Bottom-Up Inflation Nowcasting with Scanner Data: Unlocking Welfare Gains from High-Frequency Information*

Online Appendix

Günter W. Beck
University of Siegen and Miggroprices

Kai Carstensen
Kiel University

Jan-Oliver Menz
Deutsche Bundesbank

Richard Schnorrenberger
Kiel University

Elisabeth Wieland
Deutsche Bundesbank

9th February 2025

Contents

*Corresponding author: Günter Beck, Universität Siegen, Unteres Schloss 3, 57072 Siegen, Germany, tel.: +49 271 7403223. Email: guenter.beck@uni-siegen.de.

1 Supplementary empirical results

1.1 Detailed nowcasting results of the bottom-up U-MIDAS approach by COICOP-10 item

Table 1: Absolute RMSE of the U-MIDAS model for FMCG product-level inflation

Unprocessed food

Product	COICOP ID	day 7	day 14	day 21	day 28	Product	COICOP ID	day 7	day 14	day 21	day 28
Beef for boiling	0112101100	2.61	1.57	1.58	1.58	Citrus fruits	0116111000	3.92	2.63	2.57	2.58
Roulade or loin of beef	0112107100	2.33	1.80	1.80	1.79	Bananas	0116120100	2.54	2.03	2.02	2.00
Beef	0112107200	2.04	1.32	1.32	1.32	Apples	0116130200	3.54	2.19	2.12	2.09
Minced beef	0112107300	3.32	2.20	2.20	2.15	Pears	0116140100	4.16	3.02	3.02	2.99
Veal	0112109100	1.14	1.02	1.01	1.01	Grapes	0116165100	7.98	5.12	4.93	5.09
Smoked pork chop	0112200100	1.91	1.29	1.25	1.27	Kiwis. melons or the like	0116170000	3.53	2.57	2.56	2.55
Minced pork	0112200200	3.48	2.22	2.21	2.21	Butterhead Lettuce	0117111100	12.49	10.27	9.52	10.30
Roast pork	0112200300	2.30	1.58	1.57	1.57	Lambs lettuce	0117119000	7.09	4.87	4.87	4.88
Pork chop or cutlet	0112200500	2.42	1.75	1.75	1.71	Cauliflower or cabbage	0117121000	7.79	6.04	6.01	5.91
Lamb	0112300100	1.61	0.95	0.97	0.98	Tomatoes	0117131100	10.71	8.52	8.14	8.12
Fresh poultry	0112410100	2.36	1.41	1.40	1.41	Sweet peppers	0117133100	6.73	5.69	5.62	5.74
Frozen poultry	0112410200	2.66	1.66	1.65	1.66	Onion or garlic	0117141100	5.50	3.73	3.55	3.46
Rabbit or game meat	0112500100	1.25	0.91	0.92	0.89	Mushrooms	0117142100	1.76	1.45	1.47	1.46
Liver or other edible offal	0112600100	1.16	0.93	0.93	0.93	Carrots	0117145100	3.93	3.02	3.06	3.06
Eggs	0114701100	1.74	1.20	1.00	0.99	Aspargus or the like	0117149100	7.26	5.53	5.44	5.52

Processed food

Product	COICOP ID	day 7	day 14	day 21	day 28	Product	COICOP ID	day 7	day 14	day 21	day 28
Rice	0111101100	1.42	0.98	1.00	1.00	Frozen vegetables	0117209100	1.36	0.87	0.87	0.88
Flour	0111201100	3.72	2.31	2.34	2.31	Dried vegetables	0117310200	1.17	0.72	0.73	0.72
Semolina	0111203100	2.79	1.60	1.58	1.60	Tinned gherkins	0117321100	1.69	1.04	1.03	1.03
White bread	0111311100	1.06	0.68	0.68	0.68	Tinned sauerkraut	0117323100	1.87	1.43	1.39	1.41
Rye bread or brown bread	0111312100	0.92	0.57	0.57	0.57	Tinned mushrooms	0117324100	1.25	0.77	0.79	0.79
Granary bread	0111313200	1.24	0.77	0.78	0.78	Tinned peas	0117325100	1.63	0.94	0.92	0.92
Ready to bake rolls	0111320200	1.31	0.77	0.78	0.78	Asparagus	0117328400	3.21	2.20	2.25	2.23
Fresh bread rolls	0111320300	1.58	1.02	1.01	0.99	Potatoes	0117401300	5.24	3.91	3.73	3.67
Sponge flan case	0111421100	1.25	0.74	0.74	0.73	Frozen chips or the like	0117402100	1.88	1.55	1.51	1.48
Frozen cake. tart or pie	0111423100	1.37	0.93	0.93	0.94	Potato crisps	0117500200	2.11	1.47	1.48	1.48
Fresh cake. tart or pie	0111424300	1.16	0.69	0.69	0.69	Sugar	0118100100	4.91	4.38	4.16	4.17
Biscuits	0111431200	1.88	1.51	1.51	1.52	Marmalade. jam or jelly	0118201100	1.91	1.60	1.59	1.60
Muffins or waffles	0111433100	1.51	0.86	0.85	0.85	Honey	0118203100	1.44	0.93	0.90	0.91
Crisp bread	0111442100	2.13	1.34	1.33	1.38	Cocoa based spread	0118205100	1.30	1.29	1.27	1.27
Toasted bread	0111444200	1.55	1.17	1.17	1.17	Slab of chocolate	0118301100	2.31	2.11	2.09	2.09
Rusk	0111446100	2.38	1.34	1.36	1.35	Chocolate	0118309100	1.48	1.17	1.17	1.17
Savoury biscuits	0111450100	2.05	1.70	1.70	1.67	Filled chocolates	0118401100	0.78	0.66	0.66	0.66
Pizza or quiches	0111500100	2.13	1.49	1.49	1.49	Boiled sweets	0118405100	0.83	0.66	0.67	0.67
Pasta	0111610100	2.18	1.34	1.34	1.38	Ice cream	0118500100	1.47	1.23	1.21	1.22
Pasta preparations	0111621200	2.51	1.94	1.93	1.93	Sweetener	0118601100	3.66	1.82	1.74	1.76
Oatflakes	0111701100	2.16	1.42	1.41	1.42	Vinegar	0119101100	1.43	0.89	0.89	0.89
Cornflakes and muesli	0111703100	0.98	0.81	0.82	0.81	Mustard	0119102100	2.05	1.19	1.18	1.18
Cake mix	0111801100	1.73	1.11	1.11	1.11	Ketchup	0119103200	2.95	1.70	1.68	1.69
Salami or sausage	0112710200	1.22	0.81	0.80	0.80	Sauce mix	0119103300	1.84	1.56	1.52	1.51
Ham or bacon	0112710300	1.09	0.74	0.75	0.75	Mayonnaise	0119104100	1.96	1.11	1.12	1.11
Lyoner pork sausage	0112721100	2.09	1.48	1.45	1.46	Salt	0119201100	2.56	1.57	1.57	1.57
Fried sausage	0112721200	1.67	1.17	1.14	1.13	Spices	0119203100	0.79	0.50	0.51	0.51
Cold meat	0112721300	2.02	1.46	1.46	1.47	Powdered infant milk	0119302100	0.82	0.52	0.53	0.53
Liver sausage	0112722100	1.42	1.01	1.00	1.01	Food for infants	0119303100	1.18	0.70	0.72	0.72
Tinned sausage	0112723100	1.57	0.91	0.91	0.91	Meat ready meal	0119406100	0.93	0.78	0.78	0.78
Meat based speciality salad	0112801100	1.61	0.92	0.92	0.92	Instant soup	0119911100	1.94	1.50	1.49	1.48
Frozen meat	0112805100	1.69	1.08	1.09	1.09	Tinned soup	0119913100	2.04	1.55	1.55	1.54
Meat-based ready meal	0112807200	1.14	0.73	0.73	0.72	Baking powder	0119930100	3.22	1.71	1.73	1.72
Prepared minced meat	0112808200	2.30	1.54	1.51	1.50	Blancmange powder	0119940100	1.89	1.14	1.12	1.14
Smoked fish	0113500100	1.82	1.24	1.21	1.21	Vitamin tablets or the like	0119990200	0.88	0.62	0.62	0.62
Tinned fish	0113601200	1.25	0.93	0.93	0.91	Pure coffee	0121110300	1.95	1.60	1.56	1.56
Fish marinade	0113602100	1.35	0.82	0.84	0.84	Instant coffee	0121121100	1.40	1.16	1.11	1.11
Fish fingers	0113603000	2.14	1.33	1.29	1.34	Black tea or green tea	0121201100	0.52	0.41	0.41	0.41
Whole milk	0114110100	1.75	1.20	1.15	1.16	Fruit tea or herbal tea	0121203100	0.94	0.79	0.79	0.79
Low fat milk	0114210100	2.02	1.31	1.30	1.31	Cocoa powder	0121300100	1.18	0.87	0.87	0.87
Condensed milk	0114300100	1.45	1.04	1.03	1.03	Sparkling mineral water	0122100100	1.22	0.89	0.88	0.88
Yoghurt	0114400200	1.91	1.45	1.54	1.54	Still mineral water	0122100200	1.18	0.93	0.92	0.92
Hard cheese	0114501100	2.22	1.38	1.39	1.36	Cola drink	0122211100	1.58	1.32	1.34	1.34
Sliced cheese	0114502100	2.35	1.73	1.72	1.73	Soft drink	0122219100	1.76	1.26	1.27	1.27
Soft cheese	0114503100	1.14	0.90	0.89	0.88	Apple juice	0122311100	1.73	1.35	1.28	1.25
Curd	0114507100	4.09	2.96	2.89	2.84	Orange juice	0122312200	1.46	1.03	1.03	1.04
Cream	0114601100	2.79	1.91	1.89	1.90	Multi vitamin juice	0122315100	1.25	0.97	0.96	0.97
Milk based dessert	0114604100	2.20	1.67	1.65	1.66	Vegetable juice	0122320300	1.40	0.79	0.80	0.82
Butter	0115100100	3.68	2.58	2.36	2.37	Liqueur	0211110100	0.59	0.45	0.44	0.45
Margarine	0115201100	2.65	2.14	2.07	2.05	Whisky	0211120100	0.75	0.60	0.60	0.59
Vegetable fat	0115209100	2.54	1.67	1.63	1.63	Brandy or cognac	0211130100	0.64	0.54	0.52	0.51
Sunflower oil	0115400100	6.72	3.26	3.18	3.25	Other spirits	0211140100	0.53	0.42	0.41	0.40
Dried fruit	0116301100	1.68	1.03	1.03	1.03	Red wine or rose wine	0212110200	0.75	0.52	0.52	0.52
Peanuts or trail mix	0116303100	0.95	0.82	0.81	0.81	White wine	0212120100	0.64	0.41	0.41	0.41
Apple sauce	0116401100	2.16	1.35	1.31	1.31	Sparkling wine	0212140100	1.29	1.02	1.03	1.03
Sour cherries	0116402100	2.15	1.81	1.78	1.75	Pils. dark or lager beer	0213100100	1.45	1.14	1.14	1.14
Tinned pineapple	0116403100	2.11	1.45	1.42	1.44	Wheat beer or Altbier	0213200100	1.03	0.87	0.87	0.87
Frozen spinach	0117201100	1.70	1.22	1.22	1.21	Non-alcoholic beer	0213300100	1.23	0.95	0.95	0.95

