# Nowcasting Consumer Price Inflation Using High-Frequency Scanner Data: Evidence from Germany

Günter W. Beck (University of Siegen), Kai Carstensen (Kiel University), Jan-Oliver Menz (Deutsche Bundesbank), Richard Schnorrenberger (Kiel University) and Elisabeth Wieland (Deutsche Bundesbank)

Madrid, October 19, 2023

*The views expressed in this presentation are those of the authors and do not necessarily reflect those of the Deutsche Bundesbank, European Central Bank or the Eurosystem.*

# Background and motivation

- **(Reliable) Real-time** information on **aggregate and disaggregate** inflation can be very beneficial:

# Background and motivation

- **(Reliable) Real-time** information on **aggregate and disaggregate** inflation can be very beneficial:
  - Large shocks (e.g., Global Financial Crisis, Covid-19 pandemic, Russia's invasion of Ukraine) generate enormous uncertainty and require swift and decisive policy measures.
  - Real-time information can generally improve decision making (detection of turning points, information is available when decisions are made, ...)
  - Real-time information can improve the communication of central banks.
  - ...

# Background and motivation

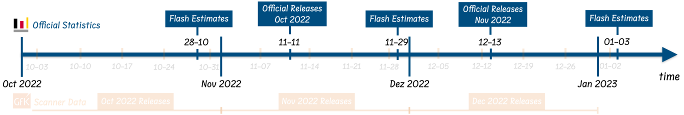
- **(Reliable) Real-time** information on **aggregate and disaggregate** inflation can be very beneficial:
  - Large shocks (e.g., Global Financial Crisis, Covid-19 pandemic, Russia's invasion of Ukraine) generate enormous uncertainty and require swift and decisive policy measures.
  - Real-time information can generally improve decision making (detection of turning points, information is available when decisions are made, ...)
  - Real-time information can improve the communication of central banks.
  - ...
- Official price statistics are generally available at a monthly frequency only and published with a certain time lag ($\approx$ two weeks after a reporting month).

# Background and motivation

- Information flow: Official data versus our scanner data in Germany

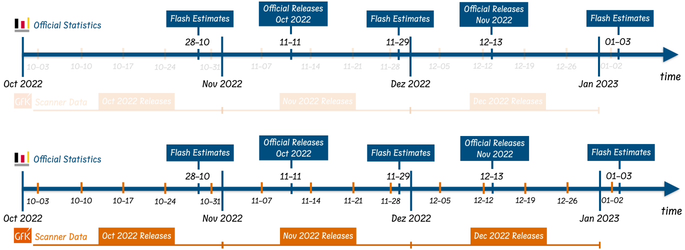
# Background and motivation

- Information flow: Official data versus our scanner data in Germany



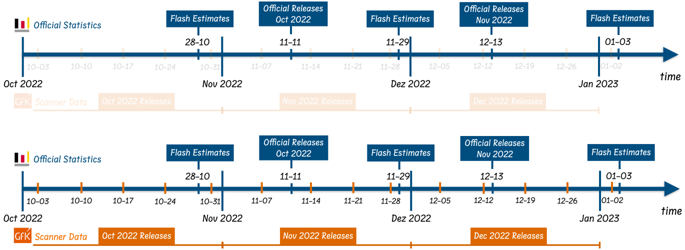
# Background and motivation

- Information flow: Official data versus our scanner data in Germany



# Background and motivation

- Information flow: Official data versus our scanner data in Germany




- Major objective of our paper: Combine non-standard high-frequency price (and quantity) data with state-of-the-art machine learning methods to provide high-quality real-time nowcast inflation

# Our approach

# Our approach

- Use non-standard high-frequency price (and quantity) data:
    - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
    - High-frequency information on energy and package-holiday prices

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)
  - Compute high-frequency price indices for these items

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)
  - Compute high-frequency price indices for these items
  - Combine the high-frequency with complementary low-frequency (= official) price indices to nowcast inflation:

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)
  - Compute high-frequency price indices for these items
  - Combine the high-frequency with complementary low-frequency (= official) price indices to nowcast inflation:
    - At the item level

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)
  - Compute high-frequency price indices for these items
  - Combine the high-frequency with complementary low-frequency (= official) price indices to nowcast inflation:
    - At the item level
    - At product-group levels used in policy analysis ((un-)processed food, ...)

