Finance and Economics Discussion Series

Federal Reserve Board, Washington, D.C.
ISSN 1936-2854 (Print)
ISSN 2767-3898 (Online)

Common and idiosyncratic inflation

Hie Joo Ahn and Matteo Luciani

2020-024

Please cite this paper as:
Ahn, Hie Joo, and Matteo Luciani (2024). "Common and idiosyncratic inflation," Finance and Economics Discussion Series 2020-024r1. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2020.024r1.

Common and Idiosyncratic Inflation

Hie Joo Ahn
Federal Reserve Board
hiejoo.ahn@frb.gov

Matteo Luciani
Federal Reserve Board
matteo.luciani@frb.gov

Abstract

We disentangle price changes due to economy-wide shocks from those driven by idiosyncratic shocks by estimating a two-regime dynamic factor model with dynamic loadings on a new large dataset of finely disaggregated monthly personal consumption expenditures price inflation indexes for 1959-2023. We find that up to the mid-1990s and after the Covid pandemic, common shocks were the primary driver of US inflation dynamics and had long-lasting effects. In between, idiosyncratic shocks were the main driver, and common shocks had short-lived effects.

JEL classification: C32, C43, C55, E31, E37

Keywords: Core inflation, Dynamic factor model, disaggregated consumer prices, monetary policy

*We are grateful to Jeremy Rudd, Andrew Figura, Lucas Moyon and seminar participants at the Federal Reserve Board, Bank of Italy, European Central Bank, Bank of England, and Reserve Bank of Australia. Any errors are our responsibility.

1 Introduction

Price stability is one part of the Federal Reserve's dual mandate. For achieving this goal, it is crucial to correctly identify the portion of price changes driven by macroeconomic shocks—that is, those shocks that affect all prices and thus change the general price level of goods and services—from that resulting from transitory or item-specific developments. Indeed, policymakers should respond only to the former while avoiding responding to the latter or—even worse—to measurement error.

With this goal in mind, we develop common core inflation, a measure that isolates price changes driven by economy-wide shocks from those led by idiosyncratic shocks. Common core inflation is based upon a dynamic factor model estimated on a new large dataset of finely disaggregated PCE price indexes suitable for factor-model analysis.

Given its importance, economists have employed similar approaches to produce measures of underlying inflation (e.g., Cristadoro et al., 2005; Stock and Watson, 2016) and these measures are currently used in policy institutions.[1] This paper differs from previous research in two key aspects. First, we estimate a two-regime dynamic factor model with dynamic loadings estimated using frequentist methods. We chose a "two-regime" dynamic factor model with "dynamic loadings" because the properties and the dynamics of the covariance and spectral density matrices of the data suggest that this is an appropriate representation of disaggregated PCE prices. Moreover, based on the evolving covariance structure of the data, we objectively and promptly identify regime changes. For example, we show that the model provided evidence of a regime switch by August 2021, just a few months after the inflation ramp-up started.

Our model is more restrictive than one with time-varying volatilities and parameters estimated using Bayesian methods (e.g., Stock and Watson, 2016) because volatilities and parameters are fixed within a regime. Nonetheless, we show that our estimate of common core inflation from the two-regime model is very close to the estimate of common trend inflation obtained by fitting Stock and Watson's 2016 model on our dataset. This observation suggests that regime changes combined with dynamic loadings capture the same important features of disaggregated price inflation captured by time-varying volatilities and parameters.

[1] For example, the Multivariate Core Trend inflation from Federal Reserve Bank of New York essentially adopts Stock and Watson's multivariate model and updates the estimates regularly. See the description `https://www.newyorkfed.org/research/policy/mct#--:overview`. Likewise, the Underlying Inflation Gauge (UIG) of Amstad et al. (2017) that the NY Fed discontinued in October 2023 adopted the Cristadoro et al. (2005) model.

Second, we construct a new and unique dataset of finely disaggregated PCE prices composed of about 140 series from 1959 to 2023 suitable for estimating dynamic factor models. Due to how the Bureau of Economic Analysis constructs the PCE price indexes, there is spurious cross-sectional correlation in the disaggregated data, which deteriorates the estimate of the model (see the discussion in Boivin and Ng, 2006 and Luciani, 2014). We take care of this issue and construct a dataset that does not exhibit spurious cross-sectional correlation while preserving the structure of the PCE consumption basket—datasets employed in previous studies are either small (e.g., Stock and Watson, 2016) or do not cover the PCE basket, hence direct aggregation based on the item weights is not possible (e.g., Reis and Watson, 2010). Having such a large and carefully cleaned dataset has the main advantage that the co-movement in the data—what we are after—and its evolution—what we leverage to estimate the probability of being in a regime or the other—are very well estimated.

Our analysis unveils additional and important characterization of the relevant features of US inflation dynamics. First, Stock and Watson (2007, 2016) show that there has been a marked reduction in the volatility of the shocks driving trend inflation before and after the 1990s. We show that this change in volatility is the result of a change in the amount and the persistency of the comovement among disaggregated prices. Specifically, we show that US inflation dynamics experienced two regimes, a long-memory regime up to the mid-1990s, and a short-memory regime from the mid-1990s to the Covid Pandemic. In the long-memory regime, inflation dynamics are primarily driven by common shocks, and the effect of these common shocks is long lasting; in the short-memory regime, inflation dynamics are primarily driven by idiosyncratic shocks, and the effect of the common shocks is short lived. Thus, when inflation is in the long-memory regime, monetary policy should promptly and decisively respond to inflation dynamics, while when inflation is in the short-memory regime, monetary policy can focus on stabilizing the real economy.

Second, Eo et al. (2023) show that since the 1990s, the dynamic of US trend inflation has been entirely dominated by services inflation. Our results add supporting evidence to this finding by showing that from the mid-1990s through the Covid pandemic, the dynamic of nearly all subcomponents of core goods price inflation became almost entirely idiosyncratic, and the commonality in the data reflected primarily the dynamics of housing services prices.

Third, we show that inflation fluctuations after the pandemic are consistent with the long-memory regime, as goods and service inflation started again to comove and show persistent dynamics. This finding is consistent with the fast and furious cycle of monetary policy tightening that occurred between March 2022 and April 2023.

Our analysis of disaggregated PCE prices presents results entirely novel in the literature and provides a new perspective for understanding how macroeconomic shocks propagate through sectoral price changes and how the nature of this propagation evolves. Specifically, we go beyond what Stock and Watson (2016) and Eo et al. (2023) did by explicitly analyzing the distribution of the commonality share—i.e., how much variance of each disaggregated price is explained by the common component. Our findings suggest that as inflation dynamics transitioned from the long-memory regime (prior to the mid-1990s) to the short-memory regime (mid-1990s to 2019), the distribution of the commonality share shifted toward zero, indicating a general increase in idiosyncratic behavior. Moreover, we find that after the onset of the pandemic, although commonality in goods prices rose across the board, idiosyncrasy in some core services, namely education and health care, increased heavily. As a result, the distribution of commonality shares became more dispersed.

The rest of the paper is organized as follows. Section 2 introduces the new dataset and summarizes the key features of the data, and Section 3 presents the model. Section 4 presents the empirical analysis, and Section 5 compares common core inflation and trend inflation from the Stock and Watson (2016) model. Section 6 concludes. This paper includes a lengthy Appendix containing material that could have be in the main body of the text. In Appendix A we provide an in depth analysis of the relevance of having two regimes and secular time-varying trends in the model. In Appendix B, we discuss inflation dynamics during the Covid lockdowns, the advantage of having a large number of variables, and the real-time reliability and the forecasting performance of common core inflation. Appendix C provide robustness results. Lastly, Appendix D provide an in-depth literature review.

2 Data

This section discusses our new dataset of disaggregated PCE prices and its main features. Section 2.1 details the construction of the dataset. Section 2.2 provides a formal statistical examination of the time-series properties of disaggregated inflation data.

2.1 A new dataset of disaggregated PCE prices

PCE price data are available at different levels of disaggregation, the highest of which includes roughly 220 price indexes, with a complete set of observations available since 1990. Our reference starting point is the disaggregation the Dallas Fed uses to produce the Trimmed Mean PCE inflation index (see Dolmas, 2005). The dataset comprises 178 disaggregated prices, the highest level of disaggregation that produces a balanced panel of data beginning from the late 1970s.

Disaggregated PCE prices can be classified as "market-based" and "nonmarket-based." According to the BEA, market-based prices are defined "as those goods and services that have been produced for sale at prices that are economically significant" and, hence, "their current market price provides a rational and viable basis for valuing" them (Bureau of Economic Analysis, 2017, pp. 2–5). Nonmarket-based prices consist of prices of "goods and of individual or collective services that are produced by nonprofit institutions and by government that are supplied for free or at prices that are not economically significant" (Bureau of Economic Analysis, 2017, pp. 2–5). Services in this category are provided by businesses either without charge or for a small fee, whose prices do not reflect the entire value of the service.[2] In other words, a "market-based" good/service can be actually bought and, hence, it is possible to record a price for it, while a "nonmarket-based" good/service cannot be bought and, hence, its price is imputed by the BEA based on the costs of production (for nonprofit institutions and government) or some other assumptions (for business).[3]

Market-based goods and services are about 87% of total PCE. Most of them are constructed by taking the corresponding (or conceptually closest) CPI, with only a few exceptions where a corresponding PPI series is used (for example, airfares and some medical prices). By contrast, most nonmarket-based prices are imputed by the BEA, with just a few exceptions constructed out of the corresponding CPIs and/or PPIs. Because there is not always a corresponding CPI or PPI for each PCE price, some disaggregated PCE prices are constructed out of the same CPI or PPI index and hence are identical (or nearly

[2]For example, education and health services provided by non-profit institutions are typically provided at below-market prices. Another example is checking account maintenance, which is often provided by banks without charge.

[3]An example here could help: one of the consumption categories is "lotteries," but what is the price for lotteries? For example, suppose John buys a scratch lottery ticket for, say, $2, and suppose John does not win. Now, John has consumed $2 in participation in a lottery, but what is the price that John paid? In this case, the BEA imputes the PCE price index for "lotteries" using the overall CPI. Another example is "standard clothing issued to military personnel,", which is imputed by using the PPI for "apparel."

so).[4]

In the level of disaggregation used by the Dallas Fed, we identify 21 price groups constructed from the same CPI/PPI. Price changes in the same group exhibit an almost perfect correlation. This data environment poses a significant challenge to estimating dynamic factor models because such models are estimated under the assumption that the idiosyncratic components are only mildly cross-sectionally correlated. If two price changes are (almost) perfectly correlated due to the data construction, as in the current PCE price data, the idiosyncratic components will also be almost perfectly correlated. The violation of this key assumption makes identifying the common and idiosyncratic components likely biased and unreliable because the excess of fictitious correlation in the idiosyncratic components is mistaken for co-movement in the data (Boivin and Ng, 2006; Luciani, 2014).

To get around this problem, we aggregate the 53 price indexes constructed from the same sources into 21 alternative price indexes exhibiting distinct variations. This operation leaves us with 146 disaggregated PCE prices. At this stage, most of the disaggregated price changes are available from 1959 except for four categories available only starting from the 1970s. For these four categories, we use higher-level aggregates available from 1959. As a result, our dataset includes 142 disaggregated PCE price inflation rates from January 1959 to December 2023. The Complementary Appendix provides the list of variables in the dataset as well as details about the data construction.[5]

2.2 Features of the data

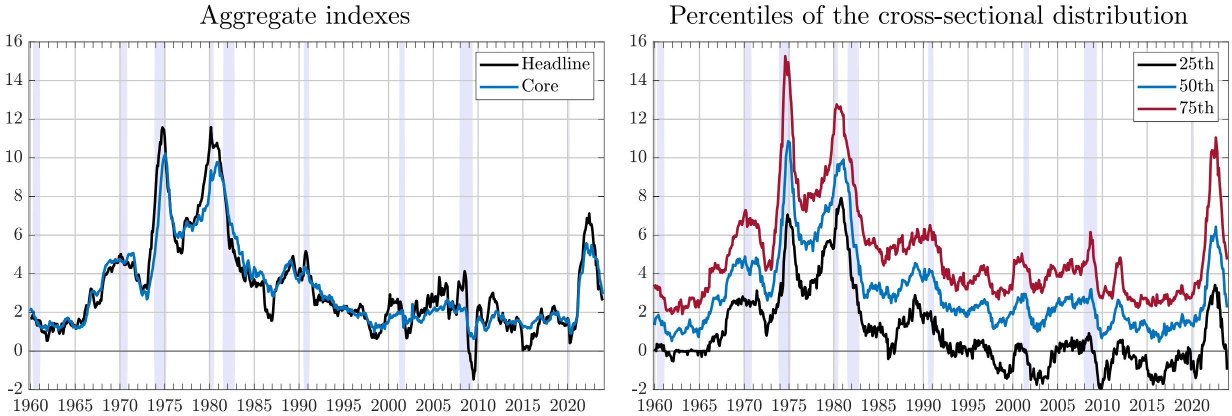
In the past several decades, the U.S. economy underwent structural changes that have altered the time-series properties of consumer price inflation. By simply looking at the time series of total and core PCE price inflation (left chart, Figure 1), we see that inflation was very persistent from (approximately) the mid-1970s to the early-1990s, while inflation has been very stable since the 1990s. We observe similar patterns from the cross-sectional distribution of disaggregated PCE prices (right chart, Figure 1).

In the macroeconomic literature, "persistency" is typically defined in terms of $I(0)$

[4]Examples are the PCE price indexes for "Bicycles and accessories," "Pleasure boats," "Pleasure aircraft," and "Other recreational vehicles," which are all constructed out of the CPI "Sports vehicles including bicycles."

[5]The price data are taken from the National Income and Product Accounts (NIPA) Table 2.4.4U, while the nominal quantity data necessary to compute the weights are taken from the NIPA Table 2.4.6U. The data were downloaded from the BEA website on March 14, 2024.

Figure 1: Year-over-year PCE price inflation



Notes: The right chart reports the percentiles of the cross-sectional distribution of 142 disaggregated PCE price inflation.

or $I(1)$ processes. However, processes can be fractionally integrated, that is, $I(d)$ with $d \in (0,1)$ (Granger and Joyeux, 1980; Hosking, 1981). A fractionally integrated process has long memory, that is, its autocorrelation function decays slowly; an $I(0)$ process has short memory, that is, its autocorrelation function decays exponentially; and an $I(1)$ process, we can say that has infinite memory, that is, its autocorrelation function is flat. The larger d is, the more persistent the process is. The remainder of this section provides formal statistical evidence indicating that disaggregated inflation data experienced a change in the order of integration from a very persistent to a much less persistent regime. In other words, the order of integration $d \in [0,1]$ has decreased meaningfully over time.[6]

The left chart in Figure 2 reports the percentage of variance explained by the 10 largest eigenvalues of the covariance matrix of disaggregated PCE price inflation, $\mathbf{\Gamma}^{\pi}$.[7] We estimate this statistic on a 25-year rolling window to keep track of low-frequency changes in the structure of the covariations between disaggregated price changes. As shown in the left panel, the share of variance explained by the largest eigenvalue has decreased dramatically over time. Two factors can explain this pattern. First, the commonality in the data has decreased over time. Second, one long-memory (and perhaps nonstationary) or $I(1)$ factor (a.k.a., common trend) accounts for the persistency in disaggregated data in the early part

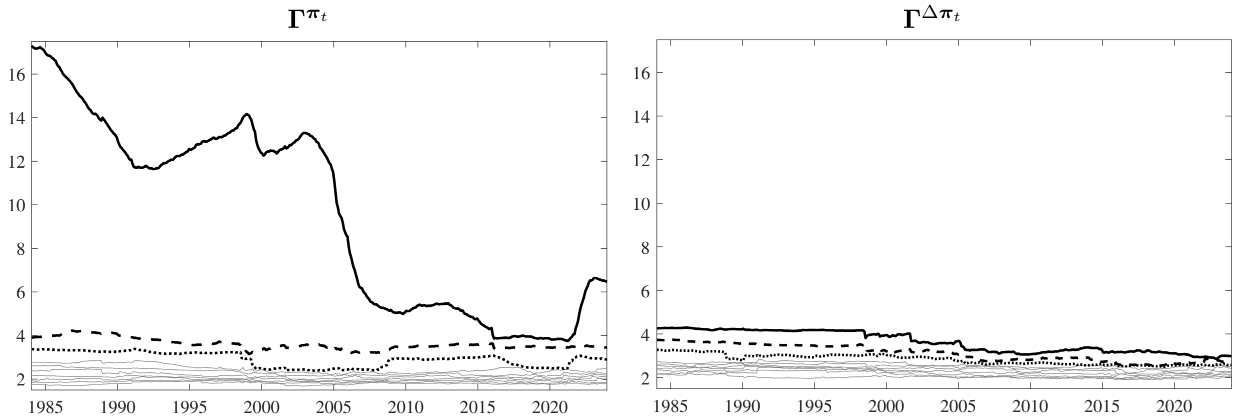
[6]In addition to $d = 0$ and $d = 1$, there are two regions of particular interest for d, $0 < d < \frac{1}{2}$ and $\frac{1}{2} \geq d < 1$. When $0 < d < \frac{1}{2}$, the process has long memory, but it is stationary; when $\frac{1}{2} \geq d < 1$, the process has long memory, and it is nonstationary meaning that Y_t has infinite variance, but mean reverts in the sense that the impulse response function is slowly decaying.

[7]We look at the largest eigenvalues of the covariance matrix because they are related to the comovement in the data. Indeed, all the criteria that we will discuss in Section 3.3 and Appendix B.6 to determine the number of factors in the model are based on the behavior of the eigenvalues of either the covariance matrix or the spectral density matrix.

of the sample, but the persistency weakened later in the sample. In other words, there has been a change in the order of integration of the common component.

To determine which of these two hypotheses is likely correct, we look at the covariation structure of the first-differenced disaggregated PCE price inflation, $\mathbf{\Gamma}^{\Delta \boldsymbol{\pi}_t}$. The rationale for first-differencing the data is to control for the potential nonstationarity in the data. If we no longer observe a declining fraction of the variance accounted for by the largest eigenvalue, we conclude that the decreased persistency in the disaggregated data accounts for the pattern shown in the left chart in Figure 2, which would be evidence that a regime change actually occurred. Otherwise, we interpret that the commonality has reduced over time, which would not necessarily be caused by a regime change. The right chart in Figure 2 suggests that the former is the likely answer. Although there has been some reduction in commonality, this reduction is far less dramatic when compared with the left chart. This observation indicates that disaggregated PCE prices were in a long-memory (and perhaps nonstationary) or $I(1)$ regime up to sometime in the early 1990s but then switched to a short-memory (i.e., $I(0)$).[8]

Figure 2: Share of variance explained by the largest eigenvalues



Notes: The left chart reports the percentage of variance explained by the 10 largest eigenvalues of the covariance matrix of disaggregated PCE price inflation. The right chart reports the percentage of variance explained by the 10 largest eigenvalues of the covariance matrix of the first difference of disaggregated PCE price inflation. In both charts, we estimated the covariance matrices over 25-year rolling windows. As we explain in Section 4.1.3 and in Appendix B.1, we excluded the data from March to August 2020 from the computation because of the Covid lockdown and the reopening.

[8]There exists a small literature that has estimated fractionally integrated models on inflation data across the world. This literature overwhelmingly concludes that inflation data is a long-memory and mean-reverting process (Baillie et al., 1996; Gadea and Mayoral, 2006; Canarella et al., 2020).

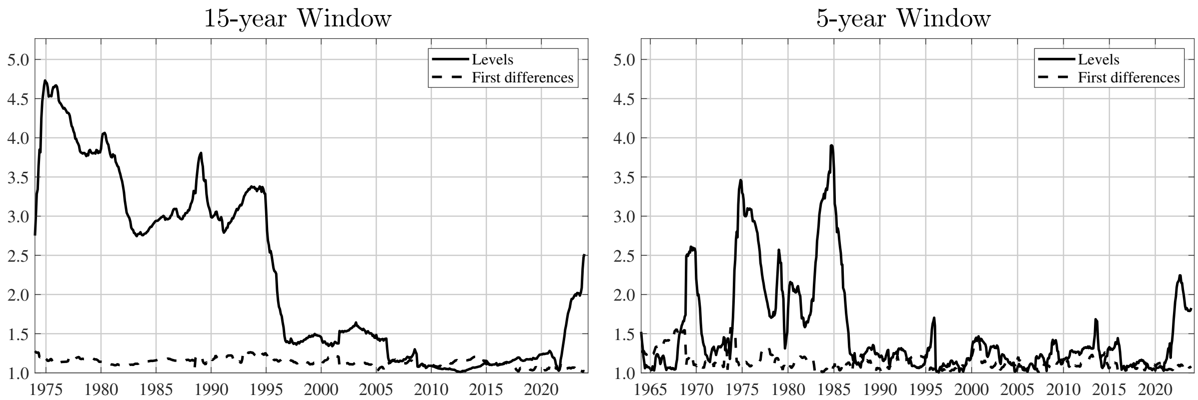
2.3 Detection of changes in the regimes

In Section 2.2, we showed that a regime change in disaggregated inflation dynamics occurred sometime in the 1990s. In this Section, we address the issue of when this change occurred.

Figure 3 shows the ratio of the largest and the second largest eigenvalue of $\mathbf{\Gamma}^{\pi}$ at each point in time estimated using windows of different lengths. We denote this statistic as z_t^w, where the superscript w indicates the window length in years. Given the results in Figure 2, we expect z_t^w to drop and remain at a low level when the disaggregated dynamics enter the short-memory regime from the long-memory regime. In contrast, if the dynamics switch from the short-memory regime to the long-memory one, we expect z_t^w to rise prominently and stay elevated.

The left panel in Figure 3 shows z_t^{15}, while the right panel shows z_t^5—we choose the 15-year window to capture low-frequency structural changes in the disaggregated inflation dynamics, while we chose the 5-year window to capture more sudden changes. As shown in Figure 3, z_t^{15} dropped suddenly in the mid-1990s and stayed low afterward, while z_t^5 spiked in 2021 and stayed elevated relative to the pre-pandemic level. This result indicates that disaggregated PCE prices became a short-memory process around the mid-1990s, and may have revert to the long-memory regime in 2021, as seen in the ramp-up in inflation from 2021.

Figure 3: Ratio of largest over second-largest eigenvalue



Notes: As we explain Section 4.1.3 and in in Appendix B.1, we excluded the data from March to August 2020 from the computation because of the Covid lockdown and the reopening.

