Constructing and evaluating core inflation measures from component-level inflation data

Edward N Gamber*
Congressional Budget Office
2nd and D Streets, SW
Washington, DC 20515
(202)-226-2777
ed.gamber@cbo.gov

Julie K Smith**
Department of Economics
Lafayette College
Easton, PA 18042
610-330-5301
smithjk@lafayette.edu
and
Research Program on Forecasting
Center of Economic Research
Department of Economics
The George Washington University
Washington, DC 20052

JEL: E31, E37
Keywords: Core inflation, inflation, forecasting, disaggregated components, principal components

*The views and analysis in this paper are the author's and should not be interpreted as the Congressional Budget Office's.
**Corresponding author

We thank two anonymous referees for useful comments and suggestions. We also acknowledge excellent research assistance by Jeromie Ballreich and Kelly Liss on earlier versions of this paper.

Abstract

This paper undertakes a comprehensive examination of ten measures of core inflation and evaluates which measure produces the best forecast of headline inflation out-of-sample. We use the Personal Consumption Expenditure Price Index as our measure of inflation. We use two set of components (17 and 50) of the Personal Consumption Expenditure Price Index to construct these core inflation measures and evaluate these measures at the three time horizons (6, 12 and 24 months) most relevant for monetary policy decisions. The best measure of core inflation for both sets of components and over all time horizons uses weights based on the first principal component of the disaggregated (component-level) prices. Interestingly, the results do vary by the number of components used; when more components are used the weights based on the persistence of each component is statistically equivalent to the weights generated by the first principal component. However, those forecasts using the persistence of 50 components are statistically worse than those generated using the first principal component of 17 components. The statistical superiority of the principal component method is due to the fact that it extracts (in the first principal component) the common source of variation in the component level prices that accurately describes trend inflation over the next 6 to 24 months.

JEL: E31, E37

Keywords: Core inflation, inflation, forecasting, disaggregated components, principal components

1. Introduction

Core inflation measures are essentially a re-weighting of the growth in component-level prices to construct an aggregate measure of inflation that achieves some objective such as maximizing the correlation of (the alternatively-weighted) inflation with money growth (Bryan and Cecchetti, 1994), maximizing economic stability in an inflation targeting regime (Mankiw and Reis, 2003) or minimizing the out-of-sample forecast error over some forecast horizon (Smith, 2004). Researchers differ in their approaches to constructing the weights. Some approaches place zero weight on certain volatile components while others replace the budget-share weights with weights chosen to achieve one of the aforementioned objectives.

Our objective is to identify the weighting method that produces the most accurate out-of-sample forecasts of the growth in the Personal Consumption Expenditures Price Index (PCEPI) over 6, 12 and 24-month horizons. We investigate three methods that place zero weight on certain components. Those methods are the weighted median, the trimmed mean and the less food and energy measures of core inflation. We also investigate three methods that are based on a re-weighting of all the component level inflation data. The first method is based on weights that measure the persistence of each of the component-level inflation rates. The second method is based on weights estimated from a linear regression of headline inflation (at some future date) on the full set of component-level inflation series. The third method uses the weights estimated from the first principal component of the component-level data. Besides identifying the weighting procedure that results in the most accurate out-of-sample forecasts, we investigate whether the number of components or the level of disaggregation used matters in creating good forecasts. We test two sets of components[1] or levels of disaggregation and compare forecasts not only within sets of components but also compare the best forecasts from each set.

Previous work has evaluated the accuracy of each of these measures of core inflation individually and relative to a handful of benchmark inflation forecasts. Our paper fills a gap in the literature by performing a comprehensive evaluation of the accuracy of all these measures.[2] Because we compare the accuracy of a large set of possible core inflation measures our findings provide valuable guidance to policy makers interested in forecasting inflation. Additionally, the deteriorating performance of Phillips-curve based inflation forecasting models since the mid-1980s (Stock and Watson (2007)) highlights the need for an assessment of the forecast performance of the various concepts of core inflation examined here.[3]

We find that core inflation constructed using weights based on the first principal component factor loadings of 17 components is statistically better than the other methods that use either 17 or 50 components. These results reinforce findings by Crone et al. (2013) that the trimmed mean inflation rate is not the best forecaster of headline inflation and contradict those by Hendry and Hubrich (2006) who find that for the United States using sectoral (component) level data aggregated using regression weights provides a good forecast of aggregate headline inflation.

---

[1] We use the 17 components suggested by Stock and Watson (2016) and the 50 suggested by Bermingham and D'Agostino (2014).

[2] Stock and Watson (2016) compare forecasts from an unobserved components measure of core inflation to forecasts produced by six benchmark methods. However, the six benchmark methods do not include the limited influence measures (such as the trimmed mean) that we examine here and that are widely followed by economic forecasters and policy makers. Previous research has found mixed results as to whether these limited-influence estimators are good forecasters (Smith, 2004; Crone et al., 2013). In addition, the Federal Reserve Bank of Cleveland produces a monthly weighted median inflation rate of the Consumer Price Index and the Federal Reserve Bank of Dallas produces a monthly trimmed mean inflation rate of the Personal Consumption Expenditure deflator indicating that policy makers are interested in the information contained in these measures. Therefore, we conduct a horserace with the other measures suggested by the literature.

[3] An exception to this general finding in the literature is Ball and Mazumder (2011) who present a modified Phillips curve that captures movements in inflation for the entire period since 1960 – up to and including the Great Recession.

The rest of this paper is outlined as follows. In Section 2, we describes previous research on core inflation. Section 3 describes the data. The empirical models and results are examined in Sections 4 and 5. Section 6 concludes.

2. Previous Research on Measuring Core Inflation

Headline inflation is constructed from individual component-level price changes using budget-shares as weights. While budget shares might be useful when comparing the cost of living over two historical time periods, they are not necessarily the optimal weights from a forecasting perspective. For example, components that are subject to frequent tax changes, weather or supply related disturbances might affect headline inflation from period to period because of their relatively large budget share but have little effect on overall trend inflation. Researchers have investigated several alternatives to budget shares in an effort to dampen the effect of these transient movements in prices.

The weighted median and trimmed mean measures of core inflation have been researched extensively.[4] Bryan and Cecchetti's (1994) idea of using disaggregated information on prices to obtain core inflation sparked great interest. With the trimmed mean and weighted median, which exploits the cross-section information in the components, core inflation is constructed period-by-period. Each component can have a different weight every period. Specifically, in the case of the weighted median a component can have a zero weight (not the median element) in period 1 and then in period 2 have a weight of one (the median element); therefore, the lack of smoothness of weights across time may disregard some information relevant to future inflation (the correlation of prices over time, for example) which may lead to inefficient forecasts.

[4] See Bryan and Cecchetti (1994), Alvarez and de los Llano (1999), Apel and Jansson (1999), Cockerell (1999), Johnson (1999), LeBihan and Sedillot (2002), and Smith (2004).

In addition to the weighted median, the trimmed mean (and closely related) less food and energy measures of core inflation, (all of which impose zero restrictions on some of the component-level prices), researchers have also examined more data-rich approaches to constructing core inflation. By data-rich we mean approaches that make use of component-level data in the construction of core inflation. First, Blinder (1997) and Cutler (2001) suggest re-weighting the component inflation rates with weights based on the time-series persistence of each component. The intuition behind this approach is that the component-level prices that are subject to more persistent shocks may be more informative about future movements in inflation than other less-persistent components regardless of their budget-share weights. This method produces less variation in weights on the components across time compared to the weighted median and trimmed mean that place different weights on each component every period.

A second data-rich approach is to directly estimate the relationship between the component-level inflation rates and the overall (headline) inflation rate at some future date using ordinary least squares. Hendry and Hubrich (2006) suggest this approach. They estimate the weights for their measure of core inflation by regressing inflation at some future date $t+h$ on inflation in the individual components at time t. A potential drawback of their approach is that including all of the component-level data on the right hand side of the regression risks overfitting and potentially compromises the out-of-sample forecasting performance. Stock and Watson (2016) also employ a data-rich method to estimate core inflation. They apply a dynamic factor model to 17 component-level price series of the personal consumption expenditure deflator index. Their method produces a core inflation rate that more accurately forecasts inflation over the 1-3 year horizon compared to several benchmark approaches.

The third data-rich method examined in this paper uses principal components to combine the information contained in the component-level data.[5] Principal components potentially overcomes the over-fitting issue associated with the Hendry and Hubrich approach by reducing the dimension of the cross section to a small number of series that are constructed to maximize the common variation in the underlying components.[6] The weights (factor loadings) assigned to each individual series in the construction of the first principal component are chosen to maximize the common element among all of the individual series. Thus, individual series which are jointly persistent – their persistence arises from a common source – will be highly correlated and will receive a greater factor loading in the principal components algorithm compared to series that are less correlated with the common (persistent) element. Principal components therefore has the potential to capture both the time-series persistence and the cross-sectional correlations between the individual component-level price data.

3. Data

The data are components of the Personal Consumption Expenditure Price Index (PCEPI)[7] from January 1959 to December 2015 but given Smith's (2005) results that the monetary regime matters for determining the best forecaster of inflation we analyze the forecasts starting in 1984, the approximate beginning of the Great Moderation in the United States. We examine two levels of disaggregation, breaking down PCE prices into both 50 and 17 components. The 50 components

[5] See Colmbra (1997), Marques et al. (2001), and Maria (2004) for examples of using principal components to measure core inflation in Portugal.

