An Phríomh-Oifig Staidrimh
Central Statistics Office

Quality adjustment in the Irish CPI

Meeting of the Group of Experts on Consumer Price Indices

7-9 May 2018

Joseph Keating, Matt Murtagh

Abstract

The recent establishment of a development unit for the Irish CPI has made it possible to give closer attention to methodological issues. The development unit is currently focusing on improving the quality adjustment methods used for different items in the CPI basket. This paper describes the different quality adjustment approaches in two areas, 1) computers and printers, and 2) clothing and footwear. It describes the combination of quality adjustment approaches necessary, including direct comparison, option pricing and hedonic pricing. It describes how these approaches are being tested by a combination of retrospective analysis and parallel runs. The recommended changes will take into account the resource issues faced by smaller national statistics institutes such as that in Ireland. The paper looks ahead to how these changes should be communicated to users of the Irish CPI.

1 Introduction

The Central Statistics Office (CSO) established a dedicated development unit for the Irish Consumer Price Index (CPI) in 2016. This has enabled more attention to be given to methodological issues. One of the most important of these issues is that of quality adjustment. This paper describes the quality adjustment research being done in two areas, 1) computers and printers and 2) clothing and footwear.

The quality adjustment methods being researched in producing indices for computers and printers are *hedonic pricing* and *option pricing*. Analysis of historic data was not sufficient to test these methods. Instead, the methods are being tested by a parallel run trial. The clothing and footwear research is focussed more extensively on using the *direct comparison* quality adjustment method. In this case it is possible to use historic data for the research.

2 Computers and Printers

2.1 Project Background

All consumer price indices are based around a fundamental concept; measuring the temporal variation of the prices of a fixed set of goods. Execution of this concept, however, is complicated by goods that change in quality and characteristic rapidly over time, such as electronic goods. For these goods, a simple monthly price comparison is insufficient. Manufacturers of electronic goods consistently find more efficient ways to improve various components such as processors, rapid access memory (RAM) and storage memory. The result is a high rate of replacement of these goods at retailers over time, and a high degree of ambiguity in the balance of pure price change and quality change in the observed price changes of electronic goods. A simple comparison of these goods would violate the basic principle of comparing like with like across time. As such, National Statistics Institutes (NSIs) have developed a variety of methods across the last two decades in order to account for this problem.

2.2 Existing Quality Adjustment Methodology in the CSO

In its effort to control for quality change in electronic goods, the CSO has utilised the bridged overlap method. Bridged overlap is a simple method for adjusting for quality across time; for any electronic good that is discontinued, a replacement good is chosen, and a price change is not recorded for that good for that month. It is assumed, therefore, that the price change between observations can be imputed as the average change for the other matched price observations within that item heading (Central Statistics Office 2016:21). However, while simple to execute, a number of flaws in the bridged overlap method have become more apparent in recent years, and have contributed to indices for electronic goods that are decreasing at a rate that is higher than what might be expected.

Firstly, a price change for an electronic good may be masked by the introduction of a new model, meaning that an actual price change goes unobserved while the overall index for a given heading remains stable or moves in such a way that fails to capture actual price trends. It is reasonable to assume that, for these goods that change in quality rapidly, manufacturers are most likely to introduce price changes when introducing a new model (Triplett 2004:27). Secondly, those goods that remain available to price across time may be steadily discounted until new, higher quality stock is brought in to replace them. Where an actual price increase exists from this replacement, it is not observed. This practice may therefore bias an index downward by only counting price changes for goods that remain in stock across time, which tend to be discounted by retailers as newer quality goods become available. The graph below demonstrates these issues in the prices for digital cameras collected by one pricer over time, where items are consistently replaced with new items that are more expensive than the previous item, which may represent a price change that is not accounted for, before being discounted and replaced again. This results in an index (the dotted line) which drops at a rate that is likely higher than reality:

