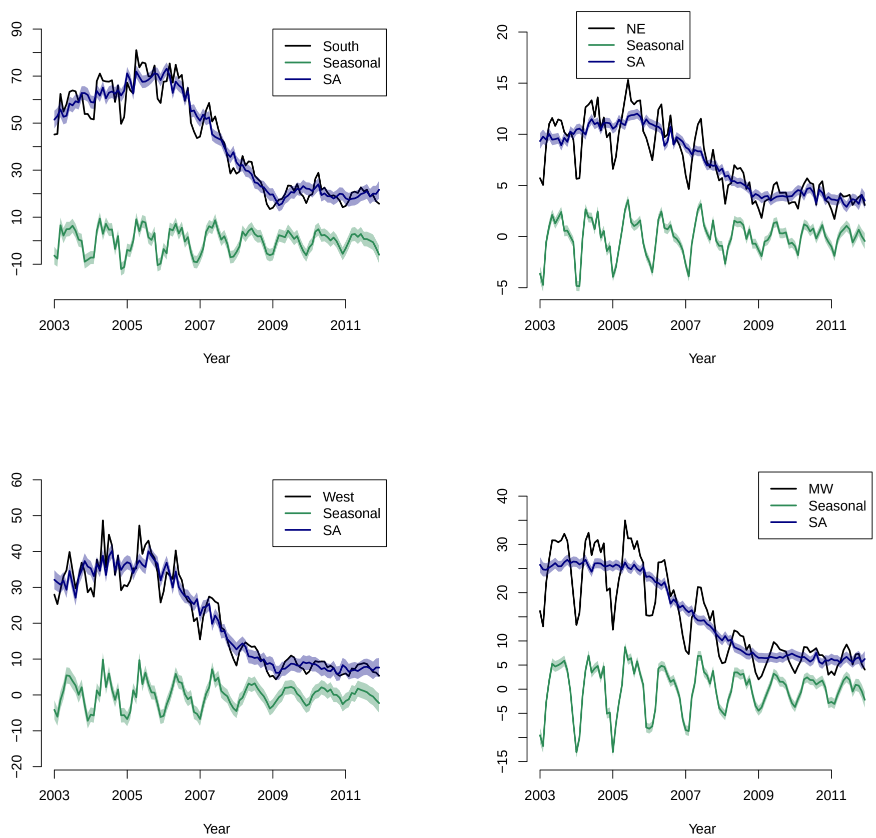

Figure 6: Housing Starts (black) for 2003-2011 period, for four regions of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates based on the best fitted multivariate model. Shaded bands corresponds to confidence intervals of width given by two standard errors.

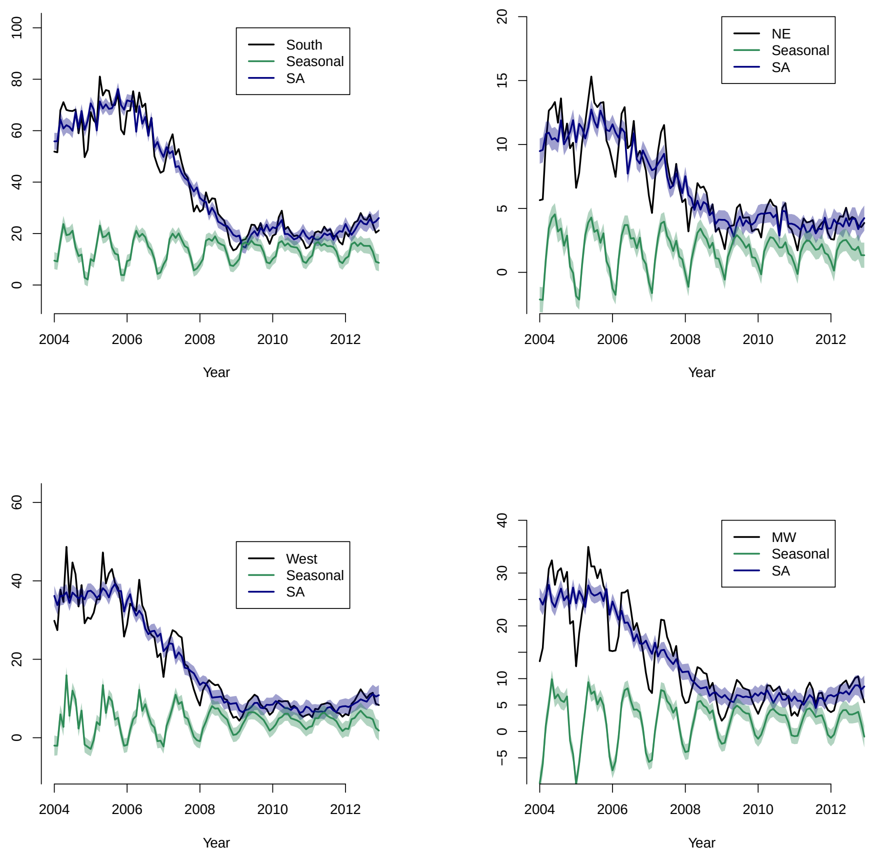

Figure 7: Housing Starts (black) for 2004-2012 period, for four regions of the U.S. (in thousands of housing units), with seasonal adjustment (blue) and seasonal (green) estimates based on the best fitted univariate models. Shaded bands corresponds to confidence intervals of width given by two standard errors.