Book of Short Papers. 2nd Italian Conference on Economic Statistics (ICES 2024) "Statistical Analysis of Complex Economic Data: Recent Developments and Applications. ISBN: 978-88-476-2950-9

Assessing food affordability in Italy: A spatial-temporal analysis using web-scraped data

Valutare l'accessibilità dei prodotti alimentari in Italia: un'analisi spazio-temporale utilizzando dati raccolti dal web

Ilaria Benedetti, Federico Crescenzi, Tiziana Laureti, and Luigi Palumbo

Abstract This study delves into the economic facet of food security, focusing on the pivotal role of food prices as significant barriers to accessing a wholesome and sustainable diet. The challenges posed by economic recessions exacerbate these barriers, amplifying the struggle of vulnerable consumers to afford food. Leveraging web-scraped data, this paper introduces a pioneering exploration of economic food access in Italy. Employing a Time-interaction-Region Dummy Model within the framework of a stochastic price index approach, the analysis of least-cost food products (LCFP) sheds light on the affordability of food products, providing valuable insights into the intricate economic dynamics of food access in Italy.

Abstract *Questo studio approfondisce l'aspetto economico della sicurezza alimentare, in particolare indagando sul ruolo chiave dei prezzi al consumo come ostacoli significativi all'accesso a una dieta sana e sostenibile. Le sfide poste dalle recessioni economiche esacerbano queste barriere, amplificando la difficoltà dei consumatori vulnerabili per l'accesso ai prodotti alimentari. Sfruttando i dati raccolti dal web, questo articolo introduce un'esplorazione innovativa sull'accesso economico ai prodotti alimentari in Italia. Utilizzando un modello ad interazioni spazio-temporali con variabili binarie nel quadro di un approccio stocastico all'indice dei prezzi, l'analisi dei prodotti alimentari a costo minimo (LCFP) fa luce sull'accessibilità economica dei prodotti alimentari, fornendo preziose informazioni sulle intricate dinamiche economiche dell'accesso ai prodotti alimentari in Italia.*

[1] Ilaria Benedetti, University of Tuscia; email: i.benedetti@unitus.it
Federico Crescenzi, University of Tuscia; email: federico.crescenzi@unitus.it
Tiziana Laureti, University of Tuscia; email: laureti@unitus.it
Luigi Palumbo, Bank of Italy; email: Luigi.palumbo@bancaditalia.it

Key words: food security, food affordability, price indices, web-scraped data

Studying least-cost food products: motivations and aim

Food security, defined as "having at all times, physical, social and economic access to sufficient, safe and nutritious food that meets dietary needs and food preferences for an active and healthy life" (FAO, 2002), is one of the most challenging topics globally (Abay et al., 2023). In the realm of economic food security, food prices emerge as a critical determinant, acting as key barriers to accessing a diet that is not only sufficient but also safe and nutritious for maintaining an active and healthy life (Herforth et al. 2020). The repercussions of food price increases are particularly concerning for low-income households, who allocate a significant portion of their income to food, thus adversely impacting the well-being of these households (Abay et al., 2023).

During economic recessions, the landscape of food security could become more challenging. The loss of income common to many households during these periods exacerbates the inability of vulnerable populations to afford food. The economic dimension of food security can be viewed as a crucial aspect of a household's well-being and it is one of the most important measures of poverty and of economic inequality (Secondi and Marchetti, 2022).

Traditional methods of assessing food affordability, usually reliant on surveys and government indices, offer valuable insights but are occasionally constrained by limitations in granularity and timeliness. Web-scraped data, harvested from diverse online platforms, offers a real-time and comprehensive snapshot of consumer behaviour and pricing trends. In particular, these data offer the potential to improve greatly the quality and coverage of consumer price indices (Benedetti et al., 2022). Web-scraped data also provides an innovative lens through which examine the food accessibility.[1] We conducted a novel exploration of economic food access, leveraging an original dataset from web-scraped food prices across grocery shops in Italy, by analysing the least-cost food products (LCFP) prices and comparing their evolution across time and space with the median and highest-cost products prices in the same category.