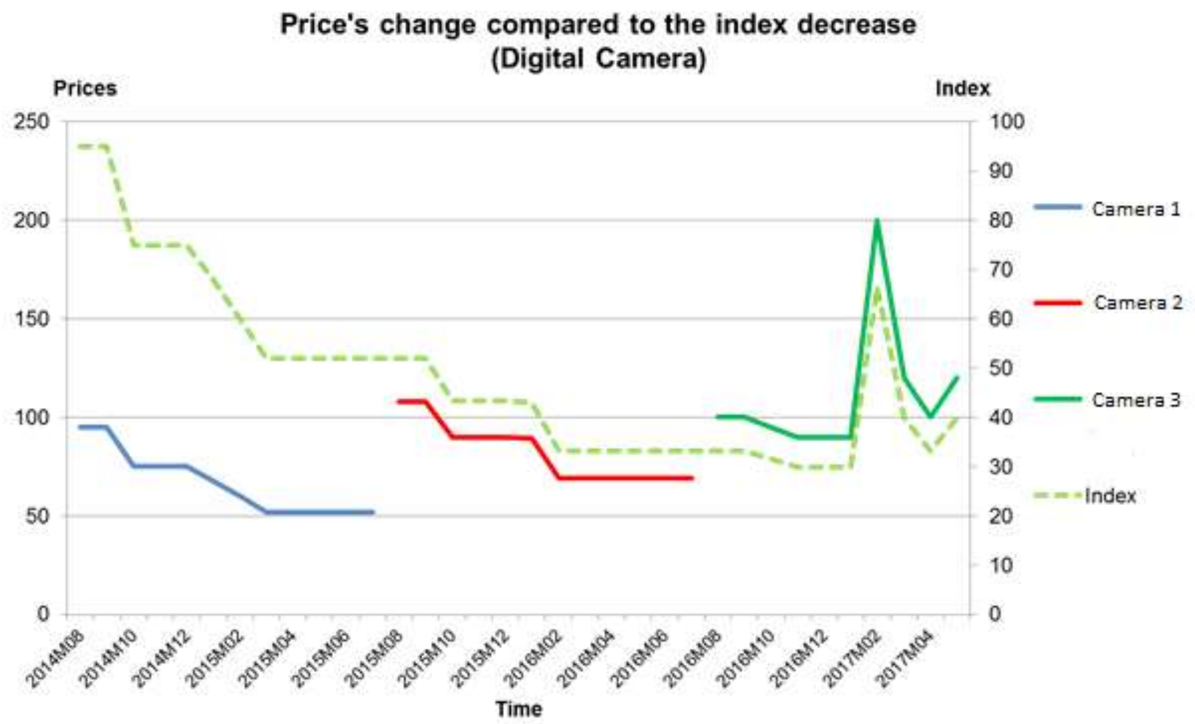


Figure 1: Replacement rate and price stability of the Digital Camera Index

While it is near impossible to judge the direction of quality adjustment bias *a priori* using the bridged overlap method, it can be shown that the direction and intensity of the bias is dependent on whether the true quality adjusted price change for the item that has been replaced is greater or less than the other price changes in a given heading used to impute the true quality change (Triplett 2004:27). Thus, given the trends in electronic goods discussed above, there exists significant evidence that the current practice of quality adjustment is a contributor to the downward bias observed in a number of indices for electronic goods in the Irish CPI. This downward trend in the indices for electronic goods, including computers and printers (as computer accessories) can be seen in the context of the ten most decreasing indices in the Irish CPI in figure 2 below:

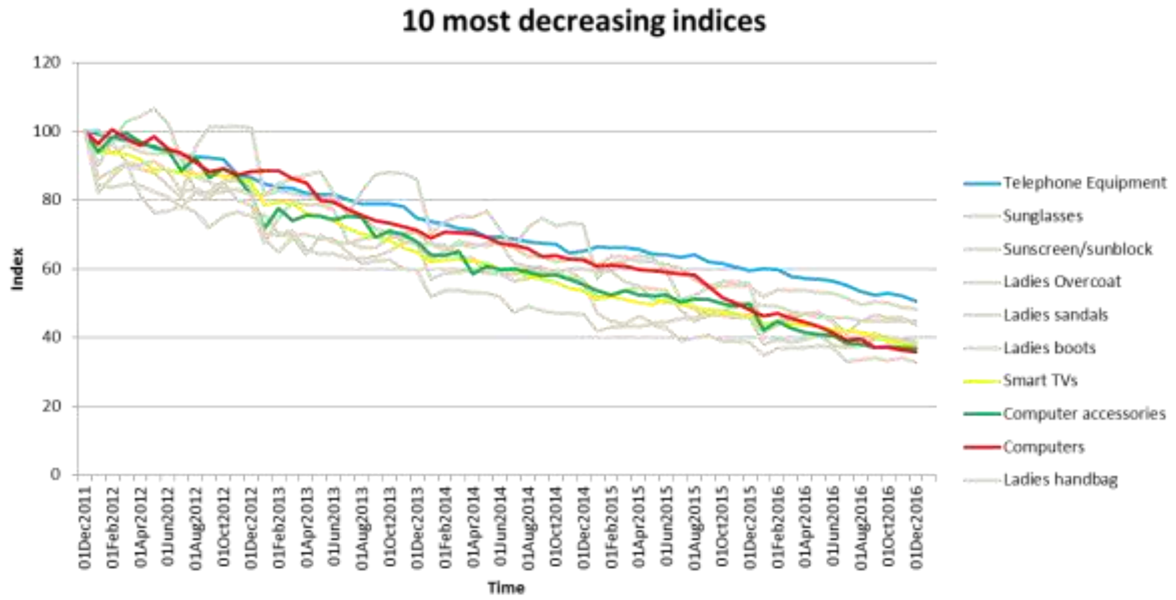


Figure 2: Most decreasing indices in the Irish CPI

2.3 Alternative Quality Adjustment Methodology

Alternative quality adjustment methodology, particularly in the area of electronic goods, has been a topic that has been explored thoroughly by NSIs in the last two decades. Two of the most prominent approaches to quality adjustment were considered for the purpose of this study: *hedonic repricing* and *option pricing.* In order to test the viability of these methods, the indices for *computers* and *computer accessories (printers)* were chosen. These indices are associated with the problems noted above and therefore serve as appropriate tests of this methodology at the CSO.

2.4 Data Collection

Both of these alternative approaches have additional data requirements that can prove difficult for smaller NSIs to resource. Additional data on a variety of product characteristics that are thought to be key contributors to price, such as how much storage or processing speed that a computer has, must be collected in order for these quality adjustment methods to be performed. While data on the prices of computers and printers are currently collected by the CSO Special Inquiries team from a variety of brick and mortar stores in three cities across Ireland, data on these product characteristics, such as processing power, are not currently collected. This unfortunately renders retrospective analysis outside the scope of this investigation, and limits the extent to which we can compare current processes at the CSO with this new methodology. Therefore, for the purpose of this trial, price and characteristic data was collected online from major chain retailers that operate across Ireland, starting from October 2017. A sample of forty computers and forty printers were chosen and priced monthly parallel to the existing indices. Additionally, a larger sample of at least one hundred computers and printers is being priced at least every six months in order to provide data of sufficient scope for the hedonic regression model.

