Stockholms universitet

Forecasting using principal components from many predictors

Ying Pang

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2011:9
Matematisk statistik
September 2011

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Master Thesis **2011:9**
http://www.math.su.se

Forecasting using principal components from many predictors

Ying Pang*

September 2011

Abstract

This paper studies macroeconomic forecasting with many predictors, and primarily uses approximate dynamic factor models, in which a small number of latent factors summarize the information of predictors. Principal component analysis has been brought into widely use to estimate factors, and four kind of principal component (PC) estimators are examined in details, which are standard PCs, weighted PCs, generalized PCs and dynamic PCs. Furthermore, targeted PC estimation is proposed as another type of estimated factors in the paper, in order to take into account the purpose of forecasting variables of interest, which is accomplished by ordering PCs with their importance to prediction and selecting a handful targeted PCs. Then all these estimates are compared in the context of forecasting macroeconomic series at various time horizons, meanwhile other many-predictors forecasting methods are considered into the evaluation.The empirical data consists of 132 U.S. monthly time series available from 1969:1 to 2003:12. The results suggest that one or two estimated factors can valuably summarize the information from many predictors; however, forecasts are different and relative to variables of study.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: ypang@math.su.se. Supervisor: Martin Sköld.

Preface

This thesis constitutes a Master's project for the degree of Master in Financial Mathematics and Finance, Department of Mathematics, Stockholm University.

Acknowledgement

I would like to give my sincere appreciation to my supervisor, Martin Sköld, for his great guidance and support.

I also would like to thank my dear family for their unconditional love.

Contents

1. Introduction

During the past decade, there has been a great demand for macroeconomic forecasts using many predictors which encourage the development of producing accurate forecasts. To forecast macroeconomic variables of interest, for example industrial production, employment rate, price index and so forth, the dataset usually covers a huge number of observed time series for a long period. Whether ignoring or considering all these relevant variables would definitely influence forecasting accuracy and may result in suboptimal forecasts. Therefore, econometricians have been developing effective ways to subtract information available among these predictors to improve the performance of forecasting both in the theoretical and empirical perspective.

Traditional econometric models, such as multivariate vector autoregressive models, or univariate autoregressive models, have limitations when handling many times series. Consequently, in order to relax restrictions to low dimensional data and improve forecasting performance, there are plenty of studies that introduce different methods and examine them both in their theoretical and empirical aspects. In general, there are three classes of available methods (Stock and Watson (2006)). The first class is combining a large panel of forecasts based on relatively simple models, which compute individual forecasts based on multivariate models and then produce a combination of these forecasts with reasonable weights depending on some historical measure. Bates and Granger (1969) originally proposed the theory of forecasts combination, and the improvement of forecasting has been successfully examined in applications, such as Timmerman (2004) and Stock and Watson (2006). The second class is taking account of several latent factors to cover the information in predictors, which certainly involves factor analyses, and then factor augmented forecasts are produced. We will detail and discuss both of methods in this paper.

The third class is to reduce the sample error and improve the forecast accuracy, such as shrinkage, model selection and model averaging methods. In terms of minimizing the mean squared sample error (MSE), many standard estimators can be improved by combining them with other information. Shrinkage is the reduction in the effects of sampling variation by shrinking them towards a fixed constant, and then improved estimators are closer to values that other information prefers. Shrinkage methods have become familiar from Stein (1955) and James and Stein (1960), and the use in the regression analysis, where there are a lot of explanatory variables, has been explained by Copas (1983). Besides, Bayesian model averaging (BMA) is, generally speaking, an extension of combined forecasting to a fully Bayesian setting, and Beyasian model selection suggest selecting models based on Bayes factor in order to use aimed model as the basis for forecasting. Originally, Leamer (1978) discussed BMA in regression models, Min and Zellner (1990) developed BMA in macroeconomic forecasting, After that, researches of prediction with large-*n* regressors in the Bayesian framework have been active lately, meanwhile other model averaging methods are studied as well, such as “bagging” (Breiman (1996)).

Along with development of methods, it is difficult to draw conclusions of measuring the performance across these methods. In one hand, derivations and relative results are certainly dependent on different modeling assumptions, which lead to difficulties of comparison in the theoretical perspective. For instance, it is complicated to identify whether assumptions always hold or not in the whole algorithms and processes. In another hand, the variability and applicability of data sets could be one of many problems in empirical aspect. Therefore, the results of the research are mixed and frail with regard to the model specification, data requirement, forecast horizons and areas, implementation, etc. It is definitely an issue to be worth investigating and analyzing in depth.

In this paper, we focus on the second class method, determining a small number of latent factors that account for variations of observed variables. Descriptive statistics can provide a simple summary of the sample, and principal component analysis (PCA) is a standard tool for descriptive analysis (Anderson (1984)). Usually, summarizing a large set of data with several factors might result in missing information. Therefore, factor-based forecasting with principal component (PC) estimation can be utilized to extract a small number of latent factors which could contain the most information of data set and improve the prediction accuracy for macroeconomic variables. According to recent literature, the dynamic factor model (DFM) has received much attention, and become a feasible solution to the forecasting problem of many potential useful predictors.

In economics, the initial contribution of this field derives from Burns and Mitchell (1946) about business cycle analysis. Factor models were oringinally extended to the DFMby Geweke (1977) and Sargent and Sims (1977), and they exploit the dynamic interrelationship of variables and provide evidence in favor of reducing the number of common factors. However, the approach does not yield estimates of factors so that cannot be used for prediction. Furthermore, it is based on a strong assumption regarding the uncorrelated idiosyncratic terms in models. Therefore, Chamberlain (1983) proposed the approximate DFM relaxing the assumption, detailed in Section 3.2.1. Under additional conditions, Engle and Watson (1981) and Stock and Watson (1989, 1991) estimated dynamic factors by Kalman smoother. The use for forecasting is proposed by Stock and Watson (1999), and they applied DFMs into U.S. macroeconomic data and suggested the consistent PC estimators of latent factors (Stock and Watson (2002a, b)). Bovin and Ng (2003) employ weighted PC estimates to improve standard PC forecasts, and suggest that factor models outperform traditional economic models with a limited number of predictors. Generalized DFMs developed PCs in the frequency domain by Forni and Lippi (2001), Forni, Hallin, Lippi, and Reichlin (2000, 2003, 2004, 2005). In addition, there are a lot of studies of macroeconomic forecasting with factor models, such as Bernanke and Boivin (2003), Bovin and Ng (2005, 2006), Bai and Ng (2006, 2008) and so forth, the DFM and PCA are seen to be complementary to each other in a growing study of macroeconomic forecasting methods.

Our goal is first to propose a method for factors estimation, named targeted PCs, in which PCs are ordered by the importance for forecasting. Recall that the goal of PCA, as a variable reduction procedure, is to simply obtain a relatively small number of factors that account for most of variations in a large number of observed variables. Rather than explaining most variations of many predictors, targeted PCs consider information from variables of forecasted, which could obtain more accurate forecasts. Aiming to forecasting, it is interesting to examine which PCs are important for prediction, and it can be realized by ordering correlations between PCs and forecasted series, or equivalently the individual *t*-values of PCs in the regressions. And, targeted PCs are examined with four kinds of estimations together, which are standard PCs, weighted PCs, generalized PCs, and dynamic PCs. The second purpose is to compare the performance of forecasting methods, which include univariate autoregression as benchmark, combined multivariate autoregression, and DFMs with all available PC estimations. To evaluate the performance, mean squared of forecast error (MSE) and a ratio of the MSE for a method to the MSE of the benchmark can be considered as the criterion, which is defined as relative mean squared error (RMSE), $RMSE = \frac{MSE(method)}{MSE(benchmark)}$.

The empirical analysis collects 131 U.S. monthly macroeconomic time series during the 1960-2003. Two of these series, index of industry production (IP) and consumer price index (CPI), are used to construct the forecasted variables, and multi-step ahead out-of-sample forecasts are estimated at horizons of 6 and 12 months. The results confirm factor models can produce substantially more accurate forecasts. IP forecasts using two factors outperform benchmark, while, performance of forecasts with one targeted factor estimate and lagged IP variables together can achieve comparable improvement. And CPI forecasts can be constructed by lagged variables and only one factor, which are sensitive to the dataset pretreatment. Forecasts with standard, generalized, and dynamic PC estimates act analogously, and targeted PCs normally perform as similar as corresponding non-targeted PCs.

The rest of the paper is organized as follows. In section 2, we introduce the forecasting procedure and methods employed. As the key point of this paper, the dynamic factor model is detailed in the way associated with four kinds of factor estimations and their targeted PC estimations. Section 3 describes the data, explains how to construct forecasted variables, and details the forecasting procedure. Section 4 provides the empirical results and illustrations for different time horizons, estimated factors and forecasted variables. In section 5, we present the conclusions and discuss the potential problems.

2. Forecasting Procedure and Methods

The forecasting experiment denotes the variable to be forecasted, Y_t, and a set of *n* predictors collected in the $n \times 1$ vector X_t. In the literature of forecasting methods, X_t and Y_t are usually assumed to be stationary, which is also an assumption we make in our paper as well. In addition, the predictors X_t need to be pretreated in a same way standardized to have zero mean and unit variance through the whole empirical comparisons. Let *h* be the forecast horizon and Y_{t+h}^h be the *h*-step-ahead forecasts. The regular forecasting regressions, in which the forecast at time *t* is denoted by $Y_{t+h|t}^h$, are the projections of an *h*-step-ahead variable Y_{t+h}^h onto *t*-dated predictors, intercept, and lagged predictors if necessary. For the multiple forecasts, let $Y_{i,t+h|t}^h$ be the i^{th} individual of the all available forecasts, such as the forecast combination described in the following. Now, we turn to the detailed description of forecasting methods.

