SERIEs
https://doi.org/10.1007/s13209-025-00309-w

ORIGINAL ARTICLE

Analyzing VAT pass-through in Spain using web-scraped supermarket data and machine learning

Nicolás Forteza[1] · **Elvira Prades**[1] · **Marc Roca**[2]

Received: 13 June 2024 / Accepted: 18 April 2025

Abstract

On December 27, 2022, the Spanish government announced a temporary value added tax (VAT) rate reduction for selected products. VAT rates were cut on January 1, 2023. Initially, the VAT was expected to go back to their previous level six months later, but several waivers led this policy to last for 21–24 months, where the VAT returned to its original level in two phases. We study the pass-through of the temporary VAT rate changes covering the daily prices of roughly 21.000 food products sold online in a Spanish supermarket. We achieve this by utilizing a dataset obtained by web-scraping and employing machine-learning methods to classify each product into a COICOP5 category. To identify the causal price effects, we compare the evolution of prices for treated items (that is, subject to the tax policy) against a control group (food items out of the policy's scope). Our findings indicate that, at the supermarket level, the pass-through was almost complete after one week. In using product characteristics, we notice differences in the rate of pass-through and in pricing strategy in the subsequent weeks.

Keywords VAT pass-through · Price rigidity · Inflation · Consumer prices · Heterogeneity · Microdata · Web-scraped data · Machine-learning

JEL Classification E31 · H22 · H25

This research was conducted during a secondment at Banque de France (BdF) under the External Working Experience program. The hospitality of the Banque de France and especially of the *Service des Analyses Microéconomiques* (SAMIC) is gratefully acknowledged.

✉ Elvira Prades
elvira.prades.illanes@gmail.com

Nicolás Forteza
nicolas.forteza@bde.es

1 Bank of Spain, Madrid, Spain

2 European Central Bank, Frankfurt, Germany

Published online: 25 June 2025

1 Introduction

In the midst of recent economic disturbances, including the COVID-19 pandemic, subsequent supply chain disruptions, and the ongoing conflict in Ukraine, governments have implemented fiscal strategies such as reducing VAT to counteract their effects on price inflation. The reduction in VAT is intended to lessen the effects of elevated inflation, particularly on lower-income households. Those in the lowest-income brackets generally spend a significant share of their overall budget on food. Recent estimates suggest that in 2021 the inflation rate for Spanish low-income families (bottom quartile) exceeded by 2 percentage points compared to the inflation of higher incomes (see Basso et al. (2023)). Consequently, the Spanish government announced on December 27, 2022 and enacted one day later, on 28 December, a temporary reduction in the VAT rate for essential food items starting on January 1, 2023. Due to ongoing inflation in 2023, further waivers of VAT changes were implemented. Finally, by mid-2024, the government announced that the reduced VAT would go back to the previous level in two phases. The effectiveness of these type of measures depends on supermarkets passing on the reduced VAT to final prices paid by consumers. Based on our back-of-the-envelope estimations, a complete VAT reduction in early 2023 would likely have decreased inflation by 0.26 percentage points. Likewise, a full pass-through of a VAT increase would likely lead to a 0.13 percentage point rise in inflation in each phase.

In this study, we assess to what degree and how quickly the pass-through to final prices was passed-through, using a novel data set on daily prices obtained via web-scraping techniques. Employing these micro-level data enhances empirical analysis by offering datasets with advantageous attributes and exceptional detail, though it also introduces certain challenges. The scraped data are massive but unstructured for analysis purposes. To fully exploit the advantages of the web-scraped data, we use state-of-the-art machine-learning techniques to closely mimic the methodology used by the statistical offices to measure inflation and understand price dynamics. In a first step, these raw web-scraped data needs to be properly classified according to the classification used by official statistical offices. The granularity of these novel data sets provides new insights in comparison with official statistics or other kinds of microdata gathered for different purposes such as scanner data or on-site sampling.[1] For that purpose, we use large language models.

We find that prices of treated products respond to the VAT cut: roughly 80% of the VAT cut are passed-through to prices after the first week and 90% after two weeks. The price differentials between the treated and the control evolve differently after these two weeks. A second VAT cut took place in July 2024 and affected only olive oil products, reducing its tax rate from 5 to 0%. In this case, the policy was transferred to final prices completely (full pass-through) within the first two days since the implementation. Then, we find that the pass-through during the VAT reversal event in October 2024 was around 70% after the first two weeks. To exploit the richness of our data we exploit

[1] Utilizing scanner data raises some concerns as regards the aggregation process, and micro-CPI data collected by National Statistical Offices might introduce bias, due to the collection of data via sampling and as products are censored and replaced. This may affect some estimations. For more details on the pros and cons of these data see Cavallo (2018).

different sources of heterogeneity, we also find it relevant to address the price pass-through analysis considering several product dimensions and characteristics through subsample tests. More specifically, we explore the pass-through of (a) processed vs. non-processed, (b) trademark vs. white label (c) imported vs. domestic food products, (d) based on price volatility, and (e) high- and low-ticket items,[2]

When exploring these dimensions, we find that the degree of pass-through is somewhat different across type of products, and in particular, we observe differences in terms of its dynamics over time. When we decompose price dynamics into the price change and the size of this change, we generally observe that this large retailer changed their prices of the targeted products within the first week, as expected. However, the price dynamics over the following weeks shows a heterogeneous pattern. We find that products under the VAT cut scheme passed-through, on average, between 70 and 100% of the tax reduction to final prices over time. As a last exercise, we evaluated the advantages of working with web-scraped daily data in comparison with alternative data sources such as the scanner data, which suffers from averaging bias, or the microdata CPI collected by INE on a monthly basis. The primary benefit of the daily web-scraped data lies in its prompt availability, unlike micro-CPI data, which is subject to delays and is collected on a monthly basis. We show how leveraging on daily data provides more accurate results, as (weekly or monthly) price aggregation leads to underestimation bias when assessing the impact of VAT changes in consumer prices.

Effective January 1, 2023, the VAT tax cut was initially scheduled to expire after six months on June 30, 2023. However, persistent elevated inflation rates required a first extension of this measure until December 31, 2023, and a further extension was announced in June 2024. So finally, the measure was reversed in a first step in October 2024 and it was be fully reversed in January 2025. Bread, flower, milk, cheese, eggs, fruits, vegetables, legumes, tubers, and cereals constituted the 4 to 0% VAT reduction scheme, while in the case of vegetable oils and pastas, the tax cut was from 10 to 5%. The Spanish government estimated that this VAT cut would result in savings of 1.32 billion Euros for Spanish households during fiscal year 2023 (according to Spanish Ministry of Economy, Commerce and Business).

There are other works that have analyzed similar VAT changes episodes in other countries and its effect on prices. In line with this literature on pass-through, our empirical strategy is a simple dynamic difference-in-differences approach. We exploit the fact that the VAT reduction was applied to certain targeted goods, while others of similar characteristics were excluded from this policy measure. For example, the VAT rate was reduced for milk but remained unchanged for yogurts. If the VAT changes had not been enacted for the affected goods, prices would have progressed in a comparable manner. The first test for this assumption is to ensure that price developments or trends were parallel before the VAT change. To efficiently capture the effect of the tax cut on retailer prices, i.e., the pass-through, we will distinguish between treated and untreated

[2] To accurately consider price variations arising from quality differences, it is essential to compare prices using consistent units. As we have the quantities detailed in the description, we can determine the unit price. However, we investigate how products that are less expensive overall yet have a higher unit price impact purchasing decisions. Consequently, our emphasis is on the total expenditure required to purchase a particular product. For example, comparing 5 L of olive oil with a 200-ml bottle, the latter may appear costlier when evaluated in standardized units, such as euros per liter.

products. We calculate the price indexes for each group and run our regression's empirical design (first, by means of Difference-in-differences and, second, using an Event study; see sect. 5), which will give us a closer look to the causal inference of the VAT policy to final consumer prices.

This paper offers three main contributions. Firstly, employing daily data provides a timely and detailed examination of the VAT pass-through, and avoids aggregation bias. We use detailed product-level and web-scraped price data to assess the degree of pass-through on Spanish supermarket prices. To do so, we use a novel data set, the Daily Price Dataset (DPD) collected by the Price-Setting Microdata Analysis Network (PRISMA) in the European Central Bank (ECB).[3] We also make use of an alternative data set provided by Datamarket[4] with the aim of checking the robustness of our results. This data set collects the same type of data, but from a different supermarket and with a different geographical location. These serve to explore whether a retailer with a different pricing strategy leads to the same results. The data collected by Datamarket follow the same principles as used by DPD Prisma. We also use the product description from this source to improve the product classification.

Data collected from web-scraping face the challenges of being difficult to handle, disorganized, and voluminous. So, a second contribution of this paper is that we show the importance of using state-of-the-art machine-learning techniques to classify products into standardized categories. To be able to work with both datasets, all available products must be classified according to the Classification of Individual Consumption according to Purpose (Coicop) at the five-digit level category.[5] At the COICOP5 level, this implies to classify each item at its sub-class level—e.g., "01.1.5.3-Olive Oil". Overall, the aim is to transform the unstructured data into a dataset that is both usable and comparable.[6]

Our third main contribution relates with the measurement of VAT pass-through using daily data. The empirical literature on pricing predominantly relies on monthly or weekly data, which, as noted by Cavallo (2018), can introduce measurement bias due to time-averaging. This bias affects the accuracy of price metrics. We show that our daily data likely provide more precise, less biased estimates compared to official monthly data. Then, we provide more insights about the nature of such time-averaging bias. We investigate the pass-through effect of a VAT cut, comparing daily and weekly data frequencies. One week after the VAT reduction, daily data shows an 87.5% pass-through rate, while weekly averages yield a lower 62.5%, these points that there is bias resulting from time averaging. After one week of the cut, daily estimates consistently exceed weekly ones by up to 5 basis points.

---

[3] For more details about the PRISMA network run by the ECB see link which is a follow-up of the "Inflation Persistence Network" (IPN).

[4] Datamarket is a company which offers web-scrapped datasets on many topics, for more information see here.

[5] We target the Classification of individual consumption by purpose (COICOP) classification by (1) division (two digit), (2) group (three digit), (3) class (four digit), (4) sub-class (five digit) and (5) product.

[6] Several initiatives are in place, but still the resources needed, both in terms of human capital and computation, are immense. As an example, the "Project Spectrum" is a joint project that reflects this need and will encompass the product classification for more supermarkets and retailers, as well as other products sold in other types of stores different from food in supermarkets. For more details, see "Project Spectrum: using generative AI to enhance inflation nowcasting".

In these contributions, we focus on food product items and use modern machine-learning techniques, specifically a pre-trained DistilBERT classifier (Sanh et al. (2020)) trained with human (but AI-boosted) annotated data to predict the categories of Spanish food products Coicop5. With regard to the labeling exercise, and to our knowledge, we explore for the first time the use of Sentence Transformers (Reimers and Gurevych 2019). For the training phase, we follow the seminal work of Hansen et al. (2023). Also, when comparing models performance, we add a new algorithm that has not yet been explored within this kind of applications, based as well on Sentence Transformers (Tunstall et al. 2022). The correct classification of products has proven to be crucial in obtaining a precise estimate of VAT pass-through. The allocation of products to the wrong category can bias the results. For example, if a milk product is assigned in a wrong category, say, under yogurts, which are not under the policy measure, we would get biased estimates, undershooting the impact of the policy. We target the Coicop5 classification as this is the one used by the National Statistical Offices (NSO) to build the Consumer Price Index and its components. The web-scraped raw data contains the product description (for example: *"Strawberry Yogurt, 4 units, BRAND"*). To map each product to the level Coicop5, we do the following. First, we label a subsample (i.e., create the training dataset) relying on semantic search and a pre-trained multilingual AI model. Then, we fine-tune another AI multilingual model using this training dataset.

Once we consider the trained model to be sufficiently optimized by comparing multiple test set metrics, we estimate the Coicop5 for all remaining unlabeled products, so that each product is labeled in a Coicop sub-class, that is, in Coicop5.

In addition to the stylized facts in terms of frequency and the size of price changes, we also check the ability of our microdata to mimic aggregate inflation developments. As a reminder, contrary to the microdata collected by the NSOs, our data are restricted to a unique retailer, and data are not collected nor stored according to the official classification. This step is crucial to properly weight each sub-class when aggregating the different statistics at the group or class level, such as the frequency of price changes or to reconstruct the CPI indices. We take advantage of machine-learning techniques to classify each product and we also take advantage of these techniques to extract certain attributes of each product, such as the country of origin and the brand. This will allow us to explore heterogeneities that cannot be exploited with the alternative data sources.

The paper is organized as follows: Section 2 summarizes the related economic literature. In Sect. 3, we describe our main data sources and briefly summarize the methodology used to map food web-scraped products to the official classification. In Sect. 4, we show evidence of price setting and its dynamics before, during, and after the temporary VAT cut period. And we compare the price index constructed with the web-scraped daily data against the official food price index. In Sect. 5, we estimate the temporary VAT pass-through level. We explore possible pass-through heterogeneities according to product characteristics in Sect. 5.4 and conclude in Sect. 6. Appendix shows a detailed explanation of data cleaning for prices (A) and the product classification approach (D).

2 Related literature

This paper relates with several strands of the literature. First, it relates to the literature that analyzes price setting using microdata. A recent study by Gautier et al. (2024) explores new insights into the evolution of price setting in all economies in the euro area. Their data cover the low inflation period and account for all the underlying prices used to compute the official CPIs, including services. Their data are collected by each National Statistical Office. In our work, we contribute by showing additional evidence at the same granular level and covering the most recent inflationary episode. As a drawback, we only analyze the prices of one components of the CPI, that is, food products, and compute this with information for a country-specific retailer. The data for this more recent period, where there has been a surge in inflation, show that with this data, we can replicate the CPI developments as well as the stylized facts in terms of frequency and size of price changes obtained from other studies in the context of high inflation. That is, as in Nakamura and Steinsson (2008), the frequency of price increases strongly covariates with inflation, whereas the frequency of price decreases or the size of price changes does not. During high inflation, the comovement of price change frequency increases, as the offsetting mechanism of price decreases disappears (see Gagnon (2009)).

Using web-scraped data, we investigate the impact of a temporary VAT rate cut on prices announced by the Spanish government by the end of 2022. While similar policies have been enacted elsewhere, the Spanish policy includes certain aspects that are particularly intriguing for our study. Primarily, we aim to evaluate the extent of pass-through. Secondly, due to various exemptions that prolonged the initial six-month time frame to 21-24 months, we seek to examine how uncertainty influences pricing dynamics. Lastly, we wish to assess the symmetry of pass-through following the policy's withdrawal, specifically determining whether retailers apply VAT reductions with the same intensity as they implement increases. As a recap of the measures taken by other Governments, Germany applied a six-month temporary VAT reduction on all goods. The German government on June 3, 2020, announced a temporary VAT reduction applied to all goods. This measure started the July 1, 2020, and went back to their previous level six months later. This episode was studied by Fuest et al. (2021). Using two supermarket web-scraped data from Germany and Austria (they used the prices of the latter as the control group), the authors find a decrease by 1.3% of prices for Germany, implying a 70% level of pass-through. They found that typically exhibit high pass-through rates; these effects vary significantly across retailers. A similar analysis has been carried out using both micro-CPI data and online prices for German supermarkets using the DPD dataset by Henkel et al. (2023).[7]

In Benzarti et al. (2022), the authors used scanner data to analyze a five-month VAT cut in selected products in Argentina and find evidence of a high degree of penetration, especially for large retailers. Almunia et al. (2023) also analyze the recent Spanish case using a data set similar to ours.[8] They found that the transmission has been almost

---

[7] For more details, see Box 1 Price setting in Germany in the light of the temporary value added tax cut in 2020: evidence from micro-price data. They found that the price reaction to the VAT cut was quick and substantial.

[8] They make use of the supermarket prices collected by Datamarket, a private initiative.

complete and analyze the implications of this broad-based policy measure that affects all income levels instead of using alternative ones, such as the provision of vouchers targeted to the most vulnerable households. Finally, Amores et al. (2023a) collected online prices from a Spanish and German supermarket and compared the average price of selected products the week before and the week after. They found that the VAT pass-through in Spain was large and almost complete for most of the products analyzed. And recently, Bernardino et al. (2024) they have analyzed the impact of the six-month VAT decrease in Portugal. The Portuguese government announced by surprise a VAT reduction on selected food products that went from 6 to 0% in mid-April 2023 and a subsequent reversal in January 2024. They have found a full, persistent, and symmetric VAT the pass-through.

This work also relies on vast empirical evidence on the passing-through of shocks that do have an impact on the cost structure. In summary, we are interested in measuring the impact of a change in prices given a cost shock on final consumer prices. These cost shocks encompass changes in the exchange rate (Burstein and Gopinath (2014)), on tariff rates (Cavallo et al. (2021)), and on input and energy costs (Lafrogne Joussier et al. (2023)) among other producer costs. Furthermore, in relation to previous work that analyzes the impact of shocks that could affect pricing behavior in specific products, Montag et al. (2020) found that the pass-through of the VAT reduction was heterogeneous between fuel types, and Gautier et al. (2022) studied the pass-through of wholesale prices, as a proxy for marginal costs, to retail prices.