[6] The methods discussed are not all the possible ways to construct core inflation measures (see for example, Cogley, (2002) who applies an exponential smoothing filter to the trimmed mean and weighted median) but cover the major ways to construct core inflation measures.

[7] These data are subject to revision and were downloaded on 2/18/2016. We recognize that using the real-time data would be more consistent with the "real-time" nature of the forecast evaluation; however, the real-time estimates of the components are not readily available. Future work will look toward providing a real-time estimate of the PCEPI core inflation measures.

and 17 components are listed in Appendix A and Appendix B, respectively.[8] The 50 components are the same as those used by Bermingham and D'Agostino (2014) and the 17 components are the same as those used by Stock and Watson (2016). We use both sets of components to forecast and then compare the forecasting results from both within and across sets of components.

We forecast the 6-month, 12-month and 24-month ahead inflation rate[9], which is calculated by

$$\pi_{t+k,t} = \left(\left(\frac{P_{t+k}}{P_t} \right) - 1 \right) * (100/(k/12)), \tag{1a}$$

k= 6,12 or 24

where P is the PCEPI. We also compute lagged inflation as the previous 12-month inflation rate of the PCEPI and lagged inflation minus food and energy from the PCEPI minus food and energy in the following manner

$$\pi^{y}_{t,t-12} = \left(\left(\frac{P^{y}_{t}}{P^{y}_{t-12}} \right) - 1 \right) * 100 \tag{1b}$$

where y denotes either PCEPI or PCEPI minus food and energy.

We find the previous 12-month inflation rates for the components from the monthly price indexes by

$$\pi^{j}_{t,t-12} = \left(\left(\frac{P^{j}_{t}}{P^{j}_{t-12}} \right) - 1 \right) * 100, \tag{1c}$$

where j denotes the component, P^j is the price index for component j and j is either 17 or 50.

[8] We choose to use these two sets of components as they have been previously used in the literature. In addition, if more components are chosen it becomes difficult to estimate as multicollinearity arises and if too few are chosen then estimating the weighted median and trimmed mean does not result in much variation in those two measures.

[9] Smith (2004) finds that there is very little difference in the ranking of core inflation measures in forecasts at the 12, 18 or 24 month time horizon.

We also compute the weighted median and trimmed mean from these components. The first step in calculating both of these measures involves sorting the component-level price changes, weighted by their respective budget shares or relative importances, w^i, in each time period. The relative importances are calculated as the fraction of the consumption basket that is attributed to each component each month. The trimmed mean[10] is calculated as

$$\pi_{t,t-12}^{tm} = \frac{1}{1-\alpha-\beta} \sum_{i=\hat{i}_t(\alpha)}^{\hat{i}_t(1-\beta)} w^i \pi_{t,t-12}^{i}$$

where $\alpha = .24$ (lower tail), $\beta = .31$ (upper tail), (1d)

w^i is the relative importance weight and

i indicates the component.

The weighted median is the most extreme version of the trimmed mean where $\alpha = \beta \approx$ ½.[11]

The reason we use smoothed inflation rates as independent variables is to reduce the noise from the monthly inflation rates. Atkeson and Ohanian (2001) use a similar smoothing in their benchmark random walk model of inflation.

4. Empirical Models

We examine several models that re-weight the component inflation rates and we compare these models to standard measures (lagged inflation, lagged inflation minus food and energy, lagged weighted median and lagged trimmed mean). Conceptually, core inflation is $\pi_t^c = \beta_1 \pi_{t-1}^1 + \beta_2 \pi_{t-1}^2 + ... + \beta_j \pi_{t-1}^j$, where π^c is core inflation based on data-rich components of inflation and *j* refers to the (in our applications, 17 or 50) components of inflation. Lagged inflation is a special case where β_j is equal to the budget shares or relative importance weights.

[10] The trimming is taken from Dolmas' (2009) technical note from the Federal Reserve Bank of Dallas (http://dallasfed.org/data/pce/tech.pdf).We use the same trimming as the Dallas Fed to be consistent with their method.

[11] In both the trimmed mean and weighted median, the relative importance weights are not normalized after trimming; this is standard procedure with limited-influence estimators.

The next question that arises is how to determine the weights (β). We use three methods to calculate the weights. In the first method, we estimate the weights that provide the best fit from a time-series regression. In the second method, we impose the weights. When imposing the weights we follow a methodology similar to Cutler (2001) and estimate the weights based on the persistence of each component and then impose the weights to forecast headline inflation. We return to discuss the details of Cutler's specification later.[12] In the third method, we combine the information from the disaggregated components by estimating the underlying factors using principal components. We then use those estimated factors to forecast inflation. Appendix Table C outlines the ten models used.

Regression weighted (disaggregated) model

The first model regresses aggregate inflation on the component inflation rates in the following regression:

$$\pi_{t+12,t} = \alpha + \beta_1 \pi^1_{t,t-12} + \beta_2 \pi^2_{t,t-12} + \ldots + \beta_j \pi^j_{t,t-12} + \varepsilon_t , \qquad (2)$$

where $\pi_{t+12,t}$ is the 12-month ahead inflation rate, $\pi^j_{t,t-12}$ is the previous 12-month component inflation rate and j is either 17 or 50. This is the **disaggregated** regression based model.

Persistence weighted models

Cutler uses an AR(1) model of the persistence of each component's inflation rate.[13] This persistence coefficient then becomes the weight for that component. To obtain forecasts she aggregates the component inflation rates by these estimated persistence weights. She allows the persistence weights to vary annually as do we.

[12] All models are estimated with Newey-West corrected standard errors.

[13] Cutler's data are for the United Kingdom and she uses eighty-one components at a monthly frequency.

To find the persistence weight for each component we estimate a recursive AR(1) with monthly inflation rates measured as year-over-year inflation rates. The following regression is estimated recursively by OLS for each component.[14]

$$\pi^j_{t+12,t} = \alpha + \beta_j \pi^j_{t,t-12} + \varepsilon_t, \quad (3)$$

where β_j is the estimated coefficient and t is inflation from period *t-12* to period *t*.[15] If β_j is positive then there is persistence in the component and the persistence coefficient is equal to β_j and if β_j is negative then the persistence coefficient is equal to zero because there is no persistence in that component. The weights are normalized to sum to one. The persistence weights do not vary monthly but annually. After obtaining the persistence weights, we transform the data to obtain the **persistence-weighted** forecast of aggregate inflation.

The next model uses a combination of **persistence weights and the budget shares**. The weight on each component equals the persistence weight multiplied by the budget share for each component. This specification prevents a component that is highly persistent but is relatively unimportant (as a share of the consumer's budget) from dominating the forecast. This measure may be more useful than the other persistence-weighted measure because it takes account of both factors: the persistence of an individual component over time and the importance of an individual component in aggregate inflation.

Another model combines the idea of persistence weights and the regression based model. In this model (**disaggregated persistence weighted**), we first calculate the persistence weights as described above (equation 3) and then we calculate the persistence weighted component inflation

[14] We use data from 1960-1982 to find the 1983 persistence weights. After 1983 the persistence weights vary by year and are estimated recursively.
[15] We use year-over-year inflation rates for the AR model to be consistent with our use of year-over-year inflation in the other regressions.

rates. With these inflation rates, we then run a regression-based model similar to the one in equation 2.

The next model combines the idea of **persistence weighting and the weighted median**. We first compute the annual persistence weights as we did for the persistence-weighted model. Then, for each month, we find the median by using the persistence weights instead of the relative importance weights to rank the component inflation rates.

Principal component model

Finally, following Maria (2004), and Hendry and Hubrich (2011) we combine the information from the component inflation rates by estimating the **principal components** of our inflation rates of our disaggregated components of the PCEPI. We use the technique proposed by Bai and Ng (2002) to determine the optimal number of factors (which in all cases equaled one).

Benchmark, traditional core inflation and limited-influence estimators

We compare these regression results to forecasts made with more standard measures. The first uses **lagged headline** inflation as the forecaster.

$$\pi_{t+12,t} = \alpha + \beta \pi_{t,t-12} + \varepsilon_t \ , \tag{4}$$

The second uses either the lagged **weighted median**, lagged **trimmed mean** or lagged inflation **less food and energy** as the forecaster.

$$\pi_{t+12,t} = \alpha + \beta x_{t,t-12} + \varepsilon_t \ , \tag{5}$$

where *x* is the weighted median, trimmed mean or the minus food and energy.

Forecast method

We perform pseudo out-of-sample forecasts[16] at three time horizons (6 months, 12 months and 24 months). We compute the monthly forecasts of the year-over-year inflation rate using recursive least squares (RLS). We limit our forecasting period to 1984 to the end of 2015 for our main set of results that look at forecasting inflation 12 months ahead; however, we use all available data since 1960 in the estimation. We conduct robustness checks by forecasting the 6-month ahead inflation rate and the 24-month ahead inflation rate as well. In addition, we break the sample (1984-2015) down into subsamples as transparency of the Federal Reserve has changed, the Great Recession occurred and possibly the end of the Great Moderation. We believe there may be structural breaks that effect which core inflation measures forecasts best due to changes in how monetary policy is conducted. Lucas (1976) notes that any change in the conduct of monetary policy will lead to changes in the reduced-form parameters in econometric model; therefore, we estimate our model over three subsamples.