2.1 Univariate Autoregression

The direct way of forecasting a time series is using linear models based on its previous observation, without any information from other variables, and it is easy to be applied by autoregressive (AR) models. Here consider a univariate AR(p) model as a benchmark,

$$Y_t^{AR} = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \varepsilon_t \quad , \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2) , \tag{1}$$

where the lag order p is determined by the information criteria of model selection, such as Akaike information criteria (AIC) or Bayes information criteria (BIC).

The h-step-ahead forecast of the AR(p) model is the projection of $Y_{t+h}^{AR,h}$ onto lagged variables with specified p orders, which is

$$Y_{t+h}^{AR,h} = \alpha + \phi(L) Y_t + \varepsilon_{t+h} , \tag{2}$$

where $\phi(L) = \phi_0 + \phi_1 L + \ldots + \phi_p L^p$ is a lag polynomials of order p, and L is the lag-operator such that $L^l Y_t = Y_{t-l}, l = 1, \ldots, p$.

Then, the forecast at time T, $Y_{T+h|T}^{AR,h}$, is estimated by

$$\hat{Y}_{T+h/T}^{AR,h} = \hat{\alpha} + \hat{\phi}(L) Y_T \tag{3}$$

where the coefficients are estimated by the ordinary least squares (OLS) based on the model (2) for t = 1,…, T-h.

2.2 Combined Multivariate Autoregression

To introducing other predictors, autoregressive distributed lag (ADL) models are considered. For each predictor, an individual forecast is obtained using the ADL model which only does involve historical data of this predictor and forecasted variable. Specifically, the i^{th} individual forecast $Y_{i,t+h}^{ADL,h}$ can be constructed with the corresponding i^{th} elements of X_t, X_{it}, and variables, Y_t. Both X_{it} and Y_t include current and lagged variables if necessary. Thus, the i^{th} forecast at time *T* is described as follows

$$\hat{Y}_{i,T+h|T}^{ADL,h} = \hat{\alpha}_i + \hat{\beta}_i(L)X_{iT} + \hat{\phi}_i(L)Y_T \qquad (4)$$

where the coefficients are estimated by the OLS regression onto data for *t* = 1,…, *T-h* as well. And then, all individual forecasts are combined with the given weights, that is the combined forecast at time *T* is constructed as follows

$$\hat{Y}_{T+h|T}^{comb,h} = w_0 + \sum_{i=1}^{n} w_i \hat{Y}_{i,T+h|T}^{ADL,h} \qquad (5)$$

where w_i is the weight on the i^{th} individual forecast of the i^{th} ADL model.

Granger and Ramanathan (1984) suggest that all the weights are estimated by OLS or restricted least squares with the constraints $w_0 = 0$ and $\sum_{i=1}^{n} w_i = 1$. However, when n is large, the estimates of the combined weights are expected to perform poorly, because estimating a large number of parameters can bring in considerable sampling uncertainty. As a matter of fact, if n is proportional to the sample size *T*, the OLS weight estimates are not consistent, and the combined forecasts based on these estimates are not asymptotically optimal, leading to large sampling errors and poor forecasts (Stock and Watson (2006)). In order to handle the case of combing weights with large *n*, several methods are introduced, depending on the various data requirement, such as simple combination, time-varying parameter weights, discounted RMSE weights, etc. In our paper, the combined forecasts based on the ADL models serve as an alternative benchmark rather than our focus, thus simple average combinations are used in the following empirical experiment, where forecasts are weighted by the mean, median, and 2-percentage-symmetrically-trimmed mean of the individual forecasts, as in Stock and Watson (2005).

2.3 Dynamic Factor Model

Now turn to the key of our paper. In this sub section, start by introducing DFMs, explain algorithms for various factor estimations, then lay out some concise information to estimate the number of factors, and end up by mentioning several other available methods of factor estimation besides PCA.

2.3.1 Specification of DFM

In DFM, each times series of a rich set of observed variables can be described as two parts, which are a common component χ_t, explained by a small number of dynamic factors, and an idiosyncratic disturbance in the following

$$X_t = \chi_t + \xi_t = \lambda(L) f_t + \xi_t \tag{6}$$

where is f_t the $q \times 1$vector of dynamic factors, $\lambda(L) = \lambda_0 + \lambda_1 L + \ldots + \lambda_s L^S$ is the $n \times q$ matrix polynomials in the lag operator *L* with finite order *s*, called dynamic factor loadings, and ξ_t is the $n \times 1$vector of idiosyncratic disturbances, which are specific for every time series. The factors and disturbances are assumed to be mutually uncorrelated at all leads and lags, and both of them are considered to be stationary like all the variables. Direct *h*-step ahead forecast is the regression of Y_{t+h}^{h} onto factors and lagged observables

$$Y_{t+h}^{h} = \alpha + \gamma(L) f_t + \phi(L) Y_t + \varepsilon_{t+h} \tag{7}$$

where $\gamma(L)$ is the lag polynomials.

Let $F_t = (f_t, f_{t-1}, \ldots, f_{t-s})'$ denote factors that include current and/or lagged dynamic factors. We call the *r* entries of F_t "static factors", corresponding to the *q* entries of f_t "dynamic factors", generally $q \leq r$. As the static representation of the DFM, models (6) and (7) can be rewritten

$$X_t = \Lambda F_t + \xi_t \tag{8}$$

$$Y_{t+h}^{h} = \alpha + \gamma' F_t + \phi(L) Y_t + \varepsilon_{t+h} \tag{9}$$

where $\Lambda = (\lambda_0, \lambda_1, \ldots \lambda_s)$ is a $n \times r$ matrix of factor loadings。

Factors are loaded contemporaneously under a condition $r = q(s+1)$, and γ is a vector of parameters composed of the elements of $\gamma(L)$, that is $\gamma = (\gamma_0, \gamma_1, \ldots, \gamma_s)'$. In the sequel, the forecasts at time *T* are estimated by

$$\hat{Y}_{T+h|T}^{DFM,h} = \hat{\alpha} + \hat{\gamma}' \hat{F}_T + \hat{\phi}(L) Y_T \tag{10}$$

where $\hat{F}_T$ is constructed from *T*-dated X_t based on the model (8), and the coefficients are OLS estimates based on the model (9) for sample period *t* = 1,…, *T*-*h*, constructed in a similar way as before.

There is a problem to be addressed about approximate DFMs. When *n* is large, different technical assumptions are made in the previous studies in order to keep a tight rein on the contribution of the idiosyncratic covariance to the total covariance. The concept of exact and approximate DFMs is introduced by Chamberlain and Rothschild (1983). The idiosyncratic terms are taken to be mutually uncorrelated at all leads and lags in the exact DFM, which is a strong underlying assumption. In macroeconomic forecasting application, the condition of exact DFMs cannot hold when dealing with large-*n* problems, and computational problems for MLE over increasing number of parameters become significant. Therefore, the approximate DFM allows for a limited amount of correlation among idiosyncratic disturbances, both serially and weakly cross-sectionally correlated, which relaxes the assumption of the exact DFM. And it provides a feasible way, in practice, to consistently estimate potential factors by employing non-parametric methods to approximate DFMs.

Facing the case of low-dimensional data, the maximum likelihood estimation (MLE), as a parametric approach, is successfully used for estimating parameters of exact DFMs (Stock and Watson (1991), Quah and Sargent (1993)), and has been further developed and implemented. Non-parametric approaches deal with large-*n* problems with a weakly correlation structure among idiosyncratic disturbances, of which several PC estimations are described in the following.

In our paper, forecasting with DFM is carried out in a two-step procedure. At first, a handful of appropriate factors are estimated, where five methods are used for factors estimation and detailed in Section 2.3.2. Secondly, we employ each kind of factor estimate into models and produce pseudo out-of-sample forecasts of our interests.

2.3.2 Estimations of Factors

In factor analysis, variability among observed variables can be described by a small number of unobserved variables, called latent factors, and variables are usually modeled as linear combinations of latent factors and "errors" terms together, as descried in models (6) or (8). A reduced number of potential factors can extracted from many variables to measure joint variations in a dataset, meanwhile, error terms cover the part of variability that cannot be explained by factors.

PCA performs an orthogonal transformation onto original data set, which takes into account all variability of variables. The new created set, named PCs, are ordered in variance, or considered as a variance-maximizing rotation of variable space. PCs are computed by seeking a matrix ***V*** consisting of the set of all eigenvectors of covariance matrix ***C*** such that $V^{-1}CV = D$, where ***D*** is the diagonal matrix of eigenvalues of ***C***. PCA and Factor analysis are closely related, and theoretically they are equivalent if error terms are assumed to all have the same variance in factor models. Moreover, similar empirical results yield in practice.

