NONSTATIONARY FACTOR MODEL APPLICATIONS of ELASTIC NET

A Master's Thesis

by
DENİZ KONAK

Department of
Economics
İhsan Doğramacı Bilkent University
Ankara
January 2013

NONSTATIONARY FACTOR MODEL APPLICATIONS of ELASTIC NET

Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

DENİZ KONAK

In Partial Fulfillment of the Requirements For the Degree
of
MASTER OF ARTS

in

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

JANUARY 2013

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Taner Yiğit
Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Fatma Taşkın
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Süheyla Özyıldırım
Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

---

Prof. Dr. Erdal Erel
Director

ABSTRACT

NONSTATIONARY FACTOR MODEL APPLICATIONS of ELASTIC NET

KONAK, Deniz

M.A., Department of Economics

Supervisor: Associate Professor Taner Yiğit

January 2013

In this thesis, we adopted Elastic Net estimators for selecting true number of factors in factor models with stationary and nonstationary factors. Elastic Net is a member of shrinkage estimators family. As a member of shrinkage estimators family, elastic net estimators are stable to changes in data and in general they do not over parametrize the models. These two properties of elastic net estimators makes elastic net more favourable than information based criterion penalty methods for estimating true factor number. Since Principal Components Analysis (PCA) based algorithms always tends to give only single factor for nonstationary data sets, we use Sparse Principal Components Analysis (SPCA) algorithm which is a regression-type optimization formulation of PCA. Simulations show the performance of Elastic Net estimator for estimation of true factor number with stationary and nonstationary factors cases .

Keywords: Elastic Net, Sparse Principal Component Analysis, Nonstationary Factor Models, True Factor Number Estimation.

ÖZET

DURAĞAN OLMAYAN FAKTÖR MODELLERDE ELASTIC NET UYGULAMALARI

KONAK, Deniz

Yüksek Lisans, Ekonomi Bölümü

Tez Yöneticisi: Doçent Taner Yiğit

Ocak 2013

Bu tez çalışmamızda, durağan olan ve durağan olmayan faktörlerin kullanıldığı faktör modeller için elastic net tahmin edicilerini kullandık. Elastic net shrinkage tahmin edicileri ailesinin bir üyesidir. Bu ailenin bir üyesi olarak elastic net tahmin edicileri, veri setlerindeki değişmelerden etkilenmezler ve genel olarak, gerçek modelde olandan fazla parametre tahmin etmezler. Belirtilen bu iki özellik faktör sayısı tahminleri icin elastic net tahmin edicilerini bilgi temelli ceza yöntemlerine göre daha tercih edilir yapmaktadır. Temel bileşenler analizi tabanlı algoritmalar durağan olmayan serilerde sadece tek faktör tahmin etmeye eğilimli oldukları için temel bileşenler analizinin regresyon tabanlı optimizasyon algoritması olan seyrek temel bileşenler analizi yöntemini kullandık. Yapılan simülasyonlar elastic net tahmin edicilerinin durağan olan ve durağan olmayan faktörlerin tahmin edilmesindeki performanslarını göstermektedir.

Anahtar Kelimeler: Elastic Net, Seyrek Temel Bileşenler Analizi, Durağan Olmayan Faktör Modeller, Doğru Faktör Sayısı Tahmini.

ACKNOWLEDGMENTS

I would like to express my deepest gratitudes to;

My father, Rıza Konak, my mother, Şerife Konak, and my brother, Alper Can Konak, for their permanent encouragements and supports as well as just being a great family for me.

My grandfather, Ahmet Konak, not only because of his endless trust, but also because of being a great role model for me.

Assoc. Prof. Taner Yiğit, for his invaluable guidance and support through out my study.

Burak Alpaslan Eroğlu for his valuable suggestions about this study.

To all my friends, who always trusted on me and showed their unlimited respect in my challenging days. Their friendship and continuous support was extremely important for my motivation.

Finally, TUBITAK, for their financial support during my study.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Panel datasets with large number of cross section and time dimensions are becoming one of the main concerns of growing econometrics literature since they are being increasingly available and mostly used in economics for many issues such as estimation, prediction and forecasting. But as the number of panel data sets are increasing the need of information extraction and dimension reduction for these data sets are increasing as well.

Factor models are one of the most common ways which provides an effective way of extracting important information from these high dimensional data sets where factors are regarded as unobservable variables that can explain co-movements of many variables. The use of factor models for deriving unobserved components is first introduced to literature by Ross (1976) for the use of arbitrage pricing theories. Then, Chamberlain and Rothschild (1983) paper developed a new factor model structure called approximate factor model which requires weaker assumptions than Ross's factor model hence Chamberlain and Rothschild (1983) approximate factor model has been more widely applicable than Ross (1976) factor model.

The use of factor models in economics in different areas such as asset pricing, forecasting, stock markets, panel co-integration and cross-sectional correlations is widely becoming popular tool. There are many studies such as Forni and Reichlin (1998), Stock and Watson (1989), Stock and Watson (1999) and Bernanke and Bovin (2000) arguing the use of factor models for defined areas. Why using factor models are important for economics is detailed in literature review section.

For all of the existing factor model literature, more than just using the factor models for various areas, the central aim of both theoretical and empirical studies have became the determination of correct factor number and estimation of these factors. Because it is a fact that explaining and extracting the unobserved component consistently and truly is the most important part of used algorithms. In other words, if we can estimate the true number of factors consistently and extract the most proper information from our data and the further analysis will be based on the most proper approximation to our data. Hence, the results of our analysis will be the true representation of main data.

Since "How many true factors are in data set?" is the main question and the focus of econometrics researchers, there are many studies developing several methodologies to estimate the true number of factors in the literature. Based on similar main algorithms in their dimension reduction such as principal component analysis (PCA) and sparse principal component analysis (SPCA), these methodologies differ in their analysis algorithms. Some of them, such as Bai and Ng (2002) and Bai(2004) are based on penalizing information criteria which is used as a kind of general factor number threshold. Some other, such as Caner (2011) is based on penalizing the loadings of the estimated factors as an alternative to information criterion penalization. On the other hand, different from penalization methods, Onatski (2009) consistently separates diverging eigenvalues from clustered ones and counts the number of separated ones as the number of estimated factor number.

As it can be understood from existing studies, all of these methodologies have several weaknesses in several issues. Some of them work only for stationary factors and fail when factors are non stationary, some other such as Caner (2009) has good small sample properties but may not be adequate for large datasets and the information criterion based ones are not robust to data changes and over parametrization can be mostly seen problem. Hence, it is hard to say one of those has absolute advantage over other.