Our samples are:

Sample	Why break?
1984 - 2015	Great Moderation begins around 1984.
1984 - 1993	Federal Reserve begins making post-FOMC announcements about stance of monetary policy.
1994 - 2007	Period of increased transparency of the Federal Reserve until start of Great Recession.
1984 - 2007	Great Moderation begins and possibly ends at the start of the Great Recession.

5. Results

Table 1a shows the results for the 12-month ahead time horizon. In our analysis, we use lagged headline as the benchmark model. This is common in the inflation forecasting literature particularly after Atkeson and Ohanian (2001) found that this simple model could beat the

[16] We use the last month of data before the month we are forecasting. These data would not have been available to forecasters in real time but this is a common way to evaluate forecasts out-of-sample.

Phillips curve in forecasting inflation. The full sample and three sub-samples are shown in Table 1a. The results use data from 1959 to 2015 and recursive estimation begins in January 1960 and ends in December 2014 due to leads and lags needed to transform the price level data into year-over-year inflation rates.

We have ten forecasting models.[17] We report the root mean squared error (RMSE) for each model, the ratio of each model to the benchmark, the Diebold-Mariano (2002) test statistic[18] that tests if the model provides significantly different forecasts from the benchmark and the mean square forecast error (MSFE) with Newey-West corrected standard errors. We focus our discussion on the models that are significantly better than the benchmark.

For the 50 component results, across subsamples of the 12-month ahead forecast in Table 1a, we find that both the persistence weighted model in 4 samples and the principal components model in 3 samples outperform the lagged inflation benchmark using a Diebold-Mariano test. The remainder of the models are statistically equivalent to the lagged headline benchmark. We also test using a Diebold-Mariano test if the persistence weighted forecast is statistically different from the principal components forecast. In Table 2, we note that for forecasting the next 12 months of inflation that both models are statistically equivalent across sub-samples.[19]

As a robustness check, we forecast both the six-month ahead and 24-month ahead inflation rates over the same samples as reported in Tables 1b and 1c. Again, the persistence weighted and principal component models are both either equivalent to the benchmark or better

[17] Appendix Table C provides a brief description of each forecasting model.
[18] Our Diebold-Mariano test statistic is also calculated correcting for autocorrelation and heteroscedasticity by using Newey-West corrected standard errors.
[19] The persistence weighted model and principal component model are equivalent in all but one case for the other two time horizons.

than the benchmark. There is only one instance when the persistence-weighted model is statistically better than the principal component model as noted in Table 2.

Similar to Clark (2006), we find that the disaggregate components are very persistent. Only one component of the 50 (insurance) has an AR(1) coefficient of less than 0.9 for every year. The evidence on the persistence of aggregate inflation has been more mixed therefore making it somewhat surprising that using the normalized persistence weights of the components provides better forecasts of inflation.[20] Due to the large AR coefficients in the persistence-weighted model, we undertake a panel unit root test to see if our panel of component inflation rates contains a common unit root. Using the Levin, Lin, and Chu (2002) test, we can reject the null of a common unit root for the 50 components and for the 17 components.[21]

For the 17 component results, across subsamples of the 12-month ahead forecast in Table 1a, we find that the principal component model is statistically better than the benchmark. Interestingly, the persistence-weighted model is no longer better than the benchmark. These results are similar at both the six-month and 24-month time horizon presented in Table 1b and 1c. With 17 components, the other models are mostly statistically similar to the benchmark as was the case with the 50 components.

Given these mixed results for the two sets of components, we test whether the principal component forecasts based on 17 components is better than both the 50 components principal component forecast and 50 components persistence-weighted forecast. We find as presented in Table 3 that the principal component forecast made from the 17 components is statistically better

[20] See Gamber, Liebner and Smith (2016) for a full discussion of the changes in aggregate inflation persistence.
[21] We run the panel unit root test on the component-level inflation rates and include an intercept. Our use of the panel unit root test is motivated by the finding that such tests help mitigate the issue of low power associated with single equation unit root tests. See Levin, Lin, and Chu (2002).

than both the principal component forecast made from the 50 components and the persistence-weighted forecast made from 50 components.

Therefore, the question that remains is why is principal components better than persistence weighted? The principal component model is capturing some comovement and common volatility that is not captured by using an AR(1) model individually on each component. In the 50 component case there are enough components that AR(1) (persistence weighted estimation method) models, which capture the persistence of each component, can accurately predict movements in aggregate inflation. Similarly, the first factor loadings in 50 components principal component model can explain enough of the common variation in the components to capture aggregate inflation movements accurately. However, when reducing the number of components to 17, the first factor loadings of the principal component model can explain more of the common variation in the components of inflation and therefore produces the best inflation forecasts.

The result that the 17-component principal component inflation measure out-performs the 50 component principal component inflation measure is consistent with the Monte Carlo simulations of Boivin and Ng (2006). In particular, they show that more data does not always produce better forecasts in factor models. To gain more insight into what is driving this result in our application, we follow Boivin and Ng (2006) by estimating the relationship between each of the component-level inflation series and their respective first principal components for both the 17 and 50 component cases. Table 4 shows the results of this exercise and is organized as follows. The categories for the 17 components of the PCE inflation rate are listed as the main bolded headings. The 50-component categories are nested below those sub-headings. The R-squared statistics are obtained by regressing each of the component level inflation series on the first principal component. For example, the R-squared from the regression of clothing and footwear price inflation on the first principal component derived from the 17-components is 0.31, and the R-squared from the regression of garments price inflation on the first principal component derived from the 50-components is 0.29. The final right-hand column of Table 4 reports the R-squared statistics averaged across the sub-components comprising each of the 17 PCE expenditure categories.

The results from this table show that the relationship between the 50-component level inflation rates and the estimated principal component is on the whole much weaker than the relationship between the 17-component level inflation series and its estimated principal component. This contrast can be seen by comparing the R-squared statistics for the 17-components to the average across the groupings of 50 components. For the majority of categories, the R-squared is higher for the 17 components than the average of the sub-components. Across all 17 components the average R-squared is roughly 11 percentage points higher than the R-squared of the 50 components.

The intuition behind the results presented in Table 4 is as follows. In the 50 component case, principal components is applied to a number of smaller sub-categories that in the 17-component case get down-weighted. For example, in the 50-component case, most of the explanatory power of the sub-components of "furnishings and durable household equipment" comes from "furniture and furnishings." In the 17-component case the remaining sub-categories are given less weight and therefore have less influence on the common factor. And because those remaining sub-categories are weakly correlated with the common factor, down-weighting those categories improves the overall fit. This is essentially an application of the weighted principal components described in Boivin and Ng (2006). They showed that down-weighting certain components results in an overall stronger fit between the common factor (the first principal component) and the individual series. Thus, the weights used to aggregate the 50 sub-components up to the 17 components illustrate the tradeoff identified by Boivin and Ng (2006) in their Monte Carlo simulations. As shown in Figure 1, there is a positive relationship between the PCE consumption share and the R-squared which indicates that by aggregating from 50 to 17 components, relatively lower weights are put on the sub-components with relatively lower correlations with the common factor (the first principal component).

In Figures 2a and 2b, we compare the 12-month ahead inflation rate to the recursive forecasts generated by both the persistence-weighted model and principal components model for both sets of components. Figure 2a shows that the forecasts based on 50 components exhibit similar degrees of dispersion around the actual inflation rate; therefore, the statistical equivalence of the two models is not unexpected. However, in the case of the 17 component models (Figure 2b), we see the persistence-weighted forecasts are missing the turning points of the headline inflation data contributing to its poor forecasting performance. The 17 component principal components forecast provides a relatively smooth forecast hence leading to its strong forecasting performance.

In addition, our results shed light on the four questions posed by Stock and Watson (henceforth SW) (2016) in their recent research on trend inflation. In particular, SW first ask whether “more precise measures of trend inflation can be obtained using disaggregated sectoral inflation measures, relative to time series smoothing.” Similar to SW, our answer to this question is yes. We find that more precise (more accurate) forecasts are obtained when we re-weight the sector-level inflation rates using principal components or persistence weights that differ from the budget shares used in the construction of headline inflation.

The second issue addressed by SW is the variation in weights across components and over time. Figure 3 shows the cross-method variation by displaying the average over time for the three methods: persistence-weighted, principal components and budget shares. Compared to budget shares, both the persistence-weighted method and principal components down weight housing and healthcare prices by roughly half.[22] Principal components, and to a lesser extent the persistence method places more weight on food services and accommodations compared to the budget weights. Principal components and persistence weights both give less weight to financial services and insurance (as well as other services) compared to the budget share weights. In terms of variation over time, the budget share weights in general appear to vary more than the principal component or persistence weights. This makes sense since the principal component and persistence weights are based on recursive estimates using progressively larger samples whereas the budget share weights are (approximately) the consumption shares of each component that can vary considerably from period to period. Nonetheless, the principal component and persistence weights do exhibit variation over time. Noteworthy are the first principal component factor loadings on food services and accommodations, and transportation which rise over the sample period. The principal component factor loading on gas and electricity falls over the sample.

