

Maximally Forward-Looking Core Inflation

Philippe Goulet Coulombe
Université du Québec à Montréal

Karin Klieber
Oesterreichische Nationalbank

Christophe Barrette
Université du Québec à Montréal

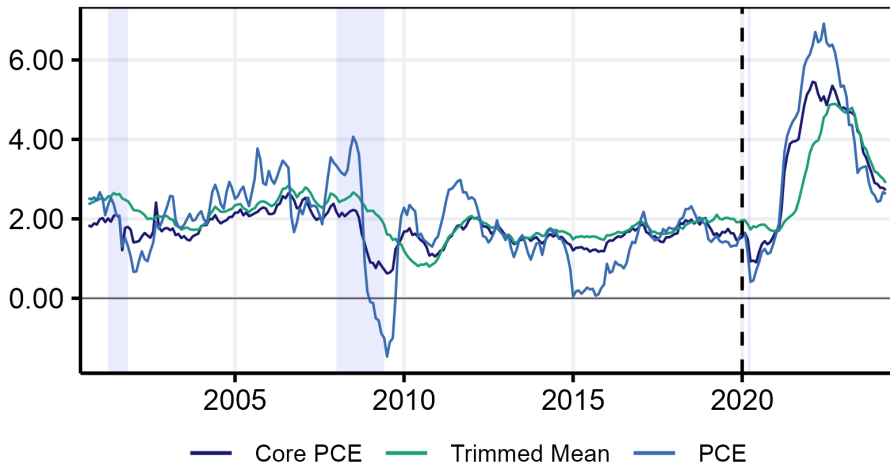
Maximilian Göbel
Bocconi University

SUERF Workshop : The use of AI in Economic Modelling and Forecasting

*The content of these slides reflects the views of the authors and not necessarily those of the OeNB or the Eurosystem.

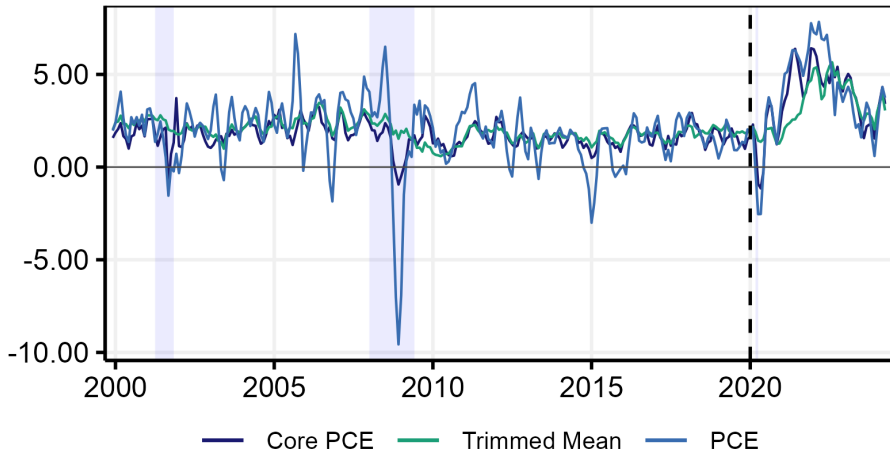
Where are we?

Some measures of inflation (YoY) for the US



Where are we?

Some measures of inflation (QoQ) for the US



Core Inflation: A Primer

How are they created?

1. Heuristics (permanent exclusion)
 - HICPX/PCE core (HICP/PCE minus food and energy)
 - PCE supercore (PCE core services minus housing)
2. Heuristics (temporary exclusion)
 - Median inflation
 - Trimmed-mean inflation
3. Model-based (unsupervised)
 - (US) Multivariate Core Trend from FRB-NY (Stock and Watson, 2016)
 - (EA) PCCI (Bańbura and Bobeica, 2020)

What are the desirable properties?

1. Low bias
2. Low volatility
3. Low error in forecasting headline inflation
4. Highly reactive to monetary policy and labor market conditions

Contribution

- We write a simple machine learning regression problem that generates, by construction, the linear aggregation of subcomponents that is maximally predictive of future inflation conditions.
- We call it the *Assemblage Regression*.
- In the context of inflation modeling, the resulting product is Albacore, for adaptive learning-based core inflation.
- Features: (3) good forecasting performance, (1) low bias, (2) low volatility.
- Criterion (4) is the subject of another ongoing project.

Assemblage Regression

Our proposed measures are estimated via a **generalized non-negative ridge regression** where the dependent variable is future headline inflation:

- **Supervised Weighting with Albacore_{comps}**

$$\hat{w}_c = \arg \min_w \sum_{t=1}^{T-h} (\pi_{t+1:t+h} - w' \Pi_t)^2 + \lambda \|w - w_{\text{PCE}}\|_2 \quad \text{st } w \geq 0, w' \iota = 1$$

where $\pi_{c,t}^* \equiv \hat{w}'_c \Pi_t$ and $h \in \{6, 12, 24\}$ months. Π_t are inflation subcomponents.