2.5 Hedonic Repricing

In producing a hedonic method of quality adjustment for these indices, a characteristic method was adopted. Under this method, a regression model is used to estimate the contributing effect of a variety of characteristics associated with the computer or printer to price. In this way, these characteristics serve as the independent variables in the model, while the natural logarithm of the price of the computer or printer acts as the dependent variable. Using these estimates, it is possible to quantify the quality change through the difference of these characteristics between the original product and its chosen replacement. Sample output from these models can be viewed in Appendix A and B. The variables used in these models can be viewed below:

Variable	**Description**
CPU_Speed	The clock speed of the processor in gigahertz
SSD	The amount of solid state drive storage in gigabytes (If present)
HDD	The amount of hard disk drive storage in gigabytes (If present)
eMMC	The amount of embedded multimedia controller storage in gigabytes (If present)
RAM	The amount of random access memory
Real_Pixel_Density	The Pixel Density of the computers monitor in pixels per inch
Graphics_Memory	The dedicated graphics memory available to the GPU in gigabytes
Battery_Hours	The stated battery life of the computer in hours
Touch	A dummy variable indicating whether the screen is a touch screen
MacOS	A dummy variable indicating 1 for MacOS and 0 for Windows

Table 1: Variables used in Hedonic Regression model for Computers

Variable	Description
Max_Yield_Black	The amount of pages that can be printed by the largest black ink cartridge available
Resolution_Width	The amount of dots printable between the longest sides of a page
PPM_Black	The pages that can be printed per minute in monochrome
Scanner_res_Length	The dots between the shortest sides of a scanned image
Fax	Dummy variable indicating whether the printer has a Fax machine
Portable	A Dummy variable indicating whether the printer is portable
Inkjet	A dummy variable taking on a value of one if the printer is inkjet and one if laser

Table 2: Variables used in Hedonic Regression model for Printers

The selection of independent variables specifically used in the model is determined by an automated model selection procedure, which tests various parameterisations, interactions and variable specifications to provide a model that generally fits the data well (Cohen 2006). The selection of characteristics of computers and printers to gather data on as variables are based on fundamental characteristics that are commonly believed to determine computer and printer quality and are unlikely to change significantly across semi-annual model updates. Thus far, the experience of modelling the prices of desktop and laptop computers has been positive. Models tend to fit well, many variables that are thought to be significant turn out to be so and there are little to no concerns about collinearity, as can sometimes be the case with hedonic regression models (Triplett 2004).

However, the experience with printers has provided less promising results. In contrast with the results for computers, model fit tends to be much poorer, many variables we would expect to have an influence on price do not and some concerns with collinearity can be justified. This might be partly due to the way in which data used for these variables are collated and presented by manufacturers. For computers, information supplied by manufacturers tends to be self-evidently reliable – a computer with eight gigabytes of RAM will most likely be more expensive and better quality than one with four. Statistics used for the performance of printers however, such as Pages Printed per Minute (PPM), are poorer indicators of performance, as manufacturers often devise their own tests to develop this measure that are not standard across the industry. Furthermore, many manufacturers follow a business model whereby printers are sold cheaper than or close to what they cost to produce, with profits being made on the sales of ink cartridges. This practice may distort the relationship between the price of the printer and its performance or components. These issues might have been expected: in a survey of ten NSIs from around the world, the ONS found that none are currently using this methodology to quality adjust for printers (Wells & Restieaux 2013: 3).

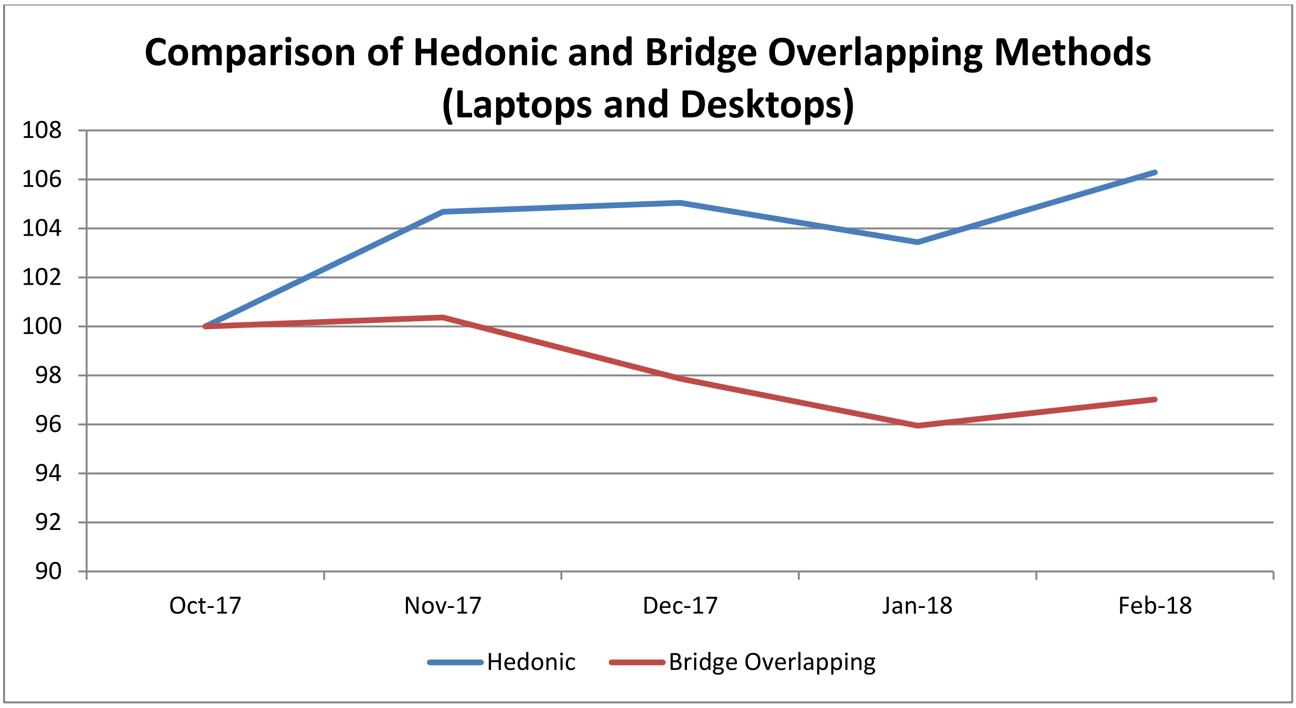


Figure 3: Comparison of Hedonic and existing Bridge Overlapping CPI Indices

	Nov-17	Dec-17	Jan-18	Feb-18
Replacements	9	11	2	7
Percent of Sample	22.5%	27.5%	5.0%	17.5%