Finally, SW ask (third) how the trend inflation measures they derive from component level data compare to conventional measures and (fourth) whether they produce more accurate forecasts compared to conventional measures. Similar to SW, we find that the comparison is quite favorable in the sense that our methods based on component-level data—particularly the

[22]In this discussion we focus on the 17 component model rather than the 50. The results for the 50 components weights are consistent with the results for the 17. In particular, the persistence weights and normalized factor loadings on housing and (two categories) of health care are significantly smaller than the budget shares for those components. Financial services also receives less weight in the principal components and persistence methods.

principal component method—produces more accurate forecasts at 6, 12 and 24-month horizons than the conventional measures of core inflation.

6. Conclusion

This paper examines whether using information in the components of inflation can lead to a better forecast of inflation. We find that using the first principal component factor loadings extracted from 17 components of inflation produces better forecasts than using standard measures of core inflation such as inflation less food and energy. In addition, this paper explores many new models such as persistence weighted and disaggregated regression weighted that utilize the varied information in the component inflation rates but ultimately those do not yield good forecasts. The results suggest that using the principal components method to generate forecasts of inflation provides a more stable forecast across both the full sample and sub-samples.

This paper is closely related to several in the literature that look at subsets of these models and compare forecasting performance relative to some benchmark. Our findings both confirm and differ from these previous studies and in particular reveal that finding a good forecaster of inflation is an empirical question that must continue to be studied as the structure of the economy evolves. The limited influence estimators such as the weighted median and trimmed mean found to be poor forecasters in Crone et al. (2013) seem to continue to be poor forecasters. In addition, we find that the disaggregated regression based weights suggested by Hendry and Hubrich (2006) perform poorly. Finally, some methods that combine specifications from two models also perform poorly. For example, weighting components by both persistence and budget shares, or finding the median based on persistence mostly results in forecasts equivalent to the benchmark or worse than the benchmark. It is difficult to determine why these forecasts are poor. More research into these methods and the optimal number of components to use to forecast inflation is an avenue for future research especially understanding how the models capture different aspects (persistence, volatility, and comovement) of the data.

The forecast comparison exercise carried out in this paper can alternatively be interpreted in the context of a signal extraction problem. To varying degrees, each component price series contains information that is useful for predicting the movement in headline inflation over some forecast horizon. The challenge is to find the weights that "best" separate the signal from the noise. Measures of core inflation such as "less food and energy" are crude attempts to separate signal from noise. The weighting schemes we examined here are more flexible than the zero-one weights, and therefore they are better able to capture the signal contained in the component series. The optimal weights, that is the weights which best capture the signal for future movements in headline inflation, are necessarily identified via a horse race exercise such as the one we conduct here. But although our set of comparison weighting schemes is quite large, there are likely other possibilities to consider in future research. A possible promising avenue for further research is to treat the individual series as separate forecasts of headline inflation and appeal to the optimal forecast combination literature to help identify optimal (from a forecasting perspective) weighting schemes (see for example, Diebold and Shin (2018)).

References

Alvarez, Luis J. and M. de los Llano Matea. (1999). "Underlying inflation measures in Spain." In *Measures of Underlying Inflation and their Role in the Conduct of Monetary Policy.* Proceeding of the Workshop of Central Bank Model Builders, Bank of International Settlements, Basel. February.

Apel, Mikael and Per Jansson. (1999). "A parametric approach for estimating core inflation and interpreting the inflation process." Working Paper S-103. Sveriges Riksbank, Stockholm.

Atkeson, A., & Ohanian, L. E. (2001). Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis. Quarterly Review-Federal Reserve Bank of Minneapolis*, *25*(1), 2.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191-221.

Ball, L. & Mazumder, S. (2011). "Inflation Dynamics and the Great Recession," *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, vol. 42(1 (Spring), pages 337-405.

Bermingham, C., & D'Agostino, A. (2014). Understanding and forecasting aggregate and disaggregate price dynamics. *Empirical Economics*, *46*(2), 765-788.

Blinder, Alan S. (1997). "Commentary" on "Measuring short-run inflation for central bankers." *Review*. Federal Reserve Bank of St. Louis, May/June.

Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis?. *Journal of Econometrics*, *132*(1), 169-194.

Bryan, Michael F. and Stephen G. Cecchetti. (1994). "Measuring core inflation." In *Monetary Policy*, ed. N. Gregory Mankiw, 195-215. Chicago: University of Chicago Press.

Clark, Todd .E. (2006). Disaggregate evidence on the persistence of consumer price inflation. *Journal of Applied Econometrics*, *21*(5), pp.563-587.

Cockerell, Lynne. (1999). "Measures of inflation and inflation targeting in Australia." In *Measures of Underlying Inflation and their Role in the Conduct of Monetary Policy.* Proceeding of the Workshop of Central Bank Model Builders, Bank of International Settlements, Basel. February.

Cogley, Timothy. (2002). "A simple adaptive measure of core inflation." *Journal of Money Credit and Banking.* 34:94-113.

Colmbra, C. Neves, P. D. (1997). “Trend Inflation Indicators,” Banco de Portugal, Economic Bulletin, March.

Crone, T. M., Khettry, N. N. K., Mester, L. J., & Novak, J. A. (2013). Core measures of inflation as predictors of total inflation. *Journal of Money, Credit and Banking*, *45*(2-3), 505-519.

Cutler, J. (2001). “Core inflation in the UK.” External MPC Unit Discussion Paper No. 3. London: Bank of England, March.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *20*(1), 134-144.

Diebold, F. X., & Shin, M. (2018). Egalitarian LASSO for Combining Economic Forecasts. https://www.sas.upenn.edu/~fdiebold/papers2/DieboldShinEgalitarianLasso2.pdf

Dolmas, J. (2009). Excluding items from personal consumption expenditures inflation. *Staff Papers*, Federal Reserve Bank of Dallas. (Jun).

Gamber, E. N., Liebner, J. P., & Smith, J. K. (2016). Inflation persistence: revisited. *International Journal of Monetary Economics and Finance*, *9*(1), 25-44.

Hendry, D. and Hubrich, K. (2006). “Forecasting economic aggregates by disaggregates.” European Central Bank Working Paper. No. 589.

Hendry, D. and Hubrich, K. (2011) "Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate." *Journal of Business & Economic Statistics*.

Johnson, M. (1999). “Core inflation: A measure of inflation for policy purposes.” In *Measures of Underlying Inflation and their Role in the Conduct of Monetary Policy.* Proceeding of the Workshop of Central Bank Model Builders, Bank of International Settlements, Basel. February.

Le Bihan, H. and Sedillot, F. (2002). “Implementing and interpreting indicators of core inflation: the case of France.” *Empirical Economics.* 27:473-497.

Levin, A., C.-F. Lin, and C.-S. J. Chu. (2002). “Unit root tests in panel data: Asymptotic and finite-sample properties.” *Journal of Econometrics* 108: 1–24.

Lucas, R. E. (1976, January). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy* (Vol. 1, pp. 19-46). North-Holland.

Mankiw, N. G., and Reis, R. (2003). “What Measure of Inflation Should A Central Bank Target,” *Journal of the European Economic Association*, vol. 1, issue 5, September, pp. 1058-1086.

Maria, J.R. (2004). “On the Use of the First Principal Component as a Core Inflation Indicator,” Economic Research Department Working Paper, No. 3-04 (Lisbon: Banco de Portugal)

Marques, Carlos Robalo, Pedro Duarte Neves and Afonso Goncalves da Silva. (2001). “Using the First Principal Component As a Core Inflation Indicator,” Banco de Portugal, Economic

Smith, J. K. (2004). “Weighted median inflation: Is this core inflation?” *Journal of Money Credit and Banking*. 36:253-263.

Smith, J. K. (2005). Inflation targeting and core inflation. *Canadian Journal of Economics/Revue canadienne d'économique*, *38*(3), 1018-1036.

Stock, J. H., & Watson, M. W. (2007). “Why has US inflation become harder to forecast?” *Journal of Money, Credit and banking*, *39*(s1), 3-33.

______________________________. (2016). Core inflation and trend inflation. *Review of Economics and Statistics*, *98*(4), 770-784.

Table 1a: 12-month forecast horizon (estimation sample 1960-2015)

