Forecasting Inflation: Disaggregating Aggregates To Have Better Forecasts

by

Yongok Choi

Dennis W. Jansen

Joon Y. Park

Abstract: We investigate methods for incorporating disaggregated components of the CPI to improve forecasts of the CPI inflation rate. We incorporate the CPI components directly in a dynamic regression model, as well as several methods of summarizing the information contained in the components. In particular, we investigate Hansen's model averaging, an application of principal components, and dynamic regression with factors chosen by either Bai's panel information criteria, the Kaiser criterion, the Scree Plot method, or a method that chooses the number of factors that explain a large fixed percent (here 96%) of the variation. We find that the dynamic regression model using BIC to select the number of factors works very well. We investigate the properties of our dynamic regression model with factors summarizing information in the CPI components, and find that it remains superior to the AR model over horizons from 1 to 6 months. Overall, we conclude that CPI inflation forecasts can be improved by using summaries of information contained in the CPI components.

1. Introduction

Many macro variables are essentially aggregates, and these aggregates are often from a cross sectional distribution, so that the data we observe is a time series from a cross sectional distribution. The Consumer Price Index, or CPI, is a measure of the cost of living, and is an aggregation of price data on individual components of the CPI. Changes in the CPI are often used as measures of the inflation rate by policy makers, by government entities, and by private sector agents.

There is a large literature devoted to forecasting inflation, and accurate inflation forecasts are important for policy makers, for planners, for financial firms, and many other agents in the economy. The idea that inflation can be modeled as a purely monetary phenomenon has led some to build models and test for cointegration between price level measures and measure of money, although this cointegration is often rejected. Stock and Watson (1999) show that including real variables in forecasting models of inflation will improve forecasts, and models based on the Phillips curve have been shown to improve forecasting power. Stock and Watson used a generalized Phillips curve including real measures other than unemployment and showed an improved forecasting power, especially when they use an index of real activity constructed using principal components analysis on 168 economic indicators.

More recently, several authors have investigated ideas of using disaggregated components to forecast aggregates. Hendry and Hubrich (2006) examine the theoretical support for two methods of using information in disaggregated components for forecasting aggregates. One method looks at including disaggregated information in the forecasting model for the aggregate by first forecasting the disaggregated components and then creating a CPI forecast by aggregating these component forecasts using the CPI relative importance weights that are used to construct the actual CPI. A second method is to forecast the aggregate using disaggregated components as explanatory variables in a model of the aggregate. Hendry and Hubrich compare these methods, and show that the second method, including disaggregated components in forecasts of the aggregate, should in theory outperform the first method in population, although they point out numerous reasons why population results may not hold in small samples, including structural change, model uncertainty, and parameter estimation uncertainty.

Hendry and Hubrich provide an application to forecasting European and U.S. inflation rates. They look at disaggregating U.S. inflation to four subcomponents, and find that for forecasts over the 1990s and into the present decade, adding disaggregates or factors derived from the disaggregated components does not provide large improvement in forecasting U.S. inflation.

Smith (2007) looks at the forecasting performance of disaggregates for forecasting the U.S. Personal Consumption Expenditure (PCE) Deflator. She finds evidence that a regression-based forecasting model in which lags of the 51 components of the PCE deflator form the explanatory variables performed far better out of sample than forecasts based on aggregates and forecasts based on various re-weighted components.

We look at forecasts of the CPI. Like Hendry and Hubrich, we investigate the potential benefits of disaggregates in forecasting U.S. inflation. Unlike them, we investigate a finer degree of disaggregation than they do in looking at four subcomponents. We also look more extensively at the benefits of using various factor methods to summarize information contained in a large number of disaggregates.[1] That is, we look at the ability of disaggregated components to help forecast the aggregate, but also at the usefulness of factor methods to summarize the large

[1] Boivin and Ng (2006) provide evidence that additional data is not always better for factor models, but our results indicate that our large data set generates superior forecasts relative to just using aggregates than Hendry and Hubrich's four subcomponents.

amount of information contained in the disaggregated components. Along the way we provide information on the performance of several alternative procedures for selecting the number of factors to include in a forecasting model, including the Kaiser method, the Scree Plot, and methods suggested by Bai (2003) and Bai and Ng (2002). We investigate a series of alternatives for dealing with a large number of disaggregated components, including Hansen's (2006) least squares forecast averaging.

The CPI is calculated in two stages. The first stage is to calculate the 8,018 basic indices for the 211 elementary items in 38 elementary areas. The second stage is to aggregate the 211 elementary items into the CPI. Thus, the CPI is essentially a weighted average index of 211 elementary items. The weights are determined by estimates of household expenditures collected in the Consumer Expenditure Survey. These weights change every two years. The prices are determined by monthly surveys of price data. Thus at each time point we have a distribution of the CPI components. In our case, we are interested in the inflation rate, so at each time point we have a distribution of inflation rates for each of the elementary items.

The CPI is a weighted average of elementary items that may have different characteristics. For example, it has long been claimed that food and energy are more volatile than other series. Various measures of so-called core inflation have been devised that remove the influence of volatile elements on the CPI. The CPI less food and energy is perhaps the oldest and most well known of these, but there are various trimmed mean and median based estimates of core inflation.[2]

Although we are interested in forecasting the mean inflation rate, we want to investigate how our forecasting power can be improved by using information from the distribution of elementary items. One way to use this information about the distribution is by adopting factor

[2] See Smith (2007) for a description of some of these methods for defining core inflation, and the references therein.

analysis.

Note that our methodology differs from that of Stock and Watson (1998, 1999). We use information about the distribution of elementary items of the CPI to generate forecasts of the mean inflation rate, while Stock and Watson considered forecasting inflation with factors constructed from a large number of macroeconomic series that included many variables other than components of the CPI.

2. Data and Modeling

We have monthly data for the period 1983:01 – 2006:12. We use the seasonally adjusted index, and for most of our analysis we use data on 134 elementary indices. As mentioned above, the CPI is a weighted average of 211 elementary indices. However, some of the indices are not available for the full period, and others are unpriced items. After eliminating such data we have 134 elementary indices that account for more than 88% of the CPI. These are listed in the Appendix, along with a description of the series and information on the weight each item gets in calculating the CPI.

In order to provide a better idea of the underlying structure of the CPI, Figure 1 graphs the distribution of the elementary indices for the CPI, and the distribution of the rates of change of these elementary items, for August 2006. The elementary items are weighted by the relative importance assigned to each item from the household expenditure survey. Clearly the elementary indices take values over a large range, and the distribution is not unimodal. The distribution of rates of change in the elementary indices is more compressed but not single-peaked.

In this paper we consider a set of forecasting models including a simple AR model of the aggregate inflation rate. An important model for our purpose is a factor model, which we use to derive factors for use in a number of our forecasting models.

A Factor Model

Let X_{it} be the observed data for the i^{th} cross-sectional data at time t, where $i = 1,...N$ and $t = 1,...,T$. Consider the following factor representation of data:

$$X_{it} = \lambda_i' F_t + e_{it} \tag{1}$$

where F_t is an (r x 1) vector of common factors, λ_i is a vector of factor loadings corresponding to F_t, and e_{it} is the idiosyncratic component of X_{it}. In our forecasting models we will derive factors from the matrix of subcomponents of CPI inflation, and use these factors to help forecast the aggregate inflation rate. In this sense the factors are summaries of the information contained in a potentially large number of subcomponents.

In our application we take the matrix of components X_{it} to be the inflation components weighted by the weights assigned to each component by the Bureau of Labor Statistics (BLS) in calculating the CPI.

Principal Component Method

A related method of representing the data is based on the Karhunen-Loeve expansion, the basis of Principal Components analysis. We can write our data X_{it} as the N x 1 vector $X_t =(X_{1t}, X_{2t}, ..., X_{Nt})'$. We then decompose X_t as:

$$X_t = \sum_{k=1}^{r} c_{kt} v_k \tag{2}$$

where c_{kt} is $<v_k,X_t>$ the k^{th} coordinate process and v_k is the corresponding eigenvector or principal component. We can use this decomposition as a method of aggregating forecasts. In particular, we will estimate the principal components and the coordinate process over our estimation period, say through period t. We then forecast the coordinate process for the following period, $t+1$, and use the forecasted coordinate process and the principal component v_k to generate a forecast of next period's X, X_{t+1}. In this way the weighted principal components will generate an inflation forecast.

