Supplemental Appendix to
“A Measure of Core Wage Inflation”

by Martín Almuzara[†], Richard Audoly[‡], and Davide Melcangi[§]

A Definition of CPS partitions

We partition workers in the CPS using the same definitions as the Atlanta Fed WGT [Atlanta Fed, 2023].

Industries (7 groups) Construction and Mining, Education and Health, Finance and Business Services, Leisure and Hospitality, Manufacturing, Public Administration, and Trade and Transportation.

Occupations (3 groups) high–skill (Managers, Professionals, Technicians), middle–skill (Office and Administration, Operators, Production, Sales), and low–skill (Food Preparation and Serving, Cleaning, individual Care Services, Protective Services).

Race (2 groups) White and Nonwhite.

Education (3 groups) High school or less, Associates degree, and Bachelor degree or higher.

Age (3 groups) 16–24 years old, 25–54, and 55+.

Gender (2 groups) Male and Female.

Wage quartiles (4 groups) The quartiles are based on the average between workers' current hourly wage and their wage 12 months prior (when their wages are last recorded).

[†]Federal Reserve Bank of New York. Email: martin.almuzara@ny.frb.org. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

[‡]Federal Reserve Bank of New York. Email: richard.audoly@ny.frb.org.

[§]Federal Reserve Bank of New York. Email: davide.melcangi@ny.frb.org.

Region (9 groups) The nine Census Divisions: New England (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont), Mid-Atlantic (New Jersey, New York, Pennsylvania), East North Central (Illinois, Indiana, Michigan, Ohio, Wisconsin), West North Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota), South Atlantic (Delaware, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, District of Columbia, West Virginia), East South Central (Alabama, Kentucky, Mississippi, Tennessee), West South Central (Arkansas, Louisiana, Oklahoma, Texas), Mountain (Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming), Pacific (Alaska, California, Hawaii, Oregon, Washington).

B Details of model and estimation approach

Recall our notation for the data $w_t = \{w_{it}\}_{i=1}^n$, the time-invariant parameters

$$\begin{aligned} \theta &= \left(\{\theta_{c\ell}\}_{1\leq \ell \leq p}, \{\theta_{i\ell}\}_{1\leq \ell\leq q, 1\leq i\leq n}\right), \\ \gamma &= \left(\{\gamma_{\alpha,m,i}\}_{m=\tau,\varepsilon,1\leq i\leq n}, \{\gamma_{\sigma,m,j}\}_{m=\Delta\tau,\varepsilon,\, j=c,1,\dots,n}\right), \end{aligned}$$

the time-varying parameters

$$\lambda_t = \left(\{\alpha_{\tau,it}, \alpha_{\varepsilon,it}\}_{i=1}^n, \sigma_{\Delta\tau,ct}, \{\sigma_{\Delta\tau,it}\}_{i=1}^n, \sigma_{\varepsilon,ct}, \{\sigma_{\varepsilon,it}\}_{i=1}^n\right),$$

and the latent components

$$\xi_t = (\tau_{ct}, \{\tau_{it}\}_{i=1}^n, \varepsilon_{ct}, \{\varepsilon_{it}\}_{i=1}^n).$$

To conduct Bayesian inference, we begin by formulating a prior on (θ, γ).

Choice of priors The MA coefficients θ are prior independent of each other with $\theta_{j\ell} \sim N(0, v_\ell^2)$ for $j = c, 1, \dots, n$. That is, we shrink the model towards one with white noise transitory errors and the strength of the shrinkage is determined by the choice of v_ℓ. In our baseline model we set $p = 0$ and $q = 3$, and we put higher penalties on the more distant lags as in the Minnesota prior of Doan, Litterman, and Sims [1983]. We achieve that by setting $v_\ell = 1/(10\ell^2)$.

The standard deviations γ control the amount of time-variation in loadings and volatilities. Unless they are small, the model may be excessively flexible causing overfitting. Our approach is to put a reasonably tight prior centered around small values to shrink the model towards no time-variation in the parameters. Specifically we use independent inverse gamma priors of the form $\gamma_{k,m,j}^2 \sim \Gamma^{-1}(d_k/2, 2/(d_k s_k^2))$ for $k = \alpha, \sigma$. The location parameters are set to $s_\alpha^2 = 0.0001$ and $s_\sigma^2 = 0.001$, and the degree-of-freedom hyperparameters are set to $d_\alpha = d_\sigma = 60$.

Estimation and filtering Inference about parameters and latent variables is implemented via Gibbs sampling. This is a type of Markov Chain Monte Carlo (MCMC) algorithm suitable to approximate the joint posterior distribution of parameters and latent variables by simulation in state-space models.

The Gibbs sampler constructs a Markov Chain $\left\{\theta^{(s)}, \gamma^{(s)}, \{\lambda_t^{(s)}\}, \{\xi_t^{(s)}\}\right\}_{s=1}^S$ having as invariant distribution the posterior

$$P\left(\theta, \gamma, \{\lambda_t\}, \{\xi_t\} \middle| \{w_t\}\right).$$

This allows us to estimate the posterior of our objects of interest, e.g.

$$P\left(\{\tilde{\tau}_t, \{\tilde{\tau}_{it}\}_{i=1}^n\} \middle| \{w_t\}\right),$$

using the draws $\left\{\theta^{(s)}, \gamma^{(s)}, \{\lambda_t^{(s)}\}, \{\xi_t^{(s)}\}\right\}_{s=1}^S$ to form $\left\{\{\tilde{\tau}_t^{(s)}, \{\tilde{\tau}_{it}^{(s)}\}_{i=1}^n\}\right\}_{s=1}^S$ and taking the simulation frequencies of the objects as estimates of posterior probabilities. If the Markov chain converges (in a suitable sense) and S is large, the approximation error will be small.

One advantage of the Bayesian approach is that posterior calculations already integrate both the sampling uncertainty from parameter estimation and the signal-extraction uncertainty about the latent components. When reporting the path over time of a latent time series in our empirical analysis, we use credible intervals with fixed credibility level pointwise in t.[1]

An alternative would be to estimate (θ, γ) by maximum likelihood. It is straightforward to modify our MCMC algorithm to approximate the maximum likelihood estimator by stochastic EM — simply replace the posterior updates of θ and γ by the solutions to the corresponding complete-data score equations. However, inferences about latent variables (and, in particular, about our objects of interest) that arise from that procedure would not necessarily account for the estimation uncertainty in (θ, γ).

[1] It is conceptually straightforward and computationally feasible to compute pathwise credible regions along the lines of, e.g., Inoue and Kilian [2016].

Gibbs sampling Our algorithm follows Stock and Watson [2016] who build on the method proposed by Del Negro and Otrok [2008] to estimate dynamic factor models with time-varying loadings and volatilities.

Relative to Del Negro and Otrok [2008], Stock and Watson [2016] incorporate outliers in the transitory shocks. Compared to Stock and Watson [2016], we allow for temporal aggregation in the persistent components and for MA dynamics in the transitory components. For simplicity, we discuss estimation of a model without outliers.[2] We find only a negligible role for them in the data we analyze.

The Gibbs sampler exploits the fact that with a careful grouping of parameters and latent variables, the conditional distributions of each block given the rest can be simulated by well-known algorithms. In our model, there are three big blocks with many sub-blocks, namely:

(A) $\boldsymbol{P(\{\xi_t\}|\{\lambda_t\}, \theta, \gamma, \{w_t\})}$. Conditional on time-varying parameters $\{\lambda_t\}$ and the MA coefficients θ, the data w_t and the latent variables ξ_t are related by a linear state-space model with time-varying matrices. We apply the simulation smoother algorithm proposed by Durbin and Koopman [2002] to efficiently sample $\{\xi_t\}$.