	50 components					17 components				
	RMSE	**Ratio**	**DM test (t-stat)**	**MSFE**	**Std. error**	**RMSE**	**Ratio**	**DM test (t-stat)**	**MSFE**	**Std. error**
1984-2014										
lagged headline (benchmark)	1.19			1.42	(0.33)	1.19			1.42	(0.33)
lagged less food and energy	1.05	0.881	-1.75	1.10	(0.21)	1.05	0.881	-1.75	1.10	(0.21)
lagged weighted median	1.18	0.989	-0.21	1.39	(0.26)	1.16	0.971	-0.55	1.34	(0.26)
lagged trimmed mean	1.15	0.967	-0.74	1.33	(0.27)	1.13	0.951	-1.17	1.29	(0.27)
disaggregated (regression)	1.03	0.861	-1.23	1.05	(0.12)	1.24	1.043	0.69	1.55	(0.32)
persistence weighted	**1.09**	**0.911**	**-4.35****	**1.18**	**(0.31)**	*1.38*	*1.158*	*3.62***	*1.91*	*(0.42)*
persistence weighted*budget shares	1.21	1.015	0.95	1.46	(0.36)	1.20	1.004	0.34	1.43	(0.35)
median weighted by persistence	1.10	0.926	-1.93	1.22	(0.26)	1.14	0.954	-1.17	1.29	(0.28)
disaggregated persistence weighted	1.03	0.868	-1.21	1.07	(0.13)	1.26	1.057	1.01	1.59	(0.31)
principal components	**1.06**	**0.892**	**-2.75****	**1.13**	**(0.26)**	**0.90**	**0.756**	**-3.54****	**0.81**	**(0.17)**
1984-1993										
lagged headline (benchmark)	1.05			1.10	(0.27)	1.05			1.10	(0.27)
lagged less food and energy	1.15	1.094	1.12	1.32	(0.34)	1.15	1.094	1.12	1.32	(0.34)
lagged weighted median	1.14	1.083	1.45	1.30	(0.32)	1.20	1.143	1.74	1.44	(0.39)
lagged trimmed mean	1.09	1.036	0.74	1.19	(0.30)	1.11	1.057	1.26	1.24	(0.30)
disaggregated (regression)	1.16	1.105	0.86	1.35	(0.18)	1.21	1.148	1.14	1.46	(0.34)
persistence weighted	**0.87**	**0.824**	**-3.47****	**0.75**	**(0.22)**	1.00	0.949	-0.63	0.99	(0.26)
persistence weighted*budget shares	1.06	1.004	0.19	1.11	(0.30)	1.04	0.990	-0.55	1.08	(0.28)
median weighted by persistence	1.01	0.958	-0.86	1.01	(0.27)	0.91	0.868	-1.80	0.83	(0.21)
disaggregated persistence weighted	1.15	1.092	0.79	1.32	(0.20)	1.19	1.129	1.04	1.41	(0.32)
principal components	0.97	0.924	-1.25	0.94	(0.24)	**0.78**	**0.740**	**-3.41****	**0.61**	**(0.14)**
1994-2007										
lagged headline (benchmark)	0.93			0.87	(0.16)	0.93			0.87	(0.16)
lagged less food and energy	**0.79**	**0.850**	**-2.44****	**0.63**	**(0.10)**	**0.79**	**0.850**	**-2.44****	**0.63**	**(0.10)**
lagged weighted median	1.05	1.122	1.51	1.10	(0.24)	0.98	1.047	0.78	0.95	(0.19)
lagged trimmed mean	0.99	1.058	0.87	0.98	(0.22)	0.95	1.019	0.33	0.90	(0.19)
disaggregated (regression)	0.78	0.833	-1.64	0.60	(0.10)	0.98	1.050	0.39	0.96	(0.21)
persistence weighted	**0.86**	**0.922**	**-2.19***	**0.74**	**(0.14)**	*1.25*	*1.342*	*4.45***	*1.57*	*(0.27)*
persistence weighted*budget shares	0.96	1.027	1.84	0.92	(0.17)	*0.96*	*1.024*	*2.23**	*0.91*	*(0.17)*
median weighted by persistence	0.92	0.990	-0.25	0.85	(0.18)	1.04	1.116	1.87	1.08	(0.23)
disaggregated persistence weighted	0.79	0.845	-1.65	0.62	(0.09)	1.02	1.096	0.87	1.05	(0.20)
principal components	**0.85**	**0.915**	**-2.21***	**0.73**	**(0.13)**	**0.79**	**0.846**	**-2.82****	**0.62**	**(0.11)**
1984-2007										
lagged headline (benchmark)	0.98			0.97	(0.15)	0.98			0.97	(0.15)
lagged less food and energy	0.96	0.974	-0.46	0.92	(0.17)	0.96	0.974	-0.46	0.92	(0.17)
lagged weighted median	1.09	1.104	1.93	1.18	(0.21)	1.08	1.094	1.73	1.16	(0.21)
lagged trimmed mean	1.03	1.048	1.07	1.06	(0.19)	1.02	1.037	0.96	1.04	(0.18)
disaggregated (regression)	0.96	0.972	-0.34	0.91	(0.11)	1.08	1.098	1.01	1.17	(0.20)
persistence weighted	**0.86**	**0.877**	**-3.74****	**0.74**	**(0.13)**	*1.15*	*1.172*	*2.66***	*1.33*	*(0.21)*
persistence weighted*budget shares	1.00	1.016	1.13	1.00	(0.17)	0.99	1.008	0.70	0.98	(0.16)
median weighted by persistence	0.96	0.975	-0.77	0.92	(0.16)	0.99	1.005	0.10	0.98	(0.17)
disaggregated persistence weighted	0.96	0.971	-0.37	0.91	(0.11)	1.09	1.112	1.29	1.20	(0.19)
principal components	**0.90**	**0.919**	**-2.22***	**0.82**	**(0.14)**	**0.78**	**0.798**	**-4.13****	**0.62**	**(0.09)**

Notes: The column labeled "ratio" shows the RMSE (root mean squared error) of each inflation measure listed in column 1 divided by the benchmark (lagged headline) inflation rate; the column labeled "MSFE" is mean square forecast error with Newey-West standard errors in parentheses. The Diebold-Mariano (DM) statistic is constructed under the null hypothesis that the RMSE for the inflation measure listed in column 1 is equal to the RMSE of the benchmark inflation rate; * indicates significance at 5% and ** indicates significance at 1%; Bold indicates significantly better than the benchmark and italics indicated significantly worse than benchmark.

Table 1b: 6-month forecast horizon (estimation sample 1960-2015)

	50 components					17 components				
	RMSE	Ratio	DM test (t-stat)	MSFE	Std. error	RMSE	Ratio	DM test (t-stat)	MSFE	Std. error
1984-2014										
lagged headline (benchmark)	1.42			2.03	(0.68)	1.42			2.03	(0.68)
lagged less food and energy	1.28	0.900	-1.53	1.64	(0.47)	1.28	0.900	-1.53	1.64	(0.47)
lagged weighted median	1.36	0.953	-0.83	1.84	(0.53)	1.36	0.953	-0.84	1.85	(0.54)
lagged trimmed mean	1.33	0.934	-1.39	1.77	(0.55)	1.33	0.930	-1.56	1.76	(0.56)
disaggregated (regression)	1.26	0.882	-1.27	1.58	(0.39)	1.42	0.994	-0.15	2.01	(0.65)
persistence weighted	**1.36**	**0.955**	**-2.59****	**1.85**	**(0.64)**	*1.61*	*1.133*	*3.59***	*2.61*	*(0.79)*
persistence weighted*budget shares	1.43	1.006	0.40	2.06	(0.72)	1.42	0.999	-0.06	2.03	(0.70)
median weighted by persistence	**1.31**	**0.918**	**-2.06***	**1.71**	**(0.56)**	1.34	0.938	-1.59	1.79	(0.58)
disaggregated persistence weighted	1.25	0.880	-1.30	1.57	(0.39)	1.43	1.002	0.05	2.04	(0.65)
principal components	**1.30**	**0.915**	**-2.28***	**1.70**	**(0.59)**	**1.16**	**0.817**	**-2.67****	**1.35**	**(0.44)**
1984-1993										
lagged headline (benchmark)	1.16			1.35	(0.36)	1.16			1.35	(0.36)
lagged less food and energy	1.21	1.042	0.51	1.47	(0.42)	1.21	1.042	0.51	1.47	(0.42)
lagged weighted median	1.19	1.026	0.40	1.42	(0.41)	1.28	1.105	1.08	1.65	(0.51)
lagged trimmed mean	1.14	0.985	-0.29	1.31	(0.38)	1.18	1.016	0.30	1.40	(0.40)
disaggregated (regression)	1.15	0.985	-0.22	1.31	(0.29)	1.21	1.042	0.41	1.47	(0.39)
persistence weighted	**1.05**	**0.899**	**-2.44****	**1.09**	**(0.31)**	1.19	1.026	0.36	1.42	(0.34)
persistence weighted*budget shares	1.16	1.001	0.03	1.35	(0.38)	1.15	0.991	-0.52	1.33	(0.37)
median weighted by persistence	1.10	0.946	-1.05	1.21	(0.36)	1.01	0.868	-1.72	1.02	(0.26)
disaggregated persistence weighted	1.11	0.958	-0.61	1.24	(0.27)	1.17	1.007	0.07	1.37	(0.33)
principal components	1.06	0.914	-1.51	1.13	(0.31)	**0.91**	**0.779**	**-3.01****	**0.82**	**(0.21)**
1994-2007										
lagged headline (benchmark)	0.94			0.89	(0.18)	0.94			0.89	(0.18)
lagged less food and energy	0.87	0.923	-1.21	0.76	(0.13)	0.87	0.923	-1.21	0.76	(0.13)
lagged weighted median	1.07	1.136	1.60	1.15	(0.29)	1.00	1.061	0.94	1.01	(0.23)
lagged trimmed mean	0.99	1.047	0.63	0.98	(0.25)	0.96	1.012	0.20	0.91	(0.22)
disaggregated (regression)	0.88	0.937	-0.82	0.78	(0.15)	1.01	1.069	0.71	1.02	(0.21)
persistence weighted	0.90	0.958	-1.07	0.82	(0.15)	*1.25*	*1.325*	*4.24***	*1.57*	*(0.29)*
persistence weighted*budget shares	0.96	1.020	1.40	0.93	(0.19)	0.96	1.020	1.74	0.93	(0.19)
median weighted by persistence	0.92	0.974	-0.80	0.85	(0.18)	1.04	1.104	1.59	1.09	(0.24)
disaggregated persistence weighted	0.89	0.938	-0.86	0.79	(0.15)	1.04	1.103	1.16	1.09	(0.20)
principal components	0.90	0.953	-1.09	0.81	(0.15)	0.86	0.909	-1.54	0.74	(0.14)
1984-2007										
lagged headline (benchmark)	1.04			1.08	(0.19)	1.04			1.08	(0.19)
lagged less food and energy	1.03	0.987	-0.24	1.05	(0.21)	1.03	0.987	-0.24	1.05	(0.21)
lagged weighted median	1.12	1.080	1.42	1.26	(0.25)	1.13	1.084	1.33	1.27	(0.27)
lagged trimmed mean	1.06	1.015	0.32	1.12	(0.23)	1.06	1.014	0.34	1.11	(0.22)
disaggregated (regression)	1.00	0.962	-0.74	1.00	(0.16)	1.10	1.055	0.75	1.21	(0.21)
persistence weighted	**0.97**	**0.928**	**-2.41****	**0.93**	**(0.16)**	*1.23*	*1.179*	*3.11***	*1.51*	*(0.23)*
persistence weighted*budget shares	1.05	1.010	0.72	1.11	(0.20)	1.05	1.005	0.43	1.09	(0.20)
median weighted by persistence	1.00	0.959	-1.26	1.00	(0.19)	1.03	0.989	-0.20	1.06	(0.18)
disaggregated persistence weighted	0.99	0.949	-1.06	0.98	(0.15)	1.10	1.054	0.80	1.20	(0.19)
principal components	0.97	0.933	-1.809	0.94	(0.17)	**0.88**	**0.844**	**-3.12****	**0.77**	**(0.12)**

