Forecasting the Present: Optimal Lags for Contemporaneous Core PCE Inflation[*]

David K Howe

Version: 0.083

Date: January 15, 2025

Abstract

Percentage changes bracketing a given date gauge current inflation best but can only be measured retrospectively. Geometrically annualized percentage changes over the previous 6 or 7 months provide the best timely approximation of underlying core PCE inflation according to a range of criteria. Decimal lags permit the optima to be stated with greater precision. Probabilistic modeling of inflation's level and direction can be conducted using estimates of the error's variance and kurtosis. A review of Federal Reserve statements during the COVID and post-COVID era suggests that greater focus on the geometrically annualized 6 month percentage change could have provided earlier warning of incipient inflation.

Keywords: Inflation, Optimal Lags, Decimal Lags, Annualization, Nowcasting, PCE Deflator, Kurtosis

1 Introduction

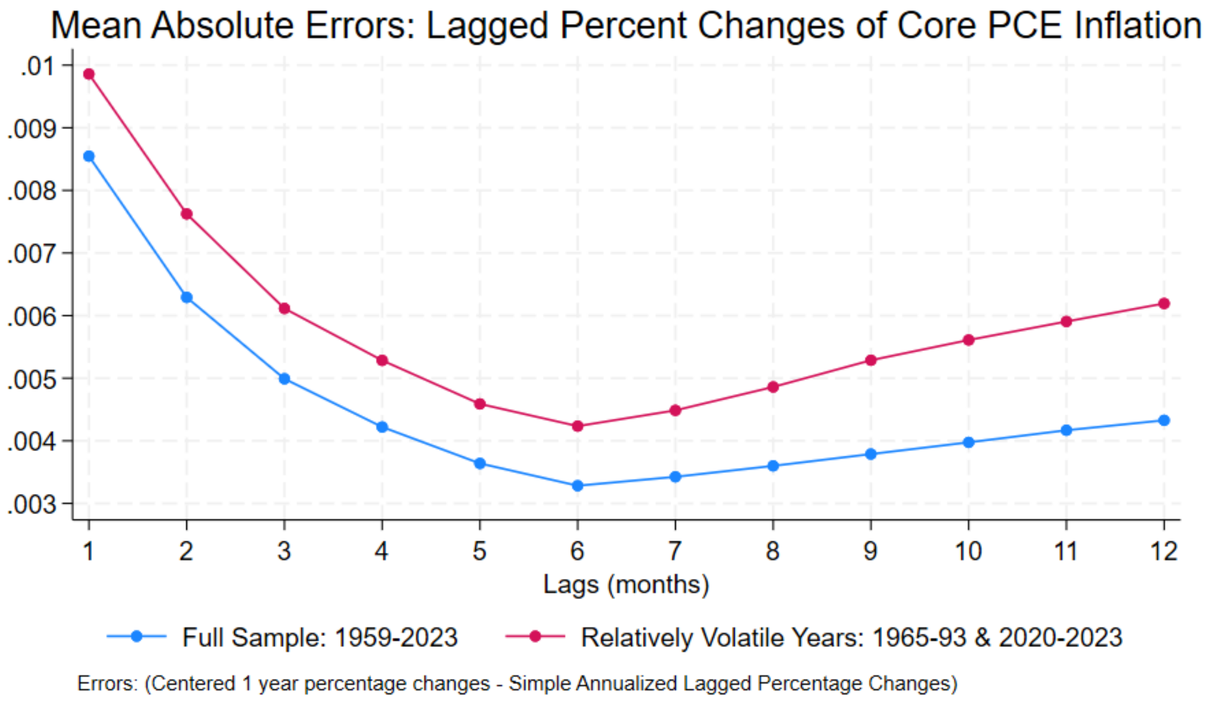
When discussing a new release of inflation data, the business press typically reports percentage changes in prices over the previous year and sometimes the previous month. Both measures are inferior. Half of the one year figure is shaped by developments occurring 6 months prior or older. The one month percentage change is difficult to interpret when not annualized and is so noisy that it is typically recommended that any single month's observation be ignored.

Figure 1 compares the mean absolute errors of some easy to calculate alternatives. The reference series is the *centered* one year percentage change in prices of the Personal Consumption Expenditure Price Index, Excluding Food and Energy (Core PCE). So for June 2000 that would be the one year return between December 1999 and December 2000. The error is the actual minus the predicted or the reference series minus annualized percentage changes over various lags. In this chart simple annualization is used; later sections will apply geometric annualization, as defined in Section 2.

---

[*] David Howe can be reached via OptLag at sixmonthlag dot com.

Figure 1



For both the full sample and the subsample covering years when inflation was higher and more volatile, annualized percent changes over the previous 6 months have the lowest average errors. Annualized one month changes perform worse than all other lags. Annualized changes between 4 and 11 months always perform better than lagged 12 month changes. For periods when inflation is volatile (and of greater note) even the 3 month percentage change performs better than 12 month changes. Later refinements providing more precise estimates will be presented. Figure 1 emphasizes that the typically presented 12 month and 1 month lagged percentage change are substandard proxies for existing inflation. The 6 month percentage change performs 23% better than the 1 year percentage change in the full sample, and 31% better in the relatively volatile subsample. The annualized one month percentage change has 2.3-2.6 times higher mean errors than the annualized 6 month.

Section 2 presents and compares annualization formulas. Simple annualization is commonly used, but geometric annualization is more accurate insofar as it adjusts for compounding. Decimal lags which include a partial weighting of the oldest lag provide greater precision. Section 3 introduces the data, roughly characterizes the underlying data generating process, and reports that simple annualization usually approximates geometric annualization fairly well. Section 4 presents four metrics for underlying inflation as viewed retrospectively. Those are the reference or outcome variables: they reflect data before and after the observation of interest. Section 5 calculates decimal lags generating the lowest errors relative to the four reference series. Results differ for periods of more or less variable inflation. During periods of more variable inflation, geometrically annualized 6.1 month percentage changes match centered one year inflation best, though the ordinary 6 month lag figure is not significantly different.

Interpreting a single new data point can be challenging. Section 6 compares top and bottom deciles in percentage change variables over a variety of lags to the longest run and most smoothed outcome variable. Conventional wisdom holds that changes represented by a single month's data should be ignored. The conventional wisdom is confirmed. Unusual percentage changes within the range of 4 to 8 lags are somewhat informative.