Although this study aims to evaluate a policy designed to alleviate inflation for low-income households, the group most affected by rising prices (see García-Miralles, (2023) or Amores et al. (2023b)), limitations in our data prevent us from evaluating its effectiveness. We restrict the analysis to evaluate the degree of a VAT cut pass-through to final prices.

3 Data and machine learning

3.1 Data sources

We used two sources of daily web-scraped prices at the product level. Primarily, we use DPD Prisma, collected by the ECB, and to validate some of the results obtained, we use the data set provided by Datamarket. We also use the second data source with the machine-learning algorithm to obtain more accurate product classifications; this is, we use also the Datamarket product description to obtain a more rich training dataset. The differences between these two data sets lie in the time span as well as in the countries and retailers they cover. It is well-known fact that each retailer follows different pricing strategies and they target different market segments based on their purchasing power, therefore, they have different price elasticities. Using both sets, we can assess the reliability and strength of these experimental data. One possible critique of the online data is if there might be substantial differences with those found online and the share of population that is exposed to these online prices. Testing this is costly, but one of the few studies that tests for this

Cavallo (2017) finds that there are no remarkable differences. To test the validity of our data, we contrast the results of basic analytics such as the reconstruction of the CPI and compare it to the official ones.

3.1.1 DPD prisma

This dataset is the result of an initiative of the Eurosystem's Price-setting Microdata Analysis Network (PRISMA) led by the European Central Bank (ECB). This data set contains daily web-scraped price data from several retailers within the Euro area that cover a wide range of goods sold in supermarkets. Automated web-scraping algorithms collect the relevant product information every day from selected online stores, including the product ID, the product name, the (sub-) category it belongs to, and the current product price at the retail level. Data collection started in April 2022 and we have information up to November 2024.[9] We restrict our analysis to the online price data of the Spanish retailer where we have the information of daily prices for around 21.000 food products, due to the entry and exit of some of the items, we observe each day the price of around 10.000 unique products.[10] We also observe the *shop* to which each product belongs. The *shop* actually refers to the same online supermarket site, but scraped from a different zipcode location. So for each price we account for the triplet: (i) product id, (ii) the retailer id and the location where this product is served and (iii) the date. Please refer to Appendix A for more information on data cleaning and treatment.

Each product contains the following information which is summarized in Table 1, the retailer id, an internal product identifier, we also have a harmonized global identifier, the GTIN-EAN[11] that allows us to proxy the origin of the product, so we can identify whether the product is imported or domestic. We also have the date, the price, the name of the product as well as a description and the section of the product being sold. Using the information in particular from these two columns, we apply GPT-based pre-trained AI models to assign a Coicop5 category that will help us organize the data to be suitable for analysis. Note that we have a very high level of disaggregation that allows for a very precise identification of products. For example, in our data, a *"Yogur de Fresa 4 units MyYogurt"* and a *"Yogur de Fresa 8 units ILoveYogurt"* are two separate items. In Sect. 3.2, we provide a detailed explanation of the classification procedure.

3.1.2 Datamarket

This proprietary data set is obtained by a third party who also uses web-scraping techniques. Although initially it was an open-source dataset, since March 2023 the dataset is available under subscription. This dataset contains information on daily

[9] For more details on the DPD dataset see box 3 "The ECB Daily Price Dataset" in Strasser et al. (2023).

[10] The original dataset contains also information on beverages, cleaning supplies, small electronic appliances and personal care products. Due to the nature of this paper, we take them out of the sample.

[11] GTIN-EAN 13 digit is the identifier of a unique product which harmonized and used by the supermarkets. This identifier allows us to proxy imported products versus domestic products.

Table 1 Basic sample of the dataset

Shop alias [1]	Product id	Date	Name and product desc	Price	Section	COICOP5 [2]	Content	Brand	Treated[3]
ES1a	01234567	01–04-2022	Yogur de Fresa 4u. brand	5.5	Dairies	01147	4 units	MyFavYogurt	0
ES1a	07654321	10–04-2024	Coca Cola zero zero 1.25 l	2.5	Beverages	012XX	1.25 L	CocaCola	0
…	…	…	…	…	…	…	…	…	
ES1a	07162534	01–03-2022	Aceite de Oliva La almazara 500 ml	7.5	Aceites	01153	500 ml	La almazara	1

[1] This is the information available in the raw data set that includes an alias for the retailer, a product identifier, the date where the info is scraped, the name of the product and some description, the section within the retailer were the product is being sold and the price

[2] These are the attributes and useful information that can be extracted using machine-learning techniques based on the product description

[3] This is the information available that we extract such as whether the product has been affected by the VAT measure based on Coicop 5. Whether the product is domestic or imported based on the GTIN-EAN, which is a globally unique 13-digit number used to identify products. Additionally, by using the brand we can identify whether the products are produced by the retailer (what we call white-label) or produced by a third-party (trademark). We identify a white label if the brand corresponds to the name of the retailer

Source: Own elaboration

online prices from three Spanish supermarkets: Carrefour, Dia, and Mercadona. For each day, we have around 5,000 daily observations for Dia, 6,000 for Mercadona, and 2,600 for Carrefour. The data collection starts at some point in mid-2021 and covers all 2022 and up to the first 2,5 months of 2023. All prices include VAT. For details on the treatment of price data, see Appendix A. Given the experimental nature of the data, we limit the analysis to one of the supermarkets because it meets the data quality criteria. We use this data set to compare the results as this supermarket has different pricing strategy to the supermarket in DPD Prisma. Among the challenges of our DPD dataset that it comes from a single chain and retail stores, including the information provided by Datamarket allows us to compare the pricing strategies of retailers facing different demand slopes and also we make use of the product description to fine-tune the language models.

3.2 Product classification

As we want to analyze different dimensions of food products that are being sold online in supermarkets, there is a great concern about the correct classification of food products in a harmonized classification system, such as the COICOP nomenclature.[12] If food products are correctly classified within their respective categories, one can group fresh foods and compare them with unprocessed foods and even identify categories whose VAT has been affected and compare it with another group whose VAT has not been affected. Furthermore, by mapping product descriptions to official statistics codes, we are able to reconstruct the CPI index with higher granularity (at the daily level) and also at the national level and track supermarket food prices in real time. This allows us to track the HICP[13] in pseudo-real time, anticipating the main trends and movements in prices with respect to the official statistics. We contribute to the literature on machine learning and the use of AI techniques in economic research by comparing novel methods that, to our knowledge, have not yet been explored, both for labeling data and for classifying data. A much more detailed walk-through of the product labeling and classification methodology can be seen in Appendix D.

3.2.1 Labeling

We leverage on Sentence Transformer embeddings to perform bulk data labeling under human supervision of Spanish product descriptions. We use a pre-trained multilingual Sentence Transformer embedding to first encode product descriptions into the transformer and second perform a semantic search based on similarity. Transformer embeddings capture the semantic meaning of sentences, and if used with a similarity measure, such as cosine similarity, we can retrieve the best N similar product descriptions for a given product description. After manually reviewing this list, we assign a given COICOP to the whole list, enhancing the labeling process. This approach

[12] Another approach would be to use the classification used by the shop as in Eichenbaum et al. (2011), but this would prevent us from comparing with the official data and would pose some difficulties in comparing between different retailers as they do not follow the same criteria.

[13] Harmonized Index of Consumer Prices.

allowed us to rapidly map product descriptions to COICOP categories in an iterative manner; that is, to go back to the labeling phase when we saw that the classifier was not performing well enough.

3.2.2 Training

Once we have a sufficiently large labeled training dataset, we rely as well on Transformer models to make inferences for the remaining not-labeled dataset. For the training phase, we find that a finely tuned multilingual Distilbert model proposed by Sanh et al. (2020) outperforms other classification methods, as in Hansen et al. (2023). This method achieves an average of 95% F1 score per category of products, which means that on average 5 of 100 COICOP products would not be classified into their true COICOP category. We believe that this 5% error rate cannot be reduced more due to the large amount of classes to be predicted (61 in total). However, this is a huge gain if we compare this method against other simpler methods, such as dictionary-based and/or other classic machine learning, natural language

3.3 Reconstructed CPI indices

To test how representative our data are, we reconstruct the Food-CPI index constructed with granular DPD Prisma data and compare it with the official indices published by the Spanish National Statistical Office (INE) for the Food category. The methodology we apply to reconstruct the price index is that of Cavallo (2013) (see appendix A.1 for technical details). The granularity and daily frequency of the data allows one to generate price indices in a quasi-real-time. To construct the price indexes, we use a combination of the online data and the official category weights. The food price index, that is, for the COICOP's class "01.1", is based on approximately 21.000 food product items. We allow product entry and exit.

Figure 1 plots the reconstructed price index using DPD Prisma against the official food price index (by INE). The price indices take the reference point as of August 2022. Both indices show quite a similar total price increase, but differ in timing. The high data frequency of DPD allow for a more detailed and accurate analysis as monthly data can dilute and underestimate the policy effects on final prices. The aggregate INE data could also hide part of the effect, probably due to its broader coverage of regions and types of retailers that might have different pricing strategies than the largest supermarkets. It is also worth noting the timeliness of our reconstructed price index, as it can cover more recent values due to its daily frequency and data availability almost in real time.

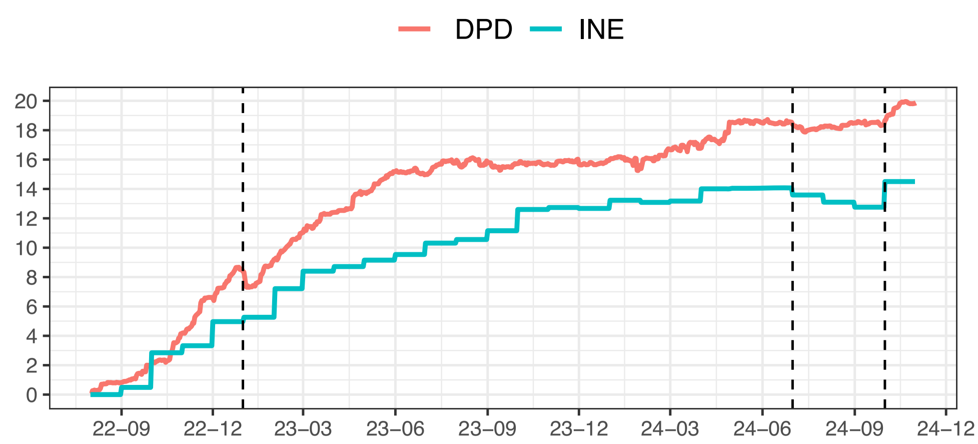


Fig. 1 Online and Official Food Price Indexes. This graph plots the daily price indices using price data from DPD PRISMA(red) compared to the monthly official index by the Spanish National Statistical Office (INE) (blue). The reconstructed index uses the cumulative product of the daily price differences, using geometric weights within each Coicop5 category, and then, each category is weighted with the share of consumption of every Coicop5 from the National Statistical Office. The date when the VAT changes took place are represented by the dashed vertical line. *Sources:* National Statistical Office (INE in Spanish) and the authors' own calculations using the Daily Price Dataset (DPD) Prisma ECB

4 Price setting during VAT changes

Before turning to our investigation on the price adjustment in response to a VAT change, we find useful to provide a description of the characteristics of our dataset as well as the main facts in terms of the frequency of price changes, the mean size of nonzero price changes as well as other moments that describe the distribution of price changes, such as the skewness and kurtosis. We briefly describe the general behavior of retail prices during normal times as well as around the dates where the VAT is modified.

Retail prices change frequently, in normal times, on average, 10.5% of the products changed their price each week. This leads to an implied duration of 10 weeks, this means that prices change every 2.6 months. Although there is a notable heterogeneity between products. For example the sub-class ”01.3.1-*Fresh or Chilled Fish*” records a frequency of 23% that implies that prices, on average, change once per month. And in the other extreme there are Coicop5 sub-classes that change prices every 4.5 months. This statistic relates to the extensive margin, which is the frequency of price changes. Additionally, it is important to study the magnitude of these price changes, known as the intensive margin. We conduct this analysis by examining histograms that depict the distribution of nonzero price changes.

On the frequency of price changes. To compute the frequency of price changes, we first calculate the fraction of all daily (or weekly) price changes over the life span of each product and then calculate the median frequency on all goods within a Coicop 5 group.[14] We also calculate the share of products that change prices on a weekly basis, to track their evolution over time.[15] In Fig. 2, we present the evolution of the weekly share of price changes, that is, the proportion of products that have changed prices in a given week within each treated group. During the first week of January 2023, among affected food products, the aggregate fraction of the price change increased to 60%, which means that a large share of the targeted products registered a decrease in price. This is a different pattern from previous and subsequent weeks, where the frequency of price changes was mainly driven by price increases and where the share of price changes hovered around 10%. During the first phase of the VAT reversal that took place in October 2024, a similar pattern is observed, with the proportion of price changes maintaining a similar level. However, the proportion of products that have increased in *treatment 2* appears more moderate.

4.1 The distribution of nonzero price changes.

The whole distribution of price changes might be important to understand the transmission of monetary policy, and therefore, we provide a description of the full distribution

[14] Given that some products life span is very short and that can bias this calculation. We remove products whose prices are available for a short period of time.

[15] It is important to point out that for product spells with identical duration, calculating the frequency of price changes will yield the same result whether done on an individual product basis with a subsequent averaging or by calculating the proportion of price changes on a specific date and then averaging across products. In contrast, when the durations of time spells vary between products, such computation could introduce a bias.

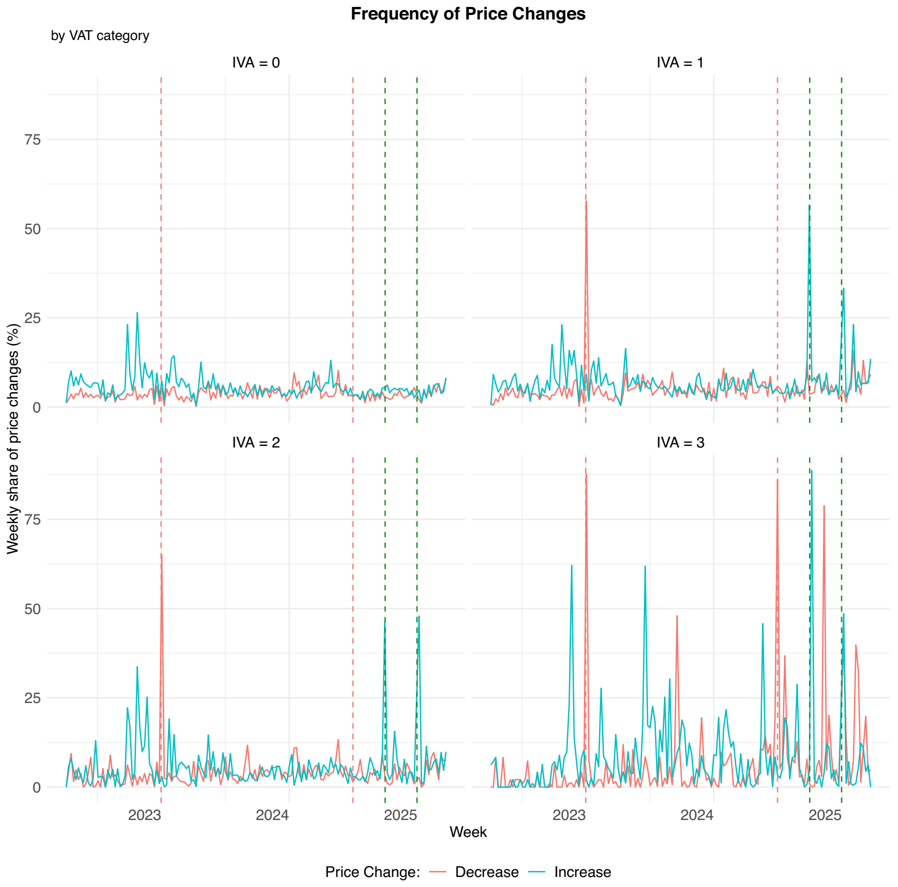


Fig. 2 Share of weekly price changes overtime. These graphs plot the share of price changes of CPI items affected by the VAT reduction within the control (T0) and each treatment group: T1, T2 and T3. The blue lines show the share of price increases and the red lines the share of products that have reduced their price in a given week. The share of price changes are computed at the shop and Coicop5 level. To aggregate we weight each Coicop5 using the INE weights and then compute the simple mean by shop. Vertical lines in red indicate the dates of the VAT cuts and green vertical lines the dates of the reversal. *Sources:* Author calculations based on DPD Prisma ECB

of price changes. One might wonder if the pattern of price changes has changed after the implementation of the policy.