Notes: The column labeled "ratio" shows the RMSE (root mean squared error) of each inflation measure listed in column 1 divided by the benchmark (lagged headline) inflation rate; the column labeled "MSFE" is mean square forecast error with Newey-West standard errors in parentheses. The Diebold-Mariano (DM) statistic is constructed under the null hypothesis that the RMSE for the inflation measure listed in column 1 is equal to the RMSE of the benchmark inflation rate; * indicates significance at 5% and ** indicates significance at 1%; Bold indicates significantly better than the benchmark and italics indicated significantly worse than benchmark.

Table 1c: 24-month forecast horizon (estimation sample 1960-2015)

	50 components					17 components				
	RMSE	**Ratio**	**DM test (t-stat)**	**MSFE**	**Std. error**	**RMSE**	**Ratio**	**DM test (t-stat)**	**MSFE**	**Std. error**
1984-2014										
lagged headline (benchmark)	1.17			1.37	(0.21)	1.17			1.37	(0.21)
lagged less food and energy	**1.02**	**0.869**	**-2.31***	**1.04**	**(0.15)**	**1.02**	**0.869**	**-2.31***	**1.04**	**(0.15)**
lagged weighted median	1.21	1.036	0.78	1.48	(0.20)	1.18	1.007	0.15	1.39	(0.19)
lagged trimmed mean	1.18	1.005	0.13	1.39	(0.19)	1.16	0.987	-0.33	1.34	(0.19)
disaggregated (regression)	1.04	0.884	-1.33	1.07	(0.11)	1.25	1.065	0.84	1.56	(0.18)
persistence weighted	**1.03**	**0.878**	**-5.91****	**1.06**	**(0.19)**	1.31	1.121	2.76**	1.73	(0.28)
persistence weighted*budget shares	1.19	1.018	1.27	1.42	(0.24)	1.18	1.005	0.39	1.39	(0.23)
median weighted by persistence	1.11	0.945	-1.68	1.23	(0.18)	1.13	0.966	-0.97	1.28	(0.20)
disaggregated persistence weighted	1.03	0.881	-1.42	1.07	(0.11)	1.25	1.070	0.98	1.57	(0.18)
principal components	**1.03**	**0.878**	**-3.40****	**1.06**	**(0.17)**	**0.87**	**0.739**	**-4.94****	**0.75**	**(0.10)**
1984-1993										
lagged headline (benchmark)	1.18			1.40	(0.24)	1.18			1.40	(0.24)
lagged less food and energy	1.28	1.079	1.34	1.63	(0.28)	1.28	1.079	1.34	1.63	(0.28)
lagged weighted median	*1.30*	*1.095*	*2.02**	*1.68*	*(0.30)*	1.34	1.128	2.58**	1.78	(0.30)
lagged trimmed mean	1.24	1.049	1.22	1.54	(0.28)	1.26	1.062	1.77	1.58	(0.27)
disaggregated (regression)	1.19	1.005	0.05	1.42	(0.19)	1.24	1.044	0.39	1.53	(0.28)
persistence weighted	**0.94**	**0.791**	**-5.62****	**0.88**	**(0.20)**	**1.02**	**0.862**	**-2.47****	**1.04**	**(0.22)**
persistence weighted*budget shares	1.20	1.010	0.64	1.43	(0.26)	1.18	0.999	-0.07	1.40	(0.25)
median weighted by persistence	1.13	0.955	-1.14	1.28	(0.24)	1.06	0.893	-2.51**	1.12	(0.23)
disaggregated persistence weighted	1.18	0.992	-0.08	1.38	(0.19)	1.21	1.025	0.24	1.47	(0.26)
principal components	1.10	0.932	-1.34	1.22	(0.23)	**0.93**	**0.782**	**-4.53****	**0.86**	**(0.15)**
1994-2007										
lagged headline (benchmark)	0.97			0.94	(0.17)	0.97			0.94	(0.17)
lagged less food and energy	**0.77**	**0.795**	**-3.61****	**0.60**	**(0.11)**	**0.77**	**0.795**	**-3.61****	**0.60**	**(0.11)**
lagged weighted median	*1.17*	*1.200*	*3.24***	*1.36*	*(0.25)*	1.09	1.124	2.66**	1.19	(0.22)
lagged trimmed mean	*1.11*	*1.141*	*2.82***	*1.23*	*(0.23)*	1.07	1.097	2.12*	1.14	(0.22)
disaggregated (regression)	0.93	0.960	-0.35	0.87	(0.15)	1.19	1.227	1.52	1.42	(0.25)
persistence weighted	**0.87**	**0.896**	**-3.13****	**0.76**	**(0.13)**	1.25	1.290	4.07**	1.57	(0.26)
persistence weighted*budget shares	*1.01*	*1.036*	*2.91***	*1.01*	*(0.18)*	1.00	1.028	3.36**	1.00	(0.18)
median weighted by persistence	1.02	1.047	1.63	1.04	(0.20)	1.10	1.137	2.78**	1.22	(0.24)
disaggregated persistence weighted	0.93	0.954	-0.43	0.86	(0.14)	1.20	1.239	1.77	1.45	(0.23)
principal components	**0.85**	**0.873**	**-3.39****	**0.72**	**(0.13)**	**0.77**	**0.788**	**-3.98****	**0.59**	**(0.09)**
1984-2007										
lagged headline (benchmark)	1.07			1.14	(0.15)	1.07			1.14	(0.15)
lagged less food and energy	1.01	0.952	-1.02	1.03	(0.16)	1.01	0.952	-1.02	1.03	(0.16)
lagged weighted median	*1.22*	*1.147*	*3.59***	*1.49*	*(0.21)*	*1.20*	*1.126*	*3.50***	*1.44*	*(0.20)*
lagged trimmed mean	*1.17*	*1.095*	*2.77***	*1.36*	*(0.20)*	*1.15*	*1.079*	*2.61***	*1.32*	*(0.19)*
disaggregated (regression)	1.05	0.983	-0.20	1.10	(0.12)	1.21	1.136	1.37	1.47	(0.20)
persistence weighted	**0.90**	**0.844**	**-5.39****	**0.81**	**(0.12)**	1.16	1.091	1.60	1.35	(0.19)
persistence weighted*budget shares	*1.09*	*1.023*	*2.10**	*1.19*	*(0.17)*	1.08	1.013	1.66	1.17	(0.16)
median weighted by persistence	1.07	1.001	0.04	1.14	(0.16)	1.09	1.018	0.48	1.18	(0.18)
disaggregated persistence weighted	1.04	0.974	-0.34	1.08	(0.12)	1.21	1.134	1.47	1.46	(0.18)
principal components	**0.96**	**0.904**	**-2.88****	**0.93**	**(0.14)**	**0.84**	**0.785**	**-5.66****	**0.70**	**(0.09)**

Notes: The column labeled "ratio" shows the RMSE (root mean squared error) of each inflation measure listed in column 1 divided by the benchmark (lagged headline) inflation rate; the column labeled "MSFE" is mean square forecast error with Newey-West standard errors in parentheses. The Diebold-Mariano (DM) statistic is constructed under the null hypothesis that the RMSE for the inflation measure listed in column 1 is equal to the RMSE of the benchmark inflation rate; * indicates significance at 5% and ** indicates significance at 1%; Bold indicates significantly better than the benchmark and italics indicated significantly worse than benchmark.

Table 2: 50 components comparison of persistence weighted and principal components

Samples	1984-2015	1984-1993	1994-2007	1984-2007
12 month ahead forecast horizon				
Is persistence weighted significantly different than principal components	No	No	No	No
6 month ahead forecast horizon				
Is persistence weighted significantly different than principal components	No	No	No	No
24 month ahead forecast horizon				
Is persistence weighted significantly different than principal components	No	Yes PW statistically better than PC.	No	No

Notes: statistical significance is based on the Diebold Mariano test; PW: persistence weighted PC: principal components.