Table 3: Replacement rate for computer index

The graph above displays the results that have been collected in this study thus far for computers and printers. Before commenting on any collected data, a caveat must be noted; with retrospective analysis in this area impractical, there are too few price comparisons to make firm conclusions on the viability of this methodology. Rather, what is inferred from this data should be seen as preliminary results pending the availability of further data produced by this methodology later in the lifetime of this study.

That being said, the data produced thus far indicates a trend whereby the index produced by the hedonic repricing methodology is consistently higher than that of the existing bridged overlap methodology, with the latter displaying a characteristic dip in the index shortly after it begins, consistent with the trends observed in the current index which provided the impetus for this investigation. Furthermore, a greater disparity in trends between the hedonic and bridged overlap indices is noted for those months where a greater amount of computers were no longer available and had to be replaced. Considering that, as discussed earlier, evidence exists where the effect of price changes are being lost on the introduction of a new model of computer or printer, a disparity such as this suggests that these price increases may be captured by the hedonic methodology where they are not captured by the existing bridged overlap methodology.

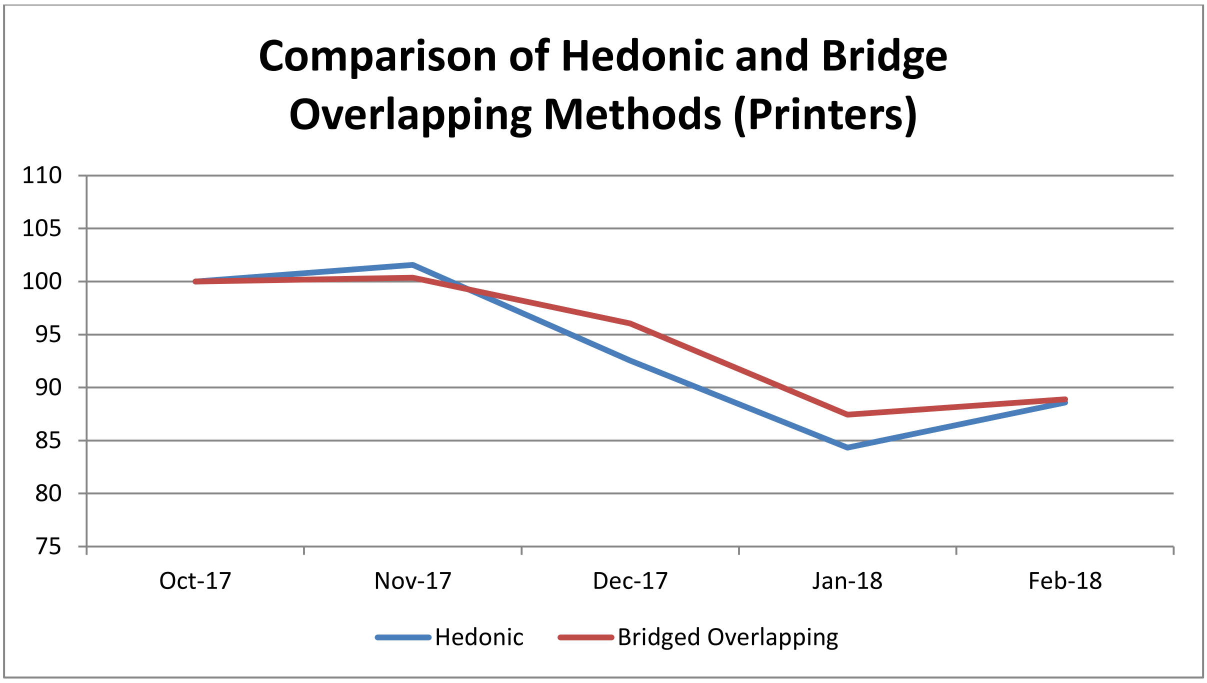


Figure 4: Comparison of Hedonic and existing Bridge Overlapping CPI Indices for Printers

The results collected thus far for printers have shown a less remarkable difference between hedonic and bridged overlap methods, with data produced by both methods largely following the same pattern of price changes. This may be due to the issues with modelling for printers, as outlined above, which could mean that the estimates gathered by the model are inaccurate for the purpose of quality adjustment. However, further comparison across indices for this item is required to understand the root cause of this result; an alternative explanation could be that bridged overlap already satisfactorily captures the quality adjustment process of printers, and what appears to be a negative bias is in fact simply a consistent reduction in the price of printers. A clearer indication of the underlying trends will hopefully become available once more data points are produced and a cross-national comparison of this item can be undertaken.

2.6 Option Pricing

Option pricing was further examined as a possible solution to quality adjustment. Similar to the characteristic hedonic method, the option pricing method estimates the contribution to the price of goods by specific characteristics associated with that good. However, instead of estimating these contributions using regression analysis, the characteristics are simply priced either from the additional cost of the same computer with an additional component, such as higher RAM, or on what these components cost to be purchased separately. The former method of referencing proved difficult to achieve in the modern computer market; most computers differ in many different components rather than a single component across models. Thus, for components such as processors, RAM and Hard Drive Disks, price data on what each of these components cost separately were collected. As with hedonic repricing, the quantifiable difference in these components between the original and replacement product, subject to a discount factor of 0.5, was used to estimate quality change.