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)
  - Compute high-frequency price indices for these items
  - Combine the high-frequency with complementary low-frequency (= official) price indices to nowcast inflation:
    - At the item level
    - At product-group levels used in policy analysis ((un-)processed food, ...)
    - At the aggregate level (headline inflation)

# Our approach

- Use non-standard high-frequency price (and quantity) data:
  - Household scanner data covering daily purchases of fast-moving consumer goods (GFK:FMCG data) at the barcode level
  - High-frequency information on energy and package-holiday prices
- Basic steps of our approach:
  - Map our scanner (and other) data to the lowest available category level of the official price statistics (=10-digit level)
  - Compute high-frequency price indices for these items
  - Combine the high-frequency with complementary low-frequency (= official) price indices to nowcast inflation:
    - At the item level
    - At product-group levels used in policy analysis ((un-)processed food, ...)
    - At the aggregate level (headline inflation)

  $\Longrightarrow$ Apply machine learning (LASSO, ridge, sg-LASSO) techniques for aggregates

# Related literature

- **Literature on the use of scanner data for economic research** including Kaplan and Schulhofer-Wohl (2017), Jaravel (2019), Butters et al., (2022), Karadi et al., (2023), Beck and Lein (2020), Jaravel and O'Connell (2020), Messner et al., (2023), ... $\Longrightarrow$ Our contribution: Use of household scanner data to nowcast headline inflation.
- **Literature that seeks to construct high-frequency measures of existing low-frequency macroeconomic series** including Anenberg and Laufer (2017), Eraslan and Götz, 2021), Buda et al., 2022, Cavallo and Rigobon (2016), Jaravel and O'Connell (2020), Alvarez and Lein (2020) ... $\Longrightarrow$ Our contribution: Construct both aggregate and disaggreagte weekly inflation measures for unprocessed food, processed food, and non-durable goods.
- **Literature on nowcasting key macroeconomic variables** including Modugno (2013), Breitung and Roling, (2015), Knotek II and Zaman (2017) Clark et al. (2022) Aliaj et al. (2023), Macias et al. (2023), Powell et al., (2018), Harchaoui and Janssen (2018), Aparicio and Bertolotto (2020) ... $\Longrightarrow$ Our contribution: Construct both aggregate and disaggreagte weekly inflation measures for

# Related literature

- **Literature examining whether it pays off to forecast headline inflation by explicitly using disaggregate price indicues** including Hendry and Hubrich (2011), Ibarra (2012) Espasa and Mayo-Burgos (2013), Bermingham and D'Agostino (2013) and more recently Joseph et al. (2022), ... $\Longrightarrow$ Our contribution: We show that combining disaggregate inflation nowcasts into an aggregate nowcast for headline inflation is a highly competitive approach
- **Literature on recent advances in machine learning that seek to improve inflation forecasts by exploiting large data sets** including Garcia et al. (2017), Medeiros et al. (2021) and Babii et al. (2022), Paranhos (2021), Li et al. (2022), Goulet Coulombe et al. (2022), Hauzenberger et al. (2023), Joseph et al. (2022), Botha et al. (2022), and Barkan et al. (2022) ... $\Longrightarrow$ Our contribution: We show that machine learning tools provide an effective solution to handle a large set of disaggregate price series in a mixed-frequency setting.

# Preview of our findings

# Preview of our findings

- Our scanner-based price indices track their official counterparts very well.

# Preview of our findings

- Our scanner-based price indices track their official counterparts very well.
- The high-frequency, scanner-based price indices improve inflation nowcasts at the **item level** considerably, notably already after the first seven days of a month.

# Preview of our findings

- Our scanner-based price indices track their official counterparts very well.
- The high-frequency, scanner-based price indices improve inflation nowcasts at the **item level** considerably, notably already after the first seven days of a month.
- Applying shrinkage estimators to exploit the large set of scanner-based price indices in nowcasting **product groups** such as processed and unprocessed food, we obtain substantial predictive gains compared to a time series benchmark model.

# Preview of our findings

- Our scanner-based price indices track their official counterparts very well.
- The high-frequency, scanner-based price indices improve inflation nowcasts at the **item level** considerably, notably already after the first seven days of a month.
- Applying shrinkage estimators to exploit the large set of scanner-based price indices in nowcasting **product groups** such as processed and unprocessed food, we obtain substantial predictive gains compared to a time series benchmark model.
- Adding high-frequency information on energy and travel services, we construct nowcasting models of **headline inflation** that are on par with, or even outperform, market-based inflation expectations.