Policy makers care about whether inflation is rising or falling. Section 7 nowcasts the change in centered one year inflation over the previous 6 months, using current 6 month lagged inflation as an estimator for current 1 year centered inflation and a probabilistic framework. Section 8 gathers the results in previous sections and applies them to the 2020-2024 spell of accelerating inflation and disinflation, noting differences with Federal Reserve perspectives as reflected in their public communication. Section 9 concludes.

2 Annualization Formulas

Formulas for simple, geometric, decimal, and log annualization follow. Figure 1 applied simple annualization which is familiar and easier to calculate by hand, though it is not the most accurate:

$$\pi_{t,L}^{sa} = (12/L)\left(\frac{P_t}{P_{t-L}} - 1\right) \tag{1}$$

Assuming monthly data, the lagged return above is multiplied by (12/L), where L is the number of lags ($L \in \mathbb{N}$), P is the price level at various times, and π is the inflation rate. Use 4 rather than 12 if the data is quarterly.

Geometric annualization works with products and is therefore a more accurate method of annualization:

$$\pi_{t,L}^{ga} = \left(\frac{P_t}{P_{t-L}}\right)^{(12/L)} - 1 \tag{2}$$

Percentage changes are a multiplicative process, not an additive one. Geometric annualization is preferable because the product of one plus consecutive monthly inflation rates is equal to the price change across the full time span. This non-annualized inflation rate formula demonstrates the compounding process:

$$\pi_{t,L} = \left(\frac{P_t}{P_{t-L}}\right) - 1 = \prod_{i=1}^{L}\left(\frac{P_{t-i+1}}{P_{t-i}}\right) - 1 = \prod_{i=1}^{L}(1 + \pi_{t-i+1}) - 1 \tag{3}$$

Sequential percentage changes compound, whether they are applied to prices or bank deposits. Geometric annualization extrapolates monthly percentage changes to an annual basis when compounding is factored in. Simple annualization ignores this compounding process: a monthly percentage change is merely multiplied by 12 to obtain the annual number.

Decimal lags permit greater precision by applying a partial weight to the oldest lag term. Redefine $L \in \mathbb{R}_{>0}$. Define intL as the integer component and fr as the fractional remainder: L = intL + fr. The geometrically annualized lagged decimal inflation rate adapts (2) with a simple weighted average of the two price levels bracketing the decimal lag:

$$\pi_{t,L}^{gad} = \left(\frac{P_t}{(frP_{t-(intL+1)} + (1-fr)P_{t-intL})} \right)^{12/L} - 1 \tag{4}$$

Multiplying identical positive and negative percent changes will result in a net decline in value, so analysts often use log changes. This paper focuses on percent changes as they are commonly used and easier to interpret. For reference geometrically annualized log changes are easily obtained from (2):

$$\pi_{t,L}^{gal} = (12/L)(lnP_t - lnP_{t-L}) \tag{5}$$

Decimal lags can be implemented in logs by substituting the weighted average for P_{t-L}.

3 Data and Data Generating Process

This paper focuses on inflation based upon the Personal Consumption Expenditures Price Index, Excluding Food and Energy, referred to as core PCE. While the Federal Reserve targets the unadjusted personal consumption deflator over the long run, food and energy are commonly removed because their movements are believed to be transient and in the case of energy particularly volatile. For 1959-2023 the ratio of the monthly percentage change standard deviations of PCE food inflation to PCE overall inflation is 1.9; for energy the ratio is a sizable 8.8. The food component has stabilized since the 1970s: in 2000-2023 its PCE volatility ratio declined to 1.5. Indeed, for the Consumer Price Index that ratio reached 0.8: headline CPI is more volatile than the food component, reflecting the fact that food prices are driven far more by the costs of processing, distribution, transportation, and marketing than by the costs of the underlying commodities.[1]

Given its unexceptional price volatility, some may wonder why food is not included in core inflation. Core inflation was routinely reported in the CPI Detailed Report only in 1978. From 1970-1979 the monthly PCE food volatility ratio was 3.1, while the energy ratio was a comparable 5.0. The definitions reflect the concerns of the era. Lombra and Mehra (1983) questioned food's exclusion from core CPI. More recently Adjemian, Li, and Jo (2023) estimate that supply side factors drove 77% of the food category from 1992-2019, falling to less than 60% during the 2020-2023 COVID pandemic. As monetary policy operates primarily on aggregate demand over the short run, it may be acceptable to remove a primarily supply-driven component from headline inflation, even if that component's volatility has stabilized over time.

---

[1] See Smith (2022) and Cowley and Scott (2022).

Table 1: Descriptive Statistics, Monthly Percentage Changes in Prices

	Core PCE	PCE	Core CPI	CPI
Mean	0.0026	0.0027	0.0030	0.0030
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Median	0.0021	0.0022	0.0025	0.0026
Standard Dev	0.0020	0.0025	0.0025	0.0031
Skewness	0.95	0.49	1.31	0.28
Kurtosis	4.48	5.86	5.94	7.18
Corr: Core PCE	1.000	0.833	0.780	0.667
Dickey-Fuller	-10.17***	-11.71***	-12.53***	-13.42***
ARMA (1,1):				
AR(1)	0.9846***	0.9500***	0.9782***	0.9055***
	(0.0066)	(0.0129)	(0.0061)	(0.0179)
MA(1)	-0.7342***	-0.6232***	-0.7459***	-0.5524***
	(0.0190)	(0.0259)	(0.0193)	(0.0287)
ar(1) + ma(1)	0.2504***	0.3268***	0.2324***	0.3531***
	(0.0157)	(0.0185)	(0.0163)	(0.0180)
Annualizations:				
Simple mean	0.0316	0.0321	0.0363	0.0365
Simple sd	0.0236	0.0298	0.0298	0.0378
Geo mean	0.0324	0.0330	0.0373	0.0378
Geo sd	0.0245	0.0310	0.0313	0.0393
Min geo:	-0.0661	-0.1328	-0.0577	-0.1929
Max geo:	0.1303	0.1578	0.1843	0.2629
Annualization Error: (geo-simple)				
mean	0.0007	0.0009	0.0010	0.0013
median	0.0003	0.0004	0.0004	0.0005
sd	0.0010	0.0014	0.0017	0.0022
Top decile cutoff	0.0020	0.0023	0.0026	0.0034
Max	0.0072	0.0104	0.0139	0.0230
Max (geo-simple)/abs(geo)	0.0552	0.0666	0.0757	0.1011

(Standard errors in parentheses)

*p<.10. **p<.05. ***p<.01

Before comparing percentage changes over various lags, the underlying data generating process of single month percentage changes will be explored in Table 1, covering 1959-2023. Core PCE is presented in column 1: subsequent columns show all items PCE, core CPI, and headline CPI for comparison.

