Integration of volatile online prices into the consumer price index

Keywords: consumer price statistics, dynamic pricing, harmonised consumer price index, web scraping

1. INTRODUCTION

The online market is increasingly gaining in importance. Consumers buy more and more goods on the online market due to the great variety of product offers, time saving and independence regarding closing hours of physical shops. For the German Consumer Price Statistics, which comprises the National Consumer Price Index (CPI) and the Harmonised Index of Consumer Prices (HICP), the Federal Statistical Office (FSO) collects approximately 10,000 individual prices for products on websites of online retailers. The share of these products on the overall basket of goods and services amounts to approximately five per cent and will probably be rising in the forthcoming years.[1]

Thanks to fairly easy to adjust prices on the internet, online retailers are able to react to market conditions or consumer's behaviour by adjusting prices automatically in short intervals, applying algorithms that take into account different parameters. This phenomenon is known as dynamic pricing. First studies investigating dynamic pricing in Germany have shown that different variants of dynamic pricing exist and are very heterogeneous and not transparent [1]. Dynamic pricing of online retailers may lead to a bias in the index calculation since the traditional way of price collection via internet is done generally at one time during the month and therefore cannot capture rapidly changing prices. Therefore, in order to display reliable price developments in the CPI/HICP, consumer price statistics needs to constantly monitor the pricing behaviour on the internet and apply methods to evaluate the large amount of data and integrate very volatile price developments into price indices.

The FSO has gathered numerous experiences through former studies in the topical subject of web scraping and also conducted a study investigating the extent of dynamic pricing on the German online market. The present paper deals with the applied techniques to monitor the pricing behaviour on the internet, includes research towards handling of dynamic pricing within the price collection for CPI/HICP and gives an overview of suitable methods when calculating indices.

1.1. Definition of dynamic pricing

Dynamic pricing is the use of automatic algorithms to change prices at short notice due to changes in market conditions or due to parameters indicating a consumer's willingness to pay. The application of a dynamic pricing pattern is not new but has become more obvious since the growing importance of internet purchases and has additionally become popular for online retailers in order to attract customers or to increase profits. Dynamic pricing must be distinguished from the phenomenon of individualised pricing which is defined as the use of automatic algorithms to change prices due to characteristics of an individual consumer, such as the device used for purchase or the location of purchase. While algorithms of dynamic pricing treat all consumers equally, individual pricing only

---

[1] In 2019, new weighting patterns for the base year 2015 will be introduced in the German CPI/HICP. When the NTTS conference takes place, the updated share of online products in the overall basket of goods and services will already be published.

affects consumers with pre-defined characteristics. Only dynamic pricing is subject to the investigation in the present study.

1.2. Introduction to consumer price statistics

For the German Consumer Price Statistics, which comprises the CPI and the HICP, the FSO collects over 600 different types of goods and services. Overall, approximately 300,000 individual prices each month. While the majority of these prices are collected in physical stores, over 10,000 prices are collected on websites of online retailers. The share of these products on the overall basket of goods and services amounts to approximately five per cent and will probably be rising in the forthcoming years. The growing importance of internet prices for the calculation of the CPI/HICP underlines the importance of investigating new phenomena on the online market, such as dynamic pricing.

2. METHODS

2.1. The application of web scraping for price collections

The technique of web scraping is successfully applied by German Consumer Price Statistics for the purpose of monthly price collections in various product groups over several online retailers. With the help of a standardized price collection program, the FSO is able to collect thousands of prices in only one hour. This way, potential pricing algorithms of online retailers may be observed and evaluated.

2.2. The dynamic pricing study at the FSO

The development of web scraping for the purpose of price collections has started in 2012 and has been largely funded by Eurostat grant agreements. The first pilot study investigating dynamic pricing on the internet started in the course of 2016. Approximately 3,000 prices at 15 major online retailers were collected hourly over the course of one year and analysed. In 2018, the FSO started a new 3-year study to experiment on new methods of including large amounts of data and products affected by dynamic pricing to the CPI/HICP.

2.3. The application of standard statistical methods

Collecting prices with the help of web scraping techniques lead to various compilation issues. These issues relate to dealing with the larger volume of data, the identification of outliers, and the imputation of missing values. Several methods such as standard deviation, variance, inclusion of weights, and calculation of various average values will be applied to the collected data and evaluated.

3. RESULTS

The study started in late 2018 and the analysis will presumably cover 3 months. Therefore, only preliminary results of the application of the methods will be given.

3.1. Volatility overall

In order to get an overview of the extent of dynamic pricing on the internet, all products and online retailers will be analysed at first. Due to the results of the former study, it is

expected that prices will especially be volatile within shops and not within product groups.

3.2. Volatility by retailer

The dynamic behaviour of online prices may be dependent on the online retailer. Some online retailers perform dynamic pricing to a high extent so that the application of techniques such as web scraping is absolutely necessary for the monthly price collection for the CPI/HICP as the traditional price collection is only capable of capturing a small amount of price changes.

3.3 Volatility by product group

Although dynamic pricing is in many cases related to the online retailer, it is vital for the quality improvement of the regular price collection to identify product groups with volatile price developments. Price collectors are usually allocated according to product groups. Investigating the volatility by product groups may help sensitize price collectors more efficiently.

4. CONCLUSIONS

The phenomenon of dynamic pricing may have a visible effect on the compilation of the CPI/HICP. As it is clear that dynamic pricing is present on the internet and may affect the calculation of the CPI/HICP, very volatile products may only be included after using statistical methods or indices.

In addition to the focus on the integration of volatile prices into the German CPI/HICP, the current study will also deal with the investigation of the pricing algorithms of online retailers, the handling of the larger amount of data, as well as presenting solutions of making price collections more efficient.

Finally, new phenomena in price statistics, such as dynamic pricing, may be caught by modernising price collection techniques (e.g. web scraping) and applying new methods. These measures will secure and increase the quality of price statistics in the near future.

REFERENCES

[1] C. Blaudow and F. Burg, Dynamic Pricing as a Challenge for Consumer Price Statistics, EURONA no. 1 (2018), 79-93.