In our case, latent factors can be estimated by PCs. Based on the model (8) and its assumptions, the covariance matrix of ***X***, is the sums of two parts, $Cov(X_t) = \Lambda \cdot Cov(F_t) \cdot \Lambda' + Cov(\xi_t)$, where $Cov(F_t)$ and $Cov(\xi_t)$ are covariance matrix of F_t and ξ_t, and it is well known as the variance decomposition in classic factor analysis. For large-*n*, the eigenvalues of $Cov(\xi_t)$ are $O(1)$ and $\Lambda'\Lambda$ is $O(n)$, and then first *r* eigenvalues of are $Cov(X_t)$ are $O(n)$ and the rest eigenvalues are $O(1)$. Connecting with the eigenvalues and eigenvectors of covariance matrix in PCA, first *r* PCs of ***X*** are suggested being the estimate of Λ, subject to $\Lambda'\Lambda = I$ .v Next, PCs are obtained from all predictors and then a specific number of PCs can be estimated factors. Several kinds of factor estimations are explained in details, which are standard PCs, weighted PCs, generalized PCs, dynamic PCs, and targeted PCs we propose. Besides, PCs depend on scaling of data, thus typically we first standardize variables.

Standard Principal Component Estimation

Stock and Watson (2002a) show that, without further conditions on the relative rates of *n* or *T*, the space of the dynamic factors is consistently estimated by the PCs estimators as *n* and *T* towards infinite in the static form of the DFM. The factor loadings, Λ, and factors, F_t, solve the following least square problem

$$\min_{F_1,\ldots,F_T,\Lambda,\Lambda'\Lambda=I_k} \sum_{t=1}^{T}(X_t - \Lambda F_t)'(X_t - \Lambda F) \tag{11}$$

The solution is to set $\hat{\Lambda}$ to be the first *k* eigenvectors of sample covariance matrix of X_t, $\hat{\Gamma} = \frac{1}{T-1}\sum_{t=1}^{T} X_t X_t'$. And estimated factor scores of *r* PCs are calculated by $\hat{F}_t = \hat{\Lambda}' X_t$, which all information of predictors can approximately transform to first *r* PCs of ***X***.

Weighted Principal Component Estimation

When PCs are estimated, original data have been standardized to have unit sample variance. In another way, it can be considered being simply weighted by the inverse sample variance. Involving different weighting schemes, weighted PC estimates are proposed, and they are examined to improve upon conventional PCs (Jones (2001)). Different weighting schemes onto the variables are developed, like Bovin and Ng (2003), Forni, Hallin, Lippi, and Reichlin (2003), and the analyses show that weighted PCA can improve upon conventional PCA, but the gains depending on the particular of series under study.

Here we propose a weighting scheme that depends on the number of predictors from same categories. For instance, there are five series of "Housing Starts" in the U.S. macroeconomic data, but two series of "S&P's Common Stock Price Index", which indicates the PCs relative to "housing starts" should obtain more weight than those relative to common stock price. Given the assumption that all category variables are of equal importance, all categories of predictors need to be downweighted, which every series is divided by the square root of the number of series in the corresponding category. More precisely, divide each "Housing Starts" series by the square root of five and each "S&P's Common Stock Price Index" series by the square root of two as examples. Then weighted PC estimations are first r PCs of weighted predictors.

Generalized Principal Component Estimation

In classical regression analysis, generalized least square (GLS) is more efficient than OLS, thus we modify formula (11) and consider the GLS problem

$$\min_{F_1,\ldots,F_T,\Lambda,\Lambda'\Lambda=I_k} \sum_{t=1}^{T}(X_t - \Lambda F_t)'\tau_{\xi}^{-1}(X_t - \Lambda F) \qquad (12)$$

where τ_{ξ} is the covariance matrix of idiosyncratic disturbance.

Because τ_{ξ} is not observable in applications, the analogous GLS estimator can be estimated by covariance matrix of sample residuals from the model with *r* factors, $\hat{\tau}_{\xi}$. However, the matrix $\hat{\tau}_{\xi}$ has the rank *n-r*, so it is not invertible, and $\hat{\tau}_{\xi}$ is not diagonal in approximate DFMs. Although optimal GLS estimates is impossible, Boivin and Ng (2003) propose a feasible approach to replace τ_{ξ} with a diagonal estimator, where off-diagonal elements of $\hat{\tau}_{\xi}$ are set to zero to overcome the instability in the generalized PCs methodology. With this restriction, generalized PC estimates can be considered as PCs of the transformed data, which Boivin and Ng (2003) found perform well in the empirical application in U.S. data. More specifically, at first, diagonal elements of $\hat{\tau}_{\xi}$ are estimated by the sample variance of the residuals from the preliminary regression of ***X*** onto the fixed *r* standard PC estimates based on the model (8), and then generalized PC estimations are first *r* PCs of the transformed data $\tilde{X}_t = (\hat{\tau}_{\xi})^{-1/2} X_t$.

The generalized PC estimates would perform better than standard PCs, because it introduces the inverse covariance matrix of idiosyncratic terms, which may be considered as the weights. In principle, the larger weights are placed on the variables with small idiosyncratic components, thus the idiosyncratic error is minimized in the estimates.

Dynamic Principal Component Estimation

Boivin and Ng (2004) and Forni, Hallin, Lippi, and Reichlin (2005) suggest a two-step estimation algorithm for the dynamic factors through dynamic PCA. The first step involves the dynamic techniques to estimate the covariance matrix of common and idiosyncratic components. In the second step, the generalized eigenvectors associated with these estimated covariances are obtained in order to construct the dynamic PC estimates, which are the linear combinations of contemporaneous ***X***'s having the minimal idiosyncratic-common variance ratio.

By introducing the lag-*k* covariance matrix of ***X***, $\Gamma^k = E\left[X_t X'_{t-k}\right]$, the variance decomposition, in the time domain, can be described as follows

$$\Gamma^k = \Gamma^k_\chi + \Gamma^k_\xi = \Lambda \Gamma^k_F \Lambda' + \Gamma^k_\xi \tag{13}$$

where Γ^k_F and Γ^k_ξ are the lag-*k* covariance matrices of F_t and ξ_t, respectively. Thus, the covariance matrix of X_t is the sum of two parts, one arisen out of the common components, and the other arising from the idiosyncratic components.

Likewise, the frequency domain counterpart of the variance decomposition of the factor model was introduced by Geweke (1977), in which the spectral density matrix of X_t at frequency θ can be decomposed into spectral densities of the common and idiosyncratic components:

$$\Sigma(\theta) = \Sigma_\chi(\theta) + \Sigma_\xi(\theta) = \lambda(e^{-i\omega})\Sigma_f(\theta)\lambda(e^{-i\omega})' + \Sigma_\xi(\theta)\ ,\ \ \theta \in \left[-\pi, \pi\right] \tag{14}$$

where $\Sigma_f(\theta)$ and $\Sigma_\xi(\theta)$ are the spectral density matrices of f_t and ξ_t at frequency θ, respectively.

In Forni, Lippi, and Reichlin (2005), dynamic PCA can be consider as a generalized orthogonal transformation process, which involves estimating eigenvalues and corresponding eigenvectors of spectral density of X_t. Generally, the spectral density is defined as a positive function of a frequency variable associated with the stationary stochastic process, and it is generated to simplify the information in a representation of frequency domain rather than time domain. Speaking of frequency domain, a given time series can be converted between the time and frequency domain with a pair of transform, and the spectral density can be obtained by applying a Fourier transform to covariance of data. Therefore, the spectral density is defined as

$$\Sigma(\theta) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \Gamma^k e^{-i\theta k} \tag{15}$$

where *i* is the imaginary unit, and $\theta = 2\pi h/H$ with the ordinary frequency h/H.

In the first step, using dynamic techniques of Forni, Hallin, Lippi and Reichlin (2000), an estimate of the lag-*k* covariance of common components, $\hat{\Gamma}_\chi^k$, is based on the frequency domain PCs. Start with estimating the auto-covariance of X_t, $\hat{\Gamma}^k$, and then calculate the spectral density of X_t by using a Bartlett window and applying the discrete Fourier transform,

$$\hat{\Sigma}(\theta) = \frac{1}{2\pi} \sum_{k=-M}^{M} w_k \hat{\Gamma}^k e^{-i\theta k} \tag{16}$$

where the auto-covariance matrix estimate $\hat{\Gamma}^k = \frac{1}{T-k-1}\sum_{t=k}^{T} X_t X_t'$, and $w_k = 1 - \frac{|k|}{M+1}$ as a Bartlett lag-window estimate weights with the window size $M = [T^{1/2}]$. In practice, the window function is usually used to weight the covariances, which can be considered averaging a given number M of covariances. The spectra are evaluated at 101 equally spaced frequencies in the interval $[-\pi, \pi]$, namely at a grid of frequencies $\theta_h = \frac{2\pi h}{100}$, $h = -50, \ldots, 50$. For each given frequency, the eigenvalues and eigenvectors of $\hat{\Sigma}(\theta)$ are computed, and then the spectral density matrix of common components driven by *q* dynamic factors is estimated

$$\hat{\Sigma}_\chi(\theta) = V_q(\theta) D_q(\theta) V_q(\theta)' \tag{17}$$

where $D_q(\theta)$ is a diagonal matrix with the diagonal of the first *q* eigenvalues of $\hat{\Sigma}(\theta)$ and $V_q(\theta)$ is a $n \times q$ matrix of corresponding eigenvectors. Meanwhile, the spectral matrix of idiosyncratic components can be seen as $\hat{\Sigma}_\xi(\theta) = \hat{\Sigma}(\theta) - \hat{\Sigma}_\chi(\theta)$. Therefore, the covariance matrices of common and idiosyncratic components are obtained using the inverse discrete Fourier transform

$$\hat{\Gamma}_\chi^k = \frac{2\pi}{101}\sum_{h=-50}^{50} \hat{\Sigma}_\chi(\theta_h) e^{i\theta_h k} \tag{18}$$

$$\hat{\Gamma}_\xi^k = \frac{2\pi}{101}\sum_{h=-50}^{50} \hat{\Sigma}_\xi(\theta_h) e^{i\theta_h k} \tag{19}$$

However, in order to calculate $\hat{\Gamma}_{\xi}^{0}$, some spurious covariance estimates are unavoidable. Then, consider $\hat{\Gamma}_{\xi}^{0}$ as an exception, which is estimated by $\hat{\Gamma}_{\xi}^{0} = \hat{\Gamma} - \hat{\Gamma}_{\chi}^{0}$ and off-diagonal elements of matrix are replaced by zeros. Under this estimate, the forecasting performance can be significantly improved, even when the actual matrix is not diagonal (Forni, Hallin, Lippi, and Reichlin (2005)).