NEIG

Product	COICOP ID	day 7	day 14	day 21	day 28	Product	COICOP ID	day 7	day 14	day 21	day 28
Baby bottle or the like	0540326100	1.39	0.98	0.98	0.99	Hair spray or gel	1213212100	2.65	1.67	1.66	1.68
Heavy duty detergent	0561101100	1.37	0.91	0.91	0.91	Toothpaste	1213213100	1.36	0.85	0.85	0.85
Mild detergent	0561101200	1.84	1.17	1.17	1.17	Mouthwash or dental floss	1213214100	1.26	0.79	0.79	0.80
Fabric softener or starch	0561101300	1.91	1.17	1.17	1.17	Shaving foam	1213215100	1.07	0.75	0.75	0.75
Dishwashing detergent	0561103100	1.63	1.01	1.00	1.01	Toilet soap	1213216100	2.29	1.44	1.44	1.45
Sanitary cleaner	0561105100	1.39	0.84	0.85	0.85	Shower gel or foam	1213217100	1.50	1.06	1.07	1.07
Glass or furniture cleaner	0561105200	0.84	0.52	0.52	0.52	Toilet tissue	1213221100	1.72	0.98	0.98	0.98
All purpose cleaners	0561105300	1.56	0.97	0.97	0.97	Paper handkerchiefs	1213222100	1.75	0.95	0.94	0.95
Shoe polish	0561107100	1.74	1.10	1.09	1.10	Nappies for babies	1213223200	1.15	0.73	0.73	0.73
Filter paper	0561211100	1.04	0.64	0.63	0.64	Tampons or facial tissues	1213229100	1.67	1.03	1.03	1.03
Aluminium foil	0561211200	2.05	1.06	1.04	1.04	Perfume	1213231100	1.22	1.06	1.05	1.06
Candles	0561241100	1.73	1.16	1.16	1.15	Lipstick or lip care	1213232100	1.21	1.14	1.15	1.14
Scrubbing brushes or brooms	0561291000	0.71	0.47	0.47	0.47	Nail varnish	1213232200	1.14	1.00	0.99	0.99
Melissengeist tonic	0611032100	0.56	0.41	0.42	0.42	Make up	1213232300	1.31	1.18	1.18	1.19
Bird food	0934201200	1.24	0.81	0.81	0.81	Kajal pencil or mascara	1213232400	1.33	1.27	1.31	1.31
Dog food or cat food	0934201400	1.32	0.84	0.85	0.86	Hand cream	1213233100	1.16	0.93	0.93	0.93
Cat litter or bird sand	0934209100	1.10	0.67	0.68	0.68	Day cream or night cream	1213233200	0.95	0.82	0.82	0.82
Non electric toothbrush	1213105200	1.11	0.76	0.76	0.76	Baby cream	1213233300	1.23	0.77	0.77	0.78
Razor blades	1213105300	1.12	0.81	0.81	0.81	Deo spray or deo roll on	1213240100	1.57	1.20	1.18	1.17
Hair shampoo	1213211100	1.56	1.20	1.25	1.21						

Sources: GfK household panel; own calculations.
Notes: The table reports the absolute RMSE values for the U-MIDAS model at nowcasting days 7, 14, 21 and 28 for all COICOP-10 items covered by the GFK:FMCG dataset.

1.2 Nowcasting results of a bottom-up OLS approach

The most basic approach to nowcasting inflation bottom-up with the help of weekly scanner-based indicators avoids mixed frequencies and uses OLS. This *bottom-up OLS approach* aggregates the weekly indicators to the monthly frequency after filling missing weeks with the last available observation. Then it estimates the following regression by OLS for each COICOP-10 item c that is matched by a weekly indicator:

$$\pi_{c,t} = \alpha_c + \beta_c x_{c,t} + \sum_{i=s}^{13} \gamma_{c,s}\, d_{s,t} + \ \varepsilon_{c,t}. \tag{1}$$

The bottom-up aggregation to product groups and headline inflation use the official HICP aggregation weights and thus work exactly as in the bottom-up U-MIDAS approach.

