ARTICLE IN PRESS

International Journal of Forecasting xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Improving inflation forecasts using robust measures[☆]

Randal Verbrugge[a,b,*], Saeed Zaman[a]

[a] Federal Reserve Bank of Cleveland, USA
[b] NBER/CRIW, USA

ARTICLE INFO

Keywords:
Median PCE inflation
Trimmed-mean PCE
Disaggregate inflation
Skewness
Forecasting

ABSTRACT

Theory and extant empirical evidence suggest that the cross-sectional asymmetry across disaggregated price indexes might be useful in forecasting aggregate inflation. Trimmed-mean inflation estimators have been shown to be useful devices for forecasting headline PCE inflation. But is this because they signal the underlying trend or implicitly signal asymmetry in the underlying distribution? We address this question by augmenting a "hard to beat" benchmark headline PCE inflation forecasting model with robust trimmed mean inflation measures and robust measures of the cross-sectional skewness, both computed using the 180+ components of the PCE price index. Our results indicate significant gains in the point and density accuracy of PCE inflation forecasts over medium- and long-term horizons, including the COVID-19 pandemic. Improvements in accuracy stem mainly from the trend information implicit in trimmed-mean estimators, but skewness information is also useful. Examining goods and services PCE inflation (using newly constructed trimmed mean and skewness measures of the same) provides similar inference.

1. Introduction

Inflation is a central concern of monetary policy. Accordingly, policymakers at most central banks monitor a range of inflation measures to come to an informed assessment of the underlying inflationary pressures. Over the past decade, increased attention has been paid to trimmed-mean inflation estimators,[1] as these provide signs of broad-based inflationary pressures or lack thereof (see Mertens, 2016; Verbrugge, 2022).

Recent research has documented the usefulness of trimmed-mean estimators in improving inflation forecasts from a variety of time-series models (e.g., Carroll & Verbrugge, 2019; Dolmas, 2005; Mertens, 2016; Meyer & Zaman, 2019; Ocampo, Schoenle, & Smith, 2022).[2] The consensus in the literature is that the superior performance of the trimmed-mean estimators in forecasting future inflation results from their ability to signal the trend in inflation. The main rationale behind this consensus is the following: when the underlying distribution is leptokurtic (fat-tailed), and the sample (i.e., the number of components or disaggregates used to compute the aggregate) is not large, as is the case for US inflation,[3] then

☆ We thank Todd Clark, Ed Knotek, Jean-Paul L'Huillier, Pierlauro Lopez, James Mitchell, Rob Rich, and conference participants at the International Institute of Forecasters' 38th International Symposium on Forecasting and the Canadian Economics Association Conference 2022. We also thank Editor Esther Ruiz, an Associate Editor, and two anonymous referees for very helpful comments and guidance. The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

* Correspondence to: Research Department, Federal Reserve Bank of Cleveland, P.O. Box 6387, Cleveland, OH 44101, USA.

E-mail addresses: randal.verbrugge@clev.frb.org (R. Verbrugge), saeed.zaman@clev.frb.org (S. Zaman).

1 In this paper, we refer to both median inflation measures (such as the median PCE) and trimmed-mean inflation measures (such as the trimmed-mean PCE) as trimmed-mean measures.

2 In related and contemporaneous work, Ocampo et al. (2022) show that trimmed mean estimators outperform core PCE, overall, but during periods when headline inflation is below 2.5 percent, core PCE performs better.

3 Technically, what matters is not the nominal number of components but rather, given the wide distribution of aggregation weights associated with the components, some appropriate notion of an *effective* number.

https://doi.org/10.1016/j.ijforecast.2023.05.003
0169-2070/

trimmed-mean estimators are likely to be more accurate estimates of central tendency, compared to the sample mean.

But there is an alternative or complementary explanation for the trimmed-mean estimators' superior predictive performance that has received little attention. In addition to being fat-tailed, as discussed in Section 2, the underlying distribution of inflation components (disaggregates) is also asymmetric, with the degree of asymmetry evolving slowly over time. Consequently, in forecasting models, when trimmed-mean estimators are added alongside headline inflation measures (as is commonly done), the differential between the two provides an implicit signal about the current degree of asymmetry in the underlying distribution of the components. Both theory and extant evidence, reviewed below, suggest that this signal may have notable predictive content. In this paper, we explore this hypothesis and determine the extent to which their implicit signal of asymmetry drives the superior forecasting performance of trimmed-mean estimators.

Accordingly, this paper examines the independent and joint predictive performance of trimmed-mean estimators and robust asymmetry (skewness) measures to forecast aggregate PCE inflation. Specifically, we make pairwise comparisons of forecast accuracy between univariate, bivariate, and tri-variate vector autoregressive (VAR) model specifications. In constructing our VAR model specifications, we build upon the "hard-to-beat" Faust and Wright (2013) model, which is a simple univariate autoregressive (AR) model in gaps, where the gap is defined as the difference between the inflation measure and long-run inflation expectations of PCE inflation. Our VAR models include a trimmed-mean inflation measure, a robust skewness statistic, and/or additional covariates. The pairwise comparisons between model specifications allow us to examine both the marginal contribution of skewness measures and trimmed-mean estimators and their joint contribution to potential improvements in the accuracy of PCE inflation forecasts (point and density) above and beyond the univariate AR model in gaps.

To complete our analysis and provide a broader perspective on forecasting performance, we also assess the accuracy of a model specification embedding the Phillips curve and a model with core PCE inflation. Finally, motivated by a growing literature exploring the predictive content of goods and services, we investigate the predictive content of robust goods and services measures.

Our main finding is that including robust measures in the AR benchmark forecasting model improves its ability to forecast aggregate PCE inflation. The statistically significant gains in the accuracy of both the point and the density forecasts are achieved for forecast horizons 1.5 years ahead and greater, which are the forecast horizons most relevant to monetary policymakers. Most of the improvements in accuracy are due to the ability of the trimmed-mean estimators to signal a trend, with only marginal improvements due to their implicit skewness signal. The statistically significant gains in accuracy are observed over periods when inflation is low, predominantly over the financial crisis and onward sample, including the COVID-19 pandemic period but prior to the inflation surge in mid-2021. Re-running our analysis separately on goods and services PCE inflation gives results largely consistent with the findings for headline PCE inflation. We also find that skewness is modestly effective in improving estimates of stochastic volatility, and our results point to a complex interplay of such forces as relative price shocks, inflation expectations, and monetary policy.

The paper is organized as follows. Section 2 describes the trimmed-mean inflation estimators, the skewness measures, and the data. Section 3 details the model specifications and the design of the forecasting exercise. Section 4 discusses the results. Section 5 explores the efficacy of skewness measures for estimating stochastic volatility. Section 6 concludes.

2. Robust measures and data

A price index is a stochastic process that is a complicated convolution of thousands of stochastic processes. For example, changes in the personal consumption expenditure price Index (PCE price index) are a weighted average of the changes in the indexes of over 180 commodities and services. The weights change over time, reflecting substitution patterns, entry and exit of goods and outlets, and so on.

The evolutions of the underlying stochastic processes are not independent. They reflect a variety of forces, such as monetary impulses, changes in transportation costs and supply constraints, transaction technologies and tastes, and productivity growth. They reflect price pressures on *groups* of goods and services. And they reflect highly idiosyncratic movements as well. Any of these influences could be transient or persistent.