- **Supervised Trimming with Albacore_{ranks}**

$$\hat{w}_r = \arg \min_w \sum_{t=1}^{T-h} (\pi_{t+1:t+h} - w' O_t)^2 + \lambda \|Dw\|_2 \quad \text{st } w \geq 0$$

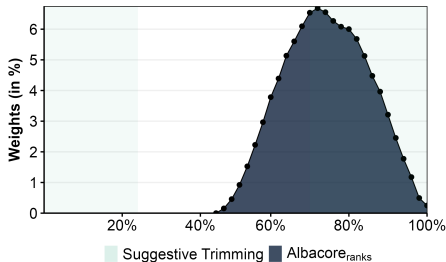
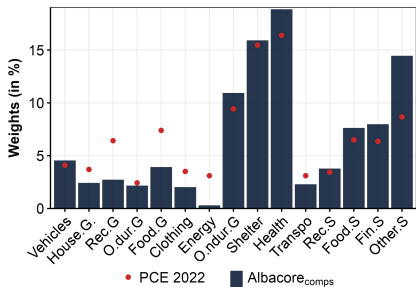
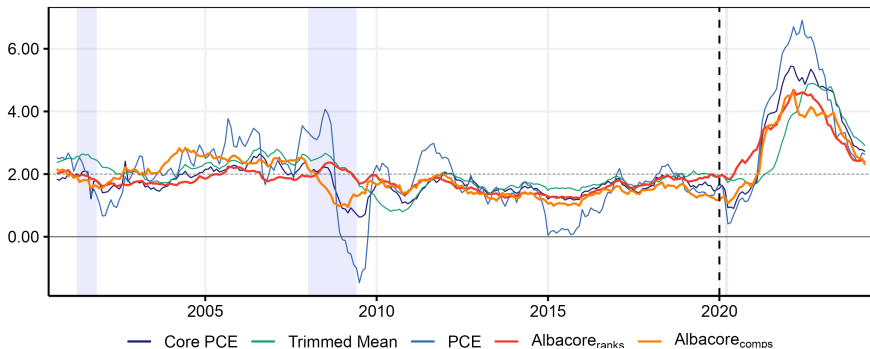
where $\pi_{r,t}^* \equiv \hat{w}'_r O_t$ and $h \in \{6, 12, 24\}$ months. O_t are "order statistics" and store inflation subcomponents sorted at each t .

Experiments and Forecasting Performance

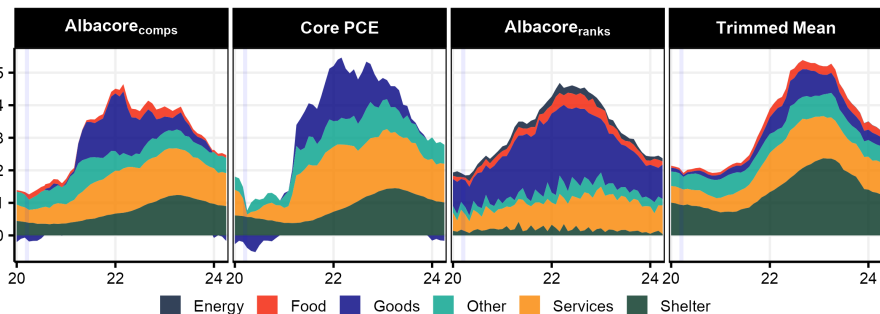
- Monthly data, predicting out-of-sample the average path of headline inflation for the next $h \in \{6, 12, 24\}$ months.
- $\Pi_t \equiv$ PCE/HICP/CPI components at 3 levels of disaggregation.
 - US: level $\in \{2, 3, 6\}$ implying $K \in \{15, 50, 215\}$
 - EA: level $\in \{2, 3, 4\}$ implying $K \in \{12, 39, 92\}$
- Benchmarks: forecast combination of headline and a wide range of existing core inflation measures
- Two test periods, with a 20 years rolling window estimation
 1. Pre-Covid: 2010-2019
 2. Post-Covid: 2020-2023

\Rightarrow $\text{Albacore}_{\text{comps}}$ and $\text{Albacore}_{\text{ranks}}$ show good forecasting performance at all levels of disaggregation.

