Fabio Rumler
Principal
Oesterreichische Nationalbank

What micro price data teach us about the inflation process: analyzing scanner data in PRISMA[1, 2]

Research conducted within the PRISMA research network draws on three types of micro price data: data underlying the official price indices (i.e. consumer price and producer price indices), scanner data and (web-scraped) online prices. This contribution discusses the role of scanner data in the analysis of price and inflation dynamics at the micro-level and presents some results from analyses based on scanner data.[3]

There are two main kinds of scanner data: retail scanner data (i.e. data of all products being scanned at the cashier's desks of supermarkets) supplied by the market research institute IRi and household scanner data (i.e. data of products bought and scanned by households) supplied by the market research companies GfK and Kantar. Each data type can be used to address different questions. They both cover only fast-moving consumer goods (FMCG), broadly speaking food, personal care and household convenience products, whose weight in the harmonized index of consumer prices (HICP) basket amounts to around 20%. For an overview of the main differences of the retail and household scanner data see table 1.

The retail scanner dataset available to PRISMA has a very high coverage of sales data of several supermarkets in four euro area countries: Germany, France, Italy and the Netherlands. It contains the revenue in euro and the units sold of about 400,000 to 800,000 products in 6,000 to 15,000 stores (depending on the country) recorded on a weekly basis from 2013 to 2017. Stores in the dataset include supermarkets, discounters and drugstores/chemists. The dataset contains information on private label products and also the location of the store (identified by the two-digit ZIP code).

The household scanner dataset covers 16 countries over several years and contains information on the goods bought by households participating in the survey during regular shopping trips. As such, the dataset comprises information on prices paid and quantities bought by households, on the retailer where the products were bought, and most importantly, it contains information on household characteristics, such as the age of the household head, the size of the household, income and the location of the household's residence.[4] The sample coverage in terms of time, products and number of households varies across countries. The main advantage of such data is that they allow to examine actual prices paid and quantities of products purchased in relation to household characteristics. Given that shopping trips of households are infrequent and made at irregular intervals, time-aggregation is needed to calculate price changes of identical products and ultimately household-level inflation rates.

The unique properties of the scanner data with information on the price, quantity and on the precise product identifier (i.e. barcode) allows PRISMA

[1] *Fabio Rumler (based on contributions to the Price-setting Microdata Analysis [PRISMA] Network)*

[2] *I would like to thank Teresa Messner (OeNB) for valuable feedback and input to this note.*

[3] *For a review of other data sources used in PRISMA, see Chiara Osbat's contribution.*

[4] *However, the availability of this demographic information is heterogenous across countries.*

researchers to analyze a range of questions that are central in the literature on price-setting and inflation determination and thus relevant for monetary policy. The questions addressed include questions on inflation heterogeneity across households and countries, on consumer behavior, product substitution, differences between online and offline prices, the degree of state dependence in price-setting, the pass-through of shocks to consumer prices, etc. (see Kiss and Strasser, 2022; Messner and Rumler, 2022; Messner et al., 2022; Strasser and Wittekopf, 2022; Ampudia et al., 2022; Karadi et al., 2022; Dedola et al., 2022). The usefulness of these data will be demonstrated by two particular PRISMA projects in the remainder of this note; there is, however, a plethora of studies drawing on scanner data. For further fields of application, analyses and research projects please refer to the PRISMA report which will be published in the fall of 2022.

Despite covering only a small part of overall household consumption (FMCG), Messner and Rumler (2022) show for Austria that inflation based on a price index calculated from these data is broadly correlated with the overall CPI and more closely correlated with the CPI for food and beverages (chart 1, left-hand panel).

Calculating inflation rates for each household in the panel, they document that households experience widely differing rates of inflation. Looking for household-specific determinants of this heterogeneity, they find that, overall, there appears to be little difference between income groups, with lower income groups facing only marginally higher inflation rates than higher income groups (chart 1, right-hand panel). While this finding is at odds with part of the existing literature on inflation heterogeneity using disaggregate CPI data, which documents a significantly negative relationship between income and inflation (see e.g. Fessler and Fritzer, 2013, for Austria), it is in line with studies also using household scanner data such as Kaplan and Schulhofer-Wohl (2017) for the US and Ampudia et al. (2022) for European countries.