Determining the Number of Factors

A critical issue is the choice of the number of factors, r, to be included in the factor model. There are various methods suggested in the literature, and we employ several alternatives in this study.

First, the Kaiser criterion can be employed. This procedure would retain only factors with eigenvalues greater than the average of all eigenvalues. Since a plot of the rank order distribution of eigenvalues typically is quite skewed, this method tends to give a number much less than half the maximum number of factors. In our case, the Kaiser criterion suggests using 9 factors.

A second procedure is based on the Scree Plot. The idea is to examine a rank-ordered plot of the eigenvalues, from highest to lowest, looking for the point where the plot of the eigenvalues appear to level off to the right of the plot. This is an 'eyeball' approach. We present the Scree Plot in Figure 2. To our eyes the Scree Plot seems to suggest an elbow in the plot at two, so we use two factors. (Note that if other eyes decided that nine factors should be chosen the Kaiser criterion and the Scree Plot procedure would give the same number of factors.)

A third procedure is to choose the number of factors needed to explain a set (high) percentage of the variation in the data. A representative threshold value is 96%, and we use that value to arrive at the number of factors to use. Figure 3 provides a graph used to determine the number of factors meeting this 96% rule.

The fourth procedure we look at is Bai's (2003) panel information criteria, which looks on selection of the number of factors as a model selection problem. This is a method similar to AIC or BIC but adapted to panel situation. We have a panel with a large cross section (*N=134*) and a large time dimension (*T=288*). To determine the number of factors, we can use Bai's panel information criterion, denoted IC_p, which is appropriate for situations where both N and T are large. Like other model selection criteria, the panel information criteria IC_p depends on a tradeoff between parsimony and goodness of fit.

$$IC_p(k) = \log(V(k, \hat{F}_k) + k\frac{N+T}{NT}\log(\frac{NT}{N+T}) \tag{3}$$

where

$$V(k, \hat{F}_k) = \min_{\Lambda}(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \lambda_i{}'\hat{F}_t^k)) \tag{4}$$

Here $IC_p(k)$ is the panel information criterion for k factors. We plot IC_p against the number of factors in Figure 4. In this case the panel information criterion picks 29 factors, the minimum of IC_p with respect to k.

Dynamic Regression Forecasting Model

We consider a number of forecasting models for using factors to summarize information in a large number of components. These include the following dynamic regression model:

$$y_t = \alpha + \sum_{k=1}^{k_1} \beta_k ' F_{t-k} + \sum_{l=1}^{l_1} \gamma_l y_{t-l} + e_t \qquad (5)$$

We do not directly observe F_t but we make use of the relationship between F_t and X_{it} from equation (1). We estimate F_t from equation (1), and denote it $\hat{F}_t$. Then we regress y_t on a constant, current and lagged $\hat{F}_t$, and lagged y_t, as specified in equation (5). Finally we forecast inflation, with the coefficients we estimate in equation (5), as:

$$\hat{y}_t = \hat{\alpha} + \sum_{k=1}^{k_1} \hat{\beta}_k ' \hat{F}_{t-k} + \sum_{l=1}^{l_1} \hat{\gamma}_l y_{t-l} \qquad (6)$$

3. Results

We check the validity of our model in several ways. First, we compare the forecasting power of our model with a set of alternatives. Second, we develop the bootstrap test for the significance of the common factors. Third, we conduct formal statistical tests of relative forecast performance.

Forecasting Evaluation

We conduct an out-of-sample forecasting exercise over the most recent 48 months. More specifically, using data from January 1983 to December 2002, we estimate the models, allowing us to calculate in-sample mean squared errors for each model. We then use the estimated model and information through December 2002 to generate a forecast of the inflation rate for January 2003. We compare this forecast to the actual inflation rate and calculate the forecast error for January 2003. We then update our estimation sample to end in January 2003, re-estimate the model coefficients, and forecast the inflation rate for February 2003. We continue iterating in this way, generating a series of 48 quasi out-of-sample forecasts and the associated forecast errors. We calculate the mean square error (MSE) of our forecast, for comparison to forecasts from other models.

We consider the following classes of forecasting models: a baseline AR model, models using weighted components, models based on aggregating forecasts of the principal components, and a dynamic regression models. We also look at some models using forecast averaging. For each class, we consider several specific versions.

Our baseline model is an AR(p) model of aggregate CPI inflation. In this model we chose lags by BIC, searching over 1 to 24 lags for the lag length that minimized BIC. Lag selection was done for each forecast and the value chosen varied over our 48 month forecasting period.

We look at two models that forecast inflation by building up forecasts of components. The first looks at eight large spending categories in the CPI, forecasting the inflation rate for each of these eight components separately, and then aggregating these forecasts by the weights used in constructing the CPI.[3] For each of these eight groups, the AR model was selected using BIC searching over lag lengths from 1 to 24. We also include a model that builds up an inflation forecast using 182 items, where we forecast the inflation rate for each of the 182 items separately and then aggregate these forecasts using the weights used in constructing the CPI.[4] The separate forecasts are each made from separate models specified using BIC searching over lag lengths from 1 to 12. Note that we use 182 elementary items for this forecasting procedure

[3] The eight aggregated components are: Food and Beverages; Housing; Apparel; Transportation; Medical Care; Recreation; Education and Communication; and Other Goods and Services. These eight groups are one hundred percent of the CPI.

[4] The CPI consists of 211 elementary items, but 29 are unpriced items that are included in the CPI as combinations of priced items. We adjust the weights to incorporate the actual underlying weights used in the CPI, those accounting for the manner in which unpriced items are included in the CPI.

but only 134 for the factor models. This is due to the problem that a large number of items were not in existence over our entire sample period. For the component AR forecasting model we can use shorter time series to estimate the forecasting model for components with shorter time series. For the factor model we calculate factors using more data over a longer time period, and it is more difficult to adjust forecasts from the factor model for data missing over part of our estimation period.

We look at a series of models based on the principal components. We decompose X_t into coordinate processes and eigenvectors (or principal components). Then X_t can be written as an exact sum of the entire set of weighted principal components, and if we choose a large enough K we can write X_t, to a close approximation, as the sum of a set of the largest K weighted principal components. We can express this as the Karhunen-Loeve expansion:

$$X_t = \sum_{k=1}^{134} \langle v_k, X_t \rangle v_k \approx \sum_{k=1}^{K} \langle v_k, X_t \rangle v_k . \qquad (7)$$

Here v_k is the k^{th} principal component, and $\langle v_k, X_t \rangle$ is the coordinate process. This approach to forecasting inflation is to forecast the sum of the weighted principal components using forecasted weights, i.e. by forecasting the coordinate process. Again the number of principal components will be critical, and we use the criteria mentioned above (Kaiser's criterion, the Scree Plot, Bai's panel information criterion, and the 96% rule). For all these models we forecast the coordinate process $\langle v_k, X_t \rangle$ in two ways. First, we use an *AR(p)* model, forecasting using own lags of $\langle v_k, X_t \rangle$. Second, we use a *VAR(p)* model, forecasting with lags of all the included parts $\langle v_k, X_t \rangle$ for $k = [1,K]$. In either case we choose the lag length p by BIC.

We consider a set of dynamic regression models. Recall that our dynamic regression model is given by:

$$y_t = \alpha + \sum_{k=1}^{k_1} \beta_k ' F_{t-k} + \sum_{l=1}^{l_1} \gamma_l y_{t-l} + e_t$$

Here we choose the number of factors (the dimension of F), the number of lags of those factors k_1, and the number of lags of aggregate inflation l_1. For our first model in this class we choose the number of factors and the lag of the factors and of aggregate inflation by BIC. We find the dimension of F is 1, and the lag specification is $k_1 = 2$, $l_1 = 13$. We hold the number of factors and the lag lengths constant over the forecasting period.

We check four other models in this class, and for these we fix the number of factors (at 2, 9, 13, or 29) and pick the lags by BIC. Recall these numbers of factors are chosen by the Kaiser method, the Scree Plot, and Bai's panel information criterion. For these models the number of factors are obviously fixed, and we also hold the lag specification fixed over the forecast period.