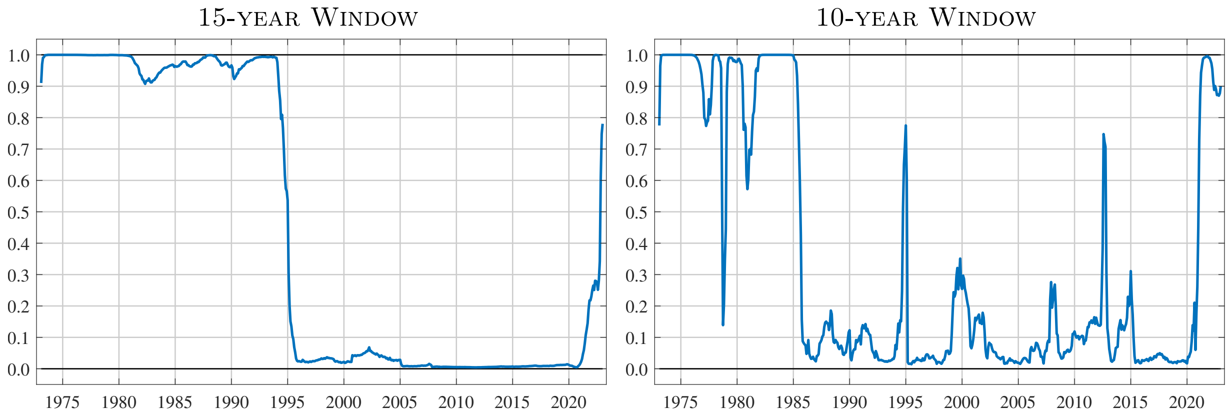
Next, we show that we can modify this heuristic approach based on z_t^w to get a formal data-driven detection of changes in the regime. We assume that there are two inflation regimes, and we model the probability that inflation is in each regime following the approach of Auerbach and Gorodnichenko (2012). Let z_t^w be our *regime determinant* variable, then the probability that inflation is in the long-memory regime is

$$p_l(z_t^w) = \frac{1}{1 + exp(-\gamma \tilde{z}_t^w)},$$

where $\tilde{z}_t^w$ is z_t^w standardized to have mean zero and variance one, and γ is set by the econometrician to achieve the desired smoothness in the estimated probability.[9] To minimize ad-hoc adjustments, we experimented with different values for γ used in the literature and settled on $\gamma = 5$.

Figure 4 reports the probability of being in a long-memory regime. There exists a clear trade-off between precision and timeliness of detection determined by the length of the window. The shorter the window, the more timely but less precise is the detection (meaning the higher the chances of a false positive); the longer the window, the less timely but more precise is the detection. Thus, our rule of thumb is to favor a longer window when trying to detect a gradual regime change, but to favor shorter windows when trying to detect a sharp regime change such as Covid.

Figure 4: Probability of inflation being in the long-memory regime



Following our rule-of-thumb, the results in Figure 4 support the conclusion in Figure 3. US inflation dynamics were in the long-memory regime up to the mid-1990s, but hen switched to the short-memory regime until the pandemic. Sometime after Covid, inflation dynamics returned to the long-memory regime.

In light of these results, we will estimate the model by imposing a regime switch in 1995.

[9]For instance, Auerbach and Gorodnichenko (2012), which use the standardized value of the output gap as *regime determinant* variable, set γ to 1.5 to capture that the economy is in a recession for 20 percent of the sample period.

However, for the post-pandemic period, we will initially be agnostic about the regime and assume that the post-pandemic could belong to either regime. We will then address the question of when the regime change could have been detected in Section 4.1.3.

3 Methodology

This section discusses the construction of common core inflation. Section 3.1 introduces the concept of common and idiosyncratic inflation. Section 3.2 illustrates the dynamic factor model used to extract these two components from disaggregated PCE price inflation data.

3.1 Defining common and idiosyncratic inflation

This analysis aims to evaluate what portion of core inflation is driven by shocks that affect all prices (macroeconomic fundamentals) and what portion is driven by idiosyncratic price movements based on a statistical method. Our methodology involves two steps.

In the first step, we decompose changes in each individual price into two components: the *common* component and the *idiosyncratic* component. The common component captures price changes attributable to economy-wide (that is, *common*) factors, such as the economic slack or movements in the input prices of goods and services. The common component has pervasive effects across disaggregated price changes, but the magnitude and dynamic features of the effect may vary cross-sectionally. The idiosyncratic component captures price changes driven by sector-specific developments or measurement errors.[10] The idiosyncratic price changes are specific to an individual price series or a particular subset of series.

Formally, let $\pi_{it} \equiv 100 \times (\frac{P_{it}}{P_{it-1}} - 1)$ be the month-over-month inflation rate, with possibly slowly moving time-varying mean (or secular trend) μ_{it}. We then have

$$\pi_{it} = \mu_{it} + \chi_{it} + \xi_{it}, \tag{1}$$

where χ_{it} is the *common* component, and ξ_{it} is the *idiosyncratic* component.

[10] An example of idiosyncratic price change is the plunge (52% at an annual rate) in the price index for wireless telephone services March 2017, which shaved off about 8 basis points from the monthly percent change in core PCE prices. The plunge was due to a methodological change to the measurement of wireless services in the CPI and the fact that in late February 2017, Verizon and AT&T (which in March 2017 accounted for nearly 70% of wireless subscriptions in the U.S.) brought back unlimited data plans.

In the second step, after estimating the common component of each series, we aggregate them to construct the common component of core inflation by using each series' weight in the core PCE price index. In this way, we estimate "common core inflation" (χ_t^c), defined to be the portion of core inflation attributed to common (macroeconomic) factors:

$$\chi_t^c = \sum_{i \in core} w_{it}(\mu_{it} + \chi_{it}) \tag{2}$$

where w_{it}^c are the "approximate" core PCE weights—the core PCE weights are computed by setting the weights for food and energy prices to zero and reweighting the PCE weights appropriately.[11]

Note that we include the time-varying mean in common core inflation. This is common practice because, by doing so, common core inflation has a level comparable to that of the published core PCE price inflation. For example, Stock and Watson, 2016 define "trend inflation" as the sum of sector-specific and common trends. Likewise, in the literature on large stationary dynamic factor models, the default strategy is to center the variables (i.e., subtracting the sample mean) before estimating the model—using our notation, we would have $\mu_{it} = \mu_i$—and then attribute back the sample mean to the common component. Now, in this literature, μ_i is just a shift parameter. In contrast, in our setting, μ_{it} is time-varying, so attributing it to the common component might be a stretch depending on how much time variation μ_{it} exhibits. However, this is not a problem for us because, as we discuss in Section 3.2, we limit the time variation in μ_{it} so that it will be a slow-moving mean (see also Appendix A.3), and hence we are fine incorporating it in the definition of common core inflation.

Finally, it is worth emphasizing that we estimate the dynamic factor model on a dataset of PCE prices that preserves the structure of PCE and thus includes food and energy prices.

[11] We use approximate weights because the PCE price index is a Fisher index, and as such, it has the drawback of the nonadditivity property (see Whelan, 2002, as well as Chapter 4 of the NIPA Handbook, Bureau of Economic Analysis, 2017). Therefore, only approximate weights can be computed. To compute the "approximate" PCE weights we follow Dolmas (2005) and set:

$$w_{it} = 0.5 \frac{Q_{it-1} P_{it-1}}{\sum_{i=1}^{N} Q_{it-1} P_{it-1}} + 0.5 \frac{Q_{i,t} P_{it-1}}{\sum_{i=1}^{N} Q_{i,t} P_{it-1}}, \tag{3}$$

where Q_{it} is real consumption of item i at time t and P_{it} is the corresponding PCE price index. Note that these weights are very similar to the weights of the Törnqvist index—the Törnqvist index weights have the same expression as in (3) but for the last term in which P_{it-1} is replaced by P_{it}. This is not surprising because Diewert (1976, 1978) shows that a Törnqvist index numerically approximates a Fisher index (see also Dumagan, 2002).

Therefore, our model can capture potential spillovers from food and energy prices to core prices, although food and energy prices are not explicitly included in the aggregation process to produce common core inflation.

3.2 Dynamic factor models

Formally, we consider the following dynamic factor model:

$$\pi_{it} = \mu_{it} + \chi_{it} + \xi_{it} \qquad i = 1, \dots, N, \ \ t = 1, \dots, T, \tag{1}$$

$$\mu_{it} = \mu_{i,t-1} + \eta_{it}, \qquad \eta_{it} \sim \mathcal{N}(0, \sigma^{2(r)}_{\eta_i}), \tag{4}$$

$$\chi_{it} = \sum_{k=0}^{s} \boldsymbol{\lambda}^{(r)}_{ik} \boldsymbol{f}_{t-k}, \tag{5}$$

$$\boldsymbol{f}_t = \sum_{\ell=1}^{p} \boldsymbol{\mathcal{A}}^{(r)}_{\ell} \boldsymbol{f}_{t-\ell} + \mathbf{u}_t, \qquad \mathbf{u}_t \overset{i.i.d.}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}^{(r)}_u), \tag{6}$$

$$\xi_{it} = \sum_{\jmath=1}^{d_i} \rho^{(r)}_{i\jmath} \xi_{it-\jmath} + e_{it} \qquad \mathbf{e}_t \overset{i.i.d.}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}^{(r)}_e), \tag{7}$$

where $\boldsymbol{f}_t = (f_{1t} \cdots f_{qt})'$ are the q common latent factors capturing co-movements across series and across time; $\boldsymbol{\lambda}_{ik} = (\lambda_{i1k} \cdots \lambda_{iqk})$ are the factor loadings for price i at lag k; $s \geq 0$, $p \geq 1$, and d_i are finite integers; $\boldsymbol{\Gamma}_u$ is a $q \times q$ positive definite covariance matrix with full rank; and $\boldsymbol{\Gamma}_e$ is an $n \times n$ positive definite covariance matrix with full rank. Shocks to the common and idiosyncratic components are denoted by $\mathbf{u}_t = (u_{1t} \cdots u_{qt})'$ and $\mathbf{e}_t = (e_{1t} \cdots e_{nt})'$. The superscript r indicates that the parameters change across different regimes.

In light of the results in Sections 2.2 and 2.3, we allow the possibility of two regimes: a pre-1995 and a post-1995 regime. That is, as we explain in Section 3.2.1, we estimate twice the model in the two samples without imposing any constraint to differentiate the two regimes so that the data will tell us what these two regimes are.[12]

In addition, we further employ the following assumptions (for a rigorous treatment of this model, see Barigozzi and Luciani, 2019, 2020a,b):

(i) The common factors $\boldsymbol{f}_t$ are pervasive—that is, they have non-negligible effects on all

[12] In Appendix A.1, we show that this assumption is crucial and that failing to switch regimes when the data moves from the long-memory to the short-memory regime leads to an incorrect representation of inflation dynamics. Moreover, in Appendix C.1, we show that our results do not change if we change the switching point by a few years.

variables at least at one lag, and they can be nonstationary and long-memory.

(ii) The idiosyncratic shocks are weakly cross-sectionally correlated—that is, they do not have a pervasive effect. Moreover, all the idiosyncratic components $\boldsymbol{\xi}_t = (\xi_{1t} \cdots \xi_{nt})'$ are stationary—that is, the (potential) higher order persistency around $\boldsymbol{\pi}_t - \boldsymbol{\mu}_t$ comes from the common component.

(iii) $\mathbf{u}_t$, $\boldsymbol{\eta}_t$, and $\boldsymbol{e}_t$ are independent of each other.

(iv) We impose $\sigma_\eta^2 = \frac{1}{600}$, which implies that the expected change in μ_{it} over 50 years has a standard deviation equal to the standard deviation of π_{it}.

Assumptions (i)–(ii) are enough to identify the common from the idiosyncratic component, while assumption (iv) is necessary to identify the idiosyncratic trend.[13] By imposing a small variance, Assumption (iv) defines the idiosyncratic trend as a slow-moving mean (see also Appendix A.3).[14]

Lastly, it is necessary to clarify that we are not imposing that the loadings in (5) are dynamic. Rather, we are allowing this possibility. Section 3.3 shows that $s > 0$ is a feature of disaggregated PCE price inflation data. Moreover, in Section 4.1.1, we show that the factor loadings are the main parameter that changes across the two regimes because of their link to the persistence of the common components.

3.2.1 Estimation

We estimate model (1), (4)–(7) by Quasi-Maximum Likelihood, implemented through the Expectation-Maximization (EM) algorithm. The EM algorithm is an iterative method to find maximum likelihood estimates of parameters in models with unobserved latent variables. In the case of model (1), (4)–(7), and abstracting from the presence of two different regimes for the moment, at any iteration $\kappa > 0$, in the E-step, given an estimate of the parameters $\widehat{\boldsymbol{\lambda}}_{ik}^{[\kappa-1]}$, $\widehat{\boldsymbol{\mathcal{A}}}_{\ell}^{[\kappa-1]}$, $\widehat{\boldsymbol{\Gamma}_u}^{[\kappa-1]}$, $\widehat{\rho}_{ij}^{[\kappa-1]}$, and $\widehat{\boldsymbol{\Gamma}_e}^{[\kappa-1]}$, the factors are estimated by running the Kalman filter and the Kalman smoother. Then, given $\widehat{\boldsymbol{f}}_t^{[\kappa]}$, in the M-step the parameters are estimated equation-by-equation by running OLS, where the OLS formulas

[13]Setting up the variance of μ_{it} is important because separating a trend and a cycle is extremely challenging using the Frequentist approach we use and present in Section 3.2.1. In particular, due to the so-called pile-up problem (see Kim and Kim, 2022), if the variance of the random walk component is left unconstrained, the estimated trend can easily overfit the data. To this end, Stock and Watson (1998) propose the median unbiased estimator as a solution. A philosophically different approach is the Bayesian one, which imposes an informative and quite tight prior on the overall variations in the estimated trend (see Del Negro et al., 2019). Our approach is in line with the Bayesian method, as one can look at our strategy as setting up a dogmatic prior.

[14]In Appendix A.2, we show that including the secular trend even with a small variance is crucial. Moreover, in C.2 we show that the estimate of common core inflation is robust to reasonable values of $\sigma_{\eta_i}^2$.

are modified to account for the estimation error in $\widehat{\boldsymbol{f}}_t^{[\kappa]}$. For a rigorous treatment of the EM algorithm in Dynamic Factor Models we refer the reader to Barigozzi and Luciani (2020b).

We discuss a few high-level details about the estimation: the estimation over the two regimes, the special treatment reserved for the secular trends, the characterization of the idiosyncratic components, and the normality assumptions.

First, we estimate the model's parameters independently for the two regimes with the EM algorithm. Next, the factors and the secular trends are estimated with one final run of the Kalman Filter and Smoother. Under the assumption that the idiosyncratic components are stationary, the EM algorithm works in the same way, independently of whether the data are stationary, long-memory, or unit root processes.

Regarding the idiosyncratic component, we forgo estimating (7); rather, we impose $\rho_{ij} = 0$. Moreover, we impose that $\boldsymbol{\Gamma}_e$ is a diagonal matrix. By imposing such a simplified structure, we are estimating a misspecified model. Nonetheless, these two assumptions have minimal effects on the efficiency of the estimator and no effect on the consistency of the estimator (see the simulations in Barigozzi and Luciani, 2020b). As emphasized earlier, the estimated idiosyncratic components are just slightly cross-correlated thanks to the carefully constructed dataset. Therefore, it is unlikely that the assumption that $\boldsymbol{\Gamma}_e$ is a diagonal matrix has meaningful effects on the consistency of the estimator.

Last, we assume that the shocks are drawn from normal distributions. However, both the model parameters and the unobserved states are consistently estimated even if the normality assumption does not hold in the data or if we relax the normality assumptions.[15]

3.2.2 Modeling disaggregated inflation during the Covid-19 pandemic

As discussed in recent studies (for example, Maroz et al., 2021), the Covid-19 recession is unique in two aspects. First, in March and April 2020, the U.S. economy was hit by an unprecedentedly large and acute shock that had pervasive effects across sectors. Second, about a year after the shock, consumer price inflation rose, reaching levels not seen in the past four decades.

The Covid-19 shock is so large that it requires special treatment. Otherwise, the dynamic factor model estimate will likely be distorted for three reasons. First, the model

[15] We could relax this assumption by allowing the shocks to be distributed from any distribution of the exponential family as long as the first four moments are defined. Therefore, the fact that disaggregated inflation rates are skewed and fat-tailed is not a problem. For instance, simulations in Barigozzi and Luciani (2020b) show that the model is consistently estimated even when the shocks come from a Skew-t distribution.

parameters change substantially to fit the extreme movements between March and April 2020, as discussed in Lenza and Primiceri (2020). Second, the smoothed estimate of the factor for the pre-pandemic period changes noticeably because the smoother interprets some of the surprises after March 2020 as informative about the pre-pandemic. Third, inflation dynamics may have returned to the long-memory regime, in which case, characterizing disaggregated price inflation as a short-memory process is misleading. All told, allowing any of these things to happen would be ill-advised.

Having this in mind, we make three assumptions. First, we will initially be agnostic about the regime and assume that the post-pandemic could belong to either regime. Second, Covid-19 is such an extreme and short-lived shock that we will treat it as not related to the standard inflation co-movement. This assumption is in line with the results of Maroz et al. (2021), who concluded that Covid dominated co-movement in the data from March-June 2020 so these data should be ignored when studying non-Covid questions. Moreover, this assumption is in line with Ng (2021), who excludes the pandemic observations when estimating the dynamic factor model to prevent the mentioned estimation problem. Third, observations after the pandemic contain no information for pre-pandemic inference because the big economic shock and its consequences in 2020:H1 were unprecedented and unexpected; hence, the surprises cannot be interpreted in the context of average pre-pandemic dynamics.

Therefore, (1) we do not re-estimate the model but continue to employ the parameters estimated over the short-memory or long-memory regime and let the data tell us which regime better describes post-pandemic inflation dynamics. (2) We run the Kalman Smoother separately for the pre- and post-pandemic periods to prevent the pandemic observations from changing the pre-pandemic inference. This procedure creates discrepancies in the smoothed factor estimates in February 2020 obtained from the two separate smoothing procedures. By adjusting the level, we match the post-pandemic smoothed estimate to the pre-pandemic estimate while letting the idiosyncratic component absorb the differences.[16]

Using this approach, we estimate the co-movement across disaggregated PCE prices had they followed their pre-pandemic short/long-memory patterns. In addition, the acute Covid-19 shock does not distort the two-sided estimate of the pre-pandemic factor. In sum, our approach produces estimates of the common component during the Covid-19 pandemic and the recovery that satisfy the desirable property.

[16] An unintended consequence of this approach is that the idiosyncratic component essentially absorbs the unprecedented dynamics driven by the pandemic-specific shock. Appendix B.1 discusses how we address this problem.

Could we have treated the Covid period differently? In other words, is there an alternative to stop estimating the model before Covid? Yes, the literature has suggested two options. The first one consists of assigning ‘NaN’ to all the observations during the Covid period (Ng, 2021). The second one is to boost the volatility of the common shock in the Covid period so that the model knows that what is going on comes from additional volatility in the common shock, not from a change in the parameter (Lenza and Primiceri, 2020). However, the complication here is the potential change of regime. If inflation reverted to the long-memory regime, then we should concatenate observations pre-1995 with those post-2021. This is certainly feasible, but it looks a little bit odd. In addition, we choose not be agnostic about the regime.

3.3 Model set-up

Before estimating the model, we must determine the number of factors q and the number of lags s in the factor loadings. To estimate the number of factors, we use the information criterion proposed by Hallin and Liška (2007), which exploits the behavior of the eigenvalues of the spectral density matrix of $\boldsymbol{\pi}_t$ averaged across all frequencies, $\boldsymbol{\Sigma}_\pi(\omega)$.[17]

In the 1959–94 sample, the Hallin and Liška (2007) criterion suggests no common factors. However, the Barigozzi et al. (2021) criterion run on the same sample selects the presence of one common trend in $\boldsymbol{\pi}_t$ (see Appendix B.6).[18] Taken together, we read these results as suggesting the presence of one common shock in the 1959–94 sample—we suspect that the Hallin and Liška (2007) criterion detects no factors because $\Delta\boldsymbol{\pi}_t$ is too noisy and at high frequencies there is no commonality. As for the 1995–2019 sample, the Hallin and Liška (2007) criterion points towards one common shock.

Having determined q, we choose s such that the share of the variance explained by the $\tilde{q} = q(s+1)$ largest eigenvalues of the covariance matrix of $\boldsymbol{\pi}_t$ coincides with the share of the variance explained by the q largest eigenvalues of the spectral density matrix of $\boldsymbol{\pi}_t$ (averaged over all frequencies)—see also D’Agostino and Giannone (2012). The rationale for this approach is that if model (1), (4)–(7) is the true data-generating process, then the

[17]To account for the different time-series properties of the pre- and post- 1990s data, we run this criterion separately in the two periods. When dealing with pre-1990s data, we run the criterion on $\Delta\boldsymbol{\pi}_t$—in the first sample $\boldsymbol{\pi}_t \sim I(1)$, and the Hallin and Liška (2007) criterion works on stationary data because it relies on the spectral density matrix of the data—while when dealing with data post-1995, we run the criterion on $\boldsymbol{\pi}_t$.

[18]The Barigozzi et al. (2021) criterion is a modification of the Hallin and Liška (2007) criterion that looks only at the eigenvalues of $\boldsymbol{\Sigma}_{\Delta\pi}(\omega)$ at frequency zero. This criterion detects the number of common trends in $\boldsymbol{\pi}_t$.

spectral density matrix of $\boldsymbol{\pi}_t$ has q eigenvalues that diverge with n, and, at the same time, the covariance matrix of $\boldsymbol{\pi}_t$ has at most $\tilde{q} = q(s+1)$ diverging eigenvalues.

Table 1 reports the cumulative variance explained by each of the 10 largest eigenvalues of the spectral density matrix (lines 1 and 3) and the covariance matrix (lines 2 and 4) in the two samples. Our approach consists of finding the value of $\tilde{q}$ in line (2) such that the variance explained by the $\tilde{q}$ largest eigenvalues of the covariance matrix is the closest to–but larger than–the variance explained in line (1) by the largest eigenvalue of the spectral density matrix (that is, $q = 1$). The same applies to lines (3) and (4).

Table 1: PERCENTAGE OF EXPLAINED VARIANCE

	Sample		1	2	3	4	5	6	7	8	9	10
(1)	1959-1994	q	7.3	13.1	18.3	23.0	27.4	31.4	35.2	38.7	42.0	45.1
(2)		$\tilde{q}$	3.8	7.0	9.8	12.2	14.6	16.8	19.0	21.1	23.2	25.1
(3)	1995-2019	q	8.6	15.0	20.6	25.7	30.2	34.3	38.1	41.6	44.9	48.0
(4)		$\tilde{q}$	3.9	7.3	9.8	12.2	14.3	16.4	18.4	20.3	22.1	23.9

NOTES: This table reports the cumulative percentage of total variance explained by the q largest eigenvalues of the spectral density matrix and the $\tilde{q} = q(s+1)$ largest eigenvalues of the covariance matrix. The spectral density matrix and the covariance matrix are estimated on $\Delta\boldsymbol{\pi}_t$ in the 1959-94 sample and on $\boldsymbol{\pi}_t$ in the 1995-2019 sample.

The results in Table 1 suggest that $\tilde{q} \simeq 3$ in the 1959–94 sample and $\tilde{q} \simeq 2$ in the 1995–2019 sample—the Bai and Ng (2002) criterion (see Appendix B.6) confirms these results. Thus, we select $\tilde{q} = 3$, that is, $s = 2$.