Table 1

Main differences of retail and household scanner data in PRISMA

	Supermarket scanner data	Household scanner data
Data provider	IRi	GfK
Collection point	store	household
Frequency	weekly average	transaction
	quite regular	infrequent
Period	2013–2017	2005–2018 (max.)
Future updates	yes, with delay	no
Number of markets	Germany, France, Netherlands, Italy	16 European countries
Region detail	two-digit zip	varies
Store coverage	participating chain	all
Product coverage	only FMCG	broader, dominated by FMCG
Barcodes	masked EANs/SKUs	EANs/SKUs
Quantities	yes	yes
Store characteristics	store ID, chain, type	chain only
Household characteristics	no	yes
Coupon use/ sales flag	no	no

Source: Author's compilation based on IRi and GfK data.

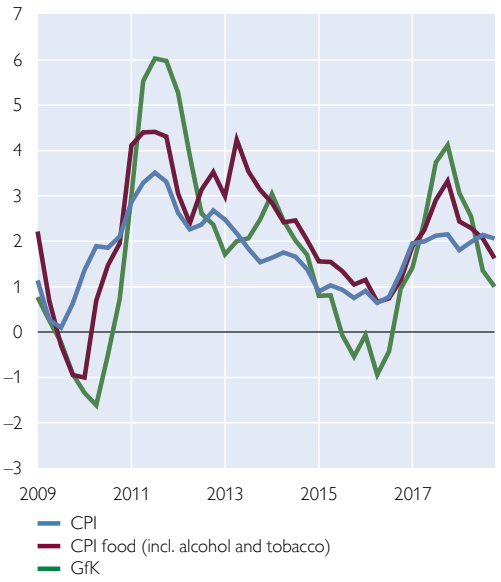
Note: EANs – European article numbers; SKUs – stock keeping units.

Chart 1

Inflation rates in Austria based on household scanner data

Overall inflation rates

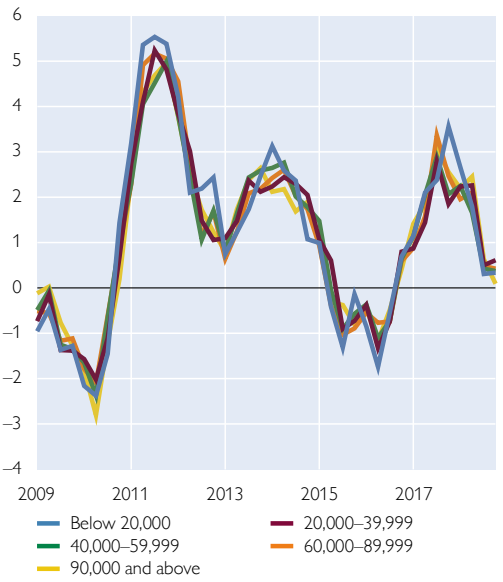
Year-on-year change in %



Source: Statistik Austria, GfK, author's calculations.

Inflation rates by income group

Year-on-year change in %



Source: GfK, author's calculations.

However, a negative relationship between income and inflation is found to arise in Austria in high-inflation periods (in line with Argente and Lee, 2021, for the US).

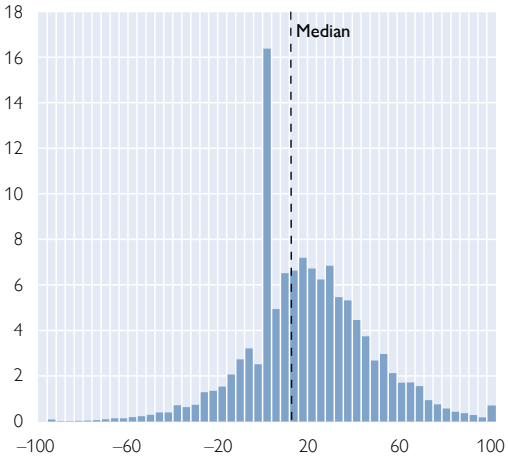
Scanner data are also informative when comparing inflation and price developments across countries. Even in an ideal setting where neighboring countries are economically and culturally perfectly integrated, like Austria and Germany, and the law of one price could be expected to apply across borders, Messner et al. (2022) find that prices and inflation rates at the product level can differ quite substantially. Comparing more than 300,000 individual products which are available on both sides of the Austrian-German border, they find that on average prices are higher on the Austrian side with a wide distribution of price differences in both directions (chart 2, left-hand panel shows the distribution of price differences subtracting German from Austrian prices). As is clearly visible from the chart, the mass of cross-border price differences is close to zero, but in absolute terms prices differ on average by about 20% across the border (right-hand panel of chart 2) which is an order of magnitude higher than price differences within each country.