We estimate a dynamic regression model in which we add lagged components of the CPI, but not lagged factors. We look at two models. One consists of adding the 8 large subgroups that we used in the weighted component regression models. The difference here is that we do not forecast each subgroup and then aggregate the forecasts, but we use the lagged subgroup variables to forecast the aggregate directly. This model is suggested by Hendry and Hubrich (2006) and also by Smith (2007). We also look at a model forecasting with all 134 elementary items that we used to calculate factors. Smith reports that a forecasting model with all 51 lagged components of the PCE deflator does well in forecasting PCE inflation, so we include this specification to see how such a method performs for forecasting the CPI. In our case, because the 134 elements is an incomplete set of the elements that make up the CPI, we also include a lag of the aggregate inflation rate.[5]

We also look at four models that use Hansen's (2006) forecast averaging methodology with lags of aggregate inflation selected by AIC. These models each start with a specification of the number of included factors and a maximum lag length for these included factors. In each of these models we also include 14 lags of the aggregate CPI inflation rate as selected by AIC. The model averaging method estimates models with lags 0 through L and constructs a weighted average using Mallow's criterion. We specify the number of factors as 2, 9, 13, and 29, corresponding to our earlier discussion of factor selection. We specify different maximum lag lengths depending on the number of factors. For 2 factors, we average over lags 0 to 6. For 9 factors, we average over lags 0 to 4, for 13 factors 0 to 2, and for 29 factors 0 to 1.

We also look at the same four models using Hansen's forecast averaging methodology with lags of aggregate inflation selected by BIC. Here BIC chooses one lag of aggregate inflation.

We present the results of forecasting with these models in Table 1. Our baseline AR model is given first, and the MSE is 1.04 E-05. For each model we also present the ratio of the MSE of that model with the MSE of our baseline AR model. We can summarize the main result by noting that all of the factor models we investigate outperform the AR model over this forecasting period and the one-month-ahead forecasting horizon. Disaggregating seems to help, often substantially. The best model has a relative MSE of 78%. We were somewhat surprised to find that even the fixed-weight component AR models result in substantial reductions in the MSE relative to the AR model. This is especially impressive because these models involve forecasting eight or 182 components and aggregating their forecasts. Such a procedure involves estimating a large number of coefficients, and such procedures are know to sometimes result in poor

[5] We tried including the 134 elements with no lag of the aggregate inflation rate but the MSE was much worse.

performance in practice, but here it results in improved forecasts. The relative MSE is 83.5% for the weighted average forecast based on eight subcategories, and 87.3% for the forecast using 182 individual items.

The forecasting models based on weighted forecasts of principal components beat the AR model, with relative MSEs ranging from 84.1% to 95.3%. The best was the two-factor model with AR forecasts, and the model with AR forecasts of the weights worked better than those with VAR forecasts of the weights.

The dynamic regression models that included lagged factors performed well, with relative MSE's ranging from 77.7% to 88.3%. In fact, the overall second best forecasting performance was turned in by the dynamic regression model with factors picked by BIC. This model's MSE was 8.078 E-06, 77.7% of the MSE of the baseline AR model.

The models based on Hansen's model averaging performed well, with relative MSE's ranging from 77.7% to 85.2%. For models with AIC-determined lags of inflation, the model with 13 factors averaging over 0 to 2 lags performs best, with a relative MSE of 80.1%. For models with BIC-determined lags of inflation, the model with 13 factors performed best, with a relative MSE of 77.7%.

We note that our dynamic regression model with BIC-selected lags and using BIC to select the number of factors beat all the forecasting models with the exception that it essentially tied the forecast averaging model with the factors fixed at 13 and lags (AR and factors) selected by BIC. Absent knowledge of the correct number of factors, the dynamic regression model with BIC-selected factors and lags is an attractive forecasting model.

In contrast to the above results, a disappointing set of forecasting results was turned in by forecasting models involving dynamic regression directly on inflation subcomponents, either the set of eight large CPI groups or the set of 134 elementary items. For the model based on the eight subcomponents, lag selection by BIC resulted in the largest lag length being chosen, so we checked results for lag lengths one through six. The best performance occurred at lag length two, but it is important to note that this would not be the lag length chosen by a lag selection criterion over the estimation period. For the 134 elementary item regression, we specified lag length one because of the large number of explanatory variables. Our forecasts for this model were particularly bad, resulting in a MSE fifty percent higher than the baseline AR model. Note that these results stand in sharp contrast to Smith's (2007) results on the PCE, where she reports a dramatic improvement from including components in dynamic regression models of PCE inflation.

We also investigated our model's performance at various alternative forecasting horizons. We compared our preferred dynamic regression model to the baseline AR model. Using the same 48-month forecasting period, we investigated forecasting performance at horizons of two through six months ahead. For these different horizons we calculated forecasts based on our baseline AR model and our dynamic regression model with the number of factors selected by BIC, i.e., 1. Our multiple step ahead forecasting models were constructed using the chain rule of forecasting, building from the one-step-ahead forecasts. For our dynamic regression model this required forecasting the (single) factor, and our factor forecasting model was an AR model with lag length 12 as determined by AIC. (Using BIC resulted in a lag length of three and almost identical inflation forecasts.)

We present these results in Table 2. Note that the dynamic regression model has a lower MSE at all horizons one through six. Second, the AR model has a MSE that jumps a lot between horizon one and two, but then declines gradually to horizon four and remains roughly constant from horizon four through six. The pattern for the MSE of the dynamic regression model is a bit different, monotonically increasing with the horizon through horizon four, then declining somewhat to horizon six. Third, the superiority of the dynamic regression model forecast fades as the horizon lengthens, although at its worse the dynamic regression model at horizon four has only 95% the MSE as the AR model.

We report a series of tests suggested by Clark and McCracken (2001) for equal MSE and for forecast encompassing. These are the MSE-F test and ENC-F test. These are appropriate for nested models, and our dynamic regression models all nest our baseline AR model, so Table 3 reports values of these test statistics and bootstrapped marginal probability values for comparisons of our dynamic regression models to the baseline AR model in the case of one-step-ahead forecasts, and comparisons of our preferred dynamic regression model (with one factor, as selected by BIC) with the baseline AR model.[6] For the one-step-ahead forecasts, our dynamic regression models are all statistically significantly better than the baseline AR model. For multiple-step-ahead forecasts, our preferred dynamic regression model is almost always statistically significantly better than the baseline AR model, although at a horizon of four months the probability value for the MSE-F test is nearly six percent.

We also conduct a test of the dynamic regression model, to test whether the common factors help explain inflation. This is an in-sample test. Formally, we are testing

$$H_0 : \beta_1 = ... = \beta_K = 0 .$$

In doing this test we have a generated regressor problem, because the factors are estimated in an earlier step. Hence the test statistic we calculate does not have the standard χ^2 distribution. To deal with this problem we use a bootstrap test.

[6] We bootstrap the distribution for the one-step-ahead forecast tests because we have a generated regressor issue due to estimation of factors for use in the dynamic regression model.

We generate the bootstrap distribution as follows. First, we estimate an AR regression model for each component:

$$X_{i,t} = \alpha + \sum_{k} \delta_{i,k} X_{i,t-k} + \varepsilon_{i,t} \tag{8}$$

From this estimation we get values for the residuals, $\hat{\varepsilon}_{i,t}$. We randomly draw from these estimated residuals to create a new series, $\hat{\varepsilon}^{*}_{i,t}$. Using this new series of residuals and the estimated model in equation (8), we generate a new series of the components of X_t, $\hat{X}^{*}_{i,t}$. Then using $\hat{X}^{*}_{i,t}$ we estimate the common factors $\hat{F}^{*}_{t-k}$ and calculate the F-statistic from the dynamic regression model,

$$y^{*}_{t} = \alpha + \sum_{k=1}^{2} \beta_k ' \hat{F}^{*}_{t-k} + \sum_{l-1}^{13} \gamma_l y^{*}_{t-l} + \varepsilon^{*}_{i,t} \tag{9}$$

These steps are repeated 1,000 times to generate the bootstrap distribution for the F-statistic.

We graph the bootstrap distribution as Figure 5. From the original data we calculate a test statistic of 12.88, which corresponds to a p-value of zero in the bootstrap distribution. Thus we reject the null hypothesis that the factors are not helpful in explaining inflation.

Finally, Figure 6 presents a graph of one-month ahead forecast errors for our baseline AR model and our preferred dynamic regression model. The forecast errors are similar, but the dynamic regression model exhibits a bit less volatility, a bit more accuracy. This is especially apparent in months 35-37 and 45, where the AR model has large errors relative to the dynamic regression model.