Among all of these problems, fail in non stationary factors is the main concern of our study. There are many reasons why nonstationarity is our concern but the main reason is that most of the available panel data is nonstationary and under this nonstationarity the estimation of true factor number always becomes most problematic issue. Hence in this study, we aim to work on the nature and conditions of an alternative algorithm which can estimate true factor number of a panel data set under nonstationary series.

In the next section, we introduce the general literature of factor number estimation and main literature of "Elastic Net" which is the main aim of our work. Section 3 gives all details about our proposed study and methodologies behind our study. Section 4 explains the details and the results of conducted simulation studies. Section 5 concludes the study with some closing remarks.

CHAPTER 2

LITERATURE REVIEW

The use of factor models in economics almost starts with Ross (1976) paper on arbitrage theory. In Ross (1976), the main aim is to set an alternative arbitrage model to existing mean variance capital asset pricing model of Sharpe, Lintner and Treynor. This study used a factor model structure for the offered alternative arbitrage model and opens a new window for the use of factors in economics.

Having some strict constraints as well as being very contributory burns new approaches to factor model. Chamberlain and Rothshild (1983) is a paper which relaxes the strict usage assumptions on implications of arbitrage theories in multi good markets. They call this new model as an approximate factor structure model which is based on principal component analysis and convergence of eigenvalues and factor loadings.

Since factor models are very useful in many areas of economics and finance they become very important in literature. Since factor models allow us to work with high dimensional data sets and allow us interpret these data set properties, the studies which require high dimensional data set analysis mostly benefit from these properties of factor models.

Stock and Watson (1989), uses factor models to extract a single, but has the power of explaining whole data set, factor among multiple macroeconomic variables. And by using extracted single factor, they aim to state any existing co-movements among macroeconomic variable indicators.

On the other hand, Stock and Watson (1999), adopts factor models to inflation forecasting literature. They use factor model to forecast 12-month horizon inflation of U.S.A. They show that extracting factors and using the extracted factors to forecast inflation does better job than forecasting inflation with all economic indices.

Moreover, Forni and Reichlin (1998) is another study which shows the use of factor models on business cycle studies for differentiating the heterogeneity between alternative shocks and cross section units.

Bernanke and Bovin (2000), tests the monetary policies of FED by using factor models. They adopt factor models to be able to use very high numbers of economic variables. They support that, large dimensional factor model allow them to work with many variables and so to test the effects of monetary policy they can consider unlimited number of indicators of economy.

The series of significant contributions to factor number determination in econometrics literature almost start with the very contributory paper of Bai and Ng(2002). In Bai and Ng (2002), by using Chamberlain's approximate factor model structure, they propose a panel criteria to estimate number of factors consistently in stationary panels. Based on some assumptions, they show that the information criteria they offered has good performance for large cross section and time dimensional finite panel data. The most important drawback of this criteria is it works only for stationary panels as well as its not being robust to data changes.

Bai (2004) examines the use of large dimension factor models with cross section common stochastic trends which are also referred as nonstationary dynamic factors. They consider the problem of determining the true number of dimension of factors and deriving the distribution of factors and factor loading. They also study generalized dynamic factor models with cross section common stochastic trends and they apply this method to the study of employment fluctuations across 60 U.S industries. They found that some number of nonstationary dynamic factors can explain much of fluctuations in the sectoral employment which is the strong support of their claim.

In a complementary way, Onatski (2010) develops a new, consistent estimator of the number of factors in the approximate factor models and this estimator is even works good when idiosyncratic terms are cross sectional correlated. This paper does not set any restriction on the speed of growth of factors cumulative effect as a difference from the literature of factor models. Moreover, this methodology has an advantage in being able to work in large samples as well as small samples.

Bai, Kao and Ng (2009) studies estimation of panel co-integrated model with cross sectional dependence generated by unobserved global stochastic trend. Since the standard least squares estimators are not giving the consistent results because of unobserved I(1) trends, they are proposing two iterative procedures that are jointly estimate the stochastic trends and slope parameter. They concluded that these both procedures are giving consistent, asymptotically unbiased and asymptotically normal estimates. And furthermore, they showed that these estimators are also working for mixed of I(0)/I(1) factors.

On the other hand, Shrinkage family estimators firstly introduced to literature by Frank and Friedman (1993) with bridge estimators. Bridge estimators are the least squares estimators with penalized objective function of classical regression equations where usual assumptions of classical regression model holds. Then, Fan and Li (2001) and Tibshirani (1996) developed ridge and lasso estimators for the shrinkage family estimators with alternative form of penalties. Since shrinkage estimators are both dimension reduction and model selection at one hand, they become favourable for factor model studies.

Different from information based criterion penalty which is used in Bai(2002), Bai(2004) and Bai, Kao and Ng (2009), Caner(2011) proposes the use of bridge estimators to determine correct number of factors consistently.Since it also allows cross sectional dependencies, approximate factor models are used similar to Bai (2004). This paper uses a penalty which is based on factor loadings as an alternative methodology to Bai(2002) for the same data generation process. The results of Caner(2011) indicate that bridge estimators show better performances for determination of correct factor numbers especially for low dimensional panels.By using bridge estimator, robustness of factor number estimators for data changes and problem of over parametrization is solved and this is supported as the contribution of that paper into econometrics literature.

PCA by Jolliffe (1986) is the most commonly used dimension reduction and data synthesizing algorithm used as a based algorithm in many factor models in literature. Aiming to explain most of variability among original data variables, PCA detects the orthogonal linear combinations of high dimensional data. But as its major disadvantage, PCA results are highly affected from the dependencies among cross sections and time dimensions units of data sets. Hence under cross sectional dependency and nonstationarity, usefulness of PCA is limited due to destroyed form of variance-covariance matrix.

To deal with drawbacks of PCA algorithm Zou, Hastie and Tibshirani(2006) is one of the main papers of Sparse Principle Component Analysis(SPCA) which is the method used by lasso, adaptive lasso, elastic net and adaptive elastic net estimators. Different from usual Principal Component Analysis(PCA), SPCA is modified using sparse loading to overcome drawbacks of PCA. They show that, PCA can be formulated as a regression type optimization problem , then imposing additional (adaptive)lasso or (adaptive)elastic net penalties to optimization problem sparse loading can be obtained. And the obtained algorithm is efficient for multivariate and high dimensional data sets. They apply SPCA algorithm to real and simulated data and provide encouraging results.

Zou and Zhang(2009) uses elastic net which is from the same family with bridge estimator in linear regression models. They use adaptive elastic net for both model selection and estimation purposes simultaneously similar to bridge estimator. But according to Zou and Zhang(2009), adaptive elastic net is better at prediction and has more stable solution paths rather than bridge and other shrinkage estimators. Also, it can handle large number of variables and so works better for correlated cases.