# Overview

1. Data: Sources and processing
   - Data sources
   - Mapping scanner data to the official CPI classification system
2. Nowcasting approach and results
   - Nowcasting item-level inflation
   - Nowcasting product group-specific inflation
   - Nowcasting headline inflation
3. Robustness analysis
4. Summary and conclusions

# Table of contents

# Table of contents

# Data sources

- High-frequency data:
  - Fast-moving consumer good (FMCG) scanner data from GfK (GFK:FMCG scanner data)
  - Energy prices
  - Package-holiday prices

# Data sources

- High-frequency data:
  - Fast-moving consumer good (FMCG) scanner data from GfK (GFK:FMCG scanner data)
  - Energy prices
  - Package-holiday prices
- Low-frequency data:
  - Monthly 10-digit COICOP index series from the German National Statistical Office

# GFK:FMCG scanner data

- The GFK:FMCG scanner data primarily cover products from the sectors food, beverages, household and personal care.

# GFK:FMCG scanner data

- The GFK:FMCG scanner data primarily cover products from the sectors food, beverages, household and personal care.
- The data are collected by participants of the GfK household panel which constitutes a **representative sample** (around 30,000 households) of the German population.

| GfK Household Panel | | | | | | |
|---|---|---|---|---|---|---|
| Household | Product Description | Barcode | Quantities | Sales | Retailer | Purchase Date |
| 1 | Green Hill Butter 250g | 400123123123 | 1 | 3.39 € | A | 28.11.2022 |
| 2 | Lovely Butter 250g | 400456456456 | 2 | 6.58 € | B | 01.12.2022 |
| 3 | Lovely Butter 250g | 400456456456 | 1 | 3.39 € | C | 01.12.2022 |
| 4 | Green Hill Butter 250g | 400123123123 | 1 | 3.29 € | B | 02.12.2022 |
| 5 | Green Hill Butter 250g | 400123123123 | 2 | 6.98 € | A | 03.12.2022 |
| 6 | Sunny Sunflower oil 1l | 100445566123 | 1 | 2.29 € | B | 01.12.2022 |
| 7 | Blossom Sunflower oil 1l | 100112233123 | 1 | 3.99 € | C | 01.12.2022 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

Note: For confidentiality reasons, “Product description” and “Barcode” are fictitious entries.

- Purchase records are available at the product level, defined by their barcode (GTIN).
- A **barcode dictionary** contains more detailed product information: manufacturer, brand, volume (package size), product category, ...
- Our sample period is 08 January 2003 to 31 December 2022.

# Table of contents

# Overview of COICOP classification system

- COICOP 2-digit level categories:

| SEA-VPI / SEA-CPI* | Bezeichnung / item | | Gewicht in Promille / Weighting in per mill |
|---|---|---|---|
| | Verbraucherpreisindex insgesamt | Consumer price index, total | 1000 |
| 01 | Nahrungsmittel und alkoholfreie Getränke | Food and non-alcoholic beverages | 96,85 |
| 02 | Alkoholische Getränke und Tabakwaren | Alcoholic beverages and tobacco | 37,77 |
| 03 | Bekleidung und Schuhe | Clothing and footwear | 45,34 |
| 04 | Wohnung, Wasser, Strom, Gas und andere Brennstoffe | Housing, water, electricity, gas and other fuels | 324,70 |
| 05 | Möbel, Leuchten, Geräte u.a. Haushaltszubehör | Furniture, lighting equipment, appliances and other household equipment | 50,04 |
| 06 | Gesundheit | Health | 46,13 |
| 07 | Verkehr | Transport | 129,05 |
| 08 | Post und Telekommunikation | Communication | 26,72 |
| 09 | Freizeit, Unterhaltung und Kultur | Recreation, entertainment and culture | 113,36 |
| 10 | Bildungswesen | Education | 9,02 |
| 11 | Gaststätten- und Beherbergungsdienstleistungen | Restaurant and accommodation services | 46,77 |
| 12 | Andere Waren und Dienstleistungen | Miscellaneous goods and services | 74,25 |

# Overview of COICOP classification system

- COICOP 2-, 3-, 4-, 5- and 10-digit level categories:

| | | | |
|---|---|---|---|
| 01 | Nahrungsmittel und alkoholfreie Getränke | Food and non-alcoholic beverages | 96,85 |
| 011 | Nahrungsmittel | Food | 84,87 |
| 0111 | Brot und Getreideerzeugnisse | Bread and cereals | 15,03 |
| 01111 | Reis, einschl. Reiszubereitungen | Rice, incl. rice preparations | 0,30 |
| 0111101100 | Reis | Rice | |
| 0111109100 | Reiszubereitung | Rice preparations | |
| 01112 | Mehl u.a. Getreideerzeugnisse | Flour and other cereals | 0,36 |
| 0111201100 | Weizenmehl | Flour | |
| 0111203100 | Grieß, Roggenmehl oder Ähnliches | Semolina, rye flour or the like | |
| 01113 | Brot und Brötchen | Bread and bread rolls | 6,27 |
| 0111311100 | Weißbrot | White bread | |
| 0111312100 | Roggenbrot oder Mischbrot | Rye bread or brown bread | |
| 0111313200 | Körnerbrot oder Vollkornbrot | Granary bread or wholemeal bread | |
| 0111320200 | Brötchen zum Fertigbacken | Ready-to-bake rolls | |
| 0111320300 | Frisches Brötchen | Fresh bread rolls | |

# Mapping scanner data into the German COICOP system

# Mapping scanner data into the German COICOP system

- To match the products of the scanner data to the corresponding COICOP- 10 items, we make use
  - of the product categories and the detailed product descriptions included in the GfK household panel and
  - the item descriptions contained in the HICP manual.