4. Conclusion

We set out to investigate the usefulness of disaggregating to improve forecasts, as applied to the CPI inflation rate. Inflation forecasting is an important topic, and recent work has looked at incorporating information from a large number of macroeconomic and financial time series into inflation forecasts. We take a different approach, and look at CPI components for information useful in predicting the aggregated CPI inflation rate. We investigate forecasting the components directly, as well as an application of Hansen's model averaging, an application of principal components, and dynamic regression with factors chosen by either Bai's panel information criteria, the Kaiser criterion, the Scree Plot method, or a method that chooses the number of factors that explain a large fixed percent (here 96%) of the variation. We find that factor models summarizing information in the subcomponents perform well compared to the baseline AR model, and that a dynamic regression model using BIC to select the number of factors as well as the number of lags appears to work very well.

We investigate the properties of our dynamic regression model over different horizons and find that it remains superior to the AR model over the horizons we investigate, 1 to 6 months ahead. We conclude that CPI inflation forecasts can be improved by disaggregating and using the information contained in subcomponents of the CPI. However, we find that summarizing the information in these CPI components by estimating a relatively small number of factors results in better forecasts than directly using the large set of components.

References

Bai, J. (2003). Inference on Factor Models of Large Dimensions, Econometrica 71(1): 135-172.

Bai, J. and Ng, S. (2002). Determining the Number of Factors in Approximate Factor Models, *Econometrica* 70(1): 191-221.

Boivin, J. and Ng, S. (2006). Are More Data Always Better for Factor Analysis? *Journal of Econometrics* 132, 169-194.

Clark, T. E. and McCracken, M. W. (2001). Tests of Equal Forecast Accuracy and Encompassing for Nested Models, *Journal of Econometrics* 105: 85-110.

Diebold, F. X., and Mariano, R. S. (1995). Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253-26.

Hansen, B. E. (2005). Least Squares Model Averaging, *University of Wisconsin working paper*.

Hansen, B. E. (2006). Least Squares Forecast Averaging, *University of Wisconsin working paper*.

Harvey, D. I., Leybourne, S. J., and Newbold, P. (1998). Tests for Forecast Encompassing, *Journal of Business and Economic Statistics* 16, 254-259.

Hendry, D. F. and Hubrich, K. (2006). Forecasting Aggregates by Disaggregates, *European Central Bank Working Paper* No. 589.

Liu, D. and Jansen, D. W. (2007). Macroeconomic Forecasting Using Structural Factor Analysis, *International Journal of Forecasting*, forthcoming.

Park, J. Y. and Qian, J. J. (2006). Autoregressive Modeling of Time-Varying Densities in Functional Space, *Rice University Working Paper*.

Smith, J. K. (2007). Better Measures of Core Inflation, *Lafayette College Working Paper*, May.

Stock, J. H., and Watson, M. W. (1998). Diffusion Indexes, *National Bureau of Economic*

Research working paper 6702.

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

Stock, J. H. and Watson, M. W. (2002). Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business & Economic Statistics*, Vol.20, No.2, 147-162.

West, K. D. (1996). Asymptotic Inference about Predictive Ability, *Econometrica* 64, 1067-84.

Table 1. Out of Sample Forecasting Results – 1 Period Ahead Forecasts

Class of Models	Specific Model	MSFE	Relative MSFE
Baseline AR	AR for Aggregate Inflation	1.04 E-05	1.000
Weighted Forecast of Components	Weighted Average of AR Forecasts of Inflation for 8 Subcategories	8.68 E-06	0.835
	Weighted Average of AR Forecasts of Inflation for 182 Individual Items	9.08 E-06	0.873
Weighted Forecast of Principal Components (Forecasted the weights and used these to aggregate Principal Components.)	2 Factors (AR Forecasts, BIC lags)	8.75 E-06	0.841
	9 Factors (AR Forecasts, BIC lags)	9.27 E-06	0.891
	13 Factors (AR Forecasts, BIC lags)	9.36 E-06	0.900
	29 Factors (AR Forecasts, BIC lags)	9.41 E-06	0.906
	2 Factors (VAR Forecasts, BIC lags)	8.86 E-06	0.852
	9 Factors (VAR Forecasts, BIC lags)	9.71 E-06	0.934
	13 Factors (VAR Forecasts, BIC lags)	9.78 E-06	0.941
	29 Factors (VAR Forecasts, BIC lags)	9.91 E-06	0.953
Dynamic Regression: Factors and Aggregates	1 Factor (selected by *BIC*); lags selected by BIC (lags=2)	8.08 E-06	0.777
	2 Factors, lags selected by BIC	9.18 E-06	0.883
	9 Factors; lags selected by BIC	8.58 E-06	0.825
	13 Factors; lags selected by BIC	8.43 E-06	0.811
	29 Factors; lags selected by BIC	8.81 E-06	0.847
Dynamic Regression: Components	8 components; 1ag length 1	1.06 E-05	1.019
	8 components; 1ag length 2	9.06 E-06	0.871
	8 components; 1ag length 3	9.73 E-06	0.936
	8 components; 1ag length 4	1.06 E-05	1.019
	8 components; 1ag length 5	1.19 E-05	1.144
	8 components; 1ag length 6	1.20 E-05	1.154
	134 items plus aggregate inflation; lag length 1	1.27 E-05	1.221
	134 items; lag length 1 (no aggregate inflation)	1.51 E-05	1.452
Model Averaging (Weights selected using Malow's criteria. Lags of inflation selected by AIC (14) or BIC (1)).	2 Factors, 0-6 Lags; 14 lags of inflation	8.32 E-06	0.801
	9 Factors, 0-4 Lags; 14 lags of inflation	8.86 E-06	0.852
	13 Factors, 0-2 Lags; 14 lags of inflation	8.30 E-06	0.798
	29 Factors, 0-1 Lag; 14 lags of inflation	8.33 E-06	0.801
	2 Factors, 0-6 Lags; 1 lags of inflation	8.50 E-06	0.827
	9 Factors, 0-4 Lags; 1 lags of inflation	8.69 E-06	0.836
	13 Factors, 0-2 Lags; 1 lags of inflation	8.08 E-06	.0777
	29 Factors, 0-1 Lag; 1 lags of inflation	8.25 E-06	0.793

Note: 1. MSFE Ratio is ratio of MSFE of model relative to MSFE of baseline AR.
2. DM test is Diebold-Mariano test statistic.

Table 2. Out of Sample Forecast Results – Multiple Period Ahead Forecasts

Horizon	MSE: Baseline AR Model	MSE: Dynamic Regression Model	MSE Ratio
1	1.04 E-05	8.08 E-06	0.78
2	1.24 E-05	1.06 E-05	0.86
3	1.19 E-05	1.07 E-05	0.90
4	1.16 E-05	1.10 E-05	0.95
5	1.16 E-05	1.08 E-05	0.93
6	1.16 E-05	1.05 E-05	0.91

Table 3. MSE-F and ENC-F Tests

A. Tests for one-step ahead forecasts

Class of Models	Specific Model	MSE- F Test	ENC-F Test
Baseline AR	AR for Aggregate Inflation	--	--
Dynamic Regression: Factors and Aggregates	1 Factor (selected by *BIC*); lags selected by BIC	13.77 (0.000)	13.50 (0.000)
	2 Factors, lags selected by BIC	6.37 (0.000)	6.39 (0.000)
	9 Factors; lags selected by BIC	10.16 (0.000)	10.12 (0.000)
	13 Factors; lags selected by BIC	11.18 (0.000)	11.31 (0.000)
	29 Factors; lags selected by BIC	8.67 (0.002)	11.70 (0.001)

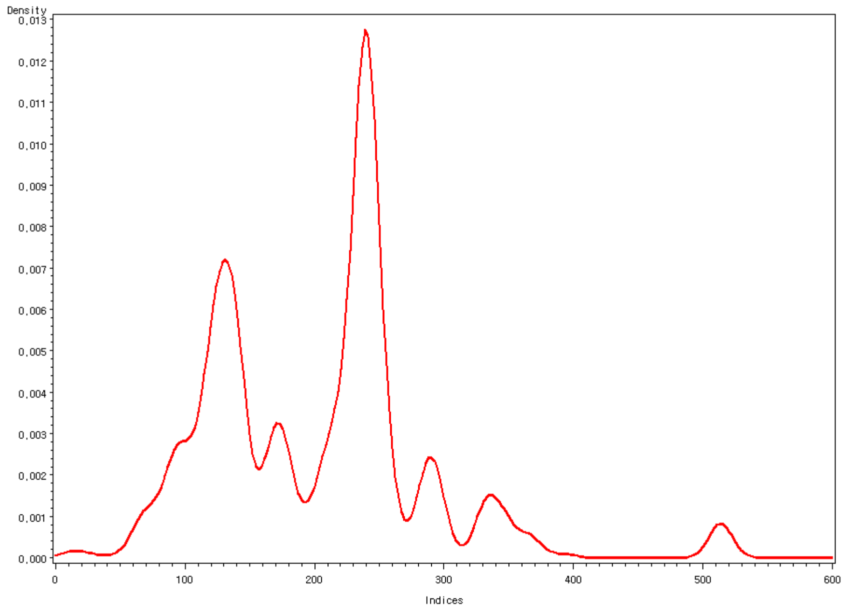
B. Tests for multiple steps ahead.