   To accommodate the MA dynamics of the common and sector-specific transitory errors, we include $\varepsilon_{ct}, \varepsilon_{c,t-1}, \dots, \varepsilon_{c,t-p+1}, \{\varepsilon_{it}, \varepsilon_{i,t-1}, \dots, \varepsilon_{i,t-q+1}\}_{i=1}^{n}$ as additional state variables.

(B) $\boldsymbol{P(\{\lambda_t\}|\{\xi_t\}, \theta, \gamma, \{w_t\})}$. This can be further partitioned into the following blocks:

   (i) $\boldsymbol{P\left(\left\{\{\alpha_{\tau,it}, \alpha_{\varepsilon,it}\}_{i=1}^{n}\right\}\middle|\left\{\{w_{it}, \tau_{it}\}_{i=1}^{n}\right\}, \{\tau_{ct}, \varepsilon_{ct}\}, \left\{\{\sigma_{\varepsilon,it}\}_{i=1}^{n}\right\}, \{\gamma_{\alpha,\tau,i}, \gamma_{\alpha,\varepsilon,i}\}_{i=1}^{n}\right)}$. It is the result of a multivariate regression with time-varying coefficients and MA error terms. It can be dealt with using linear state-space techniques. Thus, we apply the simulation smoothing algorithm of Durbin and Koopman [2002] to the corresponding state-space representation in order to sample $\{\{\alpha_{\tau,it}, \alpha_{\varepsilon,it}\}_{i=1}^{N}\}$.

   (ii) $\boldsymbol{P\left(\{\sigma_{m,jt}\}\middle|\{m_{jt}\}, \gamma_{\sigma,m,j}\right)}$ for $m = \Delta\tau, \varepsilon$ and $j = c, 1, \dots, n$. Given $\gamma_{\sigma,m,j}$ and m_{jt}, $\sigma_{m,jt}$ follows a stochastic-volatility model with observation equation $\ln m_{jt}^2 = \ln \sigma_{m,jt}^2 + \ln \eta_{m,jt}^2$ and transition equation $\Delta \ln \sigma_{m,jt}^2 = \gamma_{\sigma,m,j}\nu_{\sigma,m,jt}$. We then use the algorithm proposed in Kim, Shephard, and Chib [1998] and Omori, Chib, Shephard, and Nakajima [2007] that consists of approximating the log-χ_1^2 distribution of $\ln \eta_{\sigma,m,jt}^2$ with a 10-component normal mixture and applying linear state-space techniques to that approximation.

[2] The outliers in Stock and Watson [2016] are introduced by assuming $\eta_{\varepsilon,jt} = s_{jt} \times \tilde{\eta}_{\varepsilon,jt}$ for $j = c, 1, \dots, n$ where $s_{jt} = 1$ with probability p_j and $s_{jt} \sim U(1, 10)$ with probability $1 - p_j$ while $\tilde{\eta}_{\varepsilon,jt} \sim N(0, 1)$.

(C) $\boldsymbol{P(\theta, \gamma|\{\xi_t\}, \{\lambda_t\}, \{w_t\})}$. This can also be partitioned into subblocks:

(i) $\boldsymbol{P\left(\gamma_{\alpha,m,i}\middle|\{\Delta\alpha_{m,it}\}\right)}$ for $m = \tau, \varepsilon$. We draw the reciprocal of the square root of a gamma random variable with $d_\alpha + T$ degrees of freedom and mean $(d_\alpha s_\alpha^2 + \sum_{t=1}^T \Delta\alpha_{m,it})/(d_\alpha + T)$ for $m = \tau, \varepsilon$.

(ii) $\boldsymbol{P\left(\gamma_{\sigma,m,j}\middle|\{\Delta \ln \sigma^2_{m,jt}\}\right)}$ for $m = \Delta\tau, \varepsilon$ and $j = c, 1, \ldots, n$. We draw the reciprocal of the square root of a gamma random variable with $d_\sigma + T$ degrees of freedom and mean $(d_\sigma s_\sigma^2 + \sum_{t=1}^T \Delta \ln \sigma^2_{m,jt})/(d_\sigma + T)$ for $m = \tau, \varepsilon$ and $j = c, 1, \ldots, n$.

(iii) $\boldsymbol{P\left(\theta_j\middle|\{\varepsilon_{jt}\}, \{\sigma_{\varepsilon,jt}\}\right)}$ for $j = c, 1, \ldots, n$ where $\theta_c = (\theta_{c1}, \ldots, \theta_{cp})'$ and $\theta_i = (\theta_{i1}, \ldots, \theta_{iq})'$ for $i = 1, \ldots, n$. This problem can be treated separately for each j. We do the derivation for $j = i = 1, \ldots, n$ (the case $j = c$ is identical except that p should take the place of q). Define $v_{it} = \sigma_{\varepsilon,it}\eta_{\varepsilon,it}$. Conditioning on q initial observations $\varepsilon_{i0}, \ldots, \varepsilon_{iq-1}$ we obtain the likelihood term

$$\begin{aligned}
P\left(\{\varepsilon_{it}\}_{t=1}^T\middle|\{\theta_{i\ell}\}_{\ell=1}^q, \{\varepsilon_{i1-\ell}\}_{\ell=1}^q, \{\sigma_{\varepsilon,it}\}\right) &= \prod_{t=1}^T P\left(\varepsilon_{it}\middle|\{\theta_{i\ell}\}_{\ell=1}^q, \{\varepsilon_{it-\ell}\}_{\ell=1}^{t-1+q}, \{\sigma_{\varepsilon,it}\}\right) \\
&= \prod_{t=1}^T P\left(\varepsilon_{it}\middle|\{\theta_{i\ell}\}_{\ell=1}^q, \{v_{it-\ell}\}_{\ell=1}^q, \sigma_{\varepsilon,it}\right) \\
&= \prod_{t=1}^T \frac{1}{\sigma_{\varepsilon,it}}\phi\left(\frac{\varepsilon_{it} - \left(\sum_{\ell=1}^q \theta_{i\ell}v_{it-\ell}\right)}{\sigma_{\varepsilon,it}}\right)
\end{aligned}$$

where ϕ is the standard normal density. This is the likelihood from a regression of ε_{it} on $(v_{it-1}, \ldots, v_{it-q})'$ with heteroskedastic Gaussian errors or, equivalently, (up to a scaling constant) from a regression of $y_{it} = \varepsilon_{it}/\sigma_{\varepsilon,it}$ on $x_{it} = (v_{it-1}, \ldots, v_{it-q})'/\sigma_{\varepsilon,it}$ with i.i.d. $N(0, 1)$ errors. Since the prior is $(\theta_{i1}, \ldots, \theta_{iq})' \sim N(0_{Q\times 1}, V_\theta)$ where the variance is $V_\theta = \operatorname{diag}\left(v_1, \ldots, v_q\right)$, the posterior follows from the usual regression formula

$$\{\theta_{i\ell}\}_{\ell=1}^q\middle|\{\varepsilon_{it}\}, \{\sigma_{\varepsilon,it}\} \sim N\left(\left[V_\theta^{-1} + \sum_{t=1}^T x_{it}x_{it}'\right]^{-1}\sum_{t=1}^T x_{it}y_{it}, \left[V_\theta^{-1} + \sum_{t=1}^T x_{it}x_{it}'\right]^{-1}\right).$$

Conditioning on the initial observations $\{\varepsilon_{i1}, \ldots, \varepsilon_{i1-q}\}$ has at most a small effect when T is large. As an alternative, we can include $\{v_{it}\}$ as state variables in step (A).

Implementation, numerical accuracy and tests We do $S = 12,000$ draws retaining one every two after burning the first $6,000$. The result is a chain for the parameters (θ, γ) with low enough autocorrelations that the posterior expectations have negligible Monte Carlo standard errors. We also monitor the behavior of the latent variables and stochastic volatilities, the paths of which seem to stabilize within a small region well before the burn-in period ends. We ran the posterior simulator test suggested by Geweke [2004] and an extensive Monte Carlo simulation study, finding no indication against our implementation.