CHAPTER 3

METHODOLOGY

The main reason which leads us to this study is the mostly observed nonstationary characteristics of panel data sets. Most of the factor number estimation methods fail while working with nonstationary data due to "ill-nature" of nonstationary data and the ones not failing with nonstationary data are not robust to data changes. Since the true estimation of factor numbers is the hearth of all existing literature, it is very crucial to be able to estimate true factor number, even with nonstationary data, consistently and robust.

Before starting to talk about methodology we adopt for this study, let me introduce the main ideas and procedures about different bridge type estimators for the ease of understanding elastic net idea. The simplest idea behind shrinkage estimators is nothing but a parameter based penalty on estimation based objective function.

Assume that the model is:

$$Y_{it} = \beta_i' X_{it} + e_{it} \tag{3.1}$$

where X_{it} are independent variables, Y_{it} are dependent variables, β_i are regression coefficients and e_{it} are iid random variables with mean 0 and variance σ^2.

The penalized objective function of shrinkage family is

$$V_n(\beta) = [\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \beta_i' X_{it}) + \alpha \sum_{i=1}^{N} |\beta_i|^{\gamma}] \tag{3.2}$$

where $\gamma > 0$ and the α is called as penalty parameter.

the associated shrinkage estimate is

$$\hat{\beta}_{shrinkage} = argmin_{\beta}[\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \beta_i' X_{it}) + \alpha \sum_{i=1}^{N} |\beta_i|^{\gamma}] \tag{3.3}$$

For different values of γ, there is specific estimator names and associated properties of shrinkage estimators.

If $\gamma = 1$, it is named as lasso estimator in literature. Adopting the general objective function of shrinkage estimators with $\gamma = 1$ we obtain lasso objective function and its estimator equation as follows.

$$V_n(\beta) = [\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \beta_i' X_{it}) + \alpha \sum_{i=1}^{N} |\beta_i|] \tag{3.4}$$

$$\hat{\beta}_{lasso} = argmin_{\beta}[\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \beta_i' X_{it}) + \alpha \sum_{i=1}^{N} |\beta_i|] \tag{3.5}$$

If $\gamma = 2$, it is called ridge estimator. Very similar to lasso estimator, bridge estimator is obtained from the following equation

$$\hat{\beta}_{ridge} = argmin_{\beta}[\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \beta_i' X_{it}) + \alpha \sum_{i=1}^{N} |\beta_i|^2] \tag{3.6}$$

which is based on the main shrinkage objective in form of

$$V_n(\beta) = [\sum_{i=1}^{N}\sum_{t=1}^{T}(Y_{it} - \beta_i' X_{it}) + \alpha \sum_{i=1}^{N} |\beta_i|^2] \tag{3.7}$$

Different from these two specific forms $\forall \gamma \epsilon (0, 1)$, shrinkage estimators are called as bridge estimator which Caner (2011) uses for factor models.

As it can be understood from the equations of estimators there is single main common characteristic of all shrinkage family estimators; They are all based on similar penalized objective functions.

The role of objective functions of shrinkage estimators can be explained in terms of convergence. That is, the main idea behind objective functions is it always converge to a constant whether coefficients of interest are zero or nonzero. In other words, if factors are zero the objective function, $V_n(\beta)$ converges to zero coefficients and converges to any other constant for nonzero coefficients. The contribution of penalty parameters to penalized objective function is that it helps us to catch real nonzero parameters and differentiate it from real zeros. As penalty parameter increase, the probability of catching nonzeros increases due to minimization which shrinkage estimators are based on.

Most of the shrinkage estimators satisfy oracle properties and give consistent estimates. But due to their individual natures there are different advantages and weaknesses of each estimator over others. Bridge estimator is better than lasso in identifying zeros but lasso is easier to compute. On the other hand, the methodology we offered for this study is better than both bridge and lasso in handling larger dimensions and prediction.

The methodology that we propose for this study is called "Elastic Net" which is a convex combination of lasso and ridge estimators. In general, this method belongs to shrinkage estimators family similar to bridge estimators which is adopted to factor model literature by Caner (2011). Family of shrinkage estimators are used mostly in economics since they are robust to data changes and they allow cross correlation among variables which is one of the mostly observed characteristics on economics data sets. According to existing literature it is clear that elastic net does a better job than other shrinkage family estimators, there exist no study which aims to show the performance of elastic net estimators in factor number estimation.

Hence, this study aims to fill in this gap between elastic net and factor models-factor number estimation literature. In other words, this study extends the elastic net methodology from least squares context to factor models by using SPCA algorithm as an underlying algorithm.

To start with we consider the following approximate factor model of Chamberlain and Rothschild (1983) as our model of interest,

$$X_{it} = \lambda_i' F_t + e_{it} \tag{3.8}$$

where X_{it} are the observed data for i^{th} cross section unit for t^{th} time, λ_i is called r dimensional factor loadings, F_t is associated r dimensional vector of factors and e_{it} are stationary error terms for $\forall$ i=1,$\cdots$,N and t=1,$\cdots$,T. We call that equation as the factor representation of data where neither of the factor loading and factors are observable as well as unobservable error terms.

In that model the true number of factors are r. We aim to find r, true number of factors, given a fitted factor model with p factors where $p > r$. In other words, we aim to identify real r number of real non-zero factors and (p - r) zero factors consistently.

As we mentioned up to now, the main issue of this study is to work with nonstationary factors. According to Lansangan and Barios (2008), if the columns of time series in a panel are nonstationary, the first component may possibly combine all variables into a single factor since variance patterns are "ill-conditioned" and hence they offer two step SPCA algorithm - first is to perform ordinary PCA and then to find sparse approximations of loadings of PCA - for nonstationary time series data.

Under the model of Chamberlain and Rothschild (1983), we propose to estimate the factor loadings consistently by Zou, Hastie and Tibshirani (2006) elastic net estimation algorithm which is based on the method called SPCA.

Elastic Net with Factor Models

In our factor model set up which is based on model (3.8), the elastic net is a penalized least squares method imposing two penalties, α_1 and α_2 , on L1 and L2 norms of loading estimates respectively. For any non-negative α_1 and α_2 we define the elastic net estimates as follows

$$\hat{\Lambda}_{ENET} = (1+\alpha_2)argmin_{\Lambda} V_n(\lambda) \tag{3.9}$$

according to given objective function

$$V_n(\lambda) = [\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \lambda_i' F_t) + \alpha_2 \sum_{i=1}^{N} |\lambda_i|^2 + \alpha_1 \sum_{i=1}^{N} |\lambda_i|] \tag{3.10}$$

The elastic net penalty is a convex combination of ridge, on L2 norm, and lasso, on L1 norm, penalties. And the sparsity occurs as a result of L1 penalty on PCA for elastic net algorithm.

