Analysing the VAT Cut Pass-Through in Spain using Web Scraped Supermarkets Data and Machine Learning*

Nicolás Forteza[1], Elvira Prades[†] and Marc Roca[1]

[1]Bank of Spain.
April 18, 2024

Abstract

On 28 December 2022, the Spanish government announced a temporary value added tax (VAT) rate reduction for selected products. VAT rates were cut on January 1, 2023 and are expected to go back to their previous level by mid-2024. Using a web-scraped data set, we leverage machine learning techniques to classify each product. Then we study the price effects of the temporary VAT rate reduction covering the daily prices of roughly 10.000 food products sold on-line in a Spanish supermarket. 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. We observe differences in the speed of pass-through across different product types.
JEL Codes: E310, H220, H250.

Keywords: price rigidity, inflation, consumer prices, heterogeneity, microdata, VAT Pass-Through.

*Corresponding authors: Nicolás Forteza (nicolas.forteza@bde.es); Elvira Prades (elvira.prades@bde.es) and Marc Roca (marc.roca@bde.es). We thank Erwan Gautier (BdF) and Aitor Lacuesta (BdE) for useful comments and guidance. We also appreciate comments from participants from Internal Seminars at Bank of Spain and Banque de France, as well as from participants in the annual "Bank of Spain - CEMLA Experts Meeting on Big Data Applications and Data Analytics in Central Banking". 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 España 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 specially of the *Service des Ètudes Microéconomiques* (SAMIC) is gratefully acknowledged.

Resumen
El 28 de diciembre de 2022, el gobierno español anunció una reducción del impuesto sobre el valor añadido (IVA) con caracter temporal para determinados productos. Los tipos de IVA se redujeron el 1 de enero de 2023 y se espera que vuelvan a su nivel anterior a mediados de 2024. Utilizando un conjunto de datos obtenidos mediante técnicas de web scraping, hacemos uso de métodos de aprendizaje automático para clasificar cada producto. A continuación, estudiamos los efectos sobre los precios de la reducción temporal del tipo de IVA, analizando los precios diarios de aproximadamente 10.000 productos alimenticios vendidos online en un supermercado español. Para identificar los efectos causales sobre los precios, comparamos la evolución de los precios de los artículos sujetos a la medida (es decir, gravados por la política fiscal) con un grupo de control (productos alimenticios fuera del alcance de la política). Nuestros resultados indican que, a nivel de supermercado, la repercusión del IVA fue casi completa. Sin embargo, observamos diferencias en la traslación a precios finales entre los distintos tipos de productos.

1 Introduction

In the face of recent exogenous economic shocks, such as the COVID-19 pandemic, subsequent supply disruptions, and the ongoing war in Ukraine, governments have used fiscal measures such as VAT reductions to mitigate their impact. To address the prevailing high-inflation scenario, the Spanish government announced on 28 December 2022 a temporary reduction in the VAT rate for essential food items. Effective January 1, 2023, this 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 expected to end in June 2024.[1]

This policy was designed to alleviate the burden of high inflation, particularly for low-income households, since lower-income households allocate a larger portion of their total spending on food. According to recent estimates, in 2021, inflation for Spanish low-income households (bottom quartile) was 2 pp higher (see Basso et al. (2023)).

Consequently, 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 pasta, 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 the fiscal year 2023 (according to Spanish Ministry of Economy, Commerce and Business). However, this measure is effective to the extent that supermarkets pass the lower tax to final prices.

In this paper, 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 ECB.[2] We also used an alternative data set provided by Datamarket to check the robustness of our results using the data from a different supermarket, with a different pricing strategy and where the data collection follows the same principles of DPD. In a first step, all available products need to be classified so as to map the Classification of Individual Consumption according to Purpose (COICOP).[3] In doing so, we first use novel natural

[1]On the 27th of June of 2023 the Government announced an extension of this measure. It was already anticipated, as the Spanish Ministry of Economy announced in several interviews in previous days, that this was a likely scenario. At the same time, in the extension, the possibility of reversing the measure in advance was taken into account. 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.

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

[3]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) subclass, and (5) product.

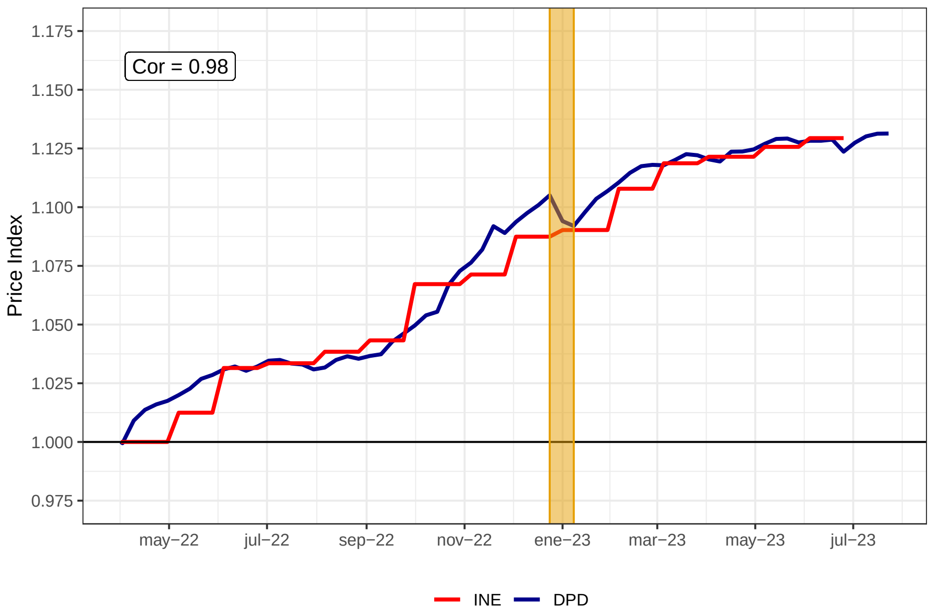


Figure 1: Food Consumer Price Index

Notes: This graph plots the aggregate re-constructed index (blue) against the Spanish National Statistical Office (INE) index (red). The reconstructed index uses the cumulative product of the weekly price difference (log), weighted by the share of consumption of every COICOP5 within the food group using web-scraped data from DPD PRISMA. The shaded area in orange represents a notable date in the underlying time period for the VAT cut in January 2023.
Sources: National Statistical Office (INE in Spanish) and the authors' own calculations using the Daily Price Dataset (DPD) PRISMA-ECB.

processing techniques to label a sub-sample and subsequently train a model to map all food products according to the official product classification. Due to the level of granularity and time frequency of the data, we reconstruct price indices on a pseudo-real-time basis and compare them against the official figures from the National Institute of Statistics (INE by its Spanish acronym), which provides data on a monthly basis. **Figure 1** exhibits both the reconstructed and the official food price indexes, which change over time very similarly, as our high Pearson correlation coefficient suggests ($\rho = 0.98$).

This timeline also indicates that the policy response may be underestimated when using aggregate monthly data, since only the DPD index captures a price drop due to the tax cut (highlighted by the orange area in **Figure 1**). The aggregate INE data could be hiding part of the effect, probably due to the lower time frequency in the panel data they provide and the broader coverage of type of retailers, regions, etc. However, given the reduced sample of information that accounts for a unique retailer located in one zipcode, the performance of the reconstructed index against the official index is quite good. 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 it is available the day after.

To efficiently capture the effect of the tax cut on retailer prices (pass-through) we will distinguish between treated and untreated products. We rebuild 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 section 5.1), which will give us a closer look to the causal inference of the matter. We also find 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 processed vs. non-processed, trademark vs. white label and imported vs. domestic food products. When exploring these dimensions, we find that the degree of pass-through is somewhat different in particular in its dynamic dimension. When we decompose price dynamics into the price change and the size of this change, overall we 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 on average the products under the VAT cut scheme passed-through, on average, between 70 and 100% of the tax reduction to final prices over time.

The paper is organized as follows. Section 2 summarizes the related economic literature. In Section 3 we describe our main data sources and briefly summarize the methodology used to map food web-scraped products to the official classification. In section 4, we show evidence of price setting and its dynamics before, during, and after the temporary VAT cut period. In Section 5, we estimate the temporary VAT pass-through level. We explore possible pass-through heterogeneities according to product characteristics in Section 5.3 and conclude in Section 6. The Appendix shows a detailed explanation of data cleaning for prices (A) and the product classification approach (C).

2 Related Literature

This paper relates with several strands of the literature. First, it contributes to the literature that analyzes price setting using microdata. A recent study by Gautier et al. (2023) 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. In our work we contribute by showing further evidence at a granular level and covering the most recent inflationary episode. As a drawback, we only analyze prices of two components of the CPI, i.e., food and beverage products, and information for a country-specific retailer. The data for this more recent period of high inflation replicate the stylized facts obtained from other studies. That is, as in Nakamura and Steinsson (2008), the frequency of price increases covariates strongly 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 these data, we investigate the impact of a specific VAT rate cut on price levels announced by the Spanish government. Although similar policies in other countries, like the temporary VAT reduction on all goods in Germany from July 1st to December 31, 2020, typically exhibit high pass-through rates, these effects vary significantly across retailers. For example, the German government on 3 June 2020 announced a temporary VAT reduction applied to all goods. This measure started the 1 July 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 latter as a counterfactual), the authors find a decrease by 1.3% of prices for Germany, implying a 70% level of passthrough. 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).[4] They found that the price reaction to the VAT cut was quick and substantial. 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 similar data set as we do.[5] They found that the pass-through has been almost 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 passthrough in Spain was large and almost complete for most of the products analyzed.

Our work adds a more detailed analysis to explore possible heterogeneities of the degree of pass-through depending on the characteristics of the products. For this, the mapping to COICOP5[6] has proven to be crucial for the analysis. This article uses state-of-the-art AI techniques to classify products to classify them into COICOP 5 categories. We focus on food and beverages and use modern machine learning techniques, specifically a pre-trained DistilBERT (Sanh et al. (2020)) classifier trained with human (but AI-boosted) annotated data to predict the categories of Spanish and Italian food products COICOP 5. With regard to the labeling exercise, and to our knowledge, we explore for the first time the use of Sentence

[4]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."

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

[6]COICOP at the fifth digit level.

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).

This work also relies on vast empirical evidence on the pass-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 passthrough of the VAT reduction was heterogeneous between fuel types, and Gautier et al. (2022) studied the passthrough of wholesale prices, as a proxy for marginal costs, to retail prices.

While this study aims to evaluate a policy designed to alleviate inflation for low-income households, the group most impacted 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

3.1 Data Sources

We used two sources of daily web-scraped prices at the product level. Primarily, we make use of DPD Prisma, collected by the ECB, and to validate some of the results obtained, we use the dataset provided by Datamarket. These two data sets differ in terms of time period, country, and retailer coverage. Using both datasets helps us to check the quality and robustness of these experimental data.