C Monte Carlo simulations

To assess how reliable our estimates of Core Wage Inflation are we conduct Monte Carlo simulations. We are interested in two questions. First, we ask whether our approach can accurately trace out the persistence pattern in monthly wage inflation from observations on 12-month wage growth rates, that is, whether we can successfully disentangle the temporal aggregation in the data. Second, we ask whether our approach has any bias — any tendency to over or understate the role of common and idiosyncratic components. To give a preview, the findings in this appendix validate the performance of the model in disentangling temporal aggregation and show that our method is not biased towards attributing an excessive role to the common component.

We simulate $n_{\text{MC}} = 200$ samples of size $N = 7$ and $T = 300$ (N, T are chosen to be similar to our sample of wage growth by industry) from the following data generating process (DGP):

$$\begin{aligned}
w_{it} &= \frac{1}{12}\sum_{\ell=1}^{12} \tilde{\tau}_{i,t+1-\ell} + \tilde{\varepsilon}_{it}, \quad i = 1,\dots,N, \quad t = 1,\dots,T, \\
\tilde{\tau}_{it} &= \alpha_{\tau,i}\tau_{ct} + \tau_{it}, \\
\tilde{\varepsilon}_{it} &= \alpha_{\varepsilon,i}\varepsilon_{ct} + \varepsilon_{it} + \theta_i \varepsilon_{i,t-1}, \\
\Delta\tau_{ct} &\overset{iid}{\sim} N(0,1), \quad \Delta\tau_{it} \overset{iid}{\sim} N(0,\sigma^2_{\Delta\tau,i}), \\
\varepsilon_{ct} &\overset{iid}{\sim} N(0,1), \quad \varepsilon_{it} \overset{iid}{\sim} N(0,\sigma^2_{\varepsilon,i}).
\end{aligned}$$

We abstract from time-variation in loadings and volatilities in the DGP and we treat sectors symmetrically, setting $\alpha_{\tau,i} = \alpha_\tau$, $\alpha_{\varepsilon,i} = \alpha_\varepsilon$, $\sigma_{\Delta\tau,i} = \sigma_{\Delta\tau}$, $\sigma_{\varepsilon,i} = \sigma_\varepsilon$. Nonetheless, we conduct estimation in each sample allowing for both time-varying parameters and heterogeneity and we use exactly the same priors we adopted in the empirical analysis of the paper.

We calibrate $\alpha_\varepsilon = 0.02$, $\sigma_{\Delta\tau} = 0.2$ and $\sigma_\varepsilon = 0.85$ (we also set $\theta_i = \theta = 0.1$) using the averages across sectors and over time of the estimates we obtained in our sample of wage growth by industry. For α_τ we try two different values, namely: $\alpha_\tau \in \{0, 0.3\}$. Since the variance of $\Delta\tau_c$ is unity, the parameter α_τ controls the importance of the common component in driving the trajectory of the total trend of each sector. The value $\alpha_\tau = 0.3$ is the average across industries and periods we found in our sample. The value $\alpha_\tau = 0$ represents a case where the common component is zero. We consider this extreme case to assess whether our method would spuriously recover a common component that does not exist.

In each sample, we run our estimation algorithm and we recover the posterior p-quantile of the total trend $\tilde{\tau}_t = N^{-1} \sum_{i=1}^{N} \tilde{\tau}_{it}$, its common part $\tilde{\tau}_{Ct} = N^{-1} \sum_{i=1}^{N} \alpha_{\tau,i} \tau_{ct}$ and its idiosyncratic part $\tilde{\tau}_{It} = N^{-1} \sum_{i=1}^{N} \tau_{it}$, that we denote $\tilde{\tau}_t(p)$, $\tilde{\tau}_{Ct}(p)$ and $\tilde{\tau}_{It}(p)$. Note that we are assuming all sectors have the same employment share and that these are constant over time, i.e., we set $s_{it} = N^{-1}$.

We use our Monte Carlo simulation to estimate the bias of the posterior median (seen as a point estimate of the latent variables) and the frequentist coverage rates of credible intervals based on the posterior. For $\tilde{\tau}_t$, for example, we have

$$\text{bias}_t = \mathbb{E}\left[\tilde{\tau}_t\left(\frac{1}{2}\right) - \tilde{\tau}_t\right],$$
$$\text{cov}_t = \mathbb{P}\left[\tilde{\tau}_t\left(\frac{\beta}{2}\right) \leq \tilde{\tau}_t \leq \tilde{\tau}_t\left(1 - \frac{\beta}{2}\right)\right]$$

where expectations and probabilities are taken with respect to repeated sampling from the DGP (and they are estimated by averaging over the n_{MC} Monte Carlo samples). A good estimator of the trend will deliver $\text{bias}_t \approx 0$ and $\text{cov}_t \approx 1 - \beta$. We can similarly define bias and coverage rates for $\tilde{\tau}_{Ct}$ and $\tilde{\tau}_{It}$.

One detail is that the location of $\tilde{\tau}_{Ct}$ and $\tilde{\tau}_{It}$ has to be decided by a normalization. In the empirical analysis of the paper, for example, we use $\tilde{\tau}_{C1} = 0$. To avoid ambiguities, below we report bias and coverage rates for $\tilde{\tau}_{Ct} - T^{-1} \sum_{s=1}^{T} \tilde{\tau}_{Cs}$ and $\tilde{\tau}_{It} - T^{-1} \sum_{s=1}^{T} \tilde{\tau}_{Is}$. Results look similar using alternative normalizations.

We display the bias calculations in Figure C1. For $\tilde{\tau}_t$, for example, we plot the sampling distribution of $\left\{\tilde{\tau}_t\left(\frac{1}{2}\right) - \tilde{\tau}_t\right\}_{t=1}^{T}$ indicating for each t the values contained between the 0.16- and 0.84-quantiles of the sampling distributions with a shaded area. We also report $\text{med}\left(\tilde{\tau}_t\left(\frac{1}{2}\right) - \tilde{\tau}_t\right)$ (blue dashed line) and $\text{bias}_t = \mathbb{E}\left[\tilde{\tau}_t\left(\frac{1}{2}\right) - \tilde{\tau}_t\right]$ (black dotted line). We do the same for $\tilde{\tau}_{Ct}$ and $\tilde{\tau}_{It}$. The figure shows that our approach has no systematic tendency to over or underestimate the trend, its common or its idiosyncratic component. This holds for both $\alpha_\tau = 0.3$ (a value representative of our sample) and, reassuringly, for $\alpha_\tau = 0$. In other words, even in the extreme case where the common component does not exist, there is no evidence to suggest that our model would spuriously find a role for a common component.

Turning to the coverage properties of posterior intervals, the performance of our method is solid. We report the average over t of estimated coverage rates $T^{-1}\sum_{t=1}^{T} \text{cov}_t$ for our two designs in Table C1.

TABLE C1. Average coverage rates for nominal rate $1 - \beta = 0.68$

	$\alpha_\tau = 0$	$\alpha_\tau = 0.3$
$\tilde{\tau}_t$	0.763	0.719
$\tilde{\tau}_{Ct}$	0.998	0.665
$\tilde{\tau}_{It}$	0.896	0.657

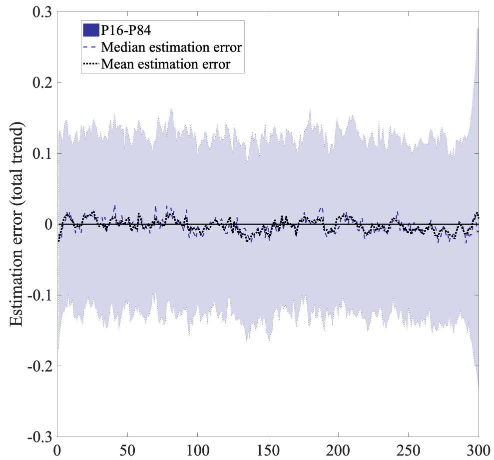
We set the probability level to $1-\beta = 0.68$, the level we use in our empirical analysis and equivalent to intervals of roughly one standard deviation radius under a normal distribution. When $\alpha_\tau = 0.3$, the average coverage rates are reasonably close to the nominal rate suggesting that our framework produces reliable inferences about the trend and its common and idiosyncratic components in repeated samples.[3]

When $\alpha_\tau = 0$, our method produces relatively conservative inferences in the sense that it overcovers both the common and idiosyncratic component. In particular, the probability bands for $\tilde{\tau}_{Ct}$ contain the zero line (its true value) in practically all samples. This agrees with our claim that our method does not have a systematic tendency to find a common component when there is none.