To assess this, in Fig. 3 we plot the distribution of the size of nonzero price changes, and we split the sample taking three weeks before and three weeks after the policy. We exclude from the sample the previous week and the first week when the policy implementation took place; in this way we avoid the exceptional bunching of prices changes around a specific value that is solely due to the VAT change. We also split the sample among the three sample groups according to the VAT treatment. A visual inspection does not provide a clear message. After the policy is implemented, there are almost no changes in the first and second moment in the distribution of nonzero

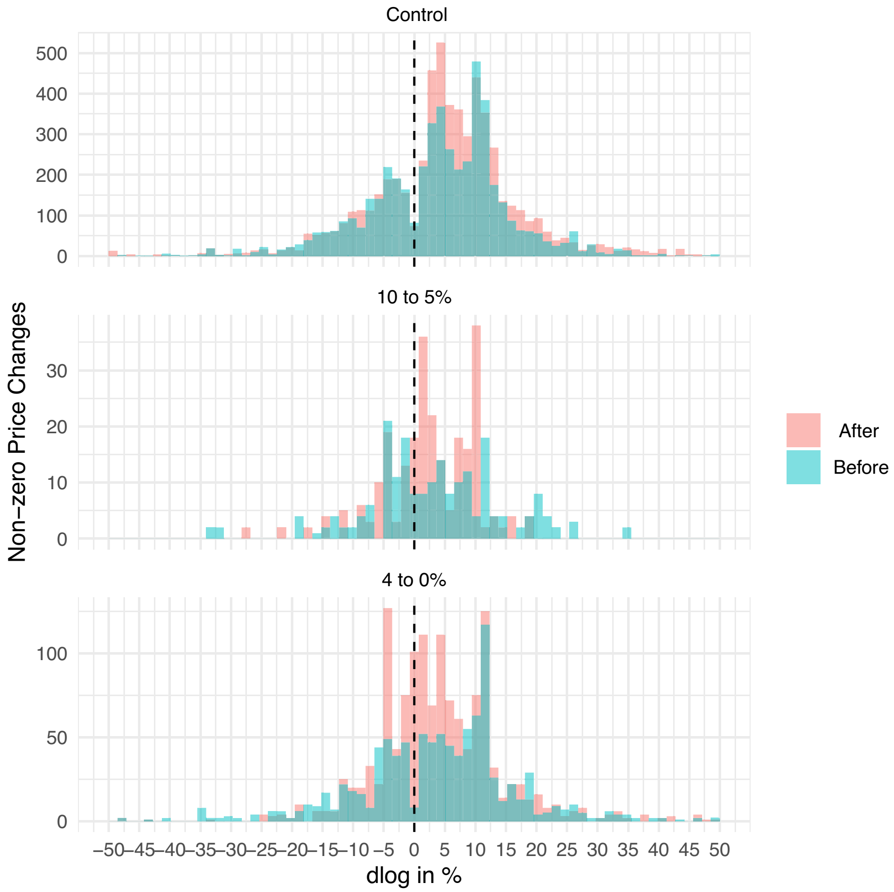


Fig. 3 Distribution of Daily Price Changes around the VAT Cut. Distribution of daily price changes during the first VAT cut: the *after* period (red shaded area, January 7, 2023, to January 30, 2023), and the *before* period (blue shaded area, December 1, 2022 to December 23, 2022). *Sources:* Authors' calculations based on DPD Prisma ECB

price changes. In Table 2, we can see that the mean increase in price in the control group lies around 10%, before and after. There is a small decline in the T2 group (from 10 to 5%) and a sharp decrease in T1 (from 4 to 0%).

4.2 On the skewness and kurtosis of price changes.

The skewness and kurtosis, that is, the third and fourth moments of the distribution, provide interesting insights. In Table 2, we report the skewness and kurtosis. The kurtosis of the distribution of nonzero price changes is a measure of the "fat tails" of the distribution. This is interpreted as follows, for a given frequency in price changes, the size of the price changes have increased. This can be the case when firms change

Table 2 Descriptive statistics of nonzero price changes

		After			Before		
		Control	10 to 5%	4 to 0%	Control	10 to 5%	4 to 0%
Decrease	Mean	−9.84	−4.98	−6.26	−9.11	−6.94	−10.96
	SD	9.79	5.22	7.59	8.15	7.87	12.15
	Skewness	−2.67	−2.94	−4.05	−2.18	−2.07	−2.60
	Kurtosis	13.70	12.85	27.91	10.82	7.03	12.05
Increase	Mean	10.53	6.37	9.74	10.13	9.97	12.20
	SD	8.59	4.67	10.16	7.45	7.59	11.53
	Skewness	2.45	0.58	2.79	2.12	0.96	3.03
	Kurtosis	14.34	2.64	13.42	12.21	3.53	15.69

This table shows descriptive statistics of the daily web-scraped, non-zero price changes distribution (in logs) of the after period (7th January, 2023, to March 1, 2023), and the before period (1st November, 2022, to 23rd December, 2022).

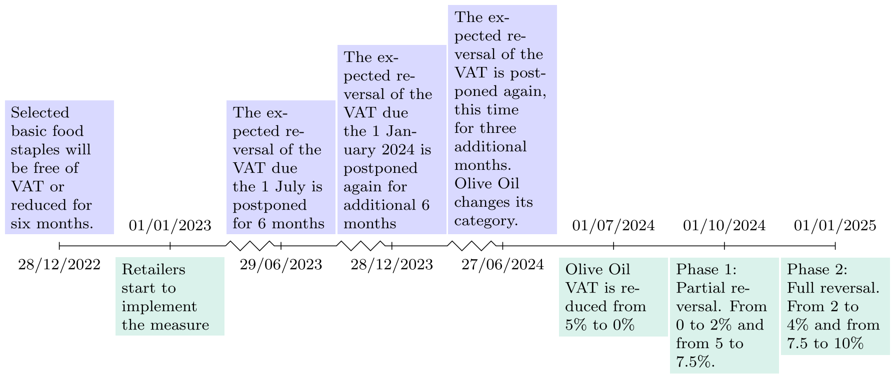


Fig. 4 Timing of VAT announcements and implementation dates. *Source:* Own elaboration based on the official Spanish Government press notice. Blue boxes refer to government announcements; green ones to measure implementation, either a VAT cut or a VAT reversal

prices with a lower frequency, and it is more likely that the current price is far from the desired price. Therefore, in the event of a price revision, the size of the change will be larger.

Comparing the two periods, the different treatment goods and the control, we observe an increase in the kurtosis for T1 when looking that the price decrease. And a generalized decrease in all the treatment groups, but with different intensities.

5 Price effects of the temporary VAT rate change

In this section, we study and quantify the VAT cut (and reversal) pass-through on final prices. First, we will describe the empirical approach and examine the overall results. Then, we will exploit the richness of our data set by providing heterogeneous effects of the treatment. This policy was initially expected to last six months, but after multiple extensions due to a persistent high inflation, the Spanish general elections that took place in July 2023, and an uncertain economic context, it ended up persisting for 21 months. The first announcement of the VAT rate cut on groceries occurred on December 28, 2022, and retailers were expected to implement it by January 1, 2023. Initially, the measure was set to last for six months, until July 2023. However, due to persistent inflation, multiple waivers were issued, causing periods of uncertainty and additional two six-month extensions, until July 2024. Then, in July 2024, instead of bringing back the VAT rates up, the Spanish government decided to cut the VAT for olive oil further and postpone the policy reversal for another three months. Finally, in October 2024 the first phase of the fiscal policy reversal policy reversal arrived, setting higher VAT for groceries under the VAT cut, but still lower than before the initial cut (Fig. 4).

The ongoing modifications to the temporary fiscal policy schedule and the varied timing of announcements could have influenced price trends and pricing approaches of retailers. In the following section, we examine the impact of this three-step fiscal policy, exploiting several heterogeneity dimensions, but we also delve into the effect of the multiple policy waivers.

5.1 General framework of the empirical strategy

In this quasi-experiment, we use three types of treatments. The first treatment group (T1) contains products whose VAT rate was reduced from 4 to 0% in December 2022 and increased to 2% in July 2024 (bread, flower, milk, cheese, eggs, fruits, vegetables, legumes, tubers, and cereals). The second group (T2) the tax cut was from 10 to 5% and increased to 7.5% in October 2024 (pastas and vegetable oils—excluding olive oil). Finally, the third group (T3) is composed of olive oil, which was subject to two VAT cuts in December 2022 (from 10 to 5%) and in July 2024 (from 5 to 0%) and then a VAT increase in October 2024 to 2%. These treated groups are compared with a control group consisting of the food products sold in this supermarket not affected by any of the VAT changes mentioned. We restricted the analysis to food products and removed beverages and other products such as household appliances and personal care products from the sample. In this way, our control is expected to be more similar to the treatment group.

The price evolution for the three treatments and control groups is shown in Fig. 5 where we also compare the performance of the official index of the Spanish statistical office, INE (as in Fig. 1). At the time of writing this article, INE has not released the October 2024 values for the Coicop 5 Consumer Price Index, so we have not been able to compare the time frame of policy reversal. This again highlights the benefits of utilizing the Daily Price Dataset, which facilitates real-time analysis of supermarket

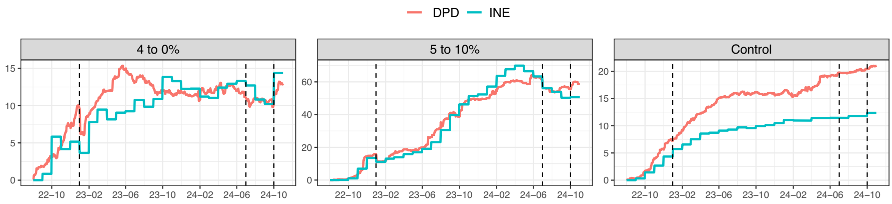


Fig. 5 Online and Official Food Price Indexes by VAT group. This graph plots the aggregate reconstructed daily index (red) with online prices and official CPI weights compared to the official monthly index for the Food category constructed by Spanish National Statistical Office (INE) (blue). The reconstructed index uses the cumulative product of the daily price differences (log) using web-scraped data from DPD PRISMA, aggregated by means of a geometric mean within each Coicop 5 category. Then each category is weighted by the share of consumption of every Coicop5 within the food group. The date when the VAT changes took place are represented by the dashed vertical line. *Sources:* National Statistical Office (INE in Spanish) and the authors' own calculations using the Daily Price Dataset (DPD) Prisma ECB

prices. The classification of each Coicop 5 into the different treatment categories together with the weight of each category in the overall CPI is summarized in Table 3.

We run the following Difference-in-differences Event Study regression[16] as in Fuest et al. (2021) to see the daily effects of the policy on consumer prices and the degree of the VAT cut pass-through to consumer prices:

$$p_{ist} = \sum_{j=-a,\neq -1}^{T} \beta_j \times b_{it}^{j} \times (\mathrm{FPT}) + \theta_t * \mathrm{COICOP}_4 + \mu_{is} + \varepsilon_{ist} \tag{1}$$

The outcome variable of interest, p_{it}, is the natural logarithm of the daily (t) price for each product (i) sold in store s. We use the day before the implementation of the VAT cut or increase ($j\ / = -1$) as the reference point for the daily coefficients, j tracks the number of days or weeks since the introduction of the VAT measure, a is the lead and T will denote the last data point observation included in the regression. The dummy variable b^j determines whether the product i sold at store s at time t is affected by the VAT change scheme or not. Since we have a three-set of treated groups with a different tax reduction and policy reversal, with the *FPT* term ("full pass-through") we standardize all price changes to interpret the two effects at the same scale. That is, we multiply the dummy with what would be the price change in the case of a full pass-through on final prices. As an illustration, the VAT reductions in December 2022 would result in a − 3.85% change for items reduced from a 4% to a 0% rate and a − 4.55% change for those with a rate reduced from 10 to 5%, assuming complete pass-through. An estimated β_j=− 1 would indicate a full pass-through VAT of the cut (or β_j=1 in the case of a VAT increase) to final prices.

The term [$\theta_t *$ $COICOP_4$] interacts with time dummies with Coicop at 4 digits to capture specific time trends, that is, at the class level (e.g., 01.1.1 "Bread and Cereals"). Finally, μ_{is} stands for product-store fixed effects, and we cluster standard errors by

[16] We opt for this specification rather than the usual one found in pass-through literature where the dependent variable is the price change (Δp_{it}), to prevent occurrences of zeros since products might not update their prices often.

Table 3 Classification of COICOP5 into treatments and controls.

Treatment 1				Control			
Temporary 0% VAT (from 4%)				Standard 10% VAT			
		CPI Weight	Type			CPI Weight	Type
01111	Rice	0.903	Proc	01115	Pizza and quiche	1.63	Proc
			Proc	01117	Breakfast cereals	0.867	Proc
01112	Flours and other cereals	0.435	Proc	01118	Other cereal products	0.667	Proc
01113	Bread	10.512	Proc	0112	X Meat	42.626	Unproc
01114	Other bakery products	8.919	Proc	0113	X Fish	19.599	Unproc
01117	Breakfast cereals	0.867	Proc	01144	Yoghurt	3.867	Proc
01141	Fresh whole milk	2.029	Proc	01146	Other milk products	1.842	Proc
01142	Fresh low fat milk	3.214	Proc	01151	Butter	0.524	Proc
01145	Cheese	6.908	Proc	01163	Dried fruit and nuts	2.864	Unproc
01147	Eggs	2.146	Proc	01164	Preserved fruit and fruit-based products	0.4	Unproc
01161	Fresh or chilled fruit	14.199	Unproc	01172	Frozen vegetables other than potatoes	0.55	Unproc
01171	Fresh of chilled vegeta-bles	9.342	Unproc	01173	Dried vegetables	4.316	Unproc
01174	Potatoes	2.439	Unproc	01175	Crisps	2.407	Unproc
Total T1		61.913		01181	Sugar	0.556	Proc
Treatment 2				01182	Jams	0.866	Proc
Temporary 5% VAT (from 10%)				01183	Chocolate	2.502	Proc

Table 3 (continued)

Treatment 1				Control			
Temporary 0% VAT (from 4%)				Standard 10% VAT			
		CPI Weight	Type			CPI Weight	Type
01116	Pasta products and cous-cous	1.701	Proc	01184	Confectionery products	1.756	Proc
01154	Other edible oils	1.074	Proc	01185	Ice cream	1.63	Proc
Total T2			2.775	01191	Sauces and condiments	2.599	Proc
				01192	Salt, spices and culinary herbs	0.812	Proc
Treatment 3				01193	Baby food	0.742	Proc
Temporary 5% VAT (from 10%) to 0%				01194	Ready-made meals	7.611	Proc
01153	Olive Oil	6.133	Proc	01199	Other food products n.e.c	3.206	Proc
Total T3		6.133					
Total treated		7.0821		Total control in %			104.439 10.44
in %		7.0821		Processed in %			7.65
				Unprocessed in %			9.87

This table shows the split of our data into treatment and control categories. Treatment categories are determined based on the "Real Decreto-ley 20/2022, de 27 de diciembre". Residual categories in our food data encompass the control products. The HICP weights correspond to the year 2023. *Sources:* Own elaboration and INE

product-store (ε_{is}). This approach provides estimates for day-specific relative price adjustments in response to VAT rate changes for the treated products compared to the products of control, taking the day before the implementation as the base period.

In this identification strategy, it is crucial that some assumptions hold so that we can establish causal relations. In a D-i-D setting, the most relevant assumption is common trends; treatment and control groups should not have statistically significant differences in trends before treatment is implemented, making sure that the outcomes for treatment and control groups moved in parallel prior to the time of treatment. We will also assume that the given treatment has no causal effect before its implementation, which in the economic literature is referred to as the no-anticipation assumption. This assumption has two main implications. First, it argues that agents do not change their behavior in anticipation of treatment in ways that would affect the outcome. Second, it states that the mechanism used to decide the treatment and control groups is not based on systematic differences in potential outcomes prior to policy intervention (Wooldridge 2021). These are examined in detail in the next section 5.2.

5.2 Main results: policy implementation and reversal

We conduct a regression analysis on all food items to compare the group of products affected by the VAT measure with the control group. The treated group accounts for 7% of the Consumer Price Index (CPI) and 21.5% of the items in our dataset. In contrast, the control group constitutes 10.9% of the CPI basket and includes 78.5% of the products analyzed. We analyze the effect of the VAT reduction initiated on January 1, 2023. Subsequently, we review the outcomes of the additional VAT reduction only for olive oil products in July 2024. Lastly, we perform a comparable assessment to evaluate the VAT partial reversal in October 2024 and its complete reversal on January 1, 2025. Figures 6a and b shows the event-time coefficients β_j for each VAT cut episode based on the regression specified in Eq. 1. Figures 7a and b presents the event-time coefficients associated with the VAT reversal events, which occurred over two distinct phases.