Table 2 shows the RMSE of the bottom-up OLS approach approach relative to the SD-AR benchmark for a selection of COICOP-10 items. Similar to the U-MIDAS approach, the benchmark is partly outperformed by a large margin. A comparison of these results with those of the U-MIDAS approach presented in the main body of our paper does not see a clear winner. For items that are already predicted exceptionally well using U-MIDAS, the OLS approach typically reduces the nowcast loss even further. This particularly applies to most items within the subgroups of unprocessed fruit and vegetables and dairy products and fat. By contrast, whenever the U-MIDAS approach yields only moderate to good improvements over the SD-AR benchmark, the OLS approach is mostly unable to further increase the nowcast precision.

Table 3 shows the nowcast results for the main six product groups and headline inflation, again in comparison to the bottom-up SD-AR benchmark. Similar to the bottom-up U-MIDAS approach, the bottom-up OLS approach improves strongly over the benchmark for the product groups of unprocessed food, processed food, and energy, for which a multitude of highly informative weekly data are available, while it is roughly on par with the benchmark for the product groups of package holidays, NEIG and services. Taken together, this yields headline nowcasts considerably more precise than those produces by the benchmark. A comparison with the bottom-up U-MIDAS approach indicates that the bottom-up OLS approach is slightly superior but the difference is small and statistically not significant.

Our takeaway from this comparison is that, while econometric methods matter to some extent, the major cause for large nowcast gains over the benchmark is rather the inclusion of the highly informative GFK:FMCG data.

Table 2: RMSE for FMCG product-level inflation: bottom-up OLS approach relative to the SD-AR benchmark

Unprocessed fruit and vegetables

	day 7	day 14	day 21	day 28
Sweet peppers	0.484**	0.496**	0.414**	0.431**
Butterhead Lettuce	0.532**	0.5**	0.452**	0.475**
Carrots	0.602**	0.722**	0.647**	0.602**
Grapes	0.672**	0.625**	0.627**	0.654**
Onion or garlic	0.587**	0.747*	0.682**	0.644**
Lambs lettuce	0.746*	0.733*	0.728*	0.736*
Cauliflower or cabbage	0.765**	0.722**	0.729**	0.755**

Processed fruit and vegetables

	day 7	day 14	day 21	day 28
Frozen chips or the like	0.543	0.601	0.51	0.454
Apple sauce	0.781	0.806	0.77	0.762
Tinned mushrooms	0.827	0.866	0.859	0.847
Potatoes	0.774**	0.938	0.944	0.911
Asparagus	0.914	0.898	0.901	0.894
Potato crisps	0.923	0.887*	0.901*	0.909
Frozen spinach	0.94	0.921	0.9	0.895

Unprocessed meat, fish and eggs

	day 7	day 14	day 21	day 28
Eggs	0.728*	0.787	0.738	0.709
Fresh poultry	0.842	0.842	0.825	0.808
Minced pork	0.84	0.834	0.827	0.819
Minced beef	0.778	0.872	0.863	0.858
Beef for boiling	0.828	0.876	0.877	0.877
Smoked pork chop	0.869	0.889	0.878	0.869
Pork chop or cutlet	0.913	0.921	0.918	0.902

Processed meat and fish

	day 7	day 14	day 21	day 28
Ham or bacon	0.65	0.691	0.669	0.655
Salami or sausage	0.711	0.759	0.716	0.695
Liver sausage	0.693	0.758	0.729	0.706
Lyoner pork sausage	0.779	0.833	0.788	0.767
Fish fingers	0.826	0.851	0.804	0.789
Prepared minced meat	0.813	0.87	0.834	0.804
Cold meat	0.857	0.868	0.855	0.848

Beverages and others

	day 7	day 14	day 21	day 28
Baking powder	0.55	0.686	0.682	0.679
Sugar	0.674	0.804	0.702	0.67
Apple juice	0.693	0.807	0.798	0.784
Orange juice	0.744	0.811*	0.776*	0.778
Ketchup	0.908	0.836	0.825	0.814
Pure coffee	0.844*	0.891	0.86*	0.857*
Food for infants	0.828**	0.886**	0.879**	0.866**

Bread and cereals

	day 7	day 14	day 21	day 28
Flour	0.594	0.561	0.55	0.554
Pasta	0.721	0.774*	0.761*	0.767*
Ready to bake rolls	0.735	0.829*	0.8**	0.778**
Rice	0.808	0.819*	0.817	0.81*
Cornflakes and muesli	0.852	0.856	0.831	0.821
Crisp bread	0.823	0.878	0.858	0.861*
Pizza or quiches	0.831	0.897	0.878	0.87

Non-durables

	day 7	day 14	day 21	day 28
Aluminium foil	0.776	0.817	0.797*	0.786*
Toilet tissue	0.82	0.844	0.813	0.792
Hair shampoo	0.744**	0.867**	0.848**	0.832**
Deo spray or deo roll on	0.812*	0.87*	0.85**	0.837**
Hair spray or gel	0.866	0.884*	0.858**	0.854**
Paper handkerchiefs	0.963	0.865	0.835	0.812
Toothpaste	0.888	0.909*	0.899**	0.888**

Dairy products and fat

	day 7	day 14	day 21	day 28
Whole milk	0.354*	0.369	0.316	0.3
Low fat milk	0.363	0.36	0.325	0.302
Condensed milk	0.332	0.376	0.356	0.361
Cream	0.355*	0.367*	0.354*	0.358*
Butter	0.353**	0.394**	0.377**	0.401**
Curd	0.356	0.518	0.468	0.499
Sliced cheese	0.582	0.648	0.606	0.588

Sources: GfK household panel; own calculations.
Notes: The figure shows heatmaps of RMSE values for the bottom-up OLS approach relative to the SD-AR benchmark at nowcasting days 7, 14, 21 and 28 for the best-performing COICOP-10 items within selected FMCG product groups. Results for the Diebold and Mariano (1995) test in the event of outperformance relative to the SD-AR model are indicated by the symbols * (5% level) and ** (1% level).

Table 3: RMSE of headline inflation and its components: bottom-up OLS approach relative to the benchmark approach

Bottom-up OLS-match

	day 7	day 14	day 21	day 28
Unprocessed food	0.677^{*}	0.695^{**}	0.671^{**}	0.667^{**}
Processed food	0.844	0.844	0.799^{*}	0.768^{*}
Energy	0.432^{**}	0.347^{*}	0.323^{*}	0.342^{*}
Package holidays	1.102	1.064	1.009	0.941^{*}
NEIG	1.014	1.003	1.001	0.998
Services	1	1	1	1
Headline	0.797	0.727	0.683	0.674

Sources: GfK household panel; European Commission's Weekly Oil Bulletin; AMADEUS; own calculations.
Note: The figure shows heatmaps of RMSEs for nowcasts based on the bottom-up OLS approach with aggregation via HICP weights relative to the benchmark approach, which is a bottom-up nowcast based on SD-AR models fitted at the COICOP-10 level. Results for the Diebold and Mariano (1995) test in the event of outperformance relative to the benchmark are indicated by the symbols * (5% level) and ** (1% level).

2 A horserace of various machine learning approaches to nowcast product-group inflation

2.1 The econometric models

A natural starting point for nowcasting inflation rates at the product-group level is to treat all weekly indicators belonging to a product group as relevant predictors only for this group. As the U-MIDAS setting is not suited to handling such a large set of predictors, we resort to machine learning methods.