One manifestation of the complexity of the evolution of the underlying price process is the cross-sectional distribution of disaggregated component price indexes. Fig. 1 depicts a histogram of the monthly inflation rates across 180+ components of the PCE price index for May 2018.

It is clear that these components experienced significantly different inflation rates in May 2018 and that there are some extreme outliers. The presence of such outliers and the sensitivity of the sample mean to outliers motivate a prominent approach to the estimation of trend inflation: the use of limited-influence inflation estimators, such as a median CPI or trimmed-mean CPI (see Bryan & Cecchetti, 1993 or a median PCE (see Carroll & Verbrugge, 2019) and trimmed-mean PCE (see Dolmas, 2005). Such measures appear to capture trend inflation in as much as they remove noise from inflation, track ex-post measures of its trend, and have been shown to improve inflation forecasting (see, e.g., Ball & Mazumder, 2011; Meyer & Zaman, 2019; Norman & Richards, 2012; Smith, 2004).

Fig. 1 also illustrates that not only is the cross-sectional distribution highly kurtotic, but it is also asymmetric — and typically left-skewed. Indeed, for this reason, the trimmed-mean PCE uses asymmetric trimming. In particular, to ensure that the trimmed-mean PCE price index is unbiased on average over long periods, 24 percent is trimmed from the lower tail, while 31 percent is trimmed from the upper tail (see Dolmas, 2005, 2009).

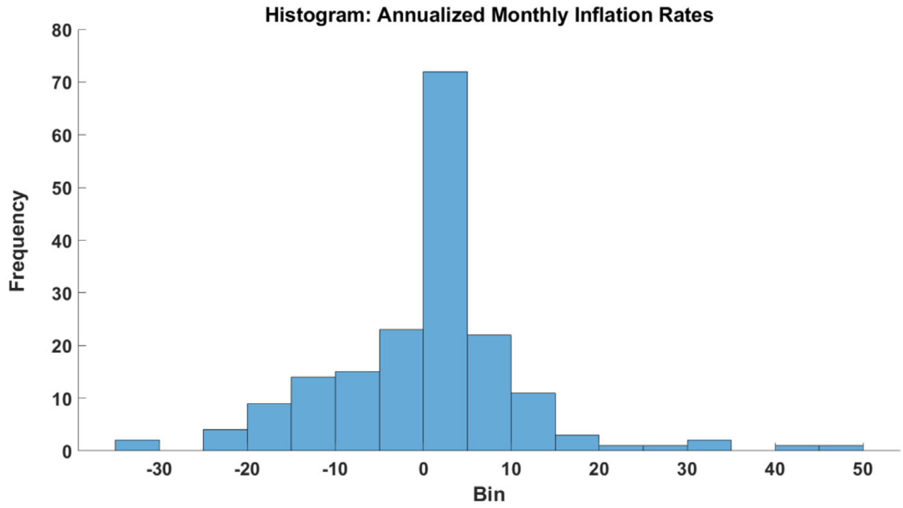

Fig. 1. Cross-sectional distribution of inflation in PCE price index components, May 2018.

However, the degree of asymmetry is not stable but changes over time. We illustrate this using two robust asymmetry estimators, Bowley skewness and Kelly skewness statistics, which we define below.[4]

Skewness Statistics: Bowley and Kelly

Following Kim and White (2004) and Dolmas (2005), we define the (weighted) Bowley and Kelly skewness:

$$Skewness\{m\} = \frac{P_a + P_b - 2P_c}{P_a - P_b} \tag{1}$$

where m refers to Bowley or Kelly skewness, and P_i is the ith percentile of the distribution of component price changes (in a given month), and we have suppressed time subscripts for clarity. When m refers to Bowley, then $a = 75$, $b = 25$, $c = 50$, and when m refers to Kelly, then $a = 90$, $b = 10$, $c = 50$.[5]

We compute *skewness* statistics for each month over the number of components available.[6] And for each skewness statistic, we calculate two measures: one based on disaggregate components' month-to-month (m-o-m) inflation rates and the other one based on those components' 12-month trailing inflation rates (y-o-y).

Fig. 2 plots Bowley and Kelly skewness measures from 1978 through June 2021 based on disaggregates' 12-month trailing inflation rates.[7] Figure A1, in the online appendix, plots the corresponding skewness measures based on disaggregates' m-o-m rates. Presented is the three-month moving average of these monthly skewness measures. Three observations stand out. First, asymmetry (skewness) displays significant medium-frequency variation. Second, most of the time, the skew is negative. Third, at times, the two measures of skewness (i.e., Bowley and Kelly) disagree with one another, especially when skewness measures are constructed using disaggregates' 12-month trailing rates. For example, in Fig. 2, between 2014 and 2018, the Kelly statistics indicate a strongly negative skew, whereas the Bowley statistics indicate periods in which the skew was positive.

Why might robust skewness measures have predictive content? There are four reasons. First, leading theories of price-setting behavior (e.g., Ball & Mankiw, 1994) indicate that inflation is linked to asymmetric price adjustment. Second, there is compelling statistical evidence that asymmetry correlates with inflation (e.g., Verbrugge, 1999). Third, as already noted, asymmetric trimming of component inflation rates is used in the trimmed mean PCE to deliver an inflation trend estimate that is unbiased *on average*. However, we demonstrated that the degree of asymmetry in the cross-sectional distribution of component inflation rates is time-varying. This implies that both the trimmed mean PCE and the median PCE have time-varying bias, with the bias in a given month related to that month's degree of asymmetry. Hence information about monthly asymmetry is likely useful for forecasting since it can provide a bias-correcting role.

Last, the time variation in asymmetry is informative about time variation in the properties of the convolution. Verbrugge (1999) indicates that asymmetry in the cross-sectional distribution is associated with the underlying conditional variance–covariance structure, which is time-varying. Accordingly, we hypothesize that a direct estimate of the asymmetry – an estimate that is a nonlinear

4 As Dolmas (2005) points out, robust asymmetry estimators are to be preferred, since moment estimators (such as the third centered moment) are all strongly influenced by outliers.

5 As implied by the formula, in its construction, the Bowley statistic uses observations in the middle 50 percent of the distribution; that is, it excludes 25 percent of the observations from each tail. Similarly, Kelly statistic uses observations in the middle 80 percent, excluding 10 percent of the observations from each tail.

6 Coverage of the PCE has increased over time, particularly in services. For example, in 1960, the Bureau of Economic Analysis (BEA) did not estimate home healthcare consumption, and services such as internet services did not exist. We compute Bowley and Kelly skew statistics using 181 categories of goods and services, which are listed in online appendix A1, Table A2.

7 Please see the online appendix Figure A2 for the profile of monthly, 3-month moving average, and 12-month moving average of the Bowley skewness measure, and Figure A3 for the corresponding figures for the Kelly measure.