Table 3: Comparison of persistence weighted and principal components forecasts at 12-month time horizon

	RMSE	MSFE	Std. error
Persistence weighted-50 components	1.09	1.18	(0.31)
Principal components-50 components	1.06	1.13	(0.26)
Persistence weighted-17 components	1.38	1.91	(0.42)
Principal components-17 components	0.90	0.81	(0.17)

	DM test stat		Ratio	Interpretation
PW 50 vs PW 17 (17 benchmark)	-4.97	**	0.79	PW 50 is better than PW 17
PC 50 vs 17 (17 benchmark)	3.21	**	1.18	PC 17 is better than PC 50
PC 50 vs PW 17 (PW 17 benchmark)	-3.65	**	0.77	PC 50 is better than PW 17
PW 50 vs PC 17 (PC 17 benchmark)	2.47	**	1.20	PC 17 is better than PW 50

Notes: the results shown are for head-to-head comparison of forecast performance of the persistence weighted (PW) and principal components (PC) measures of inflation. There are two each of PW and PC measures thus four head-to-head comparisons. The Diebold-Mariano (DM) test stat listed in the bottom portion of the table was constructed under the null hypothesis that the RMSEs in each pairwise comparison are equal; * indicates significance at 5% and ** indicates significance at 1%.

Table 4: Relationship Between Component Level Price Inflation and Estimated Principal Component

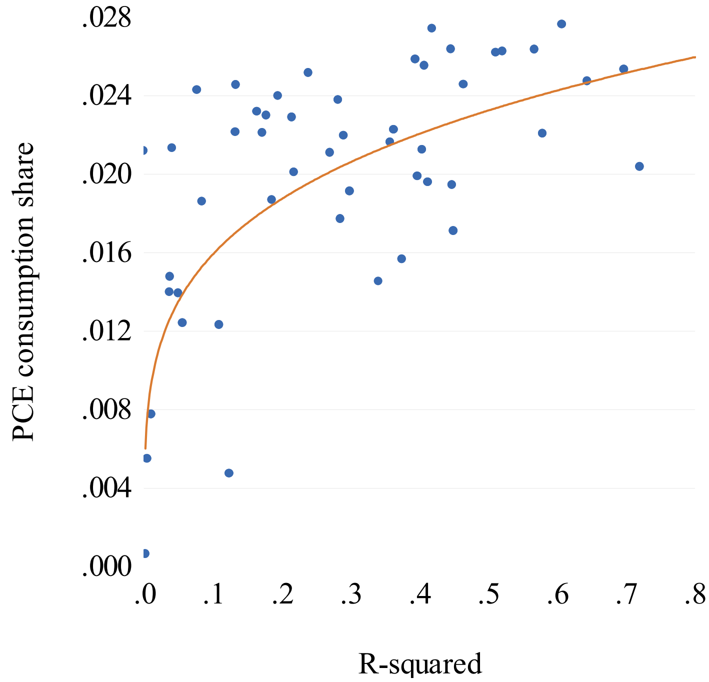
		PCE expenditure category	R-squared: 17-component	R-squared: 50-sub components	R-squared: average R-squared of sub-categories
1		**clothing and footwear**	0.31		
	1	Garments		0.29	
	2	Other clothing materials and footwear		0.22	0.25
2		**food and beverages purchased for off-premises consumption**	0.23		
	3	Food and nonalcoholic beverages purchased for off-premises consumption		0.19	
	4	Alcoholic beverages purchased for off-premises consumption		0.29	
	5	Food produced and consumed on farms		0.00	0.16
3		**motor vehicles and parts**	0.10		
	6	New motor vehicles		0.18	
	7	Used autos		0.01	
	8	Motor vehicle parts and accessories		0.00	0.06
4		**furnishings and durable household equipment**	0.46		
	9	Furniture and furnishings		0.45	
	10	Household appliances		0.13	
	11	Glassware, tableware, and household utensils		0.28	
	12	Tools and equipment for house and garden		0.04	0.23
5		**recreational goods and vehicles**	0.62		
	13	Sporting equipment, supplies, guns, and ammunition		0.40	
	14	Sports and recreational vehicles		0.27	
	15	Video, audio, photographic, and information processing equipment and media		0.41	
	16	Recreational books		0.30	
	17	Musical instruments		0.13	0.30
6		**recreation services**	0.66		
	18	Membership clubs, sports centers, parks, theaters, and museums		0.36	
	19	Audio-video, photographic, and information processing equipment services		0.38	
	20	Gambling		0.61	
	21	Other recreational services		0.20	0.39
7		**transportation services**	0.43		
	22	Motor vehicle services		0.42	
	23	Public transportation		0.09	0.25
8		**final consumption expenditures of nonprofits serving households**	0.35		0.35
	24	Final consumption expenditures of nonprofit institutions serving households			
9		**financial services and insurance**	0.01		
	25	Financial services		0.04	
	26	Insurance		0.12	0.08
10		**other nondurable goods**	0.42		
	27	Pharmaceutical and other medical products		0.45	
	28	Recreational items		0.65	
	29	Household supplies		0.47	
	30	Personal care products		0.52	
	31	Tobacco		0.01	
	32	Magazines, newspapers, and stationery		0.17	0.38

Table 4 (continued)

		PCE expenditure category	17-component	50-sub components	average R-squared of sub-categories
11		**gasoline and other energy goods**	0.03		
	33	Gasoline and other energy goods		0.04	0.04
12		**food services and accommodations**	0.23		
	34	Food services		0.51	
	35	Accommodations		0.22	0.36
13		**healthcare**			
	36	Outpatient services	0.84	0.70	
	37	Hospital and nursing home services		0.72	0.71
14		**other services**	0.42		
	38	Communication		0.11	
	39	Education services		0.45	
	40	Professional and other services		0.41	
	41	Personal care and clothing services		0.57	
	42	Social services and religious activities		0.41	
	43	Household maintenance		0.17	
	44	Foreign travel by U.S. residents		0.05	
	45	Less: Expenditures in the United States by nonresidents		0.40	0.32
15		**electricity and gas**	0.05		0.05
16		**housing expenditures excluding gas and electric utilities**	0.44		
	46	Housing		0.36	
	47	Household utilities		0.08	0.22
17		**other durable goods**	0.55		
	48	Other durable goods		0.58	
	49	Expenditures abroad by U.S. residents		0.06	
	50	Less: Personal remittances in kind to nonresidents		0.24	0.29

Notes: The R-squared statistics were obtained by regressing each component-level inflation rate on the estimated principal component derived from the 17 and 50 components respectively. The final column reports the average of the estimate R-squared statistics for each grouping of sub-components.

Figure 1: PCE shares and R-squared statistics



Notes: this figure shows the positive relationship between the consumption shares for the 50-sub-components of the PCE and the R-squared from the regression of sub-component level inflation on the common factor (first principal component).

Figure 2a: Forecast comparison of persistence weighted and principal components:
1984.01 to 2014.12, 50 components