While this method is certainly simpler in execution than the hedonic repricing method, stripping away the time intensive need to update the model with new data, the practicality of option pricing has faltered in recent years. This is primarily due to trends in consumer electronics towards devices that are less user-modifiable and contain more unreplaceable, proprietary components that are either difficult or impossible to price using option pricing. This has always been a problem for laptops and printers; components such as laptop screens have rarely, if ever, been purchasable by users. For desktop computers, separate components can usually be bought and therefore easily priced. Yet, even for desktop computers, manufacturers have been moving towards models with more proprietary parts, such as all-in-one computers, where separate component prices may be increasingly unrepresentative of the actual cost of that particular component. Some components may even be present in cheaper models and simply disenabled in order to introduce a price tier to the product, such as processors being limited to a certain clock speed.

Regardless of these concerns, this study included option pricing for desktop computers as a possible way of introducing an improved quality adjustment measure over bridged overlap that was less resource intensive than a method based wholly on hedonic repricing. Our option pricing method therefore included option pricing only for desktop computers; hedonic repricing was used for laptop computers. The reference prices for components, where required, were gathered from retailers who specialise in providing computer components for those wishing to build their own desktop computer, and were gathered from the internet. The graph below displays the outcome of this hybrid option pricing method against the fully hedonic method elucidated above and the existing bridged overlap method.

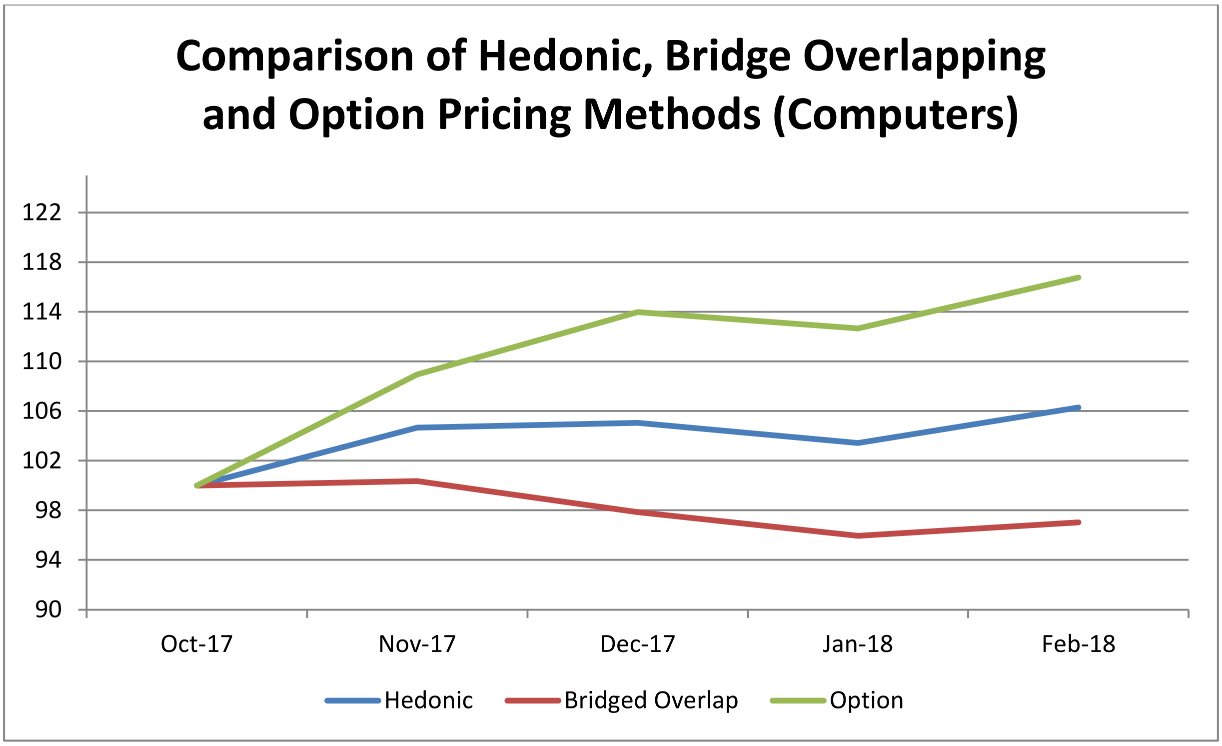


Figure 5: Comparison of Hedonic, Bridged Overlap and Hybrid Option Pricing CPI Indices for Computers

With the same reservations made about the limited data generated by this study thus far, we can see a clear positive bias in the option method compared to that of the hedonic and bridged overlap methods used in this study, consistent with past experience of this methodology (Triplett 2004:31). There may be a number of reasons for this. Option pricing may positively bias quality adjustment in that it considers a generic consumer or producer cost of components, not the actual cost to specific manufacturers for producing components. Companies may have their own methodology of producing components that is cheaper or more efficient than the reference prices we can produce via option pricing. Furthermore, when an additional option is adopted and "made standard", the cost of including that option between replacements has usually fallen to a point where this is acceptable to the manufacturer, meaning this additional option is overpriced in our adjustment and therefore quality change is over-adjusted for (Triplett 2004:31).