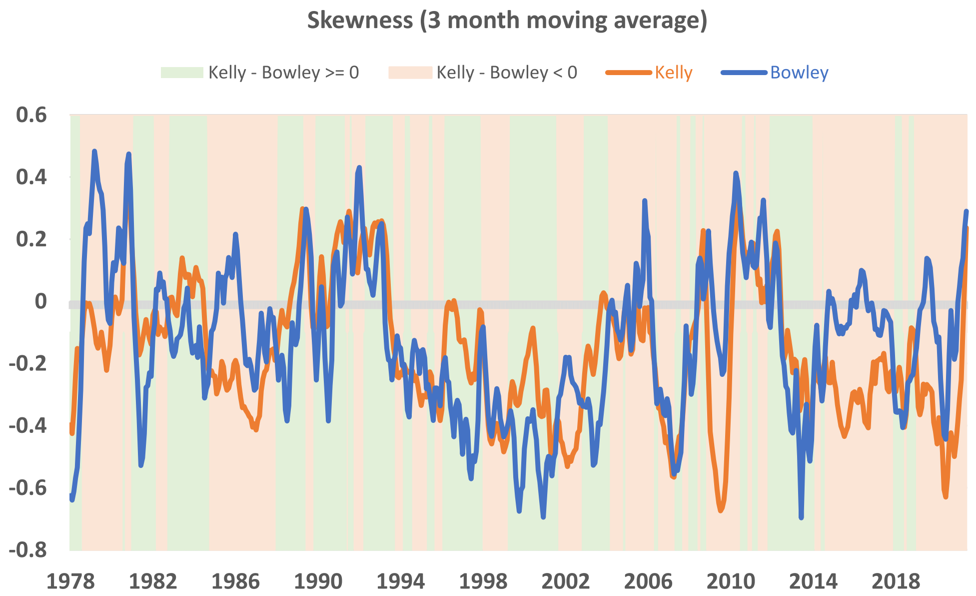

Fig. 2. Cross-sectional asymmetry in PCE inflation (12-month %).

function of the cross-sectional association or relationship of the underlying stochastic processes – may have beneficial information for inflation forecasting and separately for informing estimates of stochastic volatility in equations defining inflation dynamics.

Median and Skew by Goods and Services

A growing literature has documented the importance of forecasting inflation by separately modeling and forecasting the goods and services sub-categories of aggregate inflation (see, e.g., Tallman & Zaman, 2017 and Verbrugge, 2023b). A recent BIS report (BIS, 2022) advocates looking at more disaggregated levels to better understand the aggregate inflation dynamics. Relatedly, Schoenle and Smith (2022) show that over time US inflationary process has been increasingly driven by idiosyncratic shocks than aggregate shocks, i.e., it has become more granular. Motivated by this line of research, we examine whether gains in the accuracy of goods and services inflation forecasts are possible by computing robust measures (separately) for goods and services. Furthermore, this decomposition could provide a better understanding of the movements of the *aggregate* robust measures. Accordingly, we construct robust measures (median and skewness) for goods and services PCE. Fig. 3 plots median goods and services PCE inflation alongside median PCE inflation. A quick visual inspection indicates a striking similarity between the median PCE inflation and median services PCE inflation. This suggests that both indexes categorize the median components with similar price changes.[8] Figure A4 in the online appendix plots the skewness measures computed separately by services and goods categories to conserve space.

Data

All of the empirical analysis in the main text uses data at monthly frequency spanning January 1978 through June 2021.[9] We use data on the personal consumption expenditures price index (PCE), PCE excluding food and energy components (core PCE), and data on both price indices and nominal expenditure shares of 181 components of PCE.[10] Our target variable of interest is the 12-month PCE inflation rate.[11] Table A1, in the online appendix, provides a complete listing of all the data series retrieved from Haver Analytics.

3. Models and forecasting setup

In the inflation forecasting literature, modeling inflation in "gap" form, where the gap is defined as the deviation of inflation from its underlying long-run trend (i.e., long-run inflation), is quite helpful in improving the accuracy of inflation forecasts (e.g., Clark & Doh, 2014; Faust & Wright, 2013; Tallman & Zaman, 2017; Zaman, 2013). A simple univariate AR model of inflation in the gap is widely recognized as an "amazingly hard to beat"

8 Interestingly, in computing the median PCE, over our sample period, about 82 percent of the time (i.e., for 435 out of 533 months), the identified median component belonged to the services category.

9 In the online appendix, we report selected results based on data spanning July 2021 through December 2022, which was made available after we had completed this paper.

10 The online appendix (Table A2) lists all of the 181 disaggregated components used to construct the robust asymmetry measure. It is worth mentioning that if we instead use the 153 components that go into constructing core PCE, the resulting estimates of the asymmetry measure are similar to the one obtained with all 181 components.

11 The Federal Reserve's inflation goal is framed in terms of the 12-month inflation rate in PCE inflation.

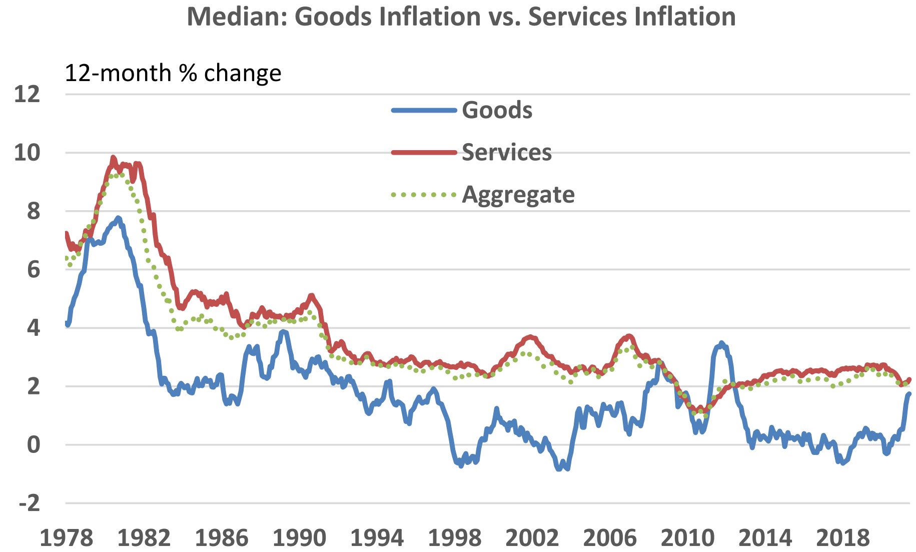

Fig. 3. Median by goods and services.

benchmark (e.g., Faust & Wright, 2013). To assess the marginal contribution of trimmed-mean estimators and skewness measures to improve the accuracy of inflation forecasts, we extend the univariate inflation in the gap model to a multivariate setup.[12]

First, we build a bi-variate Bayesian VAR[13] of headline and median PCE inflation. We denote this specification as *"BVAR: PCE + Median"*. We view median PCE inflation as capturing the "medium-term" trend in inflation,[14] and accordingly, we expect headline PCE to move towards median PCE over the medium term. We model *both* headline and median PCE inflation in deviations from PTR[15] to preserve the information implicit in the headline-median gap. We compare the accuracy of the bi-variate BVAR (*i.e., BVAR: PCE + Median)* in forecasting headline PCE inflation to that of the univariate inflation in the gap model. This comparison would give us a sense of the marginal contribution of median PCE inflation above and beyond headline PCE inflation's history in improving the forecast accuracy of headline PCE inflation. Recall that this marginal contribution of median PCE would reflect both its signal about the underlying medium-term trend and its implicit signal about the current degree of asymmetry (skewness).