Our specification's coefficients before the different events (VAT cuts and reversals) are not statistically different from 0 and do not follow a specific trend; the "0 effect threshold" is crossed on repeated occasions and inside the 95% confidence bands, which suggests that the parallel trends assumption is likely to hold. To compare the pass-through in all these four episodes: Products that were under the VAT changes scheme passed-through, on average and after a week, around 90% in the first cut in 2022, 100% in July 2024, 75% in October 2024 and 50% of the tax changes to final prices. It is also worth mentioning how in the VAT cut episodes, consumer prices dropped significantly in the first day after the policy entered into force, whereas for the first phase of the VAT reversal it was not until the third day where we see the stabilized pass-through estimate. In the case of the second phase of the reversal, we observe a certain degree of anticipation, as the price difference between groups is observed one week before.[17] This denotes a certain degree of asymmetry in the pass-through.

Throughout time, variations in pricing between the control and treated groups reveal further dynamics, suggesting that prices are initially modified in response to the VAT measure, with adjustments to their preferred level occurring a few weeks afterward.

Three treatment groups have been identified depending on the degree of the VAT change. According to the Spanish Ministry of Economy, products that became VAT-free (T1 and then T3) are considered as "staple food", while the other group (T2) is labeled as "second class" items. Among the 49 Coicop5 sub-classes, T1 includes products in 12 distinct categories, whereas T2 has items in just two categories, and T3 contains products in only a single category. Food items constitute 17.9% of the total HICP, with the processed products making up 7.08% (see Table 3).

[17] We suspect this is due to the holiday season where the non-treated goods by the VAT measure registered some differentiated price-setting dynamics.

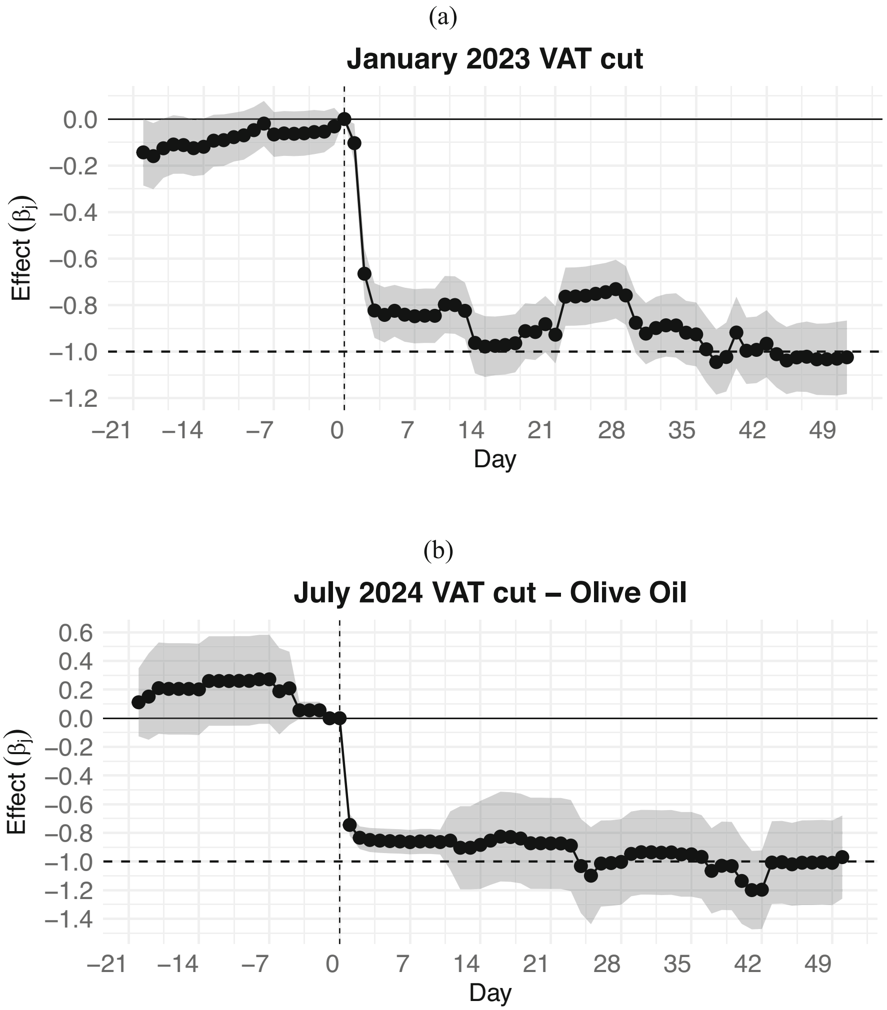
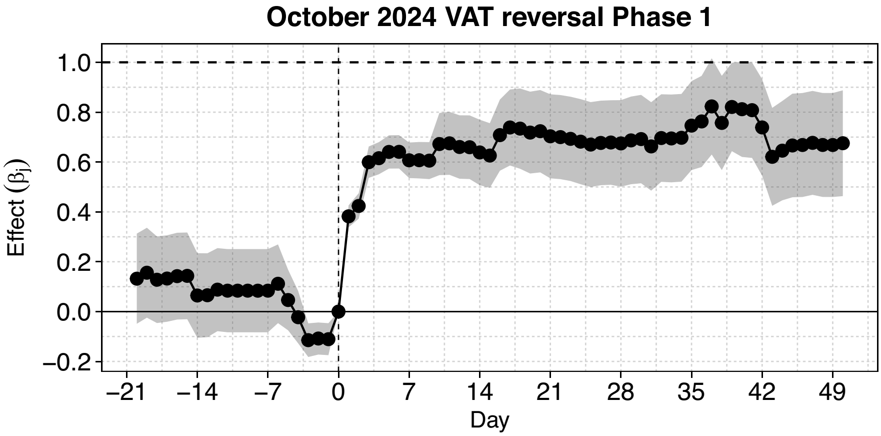


Fig. 6 Event study: VAT cut pass-through estimates (I). This figure shows the estimates of the degree of VAT pass-through. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the last day before the VAT cut (Day -1), which is the time period we take as the reference. The horizontal dashed line accounts for the full pass-through, in negative for the VAT decrease. *Sources:* Authors' calculations based on DPD ECB Prisma

Furthermore, as shown in Fig. 8, when splitting the staple and second class groups, we find almost negligible different treatment responses the first days after the VAT measure, but we observe divergent patterns after the third week. One possible explanation is that large retailers are scrutinized during the first weeks. Afterward, the price setting is determined by other shocks and pricing strategies.

On one hand, the reaction of staple food products mirrors the results seen earlier in Fig. 6, although the degree of price pass-through of the VAT reduction is lower:

(a)



(b)

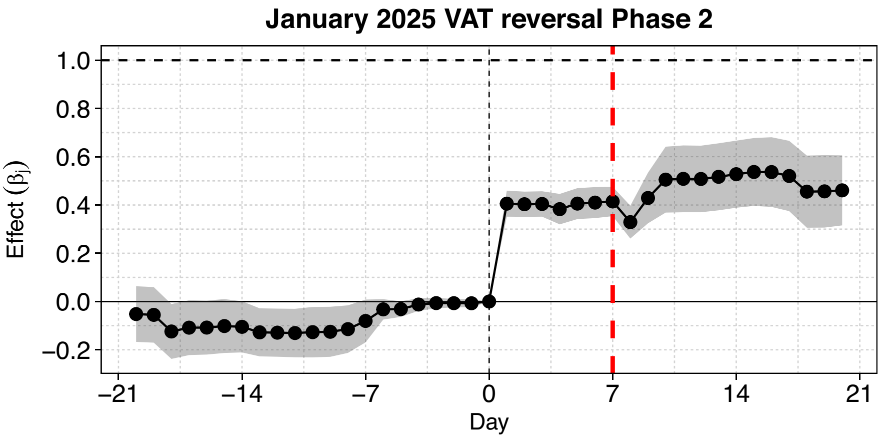


Fig. 7 Event study: VAT reversal pass-through estimates (II). This figure shows the estimates of the degree of VAT pass-through. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the last day before the VAT cut (Day -1), which is the time period we take as the reference. The horizontal dashed line accounts for the full pass-through, in positive for the VAT increase. *Sources:* Authors' calculations based on DPD ECB Prisma

approximately 85% in the first week following the VAT reduction and 60% in the initial week after the reversal. Second-class products appear to even overshoot the price decrease required by the fiscal measure, with a quick response with a 100% pass-through in the first two weeks and a further decrease after one month. Then, in the policy reversal they get a lower pass-through rate. Products under the 10- to 5% cut, identified as the T3, experienced sharp price increases, showing a sharper upward trend pattern throughout October and November 2022, which could negatively

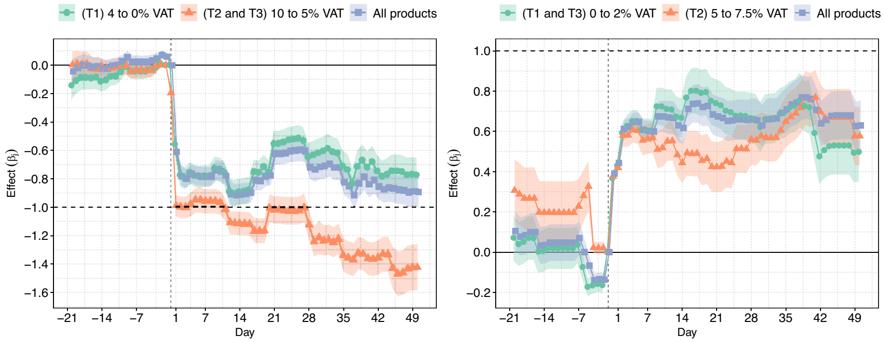


Fig. 8 Event study: Comparing VAT Pass-Through among Treated Groups. This figure shows the different estimates of the degree of VAT pass-through when we break down different treated products into the different VAT treatment groups. In orange the degree of pass-through in treated products that belong to T2 and T3. In blue the degree of pass-through in the treated products in T1. In green the baseline estimation. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

impact the validity of our model assumptions. This raises two concerns regarding the model's main assumptions: trends are less likely to be parallel between treatment and control (recall, the parallel trends assumption), and the mechanism to decide which subjects fell under the policy scope could have been based on outcomes before the VAT intervention (no anticipation assumption). The first issue is tackled using the [$\text{COICOP}_4 * \theta_w$] interaction term in Eq. 1, which captures specific product time trends. For the second obstacle, though, we assume that the Spanish government pursued helping lower-income households by alleviating the tax pressure on products that have larger weight on their consumption basket, not due to a correction to some products' price behavior.

In sum, the event-time coefficients for the second treatment group (in blue) might be upward biased due to the large trend change before the policy. That could be the reason behind these larger than 100% price pass-through β_js, which otherwise would be difficult to understand from the retailer's point of view.

5.3 The role of policy uncertainty: prices around the expected reversal

This section explores the price-setting dynamics before and after the various waivers of this policy measure. As shown in Fig. 4, the VAT "temporary" measure was extended up to three times. As a result, the policy originally intended to last six months was eventually extended to 21 months. The official announcements concerning the extension of this measure were made at different times relative to their respective deadlines: the initial waiver was revealed merely 3 days prior to its expiry; the second extension was declared 45 days before its scheduled expiration, and the third was communicated 20 days ahead. We observe a common pattern during the previous days of the official expiration. Government officials start to announce informally the likelihood of the extension until it is made official and published as a Royal Decree a few days before the expiration. In addition, the reversal was conditioned by developments in core inflation. If core inflation recorded in September 2023 an increase year on year below 5. 5%, then the reversal of the measure would have been brought forward to 31 October 2023 RDL 5/2023 on 28 June. This scenario did not occur, and the VAT reduction in selected products remained unchanged.

We apply the same specification as depicted in Eq. 1, but limit our sample to the periods surrounding the announcement and the anticipated reversal. In Fig. 9, we plot the beta coefficients that account for the different prices evolution between the treated group and the control group. Despite variations in the duration of time intervals and the different inflationary contexts, we can see that the price dynamics vary between the two groups. In the last announcement that took place in June 2024 we observe

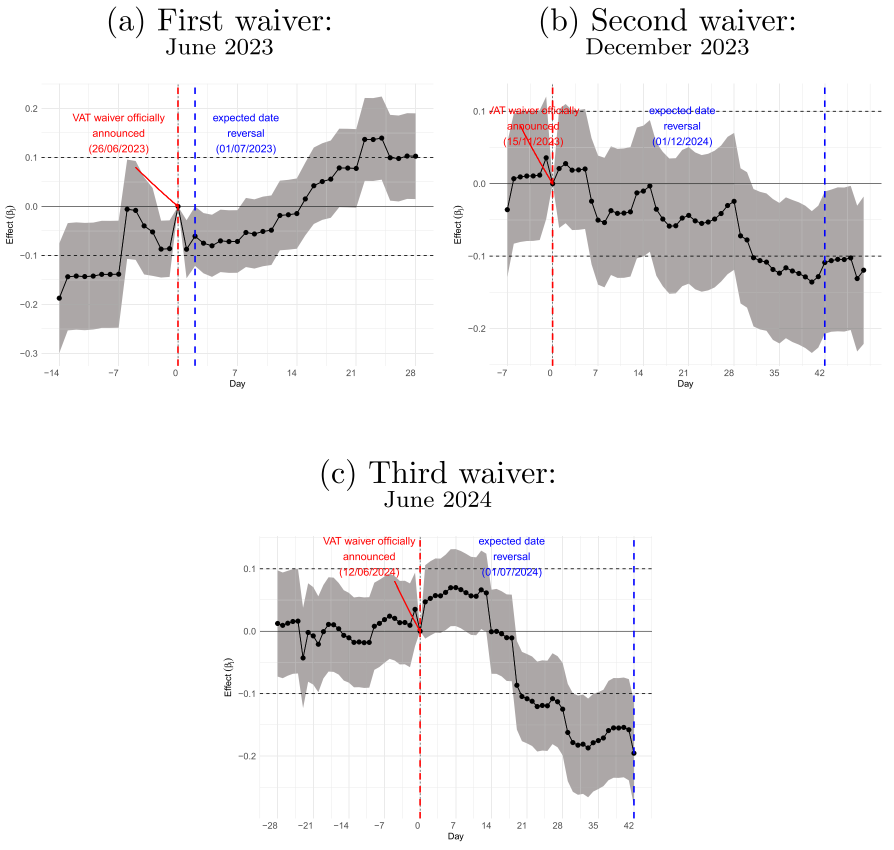


Fig. 9 The role of policy uncertainty in price dynamics. Graphs 9a, 9b and 9c explore the differential in price dynamics of the goods affected by the VAT policy. The coefficients reflect the price variation of the treated products relative to the control group setting the base the price differential on the day when an official announcement is made, stating that the measure will be extended for an additional six months with respect to the due date. The vertical red line accounts for the day of the announcement and the blue line for the due date of the policy reversal

that the price dynamics of the two groups remained unchanged. And only when the expected date arrived, the treated group started to reduce their prices.

5.4 The role of product heterogeneity and implications

The VAT pass-through may vary depending on certain characteristics of the product. In Fig. 10, we plot the estimates of the VAT pass-through for selected Coicop5 categories. We can observe that the degree of pass-through differs between categories. We observe that some categories register a full pass-through just a few days after the implementation, and in other cases not. We also observe a differentiated pattern over time. In this section, we explore the characteristics that can explain differences in the degree of pass-through. As in Eq. 1, we control for the product-store fixed effect (μ_{is}), we cannot add another product characteristic that does not change over time as a covariate, or we would run into collinearity problems. Therefore, we run multiple regressions for by splitting the samples to examine different behaviors in product characteristics that can be identified within the DPD Prisma dataset, namely Processed vs Unprocessed (in sect. 5.4.1), Trademark vs White-label (sect. 5.4.2), and Domestic vs Imported goods (sect. 5.4.3).

In Table 4, we report the number of products that fall under these different categories. Although the origin product subsamples are quite balanced, there are far fewer

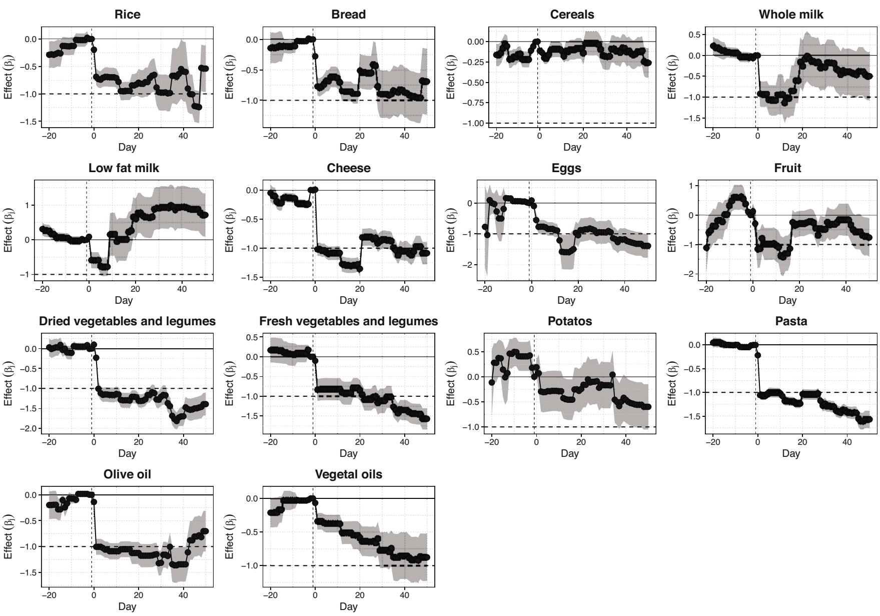


Fig. 10 Event study: VAT Pass-Through by Coicop 5. This figure show the estimates of the degree of VAT pass-through for different treated products by Coicop 5 category. Each coefficient bandwidth represents the 95% confidence interval. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

unprocessed products than processed products, which may pose some concerns of the statistical power of our estimates. In order to overcome this issue, we also explore how other time-varying product characteristics, such as those related with the price level, can lead to different pass-through behavior. Specifically, we split the sample of the food product according to its price volatility, which is ultimately related to the frequency of price changes, and whether the product is a big ticket product. This is explored in sub-sect. 5.4.4 and 5.4.5.