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 dataset contains daily web-scraped price data from several retailers within the Euro Area covering a wide range of goods sold in supermarkets. Automated webscraping 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 retail product price. Data collection started in April 2022 and we have information up to March[7], and up to July 2023.[8] We have information on online prices on a daily basis for around 10,000 products for Spanish and Italian establishments (see figures A.1).[9] For details on data treatment, see **Appendix A**. The data set contains information for 16,000 unique products[10] sold on-line by these retailers between April 2021 and July 2023. Most products are in the food and beverage categories, cleaning supplies, and personal care products. This level of disaggregation allows for a very precise identification of products. For example, in our data, a *"Coca Cola zero azúcar zero cafeína botella 1,25 l"* and a *"Coca Cola zero azúcar zero cafeína pack 4 botellas 2 l"* are two separate items in the soft beverages product group. For a detailed overview of the coverage of food products, see table C.8.

3.1.2 Datamarket

This proprietary data set is also obtained by a third party using web-scraping techniques. Initially, it was an open source dataset.[11] This dataset contains information on daily 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 (see panel (b) in **figure** A.1). 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**. For product coverage, see table C.8. Given the experimental nature of the data, we limit the analysis to one of the supermarkets as it fulfills the criteria on data quality.

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

[7] This is due to problems with firewalls set by the establishments.

[8] For more details see box 3 "The ECB Daily Price Dataset" in Strasser et al. (2023).

[9] Data is available for a French supermarket for about 12,000 products. But during a non-negligible amount of weeks there were some technical problems to obtain the data. Therefore, the French sample has been dropped from the analysis.

[10] GTIN-EAN 13 is the identifier of a unique product that is being sold in more than one country. This allows us to proxy imported products vs. domestic products.

[11] Since March 2023 the dataset is available under subscription.

[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.

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 pseudoreal 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 the Appendix C.

3.2.1 Labelling

We leverage on Sentence Transformer embeddings to perform bulk data labeling under human supervision of Spanish and Italian 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 allowed us to rapidly map product descriptions to COICOP categories in an iterative manner; this is, going 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 processing-based algorithms.

[13]Harmonized Index of Consumer Prices.

3.3 Descriptive Statistics of data classification

In table C.8 in the Appendix we list all classified products that are sold in Spanish supermarkets. Our entire sample for Spanish food products comprises approximately 11,000 products for the DPD-PRISMA database and approximately 4300 products for the proprietary Datamarket database. Note that the CPI share is expressed in tens of percentage points, which means that our coverage represents 17.9% of the consumption basket.

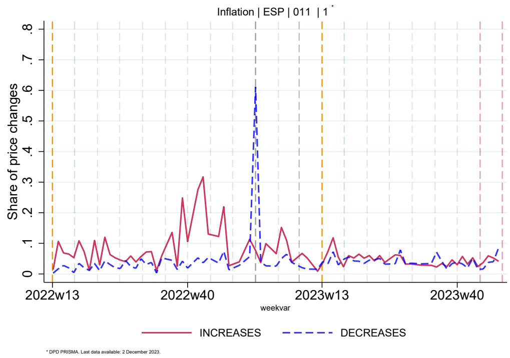
4 Price setting during a temporary VAT cut

To compute the frequency of price changes, we first, we compute the fraction of all weekly price changes over the life-time of each product, and then we calculate the median frequency over all goods within a group of COICOP5. In **figure** 2 we explore the evolution of frequency of price changes, that is, the proportion of products that has changed prices within each treated group. During the first week of January, among affected food products, the aggregate fraction of price change increased to 70% and 80%, meaning that almost all targeted products registered a decrease in price. This is a different pattern with respect to previous and consecutive weeks, where the frequency of price changes was mainly driven by price increases and the share of price changes hovers around the 10%. A similar pattern is observed when using DATAMARKET with an alternative supermarket which is known to have a very specific pricing strategy, based on keeping low prices and not relying much on special offers (see **figure** B.2). When comparing with an Italian supermarket using DPD PRISMA, since in this country there was no VAT reduction, we can observe that there were no major changes in the frequency of price decreases (see **figure** B.3). This fact makes it a good candidate to use it as a control for the underlying evolution of prices of the products targeted by the fiscal measure.

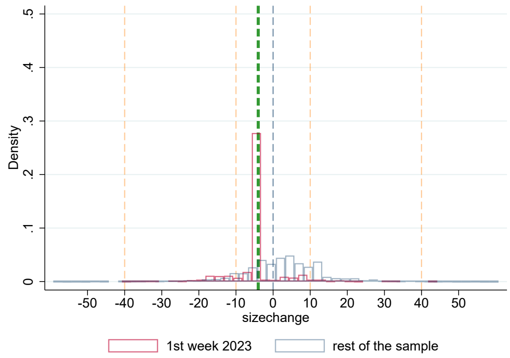
The right panels in **figure** 2 plot the size of nonzero price changes' distribution, showing that among treated goods the size of price changes are concentrated at around the -4% bin. While nontreated products show a bimodal distribution, with positive and negative changes.

(a) From 4 to 0%

(a.1) Frequency

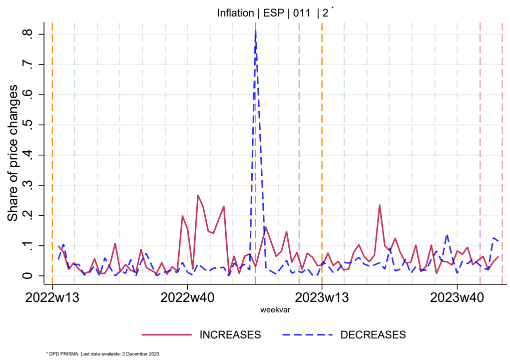


(a.2) Size

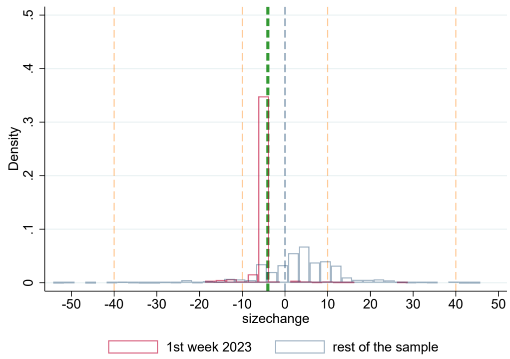


(b) From 10 to 5%

(b.1) Frequency

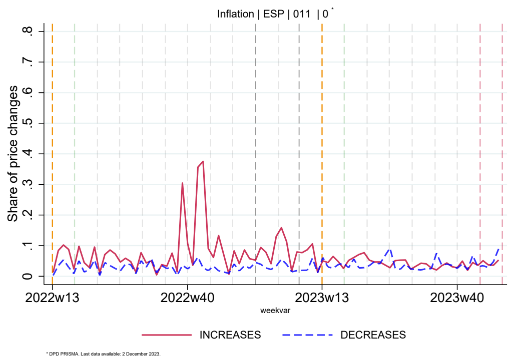


(b.2) Size



(c) Not treated

(c.1) Frequency



(c.2) Size

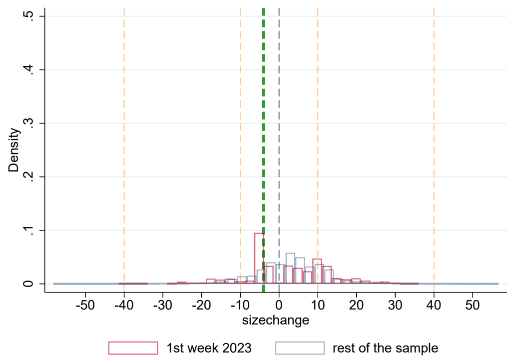


Figure 2: **Frequency and size distribution of price changes**

Spanish Supermarket. DPD PRISMA.

Notes: Panels (a.1) and (b.1) shows the frequency of price changes of CPI items affected by the VAT reduction. The share of prices that increased is depicted in red (—) and in blue (- - -) the share of prices that decrease. Panel (c.1) shows the frequency of those products that are not affected by the measure. In this case, there is no increase in the frequency of price decreases during the first week of 2023. The histograms in panels (a.2), (b.2) and (c.2) show the distribution of the nonzero price size changes (dlog in %) for the three different groups of products (bins of 2.5pp). The shaded red bars correspond to the distribution of nonzero price changes during the first week of 2023. The rest of the sample is depicted in the blue bars. For comparison with a different Spanish supermarket, see figure B.2 and with an Italian supermarket see figure B.3.

Sources: Authors' calculations based on DPD PRISMA ECB.

5 Price effects of the temporary VAT rate change

5.1 Estimating VAT pass-through

In this quasi-experiment we use two types of treatments; the first treatment group (T1) contains the products whose VAT rate was reduced from 4% to 0% (bread, flower, milk, cheese, eggs, fruits, vegetables, legumes, tubers, and cereals), while in the second group (T2) the tax cut was from 10% to 5% (pastas and vegetable oils). These treated groups are compared with a control group consisting of Spanish products not affected by the VAT change. We restrict the analysis to food products, so that our control is more similar to the treatment group.

Figure 3 shows the evolution of the official INE and the reconstructed indices for these three groups as in **Figure 1**, with the orange-shaded area highlighting the period covering one week before and after the measure. **Figure 3** shows that during this two-week period, the treated products decreased their prices by 4.24% for T1 and 3. 61% for T2, while the prices of the items in the control group increased by 1.11%, on average. In a full price pass-through scenario, T1 should cut its prices by 4.55% and T2 by 3.84%.[14] Assuming that in a counterfactual scenario (no policy implementation), treated products would have behaved as the control group, that is, by increasing its prices, by 1.11% , we would obtain a 140% pass-through for T1 and 104% for T2.[15] It is also worth mentioning how similar the trends are for the official index and the reconstructed index, as we obtain a Pearson's correlation coefficient greater than 0.9 in all three cases.

[14]This is computed as follows $(110 - 105)/110 \approx$ -4.55% and $(100 - 104)/104 \approx$ -3.85%.

[15]Price pass-through (PPT) = $\frac{Treated(\Delta\%) - \%Control(\Delta\%)}{Fullpass-through(\Delta\%)}$

PPT (T1)= $\frac{-4.24\% + 1.11\%}{-3.84\%} = 140\%$

PPT (T2)= $\frac{-4.61\% + 1.11\%}{-4.55\%} = 104\%$.