Chart 2

Price differences (in %) of identical products across Austrian and German border regions

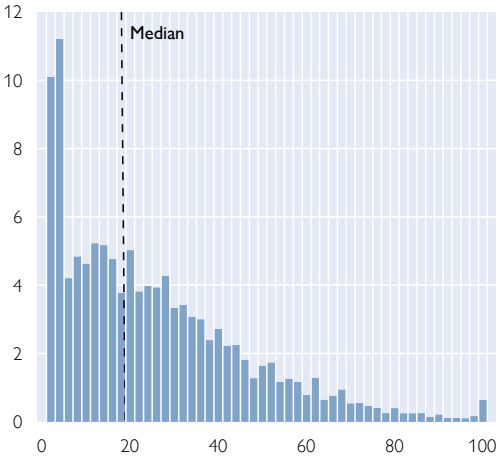
Price differences

Frequency in thousands., difference in %



Absolute price differences

Frequency in thousands., difference in %



Source: GfK, author's calculations.

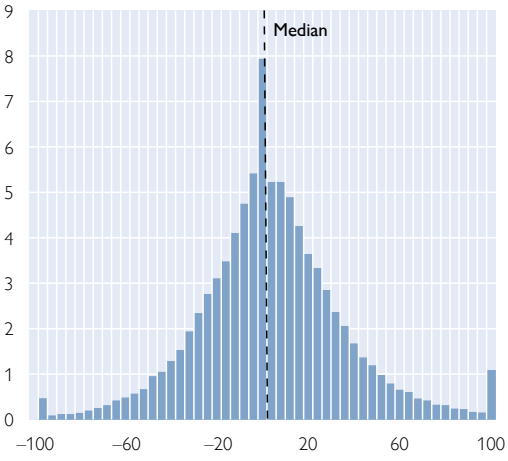
Note: Observations are winsorized at +/– 100.

Chart 3

Price change differences (year-on-year, in %) of identical products across Austrian-German border regions

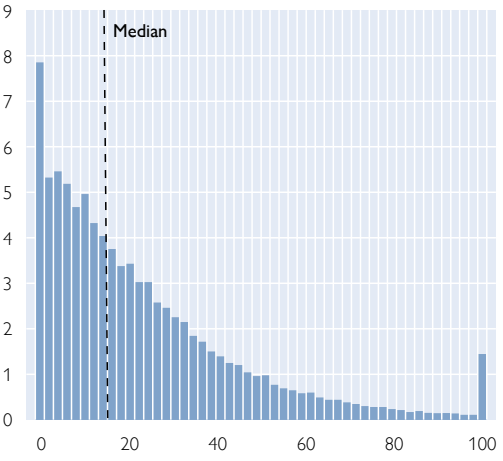
Price change differences

Frequency in thousands, difference in percentage points



Absolute price change differences

Frequency in thousands, difference in percentage points



Source: GfK, author's calculations.

Note: Observations are winsorized at +/– 100.

In contrast, differences in year-on-year price changes derived from the same products across the Austrian-German border (chart 3) tend to be more balanced than differences in price levels. This indicates that for the products considered there is no overall inflation difference found between Austrian and German border regions (left-hand panel).

References

Ampudia, M., M. Ehrmann and G. Strasser. 2022. Monetary Policy and Inflation Inequality. Mimeo.

Argente, D.O. and M. Lee. 2021. Cost of living inequality during the Great Recession. Journal of the European Economic Association 19(2). 913–952.

Dedola, L., C. Osbat and T. Reinelt. 2022. Tax thy neighbour: Corporate tax pass-through into downstream consumer prices in a monetary union. Mimeo.

Fessler, P. and F. Fritzer. 2013. The Distribution of Inflation among Austrian Households. Monetary Policy & the Economy (3). 12–28.

Kaplan, G. and S. Schulhofer-Wohl. 2017. Inflation at the household level. Journal of Monetary Economics 91(C). 19–38.

Karadi, P., J. Amann, J.S. Bachiller and J. Wursten. 2022. State-dependence in price setting on the two sides of the Atlantic - Evidence from supermarket scanner data. Mimeo.

Kiss, R. and G. Strasser. 2022. Inflation heterogeneity across households. Mimeo.

Messner, T. and F. Rumler. 2022. Inflation heterogeneity across Austrian households. Evidence from household retail scanner data. Mimeo.

Messner, T., F. Rumler and G. Strasser. 2022. Cross-country price and inflation dispersion: Retail network or national border? Mimeo.

Strasser, G. and D. Wittekopf. 2022. Online shopping, consumer behavior and inflation. Mimeo.