To quantify the effect of treatment across these categories, we first approach this question from a simpler point of view, looking at narrow time spans (three weeks before and three after the tax cut: $w \in \{-3:3\}$) to address the immediate policy response. This gives us an idea of what we can expect given some degree of heterogeneity. In Eq. 2, we regress on several outcome variables (y_{ist}): first on the level prices (in logarithms) in order to capture price differences among the compared groups. Then, we compare the pricing behavior by calculating the probability of a price change in product i at time t (Probit model). We also distinguish whether this price change is positive or negative.

$$y_{iw} = \beta_0 + \beta_1 H_i + \beta_2 T_i + \beta_3 A_w + \beta_4 H_i T_i + \beta_5 H_i A_w + \beta_6 T_i A_w + \beta_7 H_i T_i A_w + \varepsilon_{iw} \tag{2}$$

H_i is the heterogeneity characteristic under study, A_w is a dummy that takes the value 1 after the VAT cut (January 1, 2023) or 0 otherwise, and T_i indicates whether the food item is treated, that is, subject to the VAT policy.

5.4.1 Processed vs unprocessed

When discussing price inflation, the economic literature usually distinguishes between general and core inflation. The latter excludes products that are not processed and are usually more volatile, as they can be affected by international and exogenous shocks, for example, the COVID-19 pandemic or the Russian-Ukrainian war. Therefore, distinguishing between processed and unprocessed items will shed light on how this difference will be understood in our study. We explore this source of heterogeneity by categorizing each product in our sample following the Eurostat E-Coicop -HICP criterion, which assigns food products to processed or unprocessed groups using Coicop categories at the 5-digit level; 38 Coicop5 sub-classes are labeled as processed (91.8% of the products, 7.8% of CPI), and the remaining 11 belong to the unprocessed food group (8.2% of the products, 9.9% of CPI).

Table 5 presents selected coefficients from Eq. 2 results. The first takeaway is that processed goods are, on average, 54% cheaper compared to unprocessed.[18] Next, we focus on differences in the probability of a price change. In the case of a processed product it is $exp(0.136) = 1.15$ more likely to have a price change compared to a non-processed.

Having explored these differences in various outcomes, we move on to the subsample analysis to exploit this consistency of treatment heterogeneity over time as in Eq. 1. Figure 11 plots the results for processed goods (in orange): close to 80%

[18] Log-lin model coefficient interpretation: Effect $= exp(\beta) - 1$.

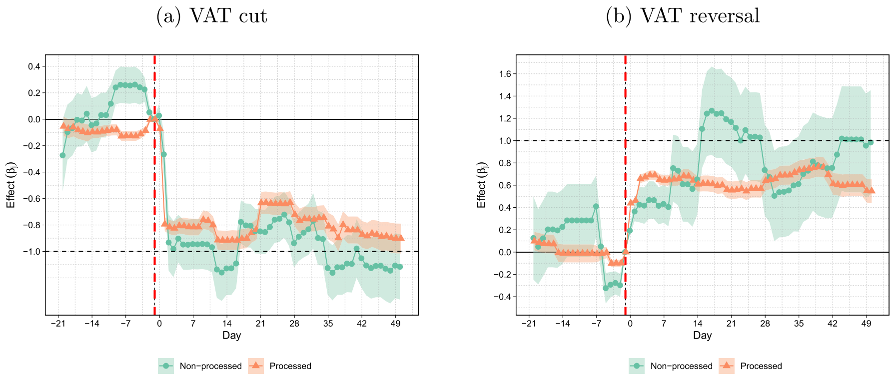


Fig. 11 Event study: Processed vs. Unprocessed. This figure shows the different estimates of the degree of VAT pass-through when we break down different treated products into processed and unprocessed foods. In orange the degree of pass-through for processed food products. In green for the unprocessed. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store id level. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

pass-through and 90% for unprocessed foods after one week. Examining the pass-through estimates for the VAT reversal event, we notice an asymmetry in the extent of pass-through compared to the VAT cut estimates. Within one week, processed foods exhibit a 70% pass-through, whereas non-processed foods achieve a 50%, along with greater price-setting volatility in subsequent weeks.

5.4.2 Trademark vs White-label

The price-setting behavior of white-labeled and trademarked products could also be quite different. Using text analysis techniques, we have been able to label 90.1% of the products to a specific brand. Then, we could assign whether they correspond to the trademarked product (61.9%) or to the supermarket white label (38.1%). Supermarket chains may have less price adjustment capability with trademarked products, due to producer contracts and agreements restricting retailers' actions. Therefore, we would expect white-labeled products to pass-through the VAT cut to final consumers more rapidly than third-party producers that have their own trademark.

Table 5 exhibits Eq. 2 results for trademark heterogeneity. The first important highlight is the difference in price levels between these two groups: trademarked items are, on average, 30% more expensive than white-labeled items. And there are not significant differences in the probability of price changes.

Figure 12 supports this last finding. Trademarked and white-labeled products both showed a large pass-through (around 85% during the first week). Although there have been some disparities over time, there is no clear pattern of divergence between these two groups. Then, during the reversal episode, the white-label products' pass-through was more stable during the 2 months than the products with trademark, who experienced a more volatile pass-through.

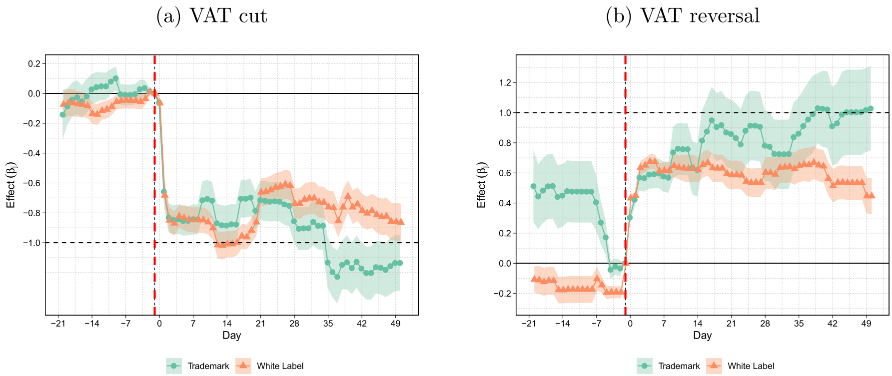


Fig. 12 Event study: Trademark vs. white label. This figure shows the different estimates of the degree of VAT pass-through when we break down different treated products into trademark vs. white-label food products. In orange the degree of pass-through for white-label food products. In green for the trademark. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

5.4.3 Domestic vs imported

It may be the case that domestically produced and imported food items demonstrate divergent behaviors in response to the underlying VAT policy. DPD PRISMA provides the GTIN-13 for each product[19] code, which indicates the country of production on its first two digits. Therefore, we are able to divide our data set between domestic (66.3% of the sampled products) and imported (33.7%). Table 4 provides the distribution of treated items in Coicop 5 for domestic and imported goods. Looking at Table 4, domestic goods are, on average, 11.5% cheaper and less volatile, since they have 3% less chances of experiencing a price reduction. Products under the VAT reduction scheme become 2.4% cheaper in the following two weeks, which would entail a price pass-through of about 60% and are 56% more likely to cut their price once the policy is enforced. We do not find evidence of a different behavior of domestic and imported products within this empirical strategy.

When we look at the dynamic impact of the policy and its evolution over time (see Fig. 13), we spot a similar degree of pass-through after one week and a remarkable divergence over time. One month after the tax change, imported products cross the "full pass-through milestone" and fluctuate below the 100% pass-through for several weeks. Domestic goods reduce their price difference with respect to the control group. As regards the reversal, as in the previous sample splits, we observe a lower degree of pass-through, and again, after several weeks there are different price-setting dynamics.

[19] GTIN stands for Global Trade Item Number. This is a 13-digit code, the first two or three digits are the country code. Note that this is a proxy, as sometimes multinationals register in different locations from the production location for some reason.

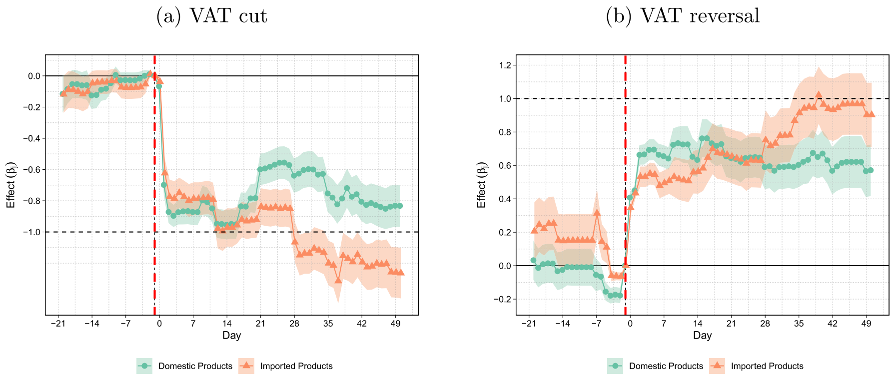


Fig. 13 Event study: Domestic vs. Imported. This figure shows the different estimates of the degree of VAT pass-through when we break down different treated products into domestic and imported foods. In orange the degree of pass-through for imported food products. In green for domestic goods. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

5.4.4 Time-varying characteristics: price volatility

We can also compare different subsets of products according to some time varying characteristics. In this subsection, we explore whether the most volatile products' pass-through dynamics are different from those that register a lower volatility ones. We define volatile products as those products for which the prices, prior to the VAT cut rate, register higher standard deviations. If the standard deviation of a product's pre-implementation prices is above the median of the standard deviation pre-policy prices' distribution, a product is defined as volatile. Note that both the most volatile and least volatile subsamples are equally sized.

Looking at Table 5, we first see that those more volatile products are on average 5% cheaper than those not volatile. More volatile products are more prone to price changes, especially increases. However, it is also important to notice how products subject to the policy with a higher historical price volatility do have more chances to see a price decrease after the policy took place.

In Fig. 14, we plot the estimates of the VAT pass-through for the two groups of products according to their price volatility. As expected, volatile products suffered a sharper and faster degree of pass-through during the first three weeks. Then, during the fourth week, such products' pass-through was heading toward the 60% pass-through area, a lower level than the not-volatile products. During the reversal episode, those not-volatile products' pass-through reached the 50% level and started to decrease slowly over time toward the 30% level, indicating a very asymmetric pass-through between the cut and the reversal episode. Those volatile products' pass-through experienced a symmetric reversal, as can be seen in the orange line that is around the 100% level during the first 50 days.

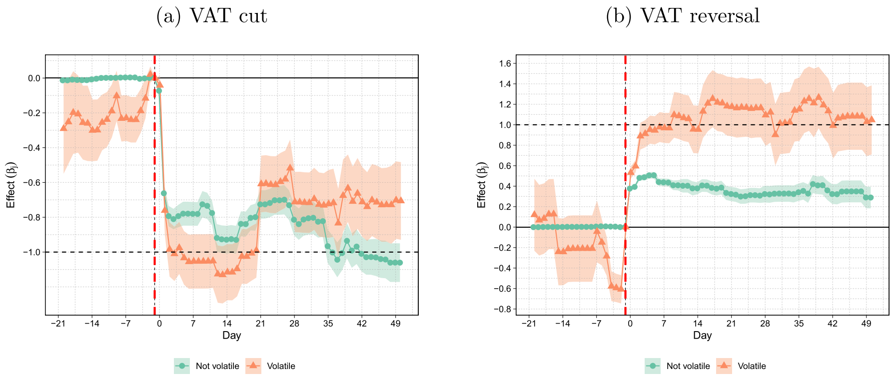


Fig. 14 Event study: Highly vs. low volatile products. This figure shows the different estimates of the degree of VAT pass-through when we break down different treated products into high vs. low volatile products. In orange the degree of pass-through for highly volatile food products. In green for low price volatility products. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

5.4.5 Big- vs. low-ticket items

We explore the degree of pass-through taking into account the price level category within each Coicop5 sub-class. To be more precise, in each category, we identify those products that are more expensive or that require a higher disbursement, what we label as the "big ticket" items, those that are medium ticket, and a third group that are in the cheap segment. So for each each Coicop 5 category, we split the subsample into three groups: the very-low-, the medium-, and the high-price ticket. So we have equally sized groups within each category. To define these groups we use the prices observed during the previous 30 days the VAT change. There are two issues with this approach; first is that we do not have the price per unit, having this information will help us to compare the VAT pass-through taking under consideration quality issues. We abstract from this, and we explore the differentiated pass-through among these three categories, as we consider that the consumer may switch toward the options that require different degrees of disbursement. As an example, they might decide to buy 200 grams of jamon iberico instead of the complete piece and not the effect of switching from a premium option to a lower quality. A second issue, that we could expect some switching of the products that lie in one of the bins after the VAT cut.

Looking at Fig. 15, we see that the cheapest items show a faster pass-through that is persistent over time. The products belonging to the low-ticket group experienced a pass-through of around 100% during the first two weeks after the first VAT cut was implemented. More expensive products experienced lower pass-through levels: around 85 and 50% after 20 days for the medium- and high-price groups. Overall, products in the upper range show a moderate pass-through and a faster reversal to the previous level. Looking at the reversal event, we could extract different conclusions: although the high-ticket group persistently adjusted their prices in response to the VAT reversal, lower-priced items did it at the same speed as those expensive ones. However, the

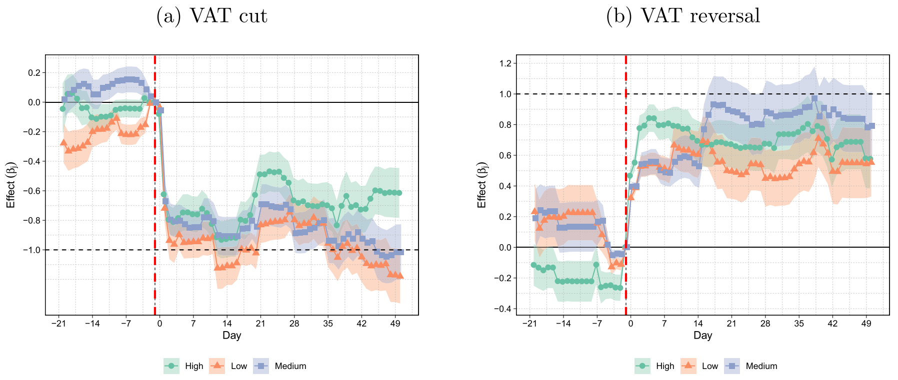


Fig. 15 Event study: Big-ticket vs. low-ticket products. This figure shows the different estimates of the degree of VAT pass-through when we break down different treated products into high-, medium- and low-ticket foods. In orange the degree of pass-through for low-ticket food products. In green for high-ticket products. In blue for medium-ticket products. Each coefficient bandwidth represents the 95% confidence interval. Standard errors are clustered at the product-store level. The vertical dashed line represents the reference period. *Sources:* Authors' calculations based on DPD Prisma ECB

pre-trends assumption is unlikely to hold within this specific event study, since except all products groups barely touch the 0 y-axis level in the pre-event period.

5.5 How measurement affects the pass-through estimates

The vast majority of the empirical literature on pricing has used either monthly or weekly data. As documented by Cavallo (2018), this can lead to measurement bias. Specifically, employing this frequency might influence the computation of various price metrics, like the occurrence and magnitude of price fluctuations, due to time-averaging bias. More specifically, statistical offices typically select one price per month for each store-location. If multiple price observations exist, they calculate an average. A similar approach is applied to the scanner data, where the prices are not observed but are derived as the price per unit, based on the information on spending and quantities. Our daily web-scraped data circumvent these issues.