# Mapping scanner data into the German COICOP system

- To match the products of the scanner data to the corresponding COICOP- 10 items, we make use
  - of the product categories and the detailed product descriptions included in the GfK household panel and
  - the item descriptions contained in the HICP manual.
- Mapping is done using a combination of machine learning and manual work

# Mapping between scanner data and the German HICP

- Mapped series:

Table 2: Mapping between high-frequency price data and the German HICP

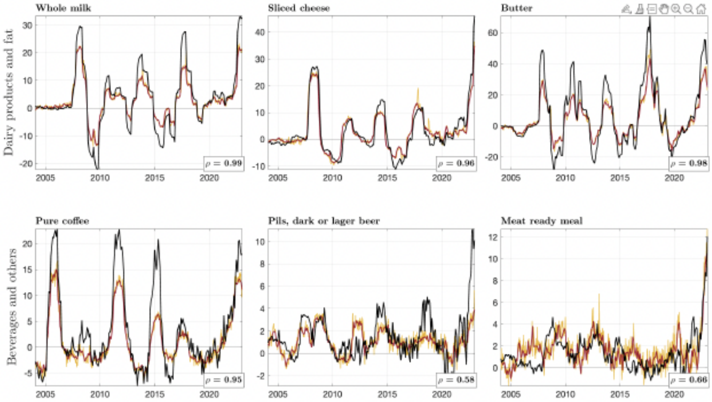
| | HICP | | Scanner data | |
|---|---|---|---|---|
| Component | COICOPs | Weight | COICOPs | Weight |
| Unprocessed food | 38 | 2.4 | 30 | 2.0 |
| Fruit | 8 | 0.7 | 6 | 0.5 |
| Vegetables | 11 | 0.7 | 9 | 0.6 |
| Meat & eggs | 15 | 0.9 | 15 | 0.9 |
| Fish | 4 | 0.1 | 0 | . |
| Processed food | 142 | 11.1 | 116 | 8.1 |
| Fruit | 7 | 0.2 | 5 | 0.1 |
| Vegetables | 12 | 0.4 | 11 | 0.4 |
| Meat | 13 | 1.1 | 11 | 0.9 |
| Fish | 7 | 0.2 | 4 | 0.1 |
| Bread & cereals | 25 | 1.5 | 23 | 1.4 |
| Dairy products & fat | 18 | 1.5 | 14 | 1.4 |
| Beverages | 29 | 2.9 | 23 | 2.7 |
| Other food products | 28 | 1.2 | 25 | 1.0 |
| Tobacco | 3 | 2.1 | 0 | . |
| NEIG | 302 | 23.0 | 39 | 1.8 |
| Non-durables | 75 | 5.9 | 36 | 1.8 |
| Semi-durables | 139 | 8.7 | 3 | 0.1 |
| Durables | 88 | 8.4 | 0 | . |
| Total HICP | 482 | 36.5 | 185 | 11.9 |

# Scanner-based and official price indices

- Scanner-based 10-digit price indices are computed using a (rolling) weighted time-product dummy price index (mean splice).

# Scanner-based and official price indices

- Scanner-based 10-digit price indices are computed using a (rolling) weighted time-product dummy price index (mean splice).
- Selected individual series:



# Table of contents

# Table of contents

# Nowcasting item-level inflation

- For each COICOP-10 item $c$ for which we have weekly GFK:FMCG data available, we estimate a (U-)MIDAS model of the monthly HICP inflation rate, $\pi^M_{c,t}$ using

$$\phi(L)\,\pi^M_{c,t+h} = \alpha_{0,h} + \beta_{c,h}\,B(L^{1/m};\theta)\,x^{(m)}_{c,t} + \sum_{i=s}^{13}\gamma_{c,s}\,d_{c,s,t+h} + \varepsilon_{c,t+h}, \tag{1}$$

  - $t = 1,\dots,T$ denotes the monthly time index
  - $m = 1,\dots,4$ denotes the weeks of a month.
- The predictors include the weekly GFK:FMCG inflation rate, $x^{(m)}_{c,t}$, sampled four times more frequently than the target variable, and a set of monthly dummy variables, $d_{c,1,t},\dots,d_{c,13,t}$.