Second, to get a rough approximation of the extent to which skewness contributes to the median PCE's marginal contribution, we construct another bi-variate BVAR of headline PCE and a skewness measure (alternatively either Bowley, "B", or Kelly, "K"). We denote this specification as *"BVAR: PCE + Skew (B)"* and *"BVAR: PCE + Skew (K)."* A forecast accuracy comparison between these BVARs and the BVAR incorporating median PCE gives a sense of how skewness contributes to the latter BVAR's forecast accuracy compared to the median PCE's signal about the trend.

Third, we construct tri-variate BVARs, which incorporate headline PCE, median PCE, and skewness (denoted *"BVAR: PCE + Median + Skew (B)"* and *"BVAR: PCE + Median + Skew (K)"*). The comparison of the tri-variate BVAR to the corresponding bi-variate BVAR would give us a sense of the marginal contribution of the "direct" measure of skewness above and beyond that of median PCE and headline PCE, noting that median PCE already embeds an *implicit* signal about the skewness (when added

12 Faust and Wright (2013) propose a quarterly AR(1) gap model and they show that a specification with a fixed slope parameter, rho = 0.46 and intercept = 0 does slightly better than the unrestricted specification whose slope and intercept are estimated from the data. Since we work with models estimated with data at a monthly frequency, we use a monthly AR(3) gap model. We *estimate* this AR model, for several reasons. First, as we estimate using Bayesian methods, this conveniently allows us to produce density forecasts. (We note that the point forecasts from the AR model with or without Bayesian estimation are identical.) Secondly, it is not obvious how a fixed quarterly parameter should be mapped into fixed monthly coefficients. Lastly, it naturally supports our extension of extending the univariate AR model to VAR model by adding the two covariates (trimmed-mean and skewness measures) and allows for more fair comparison with the univariate AR model.

13 Bayesian VARs are widely used to forecast macroeconomic variables. We use BVAR models similar to those used in Bańbura, Giannone, and Reichlin (2010) and Knotek II and Zaman (2019). We set lag length = 3 to be consistent with the AR(3) benchmark model. We relegate the BVAR model details to online appendix A2.

14 On their website, the Federal Reserve Bank of Cleveland explicitly states that the median PCE indicator is designed to capture the underlying trend in inflation, where they define underlying trend as the "medium-horizon" trend in inflation. Further, when evaluating how well trimmed-mean estimators track the underlying trend inflation, the common practice in the literature is to use the 36-month centered moving average of actual inflation to define trend inflation. These facts support the notion that trimmed-mean estimators best reflect medium-term trend inflation. We recognize that in the literature, researchers often treat trimmed-mean estimators of inflation as reflecting long-run trend in inflation; however, more recently, there is an increasing recognition that is not the case.

15 PTR is the survey-based long-run (5- to 10-years-ahead) PCE inflation expectations series from the Federal Reserve Board of Governor's FRB/US econometric model.

alongside the headline PCE). Similarly, comparing the tri-variate BVAR with the univariate model would give us a sense of the combined use of median PCE and the "direct" measure of skewness in improving the forecast accuracy of headline PCE inflation.

We repeat this exercise by replacing the median PCE with the trimmed-mean PCE, which gives us a bi-variate BVAR, *"BVAR: PCE + Trim"*, and tri-variate BVARs *"BVAR: PCE + Trim + Skew (B)"* and *"BVAR: PCE + Trim + Skew (K)"*. Then we repeat replacing the trimmed-mean PCE with core PCE, which gives us bi-variate BVAR, *"BVAR: PCE + Core"*, and tri-variate BVARs *"BVAR: PCE + Core + Skew (B)"* and *"BVAR: PCE + Core + Skew (K)"*.

Fourth, we assess the value added of our robust measures in improving the accuracy of the inflation forecasts from the Phillips curve specifications. A long list of papers has documented the inferior accuracy of forecasts from the Phillips curve models compared to forecasts from models with univariate specifications (e.g., Faust & Wright, 2013). More recently, Ball and Mazumder (2020) and Ashley and Verbrugge (2023) show the competitive accuracy of the inflation forecasts from Phillips curve models based on trimmed-mean inflation measures. Accordingly, we examine whether including median PCE (or trimmed-mean PCE) and skewness in the Phillips curve specification helps improve accuracy. If it does, are the gains large enough to make the accuracy of the forecast competitive with the univariate benchmark? We find that the gains are quite small; hence in the interest of brevity, we simply report the forecast accuracy from the Phillips curve specification without the robust measures, which we denote as *"BVAR: PCE + UR"*, where UR refers to the unemployment rate gap constructed as the difference between the unemployment rate and the CBO's estimate of the natural rate of unemployment.

Fifth, to assess the usefulness of robust measures of goods and services inflation in improving the accuracy of goods and services inflation forecasts, we perform two sets of forecasting exercises similar to those described previously. In the interest of brevity, we only summarize the results below and report these results in detail in Appendix A9.

Pseudo-Out-of-Sample Forecasting

While real-time data is available for aggregate PCE inflation and the unemployment rate, real-time data at the disaggregate component level (required to compute the median PCE and skewness) is limited; therefore, we resort to pseudo-out-of-sample forecast evaluation. We perform forecasting evaluation using a recursively expanding window of estimation. All the models, including the univariate AR gap model, are estimated using Bayesian methods, facilitating the computation of the density forecasts. The estimation sample started in January 1978, and forecast evaluation was performed over the sample from January 1994 through June 2021. At each recursive run, forecasts are produced up to three years out (i.e., the forecast horizon, h, ranges from h = 1 to h = 36 months ahead). The models produce forecasts of the PCE inflation "gap", which are then converted into forecasts of the PCE inflation rate by adding to the forecasts of the inflation "gap" the latest estimate of the PTR available at each recursive run. The point forecasts, which are the posterior mean of the density forecasts, are evaluated using the metric of the mean squared forecast error (MSE). To assess the statistical significance of gains in the accuracy of point forecasts between the two models, we use the Diebold and Mariano test (with the Newey–West correction) using the two-sided tests of the standard normal. The density forecasts are evaluated using the widely used metric of the logarithmic predictive score (parametric normal approximation), and the statistical significance is assessed using the likelihood-ratio test of Amisano and Giacomini (2007), where the test statistics use a two-sided t-test.

4. Forecasting results

Table 1 reports the results of the point forecast evaluation comparing inflation forecast accuracy across several model specifications. The results correspond to model specifications that use Kelly skewness measures constructed based on disaggregates' month-to-month inflation rates;[16] we compute the three-month moving averages as our estimates of the skewness measures that enter the models.[17]

It matters how skewness is measured. We find that Kelly skewness contains more predictive content for inflation than Bowley skewness (see online appendix Tables A4 and A5). Further, skewness constructed from the components' month-to-month inflation rates is preferred to the corresponding 12-month trailing rates, and a three-month window for the moving average of monthly skewness is preferred over other window lengths.