This paper represents a departure from existing literature on sub-national Spatial Price Indices (SPIs) by illustrating the practicality of utilizing web-scraped data for assessing intra-national spatial price dynamics. By specifying a Time-interaction-Region Dummy Model (TiRD) developed in the context of stochastic approach of price indexes, the outcomes of this study contribute valuable insights into the economic dynamics of food access in Italy, shedding light on the affordability and availability of least-cost food options.

[1] Web-scraped data is also used in many countries for official consumer price index compilation (EC, 2020).

Data and Methodology

The web-scraped dataset spans from July 2021 to February 2023, focusing on 167 outlets in 7 different regions affiliated with a major Italian retail chain. Our web scraping routines performed a daily collection of prices and details for products available in each individual outlet, leveraging a hidden Application Programming Interface (API) in the retail chain e-commerce platform. The analysis focuses on two product groups, namely Milk and Olive Oil, with individual items identified through Global Trade Item Numbers (GTINs).

The Milk and Olive Oil product groups account for 1% and 2% of the Italian household monthly expenditure, respectively. We only selected products with size equal to 1 Litre, in order to keep a direct comparability of prices amongst our sample. The pricing concept refers to the monthly price for each GTIN in each outlet. Our data collection amounted to 521 unique products for Milk and 102 unique products for Olive Oil, with a total of 189,445 and 82,405 monthly observations, respectively.

To explore the least-cost food products evolution across space and time, we applied a modified version of the TiRPD index, introduced by Aizcorbe and Aten (2004) in the context of the stochastic approach to price indexes considering a single homogeneous product across all outlets:

$$\ln P_{ijt} = \sum_{t=1}^{T} \sum_{j=1}^{J} \delta_{jt} C_{ij} T_{it} + \nu_{ijt}$$

Where lnP_{ijt} and D_{ijt} are respectively the log-price and the dummy for the product in outlet i in region j at time t (i = 1, 2, ..., N; j = 1, 2, ..., M; t = 1, 2, ..., T). C_{ij} and T_{jt} are dummy variables for each combination of area and time. The sub-national SPI for the area j at time t is given by $exp(\delta_{jt})$. In our case, all products in the same price class – least-cost, median, and highest – are considered homogeneous.

Results and conclusions

Results from TiRPD models are reported in Fig.1 and Fig. 2. Findings from the LCFP analysis revealed that, on average for these two food product groups, the prices of the least expensive items have increased at a faster pace compared to the prices of the median or most expensive products in the same category, both over time and across different regions in Italy.

Further research is needed in order to generalise such findings, however this observation highlights the potential challenges posed by escalating food commodity prices, particularly for those relying on more affordable food options.

Figure 1: Spatial and Temporal price index for Milk product group according to price band

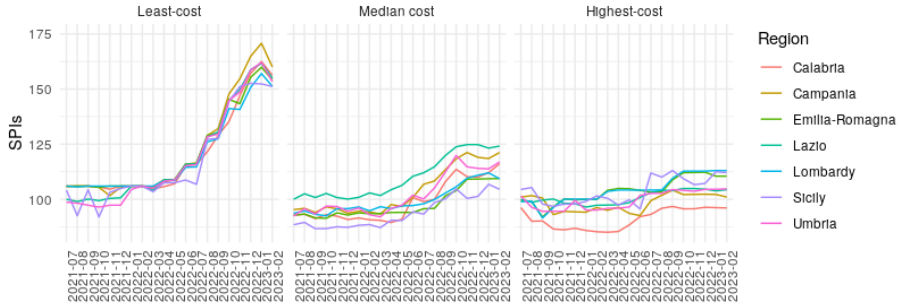
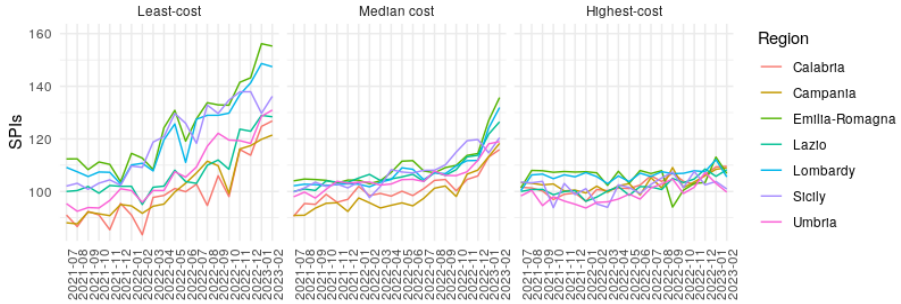