2.7 Viability of Quality Adjustment Methodology in Practice

Introducing and adapting new quality adjustment methodology to existing processes in NSIs is often as large a task as researching which methodology would best produce accurate output. While the evidence gathered by this study, in addition to the experience of other NSIs, thus far suggests that hedonic repricing may produce a stable output for computers, there has been gathering criticism in recent years over the viability of this methodology due to its resource intensiveness (Wells & Restieaux 2013:2). Of particular note in this regard is the requirement to update the model with new data on prices and item characteristics, which is often performed quarterly (Wells & Restieaux 2013:3), in addition to the monthly updating of prices and characteristics in order to quality adjust and produce an index. Adopting processes such as these therefore provide significant challenges to smaller NSIs, where resources are already limited.

At present, the collection of prices for computers and printers are under the responsibility of the special inquiries team of the CPI section at CSO, which collects the prices of 137 items that tend to need special attention due to their incompatibility with direct pricing. In most cases, special methodology is applied within spreadsheets in order to produce output. The current bridged overlap method is applied in this way; prices are gathered on a spreadsheet and those computers which are no longer available are excluded from calculation, with their replacement being chosen and priced in the next month. With this existing process already being work intensive, continuing in this format was seen as unviable for the purpose of the new methods of quality adjustment, where regression output or reference prices would have to be continually updated.

Instead, a repricing program was developed in SAS in order to minimise the additional workload associated with the new methodology. The function of this program is to take the updated list of prices, perform the hedonic regression to estimate component costs, take in the current and previous months prices, identify where models do not match between months, perform the quality adjustment procedure on these models using the output from the hedonic regression, calculate the regular and quality adjusted relatives for each model and produce the output for the index. For the purpose of this study, the program additionally calculates the output under the bridged overlap and hybrid option pricing methods outlined above for comparison. The result of this automation is that those working in Special Inquiries would only have to check which models are no longer available for a given month or where prices have changed, inputting the new price and finding a replacement computer where a replacement is necessary.

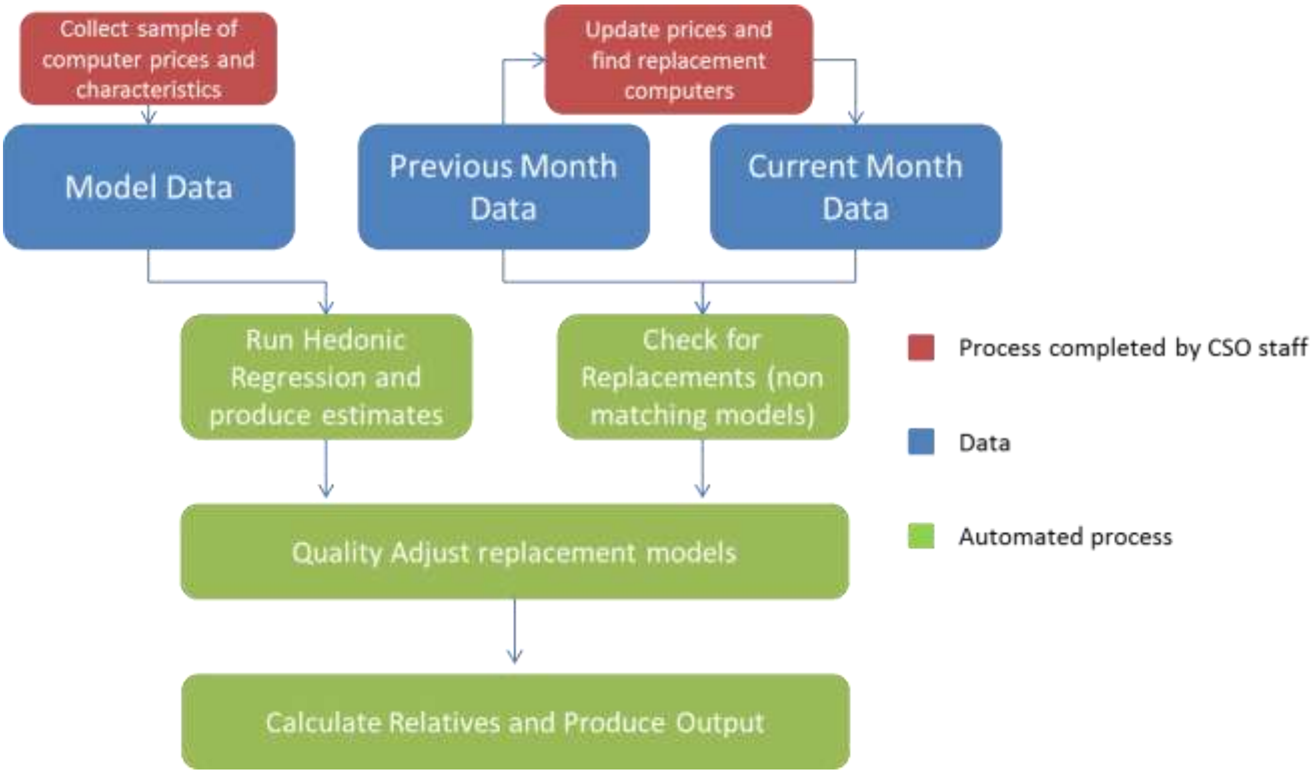
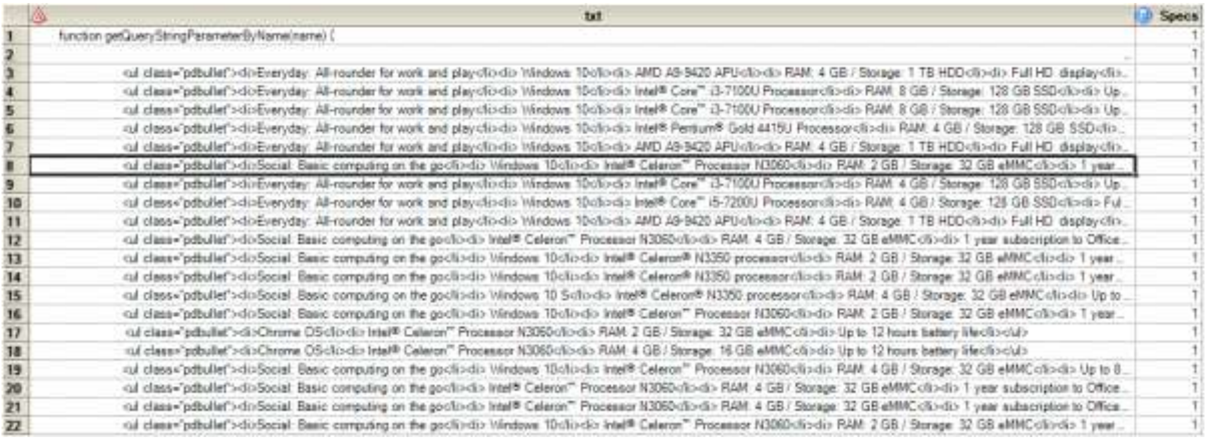