In the second step, take the *r* generalized PCs of $\hat{\Gamma}_{\chi}^{0}$ with respect to $\hat{\Gamma}_{\xi}^{0}$, which compute the generalized eignvalues, μ_j, of the couple matrices ($\hat{\Gamma}_{\chi}^{0}$, $\hat{\Gamma}_{\xi}^{0}$) along with the corresponding generalized eignvectors, Z_j, that is solution

$$Z_j \hat{\Gamma}_{\chi}^{0} = \mu_j Z_j \hat{\Gamma}_{\xi}^{0}, \qquad j = 1, 2, \ldots, n \tag{20}$$

with the normalizing constraints $Z_j \hat{\Gamma}_{\xi}^{0} Z_j{}' = 1$ and $Z_i \hat{\Gamma}_{\xi}^{0} Z_j{}' = 0$ for $i \neq j$. Ordering the eigenvalues μ_j in descending order and taking the eignvectors $\hat{Z} = (\hat{Z}_1, \ldots \hat{Z}_r)'$ corresponding to the *r* largest values, the linear combinations, $\hat{F}_t = \hat{Z} X_t$, are the dynamic PC estimations of ***X*** relative to the $\hat{\Gamma}_{\chi}^{0} \left(\hat{\Gamma}_{\xi}^{0} \right)^{-1}$, or $\hat{\Gamma}_{\chi}^{0} \big/ \hat{\Gamma}_{\xi}^{0}$, that is the linear combinations of the ***X***'s having the small idiosyncratic-common variance ratio.

One of advantages of dynamic PCs is that information of cross covariance among all the predictors are used, both lagged and contemporaneous, because estimates in the first step are based on the frequency domain. Furthermore, the second step involves the preliminary estimation of covariance of common and idiosyncratic components, as does the generalized PC estimates.

Targeted Principal Component Estimation

Hawkins (1973) suggest the idea to obtain proper PCs by considering the predictor variables and forecasted variable, Y_t, together, rather than the set of predictors alone. Actually, in order to improve the forecasting performance, it might not be sufficient to use just several PCs that account for most of the variance in a set of many observed predictors. Banerjee, Marcellino and Masten (2003) suggest one factor, which does not have the highest explanatory power for predictors, could be important for forecasting (a related fact is the first PC estimation of U.S data is systematically deleted in model selections and the sixth PC performs well in prediction). Therefore, we here propose a method reducing the influence of PCs uninformative for predicting Y_t. Consider all PCs based on available predictors, which are estimated by the four kinds of estimation methods discussed above. For each method, the estimated PCs are ordered by the importance of predicting, Y_t^h, and the first *r* PCs having most predictive power are obtained as the targeted PCs. The correlation values between Y_t and PCs, or equivalently the individual *t*-values of PCs in the regressions, are used as criteria for ordering. Thus, those PCs with small variance, which definitely are deleted in the classic PCA, may have relatively high correlation to Y_t, and then they can be retained as targeted PC estimation of factors.

2.3.3 Determination of the numbers of factors

Here we address some criteria to determine the numbers of factors. The widely used one is based on PCA, where the eigenvalues of the sample correlation, or covariance, matrix may indicate the number of common factors driving the dataset. The value of *r* should be chosen in order to explain a large fraction of total variance; however, there is no generally accepted limit for how much explained variance could suggest a sufficient fit. In the macroeconomic applications, a variance ratio of 40% is usually considered as a reasonable fit (Stock and Watson (2002)). Besides, there are also several available criteria based on statistical tests, but with the restriction for the numbers of variables. An alternative procedure is recently suggested by Onatski (2007), which is based on the few largest eigenvalues of covariance matrix of a complex-valued sample derived from the original dataset, which asymptotically distribute as a Tracy-Widom.

Traditional information criteria, AIC and BIC, are only designed for small n. With selection problems of number of factors for both large n and T, Bai and Ng (2002) propose the information criteria in six forms, which modifies AIC and BIC by the penalty function, $p(n,T)$, in order to deal with overparameterization. Let $V_r = \frac{1}{nT}\sum_{t=1}^{T}(X_t - \hat{\Lambda}\hat{F}_t)'(X_t - \hat{\Lambda}\hat{F}_t)$ be the variance of the estimated idiosyncratic terms from the *r*-factor model. And the criterion is aimed at minimizing the residual variance, V_r. Bai and Ng (2002) suggest two classes of criterion. One is the panel criteria, written as $P_{anel}C$, and the other is the panel information criteria, named as $P_{anel}IC$. Both of them are explained in the following, with three different choice of penalty function respectively

$$P_{anel}C_1(r) = V_r + r\sigma^2(\frac{n+T}{nT})\ln(\frac{nT}{n+T}); \quad (21)$$

$$P_{anel}C_2(r) = V_r + r\sigma^2(\frac{n+T}{nT})\ln(\min\{n,T\}); \quad (22)$$

$$P_{anel}C_3(r) = V_r + r\sigma^2 \frac{\ln(\min\{n,T\})}{\min\{n,T\}}; \quad (23)$$

$$P_{anel}IC_1(r) = \ln(V_r) + r(\frac{n+T}{nT})\ln(\frac{nT}{n+T}); \quad (24)$$

$$P_{anel}IC_2(r) = \ln(V_r) + r(\frac{n+T}{nT})\ln(\min\{n,T\}); \quad (25)$$

$$P_{anel}IC_3(r) = \ln(V_r) + r\frac{\ln(\min\{n,T\})}{\min\{n,T\}}. \quad (26)$$

And, the estimated number of factors, $\hat{r}$, is obtained from minimizing these criterion in the range of all possible value of r , that is $r = 0,1,\ldots,r_{\max}$,

$$\hat{r} = \underset{0\le r\le r_{\max}}{\arg\min}\, P_{anel}C_a(r), \quad (27)$$

or,

$$\hat{r} = \underset{0\le r\le r_{\max}}{\arg\min}\, P_{anel}IC_a(r) \quad \text{with a=1,2,3.} \quad (28)$$

where $r_{\max} = \min(n,T)$ is the pre-specified upper bond for the number of factors. The minimums of these criteria yield the consistent estimates of *r* (Bai and Ng (2002), and Alessi, Barigozzi and Capasso (2008)). Then, $\hat{r}$ numbers of common factors are estimated into the model (8), and forecast estimates using these common factors, based on the model (10), are examined to compare the performances.

2.3.4 Other estimations

Like PCA, partial least squares (PLS) also provide linear combinations of a reduced number of factors, however, it consider the correlation information between predictors and variables for forecasting. PLS and PCA are usually compared in theoretical and empirical aspects, and with other variables selection methods, such as Frank and Friedman (1993). The results normally suggest that PLS has fewer components than PCs, however, they perform similarly in empirical applications, and PCs are able to provide more stable forecasts. As another well-know method, ridge regression can be seen as a shrinkage by imposing a penalty on regression coefficients size, that is the sum-of-squares of parameters. Besides, many methods are widely applied in application of macroeconomic forecasting, for instance Lasso, pretest and information criterion methods and so forth.

3. Empirical Experimental Design

3.1 Data Description

The complete dataset consists of 132 monthly U.S. macroeconomic series, and spans 1960:1 to 2003:12 for a total 528 observations.[1] All 132 series are dealt with by three preliminary steps, as in Stock and Watson (2006). First, the series are transformed to be stationary, whether taking logarithms and/or differencing are made judgmentally by inspection of the data and unit root tests. Generally, the first differences are used for nominal interest rates, the first differences of

[1] Data sources from Mark Watson's website : http://www.princeton.edu/~mwatson

logarithms (growth rates) are used for real activity variables, and second differences of logarithms (changes in inflation) are used for price variables, which are detailed in Table A.1 of the Appendix. As the second step, the transformed series are screened for outliers. When deviations between observations and absolute median exceed six times the inter quartile range, these observations are replaced with the median value of the preceding five observations. Finally, all series are further standardized to have zero mean and unit variance.

3.2 Simulated Forecasting

Two of all the predictors are used to construct the forecasted variables. One is monthly IP total index, which is integrated of order one, $I(1)$, in logarithms. And the *h*-month growth rate, in percentage points at an annual rate, is explained as

$$Y_{t+h}^{h} = (1200/h)\ln(IP_{t+h}/IP_{t}), \quad (29)$$

and $Y_t = \ln(IP_t/IP_{t-1})$

where the factor 1200/*h* transforms monthly decimal growth into annual percentage growth.