We try two modeling approaches. Our first approach avoids mixed frequencies by first aggregating the weekly GFK:FMCG series to the monthly frequency and then applying standard shrinkage methods to estimate nowcasting models for the group-specific target.[1] Specifically, we use the LASSO, the ridge, and the elastic net estimators.[2]

From a mathematical standpoint, let us assume a set of monthly aggregated predictors $\mathbf{x}_t = (x_{1t}, \ldots, x_{qt})'$ such that $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_t)'$, where q denotes an abundant number of COICOP-10 series underlying a given group-specific target. Hence, conditional on official inflation data available at t, we model our group-specific target $\pi^{\mathbf{M}} = (\pi^M_{g,1}, \ldots, \pi^M_{g,t})'$ as a function of $\mathbf{X} = (\iota, \mathbf{x})$ using standard shrinkage methods such as LASSO, ridge and elastic net regression, where ι is a t-dimensional vector of

[1] Applying penalized U-MIDAS regressions to the large set of predictors defined at the weekly frequency (four weekly series for each predictor) is also feasible; however, this approach does not recognize serial dependence across high-frequency lags and thereby may be subject to random selection. Zhao and Yu (2006) show that LASSO selects the true model consistently if and (almost) only if the irrelevant covariates are not highly correlated with the predictors in the true model ("irrepresentable condition").

[2] We use the elastic net without tuning the relative weights of the L1 and L2 norms. Instead, we impose equal weights.

ones. The hybrid elastic net estimator solves the following penalized least squares problem:

$$\hat{\beta} = \min_{\hat{\beta}} ||\pi^{\mathbf{M}} - \mathbf{X}\beta||^2 + \lambda \left(\alpha \, |\beta|_1 + \frac{(1-\alpha)}{2} \, ||\beta||^2 \right), \tag{2}$$

where $\alpha \in (0, 1]$ is a weight parameter that interpolates between LASSO ($\alpha = 1$) and ridge regression (as $\alpha \to 0$) while the regularization parameter λ controls the amount of shrinkage in β. Hence, the idea is to shrink coefficient estimates $\hat{\beta}$ to or towards zero if the c-th COICOP-10 series is not relevant. Finally, we construct the monthly aggregated estimate of $\mathbf{x}$ using the latest contemporaneous weekly GFK:FMCG information and compute the nowcast for $\pi^M_{g,t+1}$ based on the estimates $\hat{\beta}$.

The second approach constructs a nowcast using the sg-LASSO-MIDAS framework (see Babii, Ghysels, and Striaukas, 2022) that handles high-dimensional mixed-frequency prediction problems. Let the matrix of covariates now be defined as:

$$\mathbf{X} = (\iota, \mathbf{X_1^{(m)}} W, \ldots, \mathbf{X_q^{(m)}} W), \tag{3}$$

where $\mathbf{X_j^{(m)}} = (X^{(m)}_{c1}, \ldots, X^{(m)}_{ct})'$ is a $t \times m$ matrix of the c-th high-frequency covariate (weekly GFK:FMCG series) and W denotes a predetermined $m \times L$ matrix of weights based on Legendre polynomials of degree L that aggregate over the high-frequency lags. Then, the sg-LASSO estimator solves the penalized least squares problem:

$$\hat{\beta} = \min_{\hat{\beta}} ||\pi^{\mathbf{M}} - \mathbf{X}\beta||^2 + 2\lambda \left(\alpha \, |\beta|_1 + (1-\alpha) \, ||\beta||_{2,1} \right), \tag{4}$$

where $||\beta||_{2,1} = \sum_{G \in \mathcal{G}} |\beta_G|_2$ is the group LASSO norm for a group structure $\mathcal{G}$ that hereby constitutes all high-frequency lags of a single covariate. Thus, in this case, $\alpha \in [0, 1]$ determines the relative importance of LASSO sparsity and the group structure.

Note that sg-LASSO has the advantage of performing shrinkage in a mixed-frequency rather than a low-frequency setting by recognizing serial dependence across different high-frequency lags, also taking into account the time series nature of the data. Hence, model (4) promotes sparsity between and within COICOP-10 items, allowing us not only to select the relevant COICOP-10 series but also the appropriate lag structure of each item. We use a Legendre polynomial of degree $L = 0$ which attributes equal weights to all high-frequency lags and delivers similar results compared to other choices of L but at a lower computational cost (see Appendix **??**).

The tuning parameters of the above machine learning approaches are determined in a data-driven manner using cross-validation to obtain optimal prediction performance. We tune the hyperparameters of sg-LASSO via expanding cross-validation splitting the in-sample data into $k = 5$ folds and tests on the $k + 1^{\text{th}}$ fold so that it accounts for the time series nature of the data, although it only uses the end of the sample as the test set. Cross-validation of the standard shrinkage methods (LASSO, ridge and elastic net) also uses a training split of $k = 5$ folds but hereby assumes independent and identically distributed samples, which is also valid in a time series context provided the models yield uncorrelated errors (Bergmeir, Hyndman, and Koo, 2018). Finally, to evaluate the nowcast precision of these machine learning approaches based on weekly GFK:FMCG information, we fit SD-AR benchmark models directly to the group-specific target $\pi^M_{g,t}$.

2.2 Empirical results

This section presents the results of the nowcasting exercise described in Appendix 2. We use machine learning shrinkage methods to estimate direct nowcasting models for high-level product groups (and some subgroups) that are matched by weekly GFK:FMCG price indicators at the COICOP-10 level. Table 4 compares the RMSE of these models relative to the SD-AR benchmark fitted to the group-specific inflation rates.

The top panel refers to the three high-level product groups. With regard to unprocessed and processed food, the shrinkage models significantly outperform the benchmark with reductions in the RMSE between 15% and 25% on all nowcasting days. By contrast, for NEIG the benchmark dominates. This outcome likely reflects the different coverage rates of the GFK:FMCG data across product groups, whereas 30 out of 38 COICOP-10 items of unprocessed food and 116 out of 142 COICOP-10 items of processed food are matched, but only 39 out of the 302 NEIG with semi-durables and durables almost lacking completely. In addition, even the relatively few matched NEIG items, mostly non-durables, do not correlate very strongly with their HICP counterparts.

Table 4: RMSE for FMCG product-group inflation: Shrinkage methods relative to the SD-AR benchmark

Unprocessed food

	day 7	day 14	day 21	day 28
LASSO	0.788*	0.788**	0.755**	0.741**
Ridge	0.838*	0.832**	0.814**	0.805**
Elastic Net	0.796*	0.784**	0.754**	0.755**
sg-LASSO	0.788*	0.775**	0.778**	0.794**

Processed food

	day 7	day 14	day 21	day 28
LASSO	0.846	0.861	0.805*	0.81*
Ridge	0.865	0.875	0.824*	0.811*
Elastic Net	0.855	0.854	0.811*	0.82*
sg-LASSO	0.829	0.779	0.774*	0.781*

Non-energy industrial goods

	day 7	day 14	day 21	day 28
LASSO	1.093	1.075	1.081	1.083
Ridge	1.134	1.081	1.076	1.069
Elastic Net	1.088	1.084	1.073	1.095
sg-LASSO	1.069	1.073	1.08	1.081

(a) High-level product groups

Unproc. fruit and vegetables

	day 7	day 14	day 21	day 28
LASSO	0.658*	0.651**	0.623**	0.62**
Ridge	0.788*	0.81**	0.794**	0.79**
Elastic Net	0.658*	0.656**	0.626**	0.63**
sg-LASSO	0.722*	0.651**	0.643**	0.668**

Unproc. meat, fish and eggs

	day 7	day 14	day 21	day 28
LASSO	0.751	0.812	0.801	0.796
Ridge	0.763	0.812	0.808	0.79
Elastic Net	0.761	0.817	0.818	0.794
sg-LASSO	0.743	0.812	0.802	0.784