Core PCE growth is lower on average than headline or core CPI; mean differences with all-items PCE are not significant at 5%. Core PCE volatility is significantly lower than the other 3 columns: the core PCE standard deviation is 21% lower than all items PCE and core CPI and 38% lower than headline CPI.[2] Annualized monthly core PCE inflation averaged 3.2% with a standard deviation of 2.5%. Monthly percentage changes in core PCE have a correlation of .78 with core CPI, indicating movements that are in less than perfect tandem.

Skew and kurtosis for all series are significantly different from normality at the 1% level of confidence. Histograms are visibly non-Gaussian with higher peaks, longer tails, and modes between the mean and zero, drifting further towards zero as skew increases. Right tails are longer than left tails, with greater moderation for the lesser skewed headline CPI.

While all of the series are serially correlated, unit roots can be decisively rejected at the 99% confidence level: prices are I(1). Monthly percentage changes are noisy so only about 25% of a single month's shock will carry through to the next period (+/- 3%), as indicated by the sum of the ar(1) and ma(1) terms. The high ar(1) term of .9846 implies substantial long run persistence in the remaining shock, corresponding to a half-life of 3.7 years.[3] By way of contrast, the corresponding calculation for headline CPI indicates a half-life of 7 months, following the first period when 35% of the shock carries over. The relevance of shocks that are 6 to 12 months in the past cannot be definitively ruled out or ruled in with this standard model.

Simple annualization is typically a good approximation for the more accurate geometric annualization. The mean error is 7 basis points while the median error is 3 basis. About a quarter of the time, the error is less than 1 basis. The cutoff for the top decile is 20 basis which is greater than *de minimus* though fairly modest. Even the maximum error of 72 basis represents a percentage error of only 5.5%. Errors for the remaining 3 inflation series are slightly worse owing to their greater dispersion and wider range.

4 Retrospective centered metrics for underlying inflation

In this section four centered reference series are compared. The one year lagged inflation rate provides a smoothed estimate of underlying inflation as of six months ago. Since it begins and ends on the same month, it can be used with not seasonally adjusted data, though it will reflect transient noise in the adjustment factor if seasonally adjusted data is used. The two year lagged inflation rate provides a smoothed estimate

---

[2] Differences in variance characterized as statistically significant in this paper are significant at the 95% level of confidence using Levene's robust test as modified by Brown and Forsythe (1974) according to both the median and trimmed mean variants.

[3] Applying a rough confidence interval of 2 standard errors produces a range of 2-26 years.

of underlying inflation as of a year ago. Table 2 centers 1 year and 2 year inflation at their midpoint, then applies two smoothers to 1 year centered inflation. The first smooths outliers, while the second smooths further to better approximate the long run trend line.

The first smoother focuses on single-observation blips. It applies Stata's 3RSRH robust nonlinear smoother, which consists of a centered 3 span moving median smoother, repeated until convergence, followed by a repeating splitting operator, and a Hanning linear smoother. Median smoothing tends to create flat-topped hills which the splitting operator smooths out to a limited extent by splitting up sequences with repeating values, modifying the endpoint values according to a function of their surrounding values, recalculating the 3 span moving median function, and then resmoothing with 3R.[4] The Hanning smoother applies this weighted moving average:

$$z_t = (y_{t-1} + 2y_t + y_{t+1})/4 \tag{6}$$

The resulting series still produces zero percent changes in 3.2% of the observations, while the base series never exhibits changes exactly equal to zero. The long run smoother eliminates changes equal to zero by applying Stata's more lengthy 3R77553RSRHHHH smoother. Numbers denote centered n-span moving medians. As before Rs denote repeated applications until the smoother produces the same series, S is the splitting operator, and H is the Hanning smoother.

Table 2: Retrospective Indicators of Core PCE Inflation

stats	**Centered 1 yr**	**Centered 2 yr**	**Outlier Smoothed 1 yr**	**LR Smoothed 1 yr**
mean	0.0325	0.0325	0.0325	0.0325
sd	0.0214	0.0206	0.0213	0.0212
detrended sd	0.0019	0.0010	0.0015	0.0012
trend reversals/yr	5.25	4.34	1.18	0.79
resid reversals/yr	6.28	5.55	2.83	2.10

Table 2 displays some of the properties of the four retrospective centered inflation indicators, covering 1960-2022. Mean and standard deviation are provided, then the more relevant detrended standard deviation subtracts the centered 13 month moving average from the series before taking the standard deviation.

Trend reversals take place when changes in the series reverse sign, so that increases in inflation follow decreases or vice versa. Residual reversals occur when the detrended series reverses sign: they are preferred since trend reversals vary with the level of inflation, decreasing at higher inflation rates.

The standard deviations in the 2nd row, which are not detrended, are statistically similar: neither robust nor Gaussian variance tests suggest different standard deviations for any pair at even a 90% significance level.

---

[4] See the smooth command in Stata's documentation which is freely available online.

All detrended standard deviations differ statistically from one another. The centered one year percentage return is the noisiest; the centered two year percentage return is the least noisy. That said, the centered two year series has many more trend reversals and residual reversals than the smoothed series, by a factor of 2.0 to 5.5. All of the trend reversal entries as well as all of the residual reversal entries are statistically different from one another at the 95% level of confidence, using an unpaired two-tailed test of proportion equality.

Overall, using the centered 1 year as the performance criteria has the advantage of familiarity: this criterion will be applied in Section 8. The 2 year centered percentage change can be used with not seasonally adjusted data without seasonal distortion. Analysts wishing to best approximate inflation's underlying trend should consider the LR smoothed 1 year series.

5 Error Minimizing Lags

Table 3 presents the optimal precision lags based upon geometrically annualized lagged percentage changes with the lowest mean absolute errors relative to the four centered outcomes detailed in the previous section over various periods. The subsample covering 1965-1993 and 2020-2022 is emphasized, because forecast accuracy matters less when inflation is low and stable, as it was during other years in the broader sample.