# RMSE for FMCG product-level inflation: U-MIDAS vs. SD-AR benchmark

**Unprocessed fruit and vegetables**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Sweet peppers | $0.515^{**}$ | $0.562^{**}$ | $0.555^{**}$ | $0.567^{**}$ |
| Butterhead Lettuce | $0.555^{**}$ | $0.628^{**}$ | $0.582^{**}$ | $0.63^{**}$ |
| Onion or garlic | $0.578^{**}$ | $0.714^{**}$ | $0.678^{**}$ | $0.662^{**}$ |
| Carrots | $0.547^{**}$ | $0.718^{**}$ | $0.727^{**}$ | $0.727^{**}$ |
| Grapes | $0.75^{**}$ | $0.72^{**}$ | $0.693^{**}$ | $0.716^{**}$ |
| Lambs lettuce | $0.755^{*}$ | $0.714^{*}$ | $0.714^{*}$ | $0.715^{*}$ |
| Cauliflower or cabbage | $0.783^{**}$ | $0.758^{**}$ | $0.754^{**}$ | $0.741^{**}$ |

**Processed fruit and vegetables**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Frozen chips or the like | 0.522 | 0.673 | 0.652 | 0.64 |
| Apple sauce | 0.802 | 0.806 | 0.785 | 0.784 |
| Tinned mushrooms | 0.803 | 0.834 | 0.852 | 0.849 |
| Potatoes | 0.813 | 0.974 | 0.929 | 0.916 |
| Potato crisps | 0.924 | 0.907 | 0.918 | 0.913 |
| Asparagus | 0.94 | 0.913 | 0.932 | 0.926 |
| Tinned pineapple | 0.923 | 0.952 | 0.931 | 0.944 |

**Unprocessed meat, fish and eggs**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Eggs | 0.666 | 0.792 | 0.662 | 0.659 |
| Fresh poultry | 0.791 | 0.808 | 0.803 | 0.808 |
| Minced pork | 0.853 | 0.855 | 0.849 | 0.85 |
| Minced beef | 0.781 | 0.885 | 0.886 | 0.867 |
| Beef for boiling | 0.83 | 0.871 | 0.876 | 0.876 |
| Pork chop or cutlet | 0.881 | 0.902 | 0.902 | 0.88 |
| Smoked pork chop | 0.881 | 0.92 | 0.894 | 0.905 |

**Processed meat and fish**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Ham or bacon | 0.622 | 0.673 | 0.675 | 0.675 |
| Salami or sausage | 0.676 | $0.727^{*}$ | $0.72^{*}$ | $0.72^{*}$ |
| Liver sausage | 0.705 | 0.771 | $0.763^{*}$ | $0.769^{*}$ |
| Fish fingers | 0.751 | 0.777 | $0.755^{*}$ | 0.785 |
| Lyoner pork sausage | 0.733 | 0.791 | 0.771 | 0.781 |
| Fried sausage | 0.777 | 0.837 | 0.816 | 0.81 |
| Prepared minced meat | 0.807 | 0.863 | 0.848 | 0.837 |

# RMSE for FMCG product-level inflation: U-MIDAS vs. SD-AR benchmark

**Beverages and others**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Baking powder | 0.544 | 0.675 | 0.685 | 0.679 |
| Orange juice | 0.713 | 0.76 | 0.763* | 0.766* |
| Sugar | 0.699 | 0.892 | 0.848 | 0.85 |
| Apple juice | 0.746 | 0.918 | 0.867 | 0.848 |
| Blancmange powder | 0.846 | 0.871* | 0.858* | 0.869 |
| Food for infants | 0.833** | 0.868** | 0.884* | 0.889* |
| Ketchup | 0.931 | 0.863 | 0.853 | 0.859 |

**Bread and cereals**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Flour | 0.582 | 0.557 | 0.565 | 0.557 |
| Pasta | 0.73 | 0.769* | 0.774* | 0.796 |
| Rice | 0.758 | 0.792* | 0.806 | 0.809* |
| Ready to bake rolls | 0.746 | 0.817* | 0.823* | 0.826* |
| Crisp bread | 0.779 | 0.846* | 0.838* | 0.869* |
| Cornflakes and muesli | 0.832 | 0.847 | 0.862 | 0.853 |
| Frozen cake, tart or pie | 0.795 | 0.872 | 0.867 | 0.875 |