The forecast accuracy is reported for select forecast horizons to conserve space. The top panel of the table reports results corresponding to the full sample (1994–2021), the middle panel corresponds to the pre-Great Recession sample (1994–2007), and the bottom panel corresponds to the financial crisis and onward sample (2008–2021). In each panel, the numbers reported in the first row are the root mean squared error (RMSE) from the benchmark univariate inflation in the gap model, denoted "*AR(3)-PCE.*" The rows below it are ratios that report *relative* MSEs (relative to MSEs from the AR(3) PCE). Thus, a ratio of more than 1 indicates that the univariate inflation in the gap model is more accurate on average than the model being compared.

The results reported in Table 1 suggest four observations. First, adding trimmed-mean estimators to the model improves the forecast accuracy of the aggregate PCE inflation forecasts for most forecast horizons. However, it worsens forecast accuracy in the near term. The gains in forecast accuracy are greater from including the median PCE than trimmed-mean PCE, which generally outperforms core PCE. In addition, a larger number of

16 The results based on model specifications in which skewness measures are constructed based on disaggregates' 12-month trailing inflation rates are found to be inferior compared to those obtained using skewness measures constructed from month-to-month inflation rates. Owing to space constraints, we do not report these results in the paper, but they are available upon request from the authors.

17 The three-month moving average was preferred to other window lengths (e.g., 1, 2, 4, 6, 8, 10, 12).

Table 1
PCE inflation out-of-sample point forecasting comparison. [Skew constructed based on month-over-month inflation rates.]

Full sample (January 1994–June 2021)							
	h = 1M	h = 6M	h = 1Y	h = 18M	h = 2Y	h = 30M	h = 3Y
AR(3)-PCE RMSE	0.265	0.858	1.126	1.064	1.075	1.077	1.044
Relative MSE							
BVAR: PCE + Skew (K)	1.028	0.957[a]	0.988	0.976	0.959[a]	0.959[a]	0.967[a]
BVAR: PCE + Median	1.046[a]	0.991	0.893	0.882[a]	0.879[a]	0.898[a]	0.887[a]
BVAR: PCE + Median + Skew (K)	1.008	0.909	0.889	0.887[a]	0.876[a]	0.897[a]	0.885[a]
BVAR: PCE + Trim	1.045[a]	0.997	0.891	0.918	0.913	0.916	0.913[a]
BVAR: PCE + Trim + Skew (K)	1.011	0.916	0.885	0.922	0.906	0.911	0.911[a]
BVAR: PCE + Core	1.045[a]	1.010	1.008	0.997	0.980	0.967[a]	0.973
BVAR: PCE + Core + Skew (K)	1.045	1.010	1.008	0.997	0.980	0.967[a]	0.973
BVAR: PCE + UR	1.109[a]	1.181	1.320[a]	1.485[a]	1.628[a]	1.634[a]	1.612[a]
Pre-financial crisis sample (January 1994–December 2007)							
	h = 1M	h = 6M	h = 1Y	h = 18M	h = 2Y	h = 30M	h = 3Y
AR(3)-PCE RMSE	0.245	0.553	0.806	0.870	0.941	0.955	0.930
Relative MSE							
BVAR: PCE + Skew (K)	1.009	1.006	0.998	0.989	0.980	0.981	0.995
BVAR: PCE + Median	1.024[a]	1.053	0.883	0.815	0.787[a]	0.795[a]	0.796[a]
BVAR: PCE + Median + Skew (K)	1.007	1.037	0.888	0.830	0.798[a]	0.804[a]	0.802[a]
BVAR: PCE + Trim	1.019	1.076	0.951	0.910	0.860	0.838[a]	0.814[a]
BVAR: PCE + Trim + Skew (K)	0.999	1.052	0.955	0.921	0.866	0.844[a]	0.818[a]
BVAR: PCE + Core	1.005	1.030	1.031	1.018	1.004	0.997	1.006
BVAR: PCE + Core + Skew (K)	1.008	1.045	1.046	1.032	1.012	1.000	1.007
BVAR: PCE + UR	1.016	1.220	1.375	1.602[a]	1.648[a]	1.769[a]	1.979[a]
The financial crisis and onward sample (January 2008–June 2021)							
	h = 1M	h = 6M	h = 1Y	h = 18M	h = 2Y	h = 30M	h = 3Y
AR(3)-PCE RMSE	0.285	1.087	1.359	1.097	0.975	0.972	0.833
Relative MSE							
BVAR: PCE + Skew (K)	1.043[a]	0.943[a]	0.982	0.953	0.908[a]	0.924[a]	0.947[a]
BVAR: PCE + Median	1.063[a]	0.976	0.906	0.932	0.793[a]	0.742[a]	0.790[a]
BVAR: PCE + Median + Skew (K)	1.009	0.877	0.901	0.931	0.771[a]	0.731[a]	0.781[a]
BVAR: PCE + Trim	1.065[a]	0.980	0.883	0.933	0.774[a]	0.709[a]	0.808
BVAR: PCE + Trim + Skew (K)	1.021	0.884	0.875	0.929	0.747[a]	0.697[a]	0.802[a]
BVAR: PCE + Core	1.076[a]	1.004	0.997	0.974[a]	0.946[a]	0.942[a]	0.954[a]
BVAR: PCE + Core + Skew (K)	1.055	0.955	0.999	0.958[a]	0.910[a]	0.927[a]	0.939[a]
BVAR: PCE + UR	1.180[a]	1.179	1.347[a]	1.603	1.913	1.807	1.894

Notes: The numbers reported in the first row of each panel are the root mean squared error (RMSE) from the univariate AR PCE inflation in gaps (3-lag specification), while the rows below it are ratios that report relative MSEs (relative to MSE from the AR(3) PCE inflation in gaps). Thus, a ratio of more than 1 indicates that the univariate inflation in gaps model is more accurate on average than the model being compared. The forecast performance is based on an expanding estimation window spanning January 1994 through June 2021 (full sample) and January 1994 through December 2007 (pre-financial crisis sample).
[a]Indicates statistical significance up to 10% level and is based on Diebold–Mariano West test.

gains in accuracy are classified as statistically significant in the case of median PCE compared to the other two, especially for forecast horizons 18 months ahead and greater.[18]

Second, including the Kelly skewness measure with or without the inclusion of trimmed-mean estimators marginally improves the forecast accuracy of the aggregate PCE inflation for most forecast horizons. But for the near-term forecast horizon, Kelly skewness plays a non-trivial role: its inclusion improves forecast accuracy, primarily by converting statistically significant losses to insignificant losses of smaller magnitude. However, in the sample before the Great Recession, skewness measures did not help improve accuracy.

A deeper examination of the errors, combined with an understanding of the behavior of median PCE and headline PCE inflation, allows us to understand how skewness is *episodically* useful. It has been shown that headline PCE inflation moves towards median PCE over time to close the gap between the two; this explains our robust finding that trimmed-mean indicators help improve forecast accuracy at the medium horizon. However, there can be *persistent* deviations between the two, likely due to persistent relative price shocks. In other words, sometimes it takes a long time for the gap to close, with the period spanning mid-2012 through late 2016 being a prominent example. During this period, forecasts from a model including median PCE are biased upwards compared to forecasts from the univariate AR model. Adding the skewness measure to this model, which is negative during this period, generates forecasts that call for less strong inflation, increasing accuracy. Similarly, the forecast from an AR model calls for inflation to move up gradually towards the end point estimate implied by the

18 Our finding that the inclusion of median PCE improves the forecast accuracy of aggregate PCE inflation over the medium- to longer-term horizons is consistent with the findings of Crone, Khettry, Mester, and Novak (2013), who find similar support for median CPI inflation in forecasting aggregate CPI inflation.