The centered one year percentage change produces broadly similar optima compared with its two smoothed variants. Errors based on the 2 year centered geometrically annualized outcome are higher and the lag with the lowest error is near 9 or 10 lags, rather than 6 or 7 for the one year variants. But these 2 year centered errors are not significantly different from one another within the 6 to 9 month range, suggesting that 6 or 7 month lags are a reasonable overall recommendation.

While the 1 year centered reference series is best approximated by a 6.1 month lag, comprised almost entirely of the overlapping portion of the two series, the 2 year centered reference series suggests a more complex data generating process. The 8.5 month optima comprises only 71% of the 12 month overlap. Absolute errors for the 8.5 month decimal lag are significantly lower than those for the 12 month lag at the 5% level of confidence, suggesting that the underlying data generating process does not follow a simple rule by which centered series are best approximated by their known lagged portion, as would be the case if for example monthly inflation was white noise. The 12 month lag has absolute errors that are 7.8% higher than the 8.5 year lag which while statistically significant is neither trivial nor impressively large in magnitude. Better rules of thumb await further investigation of other price series.

Decimal lags possess statistically significant improvement for the two smoothed centered outcomes. For the unadjusted one and two year outcomes differences with the integer lags are not statistically significant. Nonetheless, all differences in errors between integer lags and decimal lags are of trivial magnitude: for the 4 outcome categories they are always 2 basis points or less. This compares with inflation rates rounded to 10 basis points, as typically reported by the Bureau of Labor Statistics.

Table 3: Core PCE Inflation: Lowest Absolute Error Lags Relative to 4 Centered Outcomes

	Centered One Year	Centered Two Year	Outlier Smoothed One Year	Long Run Smoothed One Year
Volatile Inflation Era Subsample: 1965-93 and 2020-2022, N=384				
Best Integer Lag	6	9	6*	7**
Best decimal Lag	6.1	8.5	6.4	6.5
Average Absolute errors for various monthly lags, N=384				
Lag				
3	0.0063	0.0071	0.0062	0.0061
6	0.0042	0.0052	0.0043	0.0044
7	0.0045	0.0051	0.0044	0.0044
9	0.0052	0.0051	0.0050	0.0049
12	0.0061	0.0054	0.0059	0.0058
Decimal	0.0042	0.0050	0.0043	0.0043
Best Decimal Lags, various periods				
1965-1980	5.7	8.6	5.6	5.6
1981-1993	6.2	10.3	6.6	7.4
2020-2022	6.1	6.5	6.3	6.8
1994-2019	9.4	13.4	9.5	10.5
1960-2022, N=756	6.2	10.4	6.5	6.6
Best Integer Lags: 1960-2022	6	10	7***	7**
RMSE criteria, N=384				
Best Integer Lag	6	8	6	6
Best Decimal Lag	5.8	8.5	5.7	5.7

Asterisks indicate significant rounding error, relative to decimal lags.

*p<.10. **p<.05. ***p<.01

Dividing the relatively volatile inflation era into 3 parts shows that optimal lags don't vary dramatically. For the one year centered baseline and its variants the optimal lag in each subperiod is always within one year of the decimal lag over the entire volatility era. The 1981-1993 period had somewhat longer optimal lags than the preceding 1965-1980 period. The COVID inflation era was somewhere in between.

For the two year centered inflation baseline, each subperiod's optimal lags are within 2 months of the broader era. Optimal two year decimal lags varied from 8.6 lags in 1965-1980 to 10.3 lags in 1981-1993, before collapsing to 6.5 during the COVID era. So unsurprisingly errors based on 2 year centered inflation are less stable than errors based on 1 year centered inflation, presumably because of the greater scope for future economic developments.

During the 1994-2019 period, part of the Great Moderation, longer lags had greater predictive power. Optimal decimal lags for one year centered inflation and its variants rose by 3 to 4 months. But the shift in the optimum was driven by small errors: the optimal lags for the entire 1960-2022 period were similar to the volatility era, excepting errors based upon the two year centered baseline. Since forecast accuracy matters a great deal more during periods of unstable inflation, 6 or 7 lags are preferred.

Most of Table 3 uses mean absolute error criteria, but root mean square error (RMSE) can also be considered. Assigning heavier penalties to larger errors has the effect of shortening optimal lags to a modest extent, ranging from zero to eight tenths of a month or zero to three and a half weeks. RMSE therefore favors the 6 month lag over the 7 month lag.

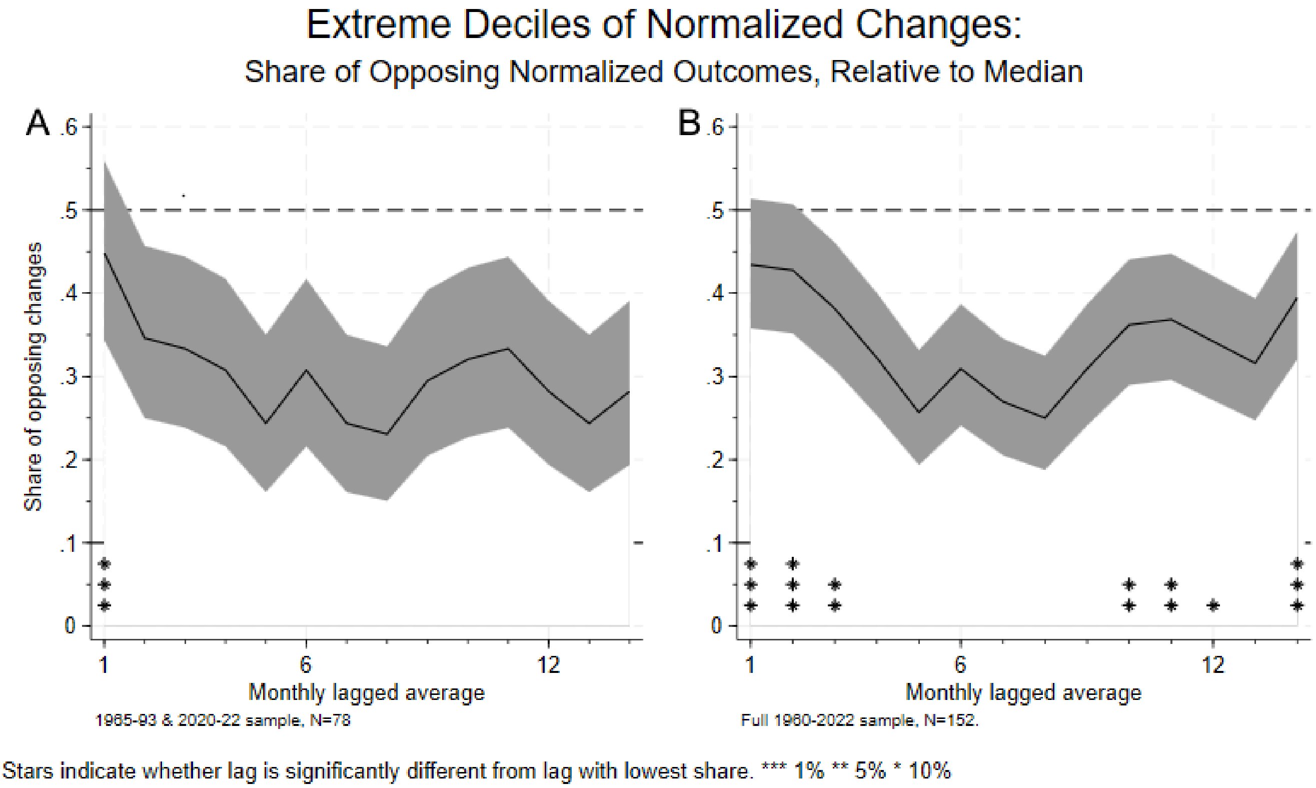
During periods of volatile inflation the standard deviation of the 6 month lag error was 59 basis, relative to the centered 1 year percentage change. The error with respect to the LR smoothed reference series was 61 basis, not statistically different. Errors are leptokurtotic: kurtosis was 5.54 for the centered 1 year percentage change and 5.20 for the LR smoothed reference series, using a specification where the normal distribution has a kurtosis of 3.