In summary, our benchmark specification features one common factor loaded by each price inflation index within three months, $q = 1$ and $s = 2$. As for the lag order of the AR model for the common factor, we set $p = 3$ based on the BIC.[19]

One final comment about the benchmark specification is in order. The model with $q = 1$ imposes that, taking aside idiosyncratic shocks and shocks to the secular trend, any difference in relative inflation will reflect differences in the factor loadings. At first glance, this specification might seem to impose very strong economic restrictions, as it postulates that one specific shock drives the co-movement in PCE prices. However, we are not identifying shocks and not labeling the common shock. As such, the model with $q = 1$ imposes that every month, one shock is the primary driver of the co-movement in the data, not a specific shock. It might be different shocks, but every month just one of them matters. If interpreted this way, it is clear that the model is not imposing any extreme economic restrictions.

[19]In Appendix C.3, we present a robustness analysis for alternative model specifications.

4 Empirical Results

This section presents the estimation results. Section 4.1 presents he aggregate results, while in Section 4.2, we dive deep into the disaggregated prices.

4.1 Commonality in aggregate inflation

This section presents the aggregate results. Section 4.1.1 Characterizing the two inflation regimes. Section 4.1.2 reports the estimates of common core inflation before the Covid-19 pandemic. Section 4.1.3 presents estimates during the pandemic.

4.1.1 Characterizing the two inflation regimes

In this section, we discuss the difference in the parameter estimates between the pre-1995 and the post-1995 regimes and their implications.

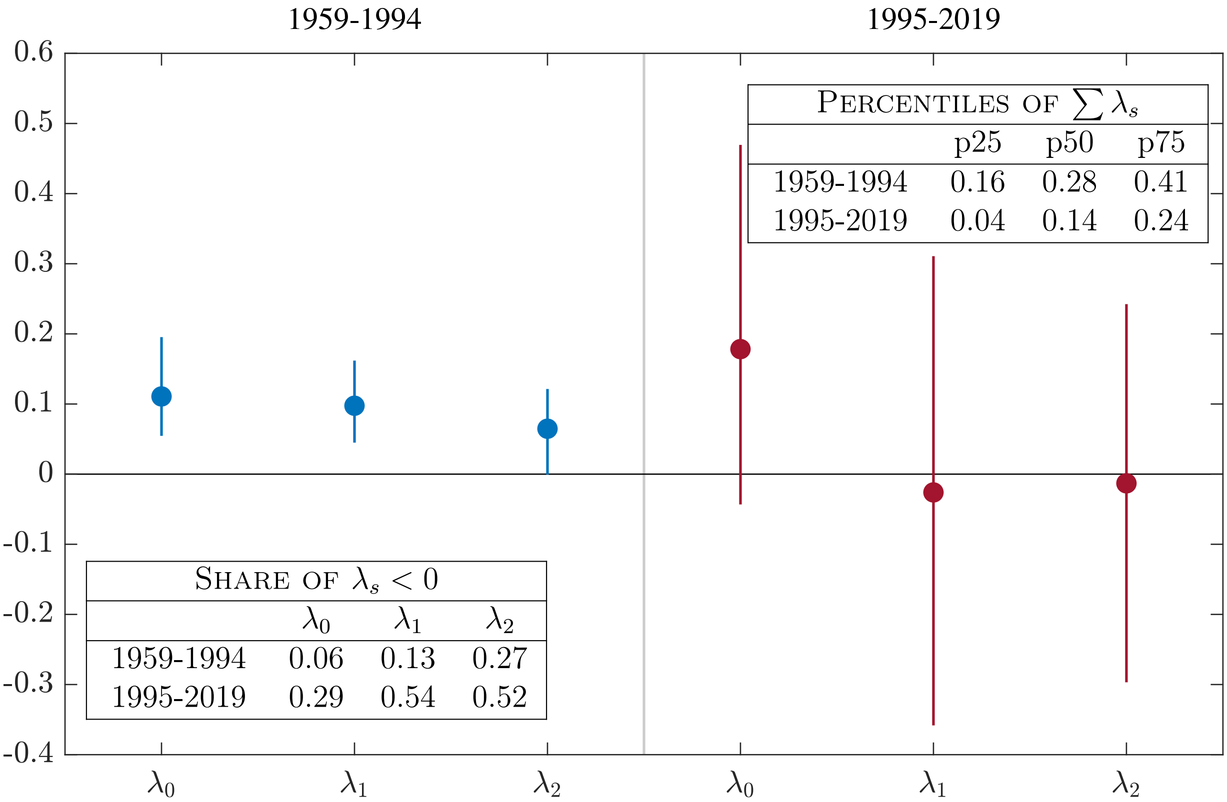
We begin by comparing the factor loadings. Figure 5 shows the median and the interquartile range of the cross-sectional distribution of the factor loadings for each regime. The interquartile range of the estimated factor loadings in the pre-1995 sample is very narrow, meaning that most disaggregated prices load the common factor similarly. Moreover, most of the contemporaneous (λ_0) and lagged loadings (λ_1 and λ_2) have a positive sign (inset box in Figure 5). In contrast, the interquartile range of the estimated factor loadings in the post-1995 sample is very large, indicating large variability around how disaggregated prices load the common factor. Moreover, the sum of the factor loadings, $\sum_{s=0}^{2} \lambda_s$, is much smaller in the post-1995 sample than in the pre-1995 sample, confirming the idea that there is much less commonality in the post-1995 sample.

Next, we move to the autoregressive parameters governing the low of motion of the common factors. We estimate $\mathcal{A}^{(1)}(L) = 0.15\mathcal{A}_1 L + 0.36\mathcal{A}_2 L^2 + 0.33\mathcal{A}_3 L^3$ in the pre-1995 sample, and $\mathcal{A}^{(2)}(L) = 0.90\mathcal{A}_1 L + 0.21\mathcal{A}_2 L^2 - 0.17\mathcal{A}_3 L^3$ in the post-1995 sample. Both $\mathcal{A}^{(1)}(L)$ and $\mathcal{A}^{(1)}(L)$ yield two short-memory processes. Surprisingly, though, $\mathcal{A}^{(2)}(L)$ generates a more persistent process than $\mathcal{A}^{(1)}(L)$, albeit the highest root of the two polynomials is almost identical (0.92 vs. 0.93).

The low of motion of common core inflation is a function of both the factor loadings and the low of motion of the common factor. If we write (6) in its MA form—$\boldsymbol{f}_t = \boldsymbol{\mathcal{A}}(L)^{-1}\mathbf{u}_t$—and substitute it into (5), we obtain

$$\chi_{it} = \boldsymbol{\lambda}_i(L)\boldsymbol{\mathcal{A}}(L)^{-1}\mathbf{u}_t \qquad \Longrightarrow \qquad \boldsymbol{\mathcal{A}}(L)\boldsymbol{\lambda}_i(L)^{-1}\chi_{it} = \mathbf{u}_t$$

Figure 5: Cross-sectional median and inter-quantile range of the estimated factor loadings

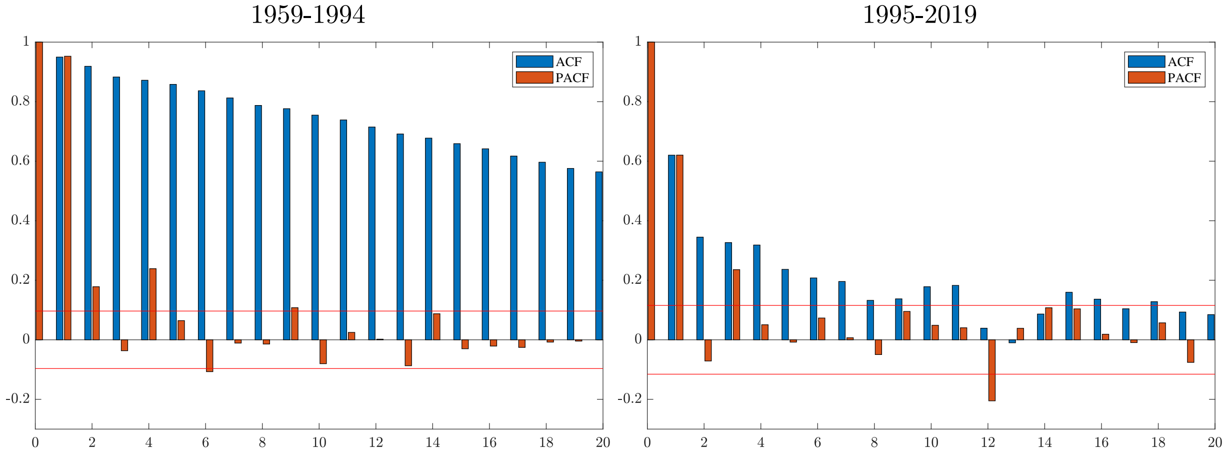


Note: The vertical bars denote the cross-sectional interquartile range, while the dots denote the cross-sectional median of the factor loading. The inset box in the chart reports the share of negative loadings.

from which it is clear that the order of integration of χ_{it} depends on the roots of $\boldsymbol{\mathcal{A}}(L)\boldsymbol{\lambda}_i(L)^{-1}$, not on those of $\mathcal{A}(L)$. Now, although $\widehat{\mathcal{A}}^{(2)}(L)$ generates a more persistent process than $\widehat{\mathcal{A}}^{(1)}(L)$, the estimates of the loadings reported in Figure 5 are such that common core inflation in the pre-1995 sample is much more persistent than in the post-1995 sample. Specifically, in the pre-1995 sample, common core inflation exhibits a slowly decreasing autocorrelation function typical of a long-memory (and potentially nonstationary) process (Figure 6, left chart). In the post-1995 sample, common core inflation exhibits a monotonically decreasing autocorrelation function typical of a short-memory process.

Looking at the autocorrelation of $\widehat{\chi}_t^c$ in the pre-1995 regime shown in the left chart of Figure 6, it is clear that $\widehat{\chi}_t^c$ is not a unit root process, but rather a long memory process, as suggested by the slow decaying of the autocorrelation function. To corroborate this conjecture, we estimate the degree of fractional integration, d, of $\widehat{\chi}_t^c$ in the persistent regime. By estimating an ARFIMA(1; d; 0) model with maximum likelihood (Beran, 1995), we get $\widehat{d} = 0.69$, which suggests that $\widehat{\chi}_t^c$ is a nonstationary fractionally integrated

Figure 6: Autocorrelation and Partial autocorrelation of $\widehat{\chi}_t^c$



process.[20] Note that a nonstationary fractional integrated process has infinite variance but mean reverts, while an $I(1)$ process has an infinite variance and does not mean revert. In other words, a shock to an $I(d = 0.69)$ process will generate transitory but very persistent dynamics, while a shock to an $I(d = 1)$ process will generate a permanent change to the level of the series.

Given these results, we dub the pre-1995 "the long-memory regime" and the post-1995 regime "the short-memory regime."

4.1.2 Inflation dynamics before the Covid pandemic

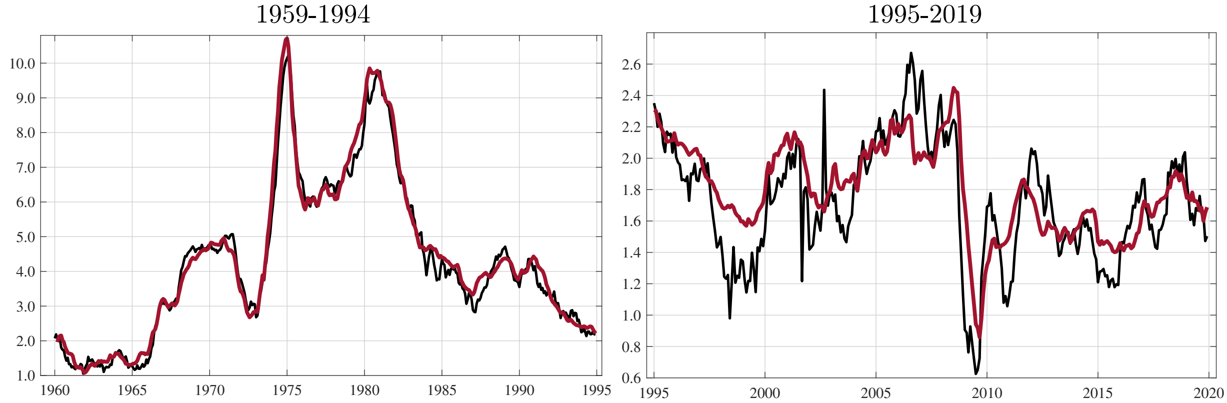
Figure 7 shows the common and idiosyncratic decomposition of year-over-year (YoY) core PCE price inflation.[21] As shown in the left chart, before the 1990s, common core inflation accounted for most of the fluctuation in core inflation, particularly low-frequency fluctuations. However, as shown in the right chart, starting from the mid-1990s, common core inflation accounts for a much smaller share of core PCE price movements. Specifically, our model classifies the 2010 downturn in core inflation as entirely idiosyncratic, and it also suggests that the 2015 and 2017 downturns in core inflation were due to idiosyncratic

[20] ARFIMA stands for Autoregressive Fractionally Integrated Moving Average. We say that $y_t \sim$ ARFIMA(p, d, q) if $x_t \sim$ ARMA(p, q), where $x_t = (1 - L)^d y_t$, d can be any real number, and $(1 - L)^d = 1 - dL + \frac{d}{2}(d-1)L^2 - \frac{d}{6}(d-1)(d-2)L^3 + \ldots$. An ARMA process is an ARFIMA process with $d = 0$, and an ARIMA process is an ARFIMA process with $d = 1$.

[21] Appendix B.2 shows how the estimate of common core inflation change when we use a smaller dataset, while Appendix B.3 shows the common and idiosyncratic decomposition of the 12-month percent change in the headline PCE price index.

dynamics. Hence, these results suggest that most of the swings in core inflation during the expansion that followed the Great Recession were mostly idiosyncratic. In other words, core inflation was restrained for about 10 years by a series of idiosyncratic shocks; hence, core inflation below 2% was not a structural feature.

Figure 7: Common and idiosyncratic decomposition
Core PCE prices — year-over-year inflation



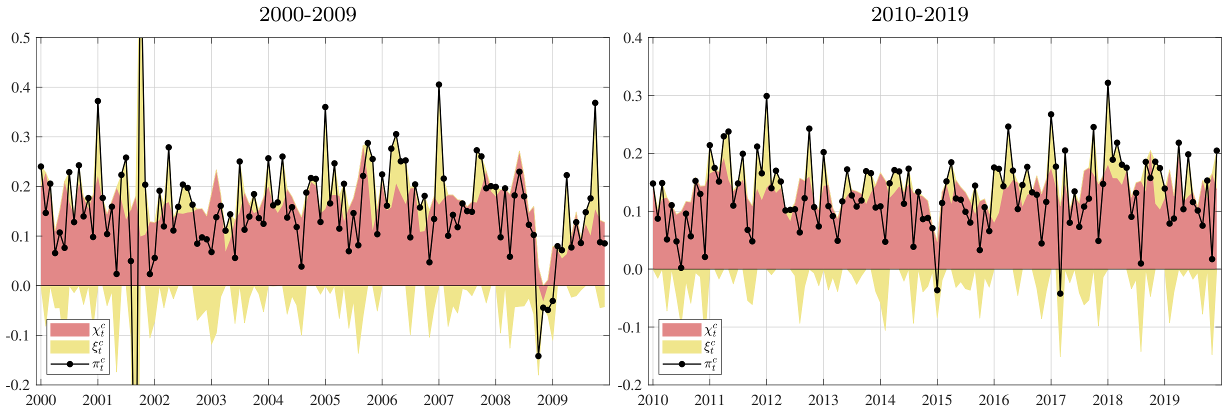
Notes: In each plot, the red line denotes year-over-year common core inflation—that is, the common component's contribution to the overall 12-month percent change of the core PCE price index (which is given by the black line). Put differently, the red line tells us what core inflation would have been had there been no idiosyncratic price shocks over the past 12 months. The black line is core PCE price inflation. The plot on the left covers the long-memory period (1959 to 1994), while the right panel covers the short-memory period (1995 to 2019). To make it easier to understand what is happening in each of the two periods, the two charts have different y-axes.

Our measure of common-core inflation speaks to the discussion on missing disinflation during the Great Recession and missing inflation during the post-Great Recession recovery (see, for example, Coibion and Gorodnichenko, 2015, and Constancio, 2015). What puzzled macroeconomists was that headline inflation measures, including core inflation, showed neither a meaningful decline during the Great Recession nor a pick up in the recovery phase in a way consistent with the large swing in economic activity. This observation suggested a clear breakdown in the Phillips correlation.

Quite differently, our common core inflation exhibits variations better aligned with the underlying economic strength during the Great Recession and the recovery than the headline data. As shown in Figure 7, YoY common core inflation dropped by 1.6 percentage points during the recession, about twice as much as the range within which it fluctuated between 1995 and 2008. Although the decrease in core inflation during the recession matched that of common core inflation, magnitude of this decline was similar to the range within which core inflation fluctuated between 1995 and 2008—in other words, it represents a much smaller deviation compared to its historical fluctuation.

In summary, the noisy idiosyncratic component masked the portion of core inflation better aligned with real economic activity, and the missing disinflation and inflation are partly attributable to the transitory component, which is less relevant to the aggregate dynamics.

Figure 8: Common and idiosyncratic decomposition
Core PCE prices — month-over-month inflation — 2000–2019



Notes: In each plot, the red area is common core inflation, while the yellow area gives the idiosyncratic component. These two components sum to overall core PCE inflation (the black line) by construction.

Figure 8 shows the common and idiosyncratic decomposition of monthly core PCE price inflation since the year 2000. As can be seen, our model effectively parses out transitory and idiosyncratic surprises in core inflation. To illustrate this and to demonstrate the relevance and usefulness of the common and idiosyncratic decomposition, we focus on episodes between the Great Recession and the Covid pandemic (right chart in Figure 8).[22] For example:

- In 2010, several clearly identifiable idiosyncratic negative shocks—the collapse of the index for luggage in January, the very low reading for Medicare hospital services prices in October, and an exceptionally long series of negative readings in the index for apparel—lowered core inflation.
- In both 2014 and 2015, two years in which medical prices were low partly due to the implementation of the Affordable Care Act, idiosyncratic factors held down core inflation.
- In March 2017, when core inflation was heavily affected by the collapse in the price index for wireless telephone services, the model correctly interpreted that these developments were idiosyncratic and transitory. This call from the model is particularly relevant because, at the time, there were concerns about inflation being constantly below 2 percent

[22] Appendix B.4 further demonstrates the usefulness of common core inflation in practice by focusing on the real-time reliability of model estimates (the properties of common core inflation as an inflation gauge are discussed in Appendix B.5).

and the possibility of inflation expectations de-anchoring on the downside.

Finally, another useful feature of the model is that it successfully detects residual seasonality in the data and attributes these movements to the idiosyncratic component. Due to residual seasonality, core inflation tends to be higher in the first half of the year.[23] The model parses this regular pattern as idiosyncratic; hence, the idiosyncratic component is positive in the first half of the year (in January in particular) and negative in the second half.

4.1.3 The post-pandemic inflation ramp-up

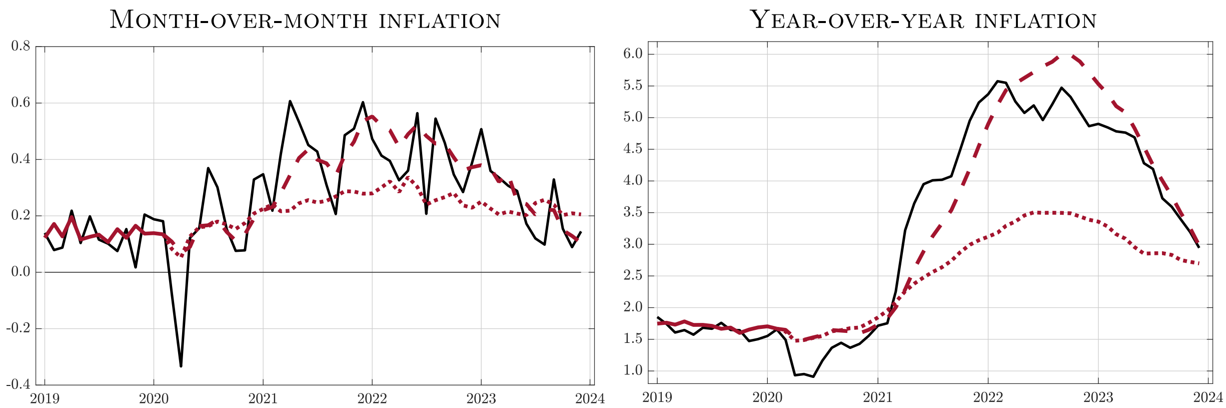
The previous section demonstrates that the model successfully parses out prominent episodes of transitory and non-pervasive price changes as idiosyncratic. This section focuses on the post-pandemic inflation ramp-up and evaluates how the model interprets this period. Correctly estimating common core inflation during the 2021-2022 inflation ramp-up and the following decrease is important given the heightened uncertainty on the inflation trajectory in the coming years. Appendix B.1 focuses on inflation dynamics during the Covid lockdowns and the reopening.

We first revisit the question we left open in Section 2.2. Has the Covid-19 shock brought back the long-memory regime? Does the 2021 ramp-up in inflation reflect this nonlinearity? To answer these questions, we estimate common core inflation after the pandemic using both sets of parameters (i.e., from the short- and the long-memory regime). Figure 9 reports estimates of month-over-month and YoY common core inflation post-pandemic in the short-memory regime (red dotted line) and in the long-memory regime (dashed line)—in this second case, the model switches regimes twice, in January 1995 from long-memory to short-memory and in March 2020 from short-memory to long-memory. The results in Figure 9 are clear: the long-memory regime model captures the 2021 inflation ramp-up and the 2023 decline much better than the short-memory regime model.

The results in Figure 9 suggest that inflation dynamics after the pandemic (and up to the end of 2023) are closer to what happened in the 1970s than in the 2000s. Would we have been able to reach this conclusion had we used the model in real time? And, when would we have been able to reach it? To answer this question, we re-estimate

[23]Peneva (2014) shows that, despite being based on data that statistical agencies seasonally adjust, core PCE price inflation exhibits a regular downward pattern from the first to the second half of the year. Moreover, Peneva and Sadée (2019) show that, although the 2018 comprehensive NIPA data revision has partially attenuated the problem, residual seasonality is still present in core PCE price inflation.