Albacore for the US

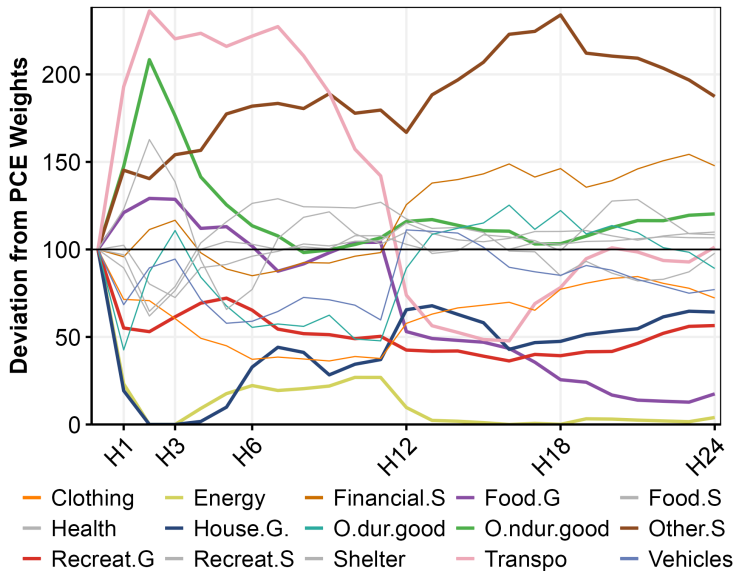


Albacore Decomposition US

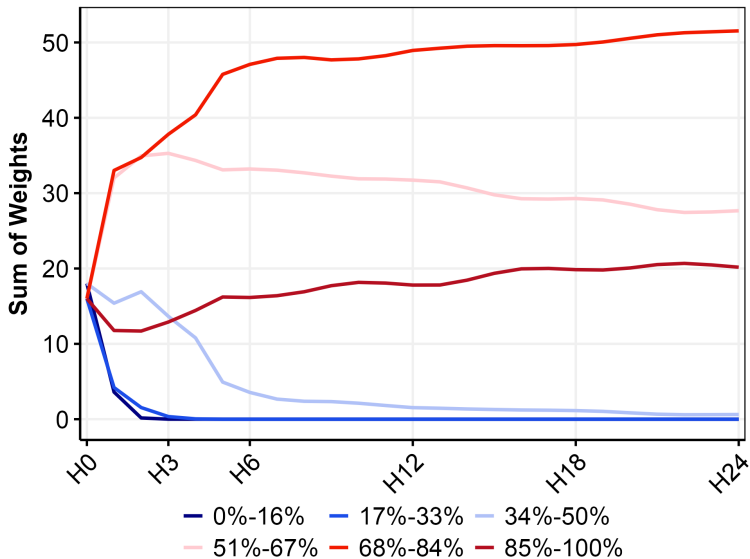


- We find high contribution of goods inflation for Albacore_{ranks} whereas the trimmed mean is driven by shelter (which is known to be lagging).
- Unlike the PCE core, Albacore_{comps} and Albacore_{ranks} do not exclude food or energy prices a priori, yet they show minimal contribution from these components.

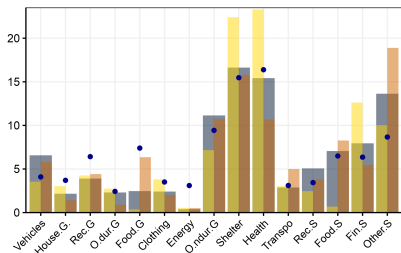
Weights Curve Across Horizons: Albacore_{comps}



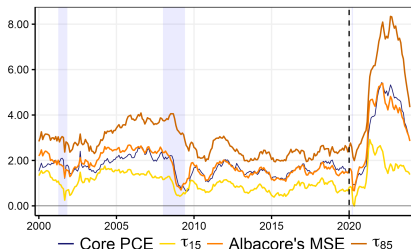
Weights Curve Across Horizons: Albacore_{ranks}



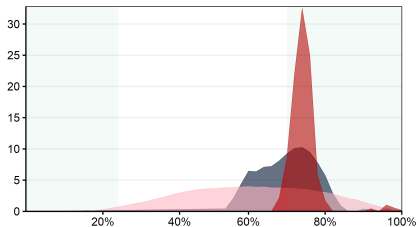
Core Inflation Measures Specialized for Tail Risks



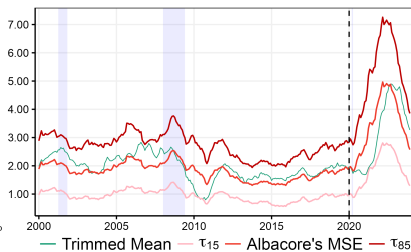
• PCE 2022 τ_{15} Albacore's MSE τ_{85}
(a) Comparison of Component Weights



(b) Time Series Comparison of Qualbacore_{comps}



Suggestive Trimming τ_{15} Albacore's MSE τ_{85}
(c) Comparison of Rank Weights



(d) Time Series Comparison of Qualbacore_{ranks}

Parting Words

- We devised 2 simple maximally forward-looking core inflation series, assembling disaggregated data to build the best leading indicator.
- This is part of a broader research agenda where macro aggregates are (partly) redefined by a machine learning algorithm so to better satisfy certain (economic or statistical) evaluation criteria.
 - Core wage growth? Core GDP?
 - Use micro-level price data?
- Circling back to the theme of this workshop, supervised aggregation is one of key building block of modern AI.
- Rather than manually building predictors from disaggregates, one can endogenously construct the features from the unstructured inputs (e.g., pixels) through multiple layers which are jointly optimized along with the parameters of the deep learning model using them.