6 Best lags for analyzing changes in inflation

Data drips in one observation at a time. This section is intended to assist in the interpretation of newly released inflation reports. Different lagged percentage changes are compared with changes in the underlying inflation trend. Specifically, percentage changes over various lags are compared with contemporaneous changes in the long run centered smoothed series described in Section 4 and analyzed in the fourth column of Tables 2 and 3. The goal is to discover the best lag for determining whether a report of atypically high or low inflation proves a useful indicator of developments in the underlying trend.

Both the lagged and centered series are normalized by subtracting the four year lagged moving average and dividing by the four year lagged moving population standard deviation, contemporaneous observation not included. To compare the reliability of different lagged averages, a sample combining the top and bottom deciles of the normalized changes in the various lagged averages is considered. Figure 2 plots 95%

Figure 2



1965-93 & 2020-22 sample, N=78

Full 1960-2022 sample, N=152.

Stars indicate whether lag is significantly different from lag with lowest share. *** 1% ** 5% * 10%

confidence intervals of the opposite outcome shares for the first 14 lagged percentage changes.[5] If the change in the normalized outcome variable is of the opposite sign of the decile (below or above the median), it is characterized on the graph as an opposing normalized outcome. If signal and outcome are independent, then expected opposite outcomes would be equal to .50, the median, indicated by the horizontal dashed line.

Graph A on the left of Figure 2 covers the subsample with relatively volatile inflation. Graph B covers the full 1960-2022 sample. Asterisks on the bottom indicate whether the lag is significantly different from the lag with the lowest opposing share: this optimum is 8 lags for both graphs.

Conventional wisdom holds that any single data point should not be relied upon on its own. This is good advice: in both samples the 95% confidence interval for one month percentage changes covers the 50% median, where the normalized signal has no predictive power for the normalized outcome. Even the best lag has limited predictive power, producing opposing centered outcomes 23-25% of the time for the partial and full samples respectively. Opposing outcomes can be far from borderline. Annualized optimal 8 lag percentage changes produced smoothed centered outcomes in the opposing quartile 13% of the time for both samples. Since a random variable would produce outcomes in the opposite quartile 25% of the time, this

---

[5] Following Brown, Cai, and DasGupta (2001), confidence intervals calculations are based on Agresti and Coull (1998). Binomial confidence intervals are qualitatively similar.

error rate is substantial.

Analysts should ignore unusually large or small 1 month percentage changes. If the 4 or 8 month lagged percentage change is a surprise, analysts could usefully devote greater attention to that month's report but should remain wary of misleading signals and focus on the lagged 6 or 7 month estimate of inflation's level itself and perhaps the probabilistic model of inflation's direction presented in the next section.

7 Nowcasting Inflation's Direction

Earlier, the lagged and geometrically annualized six month percentage change was shown to possess lower errors when estimating one year centered core PCE inflation whether smoothed or not smoothed. Market observers would also like an estimate of inflation's direction, a sense of whether it is currently rising, falling, or flat. With that in mind, a parsimonious model of the probability that centered inflation has increased over the preceding six months is constructed. The model translates a single price series into a very conditional probability, providing a baseline for more elaborate models.

Geometrically annualized lagged inflation is subtracted from one year centered inflation: this represents the error when the former is predicting the latter. A heavy tailed distribution models the error in the six month figure, using historical estimates of variance and kurtosis. If the kurtosis is significantly different from Gaussian kurtosis (set here to 3) a t-distribution is applied. Relaxing the assumption of normality attenuates extreme probabilities such as .99 or .01.[6]

For the t-distribution, the probability that one year centered inflation has increased over the past six months follows:

$$P(\hat{\Delta}_t > 0) = 1 - t(df, \hat{\Delta}_t/\sigma_{\hat{\pi}}), \tag{7}$$

where $\hat{\Delta}_t = \hat{\Pi}_t^c - \Pi_{t-6}^c$, the change in estimated one year centered inflation over the past 6 months;
df $\in \mathbb{R}_{>4}$; $\sigma_{\hat{\pi}} \in \mathbb{R}_{>0}$. The variable df targets kurtosis: for the Student's t distribution df=2*(3-2*kurt)/(3-kurt). The variable can manage leptokurtotic distributions for 3<kurt<9.[7] $\sigma_{\hat{\pi}}$ is the standard deviation of the estimation error of $\hat{\pi^c}$: it is the scale parameter inside t(df), the cumulative distribution for Student's t with df degrees of freedom.[8]

To evaluate the quality of this model's forecast, the full 63 year sample is divided into two parts. The first half provides estimates of the standard deviation and kurtosis of the estimation error of $\hat{\pi^c}$. The second

---

[6] Stoyanov et al (2011) discuss various approaches to address fat tailed distributions in financial markets.