Processed fruit and vegetables

	day 7	day 14	day 21	day 28
LASSO	0.785	0.942	0.916	0.899
Ridge	0.83	0.939	0.911	0.88
Elastic Net	0.784	0.928	0.902	0.881
sg-LASSO	0.779	0.94	0.902	0.878

Processed meat and fish

	day 7	day 14	day 21	day 28
LASSO	0.675	0.723	0.675	0.638
Ridge	0.673	0.73	0.678	0.651
Elastic Net	0.674	0.728	0.665	0.631
sg-LASSO	0.622	0.711	0.68	0.668

Bread and cereals

	day 7	day 14	day 21	day 28
LASSO	1.004	0.912	0.856	0.827
Ridge	1.061	0.965	0.885	0.856
Elastic Net	0.999	0.914	0.864	0.828
sg-LASSO	0.967	0.82	0.829	0.841

Dairy products and fat

	day 7	day 14	day 21	day 28
LASSO	0.463*	0.521*	0.455*	0.442*
Ridge	0.563*	0.595*	0.534*	0.503*
Elastic Net	0.484*	0.52*	0.45*	0.446*
sg-LASSO	0.47*	0.531*	0.514*	0.507*

Beverages and others

	day 7	day 14	day 21	day 28
LASSO	0.963	0.953	0.93	0.905
Ridge	0.949	0.957	0.926	0.919
Elastic Net	0.953	0.952	0.929	0.917
sg-LASSO	0.933	0.908	0.914	0.917

Non-durables

	day 7	day 14	day 21	day 28
LASSO	1.076	1.026	0.995	0.977
Ridge	1.064	1.041	1.001	0.981
Elastic Net	1.079	1.028	0.994	0.983
sg-LASSO	1.048	1.013	1.004	1.004

(b) Subgroups

Sources: GfK household panel; own calculations.
Notes: The figure shows heatmaps of RMSEs for nowcasts based on shrinkage machine learning methods relative to the SD-AR benchmark for product groups (and subgroups) matched by GFK:FMCG data. Results for the Diebold and Mariano (1995) test in the event of outperformance relative to the benchmark are indicated by the symbols * (5% level) and ** (1% level).

The bottom panel in Table 4 shows the results for the eight more disaggregated subgroups. In the large majority of cases, it pays off considerably to use shrinkage models that include weekly GFK:FMCG data. The advantage is particularly large for dairy products and fat (reduction in RMSE of roughly 45% to 55%), unprocessed fruit and vegetables (reduction of around 20% to almost 40%), processed meat and fish (reduction of more than 25% to almost 40%), and unprocessed meat, fish and eggs (reduction of nearly 20% to 25%). For processed fruit and vegetables, bread and cereals, and beverages and others, the nowcasting gains are more muted, but it is still generally beneficial to use shrinkage models. Only in the case of non-durables, there is no clear difference to the benchmark, which likely again reflects the low correlation of the GFK:FMCG items at the COICOP-10 level with their HICP counterparts.

The weekly flow of information affects the nowcasting performance in a way very similar to what is reported at the COICOP-10 level. Most importantly, the information available on day 7 of a given month already turns out to be highly valuable. This likely reflects the fact that at day 7 of a month t, the benchmark model includes only official inflation rates of month $t-2$, while the shrinkage approaches use the full GFK:FMCG data of month $t-1$ and the first week of month t. The additional information exploited at day 14 typically further improves the nowcasts in absolute terms, whereas this is not always the case relative to the benchmark, which on that day includes the official inflation rates of month $t-1$. Finally, the additional information gained in weeks 3 and 4 of a month is of minor qualitative importance.

Concerning the different shrinkage approaches, the nowcasting results do not favor a single method. The general conclusion is that it is important to include the weekly GFK:FMCG dataset and make it usable in an appropriate way. To this end, standard shrinkage methods (LASSO, ridge and elastic net) work generally as well as the sg-LASSO approach, which performs variable selection in a mixed-frequency setting and fully accounts for the time series nature of the dataset. Nevertheless, for processed food, which is the product group with by far the largest number of underlying COICOP-10 items that we match with GFK:FMCG data, the sg-LASSO is clearly superior. This may indicate that this approach is especially promising when it comes to very high-dimensional estimation and nowcasting settings.

2.3 Robustness

We examine whether the product group-specific nowcasting results hold irrespective of our hyperparameter choices for the estimation of shrinkage methods. We start by increasing the degree of the sg-LASSO Legendre polynomial from $L = 0$ to $L = \{1, 2\}$. Our findings suggest a slight improvement in the precision of the nowcast in only a small number of cases where statistically significant outperformance is already achieved by sg-LASSO with $L = 0$ compared to SD-AR. Hence, it represents the optimal choice given that it promotes a higher dimensionality reduction and carries a smaller computational burden when estimating sg-LASSO coefficients. Similarly, we test for different folds of the cross-validation considering the grid set $k \in \{5, 10, 15, 20, 25\}$. Overall, these choices favor similar tuning parameters and model architectures, thus not altering the results for group-specific targets.

3 Model simulation: Detailed model solution

3.1 Model setup

For our simulation exercise, we employ a standard New-Keynesian model closely following Galí (2015). The demand side is given by the IS equation:

$$x_t = -\frac{1}{\sigma}\left(i_t - E_t\left\{\pi_{t+1}\right\} - r^*\right) + E_t\left\{x_{t+1}\right\} + g_t, \tag{5}$$

where i_t denotes the nominal interest rate, x_t denotes the output gap in period t and π_t represents period t's inflation rate. r^* corresponds to the deterministic steady-state level of the natural interest rate, given by $r^* = \frac{1}{\beta} - 1$ with β denoting the subjective discount factor. g_t is a demand shock. The expression $E_t\left\{\bullet\right\}$ is used to denote the (rational) expectation of a variable given the information available to the policymaker and the public in period t. $\frac{1}{\sigma}$ denotes the intertemporal elasticity of substitution.

The supply side of the economy is represented by the following forward-looking Phillips curve:

$$\pi_t = \kappa x_t + \beta E_t\left\{\pi_{t+1}\right\} + u_t. \tag{6}$$

u_t represents a cost-push shock. The slope of the Phillips curve, κ, is given by:

$$\kappa = \lambda\left(\frac{\sigma\left(1-\alpha\right) + \varphi + \alpha}{1-\alpha}\right) \tag{7}$$

with $\lambda = \frac{(1-\beta\theta)(1-\theta)}{\theta} * \Theta$ and $\Theta = \frac{1-\alpha}{1-\alpha+\alpha\varepsilon}$. $1-\alpha$ denotes the elasticity of output with respect to labor, $\frac{1}{\varphi}$ represents the Frisch elasticity of labor supply, θ denotes the share of firms not adjusting prices in a given period and ε represents the elasticity of substitution between varieties.

The demand and cost-push shocks are modeled as AR(1) processes, as follows:

$$g_t = \rho_g g_{t-1} + \eta_t^g \tag{8}$$

$$u_t = \rho_u u_{t-1} + \eta_t^u, \tag{9}$$

with $0 < \rho_g, \rho_u < 1$ and $\eta_t^g \sim$ i.i.d. $\mathrm{N}\left(0, \sigma_{\eta_g}^2\right)$ and $\eta_t^u \sim$ i.i.d. $\mathrm{N}\left(0, \sigma_{\eta_u}^2\right)$.

Assuming discretion, the policymaker minimizes the following loss function:

$$L_t = \frac{1}{2}\left(\pi_t^2 + \delta x_t^2\right), \tag{10}$$

where δ represents the relative weight that the policymaker places on output stabilization versus inflation stabilization. We assume that the inflation target of the policymaker, π^*, is set equal to zero.