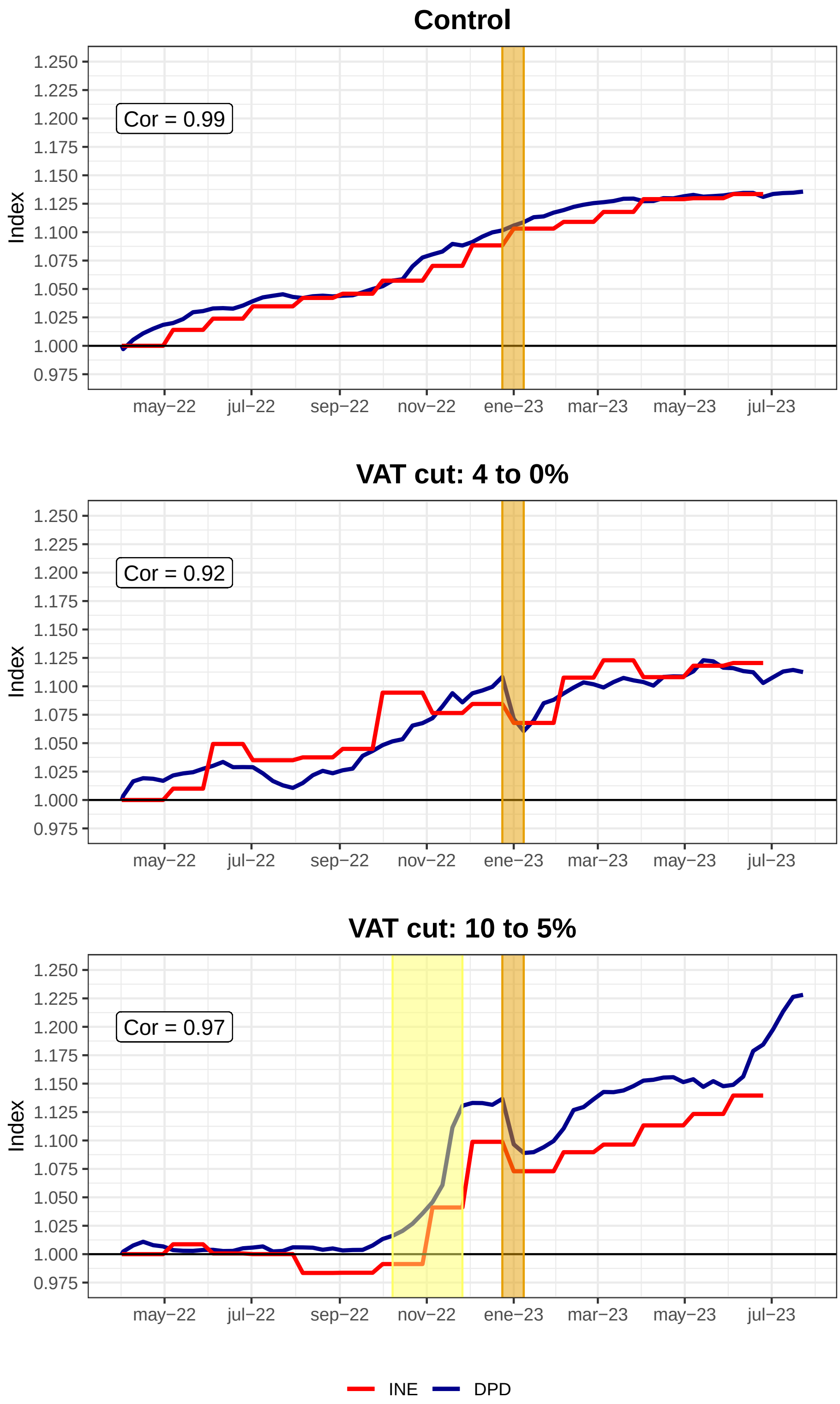


Figure 3: **Food Price Developments**

Notes: These graph plots the aggregate re-constructed index (blue) against the Spanish National Statistical Office (INE) index (red). Here, we break the sample into the (i) the control group, (ii) those products not affected by the VAT reduction and (iii) the group of products for which the VAT went from 4% to 0% and those that changed from 10% to 5%. The orange shaded area represents a notable date in the underlying time period that the VAT cut in January 2023.
Sources: National Statistical Office (INE in Spanish) and authors' calculations using the Daily Price Dataset (DPD) PRISMA-ECB.

In order to get more reliable estimates and fully exploit the granularity of DPD PRISMA we run the following Difference-in-differences Event Study regression as in Fuest et al. (2021). This will provide a more reliable estimate of the VAT pass-through (see **equation 1**):

$$p_{iw} = \sum_{j=-30,\neq -1}^{T} \beta_j \times b_{iw}^j \times (0.0385 \; or \; 0.0455) + \; \theta_w * COICOP_4 + \mu_i + \varepsilon_{iw} \qquad (1)$$

The outcome variable of interest, p_{iw}, is the natural logarithm of the mean weekly (w) price for each product (i). The time unit, weeks, is not specified as calendar weeks, that is, week one includes from day one after the policy implementation until day seven.[16] Therefore, the weeks go from Sunday to Saturday. The β_j shows the (weekly) event-time coefficients that go from 30 weeks before the public policy implementation (first week of January, 2023) until the latest available data, in this case 30 weeks after. The dummy variable b_{iw}^j determines whether the product i is affected by the VAT cut scheme or not. Since we have a two-set of treated products with a different tax reduction, we standardize both price changes to interpret the two effects at the same scale. That is, we substitute the dummy with the expected price decline, either a 3.85% or a 4.55%. Therefore, an estimated β_j=1 will indicate a full VAT cut pass-through to final prices.

The term $[\theta_w * COICOP_4]$ interacts with time dummies with COICOP at 4 digits to capture specific time trends. Finally, μ_i stands for product-fixed effects, and we cluster standard errors by product (ε_{iw}). This approach provides estimates for week-specific relative price adjustments in response to VAT rate changes taking as a base period *j=-1*, the last time period (week) before the tax cut. Thus, our coefficients of interest will be the log difference between two periods, which approximates the change rate.

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. **Figures 4** to **7** plot the event time coefficients that will allow us to check the pretreatment balance between the treatment and control group and whether the condition of parallel trends is likely to hold. 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

[16] The VAT reduction entered in force on a Sunday, the 1st of January, according to the natural calendar this day corresponds to week 52 in year 2022. To ensure that the prices on 1 January fall into week 1 in the year 2023, we shift all daily observations backwards one day.

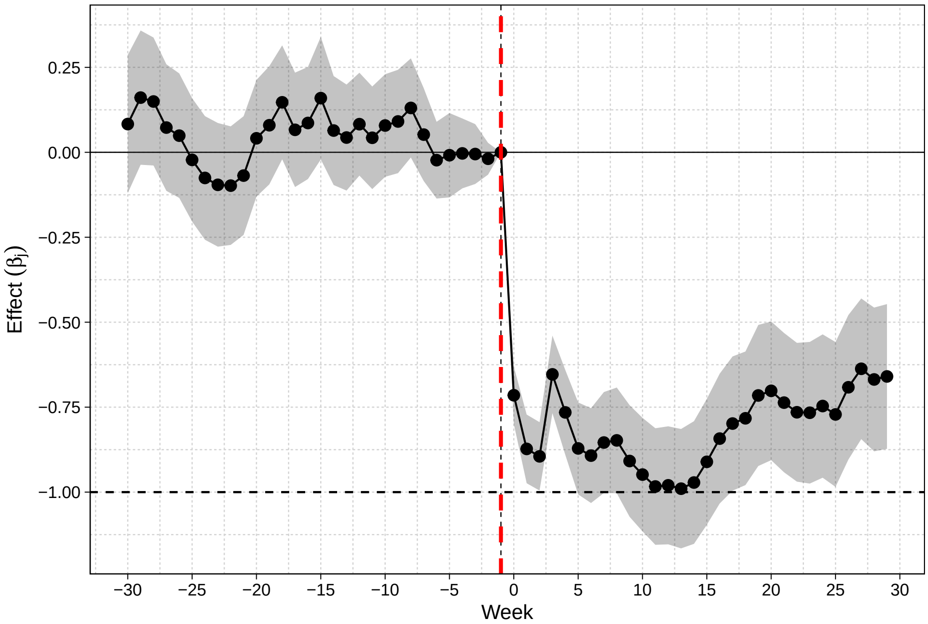


Figure 4: **Event study:** Retailer 1
Spanish Supermarket. DPD Prisma.

Notes: 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 we take as a reference.
Sources: Authors' calculations based on DPD ECB Prisma.

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 will be examined in detail when we go into the empirical findings (see section 5.2).

5.2 Results

We first regress all products, comparing the treated (21.5% of the products) and the control group (78.5%). **Figure** 4 shows the event-time coefficients β_j of the regression specified in **equation** 1.

Our coefficients before the event (VAT cut) are not statistically different from 0 and do not follow a specific trend; the "0 effect threshold" is crossed on repeated occasions, which suggests that the parallel trends assumption is likely to hold. Then, after the tax change, there is an important discontinuity, indicating that products that were under the VAT cut scheme passed-through, on average, between 70 and 100% of the tax reduction to final prices over time. Despite the quick response of the retailers, who reached an almost complete pass

90% the second week after the event ($\beta_j \approx 0.9$), it is not until two and a half months after (Week = 11) when the complete pass is reached. However, some weeks later, the effect dilutes going back to around 70% passthrough. As mentioned in Section 5.1, two treatment natures have been implemented depending on the VAT reduction: 4 to 0%, and 10 to 5%, which stand for 7% and 77% of the total weight of the CPI, respectively. Of the 2,379 food items that fell under the VAT cut scheme, 1,965 (82.6%) belong to the 4 to 0% group and the remaining 414 (17.4%) belong to the 10 to 5% group. According to the Spanish Ministry of Economy, products that became VAT-free are considered as "staple food", while the other group labeled as "second-class" items. The first group has products in eleven different $COICOP_5$ classes, the second group, in contrast, has products in only three $COICOP_5$. Furthermore, as shown in **Figure 5**, when comparing these two groups to products not affected by tax policy (control group), we find different treatment responses. On the one hand, the response of staple food products is quite similar to that stated previously in **Figure 4**, but having a lower degree of price transmission: 70% during the first 15 weeks, and then diluting to 50% (orange coefficients). However, second-class products seem to be a bit more difficult to fit our empirical design. Products under the 10- to 5% cut experienced a sharp price increase and a large trend change during October and November 2022 (yellow-shaded area in **Figure 3**), which could damage 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 (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 $[COICOP_4 * \theta_w]$ interaction term in **equation 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.

Finally, 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 different price-setting strategy. The results are shown in **Figure 6**, 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.