Appendix

Detour: A Primer on Ridge Regression

- A ridge regression solves what is more generally known as a penalized linear regression problem.
- The RR coefficients are obtained via

$$\hat{\beta}_{\text{Ridge}} = \arg \min_{\beta} \sum_{t=1}^{T-1} (y_{t+1} - \beta' \mathbf{X}_t)^2 + \lambda \|\beta\|_2$$

where $\|\cdot\|_2$ is the l_2 norm. The latter is equivalently $\sum_{k=1}^K \beta_k^2$ in summation notation, where $K = \text{ncol}(\mathbf{X})$.

- The penalty term provides regularization by bringing in the mix the a priori that each coefficient should contribute to the fit, but modestly. In other words, it is *shrinking* coefficients towards 0.
- With a suitable $\lambda > 0$, RR curbs overfitting that would plague OLS (i.e., $\lambda = 0$), especially when K is large relative to N .
- λ is typically tuned via cross-validation, which is, in essence, a pseudo-out-of-sample evaluation metric.

Assemblage Regression

Supervised Weighting with Albacore_{comps}

- Our proposed measure is estimated via a **non-negative ridge regression** where the dependent variable is future headline inflation and the regressors are inflation subcomponents.
- The permanent exclusion version, or supervised weighting of basket components, is obtained via

$$\hat{w}_c = \arg \min_w \sum_{t=1}^{T-h} (\pi_{t+1:t+h} - w' \Pi_t)^2 + \lambda \|w - w_{\text{PCE}}\|_2 \quad \text{st } w \geq 0, w' \mathbf{1} = 1$$

where $\pi_{c,t}^* \equiv \hat{w}'_c \Pi_t$ and $h \in \{6, 12, 24\}$ months.

- Note the absence of an intercept which forces (i) the random walk hypothesis and (ii) $\pi_{c,t}^*$ to have the same unconditional mean as $\pi_{t+1:t+h}$ through $w_{\text{intercept}} = \bar{\pi}_{t+1:t+h} - \bar{\pi}_t^* = \bar{\pi}_t - \bar{\pi}_t^* \approx 0$, therefore making it already denoted in headline inflation units.
- Some kind of multivariate-to-univariate Hamilton filter (Hamilton, 2018).

Assemblage Regression

Supervised Trimming with Albacore_{ranks}

- This is effectively just sorting the components at each t and inputting them in the matrix \mathbf{O}_t .
- The switch to "rank statistics space" can be formalized as $\mathbf{O}_t = \mathbf{A}_t \mathbf{\Pi}_t$ where \mathbf{A}_t is an allocation matrix.
- Rather than learning which subcomponents to include, the problem will now be learning which ranks (and with which weight) to include or exclude. Thus, we run

$$\hat{\mathbf{w}}_r = \arg \min_w \sum_{t=1}^{T-h} (\pi_{t+1:t+h} - \mathbf{w}' \mathbf{O}_t)^2 + \lambda \|\mathbf{D}\mathbf{w}\|_2 \quad \text{st } \mathbf{w} \geq 0$$

where $\pi_{r,t}^* \equiv \hat{\mathbf{w}}_r' \mathbf{O}_t$ and $h \in \{6, 12, 24\}$ months.

- Albacore_{ranks} will favor a smooth and adaptive trim, in contrast to
 - FRB Cleveland trimmed-mean CPI, which sets to 0 the first and last 16% of ranks, and then a weighted average of the center band is reported.
 - FRB Dallas trimmed-mean PCE inflation, which trims 24% from lower tail and 31% from the upper tail.
 - Median CPI, which keeps only one rank, the middle one.

Euro Area Results

	2010m1 - 2019m12			2020m1 - 2023m12		
	$h = 6$	$h = 12$	$h = 24$	$h = 6$	$h = 12$	$h = 24$
Level 2 ($K = 12$)						
Albacore _{comps}	0.94	0.84	0.75	0.96	0.99	1.02
Albacore _{ranks}	0.90	0.84	0.78	0.95	0.94	0.95
Level 3 ($K = 39$)						
Albacore _{comps}	0.93	0.88	0.81	1.01	1.00	1.07
Albacore _{ranks}	0.90	0.85	0.85	0.97	0.95	0.94
Level 4 ($K = 92$)						
Albacore _{comps}	0.96	1.00	0.85	1.06	0.97	1.04
Albacore _{ranks}	0.92	0.87	0.87	0.93	0.93	0.91
Benchmarks						
$X_t^{\text{bm}}, (w_0 = 0)$	0.95	0.90	0.78	0.98	1.02	1.09
$X_t^{\text{bm}+}$	0.92	0.98	1.00	0.93	0.97	1.00
$X_t^{\text{bm}+}, (w_0 = 0)$	0.91	0.86	0.80	1.02	0.98	1.05

Notes: Numbers in **bold** indicate the best model for each *level* and each *horizon* in each of the out-of-sample periods. Numbers highlighted in **green** show the best model per *horizon* and out-of-sample period *across levels*.