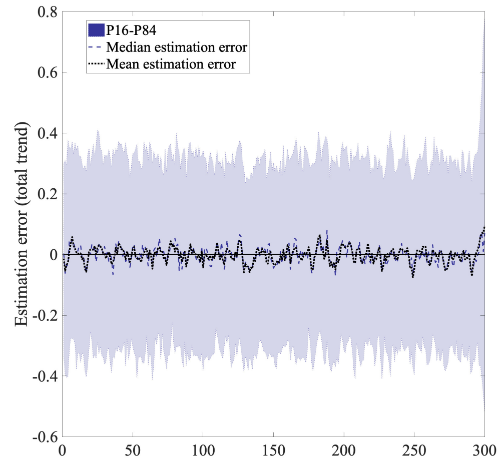
These results are important because the good coverage of our method is a frequentist property, even though the intervals we use are Bayesian credible intervals. Moreover, our estimation and inference approach uses a prior that does not center the model at the DGP, suggesting that shrinking our model away from the DGP has a negligible effect as with these parameters and sample size the prior is dominated by the sample information.

[3] Our method also achieves good coverage pointwise in t.

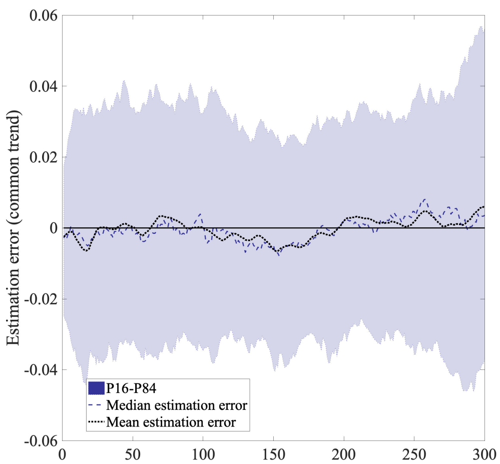
FIGURE C1. Bias of the posterior median estimate



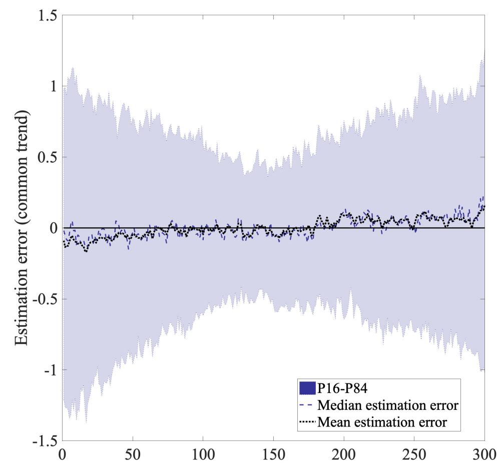
(a) Bias for $\tilde{\tau}_t$ for $\alpha_\tau = 0$



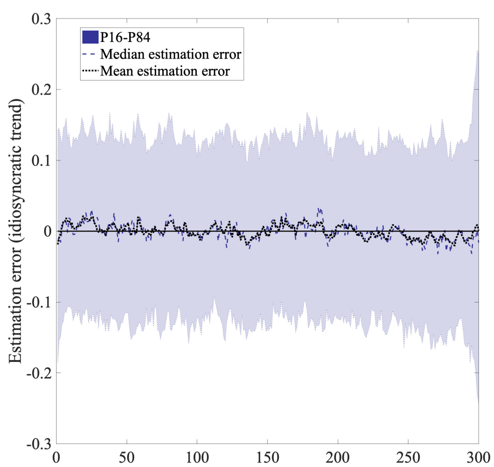
(b) Bias for $\tilde{\tau}_t$ for $\alpha_\tau = 0.3$



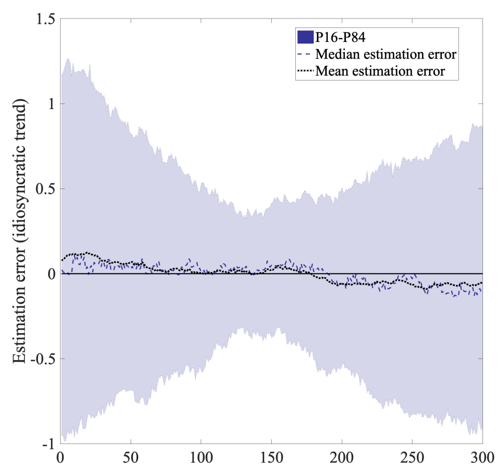
(c) Bias for $\tilde{\tau}_{Ct} - T^{-1}\sum_{s=1}^{T}\tilde{\tau}_{Cs}$ for $\alpha_\tau = 0$



(d) Bias for $\tilde{\tau}_{Ct} - T^{-1}\sum_{s=1}^{T}\tilde{\tau}_{Cs}$ for $\alpha_\tau = 0.3$



(e) Bias for $\tilde{\tau}_{It} - T^{-1}\sum_{s=1}^{T}\tilde{\tau}_{Is}$ for $\alpha_\tau = 0$

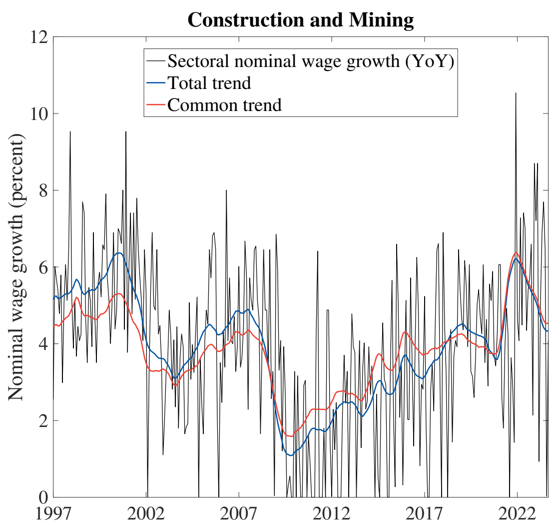


(f) Bias for $\tilde{\tau}_{It} - T^{-1}\sum_{s=1}^{T}\tilde{\tau}_{Is}$ for $\alpha_\tau = 0.3$