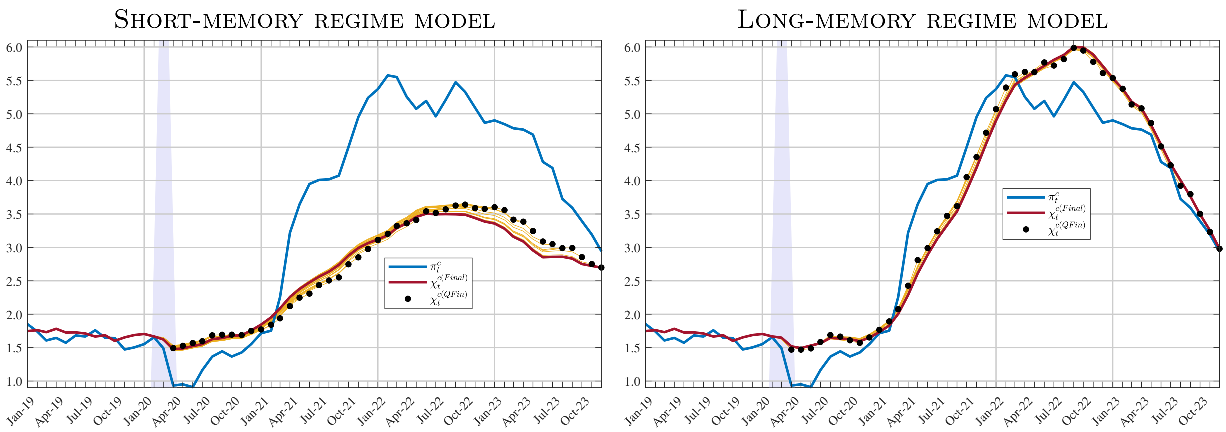
Figure 9: Common core inflation during Covid



Notes: The black line is core PCE price inflation. The solid red line is the estimate of common core inflation as discussed in Section 4.1.2. The dotted red line is common core inflation estimated by assuming that inflation remained short-memory during the Covid pandemic. The dashed line is common core inflation estimated by assuming that inflation switched back to the long-memory regime during the Covid pandemic.

common core inflation, but instead of using the whole sample, we add one observation post-February 2020 at a time. Figure 10 shows the results of this exercise.[24] The main

Figure 10: Real-time estimates of Common core inflation during Covid



Notes: In all plots, the red line is the estimate of common core inflation obtained using data up to December 2023. The yellow lines are the estimate obtained by adding one observation at a time starting with April 2020. Each black dots represent the estimate of common core inflation for month M and year Y obtained over the sample ending at month M and year Y—according to the definition of Section **??**, the black dots are the *quasi*-final estimate of core inflation. Put differently, in the figure, there is a yellow line for each black dot, and each yellow line ends with a corresponding black dot. The black line in the left plot is headline PCE price inflation.

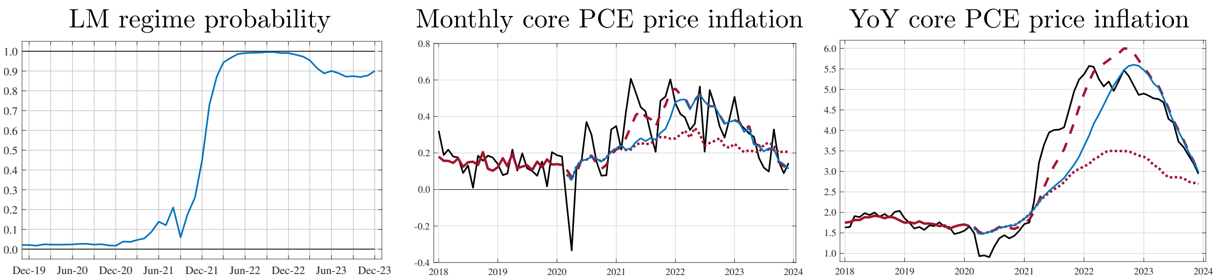
result emerging from Figure 10 is that the estimate of common core inflation revised very

[24]Section B.4 provides a thorough real-time evaluation showing how much data revisions, parameters estimation, and filtering, contribute to model revisions and calculating their relative importance. In this Section, we do not have to estimate the parameters so we ignore the data revision problem to simplify. This simplification is justified by the results in Section B.4 that shows that data revisions do not have a big impact on the revision of common core inflation.

little. Thus, in real-time, the long-memory regime model would have captured the ramp-up in inflation already in March 2021, while the short-memory regime model would have heavily discounted the high readings in core inflation. By July 2021, the difference between the estimate of YoY common core inflation from the long-memory regime model and the short-memory regime model was higher than half of a percentage point, signaling that it was likely that the ramp-up in inflation would have been not so transitory as at the time several people thought.

The left plot in Figure 11 zooms into our preferred probability estimate to detect potential abrupt regime changes. As can be seen, as of August 2021, the estimated probability of being in the long-memory regime was 20%, while, as of December 2021, it was 40%. This result, combined with those in Figure 9, shows that our model provided evidence of a regime switch by August 2021, just a few months after the inflation ramp-up started, evidence that became very robust by December 2021. Moreover, as of late 2023 (nearly two years after the start of the inflation ramp-up), we estimate that the probability of inflation dynamics still being in the long-memory regime is 90%.

Figure 11: Weighted common core inflation during Covid



Notes: The left chart shows the probability of inflation being in the long-memory regime. This is the same probability shown in Figure 4. In the middle and right plot, the black line is core PCE price inflation. The solid red line is the estimate of common core inflation as discussed in Section 4.1.2. The dotted red line is common core inflation estimated by assuming that inflation remained short-memory during the Covid pandemic. The dashed line is common core inflation estimated by assuming that inflation switched back to the long-memory regime during the Covid pandemic. The blue line is the weighted average of the dashed and dotted lines, where the weights are the estimated probability shown in the left chart.

As an alternative to the approach of determining the regime and then applying a binary approach, a practitioner can exploit the probability of being in the long-memory regime to build a weighted average of common core inflation from the two regimes (i.e., the dashed and dotted lines in Figure 9). The blue line in the middle and right chart in Figure 11 illustrates this procedure. As show in the middle chart, we would have made the right call with this approach. In 2021, we would have been able to call core PCE price changes increasingly common, and by early 2022, we would have been able to call core PCE price changes mainly common.

4.2 Commonality in disaggregated inflation

This section discusses how much disaggregated inflation indexes co-moves and how much of their fluctuations are idiosyncratic.

We begin our analysis of disaggregated PCE prices by examining how our model parses disaggregated inflation indexes into common and idiosyncratic dynamics. To this end, we look at the percentage of variations in the inflation rate for total, food, energy, and core (goods and services) explained by common dynamics (across all frequencies and at different frequencies). To do so, we first decompose each inflation index (and its common component) into six different series, each isolating fluctuations with different frequencies—call them $\pi_t^j(\omega)$ and $\chi_t^j(\omega)$, where $j = \{$*Total, Core, Core Goods, Core Services, Food, Energy*$\}$ and ω is the frequency band. To compute $\pi_t^j(\omega)$ and $\chi_t^j(\omega)$, we use cosine projections as in Müller and Watson (2017). Next, we compute a *pseudo* R^2 for each frequency band, call it $\tilde{R}_j^2(\omega)$, as

$$\tilde{R}_j^2(\omega) = 1 - \frac{\sum_t(\pi_t^j(\omega) - \chi_t^j(\omega))^2}{\sum_t(\pi_t^j(\omega))^2}. \tag{8}$$

The higher $\tilde{R}_j^2(\omega)$, the larger the share of fluctuations of index j at frequency ω explained by the common component. Henceforth, we refer to $\tilde{R}_j^2(\omega)$ as the commonality share.

Each column in the Table 2 reports the commonality share computed over a certain frequency band, where we report the period τ of each band expressed in months. Thus, the second column, $\tau \geq 60$, reports the share of variance explained by common components of fluctuations longer than 60 months, while the last column, $\tau < 6$, reports the share of variance explained by common components of fluctuations shorter than 6 months. The first column $(0 < \tau < \infty)$ reports the share overall frequencies—that is, the overall share.

The first column of Table 2 shows that the commonality in PCE prices has decreased since the mid-1990s. In particular, before the mid-1990s, core goods and core services prices were moving in sync and so the commonality in the data reflected the dynamics of both core goods and services prices (in addition to those of food prices). In contrast, since the mid-1990s, core goods prices became almost entirely idiosyncratic, and this is true for all subcomponents but to a lesser extent to motor vehicles prices (see Table 4). Our intuition is that the increased idiosyncrasy in goods prices is probably due to the increase in the share of goods produced abroad. Thus, the commonality in the data reflects the dynamics of core services prices, and food and energy prices. Specifically, common core inflation since the mid-1990s reflects primarily the dynamics of energy goods (essentially,

Table 2: Commonality shares of aggregated PCE price inflation (by frequency bands)

		$0 < \tau < \infty$	$\tau \geq 60$	$12 \leq \tau < 60$	$6 \leq \tau < 12$	$\tau < 6$
1959-1994	Total	81.1	98.9	63.2	50.7	35.7
	Core	79.7	99.1	70.3	16.5	9.1
	Goods	59.6	98.8	57.1	2.3	4.9
	Services	69.1	98.5	14.0	12.7	3.5
	Food	45.3	90.1	18.4	28.2	39.9
	Energy	25.3	91.3	15.4	12.7	-0.2
	Energy goods	18.0	84.6	13.1	10.9	0.9
	Energy svs	46.3	96.1	31.6	-0.8	2.4
1959-1994	Total	76.4	91.4	88.4	91.5	59.1
	Core	19.2	76.7	39.8	22.3	2.3
	Goods	5.8	65.7	-0.9	6.1	0.2
	Services	15.9	86.6	41.9	15.2	-2.1
	Food	26.5	90.8	59.8	-7.2	-3.1
	Energy	76.6	95.7	80.4	91.7	66.9
	Energy goods	72.5	94.1	74.7	88.6	63.2
	Energy svs	27.5	89.7	42.2	5.1	19.0

Notes: This table shows the commonality share over different frequency bands $\omega = \frac{2T}{\tau}$, where τ is the period expressed in months. The first column ($0 < \tau < \infty$) reports the share overall frequencies—that is, the overall share—the second column ($\tau \geq 60$) reports the share over fluctuations longer than 60 months, and so on. The commonality share is computed as $100 \times \tilde{R}^2$, where $\tilde{R}^2$ is defined in (8). Note that, because $\tilde{R}^2_j(\omega)$ is a *pseudo* R^2, it can be negative, in which case it signals that $\chi^j_t(\omega)$ does not explain fluctuations in $\pi^j_t(\omega)$.

To perform this exercise, we removed from the monthly inflation rate in core PCE prices the September and October 2001 observations and replaced them with the average over the previous 12 months. In September 2001, core PCE price inflation was -0.56% (-6.5 % at an annual rate), while in October 2001, it was +0.72% (+8.9 % at an annual rate). The 2001 swing in core PCE price inflation was driven by the price index for life insurance, which plunged 55 % in September 2001 and jumped 121 % in October 2001 as a result of the 9/11 terrorist attacks.

gasoline) prices and housing services prices (see Table 4, line 9).[25]

Table 3, which decomposes the variance of different aggregated inflation indexes into the contribution coming from different frequency bands,

$$\check{R}^2 = \frac{\sum_t (\pi^j_t(\omega))^2}{\sum_t (\pi^j_t)^2}, \tag{9}$$

provides additional intuition on why commonality in PCE prices has decreased since the mid-1990s. Before the mid-1990s, about 65-70% of the variability in total and core infla-

[25]One reason why housing has become more common, is that in 1985 the Bureau of Labor Statistics (BLS) changed the methodology used to compute the price of homeownership, moving from a "user cost approach" in which the services for an owned dwelling are computed by summing the cost related to home ownership, to "rental equivalence" in which the services for an owned dwelling is the rate of change in the amount an owner would need to pay to rent that dwell. As a result of this change, the time series of inflation for homeownership, nowadays called Owner Equivalent Rent (OER), has become way less volatile.

tion was driven by low-frequency fluctuations (first column), with a little over 15% of the variance accounted for by high-frequency fluctuations of less than six months (fourth column). However, since the mid-1990s, low-frequency fluctuations account for less than 10% (15%) of total (core) inflation fluctuations, whereas high-frequency fluctuations account for 45% (60%) of the variance. In sum, high-frequency variations, which are accounted for mainly by the idiosyncratic component (see Table 2), become significantly more important from the mid-1990s than in the earlier period, suggesting the increased role of idiosyncratic fluctuations in disaggregated inflation dynamics.

Table 3: VARIANCE DECOMPOSITION OF AGGREGATED PCE PRICE INFLATION BY FREQUENCY BANDS

	Frequency $\omega = \frac{2T}{\tau}$	$\tau \geq 60$	$12 \leq \tau < 60$	$6 \leq \tau < 12$	$\tau < 6$
1959-1994	Total	66.0	8.6	8.6	16.8
	Core	71.9	8.9	5.6	13.6
	Goods	47.1	20.5	6.3	26.1
	Services	67.4	6.2	9.4	17.0
	Food	22.1	13.8	23.1	41.0
	Energy	20.2	23.5	25.5	30.8
1959-1994	Total	7.6	25.1	23.1	44.2
	Core	11.7	13.1	16.2	58.9
	Goods	7.2	11.7	17.1	64.0
	Services	13.2	10.9	8.5	67.4
	Food	11.0	31.3	9.5	48.2
	Energy	3.9	21.0	23.1	52.0

NOTES: This table shows the share of variance of each inflation index explained by each frequency band $\omega = \frac{2T}{\tau}$ as defined in (9). τ is the period expressed in months. The first column ($0 < \tau < \infty$) reports the share over all frequencies—that is, the overall share—the second column ($\tau \geq 60$) reports the share over fluctuations longer than 60 months, and so on. The share is computed as $100 \times \breve{R}^2$, where $\breve{R}^2$ is defined in (9).

To perform this exercise, we removed from the monthly inflation rate in core PCE prices the September and October 2001 observations and replaced them with the average over the previous 12 months (see also the note on Table 2).

Putting together the results in Tables 2 and 3 with those in Section 4.1.1 we can further refine the characterization of the the long-memory and short-memory regime. In the long-memory regime, inflation dynamics are primarily driven by common shocks, and the effect of these common shocks is long lasting; in the short-memory regime, inflation dynamics are primarily driven by idiosyncratic shocks, and the effect of the common shocks is short lived. Given this characterization, and since monetary policy should respond only to common inflation shocks, we conclude that when inflation is in the long-memory regime, monetary policy should promptly and decisively respond to inflation dynamics, while when inflation is in the short-memory regime, monetary policy should focus on stabilizing the real economy.

Moving back to Table 2, the second column shows that the common component accounts for most of the low-frequency variations in both samples. Note that the dynamic factor model performs cross-sectional smoothing and no temporal smoothing; hence, this result is not the necessary outcome of any restrictions in the model. In contrast, the common component accounts for a small share of core inflation's mid- to high-frequency fluctuations (second to fourth columns). This portion becomes more important after the mid-1990s. As for total inflation, the common component captures mid- and high-frequency fluctuations well (second to fourth columns) because it captures energy price fluctuations. Finally, common dynamics explain about 30% (60%) of the fluctuations in energy price inflation before (after) the mid-1990s. Considering oil prices heavily influence energy prices, this result suggests that the model views oil price shocks as a common macroeconomic shock (see Conflitti and Luciani, 2019, for a detailed discussion on oil prices' effect on common and idiosyncratic core inflation).

Moving to the post-Covid inflation ramp-up, Table 4 reports the commonality shares of disaggregated inflation sub-components before 2019 and in 2021-2023. The commonality in the data in the 2021-2023 period reflects primarily the dynamics of core goods inflation (nearly all components), housing and food services prices (lines 9 and 13), and food prices (line 23). This result explains why the long-memory regime fits the data starting in 2021 and, especially, why it can do so in real time. First, one of the main drivers of core inflation in 2021 was the rise in goods prices due to pandemic-related supply chain issues. In the long-memory regime, a large share of the fluctuations in core goods is common; in contrast, core goods inflation is idiosyncratic in the short-memory regime. Therefore, when goods prices began to increase, common core inflation estimated with the long-memory regime model responded immediately, while the estimate from the short-memory model was unresponsive and started to respond later on, only once services inflation started picking up. Second, commonality in the long-memory regime also reflects housing and food services fluctuations, two large categories of core PCE price inflation that increased substantially during the inflation rump-up.

The aforementioned changes in the disaggregated dynamics are also evident in the distribution of commonality shares of each disaggregated inflation. Figure presents a histogram of the commonality share for 142 prices, which is unprecedented in the literature—previous studies like Stock and Watson (2016) and Eo et al. (2023) cannot capture such distributional changes because they lack sufficient cross-sectional information. When comparing the dis-

Table 4: COMMONALITY SHARES OF DISAGGREGATED PCE PRICE INFLATION

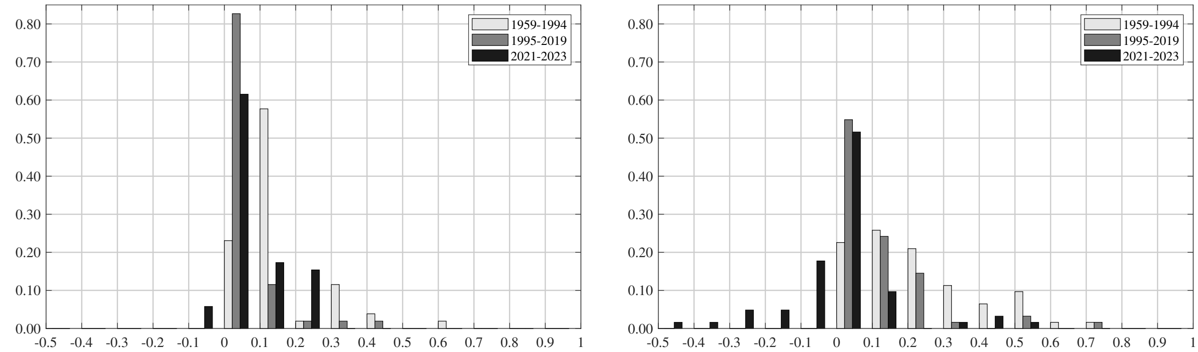
		1959-1994	1995-2019	2021-2023^L	2021-2023^S
(1)	Core goods	59.6	5.8	35.6	8.3
(2)	Motor vehicles	20.9	10.8	15.9	7.1
(3)	Furnishings and equipment	39.0	8.5	32.7	18.9
(4)	Recreational	29.6	8.8	13.3	0.5
(5)	Other durables	20.2	4.9	-0.5	-1.8
(6)	Clothing and footwear	22.8	5.2	19.1	3.5
(7)	Other nondurables	59.8	4.9	23.7	10.9
(8)	Core services	69.1	15.9	-8.8	11.1
(9)	Housing	50.3	72.5	42.2	54.7
(10)	Health care	55.4	4.9	-41.2	-7.8
(11)	Transportation	25.4	52.0	5.7	5.4
(12)	Recreation	32.9	27.5	1.1	-0.5
(13)	Food services	56.4	16.6	43.5	31.1
(14)	Accommodations	18.5	5.9	4.6	3.4
(15)	Financial and insurance	4.6	17.0	-2.1	2.9
(16)	Communication	8.2	22.9	-3.4	-9.7
(17)	Education	22.1	6.5	-94.1	-23.6
(18)	Professional and other	31.9	6.4	5.0	2.7
(19)	Personal care and clothing	55.2	10.4	11.3	7.0
(20)	Social and religious	38.8	35.9	-42.1	1.8
(21)	Household maintenance	28.8	7.8	3.5	1.5
(22)	NPISH	29.4	10.9	6.4	0.5
(23)	Food	45.3	26.5	40.2	53.2
(24)	Energy	25.4	76.5	11.1	67.1

Notes: This table shows the commonality share as defined in (8). The column "2021-2023^L" shows the share of variance explained in 2021–2023 when common core inflation is estimated using the long-memory model, while column "2021-2023^S' reports number obtained with the short-memory model.

NPISH stands for "Non-Profit Institutions Serving Household."

tribution of commonality shares over different periods, we see that as inflation dynamics transitioned from the long-memory regime (prior to the mid-1990s) to the short-memory regime (mid-1990s to 2019), the distribution of the commonality share shifted toward zero, indicating a general increase in idiosyncratic behavior. Moreover, we see that after the onset of the pandemic, although commonality in goods prices rose across the board, idiosyncrasy in some core services, namely education, health care, and some non-market services, increased heavily. As a result, the distribution of commonality shares became more dispersed.

Figure 12: Distribution of commonality shares



Notes: This figure displays the distribution of the commonality share in the 114 core (goods and services) disaggregated PCE prices in our dataset for three different sample periods: 1959-1994, 1995-2019, and 2021-2023, based on the long-regime estimates.

5 Common core inflation vs trend inflation

Our paper contributes to the literature using large-dimensional dynamic factor models to study disaggregated prices (for example, Cristadoro et al., 2005; Boivin et al., 2009; Conflitti and Luciani, 2019). While in Appendix D we provide a detailed literature review, in this section, we compare our model with that of Stock and Watson (2016), which is considered the gold standard in inflation modeling and forecasting.

Stock and Watson (2016), henceforth SW, estimate common trend inflation using a multivariate unobserved components/stochastic volatility and outlier-adjustment model (MUCSVO). The model includes common and idiosyncratic trends, stochastic volatility, time-varying factor loadings, and outlier treatment. The model is estimated with Bayesian methods.

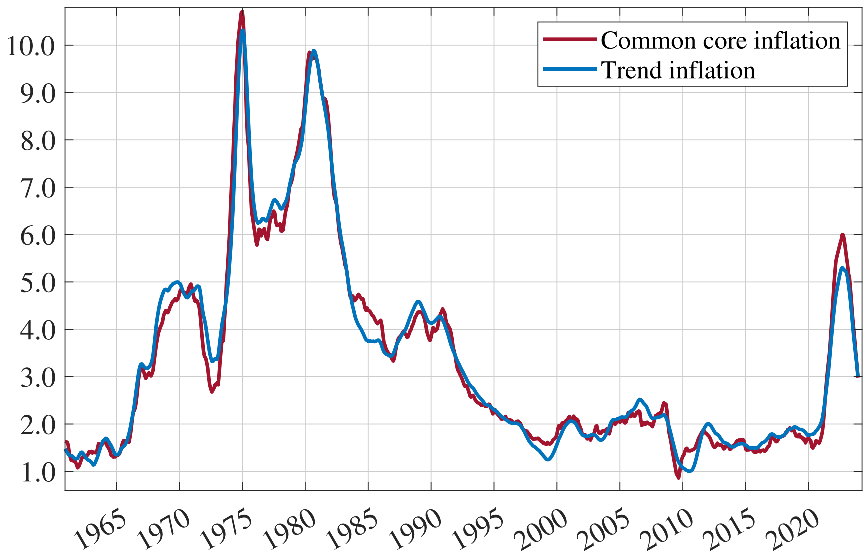
Our model differs from SW's model in two key aspects. First, SW characterize the common factor as a random-walk process with static loadings. Instead, we model it as an autoregressive process with dynamic loadings, capturing potential non-stationarity and persistency through these two sets of parameters. Second, SW have time-varying volatilities and parameters estimated. Instead, we have volatilities and parameters that differ between the two regimes but are fixed within a regime. As such, our model is more restrictive than SW's model.

The question we want to answer in this section is whether the additional flexibility in SW's model helps measuring the underlying co-movements in disaggregated inflation. To answer this question, we estimate SW's model on our dataset and compare the estimate of trend inflation from their model against the estimate of common core inflation.

As shown in Figure 13, the two estimates are very similar, including during the pandemic period. This observation suggests that regime changes combined with dynamic loadings capture the same important features of disaggregated price inflation captured by time-varying volatilities and parameters. This result is in line with the conclusion of Müller (2013) who argues that identification of parameters is not entirely clear in overparameterized time-varying parameter models, and a simpler model can capture the key dynamics crucial for forecasting. In particular, the fact that both models produce similar estimates during the Covid-19 pandemic indicates that our succinct model is flexible enough to adapt to a large shock like Covid-19 and, hence, serve as a useful indicator.

Figure 13: Common core inflation vs trend inflation

Year-over-year inflation



Notes: Trend inflation is estimated by fitting the MUCSVO model of Stock and Watson (2016) on our dataset.