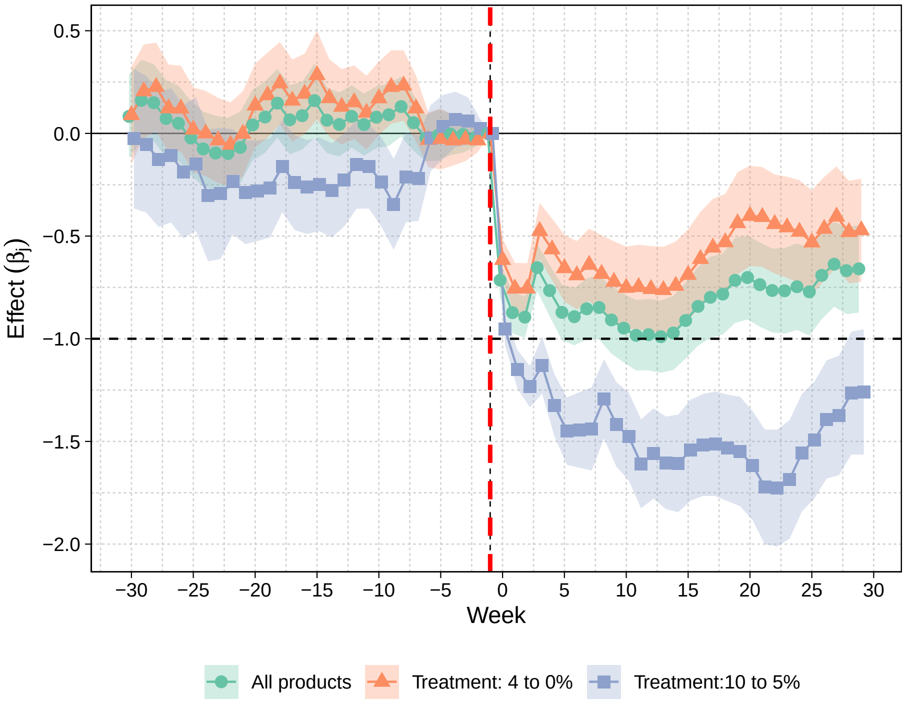


Figure 5: **Event study exploiting VAT heterogeneity**

Notes: This figure show the estimates of the degree of VAT pass-through for different treated products. In orange the degree of pass-through in treated products from 4% to 0%. In blue the degree of pass through in the treated products from 4% to 0%. In green the baseline estimation. 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 we take as a reference.

Sources: Authors' calculations based on DPD PRISMA ECB.

5.3 Pass-Through across products characteristics: Treatment heterogeneity

Treatment effects may vary depending on certain product characteristics. In **Equation 1**, we control for the product fixed effect (μ_i), so 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 sub-samples to examine different behaviors in product characteristics that can be identified within the PRISMA dataset: Processed vs Unprocessed (in section 5.3.1), Trademark vs White-label (section 5.3.2), and Domestic vs Imported goods (section 5.3.3). In **Table 2** we observe the share of products in these different categories. Although the origin product subsamples are quite balanced, there are far fewer unprocessed products than processed products, which will have an impact on this analysis.

To quantify the effect of treatment across these categories, we first approach this question from a simpler point of view, looking at fewer time periods (two weeks before and two after the tax cut: $w \in \{-1 : 2\}$) to address the immediate policy response. This gives us an idea

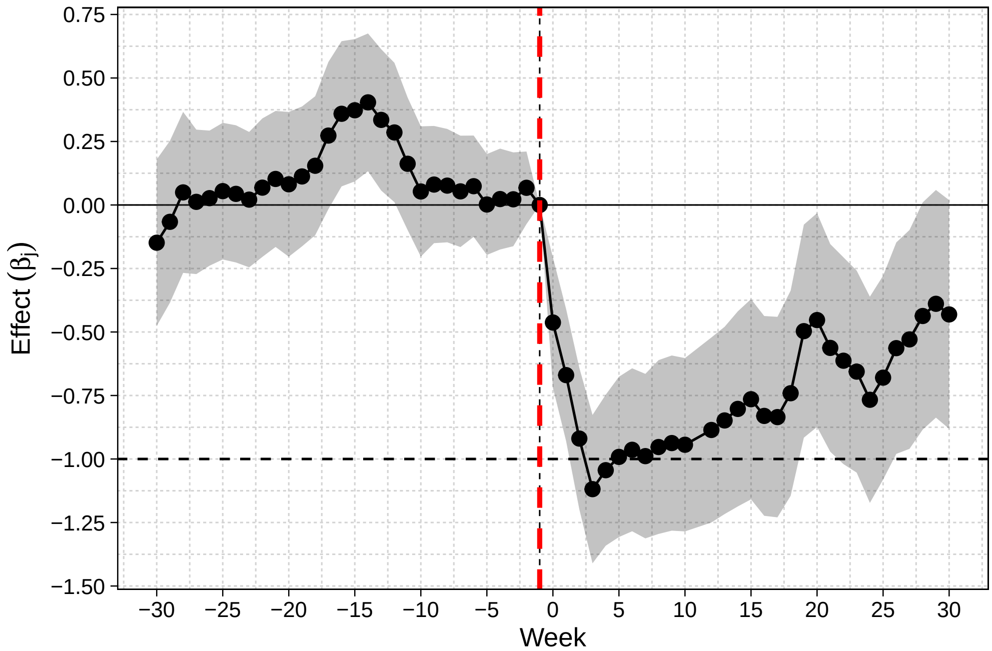


Figure 6: **Event study:** Retailer 2
Spanish Supermarket. DATAMARKET.

Notes: 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.

Table 2: Product Distribution by Category for Retailer 1

Domestic	**No**						**Yes**						
Trademark	**Yes**		**No**		**-**		**Yes**		**No**		**-**		
Processed	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	**No**	**Yes**	
Never Treated	163	719	55	1639	60	165	76	1972	44	3268	42	492	8695
From 4 to 0%	77	205	62	328	35	98	128	311	78	454	93	96	1965
From 10 to 5%	0	50	0	79	0	3	0	100	0	171	0	11	414
Total	240	974	117	2046	95	266	204	2383	122	3893	135	599	**11074**

Notes: 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 vs. non-processed and trademark vs. white-labeled products.

Sources: Authors' calculations based on DPD PRISMA ECB.

of what we can expect given some degree of heterogeneity. In **equation 2**, we regress on several outcome variables (y_{iw}): first on prices (in logarithms) and then on the probability of a price change (Probit model). In the latter, price change is measured in absolute terms and then a distinction is made between positive and negative price changes.

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

H_i is the heterogeneity parameter, A_w is a dummy that takes the value 1 after 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. In **tables 3** to **5** we provide results for β_1 to β_7 for the three sub-samples, respectively. The first estimator (β_1) will explain the difference in price or probability of price change if an item is processed, trademarked, or domestic ($H_i = 1$) compared to when it is not (H_i=0). β_2 will provide an idea on whether there is a change in the outcomes of interest for the treated items, and β_3 looks at the general change in the outcomes after the policy is implemented. Then, the two-term interactions capture more relevant effects, like how treatment affects differently the heterogeneity dimension (β_4), how the heterogeneity changes after the policy (β_5), and whether treated items behave asymmetrically after the VAT cut (β_6). Finally, the most important effect for our research purpose is addressed by the triple interaction term in β_7, which determines how the treated items react after the policy exploiting their characteristics through the heterogeneity measure.

This approach identifies changes in both the extensive (how many treated items changed its price) and intensive margin (measuring the size of the price change). It also captures price level differences between groups that are not shown in the main identification strategy (**equation 1**).

5.3.1 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, since they can be affected by international and exogenous shocks, for example, the COVID-19 pandemic or the Russo-Ukranian war. Thus, the differentiation between processed and unprocessed products will provide insight into how the difference referred to will be interpreted in our experiment. We explore this source of heterogeneity by categorizing each product in our sample following the Eurostat ECoicop -HICP criterion, which assigns food products to processed or unprocessed groups using Coicop categories at the 5-digit level (Coicop); 38 Coicop codes are labeled processed (91. 8% of the products), and the

Table 3: Processed heterogeneity

	ln(Price)	$\Delta Price$	$\Delta Price > 0$	$\Delta Price < 0$
	(1)	(2)	(3)	(4)
Processed	$\boxed{-0.774^{***}}$	0.024	$\boxed{0.062^{**}}$	-0.037
	(0.116)	(0.026)	(0.024)	(0.026)
Treatment	-0.980***	0.029	0.049	-0.019
	(0.163)	(0.041)	(0.037)	(0.026)
After	-0.025**	-0.037	-0.002	-0.035*
	(0.012)	(0.026)	(0.036)	(0.018)
Processed × Treatment	1.039***	-0.007	-0.034	0.027
	(0.243)	(0.056)	(0.047)	(0.029)
Processed × After	0.023*	0.034	-0.020	0.054**
	(0.012)	(0.034)	(0.038)	(0.023)
Treatment × After	-0.011	0.668***	-0.070*	$\boxed{0.738^{***}}$
	(0.012)	(0.050)	(0.039)	(0.033)
Processed × Treatment × After	-0.004	-0.160*	0.032	$\boxed{-0.193^{**}}$
	(0.014)	(0.091)	(0.049)	(0.095)
Observations	32,426	32,184	32,184	32,184

Notes: This table reports the results of the estimation of equation (2) for different dependent variables. Column (1) evaluates the differences in price levels. Columns (2) to (4) evaluates the probability of price changes, either increases or decreases. Prices of processed products are on average cheaper than non processed, 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.

Source: Authors' calculations based on DPD PRISMA ECB.

remaining 11 belong to the unprocessed food group (8.2%).[17] **Table 3** presents **equation 2** results. The first takeaway is that processed goods are, on average, 54% cheaper.[18] Moreover, processed and treated items are 183% more expensive than the rest, which is a relevant price difference. Furthermore, there is a statistically significant 2.5% decrease in prices after the policy is implemented. After the VAT cut, there is a 66. 8% chance that treated products will change their price: 7% fewer chances of obtaining a positive change and 74% more chances of obtaining a decrease in price. Furthermore, processed goods under the VAT cut scheme have 19. 3% fewer chances to go through a price reduction after the implementation of the policy.

Having explored these differences in various outcomes, we move to the subsample analysis to exploit this consistency of treatment heterogeneity over time as in **equation 1**. **Figure 7** plots the results for processed goods (in orange), which are quite consistent with the general

[17]In terms of CPI weights, processed foods account for 14.7% of the total basket, and unprocessed food for 4.5%.

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

model represented in **Figure 4**: close to 100% passing through that slightly dilutes over time and consistent parallel trends. However, unprocessed goods (coefficients in green) show higher volatility in price, ranging from 65 to 140% pass-through levels with a strong cyclical component. Furthermore, these results are consistent with the previous analysis, which predicted that processed items had less chances to shrink their prices within the first two weeks after the tax change. In this case, though, there is a significant difference between the control and treatment group before the policy: a positive but decreasing trend that could be upward biasing our results and hamper the experiment's casual inference validity.

5.3.2 Trademark vs White-label

The price-setting behavior of white-labeled and trademarked products might 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 4** exhibits **equation 2** results for trademark heterogeneity. The first important highlight is the difference in price levels between these two groups: trademarked items are, on average, 31.8% more expensive than white-labeled items. If trademarked and treated products are, on average, 20.1% more expensive than the rest (see column (1)).