Figure 8: Quality adjustment process currently being trialled at CSO

2.8 A note on Web Scraping

Of all of the tasks in this process, the collection of a sample of prices and characteristics is the most resource intensive. This involves manually collecting price information and characteristics relating to the variables used in the model from listings of computers available online. This requires a sample size large enough to run the hedonic regression, which takes a considerable amount of time to collect. Given the linearity and repetitive nature of the presentation of this data on online retailers, and given the resource limitations of the CSO compared to NSIs of a larger scale, using such tools in this capacity may be a viable way of significantly reducing the resource intensiveness of introducing hedonic repricing. The next step in this project is therefore to investigate the potential use of this method. Investigation is currently underway into the extraction, cleaning and utilisation of this data.



	txt	Specs
1	function getQueryStringParameterByName(name) {	1
2		1
3	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - AMD A9-9420 APU - RAM: 4 GB / Storage: 1 TB HDD - Full HD display ...	1
4	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - Intel® Core™ i3-7100U Processor - RAM: 8 GB / Storage: 128 GB SSD - Up...	1
5	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - Intel® Core™ i3-7100U Processor - RAM: 8 GB / Storage: 128 GB SSD - Up...	1
6	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - Intel® Pentium® Gold 4415U Processor - RAM: 4 GB / Storage: 128 GB SSD ...	1
7	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - AMD A9-9420 APU - RAM: 4 GB / Storage: 1 TB HDD - Full HD display ...	1
8	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 - Intel® Celeron™ Processor N3060 - RAM: 2 GB / Storage: 32 GB eMMC 1 year...	1
9	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - Intel® Core™ i3-7100U Processor - RAM: 4 GB / Storage: 128 GB SSD - Up...	1
10	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - Intel® Core™ i5-7200U Processor - RAM: 4 GB / Storage: 128 GB SSD - Ful...	1
11	<ul class="pdbullet"> - Everyday: All-rounder for work and play - Windows 10 - AMD A9-9420 APU - RAM: 4 GB / Storage: 1 TB HDD - Full HD display ...	1
12	<ul class="pdbullet"> - Social: Basic computing on the go - Intel® Celeron™ Processor N3060 - RAM: 4 GB / Storage: 32 GB eMMC 1 year subscription to Office...	1
13	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 - Intel® Celeron® N3350 processor - RAM: 2 GB / Storage: 32 GB eMMC 1 year...	1
14	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 - Intel® Celeron® N3350 processor - RAM: 2 GB / Storage: 32 GB eMMC 1 year...	1
15	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 S - Intel® Celeron® N3350 processor - RAM: 4 GB / Storage: 32 GB eMMC - Up to...	1
16	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 - Intel® Celeron™ Processor N3060 - RAM: 2 GB / Storage: 32 GB eMMC 1 year...	1
17	<ul class="pdbullet"> - Chrome OS - Intel® Celeron™ Processor N3060 - RAM: 2 GB / Storage: 32 GB eMMC - Up to 12 hours battery life	1
18	<ul class="pdbullet"> - Chrome OS - Intel® Celeron™ Processor N3060 - RAM: 4 GB / Storage: 16 GB eMMC - Up to 12 hours battery life	1
19	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 - Intel® Celeron™ Processor N3060 - RAM: 4 GB / Storage: 32 GB eMMC - Up to 8...	1
20	<ul class="pdbullet"> - Social: Basic computing on the go - Intel® Celeron™ Processor N3060 - RAM: 4 GB / Storage: 32 GB eMMC 1 year subscription to Office...	1
21	<ul class="pdbullet"> - Social: Basic computing on the go - Intel® Celeron™ Processor N3060 - RAM: 4 GB / Storage: 32 GB eMMC 1 year subscription to Office...	1
22	<ul class="pdbullet"> - Social: Basic computing on the go - Windows 10 - Intel® Celeron™ Processor N3060 - RAM: 2 GB / Storage: 32 GB eMMC 1 year...	1

Figure 6: Raw data extracted from the website of one of the retailers used indicating computer specifications and price

3.1 Project Background

There are 68 items in COICOP 03 (Clothing and Footwear) in the Irish CPI basket. These items are all priced each month by 75 pricers. A characteristic of the Irish CPI is that these items have broad descriptions (e.g. Ladies' Dress, Men's Boots) and the individual pricers are given freedom within this description to choose a specific representative product. The shop in which they choose the representative product is assigned centrally. The pricer prices this specific product each month until it becomes unavailable. When the specific product becomes unavailable the pricer chooses a replacement representative product for the item.