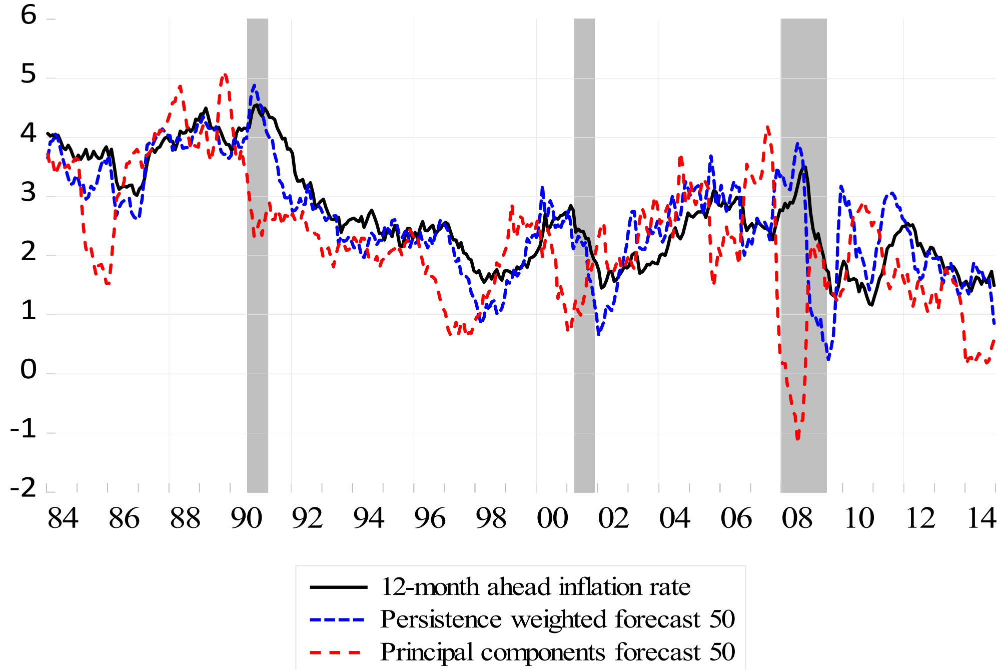
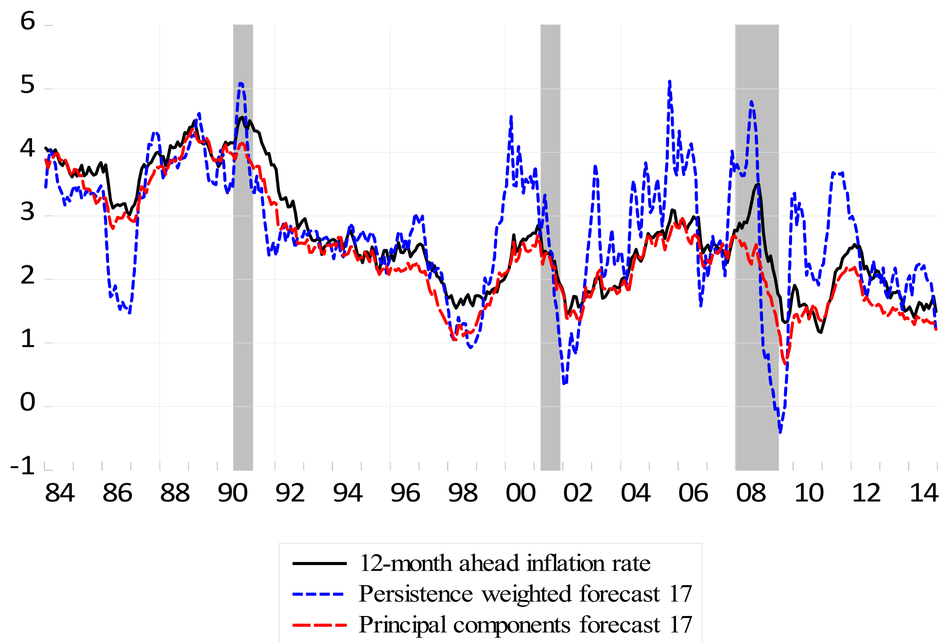
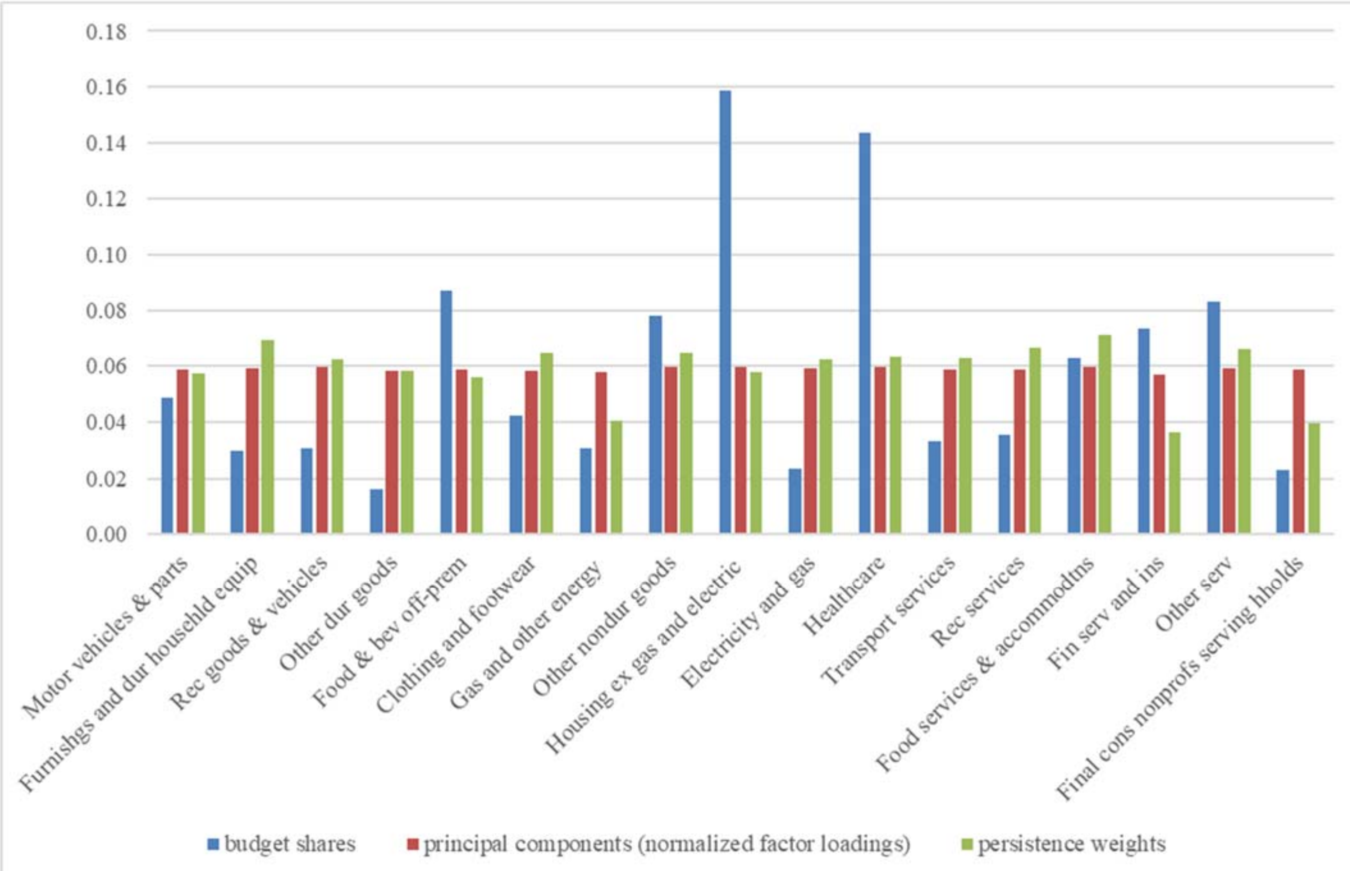


Figure 2b: Forecast comparison of persistence weighted and principal components:
1984.01 to 2014.12, 17 components



Notes: The figures displays the forecasts produced with the 50 and 17 component persistence weighted and principal component inflation rates. The actual 12-month ahead inflation rate is shown in black. The shaded regions denote NBER-dated recessions.

Figure 3: Comparison of component weights (17 components)



Notes: each bar represents the average weight over the sample 1984.01-2014.12.

Appendix A: List of 50 Components

1. New motor vehicles
2. Used autos
3. Motor vehicle parts and accessories
4. Furniture and furnishings
5. Household appliances
6. Glassware, tableware, and household utensils
7. Tools and equipment for house and garden
8. Video, audio, photographic, and information processing equipment and media
9. Sporting equipment, supplies, guns, and ammunition
10. Sports and recreational vehicles
11. Recreational books
12. Musical instruments
13. Other durable goods
14. Food and nonalcoholic beverages purchased for off-premises consumption
15. Alcoholic beverages purchased for off-premises consumption
16. Food produced and consumed on farms
17. Garments
18. Other clothing materials and footwear
19. Gasoline and other energy goods
20. Pharmaceutical and other medical products
21. Recreational items
22. Household supplies
23. Personal care products
24. Tobacco
25. Magazines, newspapers, and stationery
26. Expenditures abroad by U.S. residents
27. Less: Personal remittances in kind to nonresidents
28. Housing
29. Household utilities
30. Outpatient services
31. Hospital and nursing home services
32. Motor vehicle services
33. Public transportation
34. Membership clubs, sports centers, parks, theaters, and museums
35. Audio-video, photographic, and information processing equipment services
36. Gambling
37. Other recreational services
38. Food services
39. Accommodations
40. Financial services
41. Insurance
42. Communication
43. Education services
44. Professional and other services
45. Personal care and clothing services
46. Social services and religious activities
47. Household maintenance
48. Foreign travel by U.S. residents
49. Less: Expenditures in the United States by nonresidents
50. Final consumption expenditures of nonprofit institutions serving households

Appendix B: List of 17 Components

1. Motor vehicles and parts
2. Furnishings and durable household equipment
3. Recreational goods and vehicles
4. Other durable goods
5. Food and beverages purchased for off-premises consumption
6. Clothing and footwear
7. Gasoline and other energy goods
8. Other nondurable goods
9. Housing expenditures excluding gas and electric utilities
10. Electricity and gas
11. Healthcare
12. Transportation services
13. Recreation services
14. Food services and accommodations
15. Financial services and insurance
16. Other services
17. Final consumption expenditures of nonprofits serving households

Appendix C: Names and brief descriptions of models

1. Lagged headline (benchmark)	Traditional benchmark model based on Atkeson and Ohanian (2001)
2. Weighted median	Use relative importance weights to weight component inflation rates and select median component each month based on Bryan and Cecchetti (1997)
3. Trimmed mean	Use relative importance weights to weight component inflation rates and average the components left in the center of the distribution each month based on Dolmas (2009)
4. Less food and energy	Traditional core inflation measure
5. Persistence weights	Use an AR(1) process to construct persistence of each component and then weight each component by the normalized persistence to construct an aggregate persistence weighted inflation rate based on Cutler (2001)
6. Disaggregated (regression) weights	Use all components on right-hand side of the equation and estimate the coefficient on each component based on Hendry and Hubrich (2006)
7. Persistence and regression weights (disaggregated)	A combination of methods 5 and 6.
8. Persistence*budget share weights	A combination of methods 5 and 1. Lagged inflation is weighted by budget shares.
9. Weighted median (weighted by persistence)	Find the median component using persistence weights instead of relative importance weights
10. Principal component weights	Use principal components to find the factor loading for each component and use the first factor loading to weight the components based on Maria (2004)