6 Conclusions

This paper introduces common core inflation, a measure that isolates price changes driven by economy-wide shocks from those led by idiosyncratic shocks. Common core inflation is based upon a dynamic factor model estimated on a new large dataset of finely disaggregated PCE price indexes suitable for factor-model analysis.

Our model is designed to comprehensively characterize the changing inflation dynamics from 1959 to the Covid-19 pandemic and beyond. Based on the properties of the disaggregated PCE price data, we set-up a two-regime dynamic factor model with dynamic loadings, which we estimate using frequentist methods. Moreover, based on the evolving covariance structure of the data, we objectively and promptly identify regime changes.

We show that US inflation dynamics experienced two regimes: a long-memory regime up to the mid-1990s and after the Covid pandemic and a short-memory regime from the mid-1990s to the Covid Pandemic. In the long-memory regime, inflation dynamics are primarily driven by common shocks, and the effect of these common shocks is long-lasting; in the short-memory regime, inflation dynamics are primarily driven by idiosyncratic shocks, and the effect of the common shocks is short-lived. Thus, because monetary policy should respond only to common inflation shocks, when inflation is in the long-memory regime, monetary policy should promptly and decisively respond to inflation dynamics, while when inflation is in the short-memory regime, monetary policy should focus on stabilizing the real economy.

References

Amstad, M., Potter, S. M., and Rich, R. W. (2017). The New York Fed Staff Underlying Inflation Gauge (UIG). *Economic Policy Review*, (23-2):1–32.

Antolin-Diaz, J., Drechsel, T., and Petrella, I. (2017). Tracking the slowdown in long-run GDP growth. *The Review of Economics and Statistics*, 99:343–356.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70:191–221.

Baillie, R. T., Chung, C.-F., and Tieslau, M. A. (1996). Analysing inflation by the fractionally integrated arfima–garch model. *Journal of Applied Econometrics*, 11:23–40.

Barigozzi, M., Lippi, M., and Luciani, M. (2021). Large-dimensional dynamic factor models: Estimation of impulse-response functions with $I(1)$ cointegrated factors. *Journal of Econometrics*, 221:455–452.

Barigozzi, M. and Luciani, M. (2019). Quasi maximum likelihood estimation of non-stationary large approximate dynamic factor models. arXiv:1910.09841.

Barigozzi, M. and Luciani, M. (2020a). Measuring the output gap using large datasets. Ssrn, id 3217816.

Barigozzi, M. and Luciani, M. (2020b). Quasi maximum likelihood estimation and inference of large approximate dynamic factor models via the em algorithm. arXiv:1910.03821.v2.

Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short- and long-memory arima models. *Journal of the Royal Statistical Society B*, 57:659–672.

Boivin, J., Giannoni, M. P., and Mihov, I. (2009). Sticky prices and monetary policy: Evidence from disaggregated US data. *American Economic Review*, 99:350–384.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 127:169–194.

Bryan, M. F. and Cecchetti, S. G. (1993). The consumer price index as a measure of inflation. *Economic Review, Federal Reserve Bank*, Quarter 4:15–24.

Bureau of Economic Analysis (2017). Concepts and Methods of the U.S. National Income and Product Accounts.

Canarella, G., Gil-Alana, L. A., Gupta, R., and Miller, S. M. (2020). Modeling US historical time-series prices and inflation using alternative long-memory approaches. *Empirical Economics*, 58:1491–1511.

Carriero, A., Clark, T. E., Marcellino, M., and Mertens, E. (2022). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics*, pages 1–38.

Coibion, O. and Gorodnichenko, Y. (2015). Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7(1):197–232.

Conflitti, C. and Luciani, M. (2019). Oil price pass-through into core inflation. *The Energy Journal*, 40(6):221–247.

Constancio, V. (2015). Understanding inflation dynamics and monetary policy. *Speech at the Jackson Hole Economic Policy Symposium*, 29:455–483.

Cristadoro, R., Forni, M., Reichlin, L., and Veronese, G. (2005). A core inflation indicator for the euro area. *Journal of Money, Credit and Banking*, 37:539–560.

D'Agostino, A. and Giannone, D. (2012). Comparing alternative predictors based on large-panel factor models. *Oxford Bulletin of Economics and Statistics*, 74:306–326.

Del Negro, M., Giannone, D., Giannoni, M. P., and Tambalotti, A. (2019). Global trends in interest rates. *Journal of International Economics*, 118:248–262.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–263.

Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of Econometrics*, 4:115–145.

Diewert, W. E. (1978). Superlative index numbers and consistency in aggregation. *Econometrica*, 46:883–900.

Dolmas, J. (2005). Trimmed mean PCE inflation. Working Paper 506, Federal Reserve

Bank of Dallas.

Dumagan, J. C. (2002). Comparing the superlative Tornqvist and Fisher ideal indexes. *Economics Letters*, 76:251–258.

Eo, Y., Uzeda, L., and Wong, B. (2023). Understanding trend inflation through the lens of the goods and services sectors. *Journal of Applied Econometrics*, 38(5):751–766.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The Generalized Dynamic Factor Model: One sided estimation and forecasting. *Journal of the American Statistical Association*, 100:830–840.

Gadea, M. D. and Mayoral, L. (2006). The persistence of inflation in OECD countries: A fractionally integrated approach. *International Journal of Central Banking*, 2:51–104.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74:1545–1578.

Giannone, D. and Matheson, T. (2007). A new core inflation indicator for new zealand. *International Journal of Central Banking*, 3:145–180.

Granger, C. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1:15–39.

Hallin, M. and Liška, R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102:603–617.

Hosking, J. (1981). Fractional differencing. *Biometrika*, 68:165–176.

Kim, C. and Kim, J. (2022). Trend-cycle decompositions of real GDP revisited: classical and bayesian perspectives on an unsolved puzzle. *Macroeconomic Dynamics*, 26(2):394–418.

Lenza, M. and Primiceri, G. E. (2020). How to estimate a var after march 2020. Working Paper 27771, National Bureau of Economic Research.

Li, M. and Koopman, S. J. (2021). Unobserved components with stochastic volatility: Simulation-based estimation and signal extraction. *Journal of Applied Econometrics*, 36(5):614–627.

Luciani, M. (2014). Forecasting with approximate dynamic factor models: The role of non-pervasive shocks. *International Journal of Forecasting*, 30:20–29.

Maroz, D., Stock, J. H., and Watson, M. W. (2021). Comovement of Economic Activity During the Covid Recession. *mimeo*.

McAlinn, K., Rockova, V., and Saha, E. (2018). Dynamic sparse factor analysis.

Müller, U. and Watson, M. W. (2017). Low-frequency econometrics. In Honoré, B., Pakes, A., Piazzesi, M., and Samuelson, L., editors, *Advances in Economics and Econometrics: Eleventh World Congress*, volume 2, pages 53–94. Cambridge University Press.

Müller, U. K. (2013). Risk of bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica*, 81(5):1805–1849.

Ng, S. (2021). Modeling macroeconomic variations after covid-19. Working Paper 29060, National Bureau of Economic Research.

Peneva, E. (2014). Residual seasonality in core consumer price inflation. FEDS Notes 2014-10-14, Board of Governors of the Federal Reserve System.

Peneva, E. and Sadée, N. (2019). Residual seasonality in core consumer price inflation: An update. FEDS Notes 2019-02-19, Board of Governors of the Federal Reserve System.

Reis, R. and Watson, M. W. (2010). Relative goods' prices, pure inflation, and the Phillips correlation. *American Economic Journal Macroeconomics*, 2:128–157.

Stock, J. H. and Watson, M. W. (1998). Median unbiased estimation of coefficient variance in a time-varying parameter model. *Journal of the American Statistical Association*, 93(441):349–358.

Stock, J. H. and Watson, M. W. (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39(1):3–33.

Stock, J. H. and Watson, M. W. (2016). Core Inflation and Trend Inflation. *The Review of Economics and Statistics*, 98(4):770–784.

Whelan, K. (2002). A guide to U.S. chain aggregated NIPA data. *Review of Income and Wealth*, 48:217–233.

Appendix A Additional details about the model

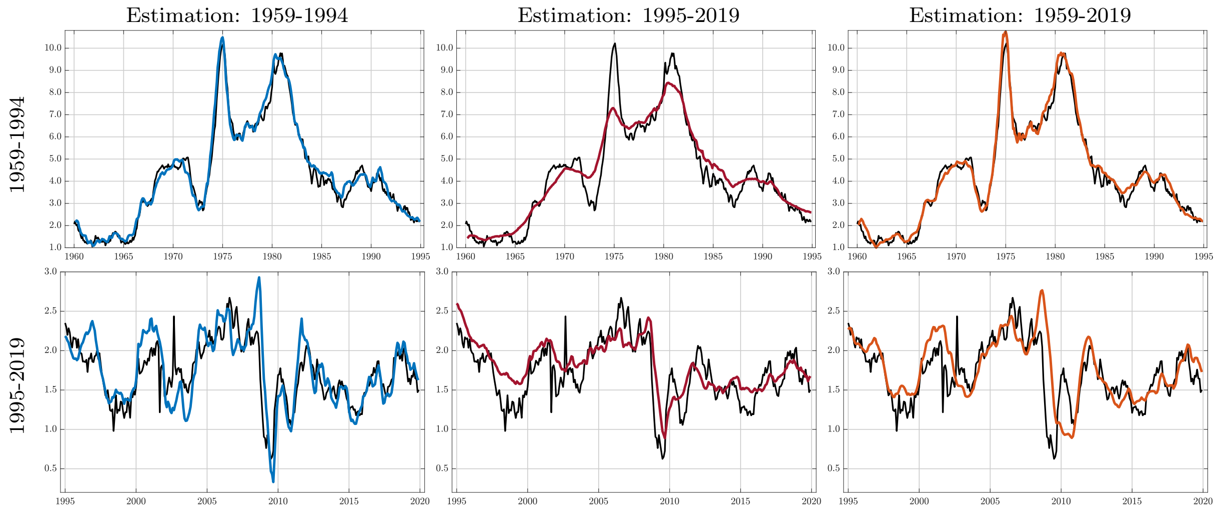
This appendix shows that some of our modeling choices are crucial to correctly representing disaggregated PCE price inflation data. In Appendix A.1, we show that the model estimated over the long-memory regime poorly fits the short-memory regime and vice versa. Likewise, ignoring the regime change and fitting the model over the full sample yields poor results. Thus, fitting a model with two regimes is necessary. Next, Appendix A.2 shows that allowing for a time-varying secular trend in each disaggregated inflation series is crucial to fit the data correctly. Lastly, Appendix A.3 decomposes common core inflation and shows that the secular trend is nothing more than a time-varying mean.

A.1 Do we need two regimes?

Figure A1 shows different estimates of common core inflation obtained from estimating the model's parameters on different samples—1959–94 (henceforth, M1), 1995–2019 (M2), and 1959–2019 (M3)—while the common factor and the time-varying means are estimated over the full sample. As shown in Figure A1, estimating the model over the pre-1995 or post-1995 period yields completely different results. Model M1 fits very well the pre-1995 data but dramatically overfits the post-1995 data. Model M2 captures the primary pre-1995 trend in the data but attributes a much larger part of the pre-1995 fluctuations to the idiosyncratic component.

What if we ignore the break and estimate the model's parameter over the full sample? As shown in the last column of Figure A1, model M3 fits the pre-1995 data well. However, although model M3 does not overfit the post-1995 data, it seems to overestimate the commonality in this period. The estimate of common core inflation during the 2008 recession supports this claim. According to model M3, common core inflation increased at the onset of the recession and fell afterward, thus lagging core inflation. In contrast, according to model M2, core inflation and common core inflation moved largely in sync in 2008 and 2009, when the economy was affected by a large macroeconomic shock, and macroeconomic variation likely dominated idiosyncratic variation in the data.

Figure A1: Common and idiosyncratic decomposition
Model estimated over different samples
Core PCE prices — year-over-year inflation

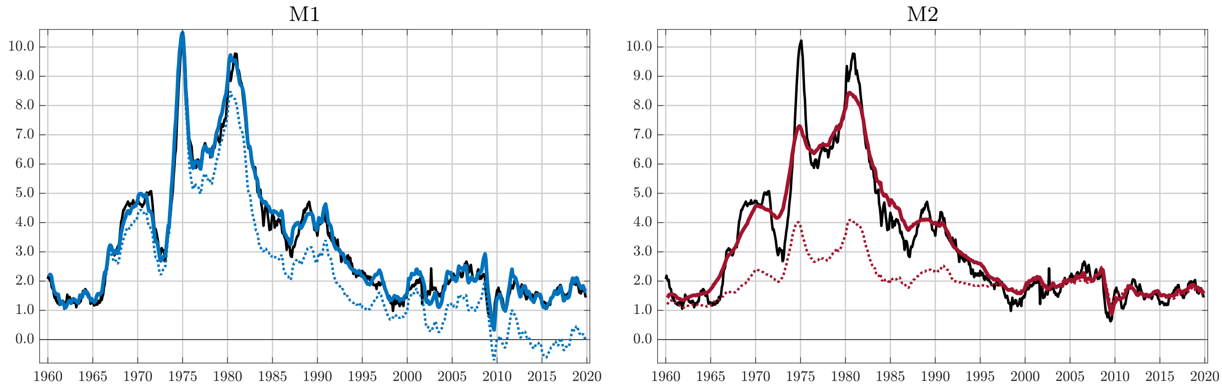


Notes: In all charts, the black line is core inflation. The blue/red/orange line is the estimate of common core inflation obtained by estimating the model on different samples.

A.2 Do we need a time-varying secular trend for each disaggregated inflation series?

Figure A2 shows estimates of M1 and M2 when there is a time-varying secular trend (thick lines) and when there is not (thin lines). The results are clear: A time-varying secular trend is necessary to fit the pre-1990s data appropriately. In contrast, a secular trend is unnecessary if the goal is to estimate a model for the 1995-2019 sample.

Figure A2: Common and idiosyncratic decomposition
Model estimated with and without time-varying secular trend
Core PCE prices — year-over-year inflation

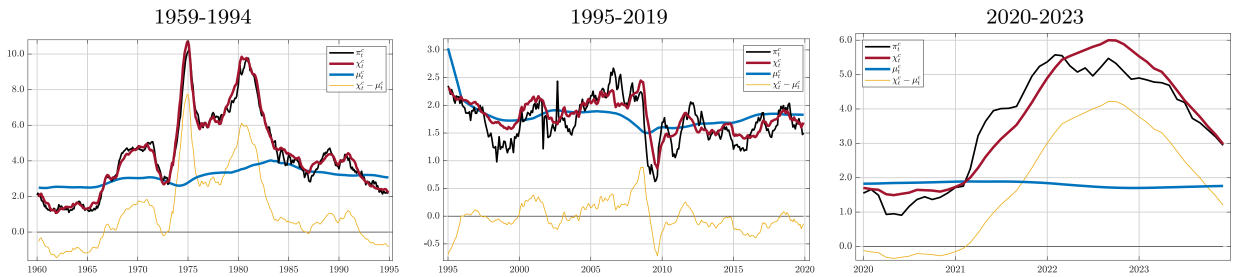


Notes: In all charts, the thick line is the estimate of common core inflation obtained by including a time-varying secular trend. The thin line is the estimate of common core inflation obtained without including a time-varying secular trend.

A.3 Decomposing common core inflation?

Figure A3 decomposes common core inflation by isolating the contribution of the secular trend, $\mu_t^c = \sum_{i \in core} w_{it}\mu_{it}$ in equation (2). As discussed in Section 3.1, we include the secular trend in common core inflation because, in this way, common core inflation has a level comparable to that of the published core PCE price inflation. Figure A3 confirms what we anticipated in the text: μ_t^c is essentially a slow-moving mean and it basically serves the function of aligning common core inflation and the data. Appendix C.2 discusses how χ_t^c and μ_t^c changes when we calibrate differently the variance of μ_t^c.

Figure A3: The components of common core inflation
Core PCE prices — year-over-year inflation



Notes: The black line is actual data. The red line is common core inflation. The blue line is the secular trend component of common core inflation, $\mu_t^c = \sum_{i \in core} w_{it}\mu_{it}$, see also (2). The estimates in the 2020-2023 chart are from the long-memory regime model.

Appendix B Additional results

B.1 Understanding inflation during the Covid lockdowns

As we show in Section 4.1.3, the model that assumes a return to the long-memory regime explains inflation dynamics well from 2021 onwards. However, neither the long-memory nor the short-memory regime model can fit the March–April inflation downturn and the subsequent rebound in the summer of 2020. This is the case because the idiosyncratic component essentially absorbs the unprecedented dynamics driven by the pandemic-specific shock. Thus, we follow Maroz et al. (2021), who suggest including an additional common factor to describe the pandemic dynamics to capture the unprecedented commonality. Following this approach, we estimate the Covid factor as the first principal component of the estimated idiosyncratic component for the pandemic period. In light of the result in the left chart of Figure 9, and given the conclusion of Maroz et al. (2021) that Covid dominated co-movement in the data from March-June 2020, we define the pandemic period as March 2020 to August 2020.

In practice, from March to August 2020, we replace the observation equation of model (1) with

$$\pi_{it} = \mu_{it} + \sum_{k=0}^{s} \boldsymbol{\lambda}_{ik}^{(r)} \boldsymbol{f}_{t-k} + \psi_i g_t + \zeta_{it} \qquad t = 2020:3, \ldots, 2020:8 \tag{B1}$$

where g_t is the Covid factor, and ψ_i is the factor loading of variable i to the Covid factor, and $\xi_{it} = \psi_i g_t + \zeta_{it}$.

Let $\widehat{\xi}_{it} = \pi_{it} - \widehat{\mu}_{it} - \sum_{k=0}^{s} \widehat{\boldsymbol{\lambda}}_{ik}^{(r)} \widehat{\boldsymbol{f}}_{t-k}$ be the idiosyncratic component estimated using equation (1), we estimate $\psi_i g_t$ as the first principal component of $\widehat{\boldsymbol{\xi}}_t$. That is, let $\boldsymbol{\Gamma}_{\widehat{\xi}}$ be the $N \times N$ covariance matrix of $\widehat{\boldsymbol{\xi}}_t$ estimated using observations from 2020:3 to 2028:8, then we estimate $\widehat{\boldsymbol{\psi}}$ as the normalized eigenvector associated to the largest eigenvalue of $\boldsymbol{\Gamma}_{\widehat{\xi}}$, and $\widehat{g}_t$ as $\widehat{g}_t = \sum_{i=1}^{N} \widehat{\psi}_i \widehat{\xi}_{it}$, that is as the weighted cross-sectional average of $\widehat{\boldsymbol{\xi}}_t$ where the weights are the factor loadings $\widehat{\boldsymbol{\psi}}$.

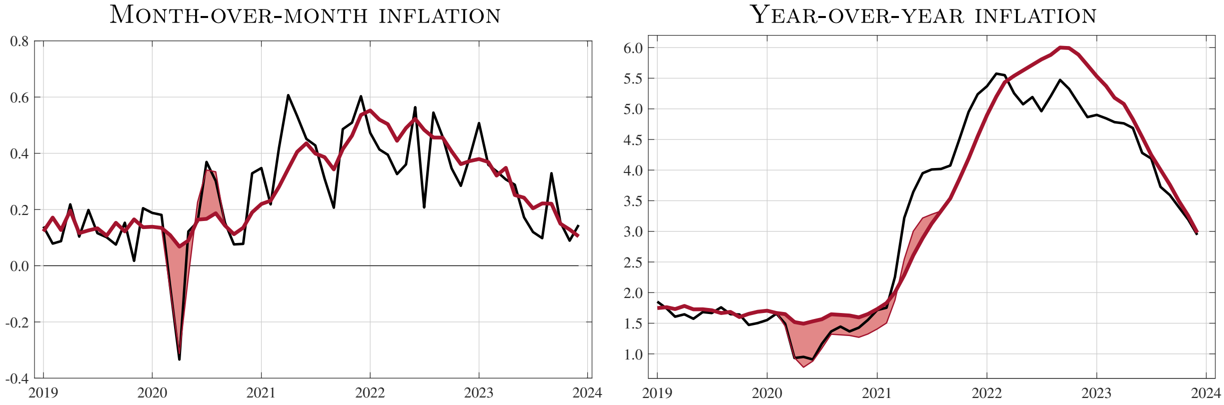
Figure B1 displays core PCE price inflation, common core inflation, and the contribution of the Covid factor. As can be seen, the Covid factor explains 100% of the decline between March and April 2020 and the rebound in the summer of 2020. During these months, the primary drivers of the Covid factor were the prices of financial services, apparel, and airfares (Table B1). Considering that the demand for these items/services was affected by pervasive pandemic-specific factors such as the lockdowns, our model correctly identifies the unusual commonality during the Covid-19 pandemic.

Table B1: Contribution of selected prices to the COVID component

Name	Mar	Apr	May	Jun	Jul	Aug
Apparel	-5.5	-10.1	-3.8	0.2	2.2	1.8
Hospitals	2.5	3.1	2.3	1.7	1.4	1.6
Motor vehicle services	-1.4	-2.9	-1.1	0.2	1.0	0.7
Air transportation	-4.3	-5.1	-2.9	-0.8	-0.2	-0.5
Hotels and motels	-5.4	-9.1	-3.5	-0.3	0.3	0.2
Financial services	-5.0	-12.5	-3.4	4.6	9.4	6.8
NPISHs	-1.2	-1.4	-1.5	-1.9	-1.7	-1.3
Total	-19.4	-38.0	-12.9	6.2	17.4	14.8

Notes: The bottom row is the total contribution of the Covid factor to core PCE price inflation. To be sure, the bottom row tells us that in April 2020, the Covid factor decreased monthly core PCE price inflation by 38.0 basis points. Of these 38.0 basis points, 10.1 were accounted for by Apparel (first line). "Apparel" is the sum of Men's & Boys' Clothing, Women's & Girls' Clothing, and Shoes & Other Footwear (ID 57,58,62).

Figure B1: Common core inflation and Covid specific factor



Notes: The black line is core PCE price inflation. The red line is common core inflation obtained by assuming that during the Covid pandemic, inflation switched back to the long-memory regime. The shaded red area is the contribution of the Covid factor when added to common core inflation.

B.2 Do we need a large number of variables?

Our model is estimated on 142 disaggregated PCE prices. Do we need all these variables? What if we use just four variables (food, energy, core goods, and core services) along the line of the early work of Bryan and Cecchetti (1993)?[26] And what if we go one level down and use 17 variables similarly to Stock and Watson (2016)?

Figure B2 compares the estimate of common core inflation obtained by estimating the model on three different datasets: our benchmark large dataset (142 variables), a medium-sized dataset (the 17 variables listed in Table B2), and a small-sized dataset (4 variables, food, energy, core goods, and core services). The results are clear. When there is enough commonality, as was the case pre-1995, then the size of the dataset does not matter. Conversely, when the idiosyncratic component dominates, as was the case post-1995, having many variables is crucial to correctly parsing out the common component.