US Results

	2010m1 - 2019m12			2020m1 - 2023m12		
	$h = 6$	$h = 12$	$h = 24$	$h = 6$	$h = 12$	$h = 24$
Level 2 ($K = 15$)						
Albacore _{comps}	1.08	1.13	1.12	0.70	0.63	0.88
Albacore _{ranks}	0.99	0.98	0.87	0.57	0.59	0.69
Level 3 ($K = 50$)						
Albacore _{comps}	1.10	1.14	1.06	0.80	0.75	1.01
Albacore _{ranks}	0.93	0.87	0.73	0.61	0.56	0.63
Level 6 ($K = 215$)						
Albacore _{comps}	1.10	1.19	1.16	0.84	0.71	0.99
Albacore _{ranks}	0.88	0.84	0.77	0.70	0.62	0.67
Benchmarks						
$X_t^{\text{bm}}, (w_0 = 0)$	1.03	1.03	0.96	0.79	0.80	0.93
$X_t^{\text{bm}+}$	1.04	1.04	1.02	1.18	1.14	1.09
$X_t^{\text{bm}+}, (w_0 = 0)$	1.07	1.07	0.97	0.86	0.87	0.96

Notes: Numbers in **bold** indicate the best model for each *level* and each *horizon* in each of the out-of-sample periods. Numbers highlighted in **green** show the best model per *horizon* and out-of-sample period *across levels*.

From One Space to Another

A time-varying parameters view

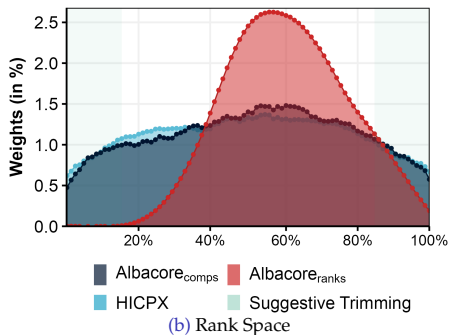
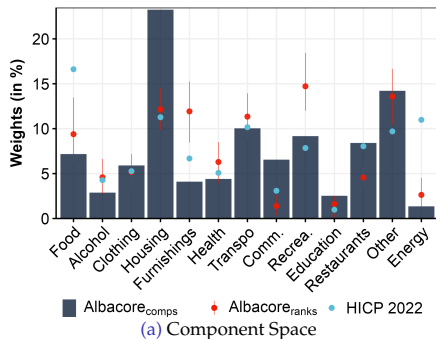
- The supervised trimming regression can be written in components space as

$$\hat{w}_r = \arg \min_w \sum_{t=1}^T (\pi_{t+1:t+h} - \underbrace{w_r' A_t}_{w_{c,t}} \Pi_t)^2 + \lambda \|w_r\|^2 \quad \text{st } w_r \geq 0$$

- Thus, fitting the regression in rank space implies *time-varying parameters* in the components space.
- Conversely, the regression in the original components space implies time-varying trimming via $w_{r,t} = w_c' A_t^{-1}$
- Traveling from one space to the other could be informative

From One Space to Another

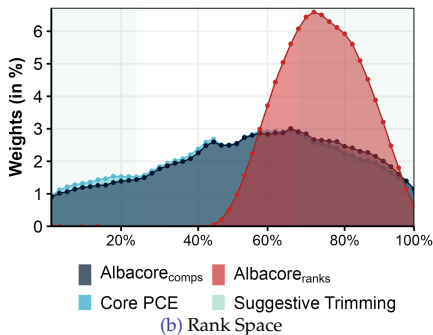
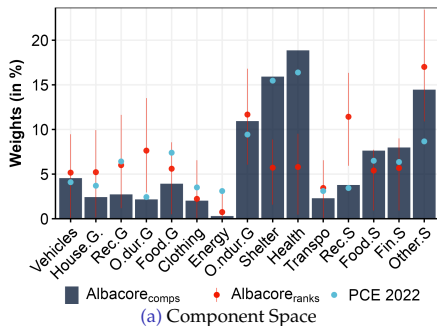
A look at our main EA models



- The asymmetric trimming of Albacore_{ranks} leads to low weights for components in the left tail.
- Albacore_{comps} covers the whole distribution; similar to HICPX, however, the latter being reweighted with actual HICP weights.