Horizon	MSE: Baseline AR Model	MSE: 1-Factor Dynamic Regression Model	MSE-F Test	ENC-F Test
1	1.04 E-05	8.08 E-06	13.77 (0.000)	13.50 (0.000)
2	1.24 E-05	1.06 E-05	7.86 (0.000)	10.01 (0.000)
3	1.19 E-05	1.07 E-05	5.61 (0.008)	7.51 (0.000)
4	1.16 E-05	1.10 E-05	2.52 (0.059)	5.15 (0.005)
5	1.16 E-05	1.08 E-05	3.40 (0.036)	5.64 (0.005)
6	1.16 E-05	1.05 E-05	4.83 (0.013)	6.18 (0.003)

Notes:
1. MSE-F is the test for equality of MSE proposed by Clark and McCracken (2001). Marginal probability values calculated from a bootstrap simulation are reported in parentheses.
2. ENC-F is the encompassing F test proposed by Clark and McCracken (2001). Marginal probability values calculated from a bootstrap simulation are reported in parentheses.

Figure 1. Distribution of price change of elementary item indices: Aug. 2006

CPI Level: Weighted average = 203.7



CPI Inflation Rate: Weighted average = 0.25% (per month)

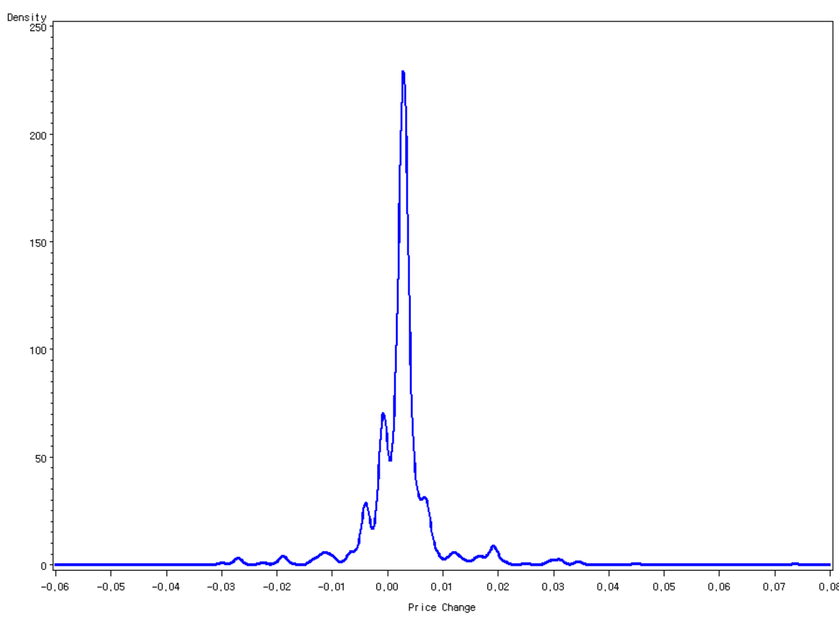


Figure 2. The Scree Plot

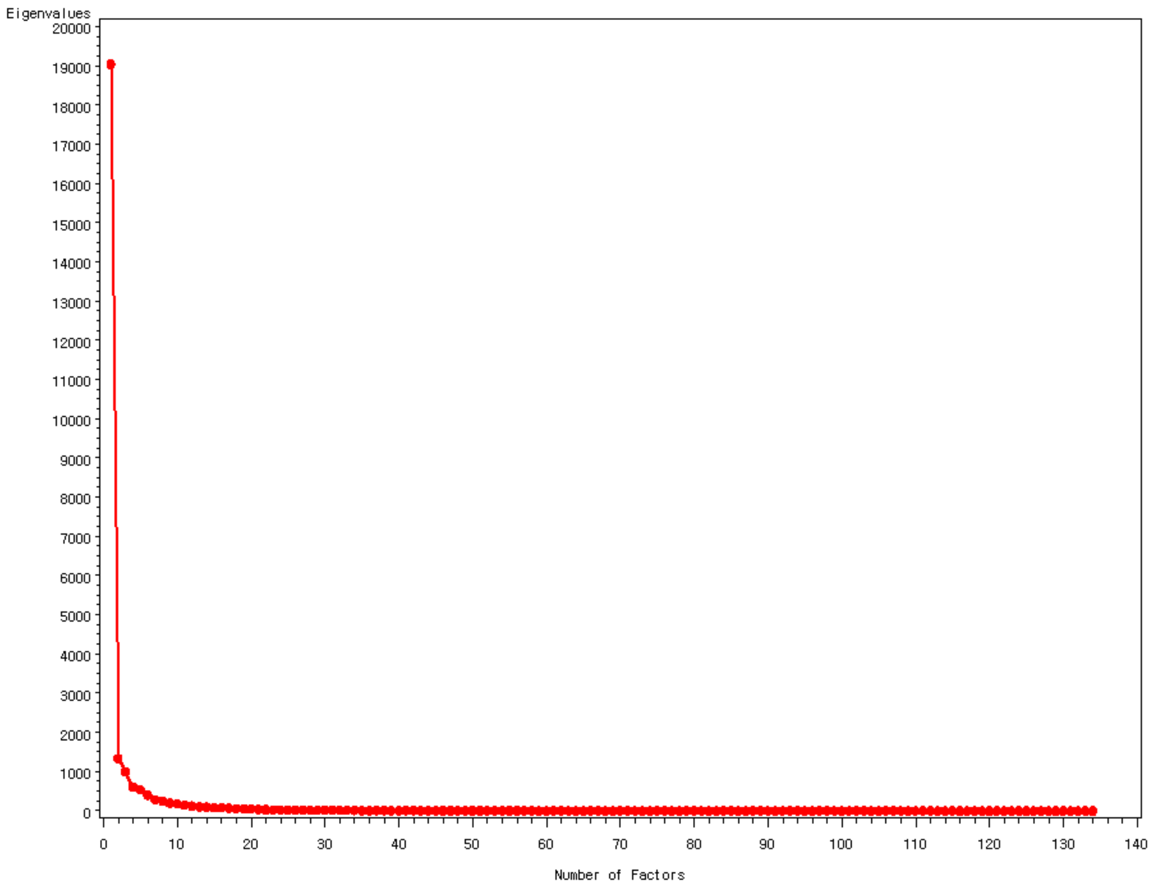


Figure 3. Determining the Number of Factors: The 96% Rule
(Graph of Percent Variance Explained versus Number of Factors)