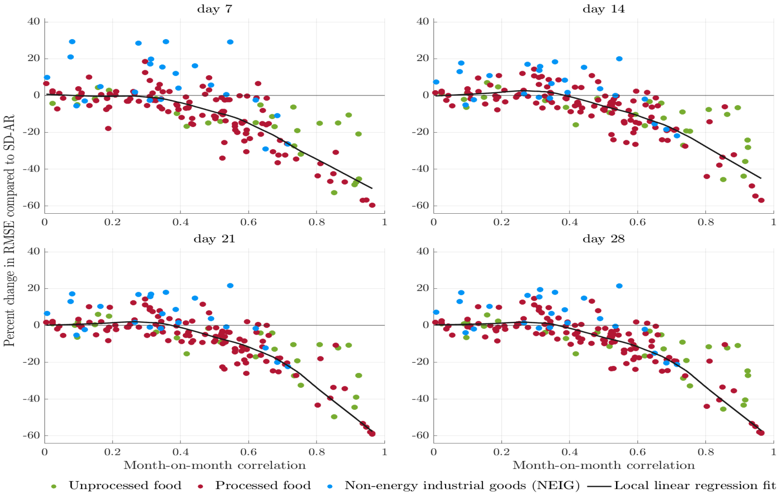
**Non-durables**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Aluminium foil | 0.72 | 0.768* | 0.752* | 0.749* |
| Toilet tissue | 0.8 | 0.8 | 0.796 | 0.803 |
| Hair shampoo | 0.725** | 0.827** | 0.857** | 0.83** |
| Deo spray or deo roll on | 0.768* | 0.873 | 0.859* | 0.851* |
| Hair spray or gel | 0.839 | 0.85* | 0.846* | 0.858 |
| Paper handkerchiefs | 0.924 | 0.832 | 0.83 | 0.834 |
| Toothpaste | 0.866 | 0.887* | 0.889* | 0.885** |

**Dairy products and fat**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Whole milk | 0.39 | 0.437 | 0.416 | 0.42 |
| Condensed milk | 0.359 | 0.442 | 0.437 | 0.436 |
| Low fat milk | 0.412 | 0.427 | 0.425 | 0.428 |
| Cream | 0.424* | 0.461 | 0.455 | 0.457 |
| Butter | 0.433** | 0.498** | 0.457** | 0.458** |
| Curd | 0.539 | 0.638 | 0.622 | 0.612 |
| Sliced cheese | 0.593 | 0.68 | 0.678 | 0.68 |

# Predictive gain of the FMCG data as a function of the in-sample fit with official counterparts



# Table of contents

# Nowcasting product group-specific inflation

- Given the restricted coverage of the GFK:FMCG data, we focus on
  - the **high-level product groups** unprocessed food, processed food and non-energy industrial goods

# Nowcasting product group-specific inflation

- Given the restricted coverage of the GFK:FMCG data, we focus on
  - the **high-level product groups** unprocessed food, processed food and non-energy industrial goods
  - the **more disaggregated low-level product groups** unprocessed fruit and vegetables; unprocessed meat, fish and eggs; processed fruit and vegetables; processed meat, fish and eggs; bread and cereals; dairy products and fat; beverages and other food products; and non-durables.

# Nowcasting product group-specific inflation

- Given the restricted coverage of the GFK:FMCG data, we focus on
  - the **high-level product groups** unprocessed food, processed food and non-energy industrial goods
  - the **more disaggregated low-level product groups** unprocessed fruit and vegetables; unprocessed meat, fish and eggs; processed fruit and vegetables; processed meat, fish and eggs; bread and cereals; dairy products and fat; beverages and other food products; and non-durables.
- A natural starting point for nowcasting at the product-group level would be to treat all COICOP-10 series belonging to that group as relevant predictors.

# Nowcasting product group-specific inflation

- Given the restricted coverage of the GFK:FMCG data, we focus on
  - the **high-level product groups** unprocessed food, processed food and non-energy industrial goods
  - the **more disaggregated low-level product groups** unprocessed fruit and vegetables; unprocessed meat, fish and eggs; processed fruit and vegetables; processed meat, fish and eggs; bread and cereals; dairy products and fat; beverages and other food products; and non-durables.
- A natural starting point for nowcasting at the product-group level would be to treat all COICOP-10 series belonging to that group as relevant predictors.

  $\Longrightarrow$ U-MIDAS setting is not suited to handle such a large set of predictors.

# Nowcasting product group-specific inflation

- Given the restricted coverage of the GFK:FMCG data, we focus on
  - the **high-level product groups** unprocessed food, processed food and non-energy industrial goods
  - the **more disaggregated low-level product groups** unprocessed fruit and vegetables; unprocessed meat, fish and eggs; processed fruit and vegetables; processed meat, fish and eggs; bread and cereals; dairy products and fat; beverages and other food products; and non-durables.
- A natural starting point for nowcasting at the product-group level would be to treat all COICOP-10 series belonging to that group as relevant predictors.