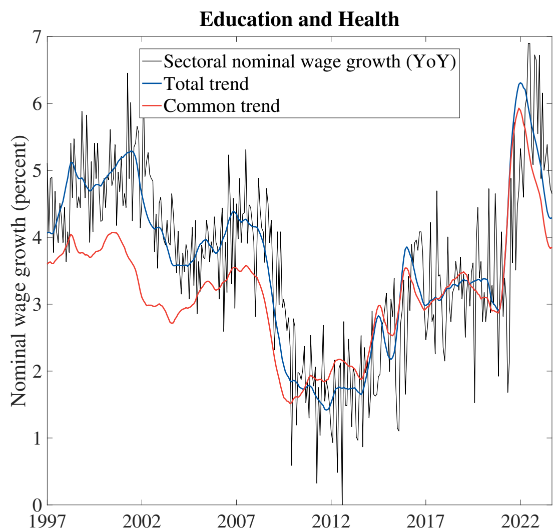
D Additional empirical results

D.1 Wage growth behavior across industries

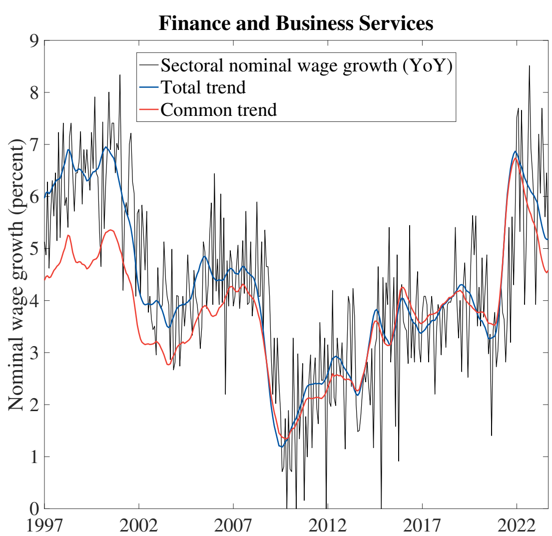
FIGURE D1. Aggregate and group-specific trend by industry



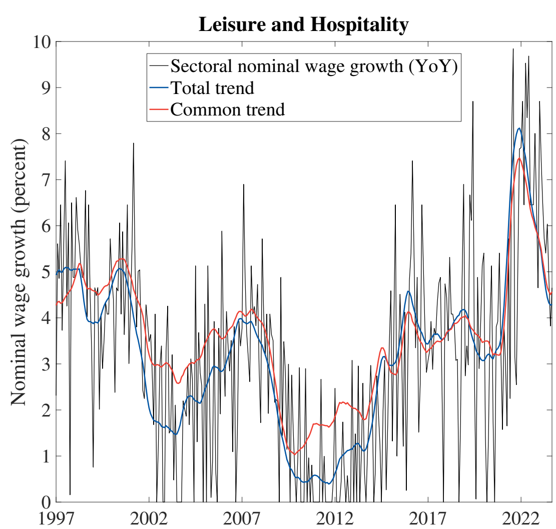
(a) Construction and Mining



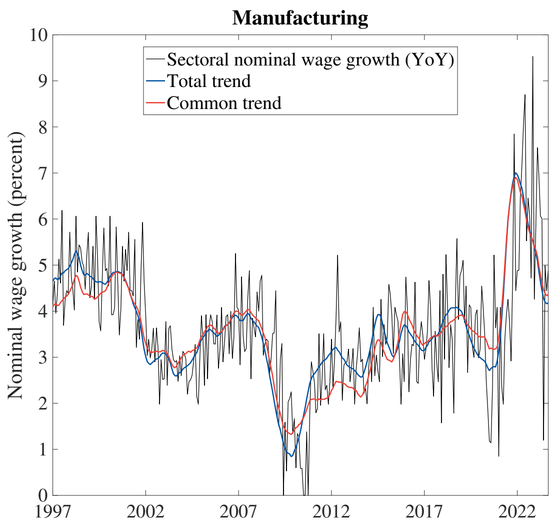
(b) Education and Health



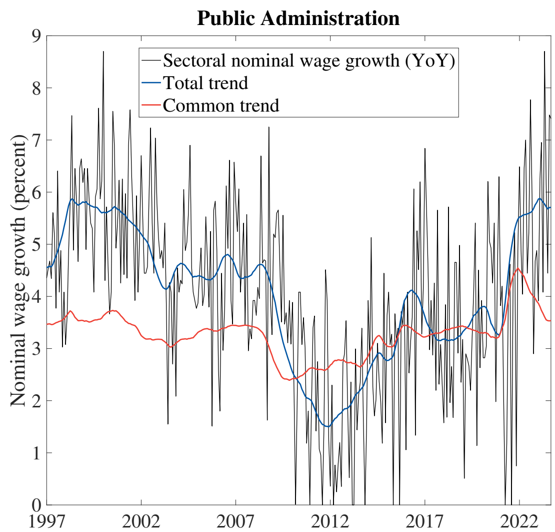
(c) Finance and Business Services



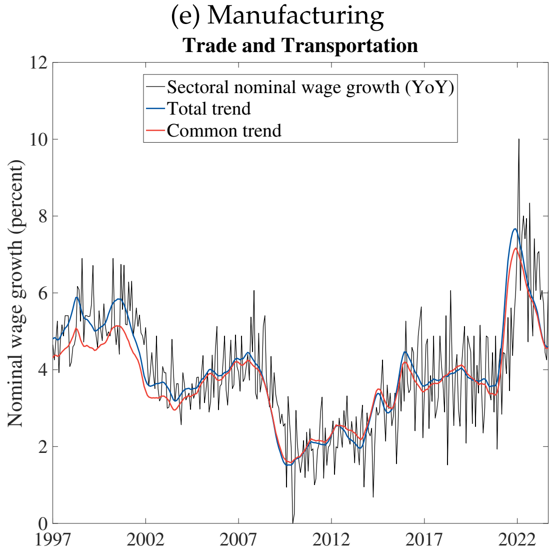
(d) Leisure and Hospitality



(e) Manufacturing



(f) Public Admininistration

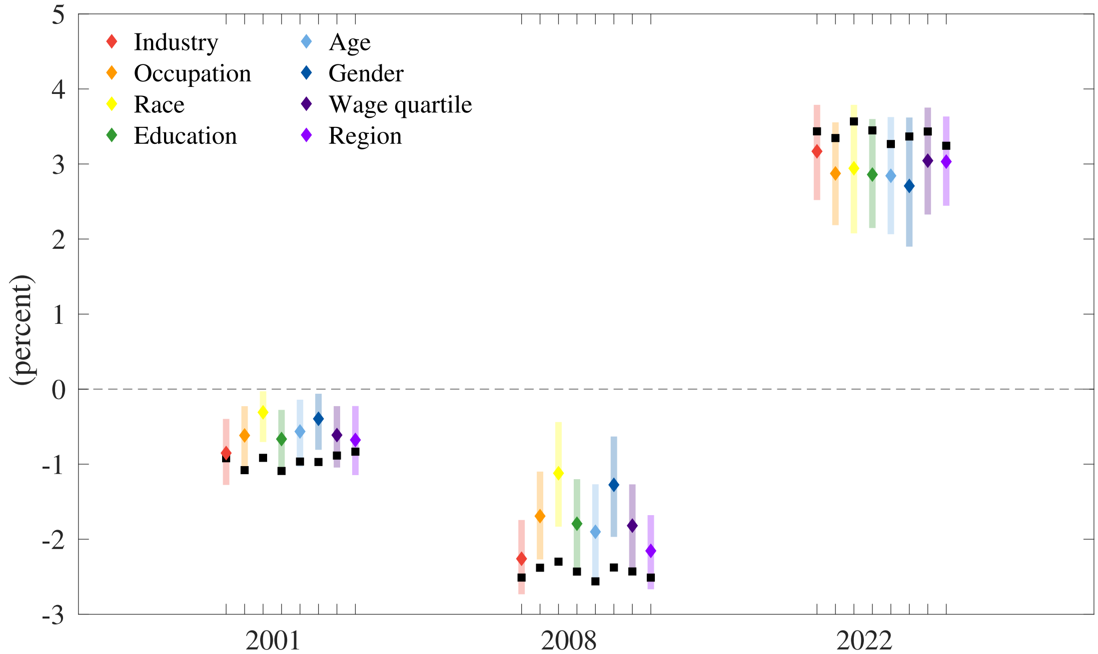


(g) Trade and Transportation

NOTES. The figure shows for each industry the raw nominal wage growth data, the common trend component ($\alpha_{\tau,it}\tau_{ct}$) and the sector specific trend (τ_{it}) over the sample period.