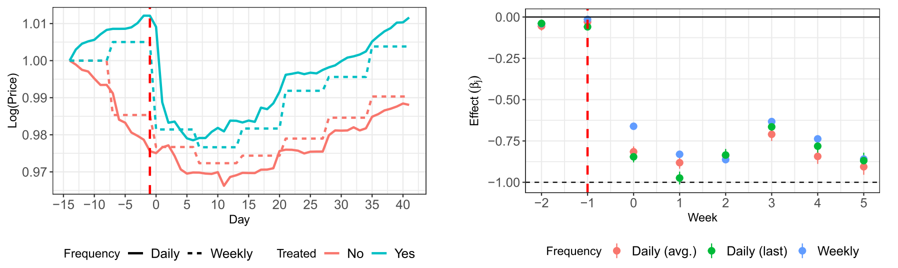
To illustrate the source of bias when using time-averaged price statistics, suppose that we are dealing with weekly data and observe three weeks. Following the example of Campbell and Eden (2014), let us say that a given product changes price in the middle of the second week. Using the weekly average price would lead to a calculation of two price changes of a small magnitude in week two and week three. Contrary to the result of using daily (or the end of period weekly) data, this will lead to just one change of bigger size. It can also be the case that there are multiple price changes during the second week, and this would produce a lower frequency of price changes and an increase in the absolute size of price changes. Therefore, in our present study, our estimates might be more precise or less biased than the estimates that rely on official, monthly data, as they do not suffer from aggregation bias.

To show the implication of data aggregation bias in this subsection, we test the implications of using our daily and granular data compared to a weekly frequency by means of averaging or the underlying micro-monthly data collected by the INE to estimate the degree of VAT pass-through. To do so, we first estimate the main results of section 5.2 based on averaged weekly observations and then perform the same exercise using the monthly micro-CPI data collected by the Spanish National Statistical Office. These micro-CPI official data has been used by Gautier et al. (2024) to document such facts in terms of frequency and size of price changes compared with other countries in the Euro Area.

5.5.1 Daily vs weekly prices

Figure 16a plots the mean (log) prices on daily and weekly frequencies. In the weekly case, prices are the weekly mean prices. Figure 16b shows the estimates of the event study in weekly vs. daily prices. Both models use the same specification of Eq. 1. To compare both models, we proceed as follows. Red dots are weekly averages of Fig. 6a point estimates, so that week 0 corresponds to days 0 to 6, week 1 corresponds to days 7 to 13, etc. Green dots are the last daily point estimate of each week, so that for week 0, the green point estimate corresponds to the estimate of day 6, for week 1 to day 13, etc. Then, each blue point in Fig. 16b shows estimates that were calculated taking average weekly prices (dashed lines of Fig. 16a).

Several conclusions can be drawn from Fig. 16. Estimates based on weekly data (blue dots) are substantially smaller than those obtained with daily data (green dots) one week and two weeks after VAT implementation.



(a) Mean (Log) Prices (b) Event Study

Fig. 16 Daily vs Weekly Frequency: 2023 VAT Cut. This figure shows the difference between using daily vs. weekly frequencies. In panel 16a we observe the evolution of the mean (log) prices during the first VAT cut by frequency and treatment group. The vertical red and dashed lines represent the last day (December 31, 2022) before the VAT cut entered into force. Then, in panel 16b, estimates of the degree of VAT pass-through are shown. The vertical red and dashed lines represent the last week before the VAT cut that took place on January 1, 2023, which is the time period we take as a reference. For the weekly specification the price is the weekly average and time fixed effects are adjusted to week FE. The horizontal dashed line accounts for the full pass-through. Daily estimates are those of Fig. 6. The green dots are each week last's daily estimation (day 6, day 13, etc.). Red dots are averaged beta estimates over each week (week 0 is days 0 (January 1, 2023) to 6 (January 7, 2023), week 1 is days 7 to 13, etc.). *Sources:* Authors' calculations based on DPD ECB Prisma

After one week (in week 0), the estimated pass-through of the VAT with daily data was around 87.5% (red dot overlaps with the green one). When using the weekly averaged data, the estimate of the pass-through is merely 62.5%. If we were to analyze what happened on average during each week, adopting a weekly frequency would provide a different, biased estimate of the pass-through effect resulting from the time-averaging problem. For example, during the second week after the VAT cut, affected products experienced an average pass-through of 87.5% (red dot for week 1 in Fig. 16b). Weekly prices, instead, delivers a pass-through of 83% (blue dot for week 1), an almost 5 basis points difference in the estimation. These results can be confirmed by looking to the left hand sided graph in Fig. 16a. Therefore, we argue that the time-averaging problem is biasing results when using weekly frequencies to estimate the pass-through. For instance, the price decrease during the first week of treated products is so abrupt, that using the average to compute weekly mean (in logs) prices entails into a loss of information. What can be seen in the solid line is that prices decrease sharply during the first days after the VAT cut, which is something to take into account when measuring the pass-through level. As we argued before, when dealing with aggregated data over time, not taking into account such price changes within weeks can impact the estimation. We believe that the use of daily frequency is advantageous in this matter.

The use of each last week's daily point estimate (green dots) is not comparable to the red and blue dots, since the green ones are daily point estimates. if we were to look at the pass-through level of treated products at the end of each week, we'd need to look at these green dots. At the end of the first week, the PT was 87.5%. The second week after the VAT cut, the pass-through achieved the full pass-through milestone. Then, it sharply decreased to the 85% PT level at the end of the third week. This sharp changes in PT levels are, by construction, less biased since we're not averaging prices.

5.5.2 Scraped vs. monthly collected prices

Another competing source of data to estimate the VAT pass-through is the microdata collected by the National Statistical Office (INE). This microdata serves as the foundation for compiling the official CPI statistics. Data collection is carried out by regular visits by pollsters to retail stores, but also through emails and phone calls. In recent years, the INE is introducing alternative sources such as scanner data provided by large retailers.[20] The INE data contains geographical and store identification; this is in which *provinces* the product is sold and in which store or supermarket it was sold.[21]

Figure 17 shows the results obtained with the same event study strategy to estimate the pass-through. Taking into account the fact that these data contain geographical and the store information, here we add to the specification in Eq. 1 fixed state and the store/retailer effects.

The measured pass-through with these monthly data and specification is estimated at 67% during the first month. We saw in Fig. 6, daily estimates move between the 65 and 100% pass-through levels. Then, we can conclude that the time-averaging

[20] For more details on the use of scanner data in constructing official CPI statistics see https://www.ine.es/metodologia/t25/ipc_scanner_data.pdf.

[21] For more details on the Spanish INE micro-CPI data, see https://assets.aeaweb.org/asset-server/files/21477.pdf.

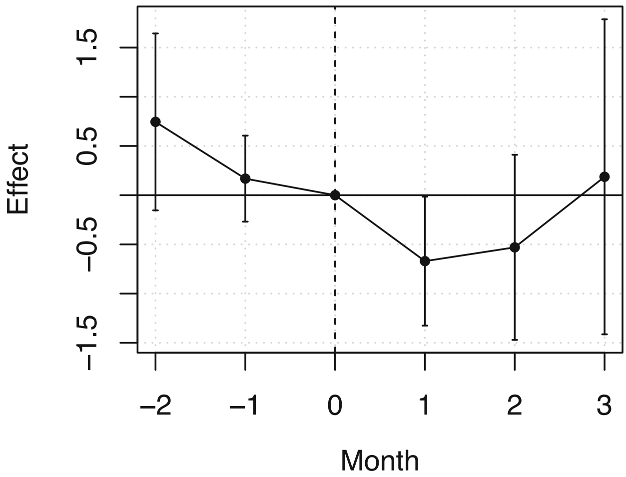


Fig. 17 Monthly Official Data: 2023 VAT cut. This figure shows the estimates of the degree of VAT pass-through with micro-monthly data collected by INE. December 2022 is the month of reference. Therefore, month 1 is January 2023, month 2 is February 2023, etc. Estimates for months 1, 2 and 3 are: -0.67, -0.53, and 0.18. We cluster standard errors by product, as in the daily and weekly estimations. *Source:* CPI microprices collected by the Spanish National Statistics Institute (INE by its Spanish acronym)

issue is present here: monthly estimates delivered the minimum of daily pass-through estimates during the first month. Web-scraped data contains daily information, so the measurement of the pass-through is more accurate (or less biased): the expected (estimated) value of pass-through is closer to the true value of the VAT pass-through.

Not only is the estimation less biased. We can argue that the estimation is more precise. We see that the standard errors of the monthly estimate are wider, indicating a less precise estimate. In fact, the bands hover around the no-pass-through after the first month of the implementation, indicating that there is a small chance that the pass-through estimate is non-reliable due to its high variance. In fact, this estimation's high variance speaks to the fact that the time-averaging bias is very present within this data source. Following the previous example rationale, one could argue that there were many price changes (those affected by the VAT cut) that are not captured due to the time-averaging. As products influenced by the cut experienced a price change in the initial week, calculating the average price throughout the month or sampling would result in less pronounced price changes. As such, average price changes would be underestimated: products that started to suffer upward price changes between the second and last week of the first month may be increasing the variance of the estimation. Overall, the estimation with monthly official data presents a slight downward bias and a notable higher variance, induced by the temporal aggregation bias.

6 Conclusions

On December 28, 2022, the Spanish government announced a temporary VAT discount on selected products to alleviate the economic impact of high inflation. They also announced an additional VAT cut to olive oil products (July 1, 2024) and the reversal of the measure in two phases, the first reversal took place in October 2024 and a second phase in January 1, 2025. We study the effect of the VAT rate changes on retail prices in a Spanish supermarket using web-scraped data collected by DPD Prisma ECB covering an average of 10,000 food product prices per day (21,000 products in total). These microdata are especially valuable for examining how VAT rate changes are reflected in final prices, as they allow us to monitor the pricing of specific products.

After a thorough description of the dataset as well as the techniques used to convert these massive unstructured data into ordered and properly classified data, we proceed as follows. In a first step, in order to check its properties and suitability for analysis, we check how representative this sample is compared to the official CPI data. DPD Prisma ECB only collects data from large retailers, while official statistics cover a variety of retailers and provide a representative geographical coverage. Therefore, it is important to assess the representativeness of these data. To check this at the CPI sub-class level, we need to correctly map each product to the official classification, that is, into Coicop5, which classifies each good according to its purpose.[22] We classify each product using machine-learning techniques. We also analyze the price-setting behavior in terms of the frequency of price changes and the distribution of size changes for the observed products. Our results match the established stylized facts found in previous research, which reassures us of our data's credibility. Once we confirm that the indices and components of the DPD Prisma ECB sample mimics quite well the official indices, we proceed to evaluate a particular policy measure introduced in Spain, that is, the above-mentioned temporary VAT reduction.

To measure the degree of pass-through of the VAT changes to final prices, we use an event-study design and compare the evolution of prices of those products targeted by the measure against those not affected. We estimate the pass-through in four different episodes: the January 2023 VAT cut, the July 2024 cut in Olive Oil products, as well as the VAT reversal that took place in two phases: the first in October 2024 and the second in January 2025. In all episodes, We find that retailers have passed on almost all of the VAT cut for affected products with certain degree of heterogeneity in their intensity. That is, most of the treated products registered a price decrease (increase), and the size of this decrease corresponded to the established VAT reduction (increase). As regards a possible asymmetry of the VAT pass-through when there is a cut or an increase, we found that the degree of pass-through is higher during cuts than during VAT increases.

When looking at price dynamics over time and exploiting observed product characteristics, we find some heterogeneities between processed vs. non-processed goods, trademark goods vs. white-labeled, domestic vs. imported goods, highly volatile price

[22] We target the Classification of individual consumption by purpose (Coicop) classification https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Classification_of_individual_consumption_by_purpose_(COICOP) by (1) division, (2) group, (3) class, (4) sub-class (to give an example, this would be"01.1.1—Rice") and (5) product.

products, and high- vs. low-ticket varieties. These differences might reflect differences in price setting, the exploitation of the different consumer price elasticities and the price-setting negotiation between producers and retailers.

We also examine the influence of policy unpredictability on pricing strategies, which arose from various waivers that postponed the reversal of this provisional measure. Originally planned as a six-month measure, the VAT reduction was extended up to three times owing to ongoing inflation throughout 2023, ultimately lasting between 24 months until its complete reversal. Finally, we outline the benefit of utilizing daily prices instead of averaged monthly data in terms of measurement bias. By contrasting our daily dataset with the microprices from the National Statistics Office, we find that our pass-through measurement displays reduced bias and enhanced precision.

Appendix

Data on prices online

Data Cleaning and treatment

To make the microdata suitable for our purposes, we implemented a number of methodological interventions to deal with missing observations, sales and promotions, and other features of daily web-scraped data. Include, the number of prices available on a daily basis from each retailer are shown in Fig. 18.

Data cleaning and price filling. We drop observations that are clearly errors (such as exorbitant prices) and exclude price changes smaller than 0.1%, as well as increases above 100.0%, to account for possible measurement error. We get rid of those products for which the number of observations is below a certain threshold of day, calculated as the number of observations out of the maximum number of possible daily observations (for DPD Prisma this is from the 1st of April 2022 up to October 2024). The sequence of price records can be interrupted because (i) it reaches the end of the observation period, (ii) an item is out-of-stock (temporally or permanently), or (iii) a technical issue (e.g., robot(s) and/or website(s) go offline). These instances can affect the length of price spells, most likely shortening them. If we observe missing observations for a certain amount of days, (21 days), these are filled forward with the most recent usable price and replace the unusable or missing observation to fill in the gaps. We assume that these short gaps are related with a problem with the scraper that collects the data that led to web-scraping routine failures, rather than stock outs, as in the case of miss-reporting the failure would be broad-based. This new price series are labeled as **filled** series.

Sales filter. Since the price information includes temporary sales and promotions, and that can alter the results of some of the statistics. We filter the series to account for temporary sales to eliminate the high frequency of price changes. A temporary sale is defined as prices that remain below the usual price and return to its previous level, we allow for a time window below 21 days (we use the filter proposed by Nakamura and Steinsson, (2008)) and we also filter temporary increases, also for 21 days (here

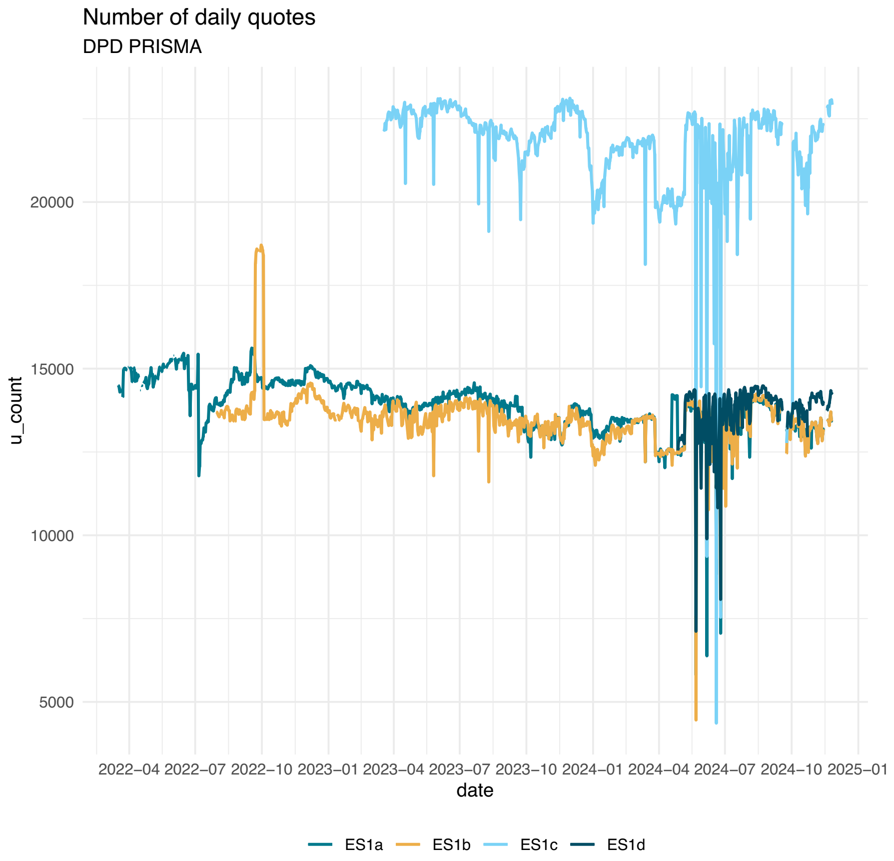


Fig. 18 Daily number of products (raw data). *Source*: DPD Prisma. Number of observed prices on a daily basis by retailer and shop. Not filtered nor classified

we use the filter proposed by Kehoe and Midrigan (2015)). In some specific cases, retailers show a particular pricing behavior, so we relax the condition where the criteria allow for small changes within a bracket instead of the condition that the price has to return to the previous level. This observed pattern of not returning to the previous price seems to be related with certain characteristics of a particular establishment that mainly operates in large cities.

We have an unbalanced panel as some product items are not always available or have missing price information.