The other is monthly CPI for all items, is integrated of order two, $I(2)$, in logarithms as following

$$Y_{t+h}^{h} = (1200/h)\ln(CPI_{t+h}/CPI_{t}) - 1200\ln(CPI_{t}/CPI_{t-1}), \quad (30)$$

and $Y_t = \Delta\ln(CPI_t/CPI_{t-1})$.

Out-of-sample forecasts are computed for each of two variables with all methods mentioned at horizons h = 1, 6, 12 months, where the forecasting period starts at 1970:1 and end at 2003:12. The forecasting procedure involves fully recursive factor estimation, parameter estimation, model selection, and so forth. For instance, in DFMs, the first simulated out-of-sample forecast is made at 1970:1. Factors are estimated using data for t = 1960:3,…, 1970:1 based on model (8), and coefficient parameters of models (9) are estimated for sample period t = 1960:3,…, 1970:1-h. Then, as defined in model (10), values of regressors at t = 1970:1, and coefficient estimates together are used to obtain the estimated forecast at 1970:1, $\hat{Y}^h_{1970:1}$. For the next forecast estimate, $\hat{Y}^h_{1970:2}$, factors are re-estimated using the full previous data, t = 1960:3,…, 1970:2, and parameters of models are obtained for sample covering t = 1960:3,…, 1970:2-h, which is called recursively estimated. Finally, to estimate final forecast, $\hat{Y}^h_{2003:12}$, factors are estimated using all data available from 1960:3 to 2003:12 and coefficients of models are estimated using data t = 1960:3,…, 2003:12–h. Especially, because latent factors are selected and estimated based on different sample periods, it is not necessary to restrict these factors to be the same for every time period. Furthermore, in targeted PC estimates, factors having the predictive power during this period might not be necessarily important for prediction in the next time period.

4. Empirical Results

The benchmark is univariate forecast based on the model (4), and usually AR lags order p is selected recursively by AIC and BIC with $1 \leq p \leq 12$ for the monthly data, or pre-specified values $p = 4$ and $p = 12$. In our case, the result of AR model selection is four lags based on the full data, and we use this for models where the lagged variables are involved. Then, fix four lags for both ***X*** and ***Y*** in ADL models, as well as the DFM if lagged ***Y*** variables are necessary to be included.

By doing a standard PCA, six PCs can explain over 40% variance of all predictors, where 40% is a reasonable fit for macroeconomic application (Stock and Watson (2002a)). On the other hand, based on the sample errors of PCA with given values $r_{\max} = 10, 20, 30$, six forms of criteria for determining number of factors can be calculated, and the results confirm six factors. The estimate of the number of factors seems to be relatively larger than results of previous literature and our following results when forecasting as well. We think the space of factors appears to be larger for the purpose of modeling. For the purpose of forecasting, our results of forecast error variance decompositions suggest that relatively small number of factors may be sufficient, which is consistent with Stock and Watson (1999, 2002a).

Now, we turn to compare the performance of different factor estimations. The RMSE of 12-month ahead forecasts of IP are summarized in Table 1, using four kinds of factor estimates, where the best result is in bold type. The results confirm the conclusion of Stock and Watson (2002a), where the factor-based forecast can substantially improve the benchmark forecast, and the two factors are able to capture most of forecasting improvement. Additionally, involving lagged IP variables does not provide better forecasts; in contrast, the forecasts only with factor estimates are more accurate, which implies IP forecasts could be estimated by only two factors.

The simple mean averaging of individual ADL models improves a little upon the AR benchmark, but the simple combining forecasts do not provide more accuracy. In another hand, these factor estimations share the similar performance of forecasting. Comparing the minima of each kind of PCs, generalized and dynamic PCs perform better than standard PCs, which is consistent with theoretical arguments.

Table 1 RMSEs of 12-month ahead Forecasts of U.S. Industrial Production Growth with r PC estimates for various methods

Factors	PC, AR	PC	WPC,AR	WPC	GPC,AR	GPC	DPC,AR	DPC
1	0.9833	0.9592	0.9841	0.9608	0.9808	0.9520	0.9814	0.9612
2	0.6254	0.6114	0.6494	0.6262	0.6584	0.6064	0.6202	0.6110
3	0.6173	0.6250	0.6313	0.6267	0.6453	0.6173	0.6180	0.6260
4	0.6142	0.6162	0.6320	0.6267	0.6458	0.6178	0.6148	0.6266
5	0.6283	0.6420	0.6285	0.6388	0.6583	0.6437	0.6379	0.6602
6	0.6700	0.6301	0.6709	0.6320	0.7013	0.6528	0.6680	0.6326

Combined ADL with mean, median, trimmed mean : 0.9478, 1.0432, 1.0264

Benchmark AR (4) MSE : 1.0314

Note: Entries are relative MSE, relative to the MSE of benchmark AR (4) that is given at the bottom, and the values less than one indicates an improvement over than benchmark. Also the RMSE of combining forecast base on ADL models are presented with three simple weighting methods. (PC, AR) indicates the forecasts are constructed by PC estimates and 4 lagged variables, meanwhile, (PC) presents the forecasts with only PC estimates. Other methods are labeled similarly.

Then, forecasts using targeted PC estimations are listed in Table 2. As same as those "original" PCs, targeted PC estimates make substantial improvement over benchmark, and estimates perform similarly among methods. In contrast to non-targeted PCs, targeted PCs with and lagged variables together perform better than only PCs themselves. Targeted PCs make little improvement to the non-targeted PCs; however, they could provide a reduced number of factors. The results suggest using lagged variables and the first targeted PCs, which is the PC most related with the forecasted variable or having most predictive power, could achieve improving forecasts as much as those with two PCs .

Table 2 RMSEs of 12-month ahead Forecasts of U.S. Industrial Production Growth with r targeted PC estimates for various methods

Factors	t-PC, AR	t-PC	t-WPC, AR	t-WPC	t-GPC, AR	t-GPC	t-DPC, AR	t-DPC
1	0.6524	0.7271	0.6662	0.7250	0.6117	0.6732	0.6484	0.7117
2	0.6117	0.7876	0.6271	0.8044	0.6585	0.7365	0.6146	0.7474
3	0.6013	0.8318	0.6357	0.8026	0.6762	0.8377	0.5906	0.7745
4	0.6329	0.8253	0.6559	0.7737	0.7063	0.9826	0.6174	0.7912
5	0.6181	0.8881	0.6233	0.8518	0.7221	1.0061	0.6092	0.8343
6	0.6148	0.8989	0.6293	0.9699	0.7203	1.0350	0.6124	0.8965

Table 3 summarizes the pseudo out-of-sample forecasting performance of another variable, CPI, for various methods at 12-month forecast horizons. The forecast with only PC estimates are worse than the benchmark, and others taking account to lagged CPI variables and PCs together are much better. Besides, one factor might be able to make the improvement of forecast. Thus, unlike IP, one factor estimation and lagged variables are used to obtain CPI forecasts, which the results are consistent with Stock and Watson (2002a).

Table 3 RMSEs of 12-month ahead Forecasts of U.S. Customer Price Inflation with r PC estimates for various methods

Factors	PC, AR	PC	WPC,AR	WPC	GPC,AR	GPC	DPC,AR	DPC
1	0.8519	1.2931	0.5657	1.2949	0.8534	1.2967	0.8569	1.2968
2	0.8435	1.2442	0.5845	1.2285	0.8565	1.2682	0.8540	1.2548
3	0.8639	1.2288	0.5967	1.2476	0.8731	1.2434	0.8673	1.2290
4	0.8409	1.1838	0.5808	1.0888	0.8535	1.1995	0.8399	1.1783
5	0.8176	1.0251	0.5812	1.0235	0.8354	1.0164	0.8176	1.0292
6	0.7919	1.0316	0.5619	1.0384	0.8104	1.0228	0.7891	1.0332

Combined ADL with mean, median, trimmed mean : 1.4678, 1.4771 ,1.4759

Benchmark AR (4) MSE : 0.8416

Interestingly, there is the evidence that some benefit might be obtained from estimating factors using weighted PCs rather than other three. And it still happens when applying targeted weighted PCs, as in Table 4. Unfortunately, compared to non-targeted PCs, targeted PC estimations can neither make improvement, even worse in the case of forecasting with only factors, nor reduce the number of factor estimates.

Table 4 RMSEs of 12-month ahead Forecasts of U.S. Customer Price Inflation with r targeted PC estimates for various methods

Factors	t-PC, AR	t-PC	t-WPC, AR	t-WPC	t-GPC, AR	t-GPC	t-DPC, AR	t-DPC
1	0.9744	2.8217	0.7540	3.1505	0.9231	2.7176	0.9472	2.7657
2	0.8285	3.0425	0.6437	3.4515	0.8679	2.8635	0.8327	2.9750
3	0.8097	3.0969	0.5745	3.4364	0.8188	2.9533	0.8057	3.0312
4	0.7846	3.2502	0.5862	3.5220	0.8088	3.0540	0.7825	3.1190
5	0.7620	3.3026	0.5789	3.5168	0.8216	3.1654	0.7727	3.1343
6	0.7488	3.3734	0.5614	3.5340	0.8427	3.3687	0.7436	3.1924

Table 5 The Rank of PCs (12-months)

	IP				CPI			
Rank	PCs #	WPCs #	GPCs #	DPCs #	PCs #	WPCs #	GPCs #	DPCs #
1	2	2	2	2	4	4	4	4
2	132	132	122	132	3	12	12	3
3	1	1	118	1	12	2	129	12
4	5	5	1	5	1	1	3	1
5	8	8	108	8	2	5	128	2
6	6	23	5	6	5	3	122	5
7	128	10	127	45	21	8	130	129
8	124	7	82	10	7	81	1	21
9	45	4	115	124	8	55	123	7
10	10	126	70	13	129	128	132	8

As proposed in our paper, for estimating targeted PCs, all “original” PCs are ordered according to the correlation with forecasted variables. Table 5 lists the top ten PCs on the fully sample for two forecasted variables. The first six targeted PCs are selected at the beginning of the model selection. As we seen, some PCs with the relative small variance are actually having predictive power, such as the 23th, 122th, and 132th PCs. Using these PCs as the targeted PCs, it is reasonable to understand why targeted PCs perform better with lagged variables One reason might be to cover the possible losses of the information from predictors.