  $\Longrightarrow$ U-MIDAS setting is not suited to handle such a large set of predictors.

  $\Longrightarrow$ We apply shrinkage estimators using two different approaches.

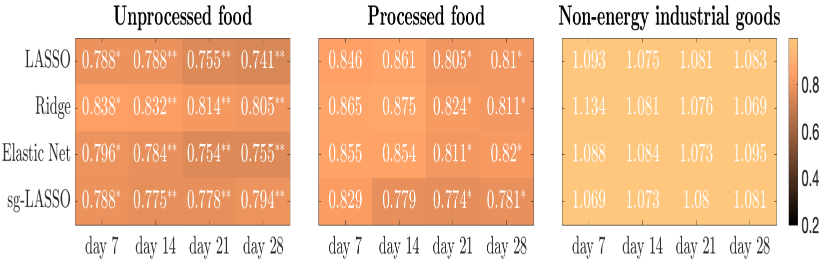
# Nowcasting product group-specific inflation

- **Our first approach** avoids mixed frequencies by aggregating the weekly GFK:FMCG indicators, $x_{c,t}^{(m)}$, $m = 1, \ldots, 4$, to the monthly frequency which yields $x_{c,t}^{(M)}$.
- We apply standard shrinkage methods to estimate nowcasting models of the group-specific target inflation rates and plug in our high-frequency disaggregate inflation measures to generate the nowcast of a month.
- **Our second approach** applies the sparse-group LASSO (sg-LASSO) estimator proposed by Babii et al. (2022) and regresses the group-specific target inflation rate, $\pi_{g,t}^{M}$, directly on the large set of weekly GFK:FMCG inflation rates

  $\Longrightarrow$ Approach performs shrinkage in a mixed-frequency rather than a low-frequency setting by recognizing serial dependence across different high-frequency lags.

# RMSE for FMCG product-group inflation: Various shrinkage methods relative to the SD-AR benchmark

- High-level product groups



# RMSE for FMCG product-group inflation: Various shrinkage methods relative to the SD-AR benchmark

- 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 |

Colour scale: 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2

# Table of contents

# Nowcasting headline inflation: Approach(es)

# Nowcasting headline inflation: Approach(es)

- To nowcast headline inflation, we split it into the following six components: unprocessed food, processed food, non-energy industrial goods, energy, package holidays, and services (w.o. package holidays).

# Nowcasting headline inflation: Approach(es)

- To nowcast headline inflation, we split it into the following six components: unprocessed food, processed food, non-energy industrial goods, energy, package holidays, and services (w.o. package holidays).
- We nowcast each part separately using two different approaches and construct a headline nowcast by using the official HICP weighting scheme.

# Nowcasting headline inflation: Approach(es)

- To nowcast headline inflation, we split it into the following six components: unprocessed food, processed food, non-energy industrial goods, energy, package holidays, and services (w.o. package holidays).
- We nowcast each part separately using two different approaches and construct a headline nowcast by using the official HICP weighting scheme.
- First nowcasting approach: *Bottom-up U-MIDAS approach*

# Nowcasting headline inflation: Approach(es)

- To nowcast headline inflation, we split it into the following six components: unprocessed food, processed food, non-energy industrial goods, energy, package holidays, and services (w.o. package holidays).
- We nowcast each part separately using two different approaches and construct a headline nowcast by using the official HICP weighting scheme.
- First nowcasting approach: *Bottom-up U-MIDAS approach*
- Second nowcasting approach: Direct machine learning approach

# Nowcasting headline inflation: Approach(es)

- To nowcast headline inflation, we split it into the following six components: unprocessed food, processed food, non-energy industrial goods, energy, package holidays, and services (w.o. package holidays).
- We nowcast each part separately using two different approaches and construct a headline nowcast by using the official HICP weighting scheme.
- First nowcasting approach: *Bottom-up U-MIDAS approach*
- Second nowcasting approach: Direct machine learning approach
- Benchmarks.
  - For the six components of headline inflation, we use aggregated nowcasts of the SD-AR model applied to each item at the COICOP-10 level as the benchmark.
  - For headline inflation, we also use market expectations provided by Bloomberg and Consensus Economics as the benchmark.