On the entry and exit of products. For a limited number of products we observe product churning. So we impose some continuity by restricting the sample to those products whose price is observed for a long period of time. We remove products that are observed for less than 100 days.

In general, we have used DPD Prisma on a daily frequency. But in some cases, we explore the results using the data on a weekly frequency. In this circumstance, we can take the last observation, the mean, or the mode of a given week. As we want to explore the implications of the bias due to mean averaging, we use the mean.

Product Grouping. To classify the products, we have used the following strategies:

- *Processed and unprocessed foods:* We classify the products according to Coicop4 class.
- *Trademark and white label:* We extract the label of each product from the description by means of machine-learning techniques. When the label coincides with the name of the supermarket or one of the identified white labels of the supermarket is attributed to be a "white label", if not is "trademark".
- *Domestic and imported products:* We base this classification on the first two digits of the GTIN-EAN code.

Index computation. We calculate the daily price indices following closely the methodology of the National Statistics, while adjusting it to the characteristics of the web-scraped daily data, as outlined in Cavallo (2013). By using web-scraped data, the number of goods within each category changes dynamically, contrary to the official methodology, where there is a fixed basket of goods and there are replacements.

To build the index, price changes are calculated at the product level, then averaged using unweighted geometric means (to avoid that the mean is driven by some products that can register large swings in the price change), and then aggregated across categories with a weighted arithmetic mean.

Geometric average of price changes in category j at each day t:

$$R^j_{t,t-1} = \prod_i \left(\frac{p_t^{i\,1/n_{jt}}}{p^i_{t-1}} \right) \tag{3}$$

The Coicop 5 category-level index at time t is:

$$I^j_t = R^j_{1,0} \cdot R^j_{2,1} \ldots R^j_{t,t-1} \tag{4}$$

Once we have a mean growth for each category, we weigh them using the weights provided by the National Statistical Office.[23]

[23] We use 2023 CPI weights by 5-digit issued by the Spanish statistical office (check this here).

$$\text{Retailer Index}_t = \sum_j \frac{w_j^i}{W} I_t \tag{5}$$

where w_j is the official CPI weight for that category and W is the sum of all the weights included in the sample (Fig. 18).

Aggregation. To create a country-level measure of each of the statistics, we use information at the product/store level, aggregated at the sub-class level, that is, with the weights at the Coicop5 digit level provided by the National Statistical Office. By doing this, we control for sampling issues as the number of products collected within each sub-class category. We compute each statistic at the product level. Then, we aggregate the results to produce a store statistic, followed by a retailer statistic, and finally a country statistic. Within each store, we weigh each product according to the HICP. For example, when calculating the frequencies of price changes by country we proceed as follows. To compute the aggregate frequencies of price changes. First, we compute the frequency of price changes of product i, in sub-class j sold in store s ($F_{i,j,store}$), we compute the mean within each sub-class *j, store*, and then we weight each sub-class by the HICP weight (w_j):

$$F_{store} = \sum_{j=1}^{N_{jstore}} w_j F_{j,store} \tag{6}$$

Results with the alternative dataset

We make use of the alternative dataset provided by Datamarket and apply the same estimation strategy. The idea is to confront the results with another Spanish supermarket that is known to have a completely different price-setting strategy. The results are shown in Fig. 19, and are quite similar to those obtained before. In this case, we can also state that the VAT pass-through to final prices was almost complete after one week.

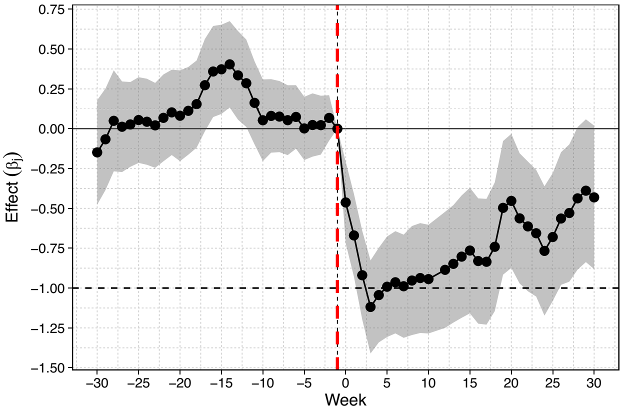


Fig. 19 Event study: Retailer 2. Spanish Supermarket. Datamarket. This figure shows the estimates of the degree of VAT pass-through. Each coefficient bandwidth represents the 95% confidence interval. The vertical red and dashed line (-.-.-) represents the last week before the VAT cut (Week -1), which is the time period that we take as a reference. *Sources:* Authors' calculations based on Datamarket

Additional tables

See Table 4.

Table 4 Product distribution in DPD data (whole period)

		Domestic				Imported				
		Trademark		White label		Trademark		White label		l
		NP	P	NP	P	NP	P	NP	P	All
Control	N	1158	1412	3586	5778	670	671	1034	3110	17419
	% row	6.6	8.1	20.6	33.2	3.8	3.9	5.9	17.9	100.0
T1	N	180	378	321	1121	165	188	260	643	3256
	% row	5.5	11.6	9.9	34.4	5.1	5.8	8.0	19.7	100.0
T2 and T3	N	0	100	0	463	0	58	0	223	844
	% row	0.0	11.8	0.0	54.9	0.0	6.9	0.0	26.4	100.0
All	N	1338	1890	3907	7362	835	917	1294	3976	21519
	% row	6.2	8.8	18.2	34.2	3.9	4.3	6.0	18.5	100.0

This table classifies each of the products of the Spanish retailer under subgroup 011, that is, Foods, into several categories to provide a sense of the number of observations under each of the groups: Treated vs. non-treated, processed (P) vs. non-processed (NP) and trademark vs. white-labeled products. *Sources:* Authors' calculations based on DPD Prisma ECB

Table 5 Heterogeneity in price levels and in price setting

	ln(Price)	$\Delta Price$	$\Delta Price > 0$	$\Delta Price < 0$
	(1)	(2)	(3)	(4)
Processed[1]	-0.774***	0.024	0.062**	-0.037
	(0.116)	(0.026)	(0.024)	(0.026)
Trademark[2]	0.276***	0.035	0.024	0.012
	(0.055)	(0.022)	(0.017)	(0.013)
Domestic[3]	-0.122**	-0.018	0.010	-0.027***
	(0.052)	(0.018)	(0.014)	(0.009)
Volatile[4]	-0.054*	0.143***	0.102***	0.041***
	(0.032)	(0.021)	(0.019)	(0.007)
Medium Ticket[5]	0.573***	-0.004	-0.003	-0.001
	(0.019)	(0.002)	(0.002)	(0.001)
High Ticket	1.192***	-0.004	-0.003	-0.001
	(0.057)	(0.003)	(0.002)	(0.001)

This table reports the results of the estimation of Eq. (2) for four different dependent variables. Column (1) evaluates the differences in price levels between the heterogeneous groups. Columns (2) to (4) eval- uates the probability of price changes, either increases or decreases.*Source:* Authors' calculations based on DPD Prisma ECB

[1] Prices of processed products are on average cheaper than non processed. Prices increase with a higher probability and, after the introduction of the policy measure, the probability of a price decrease was higher, but smaller if the product was processed and under treatment

[2] Prices of trademarked products are on average more expensive than white-label. We do not observe any differential behavior in terms of the probability of price changes, except for the specific episode after the treatment

[3] Domestic products are on average cheaper than imported, there are not substantial in- crease with a higher probability

[4] Volatile products show a higher probability to change prices, both increases and decreases

[5] The ordering of the price level is reasonable. And there are not substantial differences in terms of the probability of a size change

Product classification

Given the broad coverage of products sold by each retailer, we also need to identify and assign a COICOP code to each product, and we obtain the brands using natural language processing techniques.

Data sources

In addition to prices, both Datamarket and DPD Prisma datasets contain text information. These two data sources contain the name and description of each product being sold (e.g., "*sliced bread 500 gr.*"), and the section or location of the supermarket where it is sold (following the example above: "*breads, bakery*"). We concatenate these two dimensions into a single piece of text. This resulting text for each product is hereinafter referred to as the product description. The product description also includes other information, such as the weight of the product (for example, 1 kg, 1l, 250g) or the brand name. We do not perform any kind of pre-processing steps for the product descriptions since, as we will mention later, we will rely on novel tokenizers that are able to detect all kind of characters within the text.[24] One advantage of using web-scraped product description data is that this piece of text associated to each product is a short sentence (for example, "*3 x frozen margherita Pizza extra cheese, Frozen food / Prepared dishes / Pizzas*"). Usually, when performing a text analysis in economics, it is preferable to use short sentences in the modeling phase instead of large chunks of text (Hansen and Ash (2023)). However, the short text may contain some noise that can confuse the trained classifier. Following the previous example, the word "cheese" may suggest that this product description should be assigned to the COICOP category "Cheese". This is one of the reasons why we choose to perform an algorithmic supervised classification task over a dictionary-based strategy. Using labels in order to minimize some loss function in the training phase reduces this kind of noise-related lower model accuracy. Besides, not only would a dictionary-based perform worse; also, it would be very time-consuming to keep on updating the dictionary terms for every product entry and exit from the market.

Data labeling

For the data-labeling phase, a pre-trained Sentence Transformer (Reimers and Gurevych) is used to encode product descriptions. A Sentence Transformer is a natural language processing (NLP) model designed to convert sentences or phrases into numerical vectors in a high-dimensional space. These vectors capture the semantic meaning of sentences, enabling tasks such as similarity search, which is our case. By converting sentences into numerical vectors, similarity can be calculated between them using distance measures. This kind of models do not use tokenized text; i.e., sentences (product descriptions) are encoded within the Transformer with no prior pre-processing or text cleaning. String punctuations, stop-words, and other types of language nuances are not removed from the text. The reason is that these Transformer models have their own tokenizer model as a pre-processing step and it is not required to preprocess the data before encoding sentences. This allows our information retrieval task to be more precise when executing the search. Then, once the produce descriptions are encoded (i.e., text is converted into a vector), a semantic similarity search

[24] The exception is for the ensemble machine-learning algorithm model we propose to evidence that our preferred method outperforms this and other method, in which we run a classical text processing pipeline.

using cosine similarity is conducted. This sequence of steps can be summarized as the following:

1. Suppose that we encoded all product descriptions within the Transformer. We now have a matrix of embedding *NxL*, where *N* is the total number of product descriptions and *L* is the maximum sequence length allowed by the transformer; that is, the maximum number of tokens[25] that the Transformer can transform into a vector. We set the maximum length to 128 tokens. This embedding matrix can be seen as a "similarity" matrix.

2. Suppose we want to find products that are semantically similar to "*3 × frozen margherita Pizza extra cheese, Frozen food / Prepared dishes / Pizzas*". We look for the encoded product description and measure the cosine similarity between the product description and all remaining product descriptions. Cosine similarity is calculated as:

$$f(x, y) = \frac{xy^T}{\|x\|\|y\|} \tag{7}$$

where x and y are row vectors (x may be our product description). Euclidean (L2) normalization projects the vectors onto the unit sphere, and their dot product is then the cosine of the angle between the points denoted by the vectors.

3. Once we have the $1xN$ cosine similarities vector, we can sort that vector and yield the top K similar products to our "*3 × frozen margherita Pizza extra cheese, Frozen food / Prepared dishes / Pizzas*" product description.

We assign a Coicop 5 to each product description at once to the retrieved list after having manually reviewed each item. This allows us to rapidly label multiple similar product descriptions. A more detailed snapshot of how this task is performed can be seen in Fig. 20. This process is repeated until a sufficient number of manually tagged samples per Coicop 5 category are obtained.

The data set used for labeling purposes consisted of approximately 54000 Spanish product descriptions, resulting in a labeled sample of around 10% of product descriptions on average for both countries, as can be seen in Table 6.

Methods

Model architecture

As Hansen et al. (2023) mention, the Natural Language Processing field has been substantially transformed in the last few years. Vaswani et al. (2023) proposed a new framework (called *self-attention*) that enables a neural network to weigh the importance of different elements in an input sequence and dynamically adjust their influence on the output. This is especially important for language processing tasks, where the meaning of a word can change depending on its context within a sentence or document. In our case, the interaction of certain key words determines COICOP categories. For instance, "*3 x frozen margherita Pizza extra cheese, Frozen food /*

[25] Tokens are groups of characters, which sometimes align with words, but not always. For instance, our "*Pizza*" product description contains precisely 20 tokens.

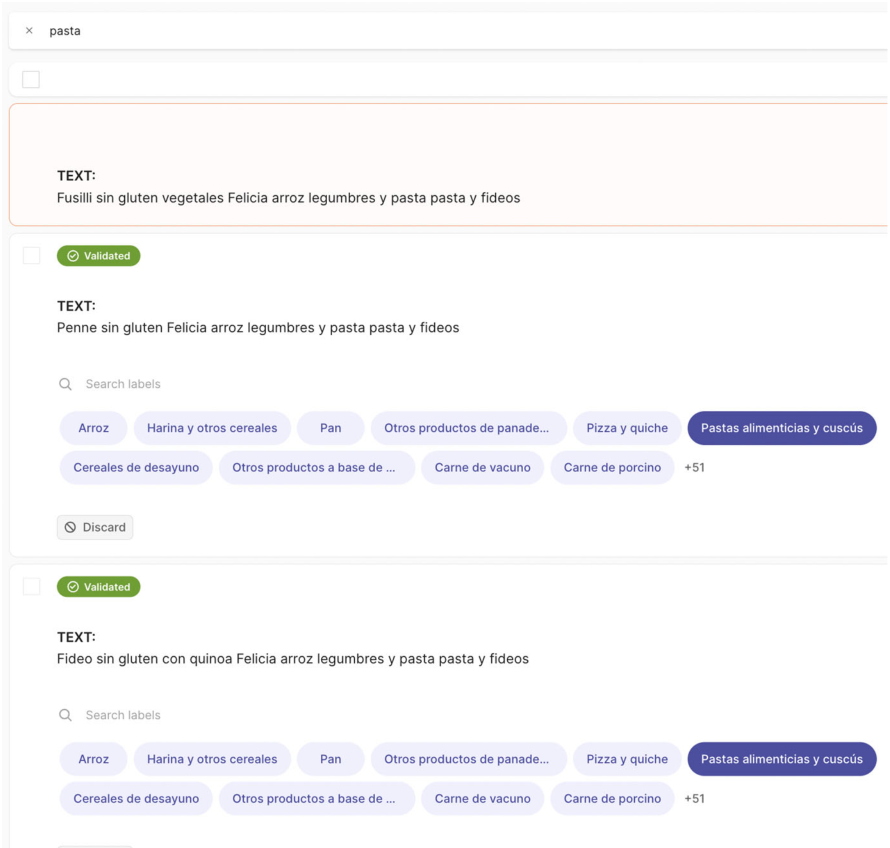


Fig. 20 Manual Bulk Labeling of similar products to 'Pasta '. We used the Argilla user interface and API https://github.com/argilla-io/argilla

Table 6 Training dataset descriptive stats

	ES
Number of products	54,449
Estimated non-food	35.17%
Labeled	9.83%
Avg. labeled prods. per COICOP	87.8

Prepared dishes / Pizzas" should be assigned a COICOP of "Pizzas and quiches". Some words, such as "*cheese*", may be interacting with its surrounding words (such as "*pizza*" or "*extra*") to indicate that it is indeed a "*pizza*". This kind of interaction is what new natural language processing models are capable of capturing. In fact, in our paper, we use a model governed by the *self-attention* mechanism.[26]

Following Hansen et al. (2023), we use a DistilBERT (Sanh et al. (2020)) model to first domain-adapt it to our product description dataset and later to train it with the labeled subsample. A DistilBERT model is a smaller, distilled version of BERT (Devlin et al. (2019)). It is created through a process called knowledge distillation, where a larger, more complex model (in this case, BERT) is used to train a smaller model with similar capabilities. The objective is to transfer the knowledge from the larger model to the smaller one while reducing its size and computational requirements. DistilBERT aims to retain as much of BERT's performance as possible while being more efficient in terms of memory and inference speed. Moreover, we choose a multilingual DistilBERT model; that is, as it name says, it can handle multiple languages. These models learn to encode and understand the context and meaning of words in multiple languages by capturing the relationships and patterns within the text during the training phase. By doing so, they acquire a cross-lingual understanding that allows them to transfer knowledge from one language to another. This means that the model can generalize its understanding of languages across different languages, even if it has not been explicitly trained on a specific language.