Based on the previous results, the targeted PC estimates provide the similar accuracy forecast, as do non-targeted PCs. Targeted PCs with lagged variables together perform better, and they are able to use a relatively small number of factors to forecast, like forecasting IP. However, targeted PCs do not have much advantage when most of information of both predictors and forecasted variable has been already summarized, for instance CPI forecasts already can be constructed by only one factor and lagged variables. Besides, when one kind of PC estimates performs better than other PCs, the same kind of targeted PCs usually does better than other targeted PCs as well. Therefore, the performance of targeted PC estimates depends on the forecasted variable of interests, as well as the corresponding PC estimations certainly.

At forecasting horizon of 6 months, generally speaking, the IP forecasts based on the PCs could improve approximately 30% accuracy to the benchmark, as shown in Table 7 and 8, meanwhile 12-month ahead forecasts reduce the benchmark by 35% to 40%. Besides, there is no significantly surprising finding comparing with the results at horizon of 12 months, thus more detailed results, CPI forecasts and rank of targeted PCs, are outlined in Table A.2, A.3, and A.4 of the Appendix.

Table 7 RMSEs of 6-month ahead Forecasts of U.S. Industrial Production Growth with r PC estimates for various methods

Factors	PC, AR	PC	WPC,AR	WPC	GPC,AR	GPC	DPC,AR	DPC
1	0.9168	0.8938	0.9176	0.8952	0.9175	0.8865	0.9194	0.8977
2	0.6722	0.6790	0.6834	0.6843	0.7124	0.6825	0.6730	0.6890
3	0.6693	0.6563	0.6710	0.6704	0.7193	0.6702	0.6690	0.6582
4	0.6739	0.6664	0.6743	0.6654	0.7261	0.6966	0.6733	0.6681
5	0.7030	0.7210	0.6920	0.7067	0.7400	0.7413	0.7089	0.7289
6	0.7986	0.7611	0.7864	0.7521	0.8258	0.7809	0.7953	0.7595

Combined ADL with mean, median, trimmed mean : 0.9872,1.0572,1.0445

Benchmark AR (4) MSE : 0.9015

Table 8 RMSEs of 6-month ahead Forecasts of U.S. Industrial Production Growth with r targeted PC estimates for various methods

Factors	t-PC, AR	t-PC	t-WPC, AR	t-WPC	t-GPC, AR	t-GPC	t-DPC, AR	t-DPC
1	0.8161	0.8500	0.7877	0.7734	0.8212	0.8449	0.7914	0.8289
2	0.7114	0.8389	0.7595	0.9400	0.7993	0.8876	0.7639	0.7898
3	0.7350	0.9417	0.7098	0.9877	0.7635	0.9766	0.7149	0.8758
4	0.7144	1.0525	0.6940	1.0799	0.7495	1.0804	0.7127	0.9509
5	0.7053	1.0525	0.6834	1.0822	0.7485	1.2659	0.7282	1.1138
6	0.6953	1.0919	0.6836	1.1128	0.7152	1.3188	0.7256	1.1827

5. Conclusion and Discussion

The empirical work in our paper has four main results. At first, most forecasts based on the DFM improve benchmark forecast, and most performance of factor estimations employed seem similar. Secondly, a small number of factors can be needed to forecast IP, which generally can be extended into real activity variables (Stock and Watson (2002a)), and furthermore a reduced number of targeted PC estimations can construct accurate forecasts as well. Thirdly, CPI variable, or saying inflation variables as in Stock and Watson (2002a), can be estimated by its lagged variables and only one factor; meanwhile, targeted PC estimates do not improve the performance as much as expected. At last, among targeted PCs, those which are associated with the most effective PC estimations are able to provide most accuracy forecasts, and the extent of forecasting improvement depends on the variables of study.

There are several issues to future. At first, our forecasting procedure uses the recursive sample, what if rolling sample? Stock and Watson (2005) suggest the recursively estimated models generally outperform rolling estimates. Therefore, we adopt recursive sample in the paper, which each forecast is constructed using dataset from 1970:1 to the point *T* of forecasting. In contrast, the rolling sample usually covers recent ten years till time *T*. Actually, the relationship between PCs and forecasted variables are involved, which varies over time, and the rolling sample might be considered to achieve effectiveness of forecasts. Next, when PCs are ordered, most of the correlation values are low, generally smaller than 0.01, thus a pretest or information criterion method could be introduced to make sure PCs with both predictive and explanatory power so that improve forecasting accuracy. Thirdly, for two variables forecasted, we use the exactly same dataset to obtain PCs, which also happens in most previous empirical analysis. However, this may raise some problems. One possible solution to improve the accuracy is the sub dataset only include variables which can explain most variance or have predictive power with respect to the variable of interest. The fourth one is that PCs are linear combinations, and forecasts are linear as well, which make the nonlinear problems open to discuss. Fourthly, BMA is definitely one of hottest issues in recent research, which we do not discuss more. In application of model selection, there are still general problems when dealing with large number of models based on many variables in Bayesian frame. However, some studies suggested using orthogonal regressors to simplify the computation problems, which bring into the new sight. Finally, we believe that performance of factor-based methods depend on macroeconomic variables of study, for example real activity or inflation, moreover, it is supposed to be related to areas and countries of interest, the underlying dataset or sub-sample, forecasting horizons, aimed short or long-term, etc. The assessment of all determinants of the forecasting performance need to be further researched.

References

Alessi, L., M. Barigozzi, and M. Capasso (2008), “A robust criterion for determing the number of static factors in approximated factor models”, Wording Paper Series No.903 (European Central Bank).

Almoy, T. (1996), “A simulation study on comparison of prediction methods when only a few components are relevant”, Computational Statistics and Data Analysis 21: 87-107.

Anderson, T.W. (1984), An Introduction to Multivariate Statistical Analysis, second edition (Wiley, New York).

Bai, J., and S. Ng (2002), “Determining the Number of Factors in Approximate Factor Models”, Econometrica, Vol. 70, No. 1 : 191-221.

Bai, J., S. Ng (2006), “Confidence intervals for diffusion index forecasts and inference with factor-augmented regressions”, Econometrica 74 (4): 1133-1150

Bai, J., S. Ng (2008), “Forecasting economic time series using targeted preditors”, Journal of Economics 146: 304-317

Banerjee, A., M. Marcellino, and I. Masten (2003), “Leading Indicators for Euro-area Inflation and GDP Growth”, Working Paper No. 235.

Bernanke, B.S., and J. Boivin (2003), “Monetary policy in a data-rich environment”, Journal of Monetary Economics 50: 525-546

Bovin, J., and S. Ng (2003), “Are more data always better for factor analysis?”, Working Paper No. 9829 (NBER).

Bovin, J., and S. Ng (2005), “Understanding and Comparing Factor-Based Forecasts”, Working Paper No. 11285 (NBER).

Bovin, J., and S. Ng (2006), “Are more data always better for factor analysis”, Journal of Economics 132: 169-194.

Breiman, L. (1996), “Bagging predictors”, Machine Learning, 36: 105-139.

Burns, A.M., and W.C. Mitchell (1946), “Measuring Business Cycles”, National Bureau of Economic Research.

Chamberlain, G. and M. Rothschild (1983), “Arbitrage factor structure, and mean-variance analysis of large asset markets”, Econometrica 51: 1281-1304.

Copas, J.B. (1983) “Regression, Prediction and Shrinkage”, Journal of the Royal Statistical Society, Series B (Methodological) 45 (3): 311-354.

Engle, R.F., and M. W. Watson (1981) “A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates”, Journal of the American Statistical Association, Vol. 76, No. 376, 774-781.

Frank, I., and J. Friedman (1993), “A statistical view of some chemometrics regression tools”, Technometrics, 35: 109-135.

Forni, M., and M. Lippi (2001), “The Generalized Dynamic Model: Representation Theory”, Economic Theory 17: 1113-1141.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), “The generalized factor model: identification and estimation”, The Review of Economics and Statistics 82: 540-554.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2003), “Do financial variables help forecasting inflation and real activity in the Euro area?”, Journal of Monetary Economics 50: 1243-1255.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2004), “The generalized factor model: consistency and rates”, Journal of Economics 119: 231-255.

Forni, M., M. Hallin, M. Lippi and L. Reichlin (2005), “The generalized factor model: one-side estimation and forecasting”, Journal of the American Statistical Association, Vol. 100, No. 471, 830-840.