The result of this optimization gives the consistent estimate of factor loadings. Hence by counting the number of non-zero estimates we obtain the estimated count of true factor number and the rest is zero loading factors.

SPCA Algorithm

To solve the objective function of elastic net, Zou, Hastie and Tibshirani (2006) presents an effective algorithm. This algorithm is based on the way of transforming PCA to regression type optimization problem to solve the elastic net optimization in linear regression models. In other words, they try to carry on the connection between PCA and the regression and use elastic net approach to obtain sparse loadings. In this thesis, we adopt their SPCA algorithm to our work.

Hence, according to approach which Zou, Hastie and Tibshirani (2006) proposes we first find the principal components and associated loadings of data and than according to Cadima and Jolliffe (1995) we regress obtained principal components on variables using elastic net model to obtain sparse loadings.

Considering all given constraints of elastic net methodology detailed in previous section, all steps of the used SPCA procedure can be given as below.

Numerical Solution Algorithm

1) First, derive the loadings of first k ordinary principal components and store these loadings as a vector, say vector **A**.
2) Given this stored vector of ordinary principal component loadings, solve elastic net objective function given above by regression ordinary principal components results and obtain the initial estimates of factor loadings, lets say **B** vector.
3) Using the initial estimates of factor loadings, compute the singular value decomposition (SVD) of $X^T XB = UDV^T$ and update loadings vector, $A = UV^T$.
4) Define a convergence algorithm for this alternating solution. (Such as, repeat each steps until difference between updated values and previous ones are smaller than 0.05.)
5) Repeat all these steps until convergence.
6) The estimated values are the final estimated loadings. Then, normalize estimated loadings.

Threshold Selection Sensitivity

After solving the objective function of elastic net by using SPCA algorithm, we obtain vectors of estimated and normalized factor loadings. So, now we must decide which of these loadings are real zero and which loadings are nonzero. Hence we need to set a threshold rule to differentiate real zeros and nonzeros.

In existing literature, Caner (2011) sets 0.06 as a threshold to decide for nonzero factors in his study of using bridge estimators for factor number determination. On the other hand, Gorsuch (1983) proposes that 0.01 can be considered as a general threshold for nonzero factor loading selection.

In our study, we use a path of many trials to determine the appropriate threshold value for our cases. We conduct two different threshold selection stage for stationary and nonstationary factor model cases since the estimated loadings form of these cases are very different.

CHAPTER 4

SIMULATION STUDY

The simulation study section of our work consists of four main steps; the first step is performance study of elastic net with stationary factors as a baseline model which allows us to determine the threshold value for non-zero factors, the second is elastic net estimation under the nonstationary factor models with same threshold values of model with stationary factors. The third step is the determination of optimal parameters for elastic net with nonstationary factors. And the fourth and the last step is performance studies with mixed factors.

1) Elastic Net with Stationary Factors

The following DGP is proposed for our baseline set up. As most of the factor number estimation studies, DGP is adopted from Bai and Ng(2002).

$$X_{it} = \sum_{j=1}^{r} \lambda_{ij} F_{tj} + \sqrt{\theta e_{it}} \tag{4.1}$$

where the factors are T matrices of N(0,1) random variables and the factor loadings are also distributed as N(0,1).The variance of the common component of X_{it} is r.

Moreover, since we also consider the cross-sectional correlation case so we also let

$$e_{it} = \rho e_{it-1} + v_{it} \tag{4.2}$$

Here v_{it} are standard normal errors and ρ determines the degree of correlation.

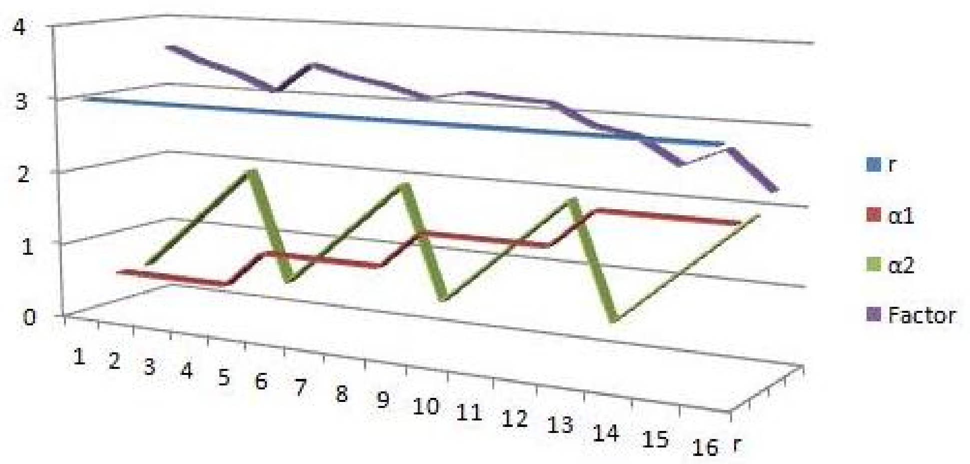
Caner(2011) also uses the same DGP for bridge estimators case.
On the other hand, we also chose different true factor numbers to see the performance of our method with different factor numbers.

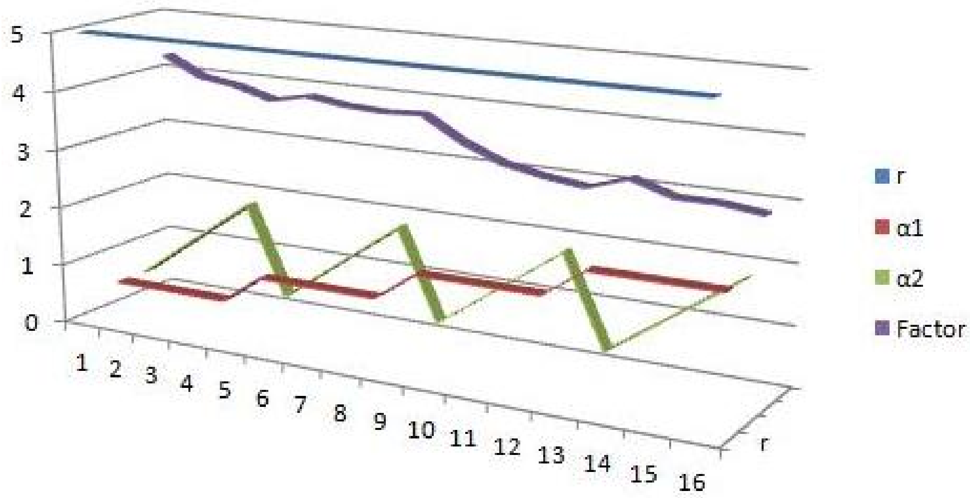
To start with, we set n=10 and t=50 for all of the simulations we work on the determination of optimal values of α_1 and α_2. We know that for a given value of α_2 algorithm of Zou, Hastie and Tibshirani (2009) estimates solution path for different values of α_1. The grid of (0, 5) is chosen for values of α_1 and the grid of (0, 5) is also used for α_2. As a result of preliminary simulations on chosen grids, (0.5, 2) is determined as the optimal range of penalty parameters for our stationary case. After range determination of optimal penalty parameter values, we try alternative choices of penalty parameters with alternative true factor number in our set up. Hence we try to find correct number of factor when r = 3, 5 respectively. For stationary factor simulation results, to determine a threshold value we analysed the results of pre- simulation results and chose $|0.01|$ as a threshold for nonzero factors.