[7] Krishnamourthy (2006) provides the kurtosis formula for the Student's t distribution. If kurtosis exceeds 9, a mixed normal distribution can be applied: see Paolella (2007). In all cases inspection of the histogram is recommended, preferably with a normally distributed overlay.

[8] See Paolella (2006), p. 241.

half provides a sample space to compare the model probability that inflation has increased over the past 6 months to the proportion of time that centered inflation has in fact increased.

Model probabilities are ordered from low to high and averaged within the five quintiles. This forecast validation method is considerably more conservative than estimating within-sample probabilities and proportions, or estimating proportions based on parameters gleaned from a rolling sample. Each observation is assigned an indicator variable equal to one if one year centered inflation increases and zero otherwise. The proportion of each quintile that inflation increased is then calculated and compared with the average model probability within the quintile.

Section 5 focused on estimating inflation levels during periods of volatile inflation, because analysts care less about those figures when inflation is stable. But the direction of inflation is always of interest, since analysts want to evaluate risks of deflation or unexpectedly high inflation.[9]

For July 1959 to September 1991, the standard deviation of the estimating error is .0056; kurtosis is 6.40. For October 1991 to October 2023 the 5 element vector of average quintile probabilities is [.241 .377 .478 .592 .781]. The vector showing proportions of the probability quintile where one year centered inflation actually increased is [.169 .325 .364 .558 .870]. The difference between the two or the error vector is [-.072 -.052 -.114 -.034 .089]. The vector of two-tailed p-values applying an exact binomial test to each 77 observation subsample is [.145 .411 .052 .564 .072]. Proportions are statistically indistinguishable from model proportions for all quintiles at the 5% level of confidence. However, we can't reject the hypothesis that they differ at the 10% level of confidence for two of five quintiles.

While model probabilities are greater than observed proportions for 4 quintiles; the highest quintile error has the opposite sign. A low powered t-test using the 5 observations doesn't reject the hypothesis that average errors are zero (p=.34). Mean absolute error is .072. Absolute errors as a share of probabilities average to 17%, with a range of 6% to 30%.

Probabilities estimated by this simple model are meaningful, though there is room for improvement. For a simulated sample of 385, equal to the number of observations from October 1991 to October 2023, mean absolute error across quintiles averaged .042 in a 500,000 draw experiment. So 58% of the mean absolute error in the model is due to expected randomness associated with probability draws for a 385 observation sample.[10]

---

[9] The Internet Appendix evaluates probabilities based on the relatively volatile inflation subsample. Perhaps unsurprisingly it shows a larger mismatch between model probability and observed proportion in the middle quintile.

[10] Stated with greater precision, the cross quintile mean absolute error of the simulation was .041818 with a standard error of .000020. This compares with observed 1991-2023 modeling error of .072376. The 58% share still obtains when +/- 2 standard errors are applied to the estimate. The standard deviation of the simulated mean absolute error was .014149: observed modeling error is 2.16 standard deviations away from the simulated .041818 mean absolute error. However, the range of plausible errors for probability draws on a 385 observation sample is quite wide: applying a range of +/- 2 standard deviations produces a share between 19 and 97 percent of the observed modeling error.

8 Centered One Year Inflation vs Lagged Inflation, Jan 2019 – Jul 2024

Many of the themes of this paper are reflected in Figure 3, showing real time geometrically annualized lagged 3, 6, and 12 month PCE inflation alongside centered and revised one year inflation in the top panel, as of 30 Aug 2024.[11] The bottom panel shows the probability that inflation has increased over the preceding 6 months, using lagged 6 month revised inflation and the model presented in Section 7. Low probabilities imply declining inflation. The combined graph brackets the February 2020 start of the COVID era, the rise in inflation in 2021, and the increases in the Federal Funds rate in March 2022 through July 2023, while also covering part of the subsequent decline in inflation.

Six month lagged inflation has a visibly superior fit to centered one year inflation, relative to 3 and 12 month lagged, consistent with Table 3. Three month lagged inflation fluctuates too much to be a reliable indicator of one year centered inflation: average absolute error in Figure 3 was 1.0 percentage point. Lagged 12 month inflation produces an error of 0.8 percentage points relative to the centered one year figure. Six month lagged inflation error was lower at 0.6 percentage points, significantly better than both 3 and 12 month lagged inflation at the 5% level of confidence.

The contemporaneously estimated current probability that centered inflation is rising is derived from Equation 7. The parameters for the t-distribution are based on the differences between centered 1 year inflation and geometrically annualized lagged 6 month inflation using revised data from July 1959 to December 2019. The standard deviation $\sigma_{\hat{\pi}}$ was .0046. Kurtosis was 8.0.[12] In the chart, 19% of the values are over .80 and 16% of the values are less than .20: the probabilities of greater than 80% and less than 20% remain comfortably inside the 95% confidence interval. The 2019-2024 era had more volatile inflation than reflected in the training data: probabilities were in the interquartile range only 29% of the time: the expected 50% share is outside the 95% confidence interval of (.20, .41).

To evaluate whether greater focus on 6 month lagged inflation and this paper's probabilistic metric for inflation's direction will produce novel or conventional views of the economy, Federal Reserve statements from January 2019 to September 2024 are reviewed. Real time unrevised data is used so that the analysis reflects data that markets and the Federal Reserve base their decisions on. PCE price data is released late in the last week of the month following the observation month. CPI data is released earlier, generally on the second week of the month after the observation month. Since CPI data won't be consulted in this narrative, the Federal Reserve's knowledge of existing prices will be underestimated on average.