B.3 Common headline inflation

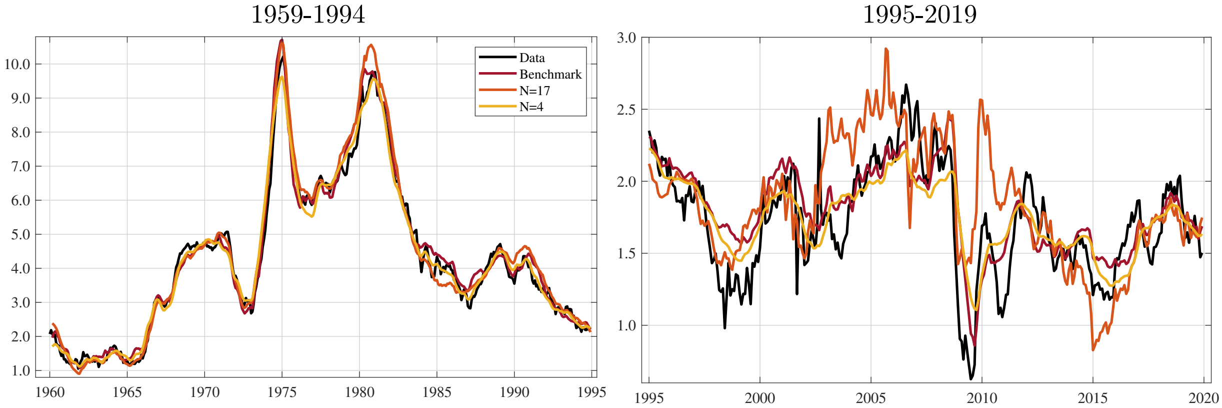
As explained in the Introduction, we focus on core inflation because, from the Fed's official speeches, we can infer that the core PCE price index has a more relevant role in the conduct of monetary policy than the total PCE price index (see also footnote 1). That said, we

[26]Bryan and Cecchetti (1993) propose using limited-information estimators, such as the median of the cross-sectional distribution of inflation, to measure core inflation. Bryan and Cecchetti essentially attempt to exclude the components that create substantial noise in the aggregate price index at high frequencies in their measures.

Figure B2: Common and idiosyncratic decomposition

Model estimated with different number of variables

Core PCE prices — year-over-year inflation



Notes: The benchmark model has 142 variables, the medium-size model has 17 variables (see Table B2, and the small model has four variables (food, energy, core goods, and core services).

Table B2: Variables included in the medium-sized model

CS	Housing	CS	Transportation services
CS	Health care	CS	Water supply and sanitation
CS	Other services	CS	Food services and accommodations
CS	Recreation services	CS	Financial services and insurance
E	Electricity and gas	E	Gasoline and other energy goods
CG	Motor vehicles and parts	CG	Recreational goods and vehicles
CG	Clothing and footwear	CG	Other durable goods
CG	Other nondurable goods	CG	Furnishings and durable household equipment
F	Food and beverages purchased for off-premises consumption		

Notes: CS = Core Services; CG = Core Goods; E = Energy; F = Food.

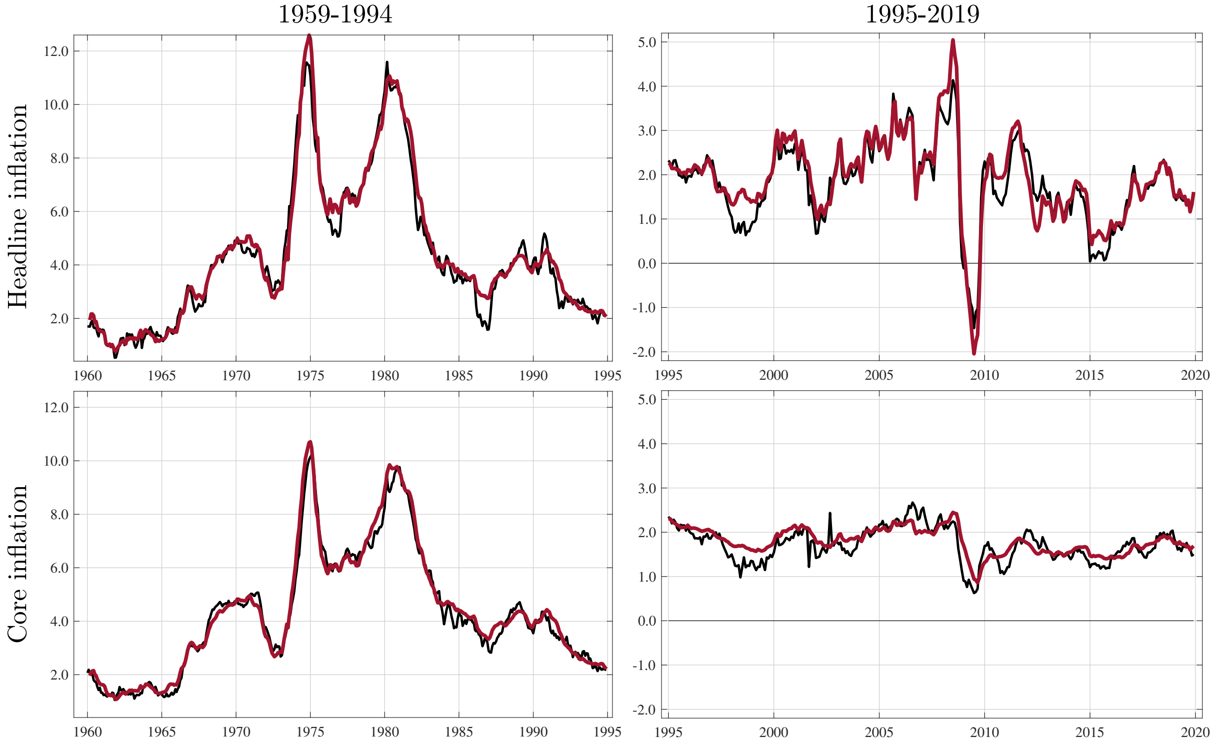
can easily construct an index of common inflation by slightly modifying equation (2) as follows:

$$\chi_t^h = \sum_{i=1}^{N} w_{it}(\mu_{it} + \chi_{it}). \tag{B2}$$

The upper charts in Figure B3 shows YoY headline inflation and common headline inflation. For comparison purposes, we report YoY core and common core inflation in the lower charts (that is, the same numbers reported in Figure 7 but on a different scale).

In a nutshell, common headline inflation fits headline inflation (upper charts in Figure B3) very well, much better than common core inflation fits core inflation (lower charts). Because the difference between core and headline inflation is primarily due to energy prices, this result confirms the intuition in Section 4.2, according to which, to a large extent, the common factor picks up oil/commodity prices. Common core inflation fits the data worse than common headline inflation because oil/commodity prices pass through core inflation indirectly by affecting input prices and through second-round effects, and this pass-through is smaller than Conflitti and Luciani (2019).

Figure B3: COMMON AND IDIOSYNCRATIC DECOMPOSITION
HEADLINE AND CORE PCE PRICES — YEAR-OVER-YEAR INFLATION



NOTES: In each plot, the red line denotes year-over-year common headline/core inflation, while the black line is year-over-year headline/core inflation. The plot on the left covers the highly persistent period (1960 to 1994), while the right panel covers the stable period (1995 to 2019). To make it easier to understand what happens in the two periods, the charts on the left have different y-axes compared to those on the right.

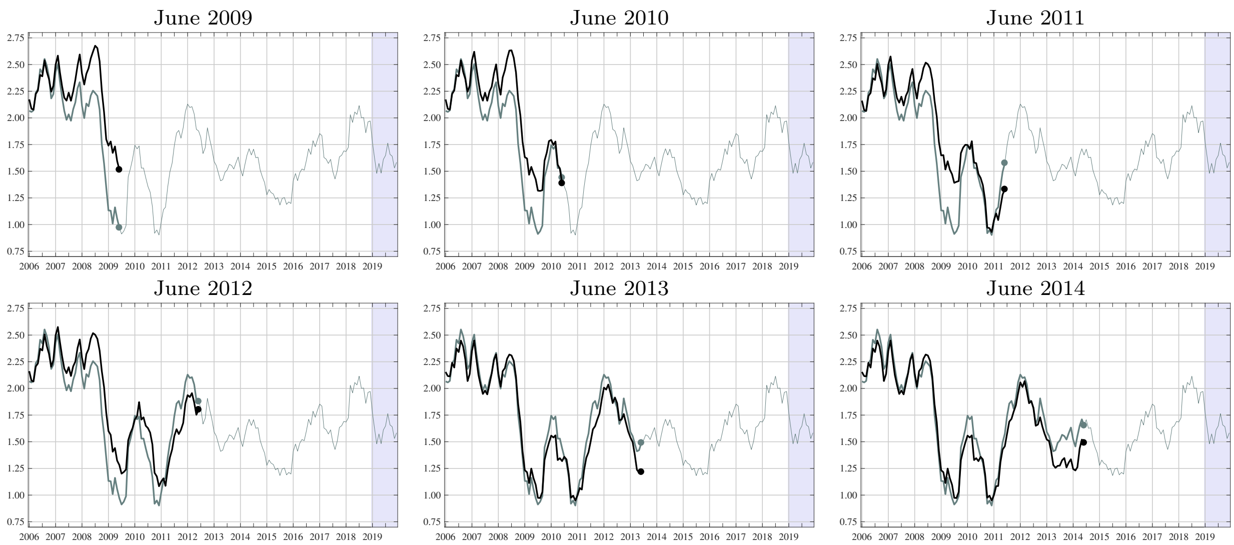
B.4 Real-time stability

In Section 5, we already demonstrate that common core inflation successfully parses out idiosyncratic and transitory disturbances in core inflation and correctly identifies the persistent portion relevant to monetary policy. However, for this indicator to be useful in practice, it is necessary not to revise too much as new observations become available and the model is re-estimated, an issue that often plagues models with unobserved variables. In other words, we do not want this model to drastically change its decomposition of current inflation after receiving a few additional data points.

There are two sources of revisions in the estimates of unobserved variables. First, the data themselves get revised. Second, new observations change the inference in history by updating the model parameters or the smoothed estimates. We examine how robust the estimates are to changes in the data. In order to estimate common core inflation in real time, we retrieved real-time data vintages for our dataset starting in August 2009—that is, after the 2009 NIPA comprehensive data revision.[27] Overall, we retrieved 127 data vintages, including the one used to produce previously reported results.

Figure B4 shows YoY core PCE price inflation computed using selected vintages of data (the black line) together with YoY core PCE price inflation computed using the "final" vintage of data (the gray line).[28] The distance between the black and the gray lines reflect the magnitude of data revisions. Looking at Figure B4, it is clear the 2013 comprehensive revision of the NIPAs brought sizable changes to the core PCE price inflation.[29] Outside of this, however, core PCE price inflation does not revise much. For this reason, we do not display the data from the post-2014 vintages.

Figure B4: CORE PCE PRICE INFLATION IN REAL-TIME YEAR-OVER-YEAR INFLATION



Note: In each plot, the black line is year-over-year core PCE price inflation computed using the data actually available at each point in time, while the gray line is year-over-year core PCE price inflation computed using the vintage of data available.

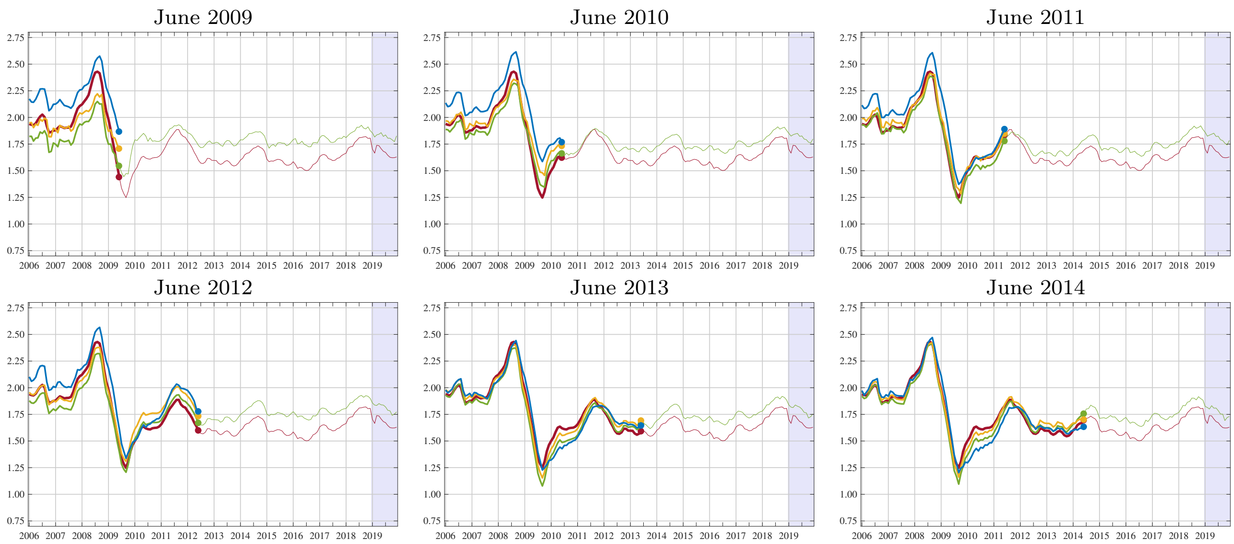
[27]Because the structure of PCE changed as a result of the 2009 comprehensive revision, it is challenging to extend the real-time analysis further back in time.

[28]Specifically, for each year, we show the vintage of data ending in June—that is, the one incorporating the annual update of the NIPAs, which normally is published at the end of July (or beginning of August) of the same year.

[29]The 2013 comprehensive revision of the NIPAs had a particularly relevant effect on the imputed price of banking services and on the price of medical and hospitalization insurance as well as income loss insurance.

Now, we examine revisions in common core inflation caused by the data revisions. Figure B5 reports the estimates of common core inflation obtained from the data vintages displayed in Figure B4. Each plot in Figure B5 has four lines. The blue line is the "real-time" estimate of common core inflation—that is, the one computed using the data actually available at each point in time. The red line is the "final" estimate of common core inflation—that is, the one computed using the latest available data (meaning the data published by the BEA on January 31, 2020). The difference between the blue and red lines is due to all the factors we mentioned at the beginning of this section, and it gives us a measure of the real-time reliability of common core inflation. The yellow line is the "*quasi*-real-time" estimate of common core inflation obtained by estimating the model on the "final" vintage of data but on the same period covered by the "real-time" vintage. The difference between the blue and yellow lines is due to data revision only. Finally, the green line is the "*quasi*-final" estimate of common core inflation obtained by estimating the parameter of the model on the "final" vintage of data but on the same period covered by the "real-time" vintage, and the smoothed estimates of the common factor on the "final" vintage up to December 2019. In other words, the difference between the green and yellow lines is solely due to the estimate of common factors. In contrast, the difference between the red and green lines is due to the parameters estimate.

Figure B5: Common core inflation in real-time year-over-year inflation



Note: In each plot, the blue line is the real-time estimate of common core inflation. The yellow line is the *quasi*-real-time estimate of common core inflation. The green line is the *quasi*-final estimate of common core inflation. The red line is the final estimate of common core inflation.

Table B3 shows the mean absolute revision for core PCE price inflation and common core inflation computed over all the 127 vintages. Here we define a "revision" as the difference between the real-time (or *quasi*-real-time, or *quasi*-final) estimate and the final estimate. Put differently, in each plot, the revision of common core inflation is the difference between the blue/yellow/green dot and the red dot—in the case of core PCE price inflation, the difference between the black dot and the gray dot. The results in the table are split into two parts, before and after the 2013 comprehensive revision, which, as discussed above, hugely impacted core PCE prices.

Table B3: Average absolute revision

	Inflation	Period	$\boldsymbol{\pi}_t^c$	$\boldsymbol{\chi}_{t,RT}^c$	$\boldsymbol{\chi}_{t,QRT}^c$	$\boldsymbol{\chi}_{t,QFin}^c$
(1)	month-over-month	2009:6 – 2013:5	5.5	1.6	1.5	1.1
(2)		2013:6 – 2019:12	3.6	1.6	1.4	1.0
(3)	year-over-year	2009:6 – 2013:5	23.1	16.1	12.6	7.9
(4)		2013:6 – 2019:12	13.1	3.5	3.7	4.6

Note: π_t^c is core PCE price inflation, $\chi_{t,RT}^c$ is the real-time estimate of common core inflation, $\chi_{t,QRT}^c$ is the *quasi*-real-time estimate, and $\chi_{t,QFin}^c$ *quasi*-final estimate. The term "revision" indicates the difference between the real-time (or quasi-real-time, or quasi-final) and the final estimate. The average absolute revision is expressed in basis points. Column "Period" indicates the period upon which the averages are computed.

Between June 2009 and May 2013, our YoY common core inflation estimate underwent sizable revisions (line 3). However, the average size of these revisions is more than 30% smaller than those for core PCE price inflation. After the 2013 comprehensive revision, although the revisions of YoY core PCE price inflation (line 4) almost halved, common core inflation revises much less than core inflation itself. Finally, the average absolute revision for the real-time estimate and the quasi-real-time estimate of month-over-month common core inflation (lines 1 and 2) is about the same, thus indicating that data revisions play a negligible role. In contrast, the model estimate really dominates. Within the model estimate, about a third of the revision is due to the smoothed estimate of the factors and two-thirds to revision to the parameter estimate.[30]

To sum up, the results in this section suggest that common core inflation is a reliable measure in real time. This result paired with the forecasting evaluation in Appendix B.5

[30] As discussed in Section 3.2.2, we introduce an additional pandemic-specific factor to model the extreme fluctuations that occurred in 2020 during the lockdowns and the reopening. Given the limited number of observations, it is nearly impossible to credibly estimate the Covid factor in real time. For this reason, we end the sample period of this analysis at 2019:M12 and exclude the Covid pandemic from the real-time exercise.

confirms once again the usefulness of common core inflation as a tool to read current inflation development, a crucial task for implementing monetary policy.

B.5 Forecasting

As we explained in the Introduction, the main goal of common core inflation measure is to help understand what is driving recent and current movements in core inflation. That said, much of the literature on inflation gauges has focused on the ability of these measures to serve as a monitoring device to gauge signals about future inflation. Therefore, in this section, we provide a real-time out-of-sample assessment of the forecasting performance of common core inflation against other well-known measures, where we run real-time exercises as Section B.4 explains.

We investigate the performance of the annualized common headline/core inflation between month $t-h$ and month t in predicting the annualized headline/core inflation rate between month t and month $t+h$. In practice, we use the following forecasting equation:

$$\pi^i_{t+h|t} = \pi^j_{t|t-h} + \varepsilon^j_{t|t-h}, \tag{B3}$$

where $i = \{h, c\}$ denotes either headline or core, j denotes different underlying inflation measures, and $\varepsilon^j_{t|t-h}$ is a forecast error.

To predict core inflation, we consider the following predictors: core inflation itself, common core inflation, and a UIG-style indicator for core inflation as explained in Appendix D. That is, when $\pi^i_{t+h|t} = \pi^c_{t+h|t}$, $\pi^j_{t|t-h} = \{\chi^c_t, UIG_t\}$. To predict headline inflation we consider the following predictors: headline inflation itself, core inflation, the Dallas Fed Trimmed Mean, common headline inflation, and a UIG-style indicator for headline inflation. That is, when $\pi^i_{t+h|t} = \pi^h_{t+h|t}$, $\pi^j_{t|t-h} = \{\pi^c_t, DTM_t, \chi^h_t, UIG_t\}$.

The results of this exercise are in Table B4 for core inflation, and Table B5 for headline inflation. The tables report the mean squared error (MSE) of each indicator relative to the MSE of the target variable forecasting itself (that is,$\pi^j_{t|t-h} = \pi^i_{t|t-h}$), as well as some tests of equal predictive ability. In a nutshell, common core inflation is a good predictor of core inflation, clearly outperforming core inflation, and slightly outperforming the UIG-style core measure (Table B4). Moreover, common core inflation is also a good predictor of headline inflation performing a little bit better than the other models at short horizons, and better than common headline inflation at all horizons.

Table B4: Forecasting core inflation

Relative mean squared error			Testing Equal Predictive Ability			
h	χ_t^c	υ_t^c	h	$\chi_t^c = \pi_t^c$	$\upsilon_t^c = \pi_t^c$	$\upsilon_t^c = \chi_t^c$
1	0.54	0.57	1	0.00	0.00	0.48
3	0.48	0.59	3	0.00	0.02	0.15
6	0.49	0.66	6	0.01	0.31	0.20
12	0.43	0.71	12	0.01	0.45	0.19
24	0.81	1.17	24	0.19	0.70	0.24
36	0.80	1.74	36	0.28	0.08	0.00

Notes: The left panel shows the mean squared error of common core inflation (χ_t^c) and UIG-style core (υ_t^c)in predicting core inflation. The MSEs are shown relative to the MSE of common core inflation predicting itself. A value smaller than 1 means that the alternative model outperforms core inflation. When comparing the numbers in the two columns, the smaller the number the better the forecasting performance of the model.

The right panel shows the p-value of a test of equal predictive ability, where the null hypothesis is in the header of the Table. The test used here is the test of "unconditional predictive ability" of Giacomini and White (2006), which is equivalent to the Diebold and Mariano (1995) test statistic. When using rolling window estimations, this test statistics has a standard normal limit distribution.

B.6 Determining the number of factors

Table B6 reports results for the Hallin and Liška (2007) and Barigozzi et al. (2021) criteria. Because these two criteria select the number of common shocks by looking at randomly selected sub-samples of the original data, we run the procedure several times for robustness. Table B7 reports the results for the Bai and Ng (2002) information criteria.

Appendix C Robustness analysis

C.1 Alternative switching point for regime change

As we wrote in Section 3.2, our model assumes that there are two regimes and that the switching point from one regime to the other was in January 1995. Figure C1 shows the estimate of common core inflation from 2000 onwards when we change the switching point—we do not show the estimates pre-2000 because they are virtually identical. As shown in the left chart, if we move backward the switching point back to January 1993 or 1994, the results do not change; if we move the switching point to January 1990-1992, the results change slightly. Likewise, as shown in the right chart, if we move forward the switching point forward to January 1996–2000, the results are virtually unchanged.

Table B5: Forecasting headline inflation

Relative mean squared error

	h	π_t^c	π_t^D	χ_t^h	χ_t^c	υ_t^h	υ_t^c
Relative mean squared error	1	0.73	0.64	0.75	0.59	0.64	0.60
	3	0.65	0.64	0.72	0.54	0.61	0.54
	6	0.67	0.67	0.80	0.52	0.68	0.51
	12	0.73	0.86	0.85	0.60	0.79	0.60
	24	0.38	0.45	1.00	0.40	0.84	0.46
	36	0.34	0.37	1.08	0.37	1.23	0.62
Test of equal predictive ability vs. π_t^h	1	0.09	0.03	0.04	0.01	0.03	0.02
	3	0.04	0.03	0.01	0.00	0.02	0.01
	6	0.03	0.02	0.11	0.00	0.05	0.00
	12	0.16	0.51	0.34	0.02	0.40	0.03
	24	0.00	0.01	0.98	0.01	0.64	0.05
	36	0.00	0.00	0.68	0.00	0.35	0.01
Test of equal predictive ability vs. χ_t^c	1	0.01	0.04	0.07		0.14	0.51
	3	0.02	0.00	0.08		0.26	0.97
	6	0.06	0.01	0.02		0.16	0.87
	12	0.16	0.02	0.17		0.28	0.99
	24	0.74	0.04	0.00		0.01	0.45
	36	0.55	0.87	0.00		0.00	0.00

Notes: The upper panel shows the mean squared error of core inflation (π_t^c), the the Dallas Fed Trimmed Mean PCE (π_t^D), common headline inflation (χ_t^h), common core inflation (χ_t^c), UIG-style headline (υ_t^h) υ_t^h, and UIG-style core (υ_t^c) in predicting headline inflation. The MSEs are shown relative to the MSE of headline inflation predicting itself. A value smaller than 1 means that the alternative model outperforms headline inflation. When comparing the numbers in the six columns, the smaller the number the better the forecasting performance of the model.