D.2 Core Wage Inflation using alternative CPS partitions

FIGURE D2. Core Wage Inflation across models



NOTES. The three episodes refer to the following periods: 2001m3-2001m11, 2007m12-2009m6, and 2020m6-2022m2. Each model uses the data cut described in the legend; the variables follow the Atlanta Fed Wage Growth tracker definition and are detailed in Appendix A. Black square markers indicate the peak-to-trough change in Core Wage Inflation for each model and episode. The diamond markers indicate the peak-to-trough change for the common component $\left(\sum_{i=1}^{n} s_{it}\alpha_{\tau,it}\right)\tau_{ct}$ in each model and episode. Vertical lines show the 68 percent probability bands.

D.3 Core Wage Inflation, labor market conditions, and price inflation

In Table D1, we benchmark our Core Wage Inflation measure, and its common component, to three commonly used measures of aggregate nominal wage growth: average hourly earnings (AHE), the Atlanta Fed Wage Growth Tracker (AWGT), and the Employment Cost Index. We report contemporaneous correlation between changes in each of these measures and changes in labor market conditions (vacancy to labor force ratio and unemployment rate) and price inflation (core PCE inflation and core services ex-housing inflation). Changes are monthly in the upper part of the table and quarterly in the lower part, given the quarterly frequency of ECI.

TABLE D1. Correlations with labor market and price inflation time series

	Vacancy to labor force ratio	Unemployment rate	Core PCE inflation	Core services ex housing PCE inflation
Monthly wage inflation measures				
Core Wage Inflation	0.331***	−0.094*	0.300***	0.242***
Core Wage Inflation (common)	0.320***	−0.077	0.310***	0.245***
AHE	−0.174***	0.588***	−0.227***	−0.226***
Atlanta Fed wage tracker	−0.003	−0.115**	0.088	0.028
Quarterly wage inflation measures				
Core Wage Inflation	0.718***	−0.295***	0.427***	0.334***
Core Wage Inflation (common)	0.717***	−0.285***	0.436***	0.334***
ECI	0.159	−0.158	0.221**	0.209**

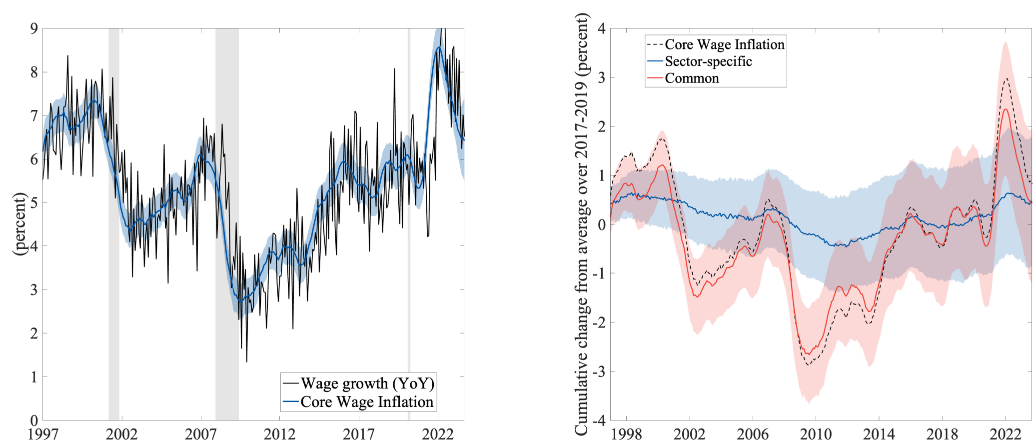
NOTES. Core Wage Inflation, Atlanta Fed wage tracker, AHE, the unemployment rate and price inflation measures are monthly series over the period 1997m1–2023m9. The vacancy ratio is over the period 2000m12–2023m9. Results are qualitatively similar prior to 2020m3. Vacancies are seasonally adjusted job openings from the Job Openings and Labor Turnover Survey (JOLTS) of the Bureau of Labor Statistics. AHE is the 12-month percent change in average hourly earnings of production and non-supervisory employees on private nonfarm payrolls, from the Current Employment Statistics of the Bureau of Labor Statistics. The Atlanta Fed wage tracker (Atlanta Fed [2023]) is the unweighted 3-month moving average of median 12-month wage growth. Core PCE inflation comes from the Bureau of Economic Analysis and excludes energy and food. ECI is a quarterly measure of the 12-month percent change in the Employment Cost Index measured by the Bureau of Labor Statistics. The time period is 1997Q1–2023Q2. When computing correlations at quarterly frequency, the price inflation measures and Core Wage Inflation are 12-month changes using the third month of the quarter as ECI. All correlations are for variables in first differences.

Changes in Core Wage Inflation (and its common component) are strongly and significantly correlated with changes in the vacancy to labor force ratio: intuitively, a tighter labor market should put upward pressure on wages. The correlation is insignificant for ECI, absent for the AWGT, and even of the opposite sign for AHE. A similar pattern is observed for the unemployment rate, although differences are less stark – with the exception of AHE. Changes in Core Wage Inflation and its common component are positively and significantly correlated with changes in inflation, consistent with the idea that wage pressures may be associated with price pressures. The ECI has a similar pattern but a weaker association; the Atlanta Fed measure is uncorrelated with inflation whereas AHE has, again, the wrong sign.

E Robustness checks

In this appendix, we verify the robustness of the main results of the paper to three choices we make in the empirical analysis. The first is to use the median instead of the mean of year-over-year wage growth as the observable w_{it} in our model. The second choice is to use the unweighted median as opposed to the median weighted by the survey weights as w_{it}. Third, we do not allow for $\tilde{\tau}_{it}$ to be itself the sum of a persistent and a transitory component. Figures E1, E2 and E3 show that both the historical behavior of the persistent component in wage growth and the relatively high importance of common variation across industries are insensitive to these choices.

FIGURE E1. Estimates based on mean year-over-year wage growth

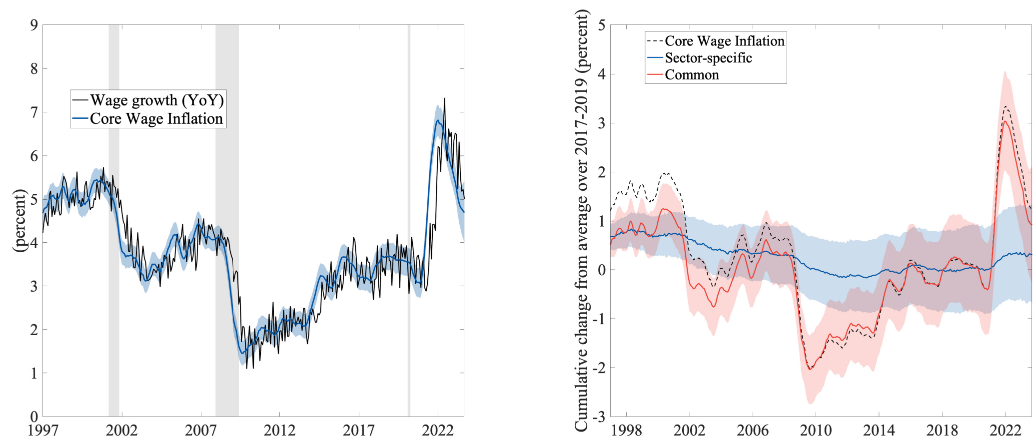


(a) Persistent component of wage growth (b) Common and sector-specific contribution

Despite mean year-over-year wage growth being more volatile than median wage growth, our model traces a remarkably similar historical evolution of the persistent component (Core Wage Inflation), with the largest swings located around the same episodes we discussed in Section 4 (i.e., the 2001 and 2008 recessions, and the post-pandemic inflation spike). Core Wage Inflation is somewhat higher when using the mean instead of the median due to the positive skewness in the wage growth distribution, but this seems to imply merely a level shift in the persistent component. The cumulative changes in panel (b) of figure E1, for example, are quantitatively very close to our baseline results.