Our Table 1 summarizes the results of elastic net estimator with stationary factors as an average of 500 replications.

Table 1 / Stationary Factors							
r	α_1	α_2	Estimated # of Factors	r	α_1	α_2	Estimated # of Factors
3	0,5	0,5	3,58	5	0,5	0,5	4,27
		1	3,38			1	3,95
		1,5	3,24			1,5	3,86
		2	3,03			2	3,68
	1	0,5	3,44		1	0,5	3,82
		1	3,3			1	3,73
		1,5	3,21			1,5	3,7
		2	3,07			2	3,74
	1,5	0,5	3,18		1,5	0,5	3,35
		1	3,14			1	3,08
		1,5	3,12			1,5	2,95
		2	2,86			2	2,87
	2	0,5	2,77		2	0,5	3,1
		1	2,44			1	2,89
		1,5	2,69			1,5	2,89
		2	2,21			2	2,81

The following graphics summarizes the relation between changes of model parameters and estimated factor number averages for true factor numbers 3 and 5, respectively.





2) Elastic Net with NonStationary Factors

The same DGP of baseline model is proposed again with a difference in factor structure. Similar to base model the DGP is as follows

$$X_{it} = \sum_{j=1}^{r} \lambda_{ij} F_{tj} + \sqrt{\theta e_{it}} \tag{4.3}$$

And the factors are generated according to

$$F_{tj} = \phi F_{(t-1)j} + w_{tj} \tag{4.4}$$

where e_{it} and w_{tj} are random error terms and $\phi = 1.3$ assumption implies factors are nonstationary.

At this step of our simulation study, we want to test the performance of elastic net estimator with nonstationary data under optimal conditions of stationary case. So using the optimal results of baseline model simulations; the optimal α_1 and α_2 grids and the optimal threshold value for non-zero coefficients, we run simulations with alternative correct factor numbers same as stationary case.

The results of elastic net estimators with nonstationary factors under the optimal conditions of stationary factors is not surprising. Independent from true factor number, elastic net algorithm estimates average factor numbers in (0.75, 1.25) range. So that is not an interesting story trying to estimate nonstationary factor numbers under optimal conditions of stationary factor estimations. The next step tries to find some ways to improve elastic net with nonstationary factors.

3) Optimality Studies for Elastic Net with NonStationary Factors

For all of the optimality studies of elastic net with nonstationary factors part of our study, the adopted DGP is exactly same with "Elastic Net with NonStationary Factors" step. Hence the model is

$$X_{it} = \sum_{j=1}^{r} \lambda_{ij} F_{tj} + \sqrt{\theta e_{it}} \tag{4.5}$$

Again, factors are generated according to

$$F_{tj} = 1.3F_{(t-1)j} + w_{tj} \tag{4.6}$$

where e_{it} and w_{tj} are random error terms as before.

After data generation, the optimality studies are performed on two alternative grounds of estimation algorithm. Since we know that lower penalty parameters may lead over estimation and that may cause to estimate factor numbers greater than 1 and that is what we need for nonstationary factors, first optimality performance for the change in grid of α_1 and α_2 is considered. And new grids for α_1 and α_2 are conducted as (0, 0.5) under nonstationarity case.

Then, alternative threshold values $|0.1|$, low, and $|0.3|$, high, for non-zero coefficients are studied.

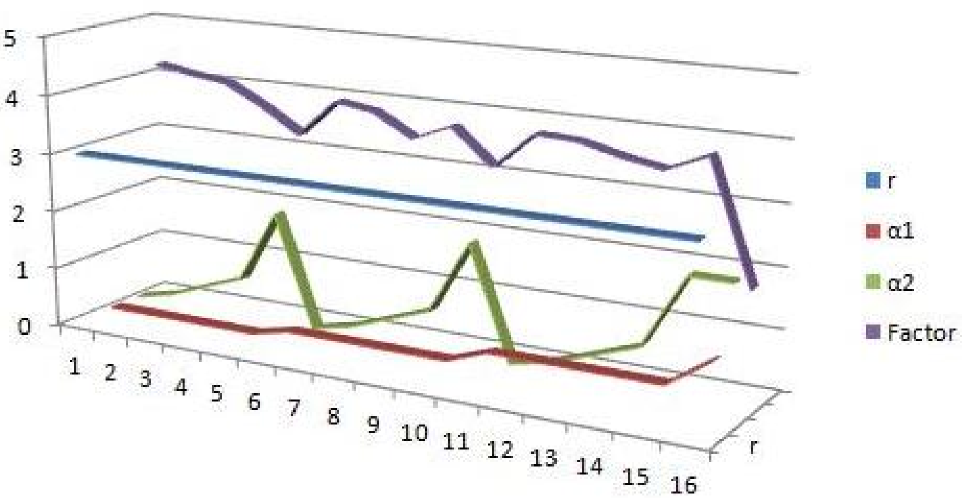
It is clear that nonstationary thresholds are absolutely higher than stationary ones to deal with the extra overestimation and high sparsity due to low penalty terms.