Domain adaption

We domain-adapt a pre-trained model with our product description data. Domain adaptation of a pre-trained model involves adjusting a model that has been trained on one type of data to work well on a different type of data. In machine learning, each type of data is called a domain, and the pre-trained model is trained on a specific domain called the source domain. However, when we want the model to work on a different domain called the target domain (product description data), we need to adapt it. Adapting the pre-trained model is necessary because the target domain may have some differences compared to the source domain. These differences could be due to variations in how the data was collected, differences in the data represents, or changes in the characteristics of the data itself. In essence, the main idea behind domain adaptation is to make the pre-trained model able to handle the differences between the source and target domains. This is done by adjusting the model so that it can transfer its learned knowledge effectively to the target domain. More specifically, this is done by removing randomly selected words from the product description data. Once words are deleted, the model updates its parameters by predicting the deleted words. We perform this task for the whole sample of product descriptionb data shown in Table 6. For this purpose, we use a DistilBERT multilingual model (Sanh et al. (2020)) pre-trained[27] on multiple languages and domain-adapt it to our product description data. We use the cased version since we do not perform ay kind of pre-processing within our product description data. Note that this could be easily extended as well to other languages, such as French, German, etc.

[26] This mechanism is also behind some famous models such as ChatGPT.

[27] We use this model.

Table 7 Hyperparameter space for cross-validation of DistilBERT

Hyperparameter	Values
Learning rate	[5e-5, 3e-5]
Epochs	[10, 20]
Batch size	[8, 16, 32]

Training

There is scarce literature on what methodology should one follow to classify food products into official statistics's categories. To our knowledge, there is only one paper relating this kind of problem. Lehmann et al. (2020) proposed using a transfer learning with a convolutional neural network trained on German human labeled data to infer and predict for French product description data. Once we have the multilingual DistilBERT model adapted to our data, we perform cross-validation over a hyperparameter space using our Spanish labeled subsample. The model we use is the same domain-adapted DistilBERT multilingual model from the previous step. We split our labeled subsample into a 70% train and 30% test sets to perform cross-validation over the training set. The hyperparameter space can be seen in Table 7.

For model evaluation, we monitor the average F1-score over the 3-folds. We select the model with highest average F1-score. Once we found the best hyperparameters,[28] we fit the model on the whole training subsample and infer for the test set sample and check the model performance over all E-COICOP labels.[29]

Model comparison

To prove that our methodology is sufficiently accurate, we compare our model with another set of models. More specifically, we choose train and validate within the same test set of the previous step, the following models:

Gradient Boosting Classifier: a gradient boosting classifier is a machine-learning algorithm used for supervised learning tasks, particularly for classification problems. It belongs to the family of ensemble methods, which combine multiple weaker models (often referred to as base learners) to create a stronger predictive model. The main difference between this model and and our chosen model is that this model isn't able to scale for other non Spanish product description data.

Moreover, some data pre-processing must be done in order to train and test for this model. In fact, we tokenize the text, deleting stopwords and removing string punctuation's, lemmatize each word and create the term-frequency matrix. This matrix constitutes the input of this model. Then, we perform again cross-validation over a hyperparameter space that can be seen in Table 8. We fit the train with the best set

[28] Learning rate of 3e-5, batch size of 8 and 20 epochs, with resulted on an average of 0.92% F1-score over the threefold subsets.

[29] An extra label or category is added: non-food products.

Table 8 Hyperparameter space for gradient boosting trees classifier

Hyperparameter	Values
Max. features	[sq. root, log, all features]
Number of trees	[50, 100, 300]
Max. depth of tree	[2, 4, 8, ∞]

Table 9 Test set metrics

	Precision	Recall	F1 Score
Gradient Boosting Trees	0.91	0.90	0.90
SetFit	0.93	0.93	0.93
DistilBERT	0.95	0.95	0.95

of hyperparameters[30] with the training set and evaluate for the test set. Note that, for model comparison, both training and test sets must be the same for all models.

SetFit (Tunstall et al.): this algorithm takes advantage of Sentence Transformers' (Reimers and Gurevych) ability to generate dense embeddings based on paired sentences. In the initial fine-tuning phase stage, it makes use of the limited labeled input data by contrastive training, where positive and negative pairs are created by in-class and out-class selection. The Sentence Transformer model then trains on these pairs (or triplets) and generates dense vectors per example. In the second step, the classification head trains on the encoded embeddings with their respective class labels. At inference time, the unseen example passes through the fine-tuned Sentence Transformer, generating an embedding that when fed to the classification head outputs a class label prediction. A nice advantage of this model it can be trained using a multilingual pre-trained model. In fact, we train this algorithm using the same domain-adapted, multilingual model for the training process with 15 iterations, 5e-5 as learning rate and a batch size of 16. We then infer for the test set to allow model comparison.

Results

Looking at Table 9, we observe that our model outperforms the other selected algorithms. We present several evaluation metrics for the 30% test set of our labeled dataset. First, the precision measures how many of the positive predictions made by the model are actually correct. It can be seen that our model predicts 95% of the cases correctly. Second, the recall quantifies how well the model captures all positive instances in the data set.

A high recall value indicates that the model effectively identifies a large proportion of positive instances correctly, minimizing the number of false negatives. A low recall value suggests that the model misses a significant number of positive instances, resulting in a high rate of false negatives. However, our model outperforms the others

[30] These were: All features, *log* and 300 trees.

Table 10 Main metrics for the best-performing model of the test set sample

	Precision	Recall	F1 score	N
Baby food	1.00	0.95	0.98	22
Beef and veal	1.00	1.00	1.00	14
Blonde beer	1.00	0.69	0.82	13
Bread	1.00	0.95	0.97	20
Breakfast cereals	1.00	0.92	0.96	26
Butter	1.00	1.00	1.00	11
Canned fruit and fruit products	0.90	0.82	0.86	22
Cheese	0.94	1.00	0.97	16
Chocolate	1.00	0.91	0.95	34
Cocoa and chocolate powder	1.00	1.00	1.00	25
Coffee	0.95	1.00	0.97	18
Confectionery products	0.82	0.90	0.86	20
Confectionery, jams and honey	1.00	0.96	0.98	26
Dried, salted or smoked meat	1.00	0.98	0.99	65
Edible offal	1.00	0.86	0.92	14
Eggs	1.00	0.92	0.96	12
Fish and shellfish, dried, smoked or salted	0.94	1.00	0.97	16
Flour and other cereals	0.75	0.75	0.75	8
Food pastes and couscous	0.96	0.96	0.96	25
Fresh or chilled fish	0.88	1.00	0.93	14
Fresh or chilled fruit	0.93	0.94	0.93	53
Fresh or chilled seafood	1.00	0.82	0.90	11
Frozen fish	0.88	1.00	0.93	14
Frozen seafood	0.89	0.94	0.92	18
Frozen vegetables other than potatoes and other tubers	1.00	0.82	0.90	17
Fruit and vegetable juices	1.00	1.00	1.00	23
Grape wine	1.00	1.00	1.00	31
Ice cream	0.88	0.96	0.92	24
Mineral or spring water	1.00	1.00	1.00	16
Non-alcoholic beer or low alcohol	0.90	1.00	0.95	9
Non-food products	0.98	0.97	0.97	148
Nuts and nuts	0.80	0.84	0.82	19
Olive oil	1.00	1.00	1.00	16

Table 10 (continued)

	Precision	Recall	F1 score	N
Other bakery products	0.93	0.98	0.96	58
Other beers with alcohol	0.81	1.00	0.90	13
Other cereal products	0.93	1.00	0.96	13
Other dairy products	0.92	0.92	0.92	51
Other edible oils	0.58	0.78	0.67	9
Other fish and shellfish preparations preserved or processed	1.00	0.91	0.95	53
Other foodstuffs	0.95	0.92	0.94	63
Other meat	0.86	1.00	0.92	6
Other meat preparations	0.97	0.93	0.95	30
Pigmeat	0.93	0.87	0.90	15
Pizza and quiche	1.00	0.95	0.97	20
Potato chips	1.00	0.93	0.97	15
Potatoes	0.94	1.00	0.97	17
Poultry meat	0.95	1.00	0.97	18
Prepared dishes	0.83	0.85	0.84	68
Refreshments	0.95	0.98	0.97	58
Rice	1.00	1.00	1.00	18
Salt, spices and culinary herbs	1.00	0.95	0.98	22
Sauces and condiments	1.00	0.95	0.97	20
Sheepmeat and goatmeat	0.86	0.86	0.86	7
	Precision	Recall	F1 score	N
Skimmed milk	1.00	1.00	1.00	29
Spirit drinks and liqueurs	1.00	1.00	1.00	19
Sugar	0.94	1.00	0.97	15
Tea	0.94	1.00	0.97	16
Vegetables, dried or otherwise preserved or processed	0.89	0.86	0.87	28
Vegetables, fresh or chilled, other than potatoes and other tubers	0.86	0.98	0.91	44
Whole milk	0.94	1.00	0.97	15
Yogurt	0.98	0.94	0.96	47
Average	0.94	0.94	0.94	1607
Weighted average	0.95	0.95	0.95	1607

The”Inflation surge” phase spans from April 2022, when DPD PRISMA data first became available, to the end of September 2023. The”Post inflation” phase commences in October 2023 and continues until November 2024. Within each retailer, the rate of price changes, excluding those from sales and promotions, is calculated, followed by averaging these rates across sub-classes within each COICOP5 category using a simple mean. Each category is then assigned a weight according to official HCPI specifications. This results in a frequency of price changes per retailer, which is then combined through a simple mean. For the EA-4 country group, a simple average is derived by equally weighting all four countries.

in terms of recall. Finally, the F1-score, measured as the harmonic mean between precision and recall, is also 95%. This means that the model has a low rate of false positives (high precision) while effectively capturing most of the positive instances (high recall). It is noteworthy that the SetFit algorithm, designed to perform well with few labels per category, underperforms our chosen method, even if this algorithm also supports a multilingual setup. To our knowledge, this is the first attempt within economics research in comparing an easy-to-implement few-shot classification method with other state-of-the-art natural language processing models. A detailed table with all precision, recall, and F1-score metrics can be seen in Table 10.

Acknowledgements We are deeply grateful to the anonymous referee for their thorough and insightful feedback, which significantly enhanced the quality and rigor of this manuscript. We also thank Erwan Gautier (Banque de France) and Aitor Lacuesta (Banco de Espan˜ a) for their valuable comments and guidance. We are especially grateful to the Daily Price Dataset (DPD) Team at the ECB for making available these data for research purposes and for guiding us. The authors also thank Datamarket for sharing the data set with us. The opinions and analyses are the responsibility of the authors and therefore do not necessarily coincide with those of the Banco de Espan˜ a or the Eurosystem. This research was conducted during a secondment at Banque de France (BdF) under the External Working Experience program. The hospitality of the Banque de France and especially of the Service de Études Microéconomiques is gratefully acknowledged.

References

Almunia M, Martínez J, Martínez Á (2023) La reducción del iva en los alimentos básicos: evaluación y recomendaciones. EsadeEcPol Brief. https://doi.org/10.56269/20230328

Amores AF, Barrios S, Speitmann R and Stoehlker D (2023a) Price Effects of Temporary VAT Rate Cuts: Evidence from Spanish Supermarkets. JRC Research Reports, JRC132542, Joint Research Centre (Seville site). https://ideas.repec.org/p/ipt/iptwpa/jrc132542.html

Ash E, Hansen S (2023) Text algorithms in economics. Ann Rev Econ 15(1):659–688

Basso HS, Dimakou O and Pidkuyko M (2023) How inflation varies across Spanish households. Occasional Papers, 2307, Banco de Espan˜a. https://doi.org/10.53479/29792

Benzarti Y, Garriga S and Tortarolo D (2022) Can vat cuts dampen the effects of food price inflation? Working paper. https://dtortarolo.github.io/WebPage/VAT_inflation.pdf

Bernardino T, Gabriel RD, Quelhas J and Silva-Pereira M (2024) The full, persistent, and symmetric pass-through of a temporary vat cut. *mimeo*. https://www.ricardoduquegabriel.com/files/BGQS_2024.pdf

Burstein, Ariel, and Gita Gopinath. (2014). *International Prices and Exchange Rates*, vol. 4. Elsevier, pp. 391–451. Preliminary January 2013. Prepared for the Handbook of International Economics, Vol. IV. https://doi.org/10.1016/B978-0-444-54314-1.00007-0

Campbell JR, Eden B (2014) Rigid prices: evidence from u.s. scanner data. Int Econ Rev 55(2):423–442. https://doi.org/10.1111/iere.12055

Cavallo A (2013) Online and official price indexes: measuring argentina's inflation. J Monetary Econ 60(2):152–165. https://doi.org/10.1016/j.jmoneco.2012.10.002

Cavallo A (2017) Are online and offline prices similar? Evidence from large multi- channel retailers. Am Econ Rev 107(1):283–303

Cavallo A (2018) Scraped data and sticky prices. Rev Econ Stat 100(1):105–119. https://doi.org/10.3386/w21490

Cavallo A, Gopinath G, Neiman B, Tang J (2021) Tariff pass- through at the border and at the store: evidence from us trade policy. Am Eco- Nomic Rev: Insights 3(1):19–34. https://doi.org/10.1257/aeri.20190536

Devlin J, Chang M-W, Lee K and Toutanova K (2019) Bert: Pre- training of deep bidirectional transformers for language understanding. 54

Eichenbaum M, Jaimovich N, Rebelo S (2011) Reference prices, costs, and nominal rigidities. Am Econ Rev 101(1):234–262. https://doi.org/10.1257/aer.101.1.234

Fuest C, Neumeier F, Stöhlker D (2025) The pass-through of temporary VAT rate cuts: evidence from German supermarket retail. Int Tax Public Financ 32(1):51–97

Gagnon E (2009) Price setting during low and high inflation: evidence from Mexico*. Q J Econ 124(3):1221–1263. https://doi.org/10.1162/qjec.2009.124.3.12217

Garc´ıa-Miralles E (2023) Support measures in the face of the energy crisis and the rise in inflation: an analysis of the cost and distributional effects of some of the measures rolled out based on their degree of targeting. *Economic Bulletin*. https://api.semanticscholar.org/CorpusID:257395075

Gautier E, Marx M, Vertier P (2023) How do gasoline prices respond to a cost shock? J Polit Econ Macroecon 1(4):707–741. https://doi.org/10.1086/727698

Gautier E, Conflitti C, Faber RP, Fabo B, Fadejeva L, Jouvanceau V, Menz JO, Messner T, Petroulas P, Roldan-Blanco P, Rumler F (2024) New facts on consumer price rigidity in the euro area. Am Econ J: Macroecon 16(4):386–431. https://doi.org/10.1257/mac.20220289

Hansen S, Lambert PJ, Bloom N, Davis SJ, Sadun R, Taska B (2023) Remote work across jobs, companies, and space. Natl Bur Econ Res. https://doi.org/10.3386/w31007

Henkel L, Wieland E, Błażejowska A, Conflitti C, Fabo B, Fadejeva L, Jonckheere J, Karadi P, Macias P, Menz JO, Seiler P (2023). Price setting during the coronavirus (COVID-19) pandemic. Occasional Paper Series, 324, European Central Bank. https://ideas.repec.org/p/ecb/ecbops/2023324.html

Kehoe P, Midrigan V (2015) Prices are sticky after all. J Monetary Econ 75:35–53. https://doi.org/10.1016/j.jmoneco.2014.12.004

Lafrogne-Joussier R, Martin J, Mejean I (2023). "Cost pass-through and the rise of inflation". mimeo. http://www.isabellemejean.com/lmm_ppi_v2.pdf

Lehmann E, Simonyi A, Henkel L and Franke J (2020) Bilingual transfer learning for online product classification. In: *Proceedings of Workshop on Natural Lan- guage Processing in E-Commerce*. Association for Computational Linguistics, pp. 21–31. https://aclanthology.org/2020.ecomnlp-1.3

Montag F, Sagimuldina A and Schnitzer M (2020) "Are temporary value- added tax reductions passed on to consumers? Evidence from germany's stimulus".

Nakamura E, Steinsson J (2008) Five facts about prices: a reevaluation of menu cost models. Q J Econ 123(4):1415–1464

Amores AF, Basso HS, Bischl JS, De Agostini P, Poli SD, Dicarlo E, Flevotomou M, Freier M, Maier S, García-Miralles E, Pidkuyko M, Ricci M and Riscado S (2023b) In- flation, Fiscal Policy and Inequality. ECB Occasional Paper No, 2023/330, European Central Bank. https://doi.org/10.2139/ssrn.4604418

Reimers N, and Gurevych I (2019) Sentence-bert: Sentence embeddings using siamese bert-networks. 5, 51, 56

Sanh V, Debut L, Chaumond J, Wolf T (2020). Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. 5, 11, 54, 55

Strasser G, Wieland E, Macias P, B±açzejowska A, Szafranek K, Wittekopf D, Franke J, Henkel L, Osbat C (2023). E-commerce and price setting: evidence from Europe. Occasional Paper Series, 320, European Central Bank. https://ideas.repec.org/p/ecb/ecbops/2023320.html 9

Tunstall L, Reimers N, Jo UE, Bates L, Korat D, Wasserblat M, Pereg O (2022). Efficient few-shot learning without prompts. 5, 56

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L and Polosukhin I (2023) Attention is all you need. 53

Wooldridge, Jeffrey M. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*. 20

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.