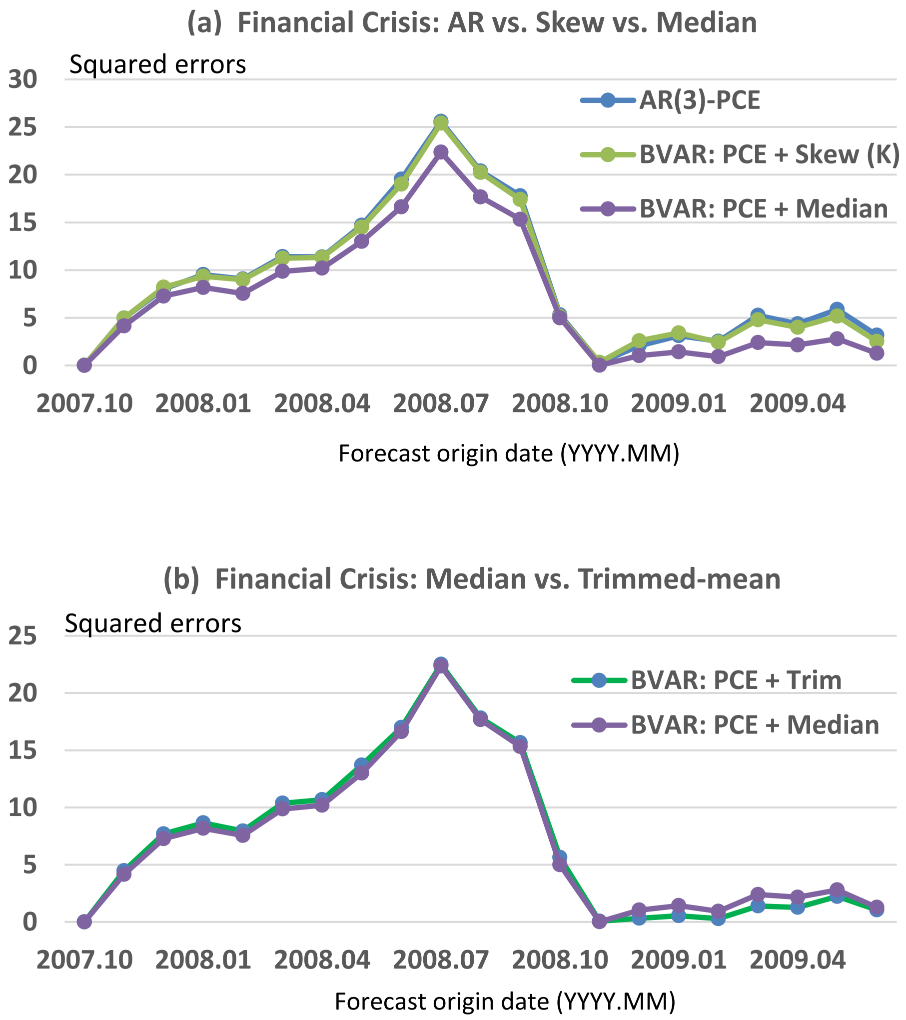

Fig. 4. Forecast errors during the Great Recession.

PTR (which during that period is higher than PCE inflation). Adding the skewness measure (which is negative) to this AR model generates a forecast with inflation moving up more slowly than the AR model, increasing accuracy. As we will discuss shortly, starting mid-2021, skewness again proved useful (especially early on) because of the persistent nature of the relative price shocks. Prior to the Great Recession period (i.e., from 1994 to 2007), deviations between headline and median PCE were relatively short-lived. Hence, skewness could play only a minor role.

Third, consistent with the findings in the literature, the bi-variate Phillips curve specification significantly underperforms. The density forecast evaluation results echo the point forecast evaluation results; in the interest of brevity, these results and discussion are relegated to the online appendix A4.

Forecasting Performance during the Great Recession and the COVID Pandemic

We next illustrate the marginal efficacy of our robust inflation measures in forecasting aggregate PCE inflation during crisis periods, normally associated with heightened uncertainty. We focus on two crisis periods: the great financial crisis (also known as the Great Recession) [GFC] and the great pandemic crisis [GPC], which is still ongoing at the time of writing. Specifically, we examine the forecasting performance of our BVAR models for 12-months-ahead forecasts generated during the GFC period spanning October 2007 through June 2009 and the GPC period spanning March 2020 through June 2020. For the latter, i.e., the GPC period, we go only through June 2020 because at the time of compiling results, the available data end in June 2021, which we need to evaluate the 12-months-ahead forecast.

Fig. 4, panel (a) plots the forecast errors over the GFC period from three models: the benchmark gap AR(3)-PCE, the BVAR: PCE + Skew (K), and the BVAR: PCE + Median.[19] As is evident by big misses, all three models generate forecasts that poorly track the actual PCE inflation during the GFC period.[20] However, the model

19 The plot for BVAR: PCE + Median + Skew (K) is almost identical to that for the BVAR: PCE + Median; therefore, we do not show it.

20 Nonlinear Phillips curve models, such as Ashley and Verbrugge (2023) and Verbrugge and Zaman (2023a), do a reasonable job of