The middle and lower panel shows the p-value of a test of equal predictive ability. In the middle panel, the null hypothesis is that the forecasting ability of a given model is the same as those of headline inflation. In the lower model, the null hypothesis is that the forecasting ability of a given model is the same as those of common core inflation. The test used here is the test of "unconditional predictive ability" of Giacomini and White (2006), which is equivalent to the Diebold and Mariano (1995) test statistic. When using rolling window estimations, this test statistics has a standard normal limit distribution.

C.2 Alternative calibration of the variance of the secular trends

As discussed in Section 3.2.1, we calibrated the variance of each secular trend as $\sigma_\eta^2 = \frac{1}{600}$. This variance implies that the expected change in μ_{it} over 50 years has a standard deviation equal to the standard deviation of π_{it}. Figure C2 shows that our results are robust to reasonable changes in this parameter, namely $\sigma_{\eta_i}^2 = \{\frac{1}{120}, \frac{1}{300}, \frac{1}{1200}, \frac{1}{1800}\}$—that is, when the expected change in μ_{it} over 10, 25, 100, and 150 years has a standard deviation equal to the standard deviation of π_{it}.

To better understand the results in Figure C2, let us rewrite common core inflation from (2) as $\chi_t^c = \mu_{it}^c + \tilde{\chi}_{it}^c$ where $\mu_{it}^c = \sum_{i \in core} w_{it}\mu_{it}$ and $\tilde{\chi}_{it}^c = \sum_{i \in core} w_{it}\chi_{it}$. Then, the results in Figure C2 show that changing $\sigma_{\eta_i}^2$ has no effect on χ_t^c but only on the share of

Table B6: Number of common shocks q

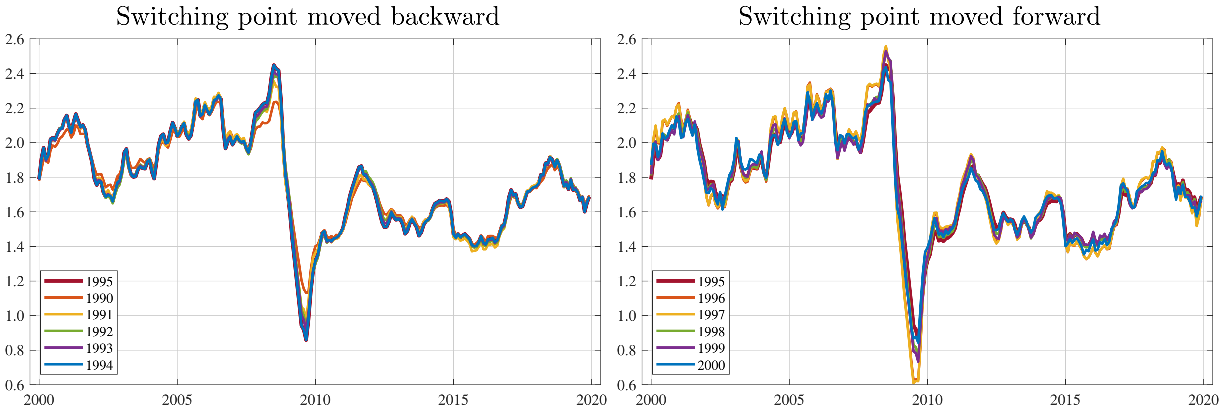
	Hallin and Liška (2007)		Barigozzi et al. (2021)	
q	0	1	0	1
1959-1994	100	0	11	89
1995-2019	0	100	85	15

Notes: Each entry reports the number of times out of 100 repetitions that a certain number of common shocks is selected.

Table B7: Bai and Ng (2002) information criteria for $\tilde{q} = q(s+1)$

	PC1	PC2	PC3
1959-1994	3	3	9
1995-2019	2	2	5

Figure C1: Common and idiosyncratic decomposition
Model estimated when changing the regime switching point
Core PCE prices — year-over-year inflation



Notes: In each plot, each line denotes the estimate of year-over-year common core inflation when the short-memory regime starts in January of the year indicated in the legend. January 1995 is our benchmark model.

χ_t^c accounted for by μ_{it}^c.

C.3 Alternative model specifications

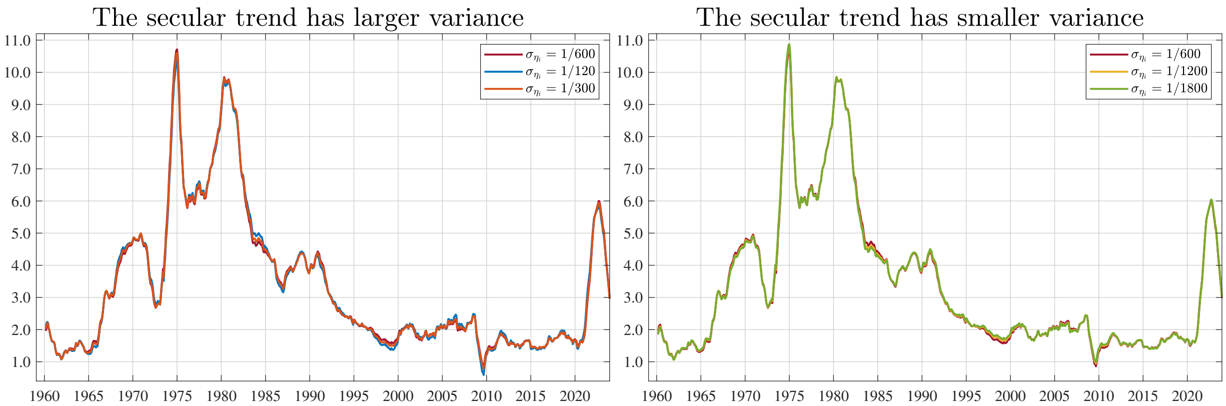
In this section, we estimate common core inflation by using three alternative specifications of the dynamic factor model:

1) Model S0, $q = 1$ and $s = 0$. In this specification, each disaggregated price loads the common factor only contemporaneously. The rationale for including this specification is to show what benefits we obtain from including lagged factor loadings.
2) Model S1, $q = 1$ and $s = 1$. Our benchmark specification has $s = 2$. However, as discussed in Section 3.3, there was evidence supporting a model with $s = 1$.

Figure C2: Common and idiosyncratic decomposition

Model estimated when changing the variance of the secular trend

Core PCE prices — year-over-year inflation



Notes: In each plot, each line denotes the estimate of year-over-year common core inflation obtained by using different values to calibrate the variance of the secular trend.

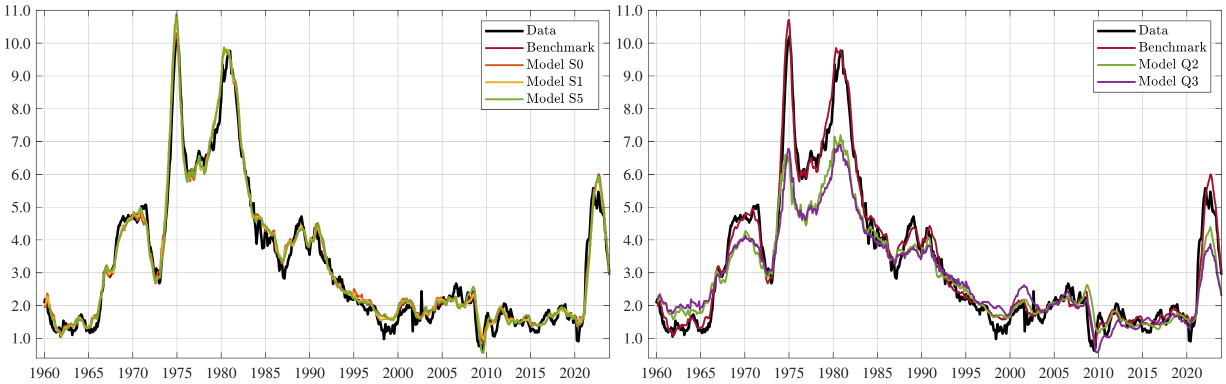
3) Model S5, $q = 1$ and $s = 5$. This is a much richer specification in which each disaggregated price can load the common factor in six months.
4) Model Q2, $q = 2$ and $s = 4$. This specification shows the effect of adding one additional common factor, where s is set in the same way as in Section 3.3.
5) Model Q3, $q = 3$ and $s = 6$, same as Model Q2.

Model S1 is the only one with support in the data; the other specifications are not supported (see Section 3.3).

As shown in Figure C3, our estimates are robust to the first three alternative model specifications, particularly those with dynamic loadings. Our benchmark model and model S1 are almost identical (the yellow line is almost above the red line). Model S0 yields an estimate of YoY common core inflation much smoother than that produced by the other models, indicating that the dynamic loadings help us capture quite a lot of the dynamics in the data. Finally, model S5 produces an estimate very similar to our benchmark estimate, just a touch more volatile.

In contrast, models Q2 and Q3 produce estimates of common core inflation quite differently than our benchmark specification. As discussed in Section 3.3, no evidence supports including more than one common factor. As a result, when we include one or two additional factors, we pick mainly idiosyncratic noise, and the fit of common core inflation worsens.

Figure C3: Common and idiosyncratic decomposition
Alternative model specifications — Year-over-year inflation



Notes: In each plot, the black line is year-over-year core PCE price inflation. The red line is the year-over-year common core inflation estimate from our benchmark model. All the other lines are year-over-year common core inflation estimates from different specifications of the model.

Appendix D Literature review

Dynamic factor models have been a popular tool for modeling inflation dynamics. However, the model specification differs across studies depending on the purpose or research question. This Appendix discusses how our approach differs from Cristadoro et al. (2005), Reis and Watson (2010), and some recent studies directly relevant for this paper. We compared our model with that of Stock and Watson (2016) in Section 5.

To simplify the comparison between the different models, we simplify and rewrite our model in its static representation. Specifically, equations (1) and (5) are collapsed into the following:

$$\pi_{it} = \tilde{\boldsymbol{\Lambda}}_i^{(r)} \boldsymbol{F}_t^{(r)} + \xi_{it}^{(r)} \tag{D1}$$

where $\tilde{\boldsymbol{\Lambda}}_{i^{(r)}} = [\boldsymbol{\lambda}_{i0}^{(r)} \; \boldsymbol{\lambda}_{i1}^{(r)} \; \dots \; \boldsymbol{\lambda}_{is}^{(r)}]$ and $\boldsymbol{F}_t^{(r)} = [\boldsymbol{f}_t^{(r)\prime} \; \boldsymbol{f}_{t-1}^{(r)\prime} \; \dots \; \boldsymbol{f}_{t-s}^{(r)\prime}]'$ are of dimension $\tilde{q} \times 1$ with $\tilde{q} = q(s+1)$.

Comparison with Cristadoro et al. (2005). We first compare our model with that of Cristadoro et al. (2005), henceforth CFRV, which is also adopted by Giannone and Matheson (2007) and Amstad et al. (2017). CFRV estimate a dynamic factor model using Generalized Principal Components (Forni et al., 2005) on a large dataset of European data. CFRV's model is similar to ours but is different in two key aspects. First, and most importantly, unlike our approach, CFRV do not allow for regime changes in the model.

Second, CFRV do not consider higher-frequency fluctuations when estimating the common component. Specifically, CFRV exclude high-frequency fluctuations in the data, which are likely to be less relevant for forecasting inflation one or more years ahead; in contrast, we take all the fluctuations on board and let the model filter out variations in disaggregated prices that are idiosyncratic. To put it another way, CFRV apply both cross-sectional and temporal smoothing, whereas we only apply cross-sectional smoothing.

Our model is more flexible and less restrictive than CFRV, and it has the comparative advantage that it can capture a signal from higher-frequency fluctuations that is pervasive across disaggregated prices and carry information useful for inflation nowcasting and forecasting common core inflation. Indeed, even without the additional temporal smoothing, our common core inflation effectively captures the persistent co-movement of disaggregated price changes.

The question we want to answer is whether the additional flexibility in our model helps measuring the underlying co-movements in disaggregated inflation. To answer this question, we compare common core inflation with the Underlying Inflation Gauge (UIG) of Amstad et al. (2017) that the NY Fed discontinued in October 2023.

When comparing common core inflation with the UIG, it is necessary to consider two key differences. First, the UIG is a measure of underlying *headline* inflation and not core inflation. Second, the UIG is a measure of underlying inflation for the CPI and not for PCE. As explained in the FAQs available on the website of the NY Fed, the NY Fed Staff used to produce a measure of underlying inflation for the PCE deflator but did not share it with the public. Therefore, we decided to estimate a custom version of the UIG by fitting the model of Cristadoro et al. (2005) (the same model underlying the UIG) on our dataset with the same parametrization as common core inflation. Doing so yields an estimate of the common component for each price index, which is then aggregated in the same way as common core inflation. This is clearly not the same thing as the UIG, but we believe it is the best apple-to-apple comparison we can do. Henceforth, with a clear abuse of terminology, we will refer to this model as UIG-style.

Figure D1 compares common core inflation with the UIG-style core measure. As can be seen, the UIG-style estimate is much less volatile than common core inflation, with small deviations from the in-sample mean. This is the result of both the different estimation method, the truncation of higher frequencies.

Comparison with Reis and Watson (2010). Reis and Watson (2010), henceforth RW, estimate a dynamic factor model on finely disaggregated PCE price changes like ours. However, there are three key differences between RW's model and ours. First, RW do not comprehensively consider long memory and/or nonstationarity in disaggregated price

Figure D1: Common core inflation vs UIG

Year-over-year inflation

Notes: the black line is core inflation, the red line is common core inflation, and the yellow line is the UIG-style core measure estimated on our dataset.

changes.[31] Specifically, RW do not allow for regime changes, while our model has both a short-memory and a long-memory regime with the possibility of switching between the two. In addition, RW do not allow for nonstationarity in the idiosyncratic component, while we do so through a time-varying mean for each disaggregated price change.

Second, RW produce pure inflation and relative price inflation, while we produce common core inflation. Pure inflation is the part of the common component that has equiproportional effects on all disaggregated prices; its purpose is to gauge the price changes that reflect anticipated policy changes. Relative price inflation is the part of the common component that has different effects on disaggregated prices; its purpose is to capture relative price changes that have Phillips correlations with real activities. These two measures are amenable to the theory on the Phillips curve that they lay out in the paper. To recover these two indices, RW impose restrictions on the structure of $\tilde{\boldsymbol{\Lambda}}_i$ and $\boldsymbol{F}_t$—specifically, $\tilde{\boldsymbol{\Lambda}}_i = [\mathbf{1}\ \boldsymbol{\Gamma}_i]$ and $\boldsymbol{F}_t = [a_t\ \boldsymbol{R}_t']'$, where $\boldsymbol{F}_t$ is an $r \times 1$ vector that follows a VAR driven by r shocks. These restrictions are unnecessary in our case because our purpose is to recover underlying co-movements in disaggregated prices, controlling for the noise in the data.

Third, we assume that of $q \leq \tilde{q}$ common shocks, where $q = 1$ and $\tilde{q} = 3$, drive the

[31] Strictly speaking, RW allow for one or more unit roots in the VAR of the common factors. However, as discussed in Section4.1.1, the regime change implies a change in the factor loadings rather than in the persistence of the factors, which is strong in both regimes.

co-movement in the data, RW assume $q = \tilde{q} = 3$. In other words, we consider dynamic loadings, while RW employ static loadings.

Other recent studies on inflation dynamics directly relevant to this paper. One strand of literature on inflation dynamics has adopted an unobserved component time-series model with stochastic volatility (for example, Stock and Watson, 2016; Li and Koopman, 2021).[32] This model is usually estimated with Bayesian methods; however, recently Li and Koopman (2021) proposed a frequentist approach based on simulated maximum likelihood estimation.

To model inflation dynamics during the pandemic and the recovery, Antolin-Diaz et al. (2017) and Carriero et al. (2022) consider outlier treatments and stochastic volatility in their dynamic models, similar to Stock and Watson, 2016. However, while Antolin-Diaz et al. (2017) consider a Bayesian dynamic factor model for a mixed-frequency dataset, Stock and Watson, 2016 analyze a fixed-frequency disaggregated data. Carriero et al. (2022) consider a monthly Bayesian VAR model, while Stock and Watson, 2016 a quarterly dynamic factor model.

Instead of adopting stochastic volatility and outlier treatments for the Covid-19 pandemic, we explicitly bring in a pandemic-specific factor following Maroz et al. (2021). All told, even though our approach is more succinct and computationally less costly than the approach based upon the time-varying parameter model with stochastic volatility, as shown in Section 5, we do not lose much information in estimating common core inflation during the pandemic.

[32] Relatedly, McAlinn et al. (2018) propose a dynamic sparse factor model where both the factor loadings and the variance of the idiosyncratic component change over time. This model is designed to characterize the changing commonality in a large disaggregated dataset. However, McAlinn et al. apply their method to a dataset of various macroeconomic variables rather than disaggregated inflation data. In spite of the novelty, McAlinn et al.'s method is more restrictive than that of SW because they do not account for long-memory in the data, outlier treatments, and stochastic volatility in the innovation of the common component.

Common and Idiosyncratic Inflation

Hie Joo Ahn
Federal Reserve
hiejoo.ahn@frb.gov

Matteo Luciani
Federal Reserve
matteo.luciani@frb.gov

Complementary Appendix

Table CA1 shows detailed information about the dataset used to estimate the common and idiosyncratic decomposition of core PCE prices. The table has five columns: columns "ID", "Item", and "Haver" report for each price index the identification number on our dataset, on NIPA Table 2.4.4U, and on the Haver USNA database, respectively. Column "PCE Component" reports the name of each PCE price component, while the column "Price index source data" reports the source that the BEA uses to construct that PCE price.[1] The sixth column of the table reports for some item a flag in four different symbols:

◇ All the entries that have a flag denoted by the "•" symbol are "PCE" price indexes that (actually) do not exist, i.e., they are not available in the NIPA Table 2.4.4U. These price indexes are constructed by us and are aggregation of PCE price indexes that are (actually) available in Table 2.4.4U. These PCE price indexes have all the same source data, and therefore they are nearly identical. There are overall 14 of such "PCE" price indexes, and specific information on each of them are available in Table CA2.

---

*We are grateful to Jeremy Rudd, Andrew Figura, Lucas Moyon and seminar participants at the Federal Reserve Board, Bank of Italy, European Central Bank, Bank of England, and Reserve Bank of Australia. Any errors are our responsibility.

[1]For detailed information on source data for PCE price index we refer the reader to the excel file that can be downloaded at `https://www.bea.gov/media/3051`.

- Suppose we have to aggregate the price and the quantity index of n items. Let q_{it} be the quantity index for item i at time t, and let p_{it} be the price index for item i at time t. Then, let P_t the aggregate price index, then for the Fisher formula we have:

$$P_t = P_{t-1}\sqrt{\frac{\sum_{i=1}^{n} p_{it}q_{it-1}}{\sum_{i=1}^{n} p_{it-1}q_{it-1}} \times \frac{\sum_{i=1}^{n} p_{it}q_{it}}{\sum_{i=1}^{n} p_{it-1}q_{it}}} \qquad \text{for } t = 1, \dots, T-1 \tag{1}$$

$$P_T = \frac{1}{n}\sum_{i=1}^{n} p_{iT} \tag{2}$$

  where (2) is necessary to fix the scale.

- Note that formula (1) is not always necessary. Indeed, in many of the aggregation that we perform, the indexes that we are aggregating are actually the same. In some other cases, the index are identical for most of the sample, but not for all the sample. This is the case because of some change of methodology in the way the BEA sourced or built the index. In those cases, formula (1).

- Once P_t is constructed, we construct the quantity index as if $p_{1t} = p_{2t} = \dots = p_{nt} = P_t$, and hence $Q_t = \sum_{i=1}^{n} q_{it}$. In other words no Fisher formula is necessary for quantities.

◇ All entries that have a flag denoted by the "⋆" symbol are PCE price indexes available on Table 2.4.4U, which are aggregation of other subindexes with the same source data. There are overall 7 of such PCE price indexes and specific information on each of them is available in Table CA3.

◇ All entries that have a flag denoted by the "∘" symbol are PCE price indexes that have multiple source data. This is the case because they are aggregation of different price indexes that have different source data.

◇ Finally, all entries that have a flag denoted by the "†" symbol are PCE price indexes constructed by the BEA by different methodologies and for which we refere the reader to the BEA website for more information.

All the PCE price indexes listed in Table CA1 are also used by the Dallas Fed for the construction of the Trimmed Mean PCE index, with the exception of the prices with a flag denoted by the "•" or "⋆" symbol. Indeed, rather than using these price indexes, the Dallas Fed uses the subcomponents listed in Table CA2 and Table CA3.