Table 4: Trademark heterogeneity

	ln(Price)	$\Delta Price$	$\Delta Price > 0$	$\Delta Price < 0$
	(1)	(2)	(3)	(4)
Trademark	$\boxed{0.276^{***}}$	0.035	0.024	0.012
	(0.055)	(0.022)	(0.017)	(0.013)
Treatment	-0.244	0.044	-0.035	0.079*
	(0.186)	(0.082)	(0.070)	(0.041)
After	0.005	0.033	-0.002	0.035
	(0.009)	(0.053)	(0.028)	(0.041)
Trademark × Treatment	$\boxed{0.183^{**}}$	-0.016	0.028	-0.044**
	(0.076)	(0.059)	(0.055)	(0.021)
Trademark × After	-0.005	-0.023	-0.013	-0.011
	(0.005)	(0.029)	(0.018)	(0.020)
Treatment × After	-0.031	0.462***	-0.005	$\boxed{0.467^{***}}$
	(0.021)	(0.139)	(0.089)	(0.100)
Trademark × Treatment × After	0.007	0.041	-0.021	0.062
	(0.010)	(0.079)	(0.065)	(0.054)
Observations	30,180	29,963	29,963	29,963

Notes: This table reports the results of the estimation of equation (2) for different dependent variables. Column (1) evaluates the differences in price levels. Columns (2) to (4) evaluates the probability of price changes, either increases or decreases. Prices of trademarked products are on average cheaper than white-label. We do not observe any differential behaviour in terms of the probability of price changes, except for the specific episode after the treatment.

Source: Authors' calculations based on DPD PRISMA ECB.

Although we find that treated products have 46.7% more chances of reducing their price after the VAT cut (column (4)), we do not find evidence supporting a different behavior for treated and trademarked products after the VAT cut scheme is applied.

Figure 7 supports this last finding. Trademarked and white-labeled products both showed a large pass-through (around 80%). Although there have been some disparities over time, there is no clear pattern of divergence between these two groups.

5.3.3 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 products) and imported (33. 7%). **Table 2** provides the distribution of treated items in $COICOP_5$ for

[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.

Table 5: Domestic heterogeneity

	ln(Price)	$\Delta Price$	$\Delta Price > 0$	$\Delta Price < 0$
	(1)	(2)	(3)	(4)
Domestic	-0.122**	-0.018	0.010	-0.027***
	(0.052)	(0.018)	(0.014)	(0.009)
Treatment	-0.029	0.038	0.024	0.014
	(0.178)	(0.031)	(0.034)	(0.018)
After	-0.003	-0.008	-0.015	0.006
	(0.003)	(0.023)	(0.014)	(0.015)
Domestic × Treatment	0.018	-0.029	-0.018	-0.012
	(0.152)	(0.027)	(0.026)	(0.025)
Domestic × After	-0.001	0.006	-0.010	0.016
	(0.006)	(0.027)	(0.019)	(0.017)
Treatment × After	-0.024**	0.521***	-0.037	0.559***
	(0.009)	(0.043)	(0.031)	(0.061)
Domestic × Treatment × After	0.005	0.010	-0.006	0.016
	(0.010)	(0.066)	(0.032)	(0.064)
Observations	32,426	32,184	32,184	32,184

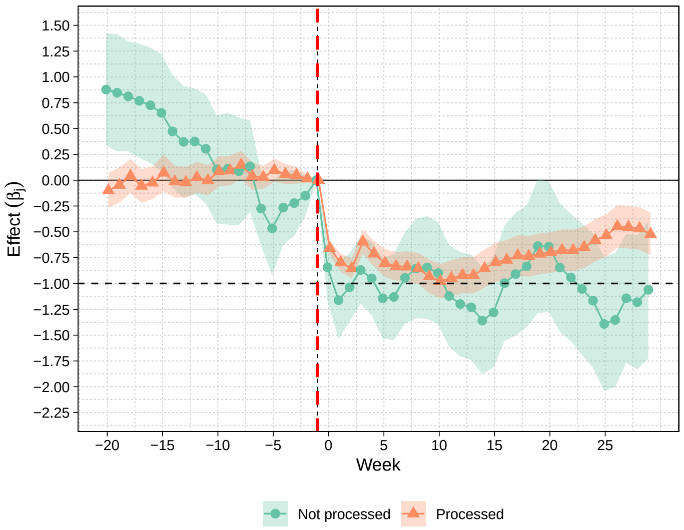
Notes: This table reports the results of the estimation of equation (2) for different dependent variables. Column (1) evaluates the differences in price levels. Columns (2) to (4) evaluates the probability of price changes, either increases or decreases. Domestic products are on average cheaper than imported, there are not substantial increase with a higher probability and, after the introduction of the policy measure, the probability of a price decrease is smaller. The probability of a price decrease is higher after the implementation but the triple interaction does not show any difference.

Source: Authors' calculations based on DPD Prisma ECB.

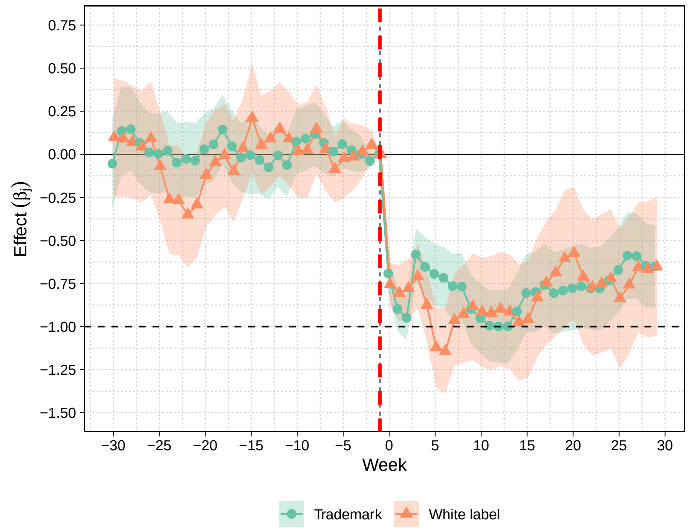
domestic and imported goods. Looking at table 5, domestic goods are, on average, 11. 5% cheaper and less volatile, since they have 3% less chances of experiencing a price reduction. Products under the VAT reduction scheme become 2.4% cheaper in the following two weeks, which would entail a price pass-through 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 **Figure 7**) we spot a notable divergence between domestically produced food items and imported ones. One month after the tax change imported products cross the "full pass-through milestone" and fluctuate below the 100% pass-through for several weeks.

Figure 7: Event study for all types of heterogeneity

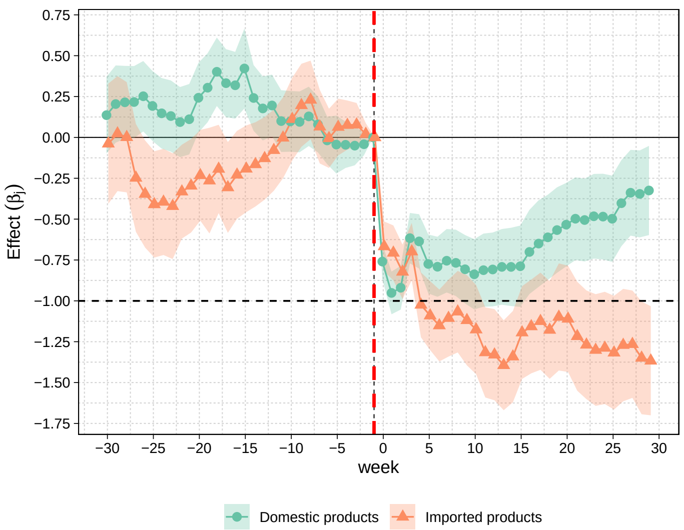
(A)



(B)



(C)



5.4 Triple diff-in-diff: Spain vs. Italy

An additional approach to quantifying the degree of pass-through of the VAT reduction to final prices is by using as control food products under the treated categories of the Italian supermarkets from DPD PRISMA. To do so, the first step was to carefully classify the products under the different COICOP5 categories as explained in Section 3.2.

Then we estimate a triple difference-in-differences specification to capture the different price developments of the Spanish treated products vs. non-treated (the control), and with the Italian goods that are under the categories of the treated COICOP5 items but that have not been under a VAT reduction.

$$\mathrm{p}_{i,week} = \gamma[\mathbb{1}(t > After)_{week} \times (TreatedVAT)_i] +$$

$$\alpha[\mathbb{1}(t > After)_{week} \times (Shop\_{ESP})_i] +$$

$$\beta[\mathbb{1}(t > After)_{week} \times (TreatedVAT)_i \times (Shop\_{ESP})_i] +$$

$$\mathrm{FE}_i + FE_{coicop4,week} + \varepsilon_{iw}$$

The index i refers to the product and the index w refers to the week of observation. The dummy variables refer to the period after ($After$) the announcement, and the variable $Shop\_{ESP}$ takes a value of 1 if the product i is sold in the Spanish supermarket. With this specification, the estimated parameter β captures the price differential with a triple interaction between the treated products, the product being sold in a Spanish supermarket

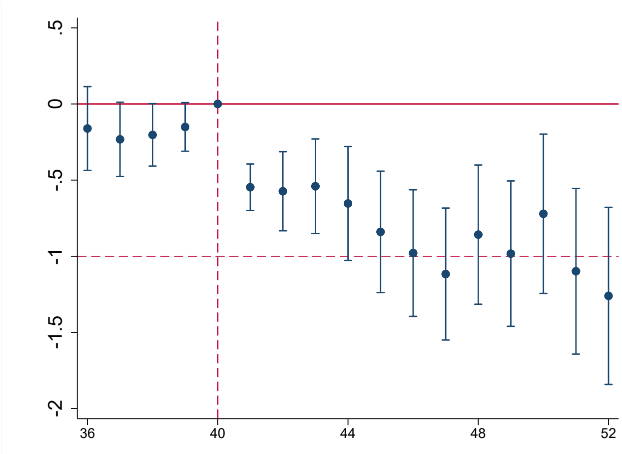


Figure 8: **Event Study based on a triple diff-in-diff**

Notes: We plot the estimated coefficients β from equation X and their 95% confidence intervals. The omitted period is week 52 of year 2022. Standard errors clustered at the product level. The coefficient estimates are transformed to obtain the measure the pass-through rate of the VAT reduction, that is, the share of the VAT rate change that was passed on to consumers. A minus one indicates a full pass through.
Sources: Authors' calculations based on DPD PRISMA ECB.

and the post-period. As before, we include fixed effects for the product and COICOP4×week.

In figure 8 we plot the event study version of the triple difference-in-differences specification. The evolution of the passing follows a very similar pattern to those obtained by restricting the sample to the Spanish products in Figure 4. The fact that the products sold by the Italian retailer are included adds more robustness to the exercise. As in this case we are comparing the price developments of products of the same sub-class, in principle, affected by similar shocks. In this case, the event study indicators show that in the first 4 weeks the pass through is around 70% and then it is completed in the following two months.