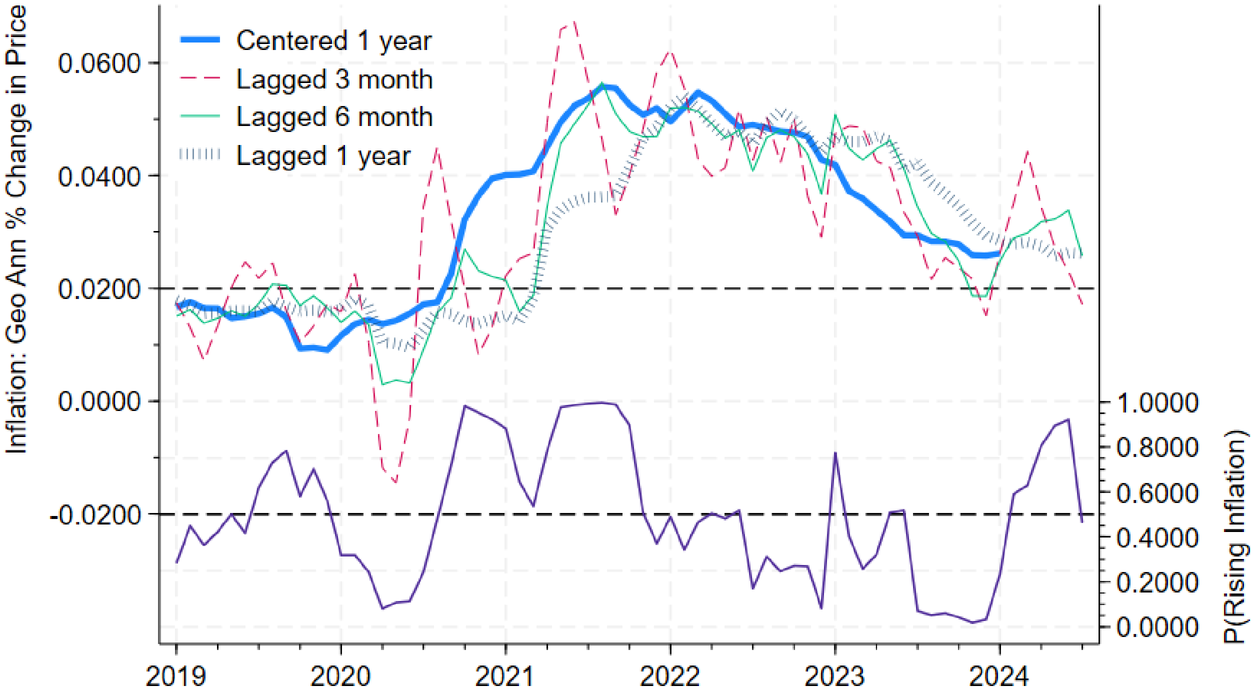
The tools developed in this paper contribute little when 6 month lagged inflation is similar to the

---

[11] The integer lag 6 is used because Table 3 showed that while the 6.1 month decimal lag fit the centered one year return best, that fit was not significantly different from the 6 month lagged percentage change. The 6.5 month decimal lag fit the smoothed 1 year better, but this section focuses on criteria discussed in Federal Reserve statements as well as the business press.

[12] A planned revision of this paper will calculate rising inflation probability using both real time and revised data. Previous sections will be updated as well. (Currently the author's real time dataset starts in 2019.)

Figure 3



Probability of rising inflation over past 6 months based on t distribution, df=5.20 (kurtosis=8.00).

conventionally reported 12 month lagged inflation. This was the case in 2019, when the average difference between the two metrics was 16 basis points (0.16%) and the maximum difference was 39 basis. This was not unusual at the time: the average difference from 1998-2018 was 26 basis, with a standard deviation of 20 basis.[13] When inflation fluctuates, the 2 metrics tend to diverge most though not all of the time.

In January 2020 to July 2024, the average divergence was 52 basis, three times higher than in 2019, while the standard deviation was 48 basis or four times higher than 2019. Median divergence was 34 basis; divergence exceeded one half percentage point in 39% of the months between January 2020 and July 2024. Divergence exceeded one percentage point only 14% of the time. When inflation fluctuates, 6 and 12 month inflation may or may not differ greatly.

In March 2020, the Federal Reserve responded to the COVID-19 pandemic by lowering the upper limit of the federal funds target range from 1.75% to 0.25% and announcing further purchases of treasury and agency securities. Subsequent disinflation over the following 3 months brought 6 month lagged inflation down to the 0.3-0.4% range; the contemporaneous probability that centered inflation had increased over the previous 6 months fell to 9-11%.

On 27 Aug 2020, the Federal Reserve announced a new consensus statement, following extensive external and internal deliberations. The symmetric 2% inflation objective was loosened such that, ”Following

---

[13] 2019 data is real time. Earlier data is revised. Revised 2019 differences were 9 basis, smaller than real time differences.

periods when inflation has been running persistently below 2 percent, appropriate monetary policy will likely aim to achieve inflation moderately above 2 percent for some time." The next month's 16 Sep 2020 statement characterized inflation as, "Running persistently below this longer run goal." That language would be frozen in place during the eight months leading to the 28 Apr 2021 meeting.

Core PCE prices followed a different trajectory. The Fed's audience would understand the 2 percent language to refer to 12 month lagged inflation. But for the more accurate though at the time unknown centered one year metric, inflation would exceed the 2% threshold that very month in September 2020 and never return, at least through the time of this writing. Six month lags would prove to be superior estimators overall: they would break the 2% threshold the following month in October 2020. The lagged 6 month percentage change would remain above 2% through Jan 2021, averaging 2.3%, before dipping to 1.6% and 1.9% in February and March 2021.

In April 2021 both measures would exceed 3% with the 6 month rising from 1.9% to 3.5% and the 12 month rising from 1.8% to 3.1%; centered inflation would be 4.5%. During the two meetings following the 28 May 2021 release of the April data, statement language would be re-oriented to the past tense: it read, "With inflation *having run* persistently below 2 percent". Between the two meetings the June 25th release of the May 2021 data would show 6 month inflation rising to 4.6%, while 12 month inflation would be reported at a more moderate 3.4%: centered inflation would eventually reveal itself to be 5.0%. Greater attention directed towards lagged six month inflation might have provided a useful warning of incipient inflation significantly above the 2% long run target in April 2021 through December 2021 and at least somewhat above target in and after October 2020.

The probability of rising inflation would remain in the top quintile from May 2021 through October 2021, before fluctuating within the middle 3 quintiles for the succeeding 8 months. The Federal Reserve would begin to signal its 16 Mar 2022 rate hike, the first since 2018, in the 15 Dec 2021 statement which finally acknowledged that inflation had, "Exceeded 2 percent for some time." The monthly average of the six month treasury bill rate would rise 8, 18, 31, and 19 basis in the months spanning December 2021 to March 2022, presumably anticipating Fed tightening. These increases came months after the peak in centered one year inflation, which occurred in August and September 2021 at a rate of 5.6%. Six month lagged inflation peaked in August 2021 at a similar 5.7%.