3.2 Model solution under full information

In the case of full information, the true sizes of the demand and cost-push shocks are known to the policymaker and the public, i.e., we have

$$g_t = E_t\{g_t\} \qquad \text{and} \qquad u_t = E_t\{u_t\}. \tag{11}$$

The policymaker's objective is to maximize the loss function

$$\min_{i_t} \frac{1}{2} E_t\left[\pi_t^2 + \delta x_t^2\right] \tag{12}$$

s.t. the IS curve (5) and the Phillips curve (6).

The associated Lagrange function is given by:

$$\begin{aligned} L \;=\; & E_t\Bigg\{\frac{1}{2}\left[\pi_t^2 + \delta x_t^2\right] + \\ & +\mu_t\left[-\frac{1}{\sigma}\left(i_t - E_t\{\pi_{t+1}\} - r^*\right) + E_t\{x_{t+1}\} + g_t - x_t\right] + \\ & + \nu_t\left[\kappa x_t + \beta E_t\{\pi_{t+1}\} + u_t - \pi_t\right]\Bigg\}. \end{aligned} \tag{13}$$

The first-order equations can be derived as follows:

- With respect to x_t:

$$\delta x_t - \nu_t + \nu_t \kappa = 0 \tag{14}$$

- With respect to π_t:

$$\pi_t - \nu_t = 0 \tag{15}$$

- With respect to i_t:

$$-\nu_t \frac{1}{\sigma} = 0 \tag{16}$$

- With respect to ν_t:

$$x_t = -\frac{1}{\sigma}\left(i_t - E_t\{\pi_{t+1}\} - r^*\right) + E_t\{x_{t+1}\} + g_t \tag{17}$$

- With respect to ν_t:

$$\pi_t = \kappa x_t + \beta E_t\{\pi_{t+1}\} + u_t \tag{18}$$

From (16), we see that

$$\nu_t = 0. \tag{19}$$

Plugging (15) into (14), using $\nu_t = 0$ and rearranging, we obtain:

$$x_t = -\frac{\kappa}{\delta}\pi_t. \tag{20}$$

Plugging this result into equation (18) and collecting terms gives:

$$\pi_t = \frac{\delta\beta}{\delta + \kappa^2} E_t \{\pi_{t+1}\} + \frac{\delta}{\delta + \kappa^2} u_t. \tag{21}$$

Solving this equation forward (and assuming that $\lim_{i \to \infty} \left(\frac{\delta\beta}{\delta+\kappa^2}\right)^i E_t \{\pi_{t+i+1}\} = 0$), we obtain for the inflation rate in the case of full information:

$$\pi_t^{full\_inf} = \frac{\delta}{\kappa^2 + \delta(1 - \beta\rho_u)} u_t = \delta q u_t \tag{22}$$

with

$$q \equiv \frac{1}{\kappa^2 + \delta(1 - \beta\rho_u)}. \tag{23}$$

Plugging this result into equation (20), gives the following solution for the output gap under full information:

$$x_t^{full\_inf} = -\frac{\kappa}{\kappa^2 + \delta(1 - \beta\rho_u)} u_t = -\kappa q u_t. \tag{24}$$

Then:

$$E_t \left\{\pi_{t+1}^{full\_inf}\right\} = \delta q E_t \{u_{t+1}\} = \delta q \rho_u u_t. \tag{25}$$

and

$$E_t \left\{x_{t+1}^{full\_inf}\right\} = -\kappa q E_t \{u_{t+1}\} = -\kappa q \rho_u u_t. \tag{26}$$

Using the IS equation, the interest rate that the policymaker will chose under full information, given by $i_t^{full\_inf}$, can be derived as follows:

$$\begin{aligned} i_t^{full\_inf} &= E_t \{\pi_{t+1}\} + r^* + \sigma \left(-x_t + E_t \{x_{t+1}\} + g_t\right) = \\ &= \delta q \rho_u u_t + r^* + \sigma \left(\kappa q u_t - \kappa q \rho_u u_t\right) + \sigma g_t = \\ &= \delta q \rho_u u_t + r^* + \sigma\kappa \left(1 - \rho_u\right) q u_t + \sigma g_t = \\ &= r^* + \left(1 + \frac{\sigma\kappa(1-\rho)}{\delta\rho_u}\right) \delta q \rho_u u_t + \sigma g_t = \\ &= r^* + \gamma_\pi \delta q \rho_u u_t + \sigma g_t \end{aligned}$$

with

$$\gamma_\pi = 1 + \frac{\sigma\kappa(1 - \rho_u)}{\delta\rho_u}. \tag{27}$$

If $\delta = 0$ (case of strict inflation targeting), this expression can be rewritten as:

$$i_t^{full\_inf} = r^* + \left(\frac{\sigma\left(1 - \rho_u\right)}{\kappa}\right) u_t + \sigma g_t, \tag{28}$$

given that

$$
\begin{aligned}
&\gamma_\pi \delta q \rho_u u_t + \sigma g_t = \\
\left(1+\frac{\kappa\sigma(1-\rho_u)}{\delta\rho_u}\right)\left(\frac{1}{\kappa^2+\delta(1-\beta\rho)}\right)&\delta\rho_u u_t + \sigma g_t = \\
\left(\delta+\frac{\kappa\sigma(1-\rho_u)}{\rho_u}\right)\left(\frac{1}{\kappa^2+\delta(1-\beta\rho_u)}\right)&\rho_u u_t + \sigma g_t \underset{\delta=0}{=} \\
\left(\frac{\kappa\sigma(1-\rho_u)}{\rho_u}\right)\left(\frac{1}{\kappa^2}\right)&\rho_u u_t + \sigma g_t = \\
\left(\frac{\sigma\,(1-\rho_u)}{\kappa}\right)&u_t + \sigma g_t.
\end{aligned}
$$

If $\delta > 0$ (case of flexible inflation targeting), the expression for the optimal interest rate under full information can be also written as:

$$
i_t^{full\_inf} = r^* + \gamma_\pi \delta q \rho_u u_t + \sigma g_t = r^* + \gamma_\pi \rho_u \pi_t + \sigma g_t = r^* + \gamma_\pi E_t\left\{\pi_{t+1}\right\} + \sigma g_t. \quad (29)
$$

3.3 Model solution under incomplete information

Modeling demand and cost-push shock uncertainty

We assume that the policymaker cannot observe demand and cost-push shocks with any certainty and needs to rely on estimates. We use the expectation operator E_t to refer to the policymaker's estimates or perceptions of such unobservable variables. Thus, $E_t\{g_t\}$ corresponds to the policymaker's estimate of the demand shock in period t, given the information available at that point in time and $E_t\{u_t\}$ refers to the corresponding estimate of the cost-push shock in period t. We assume that these perceptions represent the best available estimates of the unobservable variables from the perspective of the policymaker. Similarly, we use E_t to denote the best available estimate or forecast of output and inflation. For example, $E_t\{\pi_t\}$ represents the policymaker's best forecast of inflation at the point in period t when it decides on its policy, i.e., before it can observe the joint consequences of demand, cost-push and its policy choice on inflation.

Fortunately, the optimal policy under uncertainty can be determined in a straightforward manner, given that our model framework fulfills the following conditions: the model is linear, the parameters are known and uncertainty is additive. In this case, certainty-equivalence applies, i.e., the optimal policy must satisfy the decision maker's first-order conditions, equations (14) to (18), in expectation (see, for example, Svensson and Woodford, 2003).