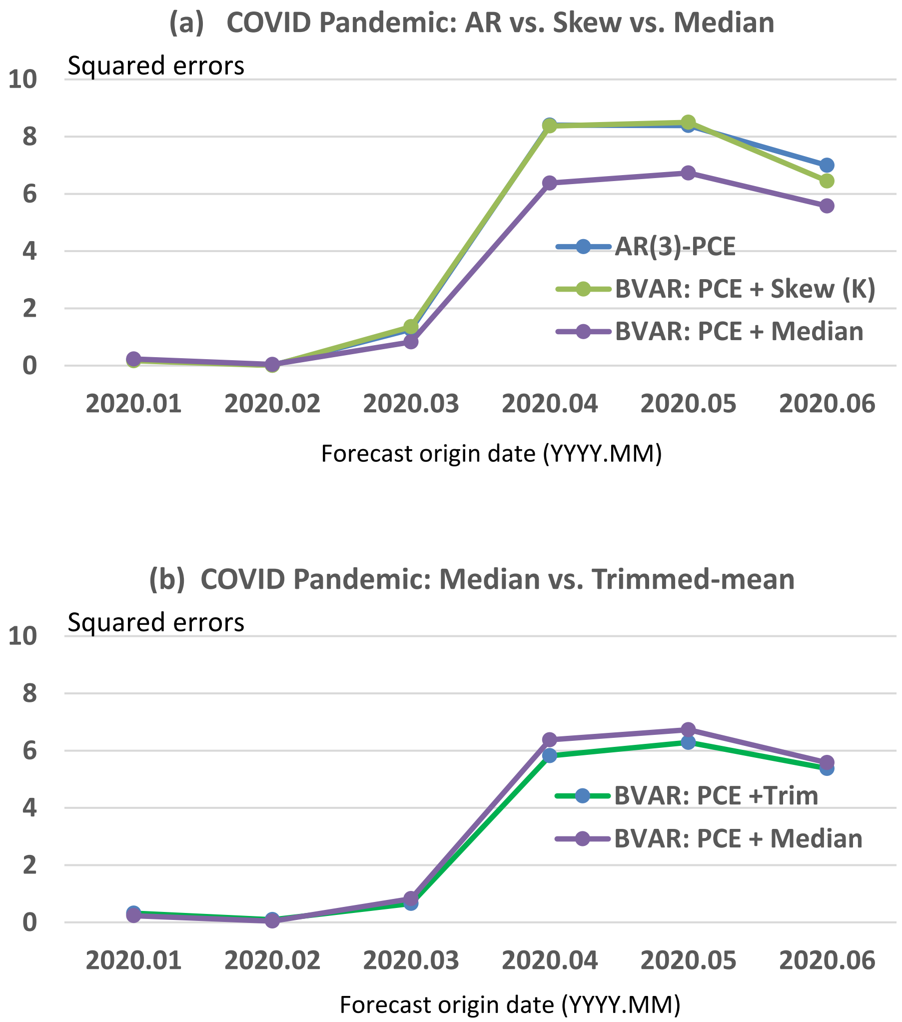

Fig. 5. Forecast errors during the Great Pandemic (COVID-19).

that includes median PCE inflation experiences relatively smaller errors than the univariate benchmark. During that period, actual PCE inflation came in well below the models' projections, resulting in large errors. Panel (b) in Fig. 4 plots the forecast errors from the BVAR: PCE + Trim model alongside the PCE + Median model. Both models performed comparably during this period.

Fig. 5, panel (a) plots the forecast errors over the GPC period from the same three models. Again, there is evidence of big misses: all three models generate forecasts during the GPC period that do an inferior job of tracking the actual PCE inflation. However, the model that includes median PCE inflation experiences relatively smaller errors than the univariate benchmark. During this period, PCE inflation exceeded the models' projections, resulting in large errors. Panel (b) in Fig. 5 plots the forecast errors from the BVAR: PCE + Trim model alongside the PCE + Median model. Both models performed comparably during this period, with the model that includes the trimmed-mean measure performing slightly better than the model with the median measure. It is worth noting that the models' forecast errors during the GPC period are smaller in magnitude than during the GFC period.

Since the writing of this paper, additional data covering the period mid-2021 through December 2022 have become available. In the online appendix A8, we briefly discuss and illustrate the forecasting performance of our selected models over this recent period. This period provides a nice illustration of the forecasting benefits that robust measures can provide under particular circumstances. Early in this period, when inflation began to pick up, it was driven by price spikes in a few sectors (see Almuzara & Sbordone, 2022); thus, trimmed mean indicators were somewhat late to recognize the persistent nature of the inflation surge.[21] Later in the period, inflationary pressures became broad-based and elevated.

tracing out the dynamics of trimmed mean PCE inflation over this period.

21 Verbrugge and Zaman (2023a) find that, over this period, inclusion of a supply-shock variable greatly improved the forecast from a trimmed-mean model.

Schoenle and Smith (2022) and Ocampo et al. (2022) find that during periods when inflation is broad-based and elevated, trimmed-mean estimators behave similarly to headline PCE; hence, their inclusion does little to improve forecast accuracy. In keeping with these facts, we find that models with or without trimmed-mean estimators generally perform comparably over this period. However, models with "direct" skew estimates provided more accurate forecasts. Skewness was unusually positive over this period, and model specifications, including skewness, projected higher inflation than those that did not.

Overall, the forecast results for the GFC, GPC, and post-pandemic (mid-2021 through December 2022) periods highlight the difficulties in accurately forecasting aggregate PCE inflation. One is better off incorporating information from trimmed-mean estimators and Kelly skewness in constructing forecasts of PCE inflation using popular time-series models.

Breakdown by Goods and Services

Similar to the results for headline PCE inflation, we find evidence that for both goods and services inflation measures, the addition of robust measures to their respective univariate gap models improves the forecast accuracy of these inflation measures. Interestingly, we find modest evidence for services inflation that skewness per se has predictive content. We refer the reader to online appendix A9 for a detailed discussion of the results.

5. The usefulness of skewness for stochastic volatility modeling

As noted in Section 2 above (and in Verbrugge, 1999), asymmetry in the cross-sectional distribution is associated with the underlying (time-varying) conditional variance–covariance structure. This suggests that estimates of skewness could help improve the (quarterly) estimates of stochastic volatility in model equations defining inflation dynamics. To help answer this question, we use the state-of-the-art stochastic volatility in the mean model developed by Chan (2017). In the interests of brevity, we refer the reader to online appendix A10 for a detailed discussion of model details, estimation, and results. Our main finding is that the skewness measure is modestly effective in refining the *contemporaneous* estimates of stochastic volatility in the innovations to the equation defining the goods PCE inflation and, in turn, headline PCE inflation. We also find important differences in coefficient estimates across headline inflation, goods inflation, and services inflation. We conjecture that these differences stem from the interaction of several factors: relative price shocks of various types, inflation, inflation expectations and uncertainty, and monetary policy.

6. Conclusion

This paper explores the usefulness of the trimmed-mean estimators and robust skewness statistics in improving the point and density accuracy of aggregate PCE inflation forecasts. Trimmed-mean estimators have been shown to do well in forecasting aggregate inflation, with the forecast accuracy gains thought to be due to their prowess in tracking the underlying trend. However, we illustrate strong evidence of time variation in the cross-sectional asymmetry computed using the 180+ components of the PCE price index. Such asymmetry correlates with inflation, suggesting a second reason that trimmed-mean estimators have predictive content: the gap between headline and trimmed-mean inflation provides an implicit signal about skewness. We assess the predictive content of skewness, independent of the information about the future trend embedded within trimmed-mean estimators.

We examine both the joint contribution of these measures and their marginal contributions in possibly improving the point and density forecast accuracy of PCE inflation. Among the trimmed-mean estimators, median PCE inflation's ability to forecast future headline PCE inflation has barely been explored. So, an important secondary contribution of this paper is to examine the usefulness of median PCE in forecasting aggregate PCE inflation. A third important contribution of this paper is introducing and examining the use of median goods PCE and median services PCE – and their respective robust skewness estimates – for forecasting goods PCE and services PCE. Finally, we explore whether robust measures are useful in stochastic volatility modeling.

Based on a forecast evaluation sample covering the period from January 1994 through June 2021, a period that includes large volatility in oil prices, a financial crisis and deep recession, and a severe global pandemic, our results indicate significant gains in the point and density accuracy of PCE inflation forecasts for horizons 18 months ahead and longer. Most of the improvements come from the inclusion of trimmed-mean estimators, with only marginal improvements from the addition of robust skewness estimators. A split sample examination suggests that most of the gains in accuracy are concentrated in the sample spanning the Great Recession and onward, i.e., January 2008 through June 2021, a period where inflation has remained low.

We find slightly stronger support for median PCE over trimmed-mean PCE, and both outperform the exclusion estimator, core PCE. The skewness estimator matters. We find strong support for Kelly skewness over Bowley skewness and for skewness measures constructed based on components' month-over-month rates. In contrast, skewness measures based on 12-month rates marginally worsened accuracy, even though aggregate PCE and trimmed-mean estimators enter the models as 12-month trailing rates.