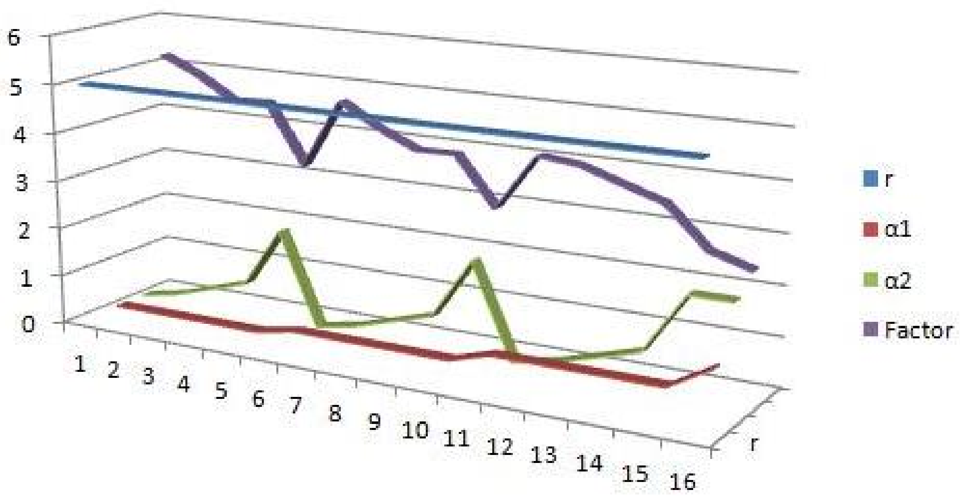
Table 2 and Table 3 reports the results of alternative improvement studies on elastic net estimators for nonstationary factors with different factor numbers.

Table 2 / Nonstationary Factors and Low Threshold

r	α_1	α_2	Estimated # of Factors	r	α_1	α_2	Estimated # of Factors
3	0,1	0,1	4,19	5	0,1	0,1	5,2
		0,25	4,07			0,25	4,85
		0,5	3,99			0,5	4,39
		0,75	3,65			0,75	4,43
		2	3,24			2	3,19
	0,25	0,1	3,88		0,25	0,1	4,63
		0,25	3,78			0,25	4,16
		0,5	3,41			0,5	3,81
		0,75	3,69			0,75	3,83
		2	3,1			2	2,84
	0,5	0,1	3,72		0,5	0,1	4
		0,25	3,69			0,25	3,92
		0,5	3,52			0,5	3,65
		0,75	3,4			0,75	3,39
		2	3,71			2	2,59
	1	2	1,69		1	2	2,34

The associated graphics for low threshold / nonstationary case are as follows.



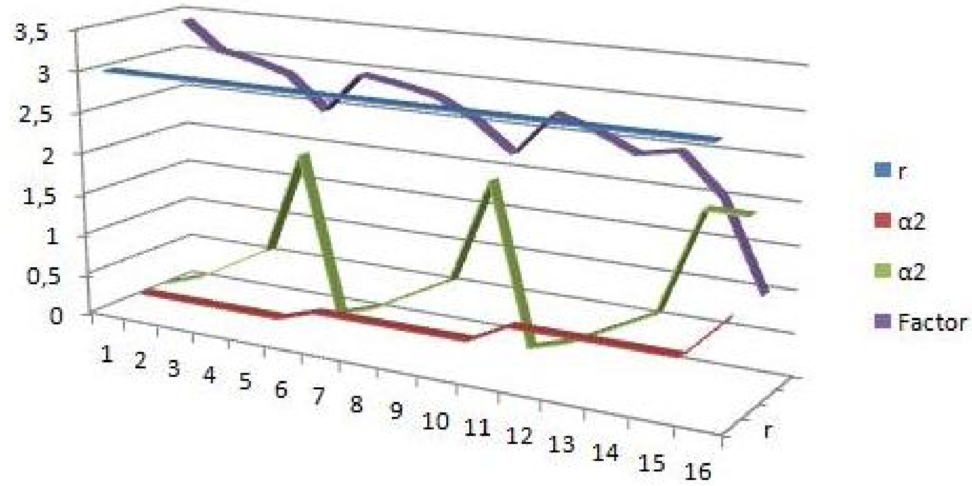


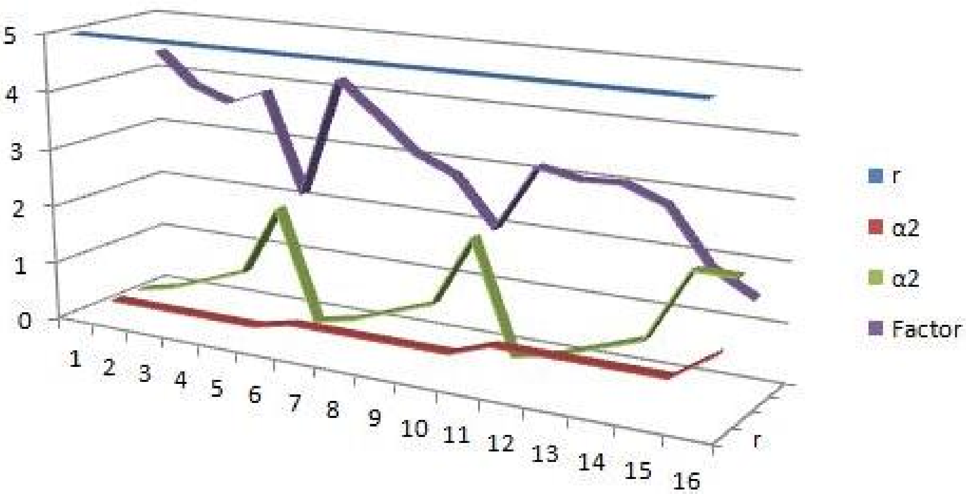
Similar to stationary factors case, the results of Table 2 and Table 3 are averaged over 500 replications for each combinations.

Table 3 / Nonstationary Factors and High Threshold

r	α_1	α_2	Estimated # of Factors	r	α_1	α_2	Estimated # of Factors
3	0,1	0,1	3,39	5	0,1	0,1	4,36
		0,25	3,05			0,25	3,79
		0,5	2,97			0,5	3,57
		0,75	2,84			0,75	3,83
		2	2,45			2	2,1
	0,25	0,1	2,94		0,25	0,1	4,19
		0,25	2,87			0,25	3,64
		0,5	2,78			0,5	3,07
		0,75	2,56			0,75	2,78
		2	2,21			2	1,97
	0,5	0,1	2,73		0,5	0,1	3,11
		0,25	2,59			0,25	2,98
		0,5	2,38			0,5	3,01
		0,75	2,47			0,75	2,75
		2	2,01			2	1,84
	1	2	0,98		1	2	1,45

The graphics summarizing high threshold / nonstationary case are as follows.