Differences in our estimates when using the weighted instead of the unweighted median of wage growth as w_{it} are imperceptible, as shown in figure E2.

FIGURE E2. Estimates based on weighted median wage growth



(a) Persistent component of wage growth (b) Common and sector-specific contribution

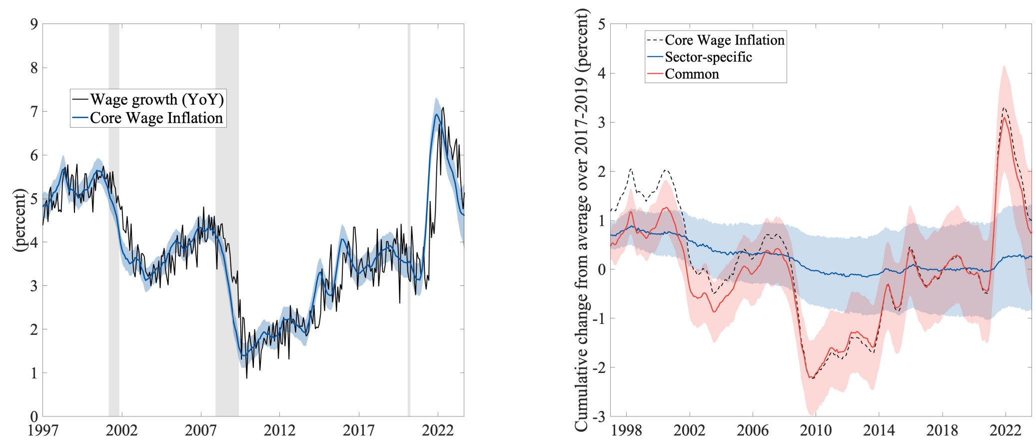
Figure E3 illustrates a point made in section 3. Our empirical analysis interprets the transitory component of year-over-year wage growth $\tilde{\varepsilon}_{it}$ as being largely measurement error. Therefore, $\tilde{\tau}_{it}$ is interpreted as the unobservable monthly growth rate of nominal wages that could be recovered with a perfect error-free survey. Because we rely on time series smoothing techniques, the assumption that $\tilde{\tau}_{it}$ is well approximated by a random walk is important in order to filter the survey measurement error out. If instead $\tilde{\tau}_{it}$ is the sum of two components,

$$\tilde{\tau}_{it} = \tilde{\tau}_{it}^{\text{pers}} + \tilde{\tau}_{it}^{\text{tr}}$$

where $\tilde{\tau}_{it}^{\text{pers}}$ is now a random walk and $\tilde{\tau}_{it}^{\text{tr}}$ is white noise, our baseline model with the choice of moving average orders $p = q = 12$ would estimate $\tilde{\tau}_{it}^{\text{pers}}$ instead of $\tilde{\tau}_{it}$. Comparing estimates from this extended model and the results in section 4 provides a sense of how important the genuine transitory shock $\tilde{\tau}_{it}^{\text{tr}}$ is. Figure E3 indicates that $\tilde{\tau}_{it}^{\text{tr}}$

plays at most a minor role and that the more parsimonious model used in our paper captures sufficiently well the most salient movements in aggregate wage growth, which tend to be very persistent.

FIGURE E3. Estimates based on a more flexible specification



(a) Persistent component of wage growth (b) Common and sector-specific contribution

A final piece of evidence supporting our interpretation of $\tilde{\varepsilon}_{it}$ as measurement error is shown in figure E4. Consider the transitory component of aggregate wage growth, which we define as

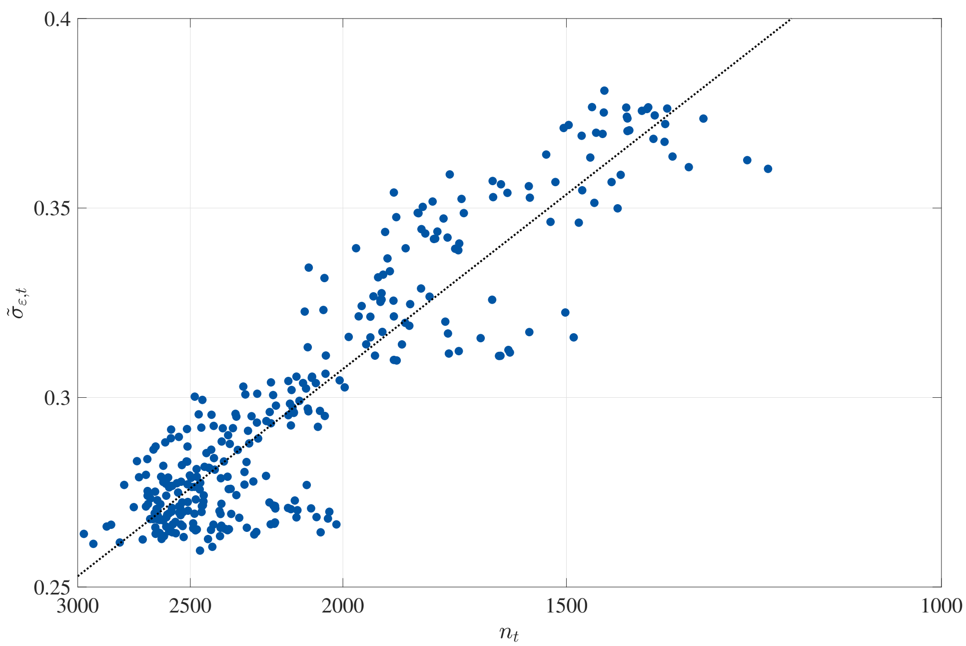
$$\tilde{\varepsilon}_t = \sum_{i=1}^{n} s_{it}\tilde{\varepsilon}_{it} = \sum_{i=1}^{n} s_{it}w_{it} - \tilde{\tau}_t$$

where s_{it} is the employment share of cross-section i in month t. The variance of $\tilde{\varepsilon}_t$ is given by

$$\tilde{\sigma}^2_{\varepsilon,t} = \left(\sum_{i=1}^{n} s_{it}\alpha_{\varepsilon,it}\right)^2 \sigma^2_{\varepsilon,ct} + \sum_{i=1}^{n} s^2_{it}\sigma^2_{\varepsilon,it}.$$

If $\tilde{\varepsilon}_{it}$ is the error made in using the sample median of year-over-year wage growth w_{it} from a sample of n_{it} workers to estimate the population growth rate $\sum_{\ell=1}^{12} \tilde{\tau}_{it+1-\ell}/12$ in each sector i, then the standard deviation $\tilde{\sigma}_{\varepsilon,t}$ should be proportional to $1/\sqrt{n_t}$ where $n_t = \sum_{i=1}^{n} n_{it}$ is the survey sample size in month t. Figure E4 shows precisely that: a scatter plot of (the posterior median estimate of) $\tilde{\sigma}_{\varepsilon,t}$ against n_t in which most of the points lie close to the line $\tilde{\sigma}_{\varepsilon,t} = \hat{c}/\sqrt{n_t}$.[4] We find a similar pattern if we consider the correlation between sample size n_{it} and the standard deviation $\sigma_{\varepsilon,it}$ for a specific industry i.

FIGURE E4. Standard deviation of transitory shocks and survey sample sizes



We also find, consistent with our interpretation, that for every i the path of $\alpha_{\varepsilon,it}\sigma^2_{\varepsilon,ct}$ always contains the zero line, indicating a negligible role for cross-sectional correlation across $\tilde{\varepsilon}_{it}$.