6 Conclusions

On 28 December 2022, the Spanish government announced a temporary VAT discount on selected products to alleviate the economic impact of high inflation. We study the effect of the VAT rate cut on retail prices in a Spanish supermarket using web-scraped data collected by DPD PRISMA ECB covering an average 10,000 food product prices per day. These microdata are particularly useful for comparing the pass-through rates of VAT changes, as we can track the prices of individual goods.

As a first step, we check how representative this sample is compared to the official CPI data. As DPD PRISMA ECB only collects data from large retailers, while official statistics cover a variety of retailers and provide a representative geographical coverage. First, to check this at the subclass level we need to correctly map each product to the official classification, that is, into COICOP5, which classifies each good according to its purpose.[20] We classify each product using Machine Learning techniques.

After confirming 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 reduction to final prices, we use an event-study design and compare products targeted by the measure to those not affected. We find that retailers have passed on almost all of the VAT cut for affected products. That is, most of the treated products registered a price decrease, and the size of this decrease corresponded to the established price reduction, either from 4% to 0% or from 10% to 5%.

When looking at price dynamics over time and exploiting observed product characteristics,

[20]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) subclass, and (5) product.

we find some heterogeneities between processed vs. nonprocessed goods, trademark goods vs. white labeled, and domestic vs. imported goods. These differences might reflect differences in price setting and negotiation between producers and retailers.

Initially intended as a temporary measure with a planned six-month duration, the VAT cut has been extended twice due to persistent inflation in 2023. As the reversal hinges on the core inflation trajectory, uncertainty remains. A natural extension of this work would be to track the pricing behavior of retailers and to analyze whether their response to the reversal is symmetrical to the cut.

References

Almunia, Miguel, Javier Martinez and Angel Martinez. (2023). "La reducción del iva en los alimentos básicos: evaluación y recomendaciones". *EsadeEcPol Brief*, 36. 6

Amores, Antonio F., Salvador Barrios, Raffael Speitmann and Daniel Stoehlker. (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 6

Amores, Antonio F., Henrique S. Basso, Johannes Simeon Bischl, Paola De Agostini, Silvia De Poli, Emanuele Dicarlo, Maria Flevotomou, Maximilian Freier, Sofía Maier, Esteban García-Miralles, Myroslav Pidkuyko, Mattia Ricci and Sara Riscado. (2023b). "Inflation, Fiscal Policy and Inequality". ECB Occasional Paper No., 2023/330, ECB.
https://doi.org/http://dx.doi.org/10.2139/ssrn.4604418 7

Basso, Henrique S, Ourania Dimakou and Myroslav Pidkuyko. (2023). "How inflation varies across Spanish households". Occasional Papers, 2307, Banco de España.
https://doi.org/https://doi.org/10.53479/ 3

Benzarti, Youssef, Santiago Garriga and Dario Tortarolo. (2022). "Can vat cuts dampen the effects of food price inflation?" Working paper.
https://dtortarolo.github.io/WebPage/VAT_inflation.pdf 6

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.
http://dx.doi.org/10.1016/B978-0-444-54314-1.00007-0 7

Cavallo, Alberto, Gita Gopinath, Brent Neiman and Jenny Tang. (2021). "Tariff pass-through at the border and at the store: Evidence from us trade policy". *American Economic Review: Insights*, 3 (1), pp. 19–34.
https://doi.org/10.1257/aeri.20190536 7

Devlin, Jacob, Ming-Wei Chang, Kenton Lee and Kristina Toutanova. (2019). "Bert: Pre-training of deep bidirectional transformers for language understanding". 41

Eichenbaum, Martin, Nir Jaimovich and Sergio Rebelo. (2011). "Reference prices, costs, and nominal rigidities". *American Economic Review*, 101 (1), pp. 234–62.
https://doi.org/10.1257/aer.101.1.234 8

Fuest, Clemens, Florian Neumeier and Daniel Stöhlker. (2021). "The Pass-Through of Temporary VAT Rate Cuts: Evidence from German Supermarket Retail". CESifo Working Paper Series, 9149, CESifo.
https://ideas.repec.org/p/ces/ceswps/_9149.html 6, 14

Gagnon, Etienne. (2009). "Price Setting during Low and High Inflation: Evidence from Mexico*". *The Quarterly Journal of Economics*, 124 (3), pp. 1221–1263.
https://doi.org/10.1162/qjec.2009.124.3.1221 6

García-Miralles, Esteban. (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 7

Gautier, E., C. Conflitti, , Fabo B. Faber, R. P., L. Fadejeva, V. Jouvanceau, J.-O. Menz, T. Messner, P. Petroulas, P. Roldan-Blanco, F. Rumler, S. Santoro, E. Wieland and H. Zimmer. (2023). "New Facts on Consumer Price Rigidity in the Euro Area". *American Economic Journal: Macro*, (Forthcoming). 5

Gautier, Erwan, Magaly Marx and Paul Vertier. (2022). "How do gasoline prices respond to a cost shock". *Journal of the European Economic Association*.
https://doi.org/https://doi.org/10.1086/727698 7

Hansen, Stephen, and Elliott Ash. (2023). "Text Algorithms in Economics". CEPR Discussion Papers, 18125, C.E.P.R. Discussion Papers.
https://ideas.repec.org/p/cpr/ceprdp/18125.html 37

Hansen, Stephen, Peter John Lambert, Nicholas Bloom, Steven J Davis, Raffaella Sadun and Bledi Taska. (2023). "Remote work across jobs, companies, and space". (31007).
https://doi.org/10.3386/w31007 7, 9, 40

Henkel, Lukas, Elisabeth Wieland, Aneta Błażejowska, Cristina Conflitti, Brian Fabo, Ludmila Fadejeva, Jana Jonckheere, Peter Karadi, Paweł Macias, Jan-Oliver Menz and Pascal and Seiler. (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 6

Kehoe, Patrick, and Virgiliu Midrigan. (2015). "Prices are sticky after all". *Journal of Monetary Economics*, 75 (C), pp. 35–53.
https://doi.org/10.1016/j.jmoneco.2014.12 31

Lafrogne Joussier, R., J. Martin and I. Mejean. (2023). "Cost pass-through and the rise of inflation". mimeo.
http://www.isabellemejean.com/lmm_ppi_v2.pdf 7

Lehmann, Erik, András Simonyi, Lukas Henkel and Jörn Franke. (2020). "Bilingual transfer learning for online product classification". In *Proceedings of Workshop on Natural Language Processing in E-Commerce*. Association for Computational Linguistics, pp. 21–31.
https://aclanthology.org/2020.ecomnlp-1.3 42

Montag, Felix, Alina Sagimuldina and Monika Schnitzer. (2020). "Are temporary value-added tax reductions passed on to consumers? evidence from germany's stimulus". 7

Nakamura, Emi, and Jón Steinsson. (2008). "Five Facts about Prices: A Reevaluation of Menu Cost Models". *The Quarterly Journal of Economics*, 123 (4), pp. 1415–1464.
https://ideas.repec.org/a/oup/qjecon/v123y2008i4p1415-1464..html 5, 31

Reimers, Nils, and Iryna Gurevych. (2019). "Sentence-bert: Sentence embeddings using siamese bert-networks". 7, 37, 43

Sanh, Victor, Lysandre Debut, Julien Chaumond and Thomas Wolf. (2020). "Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter". 6, 9, 40, 41

Strasser, Georg, Elisabeth Wieland, Paweł Macias, Aneta Błażejowska, Karol Szafranek, David Wittekopf, Jörn Franke, Lukas Henkel and Chiara Osbat. (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 8

Tunstall, Lewis, Nils Reimers, Unso Eun Seo Jo, Luke Bates, Daniel Korat, Moshe Wasserblat and Oren Pereg. (2022). "Efficient few-shot learning without prompts". 7, 43

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin. (2023). "Attention is all you need". 40

Wooldridge, Jeffrey M. (2021). "Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators". *Available at SSRN 3906345*. 15

Appendix

A Data on prices on-line

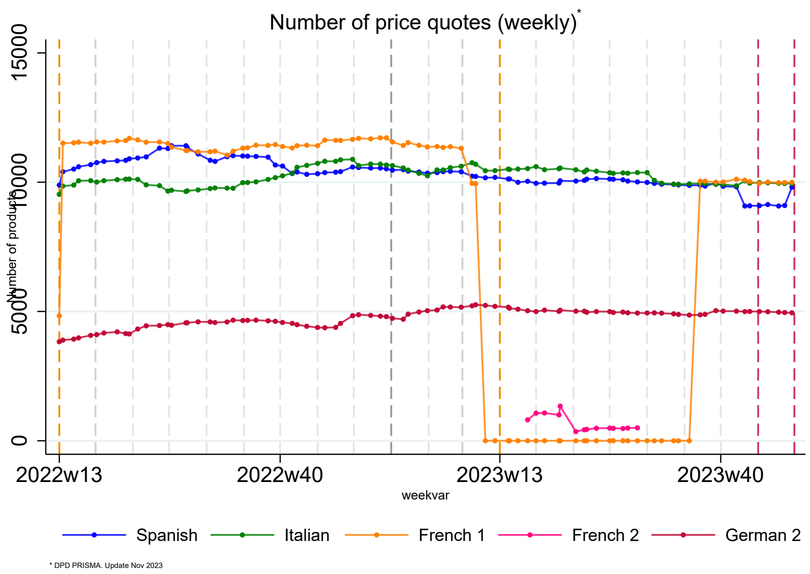
A.1 Data Cleaning and treatment

To correctly treat the data we clean the datasets as follows, we drop observations that are clearly errors (such as exorbitant prices) and we 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 46 days (percentile 1%) out of the maximum number of possible daily observations (for DPD Prisma this is from the 1st of April 2022 up to the October 2023). If we observe missing observations for a certain amount of days, (15 days), these are filled forward with the most recent usable price and replace the unusable or missing observation to fill in the gaps. These gaps are related with a problem with the bot 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. The price series will be labelled as **filled** series.

Then, we filter the series to account for temporary sales, 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 we use the filter proposed by Kehoe and Midrigan (2015)). In some specific cases the retailers show a particular pricing behavior so we relax the condition where the criteria are to 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.

For the regression analysis, we work with weekly data instead of daily data. We take the last observation of a given week. We also work with monthly data to report some statistics to make them comparable with the existing literature. In this case we also take either the mean, the mode, or the last observation.