Formally, we model perceived shocks, $E_t\{g_t\}$ and $E_t\{u_t\}$ as follows (see Beck and Wieland, 2008):

$$
g_t = E_t\{g_t\} + \varepsilon_t^g, \text{ with } \varepsilon_t^g \sim \text{i.i.d. N}\left(0, \sigma_{\varepsilon_g^2}\right) \quad (30)
$$

$$
u_t = E_t\{u_t\} + \varepsilon_t^u, \text{ with } \varepsilon_t^u \sim \text{i.i.d. N}\left(0, \sigma_{\varepsilon_u^2}\right). \quad (31)
$$

ε_t^g and ε_t^u denote the nowcast error of the policymaker with respect to the size of the demand and cost-push shock in period t, respectively. $\sigma_{\varepsilon_g^2}$ and $\sigma_{\varepsilon_u^2}$ determine the degree of uncertainty of the demand and cost-push shock nowcasts.

Model solution

Since certainty-equivalence applies, we can replace the true shock values in the decision maker's first-order conditions by their perceived values and solve for the realized values of the choice variables following the steps outlined above (Section 3.2) in an analogous manner. For the optimal interest rate under incomplete information, we then obtain:

$$\begin{aligned} i_t^{incomplete\_inf} &= r^* + \gamma_\pi \kappa q \rho_u E_t\{u_t\} + \sigma E_t\{g_t\} = \\ &= r^* + \gamma_\pi E_t\{\pi_t\} + \sigma E_t\{g_t\} \end{aligned} \tag{32}$$

with γ_π given by equation (27).

Alternatively, the optimal interest rate chosen under incomplete information can also be expressed as:

$$\begin{aligned} i_t^{incomplete\_inf} &= r^* + \gamma_\pi \delta q \rho_u E_t\{u_t\} + \sigma E_t\{g_t\} = \\ &= r^* + \gamma_\pi \delta q \rho_u (u_t - \varepsilon_t^u) + \sigma (g_t - \varepsilon_t^g) = \\ &= r^* + \gamma_\pi \delta q \rho_u u_t + \sigma g_t - \gamma_\pi \delta q \rho_u \varepsilon_t^u - \sigma \varepsilon_t^g = \\ &= i_t^{full\_inf} - \gamma_\pi \delta q \rho_u \varepsilon_t^u - \sigma \varepsilon_t^g. \end{aligned} \tag{33}$$

The intended values for the output gap is given by:

$$E_t\{x_t\} = -\frac{\kappa}{\kappa^2 + \delta(1-\beta\rho)} E_t\{u_t\} = -\kappa q E_t\{u_t\}. \tag{34}$$

Plugging the interest rate chosen under uncertainty into the IS curve, the actual value for the output gap can be derived as follows:

$$\begin{aligned} x_t^{incomplete\_inf} &= -\frac{1}{\sigma}\left(i_t^{incomplete\_inf} - E_t\{\pi_{t+1}\} - r^*\right) + E_t\{x_{t+1}\} + g_t \\ &= -\frac{1}{\sigma}\left(i_t^{full\_inf} - \gamma_\pi \delta q \rho_u \varepsilon_t^u - \sigma \varepsilon_t^g - \delta q \rho_u E_t\{u_t\} - r^*\right) - \kappa q \rho_u E_t\{u_t\} + g_t \\ &= -\frac{1}{\sigma}\left(i_t^{full\_inf} - \underbrace{\delta q \rho_u u_t}_{=E_t\left\{\pi_{t+1}^{full\_inf}\right\}} - r^*\right) \underbrace{-\kappa q \rho_u u_t}_{=E_t\left\{x_{t+1}^{full\_inf}\right\}} + g_t - \frac{1}{\sigma}\left(-\gamma_\pi \delta q \rho_u \varepsilon_t^u - \sigma \varepsilon_t^g + \delta q \rho_u \varepsilon_t \right. \\ &= x_t^{full\_inf} + \left(\frac{1}{\sigma}\gamma_\pi \delta q \rho_u - \frac{1}{\sigma}\delta q \rho_u + \kappa q \rho_u\right)\varepsilon_t^u + \varepsilon_t^g \\ &= x_t^{full\_inf} + \kappa q \varepsilon_t^u + \varepsilon_t^g, \end{aligned}$$

where we used that $E_t\left\{\pi_{t+1}^{full\_inf}\right\} = \delta q \rho_u E_t\{u_t\}$ and $E_t\left\{x_{t+1}^{full\_inf}\right\} = -\kappa q \rho_u E_t\{u_t\}$

(see equations (26) and (25)) and the fact that

$$\begin{aligned}
\frac{1}{\sigma}\gamma_\pi\delta q\rho_u - \frac{1}{\sigma}\delta q\rho_u + \kappa q\rho_u &= q\left[\frac{1}{\sigma}\delta\rho_u\left(\gamma_\pi - 1\right) + \kappa\rho_u\right] \\
&= q\left[\frac{1}{\sigma}\delta\rho_u\left(1 + \frac{\kappa\sigma(1-\rho_u)}{\delta\rho_u} - 1\right) + \kappa\rho_u\right] \\
&= q\left[\frac{1}{\sigma}\delta\rho_u\left(\frac{\kappa\sigma(1-\rho_u)}{\delta\rho_u}\right) + \kappa\rho_u\right] \\
&= \kappa q.
\end{aligned}$$

The latter expression can also be written as:

$$\begin{aligned}
x_t^{incomplete_inf} &= x_t^{full_inf} + \kappa q\varepsilon_t^u + \varepsilon_t^g = && (35) \\
&= -\kappa q u_t + \kappa q\varepsilon_t^u + \varepsilon_t^g = \\
&= -\kappa q E_t\left\{u_t\right\} + \varepsilon_t^g. && (36)
\end{aligned}$$

The intended value for inflation is given by:

$$E_t\left\{\pi_t\right\} = \frac{\delta}{\kappa^2 + \delta(1-\beta\rho)}E_t\left\{u_t\right\} = \delta q E_t\left\{u_t\right\} \tag{37}$$

Plugging the expression obtained for actual output under incomplete information into the Phillips curve, the actual value for the inflation rate under incomplete information can be derived as follows:

$$\begin{aligned}
\pi_t^{incomplete_inf} &= \kappa x_t^{incomplete_inf} + \beta E_t\left\{\pi_{t+1}\right\} + u_t = \\
&= \kappa\left[-\kappa q E_t\left\{u_t\right\} + \varepsilon_t^g\right] + \beta\delta q\rho_u E_t\left\{u_t\right\} + u_t = \\
&= \left[-\kappa^2 + \beta\delta\rho\right] q E_t\left\{u_t\right\} + u_t + \kappa\varepsilon_t^g = \\
&= \left[-\kappa^2 + \beta\delta\rho\right] q E_t\left\{u_t\right\} + \left[\kappa^2 + \delta - \beta\delta\rho)\right] q u_t + \kappa\varepsilon_t^g = \\
&= \delta q u_t + \left[\kappa^2 - \beta\delta\rho\right] q\left[u_t - E_t\left\{u_t\right\}\right] + \kappa\varepsilon_t^g = \\
&= \delta q\left[E_t\left\{u_t\right\} + \varepsilon_t^u\right] + \left[\kappa^2 - \beta\delta\rho\right] q\left[u_t - E_t\left\{u_t\right\}\right] + \kappa\varepsilon_t^g = \\
&= \delta q E_t\left\{u_t\right\} + \left[\kappa^2 - \beta\delta\rho\right] q\left[u_t - E_t\left\{u_t\right\}\right] + \delta q\varepsilon_t^u + \kappa\varepsilon_t^g = \\
&= \delta q E_t\left\{u_t\right\} + q\left[\kappa^2 - \beta\delta\rho_u + \delta\right]\varepsilon_t^u + \kappa\varepsilon_t^g = \\
&= E_t\left\{\pi_t\right\} + \varepsilon_t^u + \kappa\varepsilon_t^g. && (38)
\end{aligned}$$