F Additional evidence using CES

This appendix presents estimates of the persistent component of month-on-month growth rates in nominal wages using data from the Current Employment Statistics (CES).[5] Because the data already provides month-on-month changes, denoted by W_{it} below, we estimate our model without temporal aggregation. In other words, instead of (2), our measurement equation is

$$W_{it} = \tilde{\tau}_{it} + \tilde{\varepsilon}_{it}$$

[4]In fact, the correlation between $\tilde{\sigma}_{\varepsilon,t}$ and $1/\sqrt{n_t}$ is 0.9.

[5]We use average hourly earnings of production and non-supervisory employees on private nonfarm payrolls.

with the persistent component $\tilde{\tau}_{it}$ and the transitory component $\tilde{\varepsilon}_{it}$ modeled as in Section 3. The cross-sectional dimension is industries since the CES is a survey of establishments.[6]

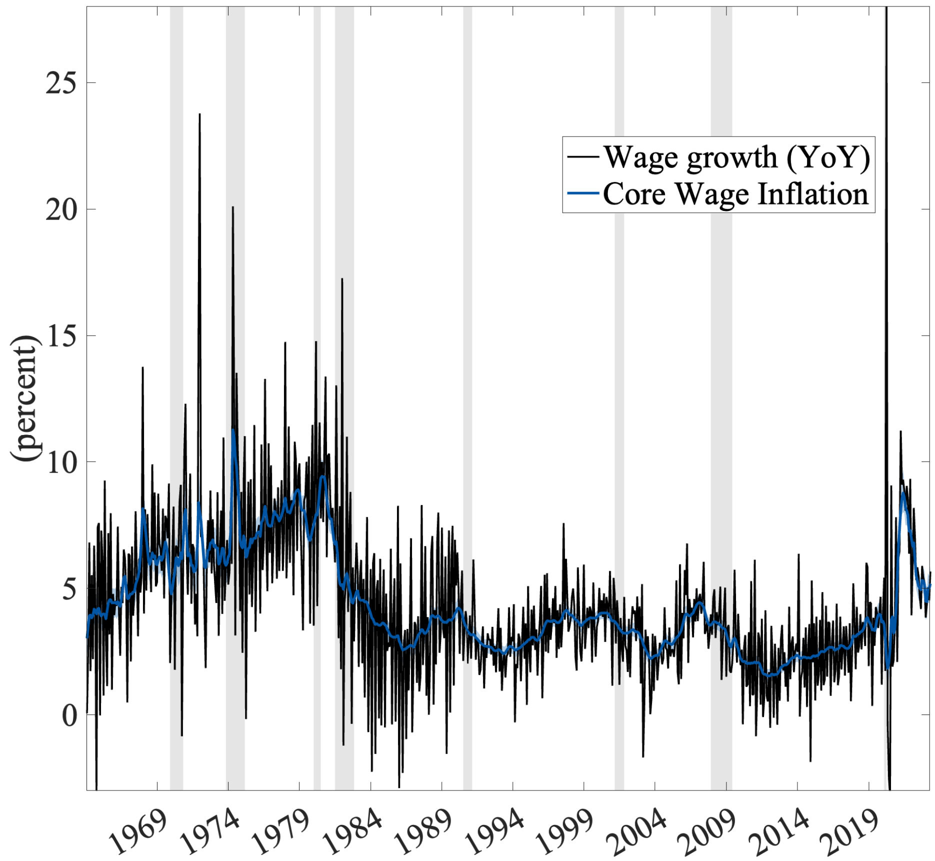
As noted in Section 2, the CES measure of wages is subject to compositional issues. However, the CES spans a longer period (in this case beginning in 1964), which allows us to empirically study additional recessions and the inflationary episodes of the late 1960s and 1970s. Figure F1 contains the trend estimates and its decomposition into common and sector-specific drivers. Figure F1a is the CES equivalent to Figure 2a and Figure F1b is comparable to Figure 2b.

Figure F1a shows that the model attributes most of the high-frequency variation in nominal wage growth in the CES to the transitory variation term $\tilde{\varepsilon}_{it}$. The two largest changes in the persistent component of wage inflation correspond to the inflation episodes in the 1970s and the post-pandemic inflation surge. From the 1980s, most NBER recessions tend to be associated with a drop in Core Wage Inflation.

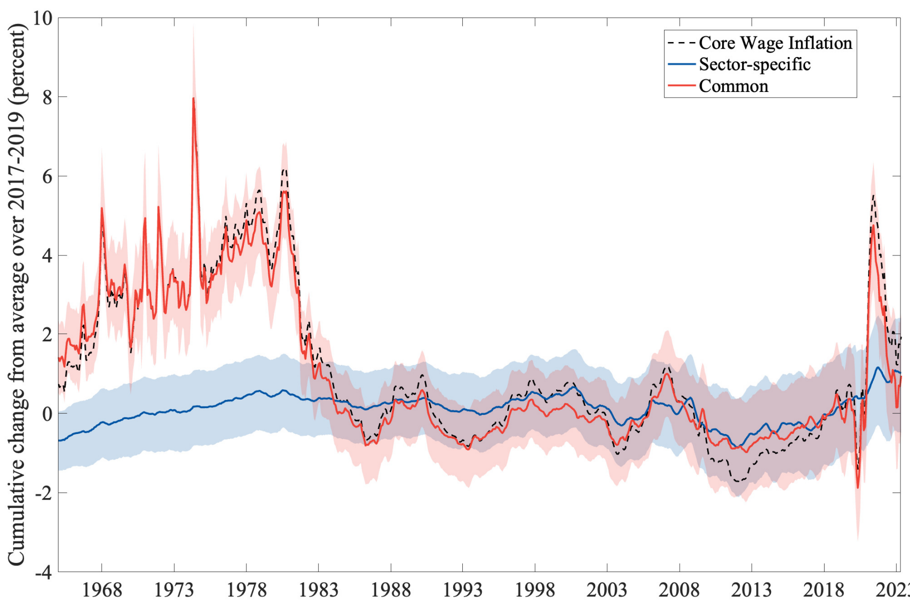
In addition, Figure F1b confirms our findings that the sector-specific persistent component captures very low frequency movements. In contrast, the common latent factor plays a prominent role during large swings in aggregate nominal wage growth, and especially in the inflationary periods. This is visually clear in the 1970s, thus suggesting that our results are not specific to the post-pandemic period. For example, about 80% of the 4.5 percentage point increase in wage inflation between 1965 and 1981 is common across industries. Looking at shorter periods, the common factor explains 78% and 85% of the increase in wage inflation between 1973 and 1975 and between 1980 and 1982, respectively.

[6] We consider 10 industries: Construction, Financial Activities, Information, Leisure and Hospitality, Manufacturing, Mining and Logging, Other Services, Private Education and Health Services, Professional and Business Services, Trade-Transportation-Utilities.

FIGURE F1. Estimates using CES data



(a) Persistent component of wage growth



(b) Common and sector-specific contribution

References: Supplemental Appendix

M. Del Negro and C. Otrok. Dynamic factor models with time-varying parameters: measuring changes in international business cycles. Technical report, 2008.

T. Doan, R. B. Litterman, and C. A. Sims. Forecasting and Conditional Projection Using Realistic Prior Distributions. NBER Working Papers 1202, National Bureau of Economic Research, Inc, Sept. 1983. URL https://ideas.repec.org/p/nbr/nberwo/1202.html.

J. Durbin and S. J. Koopman. A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89(3):603–615, 2002.

J. Geweke. Getting it right: Joint distribution tests of conditional simulators. *Journal of the American Statistical Association*, 99:799–804, 2004.

A. Inoue and L. Kilian. Joint confidence sets for structural impulse responses. *Journal of Econometrics*, 192(2):421–432, 2016. doi: 10.1016/j.jeconom.2016.02. URL https://ideas.repec.org/a/eee/econom/v192y2016i2p421-432.html.

S. Kim, N. Shephard, and S. Chib. Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65:361–393, 1998.

Y. Omori, S. Chib, N. Shephard, and J. Nakajima. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140:425–449, 2007.

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