From One Space to Another

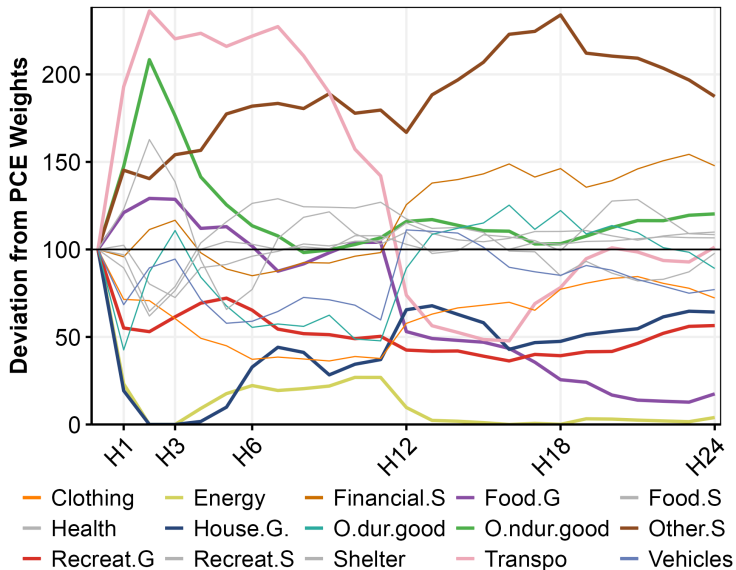
A look at our main US models



- The asymmetric trimming of Albacore_{ranks} leads to low weights for components in the left tail (which are mostly Energy, Food, Clothing).
- Albacore_{comps} covers the whole distribution but left-skewed; similar to Core PCE, however, the latter being reweighted with actual PCE weights.

Weights Curve Across Horizons: US

Albacore_{comps}



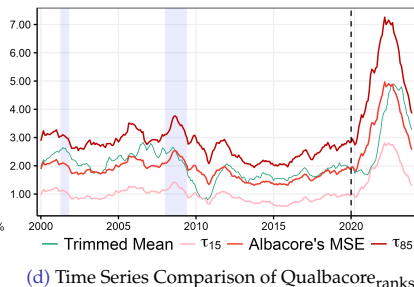
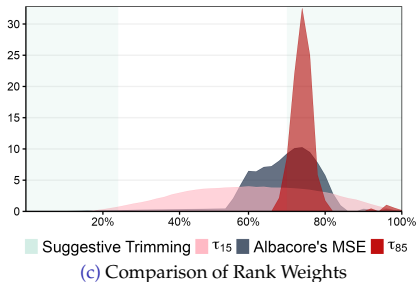
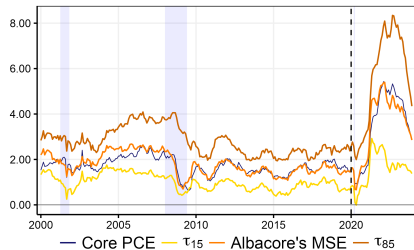
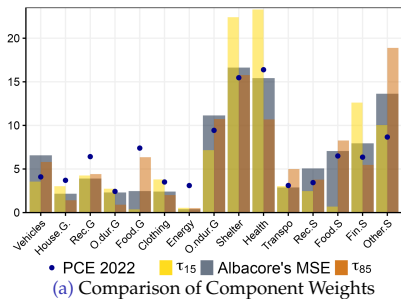
Qualbacre US Results

Table: Quantile Forecasting Performance of Albacore for the US

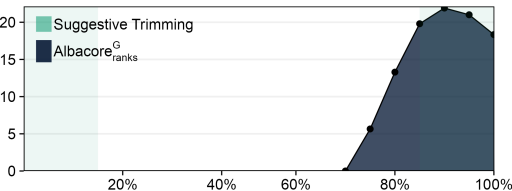
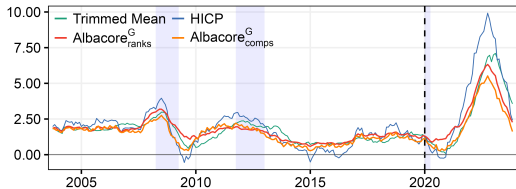
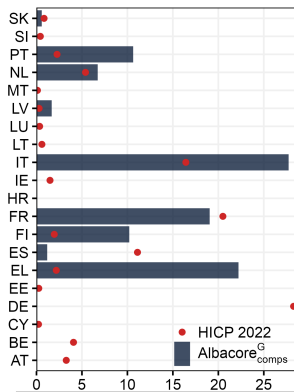
	2010m1-2019m12						2020m1-2023m12					
	$h = 3$			$h = 6$			$h = 3$			$h = 6$		
	τ_{15}	MSE	τ_{85}	τ_{15}	MSE	τ_{85}	τ_{15}	MSE	τ_{85}	τ_{15}	MSE	τ_{85}
Qualbacre _{comps}	0.97	1.06	1.29	1.10	1.09	1.21	1.24	1.07	0.75	1.23	0.97	0.71
Qualbacre _{ranks}	0.94	0.92	0.95	0.94	0.93	0.86	0.99	0.80	0.36	0.97	0.72	0.40
Benchmarks												
$X_t^{\text{bm}}, (w_0 = 0)$	0.96	0.99	1.12	0.97	1.01	1.11	1.04	0.96	0.81	1.00	0.94	0.90
$X_t^{\text{bm}+}$	1.03	1.03	1.13	1.07	1.03	1.11	1.14	1.22	0.81	1.01	1.07	0.83
$X_t^{\text{bm}+}, (w_0 = 0)$	1.02	1.02	1.13	1.07	1.04	1.13	1.14	1.02	0.81	1.01	0.95	0.90