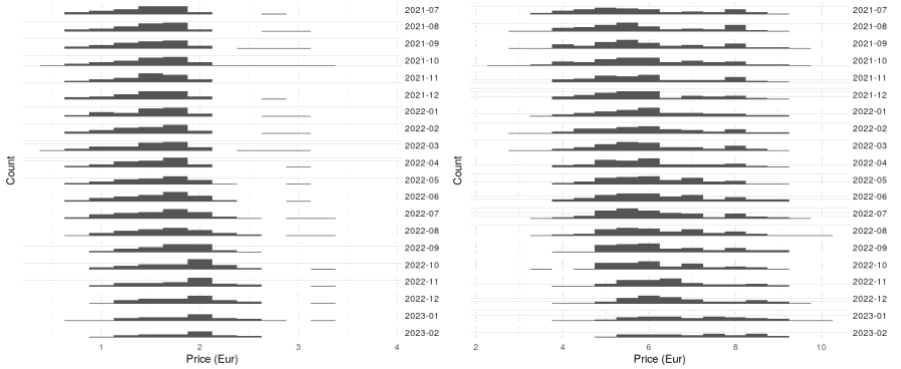
Figure 2: Spatial and Temporal price index for Olive Oil product group according to price band



This rise in LCFP prices could have significant implications for global food security, particularly affecting vulnerable countries in the European Union and specific regions in Italy, and the faster increase compared to higher cost options may also contribute to increase the inequality in living standards across different income classes.

The overall distribution of prices in both Milk and Olive oil categories, presented in Figure 3, confirms our findings. Indeed the left side of the distribution, where LCFP are positioned, presents a substantial shift towards highest prices over time. More expensive products, instead, were substantially less affected.

Figure 3: Overall price distribution for Milk (left) and Olive Oil (right) product groups



References

1. Aizcorbe, A., & Aten, B. (2004). An approach to pooled time and space comparisons.
2. Abay, K. A., Breisinger, C., Glauber, J., Kurdi, S., Laborde, D., & Siddig, K. (2023). The Russia-Ukraine war: Implications for global and regional food security and potential policy responses. Global Food Security, 36, 100675.
3. Benedetti, I., Laureti, T., Palumbo, L., & Rose, B. M. (2022). Computation of High-Frequency Sub-National Spatial Consumer Price Indexes Using Web Scraping Techniques. Economies, 10(4), 95.
4. European Commission (EC). (2020). Practical guidelines on web scraping for the HICP. Harmonised Indices of Consumer Prices. https://ec.europa.eu/eurostat/documents/272892/12032198/Guidelines-web-scraping-HICP-11-2020.pdf
5. FAO, (2002), Report of the World Food Summit Five Years After, FAO, Rome, 2002
6. Herforth, A., Bai, Y., Venkat, A., Mahrt, K., Ebel, A., & Masters, W. A. (2020). Cost and affordability of healthy diets across and within countries: Background paper for The State of Food Security and Nutrition in the World 2020. FAO Agricultural Development Economics Technical Study No. 9 (Vol. 9). Food & Agriculture Org.
7. Marchetti, S., & Secondi, L. (2022). The economic perspective of food poverty and (in) security: An analytical approach to measuring and estimation in Italy. Social indicators research, 162(3), 995-1020.