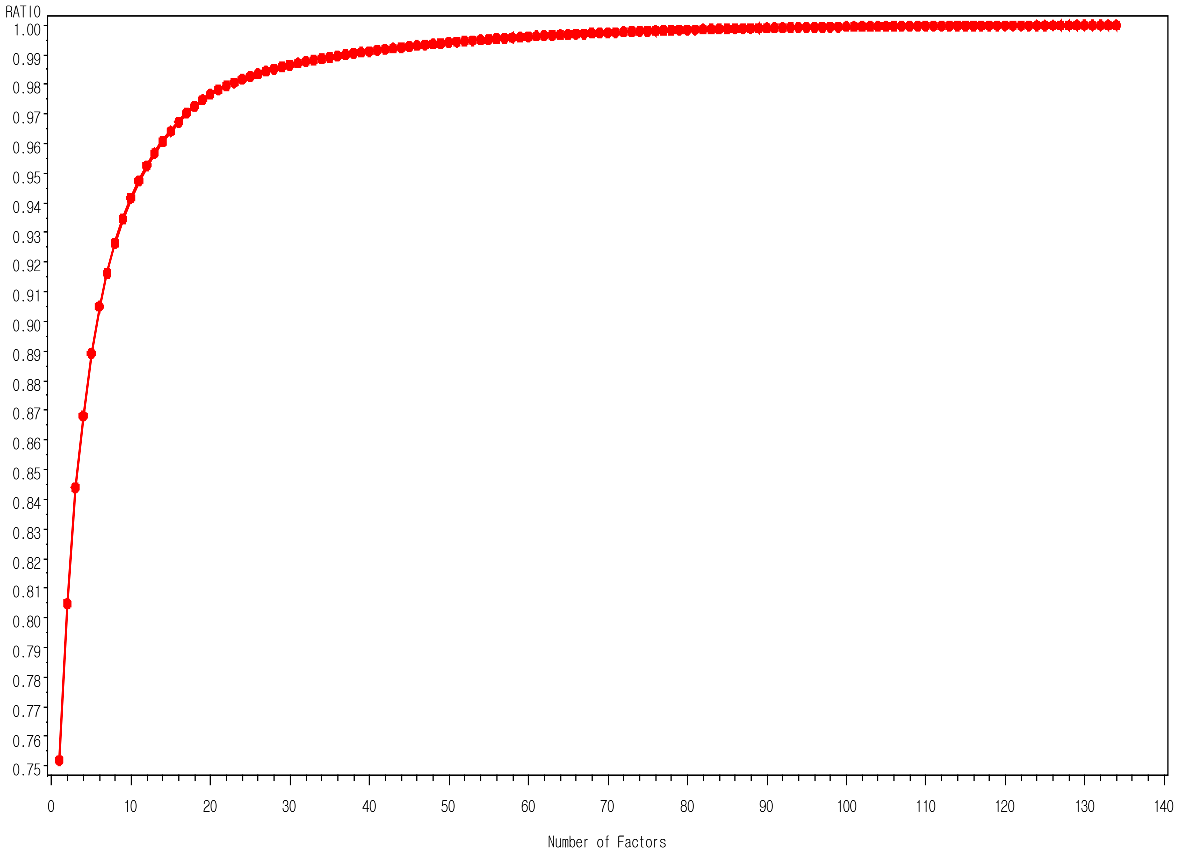


Figure 4. IC_p Plot

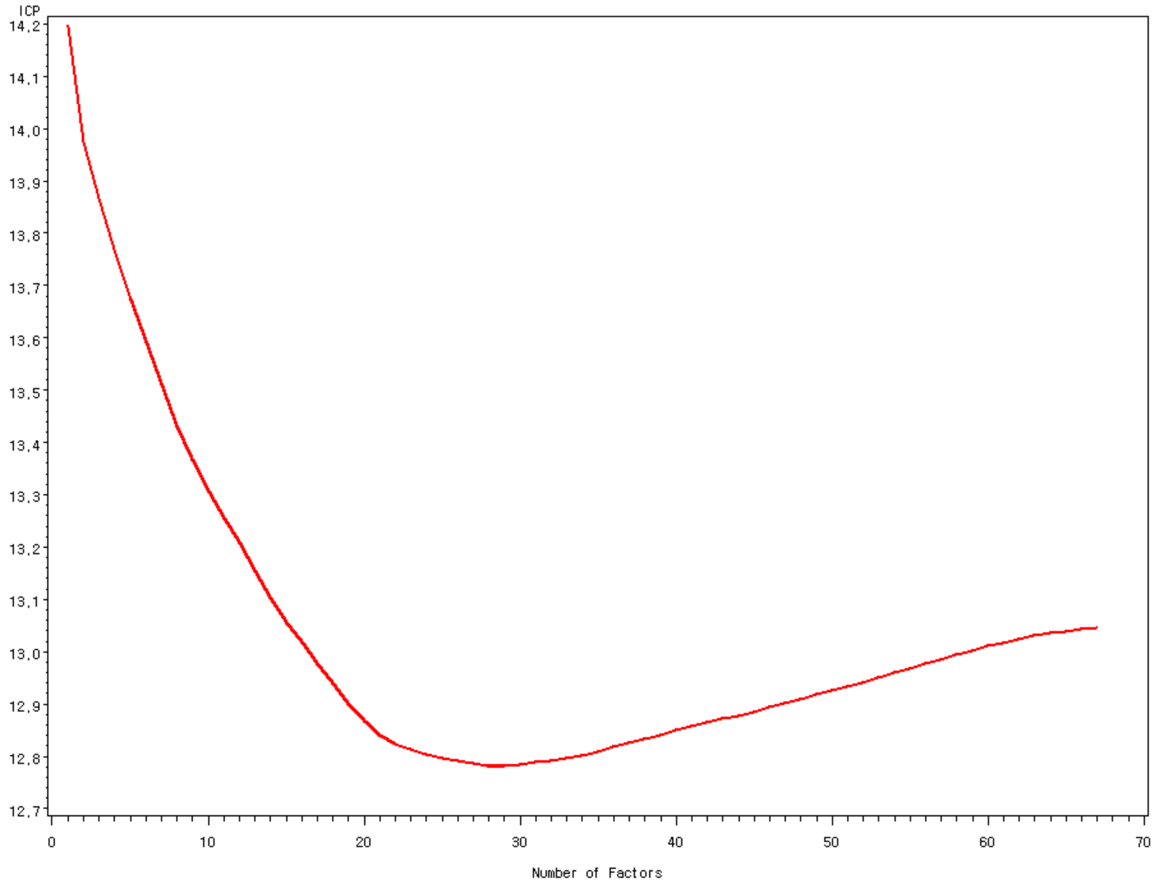


Figure 5. The Bootstrap Distribution for a Test that the Factors Do Not Help Explain Inflation