(a) DPD Prisma



(b) Datamarket

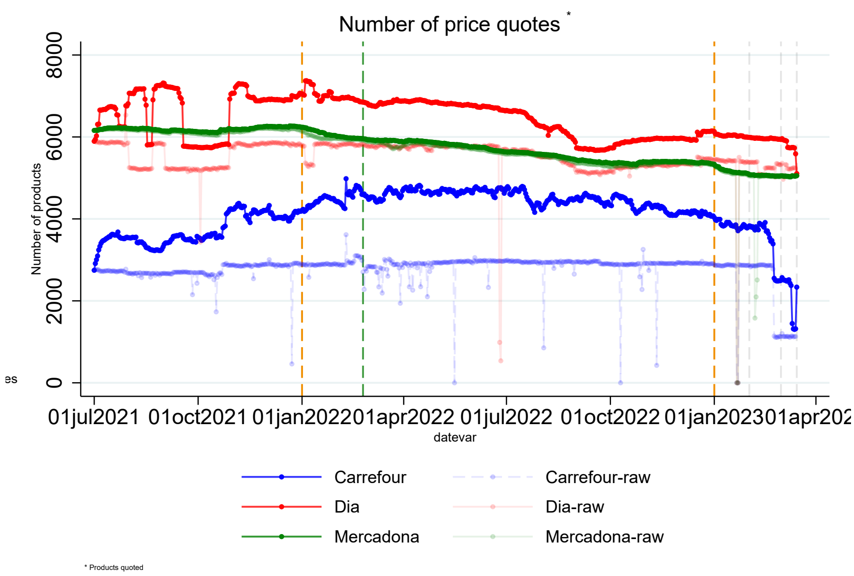
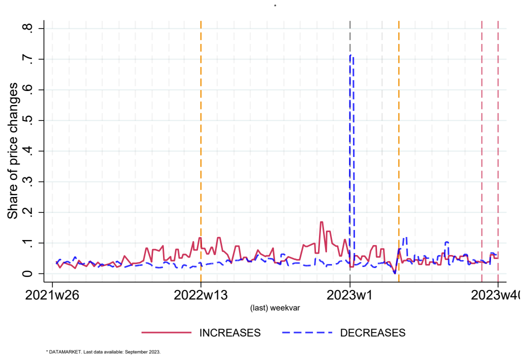


Figure A.1: Weekly number of products

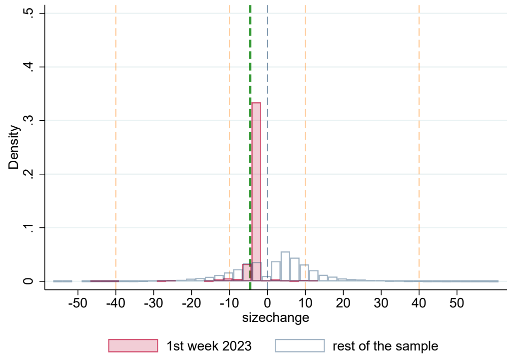
B Additional graphs

(a) From 4 to 0%

(a.1) Frequency

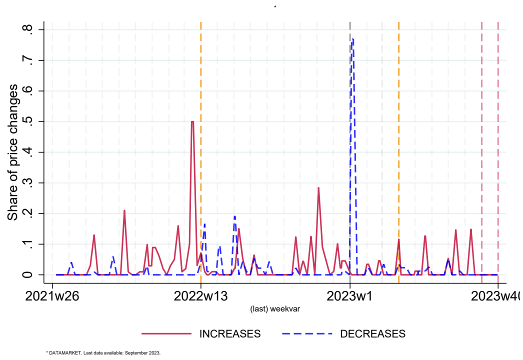


(a.2) Size

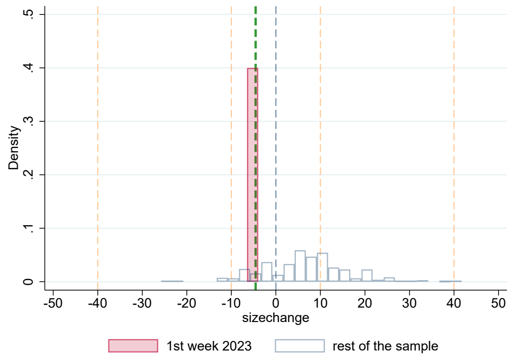


(b) From 10 to 5%

(b.1) Frequency

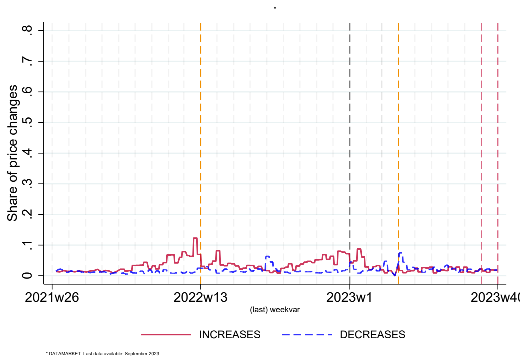


(b.2) Size



(c) Not treated

(c.1) Frequency



(c.2) Size

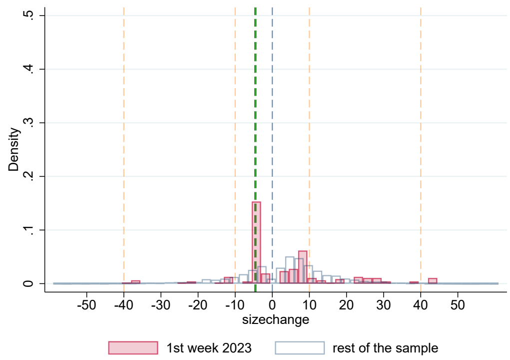


Figure B.2: **Frequency of price changes and size distribution**

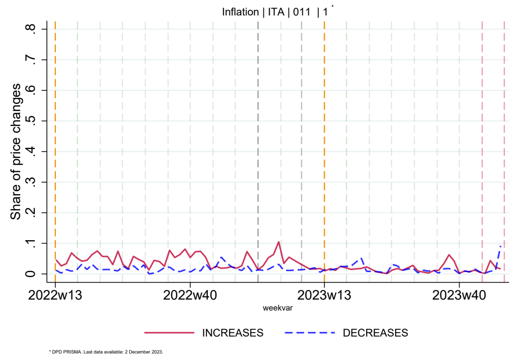
Spanish Supermarket. DATAMARKET.

Notes: Panels (a.1) and (b.1) shows the frequency of price changes of CPI items affected by VAT reduction. The share of prices increases is depicted in red (—) and in blue (- - -) the share of prices decreasing. Panel (c.1) shows the frequency of those products not affected by the measure. In this case, there is not an increase in the frequency of price decreases during the first week of 2023. The histograms in panels (a.2), (b.2) and (c.2) show the distribution of the nonzero price change size (dlog in%) for the three different groups of products (bins of 2.5pp). The shaded red bars correspond to the distribution of nonzero price changes during the first week of 2023. The rest of the sample is depicted in the blue bars. Back to figure **??**.

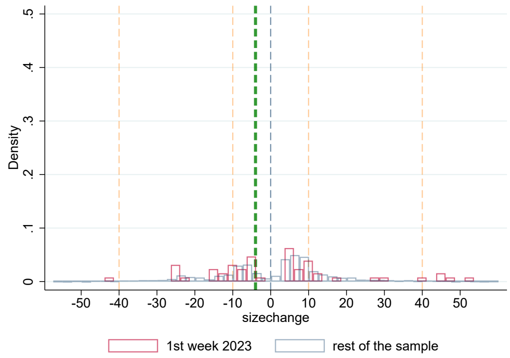
Sources: Authors' calculations based on DPD PRISMA ECB.

(a) From 4 to 0%

(a.1) Frequency

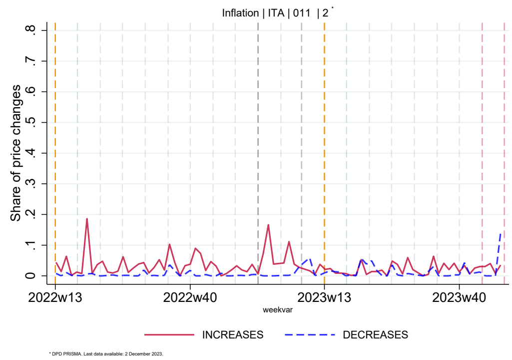


(a.2) Size

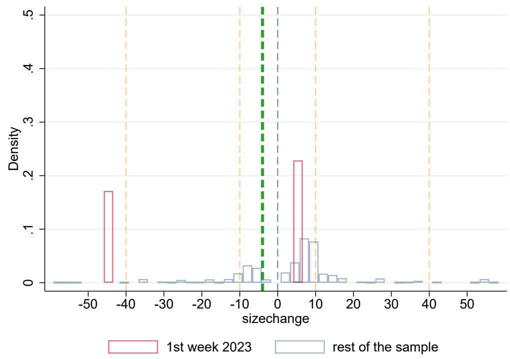


(b) From 10 to 5%

(b.1) Frequency

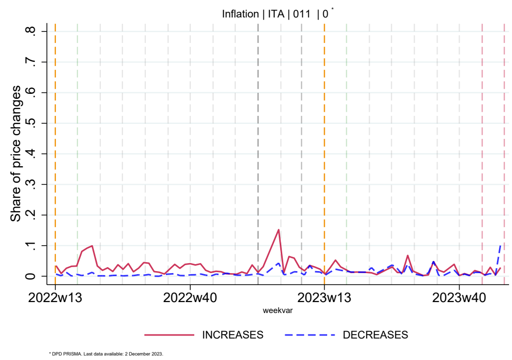


(b.2) Size



(c) Not treated

(c.1) Frequency



(c.2) Size

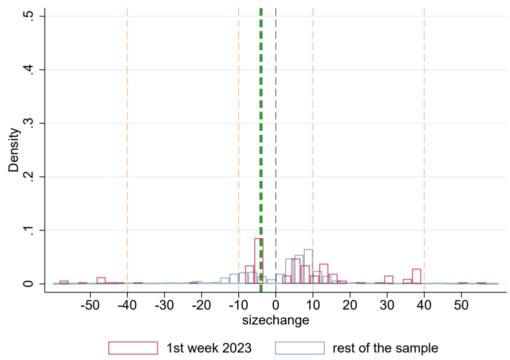


Figure B.3: **Frequency of price changes and size distribution.**
Italian Supermarket. DPD Prisma.

Notes: Panels (a.1) and (b.1) shows the frequency of price changes of CPI items affected by VAT reduction. The share of prices increases is depicted in red (—) and in blue (- - -) the share of prices decreasing. Panel (c.1) shows the frequency of those products not affected by the measure. In this case there is not an increase in the frequency of price decreases during the first week of 2023. The histograms in panels (a.2), (b.2) and (c.2) show the distribution of the nonzero price change size (dlog in%) for the three different groups of products (bins of 2.5pp). The shaded red bars correspond to the distribution of nonzero price changes during the first week of 2023. The rest of the sample is depicted in the blue bars. Back to figure **??**.
Sources: Authors' calculations based on DPD Prisma, ECB.

C 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.

C.1 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*"), 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.[21] 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.