Notes: The table presents root mean square error (RMSE) relative to the AR model. Our benchmarks are: $X_t^{\text{bm}} = [\text{PCE}_t \text{ PCEcore}_t \text{ PCEtrim}_t]$ without an intercept (i.e., $w_0 = 0$), $X_t^{\text{bm}+}$ with and without an intercept. Numbers in **bold** indicate the best performing model for each pair of horizon and loss function. Numbers in **green** highlights the loss function for which there is the largest improvement with respect to the benchmark, if applicable. In the results Section, we present level 6 for Qualbacre_{comps} (with $K = 215$) and level 3 for Qualbacre_{ranks} (with $K = 50$). Note that unlike previous results, we consider an expanding window from 1990 to increase the inevitably scarce number of observations in the tails.

Quantile regression extension for the US ($h = 6$)



Geographic Assemblage



	2010m1-2019m12			2020m1-2023m12		
	$h = 6$	$h = 12$	$h = 24$	$h = 6$	$h = 12$	$h = 24$
Albacore ^G _{comps}	0.96	0.81	0.70	0.90	0.96	1.12
Albacore ^G _{ranks}	0.86	0.74	0.66	0.80	0.87	0.93
Best Core Benchmark	0.93	0.87	0.75	0.92	0.97	0.99

Properties of Underlying Inflation Measures for US

Inflation series	Bias		Volatility		Coefficient of var.		Lead/lag corr.	
	Full	Pre-cov	Full	Pre-cov	Full	Pre-cov	Full	Pre-cov
3Mo3M								
Core PCE	-0.12	-0.09	0.57	0.36	0.59	0.39	0	0
Trimmed Mean (FedDallas)	0.02	0.14	0.42	0.29	0.41	0.27	-8	0
Core excl. housing	2.35	2.37	5.45	1.83	2.52	0.80	1	2
Median CPI (FedCleveland)	0.59	0.59	0.61	0.38	0.46	0.29	-8	-9
Trimmed Mean (FedCleveland)	0.33	0.31	0.63	0.40	0.53	0.34	0	0
Sticky Core (FedAtlanta)	0.51	0.56	0.57	0.37	0.45	0.28	-9	-11
Albacore _{ranks}	-0.03	0.03	0.40	0.22	0.40	0.21	0	0
Albacore _{comps}	-0.08	0.00	0.48	0.34	0.49	0.34	0	0
YoY								
Core PCE	-0.13	-0.11	0.69	0.39	0.49	0.22	0	0
Trimmed Mean (FedDallas)	0.01	0.13	0.56	0.46	0.37	0.23	-7	-9
Core excl. housing	2.42	2.43	2.93	2.03	0.91	0.46	2	3
Median CPI (FedCleveland)	0.55	0.56	0.82	0.61	0.43	0.25	-8	-11
Trimmed Mean (FedCleveland)	0.30	0.29	0.82	0.55	0.47	0.25	-4	-6
Sticky Core (FedAtlanta)	0.48	0.53	0.74	0.59	0.40	0.24	-8	-12
Albacore _{ranks}	-0.04	0.01	0.54	0.28	0.36	0.15	0	-1
Albacore _{comps}	-0.09	0.00	0.60	0.50	0.41	0.26	0	2

Notes: Cells shaded in green indicate the three best performing models for each criterion. The evaluation is based on the sample ranging from 2000m1 to 2023m12 for the full sample and from 2000m1 to 2019m12 for the pre-Covid case. Bias is determined as the difference between the long-run average of the respective series and PCE headline. Volatility refers to the standard deviation of each series relative to the standard deviation of headline. Coefficient of variation is the ratio between standard deviation and long-run average of each measure. Positive (negative) numbers for the lead/lag correlation refer to leads (lags) with the highest cross-correlation between each measure and PCE headline.