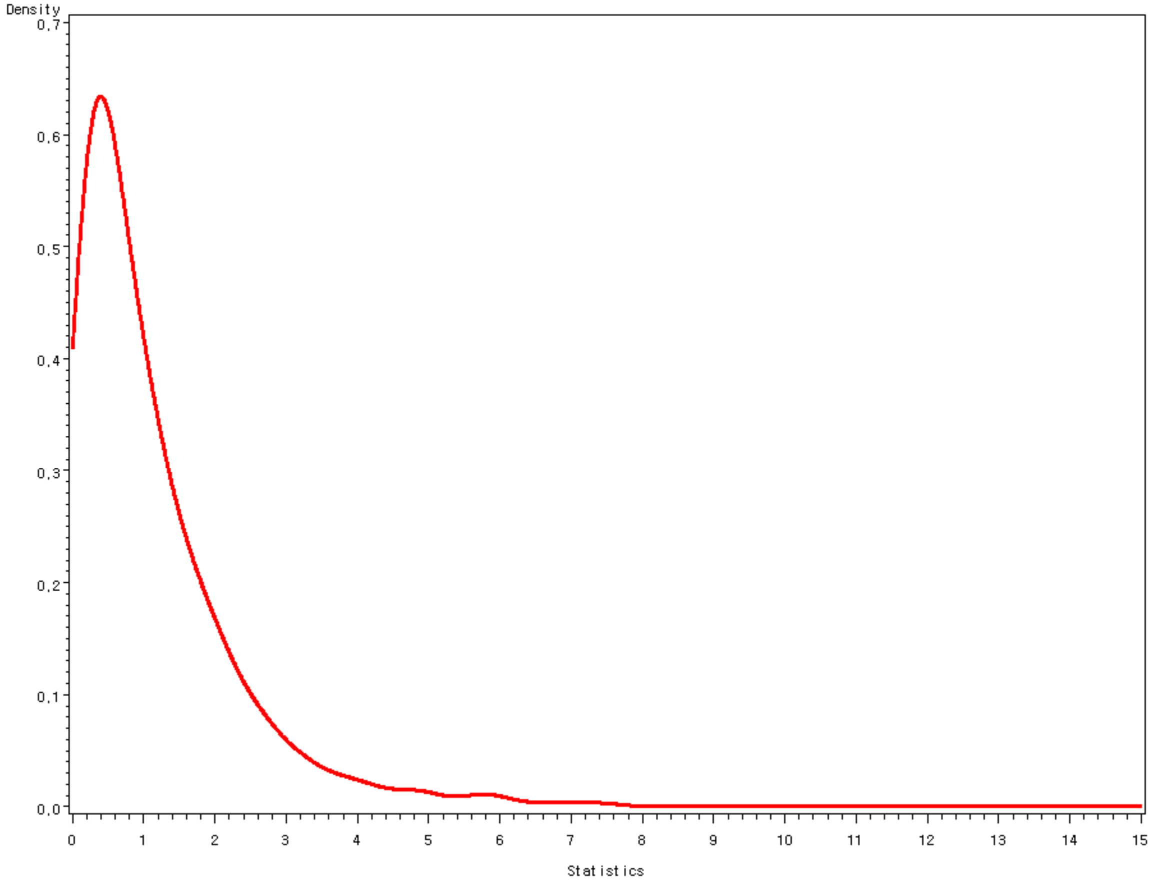
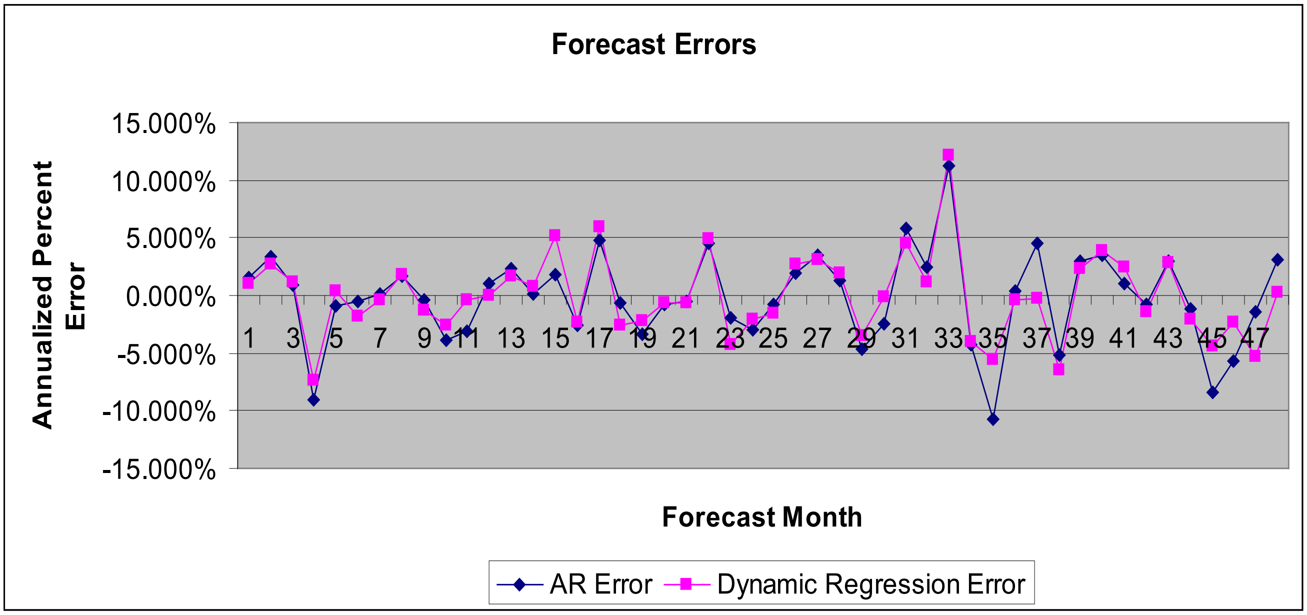


Figure 6. Graph of One-Period Ahead Forecast Errors



Appendix: List of Elementary Items

The following table lists the 134 elementary items included in most of our analyses. We list the Series ID codes assigned by the BLS, the start date for the series, a description of the series, and the weights that hold during our forecasting period. These weights vary over time. W2001 are the weights that hold up to December 2001. W2003 are the weights from January 2002 through December 2003. W2005 are the weights from January 2004 through December 2005, and W2007 are the weights from January 2006 through December 2007.

Our forecasting period is January 2003 to December 2006, so over this period we use W2003, W2004, and W2005 when calculating the inflation forecast based on these 134 elementary items.

The weights changed during our estimation period as well, with weights changing beginning January 1987 and January 1998. The January 1998 weights are the ones that we call W2001 since they hold through December 2001. Note that the changing weights during the forecast period are accounted for in our component AR models and otherwise do not directly impact our forecasts. The changing weights would exert a small impact on our factor estimates made during the estimation period, because we estimate factors from the weighted elementary items. We estimated factors over the estimation period using the weights we label W2001, the weights that held over the end of our estimation period.

series_id	description	W2001	W2003	W2005	W2007
CUSR0000SEFA01	Flour and prepared flour mixes	0.068	0.06	0.051	0.046
CUSR0000SEFA02	Breakfast cereal	0.309	0.24	0.209	0.203
CUSR0000SEFA03	Rice, pasta, cornmeal	0.151	0.137	0.126	0.111
CUSR0000SEFB01	Bread	0.304	0.247	0.224	0.215
CUSR0000SEFB02	Fresh biscuits, rolls, muffins	0.148	0.114	0.108	0.1
CUSR0000SEFB03	Cakes, cupcakes, and cookies	0.269	0.241	0.218	0.209
CUSR0000SEFB04	Other bakery products	0.287	0.254	0.222	0.213
CUSR0000SEFC02	Uncooked beef roasts	0.143	0.137	0.128	0.108
CUSR0000SEFC03	Uncooked beef steaks	0.293	0.34	0.263	0.245
CUSR0000SEFC04	Uncooked other beef and veal	0.053	0.052	0.057	0.051
CUSR0000SEFD01	Bacon, breakfast sausage, and related products	0.199	0.148	0.137	0.133
CUSR0000SEFD02	Ham	0.128	0.102	0.095	0.097
CUSR0000SEFD04	Other pork including roasts and picnics	0.118	0.102	0.102	0.108
CUSR0000SEFF01	Chicken	0.378	0.33	0.32	0.308
CUSR0000SEFF02	Other poultry including turkey	0.111	0.082	0.08	0.075
CUSR0000SEFG01	Fresh fish and seafood	0.187	0.182	0.185	0.205
CUSR0000SEFG02	Processed fish and seafood	0.139	0.123	0.125	0.129
CUSR0000SEFJ01	Milk	0.42	0.323	0.324	0.309
CUSR0000SEFJ02	Cheese and related products	0.324	0.261	0.245	0.264
CUSR0000SEFJ03	Ice cream and related products	0.193	0.166	0.143	0.143
CUSR0000SEFJ04	Other dairy and related products	0.157	0.139	0.124	0.137
CUSR0000SEFK03	Citrus fruits	0.13	0.085	0.098	0.091
CUSR0000SEFK04	Other fresh fruits	0.236	0.218	0.226	0.24
CUSR0000SEFL04	Other fresh vegetables	0.274	0.257	0.247	0.251
CUSR0000SEFN01	Carbonated drinks	0.399	0.353	0.319	0.332
CUSR0000SEFR01	Sugar and artificial sweeteners	0.069	0.056	0.054	0.054
CUSR0000SEFR02	Candy and chewing gum	0.219	0.196	0.19	0.196