Geweke, J. (1977), “The Dynamic Factor Analysis of Economic Time Series”, in: D.J Aigner and A.S. Goldberger, eds., Latent Variables in Socio-Economic Models, (North-Holland, Amsterdam).

Granger, C.W.J., and R. Ramanathan (1984), “Improved methods of combining forecasting”, Journal of Forecasting 3: 197-204.

Hawkins, D.M. (1973), “On the Investigation of Alternative Regressions by Principal Component Analysis”, Journal of the Royal Statistical Society. Series C (Applied Statistics) Vol. 22, No. 3: 275-286.

Helland, I., and T. Almoy (1994), “Comparison of prediction methods when only a few components are relevant”, Journal of the American Statistical Association 89: 583-591.

Hoerl, A.E., and R.W. Kennard (1970), “Ridge regression: biased estimation for nonorthogonal problems”, Technometrics 8: 27-51.

Hwang, J.T., and D. Nettleton (2003), “Principal components regression with data-chosen components and related methods”, Technometrics 45: 70-79.

James, W., and Stein, C. (1960), “Estimation with quadratic loss”, Proccedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1: 361-379.

Jones, C.S. (2001), “Extracting factors from heteroskedastic asset returns”, Journal of Financial Economics 62: 293-325.

Leamer, E.E. (1978), Specification Searches, New York: Wiley.

Min, C., and A. Zellner (1993), “Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates”, Journal of Econometrics 56: 89-118.

Naes, T., and I. Helland (1993), “Relevant components in regression”, Scandinavian Journal of Statistics 20: 239-250.

Naik, P., and C. Tsai (2000), “Partial least squares estimator for single-index models”, Journal of Royal Statistical Society B 62: 763-771.

Onatski, A. (2007), “A formal statistical test for the number of factors in the approximate factor models”, Mimeo, Columbia University, Department of Economics.

Quah, D., and P. Poncela (2004), “A Dynamic Index Model for Large Cross Sections”, in: J.H. Stock and M.W. Watson, eds., Business Cycle, Indicators, and Forecasting (University of Chicago Press for the NBER, Chicago) CH. 7.

Sargent, T.J., and C.A. Sims (1977), “Business cycle modeling without pretending to have too much a priori economic theory”, in: C. Sims et al., eds., New Methods in Business Cycle Research (Federal Reserve Bank of Minneapolis, Minneapolis).

Stein, C. (1955), “Inadmissibility of the usual estimator for the mean of the multivariate normal distribution”, Proceeding of the Third Berkeley Syposium on Mathematical Statistics and Probability, 1: 197-206.

Stock, J.H., and M.W. Watson (1989), “New indexes of coincident and leading economic indicators”, NBER Macroeconomics Annual 351-393.

Stock, J.H., and M.W. Watson (1991), “A probability model of the coincident economic indictors”, in: G. Moore and K. Lahiri, eds., The leading Economic Indicators: New Approaches and Forecasting Records (Cambridge University Press, Cambridge) 63-90.

Stock, J.H., and M.W. Watson (1999), “Forecasting Inflation”, Journal of Monetary Economics 44: 293-335.

Stock, J.H., and M.W. Watson (2002a), “Macroeconomic forecasting using diffusion indexes”, Journal of Business and Economic Statistics 20: 147-162.

Stock, J.H., and M.W. Watson (2002b), “Forecasting using principal components form a large number of predictors”, Journal of the American Statistical Association 97: 1167-1179.

Stock, J.H., and M.W. Watson (2005), “An empirical comparison of methods for forecasting using many predictors”, manuscript.

Stock, J.H., and M.W. Watson (2006), “Forecasting with many predictors”, in: Elliott, Graham, Granger, Clive, Timmermann, Allan (Eds.), Handbook of Forecasting, 2006.

Stone, M., and R. Brooks (1990), “Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression”, Journal of Royal Statistical Society B 52: 237-269.

Appendix

Table A.1 Data

Series	Trans	Description
a0m052	△ln	Personal Income (AR, Bil. Chain 2000＄) (TCB)
a0m051	△ln	Personal Incom Less Transfer Payments (AR, Bil. Chain 2000$) (TCB)
a0m224_r	△ln	Real Consumption (AC) a0m224/gmdc (a0m224 in from TCB)
a0m057	△ln	Manufacturing And Trade Sales (Mil. Chain 1996 ＄) (TCB)
a0m059	△ln	Sales of Retail Stores (Mil. Chain 2000 ＄) (TCB)
ips10	△ln	Industrial Production Index – Total Index
ips11	△ln	Industrial Production Index – Products, Total
Ips299	△ln	Industrial Production Index – Final Products
ips12	△ln	Industrial Production Index – Consumer Goods
ips13	△ln	Industrial Production Index – Durable Consumer Goods
ips18	△ln	Industrial Production Index – Nondurable Consumer Goods
ips25	△ln	Industrial Production Index - Business Equipment
ips32	△ln	Industrial Production Index - Materials
ips34	△ln	Industrial Production Index – Durable Goods Materials
ips38	△ln	Industrial Production Index – Nondurable Goods Matertials
ips43	△ln	Industrial Production Index – Manufacturing (Sic)
ips307	△ln	Industrial Production Index – Residential Utilities
ips306	△ln	Industrial Production Index - Fuels
pmp	lv	Napm Production Index (Percent)

a0m082	△lv	Capacity Utilization (Mfg) (TCB)
lhel	△lv	Index of Help-Wanted Advertising In Newspaper (1967=100;Sa)
lhelx	△lv	Employment: Ratio; Help-Wanted Ads: No. Unemployed Clf
lhem	△ln	Civilian Labor Force: Employed, Total (Thous.,Sa)
lhnag	△ln	Civilian Labor Force: Employed, Nonagric. Indutries (Thous.,Sa)
lhur	△lv	Unemployment Rate: All Workers, 16 Years & Over (%,Sa)
lhu680	△lv	Unemploy.By Duration: Average (Mean) Duration In Weeks (Sa)
lhu5	△ln	Unemploy.By Duration: Personal Unempl.Less Than 5 Wks (Thous,Sa)
lhu14	△ln	Unemploy.By Duration: Personal Unempl.5 To 14 Wks (Thous,Sa)
lhu15	△ln	Unemploy.By Duration: Personal Unempl.15 Wks + (Thous,Sa)
lhu26	△ln	Unemploy.By Duration: Personal Unempl.15 To 26 Wks (Thous,Sa)
lhu27	△ln	Unemploy.By Duration: Personal Unempl.27 Wks + (Thous,Sa)
a0m005	△ln	Average Weekly Initial Claims, Unemploy. Insurance (Yhous.) (TCB)
ces002	△ln	Employees On Nonfarm Payrolls – Total Private
ces003	△ln	Employees On Nonfarm Payrolls – Good-producting
ces006	△ln	Employees On Nonfarm Payrolls – Mining
ces011	△ln	Employees On Nonfarm Payrolls – Construction
ces015	△ln	Employees On Nonfarm Payrolls – Manufacturing
ces017	△ln	Employees On Nonfarm Payrolls – Durable Goods
ces033	△ln	Employees On Nonfarm Payrolls – Nondurable Goods
ces046	△ln	Employees On Nonfarm Payrolls – Service-Providing
ces048	△ln	Employees On Nonfarm Payrolls – Trade, Transportation, and Utilities
ces049	△ln	Employees On Nonfarm Payrolls – Wholesale Trade

ces053	△ln	Employees On Nonfarm Payrolls – Retail Trade
ces088	△ln	Employees On Nonfarm Payrolls – Financial Activities
ces140	△ln	Employees On Nonfarm Payrolls – Government
a0m048	△ln	Employee Hours In Nonag. Establishments (AR, Bil. Hours) (TCB)
ces151	lv	Avg Weekly Hrs of Prod or Nonsup Workers On Private Nonfarm Payrolls-Goods-Producing
ces155	△lv	Avg Weekly Hrs of Prod or Nonsup Workers On Private Nonfarm Payrolls-Mfg Overtime Hours
aom001	lv	Average Weekly Hours, Mfg. (Hours) (TCB)
pmemp	lv	Napm Employment Index (Percent)
hsfr	ln	Housing Starts: Nonfarm (1947-58); Total Farm&Nonfarm (1959-) (Thous.,Saar)
hsne	ln	Housing Starts: Northeast (Thous.U.) S.A.
hsmw	ln	Housing Starts: Mideast(Thous.U.) S.A.
hssou	ln	Housing Starts: South(Thous.U.) S.A.
hswst	ln	Housing Starts: West (Thous.U.) S.A.
hsbr	ln	Housing Authorized: Total New Priv Housing Units (Thous.,Saar)
hsbne*	ln	Housing Authorized By Built. Permits: Northeast (Thous.U.)S.A.
hsbmw*	ln	Housing Authorized By Built. Permits: Mideast (Thous.U.)S.A.
hsbsou*	ln	Housing Authorized By Built. Permits: South (Thous.U.)S.A.
hsbwst*	ln	Housing Authorized By Built. Permits: West (Thous.U.)S.A.
pmi	lv	Purchasing Managers' Index (Sa)
pmno	Lv	Napm New Orders Index (Percent)
pmdel	lv	Napm Vendor Deliveries Index (Percent)
pmnv	Lv	Napm Inventories Index (Percent)