Using state-of-the-art stochastic volatility in the mean model, we illustrate the modest efficacy of the skewness measure in refining the *contemporaneous* estimates of stochastic volatility in the innovations to the equation defining the goods PCE inflation and, in turn, headline PCE inflation.

Over time, the reliance on trimmed-mean inflation estimators to obtain a signal about both the underlying trend in inflation and future inflation has increased globally. Hence, our empirical findings are useful for a broad swath of practitioners interested in forecasting inflation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2023.05.003.

References

Almuzara, Martin, & Sbordone, Argia (2022). Inflation persistence: How much is there and where is it coming from? In *Liberty Street Economics 20220420*. Federal Reserve Bank of New York.

Amisano, Gianni, & Giacomini, Raffaella (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, *25*(2), 177–190. http://dx.doi.org/10.1198/073500106000000332.

Ashley, Richard, & Verbrugge, Randal J. (2023). *The intermittent Phillips curve: Finding a stable (but persistence-dependent) phillips curve model specification*: *Working Paper No. 19-09R2*, Federal Reserve Bank of Cleveland, http://dx.doi.org/10.26509/frbc-wp-201909r2.

Ball, Laurence, & Mankiw, N. Gregory (1994). Asymmetric price adjustment and economic fluctuations. *The Economic Journal*, *104*(423), 247. http://dx.doi.org/10.2307/2234746.

Ball, Laurence, & Mazumder, Sandeep (2011). Inflation dynamics and the great recession. *Brookings Papers on Economic Activity*, *2011*(1), 337–381. http://dx.doi.org/10.1353/eca.2011.0005.

Ball, Laurence, & Mazumder, Sandeep (2020). The nonpuzzling behavior of median inflation. In Gonzalo Castex, Jordi Galí, & Diego Saravia (Eds.), *Series on Central Banking Analysis and Economic Policies 27*, *Changing inflation dynamics, evolving monetary policy* (1st ed.). (pp. 49–70). Santiago, Chile: Banco Central de Chile, https://hdl.handle.net/20.500.12580/4880.

Bańbura, Marta, Giannone, Domenico, & Reichlin, Lucrezia (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics*, *25*(1), 71–92. http://dx.doi.org/10.1002/jae.1137.

Bryan, Michael, & Cecchetti, Stephen (1993). *Measuring core inflation. W4303*. Cambridge, MA: National Bureau of Economic Research, http://dx.doi.org/10.3386/w4303.

Carroll, Daniel R., & Verbrugge, Randal J. (2019). *Behavior of a new median PCE measure: A tale of tails*: *Economic Commentary (Federal Reserve Bank of Cleveland), July*, http://dx.doi.org/10.26509/frbc-ec-201910.

Chan, Joshua C. C. (2017). The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. *Journal of Business & Economic Statistics*, *35*(1), 17–28. http://dx.doi.org/10.1080/07350015.2015.1052459.

Clark, Todd E, & Doh, Taeyoung (2014). Evaluating alternative models of trend inflation. *International Journal of Forecasting*, *30*(3), 426–448. http://dx.doi.org/10.1016/j.ijforecast.2013.11.005.

Crone, Theodore M., Khettry, N. Neil K., Mester, Loretta J., & Novak, Jason A. (2013). Core measures of inflation as predictors of total inflation. *Journal of Money, Credit and Banking*, *45*(2–3), 505–519. http://dx.doi.org/10.1111/jmcb.12013.

Dolmas, Jim (2005). *Trimmed mean PCE inflation*: *Working paper 0506*, Federal Reserve Bank of Dallas, https://ideas.repec.org/p/fip/feddwp/05-06.html.

Dolmas, Jim (2009). *The 2009 revision to the trimmed mean PCE inflation series*: *Technical note*, Federal Reserve Bank of Dallas.

Faust, Jon, & Wright, Jonathan H. (2013). Forecasting inflation. *vol. 2*, In *Handbook of economic forecasting* (pp. 2–56). Elsevier, http://dx.doi.org/10.1016/B978-0-444-53683-9.00001-3.

Kim, Tae-Hwan, & White, Halbert (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters*, *1*(1), 56–73. http://dx.doi.org/10.1016/S1544-6123(03)00003-5.

Knotek II, Edward S., & Zaman, Saeed (2019). Financial nowcasts and their usefulness in macroeconomic forecasting. *International Journal of Forecasting*, *35*(4), 1708–1724. http://dx.doi.org/10.1016/j.ijforecast.2018.10.012.

Mertens, Elmar (2016). Measuring the level and uncertainty of trend inflation. *The Review of Economics and Statistics*, *98*(5), 950–967. http://dx.doi.org/10.1162/REST_a_00549.

Meyer, Brent, & Zaman, Saeed (2019). The usefulness of the median CPI in Bayesian VARs used for macroeconomic forecasting and policy. *Empirical Economics*, *57*(2), 603–630. http://dx.doi.org/10.1007/s00181-018-1472-1.

Norman, David, & Richards, Anthony (2012). The forecasting performance of single equation models of inflation. *Economic Record*, *88*(280), 64–78. http://dx.doi.org/10.1111/j.1475-4932.2011.00781.x.

Ocampo, S., Schoenle, Raphael, & Smith, Dominic (2022). *How robust are robust measures of PCE inflation*: *Economic Working Paper WP-552*, US Bureau of Labor Statistics.

Schoenle, Raphael, & Smith, Dominic (2022). What can we learn from 60 years of PCE inflation data? In *Mimeo, Presented at the conference inflation: drivers and dynamics 2022*. The Cleveland Fed's Center for Inflation Research and the European Central Bank.

Smith, Julie K. (2004). *Journal of Money, Credit, and Banking*, *36*(2), 253–263. http://dx.doi.org/10.1353/mcb.2004.0014.

Tallman, Ellis W., & Zaman, Saeed (2017). Forecasting inflation: Phillips curve effects on services price measures. *International Journal of Forecasting*, *33*(2), 442–457. http://dx.doi.org/10.1016/j.ijforecast.2016.10.004.

Verbrugge, Randal J. (1999). Cross-sectional inflation asymmetries and core inflation: A comment on bryan and cecchetti. *The Review of Economics and Statistics*, *81*(2), 199–202. http://dx.doi.org/10.1162/003465399558166.

Verbrugge, Randal J. (2022). Is it time to reassess the focal role of core PCE inflation in assessing the trend in PCE inflation? *Economía*, *45*(89), 73–101. http://dx.doi.org/10.18800/economia.202201.004.

Verbrugge, Randal J. (2023b). *Post-COVID inflation dynamics: Higher for longer*: *Working Paper No. 23-06*, Federal Reserve Bank of Cleveland, http://dx.doi.org/10.26509/frbc-wp-202306.

Verbrugge, Randa J., & Zaman, Saeed (2023a). *The hard road to a soft landing: Evidence from a nonlinear structural model*: *Working Paper No. 23-03*, Federal Reserve Bank of Cleveland, http://dx.doi.org/10.26509/frbc-wp-202303.

Zaman, Saeed (2013). *Improving inflation forecasts in the medium to long term*: *Economic Commentary (Federal Reserve Bank of Cleveland), No. 2013–16 (November)*, http://dx.doi.org/10.26509/frbc-ec-201316.