# RMSE of headline inflation and its components: Bottom-up U-MIDAS and direct machine learning approaches relative to the benchmark approach

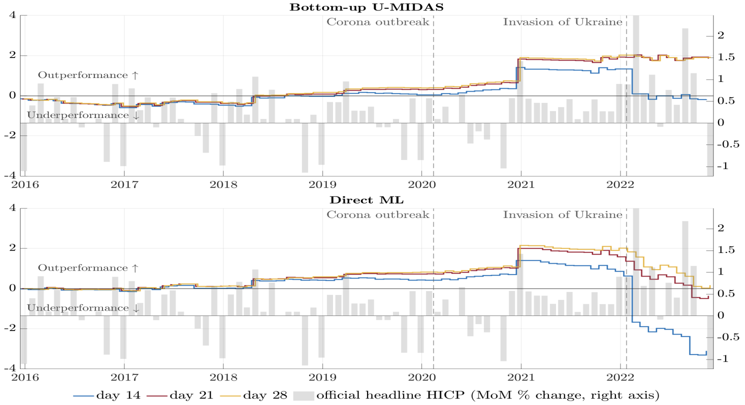
**Bottom-up U-MIDAS**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Unprocessed food | 0.696* | 0.735* | 0.724* | 0.721* |
| Processed food | 0.743 | 0.767* | 0.763* | 0.76* |
| Energy | 0.532** | 0.357* | 0.32* | 0.32* |
| Package holidays | 1.043 | 1.044 | 0.951 | 0.882** |
| Non-energy industrial goods (NEIG) | 1.005 | 0.998 | 0.998 | 0.998 |
| Services | 1 | 1 | 1 | 1 |
| Headline | 0.832 | 0.707 | 0.675 | 0.668 |

**Direct ML**

| | day 7 | day 14 | day 21 | day 28 |
|---|---|---|---|---|
| Unprocessed food | 0.786 | 0.735** | 0.719** | 0.706** |
| Processed food | 0.996 | 0.938 | 0.889 | 0.849 |
| Energy | 0.625* | 0.412* | 0.466 | 0.331* |
| Package holidays | 1.071 | 1.083 | 0.979 | 0.894** |
| Non-energy industrial goods (NEIG) | 1.134 | 1.129 | 1.119 | 1.128 |
| Services | 1 | 1 | 1 | 1 |
| Headline | 0.877 | 0.706 | 0.715 | 0.66 |

# Cumulative sum of the squared forecast error differentials: models versus Bloomberg market expectations



# Table of contents

# Table of contents

# Summary and conclusions

# Summary and conclusions

- We demonstrate how combining millions of household scanner data with state-of-the-art machine learning techniques yields highly competitive real-time inflation nowcasts for Germany at the disaggregate and aggregate level.

# Summary and conclusions

- We demonstrate how combining millions of household scanner data with state-of-the-art machine learning techniques yields highly competitive real-time inflation nowcasts for Germany at the disaggregate and aggregate level.
- Our strategy to combine fixed economic structure inherent in the COICOP classification system with flexible machine learning tools turned out to provide two major merits:

# Summary and conclusions

- We demonstrate how combining millions of household scanner data with state-of-the-art machine learning techniques yields highly competitive real-time inflation nowcasts for Germany at the disaggregate and aggregate level.
- Our strategy to combine fixed economic structure inherent in the COICOP classification system with flexible machine learning tools turned out to provide two major merits:
  - It exploits the virtue of granular data to provide an understanding of the disaggregate dynamics underlying overall inflation.

# Summary and conclusions

- We demonstrate how combining millions of household scanner data with state-of-the-art machine learning techniques yields highly competitive real-time inflation nowcasts for Germany at the disaggregate and aggregate level.
- Our strategy to combine fixed economic structure inherent in the COICOP classification system with flexible machine learning tools turned out to provide two major merits:
  - It exploits the virtue of granular data to provide an understanding of the disaggregate dynamics underlying overall inflation.
  - At the same, time, it delivers valuable high-frequency real-time information about price developments of aggregates closely monitored by policy makers and market participants.

# Summary and conclusions

- We demonstrate how combining millions of household scanner data with state-of-the-art machine learning techniques yields highly competitive real-time inflation nowcasts for Germany at the disaggregate and aggregate level.
- Our strategy to combine fixed economic structure inherent in the COICOP classification system with flexible machine learning tools turned out to provide two major merits:
  - It exploits the virtue of granular data to provide an understanding of the disaggregate dynamics underlying overall inflation.
  - At the same, time, it delivers valuable high-frequency real-time information about price developments of aggregates closely monitored by policy makers and market participants.
- Given our findings, we think it is highly worthwhile to identify and exploit high-frequency information concerning those parts of the consumption basket underlying German (and other countries') inflation that are not covered by scanner data on fast-moving consumer goods.