4) Performance Studies for Elastic Net with Mixed Factors

In performance studies of elastic net with mixed factors part, we used the DGP of stationary and nonstationary sections. Again the model is

$$X_{it} = \sum_{j=1}^{r} \lambda_{ij} F_{tj} + \sqrt{\theta e_{it}} \tag{4.7}$$

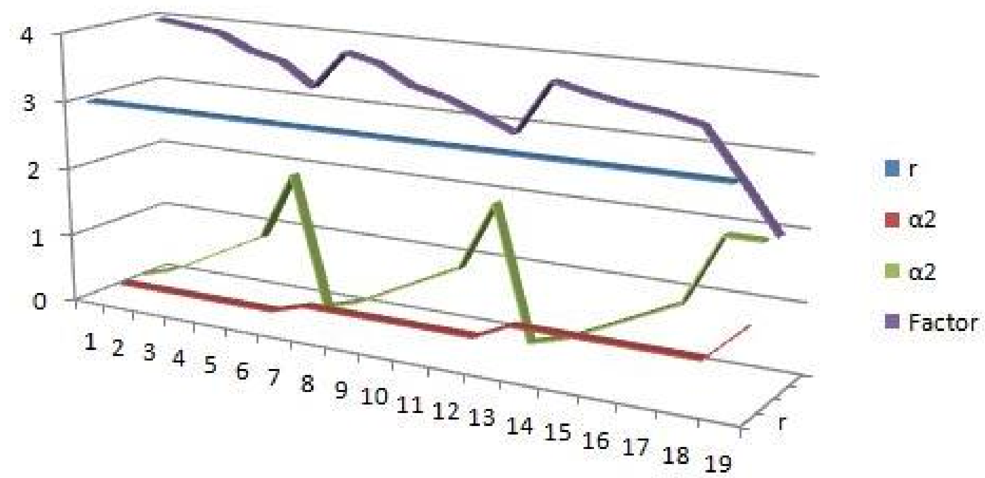
where factors are mixed of stationary and nonstationary factors generated according to previously outlined processes and e_{it} and w_{tj} are random error terms as before.

After data generation process, the performance studies for mixed factors are conducted on the same set up of optimal nonstationary factors case.

Table 4 and Table 5 report the results of performance studies on elastic net estimators with mixed factors with different factor numbers.

And the associated graphics for following tables show the relation of average factor numbers and model parameters for mixed factor simulations when nonstationary factor numbers are 1 and 2 for true factor number 3.

Table 4 / Mixed Factors (r=3)							
ns=1/s=2				ns=2/s=1			
r	α_1	α_2	Estimated # of Factors	r	α_1	α_2	Estimated # of Factors
3	0,1	0,1	3,97	3	0,1	0,1	3,42
		0,25	3,91			0,25	3,34
		0,5	3,85			0,5	3,28
		0,75	3,63			0,75	3,21
		1	3,52			1	3,15
		2	3,19			2	2,84
	0,25	0,1	3,75		0,25	0,1	3,26
		0,25	3,64			0,25	3,08
		0,5	3,37			0,5	2,99
		0,75	3,25			0,75	2,78
		1	3,06			1	2,73
		2	2,87			2	2,58
	0,5	0,1	3,65		0,5	0,1	2,83
		0,25	3,53			0,25	2,8
		0,5	3,44			0,5	2,74
		0,75	3,38			0,75	2,69
		1	3,27			1	2,65
		2	2,61			2	2,19
	1	2	1,92		1	2	2,01



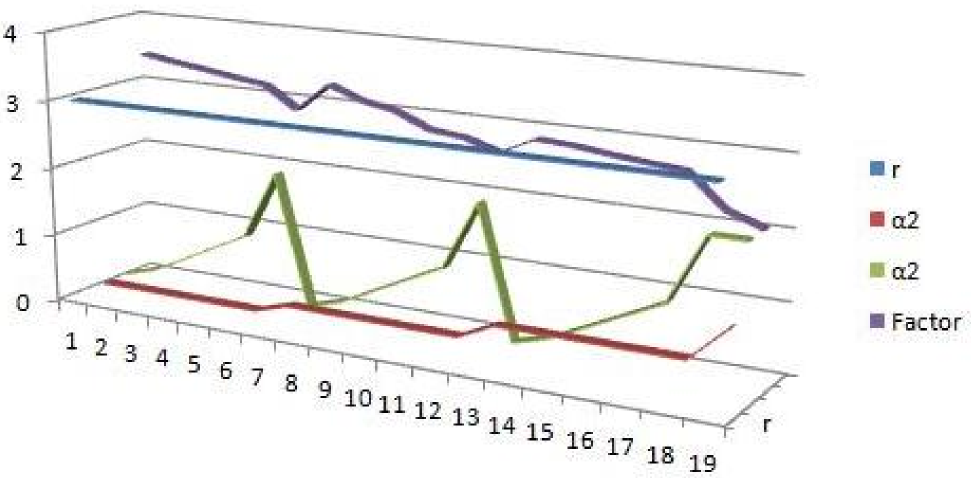
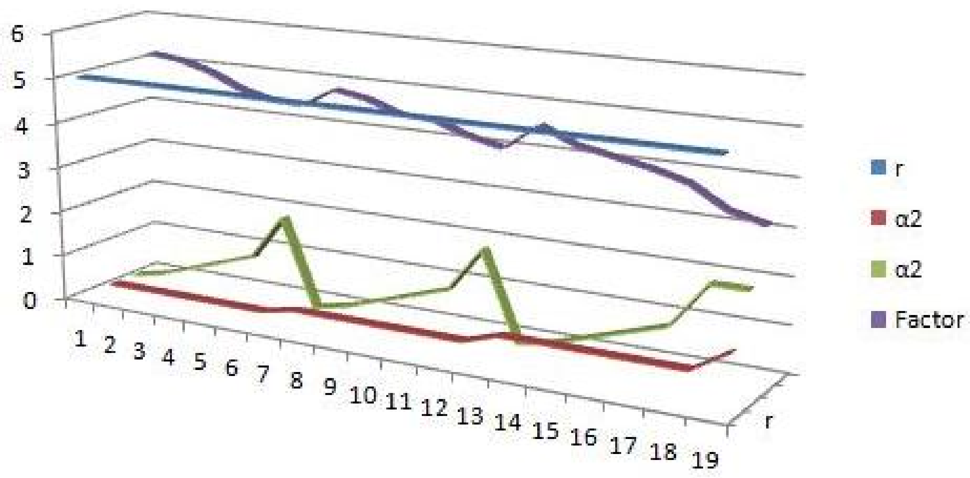
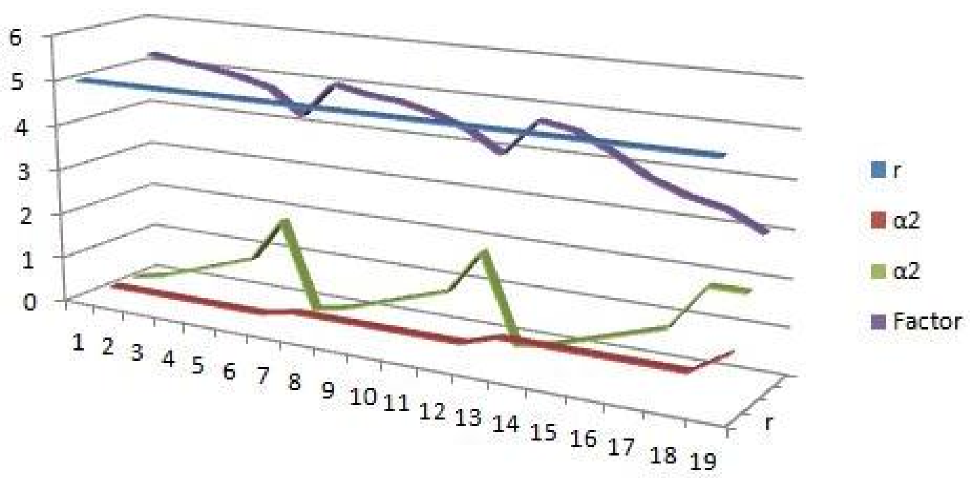