Alternatively, the following expression can be derived:

$$\begin{aligned}
\pi_t^{incomplete_inf} &= \delta q E_t\left\{u_t\right\} + \varepsilon_t^u + \kappa\varepsilon_t^g = && (39) \\
&= \delta q u_t - \delta q\varepsilon_t^u + \varepsilon_t^u + \kappa\varepsilon_t^g = \\
&= \pi_t^{full_inf} + \left(1 - \delta q\right)\varepsilon_t^u + \kappa\varepsilon_t^g.
\end{aligned}$$

4 Model simulation: Supplementary simulation results

On the following pages, supplementary simulation outcomes are presented. Specifically, the following results are presented:

a) **Table 5 Variances of inflation and the output gap: Low weight on the output gap ($\delta = \frac{0.02}{3}$), all nowcast scenarios**: Variances of inflation and the output gap obtained for the case of the calibration using default values and a low weight for output stabilization in the loss function employing the forecast results for all model specifications used in the empirical part (SD AR forecast, bottom-up OLS-match, bottom-up U-MIDAS, direct ML). These results supplement numbers reported in Table E5 of the main paper which only focuses on outcomes obtained for the SD AR forecast and Bottom-up U-MIDAS models.

b) **Table 6 Variances of inflation and the output gap: High weight on the output gap ($\delta = \frac{0.25}{3}$), all nowcast scenarios**: Variances of inflation and the output gap obtained for the case of the calibration using default values and a high weight for output stabilization in the loss function employing the forecast results for all model specifications used in the empirical part (SD AR forecast, bottom-up OLS-match, bottom-up U-MIDAS, direct ML). These results supplement numbers reported in Table E6 of the main paper which only focuses on outcomes obtained for the SD AR forecast and Bottom-up U-MIDAS models.

Table 5: Variances of inflation and the output gap: Low weight on the output gap ($\delta = \frac{0.02}{3}$), all nowcast scenarios

Forecast scenario	Week 1	Week 2	Week 3	Week 4
Variance of inflation				
Medium economic uncertainty scenario				
Baseline scenario (SD AR forecast)	5.76	4.71	4.71	4.71
Bottom-up OLS-match	4.94	4.29	4.28	4.27
Bottom-up U-MIDAS	4.99	4.31	4.28	4.28
Direct ML	5.08	4.32	4.34	4.27
High economic uncertainty scenario				
Baseline scenario (SD AR forecast)	9.62	8.29	8.29	8.29
Bottom-up OLS-match	7.44	6.82	6.8	6.84
Bottom-up U-MIDAS	7.75	6.87	6.82	6.84
Direct ML	8.6	7.12	7.21	6.95
Low economic uncertainty scenario				
Baseline scenario (SD AR forecast)	2.05	1.68	1.68	1.68
Bottom-up OLS-match	1.72	1.62	1.6	1.59
Bottom-up U-MIDAS	1.72	1.62	1.6	1.57
Direct ML	1.66	1.59	1.58	1.58
Variance of the output gap				
Medium economic uncertainty scenario				
Baseline scenario (SD AR forecast)	97.21	70.4	70.4	70.4
Bottom-up OLS-match	76.3	59.63	59.27	59.05
Bottom-up U-MIDAS	77.48	60.05	59.34	59.25
Direct ML	79.71	60.28	60.82	59.09
High economic uncertainty scenario				
Baseline scenario (SD AR forecast)	162.98	128.8	128.8	128.8
Bottom-up OLS-match	107.17	91.04	90.65	91.73
Bottom-up U-MIDAS	114.95	92.58	91.28	91.69
Direct ML	136.9	98.75	101.2	94.53
Low economic uncertainty scenario				
Baseline scenario (SD AR forecast)	33.91	24.46	24.46	24.46
Bottom-up OLS-match	25.61	22.86	22.33	22.09
Bottom-up U-MIDAS	25.66	22.88	22.33	21.74
Direct ML	23.99	22.32	21.98	22.03

Notes: Table 5 reports variances of the inflation rate and the output gap obtained for the case of a low weight ($\delta = \frac{0.02}{3}$) on output stabilization when nowcast error statistics from both the baseline model (SD-AR model) and all scanner-data based nowcast models are used.

Table 6: Variances of inflation and the output gap: High weight on the output gap ($\delta = \frac{0.25}{3}$), all nowcast scenarios

Forecast scenario	Week 1	Week 2	Week 3	Week 4
Variance of inflation				
Medium economic uncertainty scenario				
Baseline scenario (SD AR forecast)	11.39	8.97	8.97	8.97
Bottom-up OLS-match	9.5	7.99	7.96	7.94
Bottom-up U-MIDAS	9.6	8.03	7.97	7.96
Direct ML	9.81	8.05	8.1	7.94
High economic uncertainty scenario				
Baseline scenario (SD AR forecast)	19.16	16.07	16.07	16.07
Bottom-up OLS-match	14.12	12.67	12.63	12.73
Bottom-up U-MIDAS	14.82	12.8	12.69	12.72
Direct ML	16.81	13.36	13.58	12.98
Low economic uncertainty scenario				
Baseline scenario (SD AR forecast)	3.9	3.04	3.04	3.04
Bottom-up OLS-match	3.15	2.9	2.85	2.83
Bottom-up U-MIDAS	3.15	2.9	2.85	2.8
Direct ML	3.0	2.85	2.82	2.83
Variance of the output gap				
Medium economic uncertainty scenario				
Baseline scenario (SD AR forecast)	1.11	0.71	0.71	0.71
Bottom-up OLS-match	0.8	0.55	0.55	0.54
Bottom-up U-MIDAS	0.82	0.56	0.55	0.54
Direct ML	0.85	0.56	0.57	0.54
High economic uncertainty scenario				
Baseline scenario (SD AR forecast)	1.88	1.37	1.37	1.37
Bottom-up OLS-match	1.04	0.8	0.8	0.81
Bottom-up U-MIDAS	1.16	0.83	0.81	0.81
Direct ML	1.49	0.92	0.95	0.85
Low economic uncertainty scenario				
Baseline scenario (SD AR forecast)	0.38	0.23	0.23	0.23
Bottom-up OLS-match	0.25	0.21	0.2	0.2
Bottom-up U-MIDAS	0.25	0.21	0.2	0.19
Direct ML	0.23	0.2	0.2	0.2

Notes: Table 6 reports variances of the inflation rate and the output gap obtained for the case of a hight weight ($\delta = \frac{0.25}{3}$) on output stabilization when nowcast error statistics from both the baseline model (SD-AR model) and all scanner-data based nowcast models are used.

References

Babii, A., E. Ghysels, and J. Striaukas (2022). Machine Learning Time Series Regressions with an Application to Nowcasting. *Journal of Business & Economic Statistics 40*(3), 1094–1106.

Beck, G. W. and V. Wieland (2008). Central Bank Misperceptions and the Role of Money in Interest Rate Rules. *Journal of Monetary Economics 55*, 1–17.

Bergmeir, C., R. J. Hyndman, and B. Koo (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis 120*, 70–83.

Diebold, F. X. and R. S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics 13*(3), 253–263.

Galí, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications. Second edition.* Princeton University Press.

Svensson, L. E. O. and M. Woodford (2003). Indicator Variables for Optimal Policy. *Journal of Monetary Economics 50*, 691–720.

Zhao, P. and B. Yu (2006). On Model Selection Consistency of Lasso. *The Journal of Machine Learning Research 7*, 2541–2563.