CUSR0000SEFR03	Other sweets	0.076	0.06	0.054	0.051
CUSR0000SEFS01	Butter and margarine	0.105	0.081	0.078	0.071
CUSR0000SEFS02	Salad dressing	0.08	0.074	0.065	0.062
CUSR0000SEFS03	Other fats and oils including peanut butter	0.112	0.099	0.103	0.099
CUSR0000SEFT01	Soups	0.106	0.099	0.091	0.088
CUSR0000SEFT02	Frozen and freeze dried prepared foods	0.218	0.241	0.235	0.257
CUSR0000SEFT03	Snacks	0.276	0.258	0.254	0.278
CUSR0000SEFT04	Spices, seasonings, condiments, sauces	0.29	0.229	0.208	0.221
CUSR0000SEFT05	Baby food	0.104	0.085	0.072	0.073
CUSR0000SEFT06	Other miscellaneous foods	0.307	0.28	0.295	0.328
CUSR0000SEFV01	Full service meals and snacks	3.025	2.652	2.666	2.594
CUSR0000SEFV02	Limited service meals and snacks	1.682	2.746	2.664	2.702
CUSR0000SEFV03	Food at employee sites and schools	0.642	0.305	0.3	0.254
CUSR0000SEFV04	Food from vending machines & mobile vendors	0.212	0.15	0.134	0.126
CUSR0000SEFV05	Other food away from home	0.182	0.391	0.338	0.277
CUSR0000SEFW02	Distilled spirits at home	0.112	0.108	0.105	0.125
CUSR0000SEFW03	Wine at home	0.175	0.214	0.191	0.254
CUSR0000SEFX	Alcoholic beverages away from home	0.37	0.353	0.359	0.393
CUSR0000SEHA	Rent of primary residence	7.3	6.516	6.116	5.832
CUSR0000SEHB01	Housing at school, excluding board	0.251	0.259	0.214	0.151
CUSR0000SEHC	Owners' equivalent rent of primary residence	21.064	22.261	22.951	23.442
CUSR0000SEHD	Tenants' and household insurance	0.366	0.364	0.366	0.375
CUSR0000SEHE02	Other household fuels	0.083	0.076	0.108	0.107
CUSR0000SEHJ02	Living room, kitchen, and dining room furniture	0.564	0.514	0.465	0.473
CUSR0000SEHJ03	Other furniture	0.154	0.17	0.174	0.197
CUSR0000SEHK01	Major appliances	0.194	0.179	0.169	0.192
CUSR0000SEHN01	Household cleaning products	0.408	0.365	0.34	0.372
CUSR0000SEHN02	Household paper products	0.228	0.188	0.189	0.203
CUSR0000SEHN03	Miscellaneous household products	0.253	0.257	0.244	0.266
CUSR0000SEAA01	Men's suits, sport coats, and outerwear	0.218	0.193	0.178	0.145
CUSR0000SEAA02	Men's furnishings	0.249	0.189	0.172	0.182
CUSR0000SEAA03	Men's shirts and sweaters	0.267	0.243	0.214	0.197
CUSR0000SEAA04	Men's pants and shorts	0.189	0.177	0.156	0.177
CUSR0000SEAB	Boys' apparel	0.25	0.213	0.183	0.196
CUSR0000SEAC01	Women's outerwear	0.104	0.103	0.103	0.116
CUSR0000SEAC03	Women's suits and separates	0.736	0.694	0.703	0.725
CUSR0000SEAC04	Women's underwear, nightwear, sportswear and accessories	0.387	0.367	0.33	0.364
CUSR0000SEAD	Girls' apparel	0.251	0.273	0.247	0.247
CUSR0000SEAE01	Men's footwear	0.241	0.269	0.232	0.228
CUSR0000SEAE02	Boys' and girls' footwear	0.183	0.164	0.165	0.168
CUSR0000SEAE03	Women's footwear	0.344	0.389	0.35	0.362
CUSR0000SEAF	Infants' and toddlers' apparel	0.25	0.18	0.177	0.183
CUSR0000SETC01	Tires	0.254	0.223	0.211	0.217
CUSR0000SETC02	Vehicle accessories other than tires	0.279	0.188	0.155	0.146
CUSR0000SETD03	Motor vehicle repair	0.992	0.841	0.766	0.602
CUSR0000SETE	Motor vehicle insurance	2.548	2.5	2.412	2.301
CUSR0000SETG01	Airline fare	0.873	0.711	0.623	0.673
CUSR0000SETG02	Other intercity transportation	0.162	0.174	0.139	0.163
CUSR0000SETG03	Intracity transportation	0.317	0.275	0.288	0.248
CUSR0000SEMB02	Nonprescription medical equipment & supplies	0.127	0.109	0.112	0.136

CUSR0000SEMC01	Physicians' services	1.502	1.521	1.551	1.631
CUSR0000SEMC02	Dental services	0.879	0.781	0.738	0.704
CUSR0000SEMC03	Eyeglasses and eye care	0.28	0.279	0.236	0.226
CUSR0000SEMC04	Services by other medical professionals	0.273	0.251	0.251	0.254
CUSR0000SEMD01	Hospital services	1.447	1.429	1.477	1.49
CUSR0000SERA01	Televisions	0.138	0.11	0.109	0.164
CUSR0000SERC01	Sports vehicles including bicycles	0.206	0.266	0.317	0.399
CUSR0000SERC02	Sports equipment	0.215	0.316	0.283	0.274
CUSR0000SERD01	Photographic equipment and supplies	0.09	0.094	0.084	0.092
CUSR0000SERE01	Toys	0.323	0.3	0.273	0.264
CUSR0000SERF02	Admissions	0.825	0.837	0.734	0.688
CUSR0000SERG01	Newspapers and magazines	0.39	0.264	0.219	0.187
CUSR0000SERG02	Recreational books	0.189	0.165	0.141	0.13
CUSR0000SEEB01	College tuition and fees	1.218	1.309	1.416	1.462
CUSR0000SEEB02	Elementary and high school tuition and fees	0.38	0.368	0.381	0.395
CUSR0000SEEB03	Child care and nursery school	0.965	0.879	0.816	0.716
CUSR0000SEED02	Land-line telephone services, long distance charges	0.969	0.773	0.645	0.681
CUSR0000SEGA01	Cigarettes	1.327	0.902	0.758	0.661
CUSR0000SEGA02	Tobacco products other than cigarettes	0.068	0.061	0.061	0.046
CUSR0000SEGB01	Hair, dental, shaving, and misc. personal care products	0.355	0.352	0.337	0.369
CUSR0000SEGD01	Legal services	0.368	0.35	0.355	0.298
CUSR0000SEGD02	Funeral expenses	0.361	0.236	0.247	0.187
CUSR0000SEGD03	Laundry and dry cleaning services	0.218	0.388	0.333	0.283
CUSR0000SEGD04	Apparel services other than laundry and dry cleaning	0.067	0.05	0.044	0.033
CUSR0000SEGD05	Financial services	0.328	0.271	0.238	0.185
CUSR0000SEGD06	Care of invalids and elderly at home	0.117	0.181	0.136	0.105
CUSR0000SEHJ01	Bedroom furniture	0.286	0.288	0.299	0.336
CUSR0000SEHL01	Clocks, lamps, and decorator items	0.184	0.372	0.317	0.356
CUSR0000SETD01	Motor vehicle body work	0.121	0.082	0.084	0.085
CUSR0000SETD02	Motor vehicle maintenance and servicing	0.522	0.483	0.472	0.423
CUSR0000SERA05	Audio equipment	0.131	0.101	0.092	0.079
CUSR0000SEED01	Land-line telephone services, local charges	1.137	1.176	0.891	0.749
CUSR0000SEGB02	Cosmetics, perfume, bath, nail preparations and implements	0.36	0.311	0.304	0.336
CUSR0000SEFW01	Beer, ale, and other malt beverages at home	0.333	0.356	0.32	0.336
CUSR0000SEAC02	Women's dresses	0.189	0.207	0.151	0.132
CUSR0000SEMA	Prescription drugs and medical supplies	0.913	0.984	1.103	1.025
CUSR0000SEFP01	Coffee	0.127	0.096	0.108	0.1
CUSR0000SEHB02	Other lodging away from home including hotels and motels	2.055	2.435	2.795	2.46
CUSR0000SETB01	Gasoline (all types)	2.559	3.241	4.418	4.148
CUSR0000SEEA	Educational books and supplies	0.215	0.246	0.223	0.196
CUSR0000SEMB01	Internal and respiratory over-the-counter drugs	0.255	0.297	0.273	0.296
CUSR0000SEFE	Other meats	0.338	0.301	0.275	0.265
CUSR0000SEFL03	Tomatoes	0.096	0.097	0.107	0.102
CUSR0000SETA01	New vehicles	4.604	4.686	4.521	5.155
CUSR0000SETA02	Used cars and trucks	1.824	1.753	1.998	1.799
CUSR0000SEGC	Personal care services	1	0.901	0.647	0.675
CUSR0000SEHE01	Fuel oil	0.193	0.143	0.251	0.232

CUSR0000SEHF01	Electricity	2.563	2.434	2.575	2.625
CUSR0000SEHF02	Utility (piped) gas service	1.159	1.134	1.658	1.53
CUSR0000SEEC01	Postage	0.199	0.188	0.164	0.169
CUSR0000SEFC01	Uncooked ground beef	0.303	0.296	0.268	0.237
CUSR0000SEFD03	Pork chops	0.127	0.109	0.1	0.083
CUSR0000SEFH	Eggs	0.104	0.127	0.093	0.089
CUSR0000SEFK01	Apples	0.099	0.089	0.087	0.08
CUSR0000SEFK02	Bananas	0.1	0.083	0.078	0.073
CUSR0000SEFL01	Potatoes	0.099	0.08	0.086	0.074
CUSR0000SEFL02	Lettuce	0.064	0.073	0.056	0.058
TOTALS:		89.623	88.896	89.078	88.439

Inflation April – August 2006: Aggregate Level and Distribution of Component Inflation Rates