Table 5 / Mixed Factors (r=5)							
ns=2/s=3				ns=4/s=1			
r	α_1	α_2	Estimated # of Factors	r	α_1	α_2	Estimated # of Factors
5	0,1	0,1	5,14	5	0,1	0,1	5,21
		0,25	5,06			0,25	5,1
		0,5	4,87			0,5	5,02
		0,75	4,56			0,75	4,9
		1	4,38			1	4,72
		2	4,36			2	4,19
	0,25	0,1	4,78		0,25	0,1	4,97
		0,25	4,66			0,25	4,82
		0,5	4,41			0,5	4,75
		0,75	4,37			0,75	4,58
		1	4,12			1	4,33
		2	3,96			2	3,89
	0,5	0,1	4,53		0,5	0,1	4,69
		0,25	4,21			0,25	4,57
		0,5	4,06			0,5	4,22
		0,75	3,92			0,75	3,78
		1	3,72			1	3,51
		2	3,27			2	3,35
	1	2	3,09		1	2	2,97

Similar to previous graphics, the following graphs show the relation of average factor numbers and model parameters for mixed factor simulations when nonstationary factor numbers are 2 and 4 for true factor number 5, respectively.


6
5
4
3
2
1
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
r
r
α2
α2
Factor


6
5
4
3
2
1
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
r
r
α2
α2
Factor

CHAPTER 5

CONCLUSION

In literature, factor models are widely used in finance and economics in areas such as asset pricing, cross-correlation analysis, cross section co-integration and forecasting. Parallel to their wide use, there are increasing number of studies aiming to estimate factor number consistently as the focus of factor models studies. As a result, there are alternative methodologies offered to estimate true factor numbers but none of these methodologies have absolute advantage which make it best among others. Hence, existing drawbacks of each method makes factor number estimation still an hot and challenging area.

In this thesis work, we studied the estimation of true factor numbers in factor models with stationary, nonstationary and fixed factors. We used elastic net estimator which is a member of shrinkage estimators family whom some other members like bridge estimators are already used in factor number estimation before. In our study, we adopted elastic net estimator from least squares literature to factor model concept. The estimation method is applied to both simulated stationary and nonstationary data and also mixed of stationary and nonstationary data sets with different true factor numbers. As a result of simulation studies, it is found that the ranges of penalty parameters for stationary and nonstationary factors are different from each other.

That is why for nonstationary models there is always an underestimation problem similar to other existing methodologies and we try to deal with this problem by using low penalties for which overestimation is expected. Hence, using lower penalties causing overestimation we aim to balance underestimation problem by pushing estimated factor numbers up. On the other hand, it is detected that using low penalties turns underestimation problem to over estimation problem for lower factor numbers. To deal with generated overestimation problem we alternatively tried different threshold values for selecting zero and nonzero factors.

According to all stationary and nonstationary simulation results and mixed factor structure simulation results of our study, we can summarize the findings that using different penalty parameter considering stationary/nonstationary nature of factors and working with changing penalty parameters based on the true factor numbers may lead more stable and consistent results for elastic net estimator for factor models. In other words, setting a general simultaneous penalty selection and threshold rules based on number of true factors and status of stationary being of a dataset elastic net estimators can work for most of the factor number estimation studies in economics.

BIBLIOGRAPHY

Bai, Jushan, 2004. "Estimating cross-section common stochastic trends in nonstationary panel data, "'Journal of Econometrics, Elsevier, vol. 122(1), pages 137-183, September.

Bai, Jushan Kao, Chihwa Ng, Serena, 2009. "Panel cointegration with global stochastic trends," Journal of Econometrics, Elsevier, vol. 149(1), pages 82-99, April.

Bai, Jushan Ng, Serena, 2002. "Determining the Number of Factors in Approximate Factor Models," Econometrica, Econometric Society, vol. 70(1), pages 191-221, January.

Bernanke, Ben S. Boivin, Jean, 2003. "Monetary policy in a data-rich environment," Journal of Monetary Economics, Elsevier, vol. 50(3), pages 525-546, April.

Caner, 2011. Using Bridge Estimators to Determine Number of Factors in Multifactor Models: Case of Large Panel Data, Paper Under Submission.

Chamberlain, G. and Rothschild, M. 1983. Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets, Econometrica 51, 13051324.

Forni, M. and Reichlin, L. , 1998. "Lets Get Real: a Factor-Analytic Approach to Disaggregated Business Cycle Dynamics," Review of Economic Studies 65, 453473.

Onatski, Alexei, 2009. "Testing Hypotheses About the Number of Factors in Large Factor Models," Econometrica, Econometric Society, vol. 77(5), pages 1447-1479, 09.

Pesaran, M. Hashem, 2006. "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," Econometrica, Econometric Society, vol. 74(4), pages 967-1012, 07.

Ross, S. , 1976. "The Arbitrage Theory of Capital Asset Pricing", Journal of Finance 13, 341360.

Stock, J. H. and Watson, M. ,1989. "New Indexes of Coincident and Leading Economic Indications,in O. J. Blanchard and S. Fischer (eds)", NBER Macroeconomics Annual 1989, M. I. T. Press,Cambridge.

Stock, J. H. and Watson, M. W. ,1999. "Forecasting Inflation", Journal of Monetary Economics 44, 293335.

Zou, Hastie Tibshirani, 2006. Sparse Principal Component Analysis, Journal of Computational and Graphical Statistics,ASA, vol. 15(2), pages 265-286, June 1, 2006.

Zou Hastie, 2003. Regression Shrinkage and Selection via the Elastic Net, with Applications to Microarrays, Technical report, Department of Statistics, Stanford University.