a0m008	$\triangle$ln	Mfrs' New Orders, Consumer Goods and Materials (Bil. Chain 1982 $) (TCB)
a0m007	$\triangle$ln	Mfrs' New Orders, Durable Goods Industries (Bil. Chain 2000 $) (TCB)
a0m027	$\triangle$ln	Mfrs' New Orders, Nondefense Capital Goods (Bil. Chain 1982 $) (TCB)
a0m092	$\triangle$ln	Mfrs' New Orders, Durable Goods Indus. (Bil. Chain 2000 $) (TCB)
a0m070	$\triangle$ln	Manufacturing And Trade Inventories (Bil. Chain 2000 $) (TCB)
a0m077	$\triangle$lv	Ratio, Mfg. and Trade Inventories To Sales (Based On Chain 2000 $) (TCB)
fm1	$\Delta^2 \ln$	Money Stock: M1 (Curr, Trav.Cks, Dem Dep, Other Ck'ale Dep)(Bil $,Sa)
fm2	$\Delta^2 \ln$	Money Stock: M2 (M1+O'nite Rps, Euro $, G/P&B/D Mmmfs&Sav&Sm Time Dep)(Bil $,Sa)
fm3	$\Delta^2 \ln$	Money Stock: M3 (M2+Lg Time Dep, Term Rp's&Inst Only Mmmfs) (Bil $,Sa)
fm2dq	$\triangle$ln	Money Supply – M2 In 1996 Dollars (Bci)
fmfba	$\Delta^2 \ln$	Monetary Base, Adj For Reserve Requirement Changes (Mil $,Sa)
fmrra	$\Delta^2 \ln$	Depository Inst Reserves: Total, Adj For Reserve Reg Chgs (Mil $,Sa)
fmrnba	$\Delta^2 \ln$	Depository Inst Reserves: Nonborrowed,Adj Res Reg Chgs (Mil $,Sa)
fclng	$\Delta^2 \ln$	Commercial & Industrial Loans Outstanding In 1996 Dollars (Bci)
fclbmc	lv	Weekly Rp Lg Com'l Banks: Net Change Com'l & Indus Loans (Bil $, Saar)
cinrv	$\Delta^2 \ln$	Consumer Credit Outstanding – Nonrevolving (G19)
a0m095	$\triangle$lv	Ratio, Consumer Installment Credit To Personal Income (Pct.)(TCB)
fspcom	$\triangle$ln	S&P's Common Stock Price Index: Composite (1941-43=10)
fspin	$\triangle$ln	S&P's Common Stock Price Index: Industries (1941-43=10)
fsdxp	$\triangle$lv	S&P's Composite Common Stock: Dividend Yield (% Per Annum)

fspxe	△ln	S&P's Composite Common Stock: Price-Earnings Ratio (%, Nsa)
fyff	△lv	Interest Rate: Federal Funds (Effective) (% Per Annum, Nsa)
cp90	△lv	Commercial Paper Rate (AC)
fygm3	△lv	Interest Rate: U.S. Treasury Bills, Sec Mkt, 3-Mo. (% Per Ann, Nsa)
fygm6	△lv	Interest Rate: U.S. Treasury Bills, Sec Mkt, 6-Mo. (% Per Ann, Nsa)
fygt1	△lv	Interest Rate: U.S. Treasury Const Maturities, 1-Yr. (% Per Ann, Nsa)
fygt5	△lv	Interest Rate: U.S. Treasury Const Maturities, 5-Yr. (% Per Ann, Nsa)
fygt10	△lv	Interest Rate: U.S. Treasury Const Maturities, 10-Yr. (% Per Ann, Nsa)
fyaaac	△lv	Bond Yield: Moody's Aaa Corporate (% Per Annum)
fybaac	△lv	Bond Yield: Moody's Baa Corporate (% Per Annum)
Scp90	lv	cp90-fyff
sfygm3	lv	fygm3-fyff
sfygm6	lv	fygm6-fyff
sfygt1	lv	fygt1-fyff
sfygt5	lv	fygt5=fyff
sfygt10	lv	fygt10-fyff
fyaaac	lv	fyaaac-fyff
fybaac	lv	fybaac-fyff
exrus	△ln	United States; Effective Exchange Rate (Merm)(Index No.)
exrsw	△ln	Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S. ＄)
exrjan	△ln	Foreign Exchange Rate: Japan (Yan Per U.S. ＄)
exryk	△ln	Foreign Exchange United Kingdom (Cents Per Pound)
exrcan	△ln	Foreign Exchange Rate: Canada (Canadian Per U.S. ＄)

pwfsa	$\Delta^2 \ln$	Producer Price Index: Finished Goods (82=100,Sa)
pwfcsa	$\Delta^2 \ln$	Producer Price Index: Finished Consumer Goods (82=100,Sa)
pwimsa	$\Delta^2 \ln$	Producer Price Index: Intermed Mat. Supplies & Components (82=100,Sa)
pwcmsa	$\Delta^2 \ln$	Producer Price Index: Crude Materials (82=100,Sa)
pwm99q	$\Delta^2 \ln$	Index of Sensitive Materials Prices (1990=100)(Bci-99a)
pmcp	lv	Napm Commodity Price Index (Percent)
punew	$\Delta^2 \ln$	Cpi-U: All Items (82-84=100,Sa)
pu83	$\Delta^2 \ln$	Cpi-U: Apparel & Upkeep (82-84=100,Sa)
pu84	$\Delta^2 \ln$	Cpi-U: Transportation (82-84=100,Sa)
pu85	$\Delta^2 \ln$	Cpi-U: Medical Care (82-84=100,Sa)
puc	$\Delta^2 \ln$	Cpi-U: Commodities (82-84=100,Sa)
pucd	$\Delta^2 \ln$	Cpi-U: Durables (82-84=100,Sa)
pus	$\Delta^2 \ln$	Cpi-U: Serivces (82-84=100,Sa)
puxf	$\Delta^2 \ln$	Cpi-U: All Items Less Food (82-84=100,Sa)
puxhs	$\Delta^2 \ln$	Cpi-U: All Items Less Shelter(82-84=100,Sa)
puxm	$\Delta^2 \ln$	Cpi-U: All Items Less Medical Care(82-84=100,Sa)
gmdc	$\Delta^2 \ln$	Pcs,Impl Pr Defl: Pce (1987=100)
gmdcd	$\Delta^2 \ln$	Pcs,Impl Pr Defl: Pce; Durables (1987=100)
gmdcn	$\Delta^2 \ln$	Pcs,Impl Pr Defl: Pce; Nondurables (1987=100)
gmdcs	$\Delta^2 \ln$	Pcs,Impl Pr Defl: Pce; Services (1987=100)
hhsntn	$\triangle$lv	U. Of Mich. Index Of Consumer Expectations (Bcd-83)

ces275	$\Delta^2 \ln$	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls – Goods-Producing
ces277	$\Delta^2 \ln$	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls – Construction
ses278	$\Delta^2 \ln$	Avg Hourly Earnings of Prod or Nonsup Workers On Private Nonfarm Payrolls – Manufacturing

Table A.2 RMSEs of 6-month ahead Forecasts of U.S. Customer Price Inflation with r PC estimates for various methods

Factors	PC, AR	PC	WPC,AR	WPC	GPC,AR	GPC	DPC,AR	DPC
1	0.9011	1.4146	0.6035	1.4158	0.9128	1.4176	0.9127	1.4161
2	0.9243	1.3820	0.6121	1.3724	0.9360	1.4003	0.9282	1.3871
3	0.9364	1.3656	0.6143	1.3789	0.9466	1.3800	0.9360	1.3644
4	0.9321	1.3023	0.6148	1.2096	0.9463	1.3114	0.9304	1.2966
5	0.9236	1.1458	0.6186	1.1453	0.9409	1.1340	0.9218	1.1482
6	0.9254	1.1633	0.6166	1.1662	0.9432	1.1533	0.9212	1.1634

Combined ADL with mean, median, trimmed mean : 1.5149, 1.5184, 1.5168

Benchmark AR (4) MSE : 0.7788

Table A.3 RMSEs of 6-month ahead Forecasts of U.S. Customer Price Inflation with r targeted PC estimates for various methods

Factors	t-PC, AR	t-PC	t-WPC, AR	t-WPC	t-GPC, AR	t-GPC	t-DPC, AR	t-DPC
1	0.9857	3.2077	0.7506	3.5340	0.9816	3.0741	0.9928	3.1216
2	0.8882	3.4304	0.6992	3.6251	0.9276	3.3217	0.9525	3.1783
3	0.8430	3.5902	0.6582	3.6711	0.8822	3.4657	0.9012	3.3327

4	0.8127	3.6211	0.6245	3.7136	0.8607	3.5619	0.8759	3.4682
5	0.8280	3.6178	0.5959	3.7475	0.8629	3.6175	0.8538	3.4915
6	0.8343	3.6879	0.5883	3.7784	0.8656	3.6862	0.8490	3.4821

Table A.4 Rank of four kinds of PCs (6-months)

	IP				CPI			
Rank	PCs #	WPCs #	GPCs #	DPCs #	PCs #	WPCs #	GPCs #	DPCs #
1	1	1	122	132	4	4	4	4
2	2	2	1	1	12	12	12	12
3	132	132	2	2	3	2	3	3
4	8	8	132	8	7	1	110	7
5	5	5	127	5	1	57	33	1
6	56	56	118	56	2	8	78	2
7	6	114	108	6	21	19	129	21
8	33	36	82	44	33	7	128	33
9	44	67	116	33	35	81	92	35
10	62	54	129	124	65	22	61	65