Table CA1: Data and data sources

ID	Item	PCE Component	Price index source data	
1	6	New Autos		⋆
2	9	New Light Trucks	CPI New trucks	
3	10	Net Purchases of Used Motor Vehicles		⋆
4	19	Tires	CPI Tires	
5	20	Accessories & Parts	CPI Vehicle parts and equipment other than tires	
6	23	Furniture	CPI Furniture and bedding	
7	24	Clocks, lamps, lighting fixtures, and other household decorative items	CPI Clocks, lamps, and decorator items	
8	25	Carpets & Other Floor Coverings	CPI Floor covering	
9	26	Window Coverings	CPI Window coverings	
10		Major appliances		•
11	29	Small Electric Household Appliances	CPI Other appliances	
12	31	Dishes and Flatware	CPI Dishes and flatware	
13	32	Nonelectric Cookware & Tableware	CPI Nonelectric cookware and tableware	
14	34	Tools, Hardware & Supplies	CPI Tools, hardware, and supplies	
15	35	Outdoor Equipment & Supplies	CPI Outdoor equipment and supplies	
16	39	Televisions	CPI Televisions	
17	40	Other Video Equipment	CPI Other video equipment	
18	41	Audio Equipment	CPI Audio equipment	
19	43	Recording media		‡
20	45	Photographic Equipment	CPI Photographic equipment	
21	47	Information processing equipment	CPI Personal computers and peripheral equipment	‡
22	50	Sporting Equipment, Supplies, Guns & Ammunition	CPI Sports equipment	

• See Table CA2 ⋆ See Table CA3 ∘ See Table CA4 ‡ See Table CA4 † See the excel file downloadable from the BEA website at `https://www.bea.gov/media/3051`

Table CA1: Data and data sources (continuted)

ID	Item	PCE Component	Price index source data	
23	51	Sports & Recreational Vehicles		⋆
24	58	Recreational Books	CPI Recreational books	
25	59	Musical Instruments	CPI Music instruments and accessories	
26	62	Jewelry	CPI Jewelry	
27	63	Watches	CPI Watches	
28		Medical equipment and supplies		•
29	66	Corrective Eyeglasses & Contact Lenses	CPI Eyeglasses and eyecare	
30	67	Educational Books	CPI Educational books and supplies	
31	68	Luggage & Similar Personal Items	CPI Miscellaneous personal goods	
32		Telephone hardware, calculators, and other consumer items	CPI Telephone hardware, calculators, and other consumer items	
33	75	Cereals	CPI Cereals and cereal products	
34	76	Bakery Products	CPI Bakery products	
35	78	Beef and Veal	CPI Beef and veal	
36	79	Pork	CPI Pork	
37	80	Other Meats	CPI Other meats	
38	81	Poultry	CPI Poultry	
39	82	Fish and Seafood	CPI Fish and seafood	
40	84	Fresh Milk	CPI Milk	
41	85	Processed Dairy Products	BEA Composite index of various CPIs	†
42	86	Eggs	CPI Eggs	
43	87	Fats and Oils	CPI Fats and oils	
44	89	Fresh Fruit	CPI Fresh fruits	
45	90	Fresh Vegetables	CPI Fresh vegetables	
46	91	Processed Fruits & Vegetables	CPI Processed fruits and vegetables	
47	92	Sugar and Sweets	CPI Sugar and sweets	
48	93	Food Products, Not Elsewhere Classified	CPI unpublished detailed categories	†

• See Table CA2 ⋆ See Table CA3 ◦ See Table CA4 ‡ See Table CA4 † See the excel file downloadable from the BEA website at `https://www.bea.gov/media/3051`

Table CA1: Data and data sources (continued)

ID	Item	PCE Component	Price index source data	
49	95	Coffee, Tea & Other Beverage Materials	CPI Beverage materials including coffee and tea	
50	96	Mineral Waters, Soft Drinks & Vegetable Juices	CPI Juices and nonalcoholic drinks	
51	98	Spirits	CPI Distilled spirits at home	
52	99	Wine	CPI Wine at home	
53	100	Beer	CPI Beer, ale, and other malt beverages at home	
54	101	Food Produced & Consumed on Farms	BEA Composite of USDA prices received by farmers	†
55	104	Womens & Girls Clothing	CPI Womens and girls apparel	
56	105	Mens & Boys Clothing	CPI Mens and boys apparel	
57	106	Childrens & Infants Clothing	CPIs Infants and toddlers apparel	
58		Sewing machines, fabrics, and supplies		•
59	109	Standard Clothing Issued to Military Personnel	PPI Apparel	
60	110	Shoes & Other Footwear	CPI Footwear	
61	113	Gasoline & Other Motor Fuel	CPI Motor fuel	
62	114	Lubricants & Fluids	CPI Motor oil, coolant, and fluids	
63	116	Fuel Oil	CPI Fuel oil	
64	117	Other Fuels	CPI Propane, kerosene, and other firewood	
65	121	Prescription Drugs	CPI Prescription drugs	
66	122	Nonprescription Drugs	CPI Nonprescription drugs	
67	125	Games, Toys & Hobbies	CPI Toys	
68	126	Pets & Related Products	CPI Pets and pet products	
69	127	Flowers, Seeds & Potted Plants	CPI Indoor plants and flowers	
70	128	Film & Photographic Supplies	CPI Film and photographic supplies	
71	130	Household Cleaning Products	CPI Household cleaning products	
72	131	Household Paper Products	CPI Household paper products	
73	132	Household Linens	CPI Other linens	

• See Table CA2 ⋆ See Table CA3 ∘ See Table CA4 ‡ See Table CA4 † See the excel file downloadable from the BEA website at `https://www.bea.gov/media/3051`

Table CA1: Data and data sources (continued)

ID	Item	PCE Component	Price index source data	
74	134	Miscellaneous Household Products	CPI Miscellaneous household products	
75	135	Personal Care Products		⋆
76	139	Tobacco	CPI Tobacco and smoking products	
77	141	Newspapers & Periodicals	CPI Newspapers and magazines	
78	142	Stationery & Miscellaneous Printed Materials	CPI Stationery, stationery supplies, and gift wrap	
79		Rent of primary residence		•
80	156	Imputed Rental of Owner-Occupied Nonfarm Housing		⋆
81	159	Rental Value of Farm Dwellings	BEA extrapolation	†
82	163	Water Supply & Sewage Maintenance	CPI Water and sewage maintenance	
83	164	Garbage & Trash Collection	CPI Garbage and trash collection	
84	166	Electricity	CPI Electricity	
85	167	Natural Gas	CPI Utility (piped) gas service	
86	170	Physician Services	PPI Offices of physicians	
87	171	Dental Services	CPI Dental services	
88	172	Paramedical Services		◦
89		Hospitals		•
90	183	Nursing Homes	PPI Nursing care facilities	
91	188	Motor Vehicle Maintenance & Repair	CPI Motor vehicle maintenance and repair	
92	190	Other motor vehicle services		‡
93	197	Railway Transportation	CPI Intercity train fare	
94		Intercity bus fare		•
95		Intracity mass transit		•
96	203	Air Transportation	PPI Domestic scheduled passenger air transportation	

• See Table CA2 ⋆ See Table CA3 ◦ See Table CA4 ‡ See Table CA4 † See the excel file downloadable from the BEA website at `https://www.bea.gov/media/3051`

Table CA1: Data and data sources (continued)

ID	Item	PCE Component	Price index source data	
97	204	Water Transportation	CPI Ship fare	
98	207	Membership Clubs & Participant Sports Centers	CPI Club dues and fees for participant sports and group exercises	
99	208	Other recreation services		•
100	210	Admission to movies, theaters, and concerts		•
101	212	Spectator Sports	CPI Admission to sporting events	
102	214	Audio-Video, Photographic & Info Processing Services		◦
103	222	Gambling		⋆
104	227	Veterinary & Other Services for Pets	CPI Pet services including veterinary	
105	229	Maintenance & Repair of Rec Vehicles & Sports Equipment	CPI Sporting goods	
106		Food at employee sites and schools		•
107	237	Other Purchased Meals		◦
108	241	Alcohol in Purchased Meals	CPI Alcoholic beverages away from home	
109	246	Hotels and Motels	CPI Other lodging away from home including hotels and motels	
110	247	Housing at Schools	CPI Housing at school, excluding board	
111	251	Commercial Banks	BEA extrapolation	†
112	252	Other Depository Instns & Regulated Invest Companies	BEA annual composite index.	†
113	253	Pension Funds	BEA input cost index	†
114	254	Financial Service Charges, Fees & Commissions		◦
115	267	Life Insurance	BEA input cost index	†
116	268	Net Household Insurance	PPI Homeowners insurance	
117	271	Net Health Insurance		◦
118	275	Net Motor Vehicle & Other Transportation Insurance	PPI Private passenger auto insurance	†
119	277	Communication		◦

• See Table CA2 ⋆ See Table CA3 ◦ See Table CA4 ‡ See Table CA4 † See the excel file downloadable from the BEA website at `https://www.bea.gov/media/3051`

Table CA1: Data and data sources (continued)

ID	Item	PCE Component	Price index source data	
120	287	Higher Education		⋆
121	291	Elementary & Secondary Schools	CPI Elementary and high school tuition and fixed fees	
122	292	Day Care & Nursery Schools	CPI Day care and nursery school	
123	293	Commercial & Vocational Schools	CPI Technical and business school tuition and fees	
124		Legal services		•
125	297	Tax Preparation & Other Related Services	CPI Tax preparation and other accounting fees	
126	298	Employment Agency Services	PPI Employment placement services	
127	299	Other Personal Business Services	CPI Miscellaneous personal services	
128	300	Labor Organization Dues	BEA input cost index	†
129	302	Funeral & Burial Services	CPI Funeral expenses	
130	305	Hairdressing Salons & Personal Grooming Estab	CPI Haircuts and other personal care services	
131		Apparel services other than laundry and drycleaning		•
132	308	Laundry & Dry Cleaning Services	CPI Laundry and drycleaning services	
133	312	Child Care	CPI Child care and nursery school	
134	313	Social Assistance	BEA input cost index	†
135	320	Social Advocacy & Civic & Social Organizations	BEA input cost index	†
136	321	Religious Organizations Services to Households	BEA input cost index	†
137	322	Foundations and grantmaking and giving services to households	BEA input cost index	†
138	324	Domestic Services	BEA Composite index of various CPIs	†
139	325	Moving, Storage & Freight Services	CPI Moving, storage, and freight expenses	
140		Repair of household items		•
141	328	Other Household Services	CPI Household operations	
142	338	Final Consumption Expenditures of Nonprofit Institutions Serving Households	BEA input cost index	†

• See Table CA2 ⋆ See Table CA3 ◦ See Table CA4 ‡ See Table CA4 † See the excel file downloadable from the BEA website at `https://www.bea.gov/media/3051`

Table CA2: Notes to Table 1 for entries with • symbol

ID	Note
10	The price index for "Major Appliances" is the aggregation of the PCE price index for "Major Household Appliances" (Item 28, Haver JCDFKKM) and the PCE price index for "Tenant Landlord Durables" (Item 155, Haver JCSHTDM@USNA), which are both constructed out of the CPI "Major Appliances"
28	The price index for "Medical equipment and supplies" is the aggregation of the PCE price index for "Therapeutic Medical Equipment" (Item 65, Haver JCDOOTM) and the PCE price index for "Other Medical Products" (Item 123, Haver JCNODOM), which are both constructed out of the CPI "Medical equipment and supplies".
58	The price index for "Sewing machines, fabrics, and supplies" is the aggregation of the PCE price index for "Clothing Materials" (Item 108, Haver JCNLOLM) and the PCE price index for "Sewing Items Price" (Item 133, Haver JCNOLSM), which are both constructed out of the CPI "Sewing machines, fabrics, and supplies".
79	The price index for "Rent of primary residence" is the aggregation of the PCE price index for "Tenant-Occupied Mobile Homes Price Index" (Item 153, Haver JCSHTBM), the PCE price index for "Tenant-Occupied Stationary Homes" (Item 154, Haver JCSHTSM), and the PCE price index for "Group Housing" (Item 160, Haver JCSHOM), which are all constructed out of the CPI "Rent of primary residence".
80	The PCE price index for "Imputed Rental of Owner-Occupied Nonfarm Housing" has two subcomponents, (1) "Owner-Occupied Mobile Homes" (Item 157, Haver JCSHRBM), (2) "Owner-Occupied Stationary Homes" (Item 158, Haver JCSHRSM), which are both constructed using the "CPI Owners' equivalent rent of primary residence". In the disaggregation used by the Dallas Fed instead of the index for "Imputed Rental of Owner-Occupied Nonfarm Housing" the single components are included.
89	The price index for "Hospitals" is the aggregation of the PCE price index for "Nonprofit Hospitals' Services to Households" (Item 180, Haver JCSMPNM), the PCE price index for "Proprietary Hospitals" (Item 181, Haver JCSMPPM), and the PCE price index for "Govt Hospitals Price" (Item 182, Haver JCSMPPM), which are all constructed out of the PPI "Hospitals".

Table CA2: Notes to Table 1 for entries with ● symbol (continued)

ID	Note
98	The price index for "Intercity bus fare" is the aggregation of the PCE price index for "Intercity Buses" (ID 199, Haver JCSTIBM) and the PCE price index for "Other Road Transportation Service" (ID 202, Haver JCSTIOM), which are both constructed out of the "CPI Intercity bus fare".
94	The price index for "Intracity mass transit" is the aggregation of the PCE price index for "Taxicabs" (Item 200, Haver JCSTLBM) and the PCE price index for "Intracity Mass Transit" (Item 201, Haver JCSTLTM), which are both constructed out of the "CPI Intercity bus fare".
95	The price index for "Other recreation services" is the aggregation of the PCE price index for "Amusement Parks, Campgrounds & Related Recreational Services" (Item 208, Haver JCSRCPM) and the PCE price index for "Package Tours" (Item 228, Haver JCSRKM), which are both constructed out of the CPI "Other recreation services".
100	The price index for "Admission to movies, theaters, and concerts" is the aggregation of the PCE price index for "Motion Picture Theaters" (Item 210, Haver JCSRSPM), the PCE price index for "Live Entertainment, ex Sports" (Item 211, Haver JCSRSTM), and the PCE price index for "Museums & Libraries" (Item 213, Haver JCSOSLM), which are all constructed out of the CPI "Admission to movies, theaters, and concerts".
206	The price index for "Food at employee sites and schools" is the aggregation of the PCE price index for "Elementary & Secondary School Lunches" (Item 235, Haver JCSFPGM), the PCE price index for "Higher Education School Lunches" (Item 236, Haver JCSFPUM), the PCE price index for "Food Supplied to Civilians" (Item 243, Haver JCSFEVM), and the PCE price index for " Food Supplied to Military" (Item 244, Haver JCSFEAM), which are all constructed out of the CPI "Food at employee sites and schools".
124	The price index for "Legal services" is the aggregation of PCE price index for "Legal Services" (Item 295, Haver JCSOBLM) and PCE price index for "Prof Assn Dues" (Item 301, Haver JCSBOPM), which are both constructed out of the CPI "Legal services".
131	The price index for "Apparel services other than laundry and drycleaning" is the aggregation of the PCE price index for "Miscellaneous Personal Care Services" (Item 308, Haver JCSOPOM), the PCE price index for "Clothing Repair, Rental& Alterations" (Item 309, Haver JCSOPRM), and the PCE price index for "Repair & Hire of Footwear" (Item 310, Haver JCSOPSM), which are all constructed out of the CPI " Apparel services other than laundry and drycleaning".

Table CA3: Notes to Table 1 for entries with ⋆ symbol

ID	Note
1	The PCE price index for "New Autos" has two subcomponents, (1) "New Domestic Autos" (Item 7, Haver JCDMNDM), and (2) "New Foreign Autos" (Item 8, Haver JCDMNFM), which are both constructed using the "CPI New cars". In the disaggregation used by the Dallas Fed instead of the index for "New Autos" the single components are included.
3	The PCE price index for "Net purchases of used motor vehicles" has two subcomponents, (1) "Used autos" (Item 11, Haver JCDMUM), which in its turn has three subcomponents (1a) "Net transactions in used autos" (Item 12, Haver JCDMUNM), (1b) "Used auto margin" (Item 13, Haver JCDMUGM), and (1c) "Employee reimbursement" (Item 14, Haver JCDMURM); and (2) "Used light trucks" (Item 15, Haver JCDMTUM), which in its turn has two subcomponents (2a) "Net transactions in used truck" (Item 16, Haver JCDMTUNM), and (2b) "Used truck margin" (Item 17, Haver JCDMTUGM). Item 12 and 16 are constructed out of the ("CPI Used cars and trucks"), and similarly Item 13 and 17 ("PPI Used vehicle sales at new car dealers"), whereas Item 14 is sourced from the "CPI Car and truck rental". In the disaggregation used by the Dallas Fed instead of the index for "SNet purchases of used motor vehicles" the two subcomponent components (Item 11 and 15) are included.
23	The PCE price index for "Sports and recreational vehicles" has three subcomponents, (1) "Motorcycles" (Item 52, Haver JCDOWLM), (2) "Bicycles and accessories" (Item 53, Haver JCDOWBM), and (3) "Pleasure boats, aircraft, and other recreational vehicles" (Item 54, Haver JCDBBM), which in its turn can be further decomposed in (3a) "Pleasure boats" (Item 55, Haver JCDBBBM), (3b) "Pleasure aircraft" (Item 56, Haver JCDBBPM), and (3c) "Other recreational vehicles" (Item 57, Haver JCDBBOM). The source of all these components is the same ("CPI Sports vehicles including bicycles"), the only exception being the PCE price index for "Motorcycles" that is sourced from the "CPI New motorcycles". In the disaggregation used by the Dallas Fed instead of the index for "Sports and recreational vehicles" the single components (Item 52, 53, 55, 56, and 57) are included.
75	The PCE price index for "Personal Care Products" has three subcomponents, (1) "Hair/Dental/Shave/Miscellaneous Pers Care Prods ex Elec Prod" (Item 136, Haver JCNOPPM), (2) "Cosmetic/Perfumes/Bath/Nail Preparatns & Implements" (Item 137, Haver JCNOPCM), and (3) "Elec Appliances for Personal Care" (Item 138, Haver JCNOPEM). Item 136 and 138 are both constructed out of the "CPI Hair, dental, shaving, and miscellaneous personal care products", while Item 137 is constructed out of the "CPI Cosmetics/perfumes/bath/nail preparations and implements". In the disaggregation used by the Dallas Fed instead of the index for "Personal Care Products" the single components are included.

Table CA3: Notes to Table 1 for entries with ⋆ symbol (continued)

ID	Note
103	The PCE price index for “Gambling” has three subcomponents, (1) “Lotteries Price Index” (Item 221, Haver JCSROGM), (2) “Casino Gambling” (Item 222, Haver JCSROLM), and “Pari-Mutuel Net Receipts” (Item 223, Haver JCSROBM), which are both constructed using the “CPI All Items”. In the disaggregation used by the Dallas Fed instead of the index for “Gambling” the single components are included.
120	The PCE price index for “Higher Education” has two subcomponents, (1) “Proprietary & Public Higher Education” (Item 286, Haver JCSOEUPM), and (2) “Nonprofit Pvt Higher Education Services to Households” (Item 287, Haver JCSOEUNM), which are both constructed using the “CPI College tuition and fees”. In the disaggregation used by the Dallas Fed instead of the index for “Higher Education” the single components are included.
140	The price index for “Repair of household items” is the aggregation of PCE price index for “Repair of Furniture, Furnishings & Floor Coverings” (Item 326, Haver JCSLORM) and the PCE price index for “Repair of Household Appliances” (Item 327, Haver JCSLOPM), which are both constructed out of the “Repair of household items”

Table CA4: Notes to Table 1 for entries with ○ symbol

ID	Note
86	The PCE price index for "Paramedical services" has three subcomponents, (1) "Home health care" (ID 173, Haver JCSMOAM), which is constructed out of the PPI "Home health care services"; (2) "Medical laboratories" (ID 174, Haver JCSMOLM), which is constructed out of the by the BEA as a composite index of fixed-weighted PPIs for "Medical laboratories" and for "Diagnostic imaging centers"; and (3) "Other professional medical services" (ID 175, Haver JCSMOLM), which in its turn has two subcategories both constructed out of the CPI "Services of other medical professionals". Note that also in the disaggregation used by the Dallas Fed index for "Paramedical services", rather then the components, is included.
102	The PCE price index for "Audio-video, photographic, and information processing equipment services" has five subcomponents, (1) "Cable & Satellite Television & Radio Services" (Item 215, Haver JCSROTM), which is constructed out of the CPI "Cable and satellite TV and radio services"; (2) "Photo Processing" (Item 216, Haver JCSRODM), which is constructed out of the CPI "Film processing"; (3) "Photo Studios" (Item 217, Haver JCSROUM), which is constructed out of the CPI "Photographer fees"; (4) "Repair of Audio-Visual, Photo & Info Process Equipment" (Item 218, Haver JCSREEM), which is constructed out of the CPI "Video and audio"; and (5) "Video Media Rental Price" (Item 219, Haver JCSROYM), which is constructed out of the CPI "Rental of video or audio discs and other media". Note that also in the disaggregation used by the Dallas Fed index for "Audio-video, photographic, and information processing equipment services", rather then the components, is included.
107	The PCE price index for "Other Purchased Meals" has three subcomponents, (1) "Meals at Limited Service Eating Places" (Item 236, Haver JCSFPLM), which is constructed out of the CPI "Limited service meals and snacks"; (2) "Meals at Other Eating Places" (Item 237, Haver JCSFPEM) and (3) "Meals at Drinking Places" (Item 238, Haver JCSFPDM), which are both constructed out of the CPI "Full service meals and snacks". Note that also in the disaggregation used by the Dallas Fed index for "Other Purchased Meals", rather then the components, is included.

Table CA4: Notes to Table 1 for entries with ∘ symbol (continued)

ID	Note
114	The PCE price index for “Financial service charges, fees, and commissions” has four subcomponents: (1) “Financial service charges and fees” (Item 253, Haver JCSNFVM), which is constructed out of the CPI “Checking account and other bank services”; (2) “Securities commissions” (Item 254, Haver JCSNFSM); (3) “Portfolio management and investment advice services” (Item 262, Haver JCSNFPM), which is constructed as a fixed weighted average of the PPI “Portfolio Management” and the PPI “Investment advice”; and (4) “Trust, fiduciary, and custody acitivities” (Item 263, Haver JCSNFTM), which is constructed out of the PPI “Commercial bank trust services”. The subcomponent “Securities commissions” has three subcomponents: (2.1) “Direct commissions” (Item 255, Haver JCSNFSDM), which in its turn has two subcomponents (2.1.1) “Exchange-listed equities” (Item 256, Haver JCSNFSEM), which is constructed out of the PPI “Brokerage services, equities and ETFs”, and (2.1.2) “Other direct commissions” (Item 257, Haver JCSNFSOM), which is constructed out of the PPI “Brokerage services, all other securities”; (2.2) “Indirect commissions” (Item 258, Haver JCSNFIM), which in its turn has two subcomponents (2.2.1) “Over-the-counter equity securities” (Item 259, Haver JCSNFIVM), which is constructed out of the PPI “Dealer transactions, equities securities”, and (2.2.2) “Other imputed commissions” (Item 260, Haver JCSNFIOM), which is constructed out of the “Dealer transactions, debt securities and all other trading”; and (2.3) “Mutual fund sales charges” (Item 261, Haver JCSBKFM), which is constructed by the BEA as an Implicit price index. Note that also in the disaggregation used by the Dallas Fed index for “Financial service charges, fees, and commissions”, rather then the components, is included.
117	The PCE price index for “Net Health Insurance” has three subcomponents: (1) “Health Insurance: Medical Care & Hospitalization ” (Item 270, Haver JCSMHIM), which is constructed out of the PPI “Homeowner’s insurance”; (2) “Health Insurance: Income Loss” (Item 271, Haver JCSMIIM), which is constructed out of the CPI “All items”; and (3) “Health Insurance: Workers’ Compensation”, which is constructed out of the PPI “Worker’s compensation insurance”. See also BEA. Note that also in the disaggregation used by the Dallas Fed index for “Net Health Insurance”, rather then the components, is included.

Table CA4: Notes to Table 1 for entries with ‡ symbol

ID	Note
19	"Recording media" (42) is the aggregate of "Audio discs, tapes, vinyl, and permanent digital downloads" (43) and "Video discs, tapes, and permanent digital downloads" (44). We took the aggregate because 44 is available only starting in 1976. In the disaggregation used by the Dallas Fed instead of the aggregate index are the three subindexes.
21	"Information processing equipment" (46) is the aggregate of "Personal computers/tablets and peripheral equipment" (47), "Computer software and accessories" (48), and "Calculators, typewriters, and other information processing equipment" (49). We took the aggregate because 47 and 48 are available only starting in 1976. In the disaggregation used by the Dallas Fed instead of the aggregate index are the three subindexes.
92	"Other motor vehicle services" (191) is the aggregate of "Motor vehicle leasing" (192), "Motor vehicle rental" (195), and "Parking fees and tolls" (196). We took the aggregate because 192 is available only starting in 1976. In the disaggregation used by the Dallas Fed instead of the aggregate index are the three subindexes.