Federal Reserve tightening would occur throughout 2022, with the last increase occurring during the 26 July 2023 meeting. Differences between 6 and 12 month lagged inflation averaged 0.2 percentage points in the 18 months from January 2022 to June 2023. Only once was the difference greater than 0.5 percentage points. With such small differences, 6 month lagged inflation outperformed 12 month lagged inflation only one half of the time: the framework in this paper had less to contribute during the 2022-2023 tightening phase itself. The probability metric would provide some reassurance, staying below 50% in all but 3 months and never entering the top quintile. The mean probability of increasing inflation during the Jan 2022 to Jun

2023 period averaged .39, while its standard deviation was .16, consistent with declining 6 and 12 month lagged inflation.

During the latest period from July 2023 to July 2024, the median difference between 6 and 12 month lagged inflation increased to 75 basis points, creating potential for this paper's framework to add value. In 2023, the 6 month lag implied a faster decline in inflation. Annualized 6 month inflation was more accurate than 12 month lagged inflation in July through October. During the last 2 months of 2023, the 6 month figure dropped to 1.9%, which was potentially misleading as centered one year inflation would resolve to 2.6%: 12 month lagged inflation provided a somewhat more accurate signal of 3.0%. Probabilities of rising inflation hovered near 2.5% and were eventually validated by the centered one year data, which showed declines near one half percent.

The divergence between the lagged metrics narrowed in the first quarter of 2024, before widening in Q2, when the 6 month lag averaged 3.3% and the 12 month lag averaged 2.7%. The 6 month lag implied that inflation was not quite in the vicinity of the Federal Reserve's long run 2% goal, at least until the July PCE price data was released on August 30, 2024. On that date 6 and 12 month inflation were tied at 2.6% and the probability metric was comfortably within the center quintile. On September 18, 2024, the Federal Reserve announced a rate cut of 50 basis points, the first cut since March 2020.

The intent here is not to second guess Federal Reserve policy, which follows a dual mandate of maximum employment and price stability and was addressing extraordinary economic developments during the COVID-19 pandemic. No single time series determines policy and regardless any shortcomings apply equally to the business press. Rather the comparison with Federal Reserve statements is intended to suggest both the potential value added and the limitations of this paper's methods. Overall, use of the 6 month lag and the probability metric might have usefully focused the Federal Reserve on incipient inflation from late 2020 through November 2021, up until higher inflation was acknowledged in the 15 Dec 2021 statement. During the tightening of policy from May 2022 to July 2023 this paper's framework offered less assistance. The final verdict on subsequent months awaits updates in centered 1 year inflation beyond January 2024.

9 Conclusion

Existing inflation consists of price changes between periods equidistant from the present, so analysts are forced to nowcast. For the Personal Consumption Deflator Excluding Food and Energy, percentage changes relative to 6 or 7 months ago best approximate 1 year centered inflation after geometric annualization is applied. They also approximate smoothed 1 year centered inflation well, and have mean absolute errors only a little worse than the optimum for the 2 year centered reference series. Simple annualization inflates with a multiplicative factor, while geometric annualization captures the compounding process. A sequence of geometrically annualized and identical monthly percentage changes will thus produce the same figure as the resulting annual percentage change. Simple annualization is usually a decent approximation for geometric annualization, though the error exceeded 20 basis points in 10% of the sample.

Centered 1 year inflation data is available 6 months afterwards: combining that with the estimates of existing centered inflation provides a probabilistic estimate of whether inflation has increased or decreased over the prior 6 months. Use of the Student's t distribution rather than the Gaussian attenuates extreme probability predictions. Averaging predictions within quintiles permits validation of the univariate model. The conventional wisdom that a single unusually large or small data point should be ignored is upheld. An unusually large change in 4 to 8 month lagged annualized inflation deserves some attention though it is far from a conclusive indicator.

When 6 and 12 month lagged percentage changes are close to one another, the 6 month metric will naturally contribute little additional insight. This was the case in 2019, during most of the Great Moderation, and even parts of the 2021-2024 era when inflation increased, plateaued, and decreased. In late 2020 to late 2021, greater attention directed at the 6 month lag could have provided warning of incipient inflation.

A planned revision of this paper will calculate rising inflation probability using both real time and revised data for years before 2019.

The length of the 1 year reference series was chosen to match convention. Future work will compare 5 to 13 month centered reference series. In addition, the robustness of the 6 or 7 month lag guideline will be tested against other price series in the US and other countries. The author is also interested in extending the framework to monthly output series.

References

[1] Adjemian, Michael K., Li, Qingxiao and Jo, Jungkeon, 2023. ”Decomposing Food Price Inflation into Supply and Demand Shocks” Available at https://tinyurl.com/3fnzek

[2] Brown, M. B., and A. B. Forsythe. 1974. Robust tests for the equality of variances. Journal of the American Statistical Association 69: 364–367. https://www.jstor.org/stable/2285659

[3] Cowley, Cortney and Scott, Francisco, 2022. ”Commodity Prices Have Limited Influence on U.S. Food Inflation,” Federal Reserve Bank of Kansas City Economic Bulletin.

[4] Federal Reserve. *Federal Open Market Committee announces approval of updates to its Statement on Longer-Run Goals and Monetary Policy Strategy*, August 27, 2020, retrieved from Federal Reserve; https://www.federalreserve.gov/newsevents/pressreleases/monetary20200827a.htm, September 23, 2024.

[5] Krishnamoorthy, K. *Handbook of Statistical Distributions with Applications*, Chapman and Hall. 2006.

[6] Lombra, Raymond E. and Mehra, Yash P., 1983. ”Aggregate demand, food prices, and the underlying rate of inflation,” Journal of Macroeconomics, Elsevier, vol. 5(4), pages 383-398.

[7] Luciani, Matteo, and Riccardo Trezzi, 2019. ”Comparing Two Measures of Core Inflation: PCE Excluding Food and Energy vs. the Trimmed Mean PCE Index,” FEDS Notes. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/2380-7172.2390. W

[8] Paolella, Marc S. *Fundamental Probability: A Computational Approach*, Wiley. 2006.

[9] Paolella, Marc S. *Intermediate Probability: A Computational Approach*, Wiley. 2007.

[10] Smith, Aaron David., 2022 ”Agricultural Prices Aren't Driving Inflation”. Ag Data News, https://asmith.ucdavis.edu/news/agricultural-prices-arent-driving-food-inflation

[11] StataCorp. 2023. *Stata Statistical Software: Release 18*. StataCorp LLC.

[12] Stoyanov, Stoyan V., Rachev, Svetlozar T., Racheva-Yotova, Boryana, Fabozzi, Frank J., 2011. ”Fat-Tailed Models for Risk Estimation,” The Journal of Portfolio Management.