Properties of Underlying Inflation Measures for EA

Inflation series	Bias		Volatility		Coefficient of var.		Lead/lag corr.	
	Full	Pre-cov	Full	Pre-cov	Full	Pre-cov	Full	Pre-cov
PoP								
HICPX	-0.46	-0.29	0.53	0.37	0.72	0.36	-6	0
Trimmed Mean 30%	-0.22	-0.06	0.68	0.60	0.81	0.51	0	0
HICP excl. energy	-0.19	-0.11	0.67	0.50	0.78	0.44	-6	0
HICP excl energy & unpr food	-0.25	-0.15	0.64	0.45	0.77	0.40	-6	0
Supercore	-0.25	-0.09	0.55	0.37	0.68	0.33	-6	-2
PCCI excl. energy	5.28	4.89	1.43	1.02	0.43	0.21	-5	-1
PCCI	5.91	5.43	1.68	1.27	0.47	0.24	-1	0
Albacore _{ranks}	-0.17	-0.01	0.48	0.29	0.59	0.25	-2	-1
Albacore _{comps}	-0.23	-0.06	0.52	0.43	0.65	0.38	0	0
YoY								
HICPX	-0.48	-0.29	0.55	0.46	0.63	0.31	-5	-6
Trimmed Mean 30%	-0.24	-0.07	0.72	0.70	0.72	0.41	-3	-3
HICP excl. energy	-0.21	-0.13	0.73	0.59	0.73	0.36	-4	-2
HICP excl energy & unpr food	-0.26	-0.15	0.70	0.56	0.71	0.34	-4	-2
Supercore	-0.27	-0.11	0.61	0.46	0.64	0.29	-4	-5
PCCI excl. energy	-0.25	-0.04	0.44	0.36	0.44	0.21	0	0
PCCI	-0.09	0.10	0.51	0.46	0.48	0.24	2	2
Albacore _{ranks}	-0.18	-0.01	0.55	0.36	0.55	0.22	-3	-4
Albacore _{comps}	-0.25	-0.06	0.57	0.53	0.60	0.33	-2	1

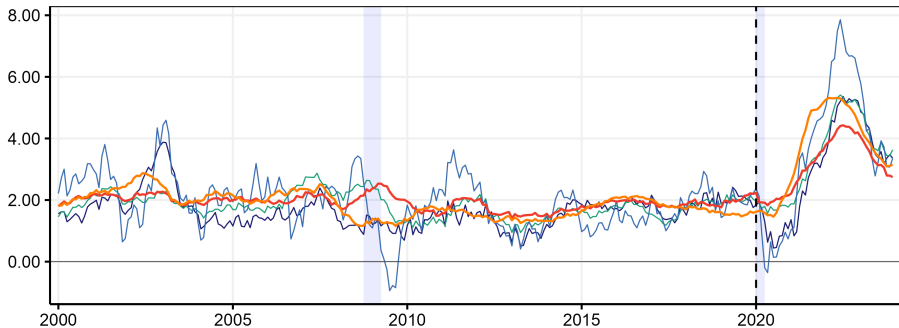
Notes: Cells shaded in green indicate the three best performing models for each criterion. The evaluation is based on the sample ranging from 2000m1 to 2023m12 for the full sample and from 2000m1 to 2019m12 for the pre-Covid case. Bias is determined as the difference between the long-run average of the respective series and HICP. Volatility refers to the standard deviation of each series relative to the standard deviation of headline. Coefficient of variation is the ratio between standard deviation and long-run average of each measure. Positive (negative) numbers for the lead/lag correlation refer to leads (lags) with the highest cross-correlation between each measure and HICP.

Canada Results

	2010m1 - 2019m12			2020m1 - 2023m12		
	$h = 6$	$h = 12$	$h = 24$	$h = 6$	$h = 12$	$h = 24$
Level 3 ($K = 19$)						
Albacore _{comps}	1.05	1.13	1.37	0.68	0.63	0.85
Albacore _{ranks}	1.05	1.13	1.15	0.74	0.74	0.78
Level 4 ($K = 49$)						
Albacore _{comps}	1.12	1.23	1.24	0.54	0.47	0.72
Albacore _{ranks}	1.07	1.15	1.18	0.68	0.67	0.76
Level 5 ($K = 87$)						
Albacore _{comps}	1.10	1.22	1.24	0.69	0.55	0.71
Albacore _{ranks}	1.05	1.12	1.27	0.76	0.71	0.74
Benchmarks						
$X_t^{\text{bm}}, (w_0 = 0)$	1.10	1.34	1.51	0.72	0.78	0.86
$X_t^{\text{bm}+}$	1.00	1.00	1.00	1.00	1.00	1.00
$X_t^{\text{bm}+}, (w_0 = 0)$	1.08	1.24	1.35	0.74	0.80	0.85

Notes: Numbers in **bold** indicate the best model for each *level* and each *horizon* in each of the out-of-sample periods. Numbers highlighted in **green** show the best model per *horizon* and out-of-sample period *across levels*.

Canada



— Core CPI — Trimmed Mean — CPI — Albacore_{ranks} — Albacore_{comps}