C.2 Data Labelling

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 pro-

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

cessing (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. We use two different pre-trained Sentence Transformers embeddings, one for each country[22]. 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 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[23] 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 x 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:

$$f(x, y) = \frac{xy^\top}{\|x\|\|y\|} \tag{C.1}$$

   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 x frozen margherita Pizza extra cheese, Frozen food / Prepared dishes / Pizzas*" product description.

[22]For the Italian embedding, we use this model and the one for the Spanish case this one.

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

We assign a COICOP 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 figure C.4. This process is repeated until a sufficient number of manually tagged samples per COICOP category is obtained.

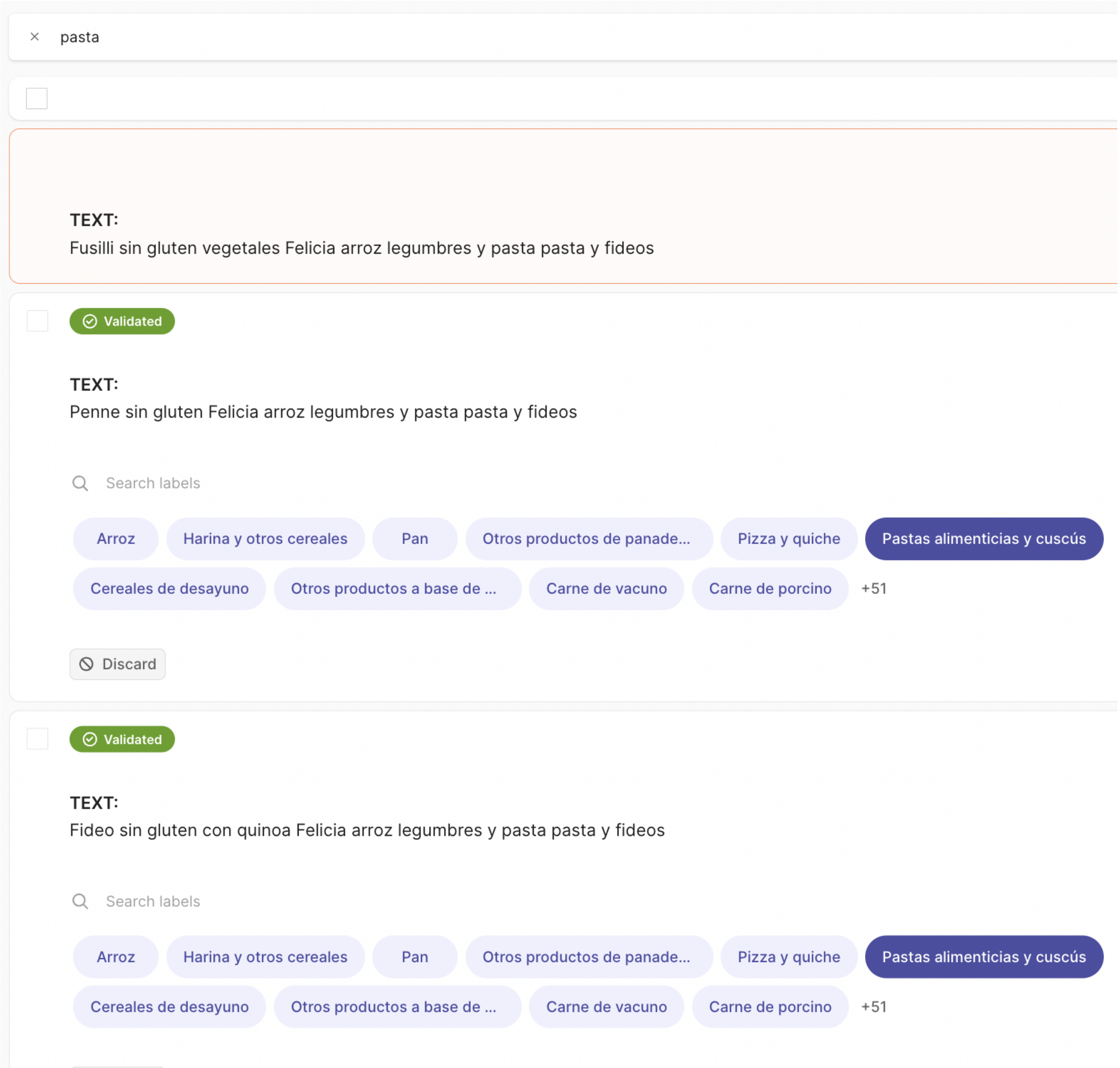


Figure C.4: Manual Bulk Labeling of similar products to "Pasta". We used the Argilla user interface and API [24]

The data set used for labeling purposes consisted of approximately 81000 Spanish and Italian product descriptions, resulting in a labeled sample of around 10% of product descriptions on average for both countries, as can be seen in Table C.3.

For Spain, we managed to have around an average of 88 product descriptions per COICOP

Table C.3: Training Dataset Descriptive Stats

	ES	IT
Number of Products	54449	27303
Estimated Non-food	35.17%	44.31%
Labeled	9.83%	11.15%
Avg. labeled prods. per COICOP	87.8	49.95

category; while for Italy, around 50 products per category. This difference can be explained by the fact that, as we will see later, our multilingual model can infer also reasonably well for other non-Spanish nor Italian product descriptions. We observed that, once the model had been fitted with Spanish data, inference with Italian data seemed to work well, so labeling with some Italian products not belonging to the Spanish market raised the posterior's inference accuracy substantially. In fact, the proposed methodology of using Sentence Transformers to enable semantic similarity search allowed us to rapidly obtain a labeled sample for both countries once we were fitting with the labeled sample and inferring for the rest of the sample. To our knowledge, no other study has used these kinds of methods to label data in such a manner.

C.3 Methods

C.3.1 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 / 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.[25]

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

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

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.

C.3.2 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 C.3. For this purpose, we use a DistilBERT multilingual model (Sanh et al. (2020)) pre-trained [26] on multiple languages and domain-adapt it to our product description data, both Spanish and Italian. 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] We use this model.

C.3.3 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 and Italian labeled subsample. The model we use is the same domain-adapted DistilBERT multilingual model from the previous step. We split our labeld 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 C.4.

Table C.4: Hyperparameter Space for Cross-Validation of DistilBERT

Hyperparameter	Values
Learning Rate	`[5e-5, 3e-5]`
Epochs	`[10, 20]`
Batch Size	`[8, 16, 32]`

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[27], 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.[28]

C.3.4 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 nor Italian product description data.

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

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

Table C.5: 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, ∞]

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 C.5. We fit the train with the best set of hyperparameters[29] 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.

C.4 Results

Looking at table C.6, 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

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

Table C.6: 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

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 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 C.7.

Table C.7: 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

Continued on next page

Table C.7: *(Continued)*

	Precision	**Recall**	**F1 Score**	**N**
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
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
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

Continued on next page

Table C.7: *(Continued)*

	Precision	**Recall**	**F1 Score**	**N**
Average	0.94	0.94	0.94	1607
Weighted Average	0.95	0.95	0.95	1607

Table C.8: Coverage of whole sample and Share in Consumption Basket

		Retailer 1		Retailer 2	
Label	**CPI %**	**N**	**Rel. Share**	**N**	**Rel. Share**
Baby food	0.008	302.0	0.027	101.0	0.023
Beef and veal	0.066	83.0	0.007	53.0	0.012
Bread	0.113	263.0	0.024	143.0	0.033
Breakfast cereals	0.009	210.0	0.019	71.0	0.017
Butter	0.005	70.0	0.006	22.0	0.005
Canned fruit and fruit products	0.004	62.0	0.006	34.0	0.008
Cheese	0.072	837.0	0.076	197.0	0.046
Chocolate	0.026	408.0	0.037	167.0	0.039
Confectionery products	0.016	275.0	0.025	214.0	0.050
Confectionery, jams and honey	0.008	169.0	0.015	52.0	0.012
Dried, salted or smoked meat	0.167	911.0	0.082	184.0	0.043
Edible offal	0.006	23.0	0.002	34.0	0.008
Eggs	0.021	28.0	0.003	15.0	0.003
Fish and shellfish	0.011	78.0	0.007	21.0	0.005
Flour and other cereals	0.004	66.0	0.006	16.0	0.004
Food pastes and couscous	0.018	248.0	0.022	92.0	0.021
Fresh or chilled fish	0.082	52.0	0.005	108.0	0.025
Fresh or chilled fruit	0.150	163.0	0.015	121.0	0.028
Fresh or chilled seafood	0.027	60.0	0.005	18.0	0.004
Frozen fish	0.018	102.0	0.009	78.0	0.018
Frozen seafood	0.013	125.0	0.011	60.0	0.014
Frozen vegetables	0.006	37.0	0.003	71.0	0.017
Ice cream	0.014	272.0	0.025	120.0	0.028
Nuts and nuts	0.031	257.0	0.023	78.0	0.018
Olive oil	0.050	108.0	0.010	13.0	0.003
Other bakery products	0.089	947.0	0.086	318.0	0.074
Other cereal products	0.007	106.0	0.010	22.0	0.005
Other dairy products	0.018	276.0	0.025	102.0	0.024
Other edible oils	0.011	58.0	0.005	13.0	0.003
Other fish and shellfish	0.063	415.0	0.037	113.0	0.026
Other foodstuffs n.e.c.	0.034	250.0	0.023	103.0	0.024
Other meat	0.007	13.0	0.001	10.0	0.002
Other meat preparations	0.036	210.0	0.019	87.0	0.020
Pigmeat	0.062	94.0	0.008	128.0	0.030
Pizza and quiche	0.016	200.0	0.018	56.0	0.013
Potato chips	0.024	281.0	0.025	76.0	0.018
Potatoes	0.022	60.0	0.005	22.0	0.005
Poultry meat	0.085	107.0	0.010	95.0	0.022
Prepared dishes	0.069	752.0	0.068	213.0	0.050
Rice	0.009	84.0	0.008	28.0	0.007
Salt, spices and culinary herbs	0.009	142.0	0.013	79.0	0.018
Sauces and condiments	0.026	336.0	0.030	120.0	0.028
Sheepmeat and goatmeat	0.016	8.0	0.001	28.0	0.007
Skimmed milk	0.034	104.0	0.009	55.0	0.013
Sugar	0.004	58.0	0.005	22.0	0.005
Vegetables, dried	0.045	526.0	0.047	180.0	0.042
Vegetables, fresh	0.103	282.0	0.025	120.0	0.028
Whole milk	0.020	41.0	0.004	24.0	0.006
Yogurt	0.041	515.0	0.047	202.0	0.047
Total	1.79	11074	1	4299	1

Notes: This table shows the whole food products sample for the DPD - PRISMA and Datamarket database, as explained in section 3. We exclude beverages. The composition of the average consumption basket in Spain can be accesed online at: `https://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica_P&cid=1254735976607`. CPI share is expressed in tens of percentage points, meaning that our coverage accounts for the 17.9% of the consumption basket.