The pricer codes the replacement representative product as Comparable or Non-Comparable to the previous representative product of the previous month. This choice can be overruled centrally in the CSO. When a replacement is coded as comparable the price change from the old representative product will be used in calculating the elementary index for that item. This results in the quality adjustment method known as *direct comparison.* When a replacement is coded as non-comparable the price change from the old representative product will **not** be used in calculating the elementary index for that item. This results in the quality adjustment method known as *bridged overlap.*

The CSO has historically favoured Non-Comparability over Comparability when replacements are chosen. In other words, bridged overlap is used much more often than direct comparison. The criteria for deciding whether a replacement is comparable to the previous representative product are somewhat less strict for clothing and footwear, but even here bridged overlap is used more often.

	Comparable replacement	Non-comparable replacement	Same representative product as previous month
All items	2%	7%	91%
Clothing and Footwear	9%	14%	77%

Table 4: Replacement in the Irish CPI for April 2017

There is evidence that, as with electronic goods, this is leading to a downward bias in our clothing and footwear index. In the following graph each of the 68 items is represented by a point on a scatter plot. The y-axis is the index growth for that item between September 2013 and December 2015. The x-axis shows the total number of non-comparable replacements for that item in this time period. The graph shows that items which are frequently replaced by representative products coded as non-comparable tend to have lower prices indices than those that are not. These items are typically those in which fashion is more important.

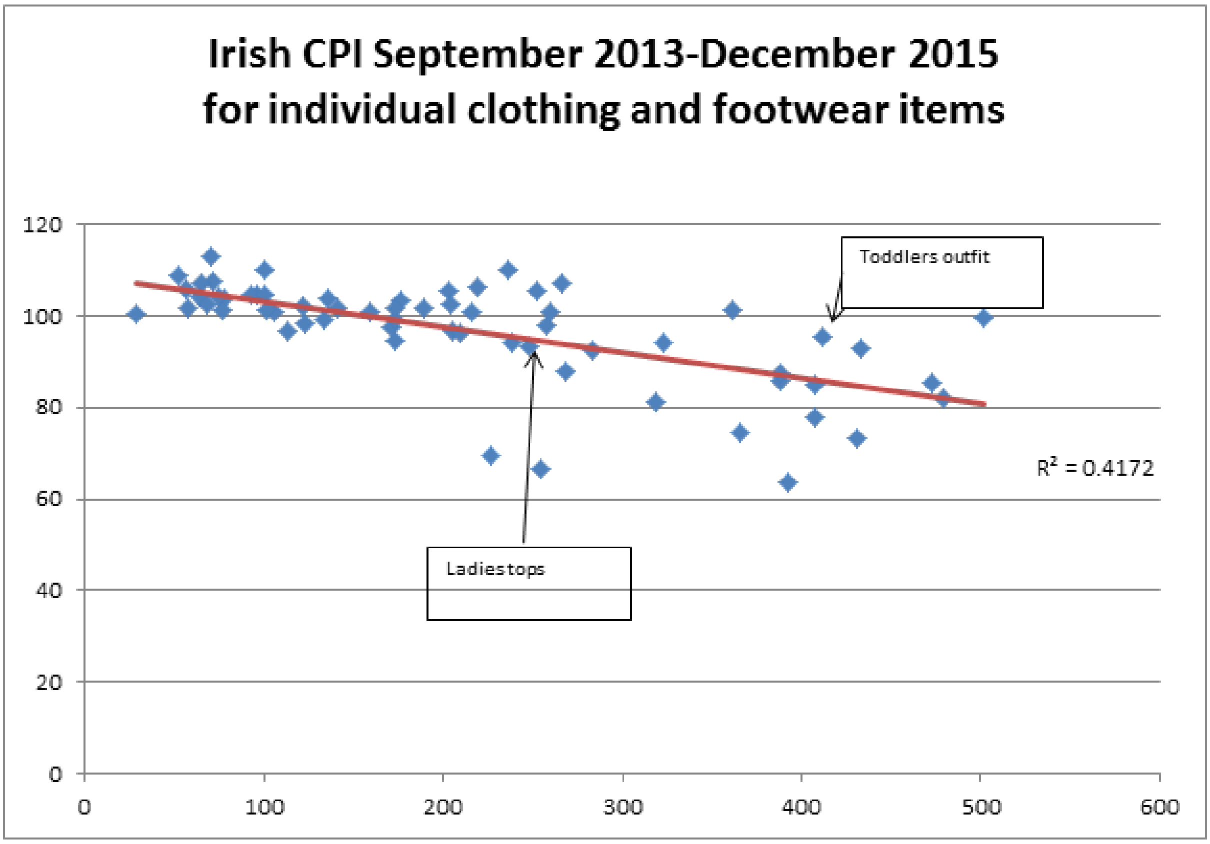


Figure 7: Irish CPI for individual clothing and footwear items from September 2013 to December 2015

The mechanism by which this happens also seems clear. Representative products for fashion items typically enter the market at their highest price. After a few months they become less fashionable and are put on sale. They then become unavailable and are replaced by a newly fashionable representative product, also entering the market at its highest price. If we decide this representative product is non-comparable to the previous one, we will miss any price increase. The problem is similar to that described in section 2.2 for digital cameras. In the case of digital cameras and other electronic goods, changes in technology drive this pattern. In the case of clothing and footwear the driver is changes in fashion.

We can also compare our clothing and footwear indices with those of other countries. The following graph shows the Harmonized Index of Consumer Prices (HICP) for clothing and footwear for various countries from January 2015 to December 2017 (Base year 2015). Nearly all countries have a pattern of 2 sales months every year. The Irish HICP shows that the recovery following the sales month tends to be incomplete, resulting in a downward index over the period. While a few other countries have a similar pattern, the most common pattern is a flat index, with some countries showing an increase in clothing and footwear prices over the period.

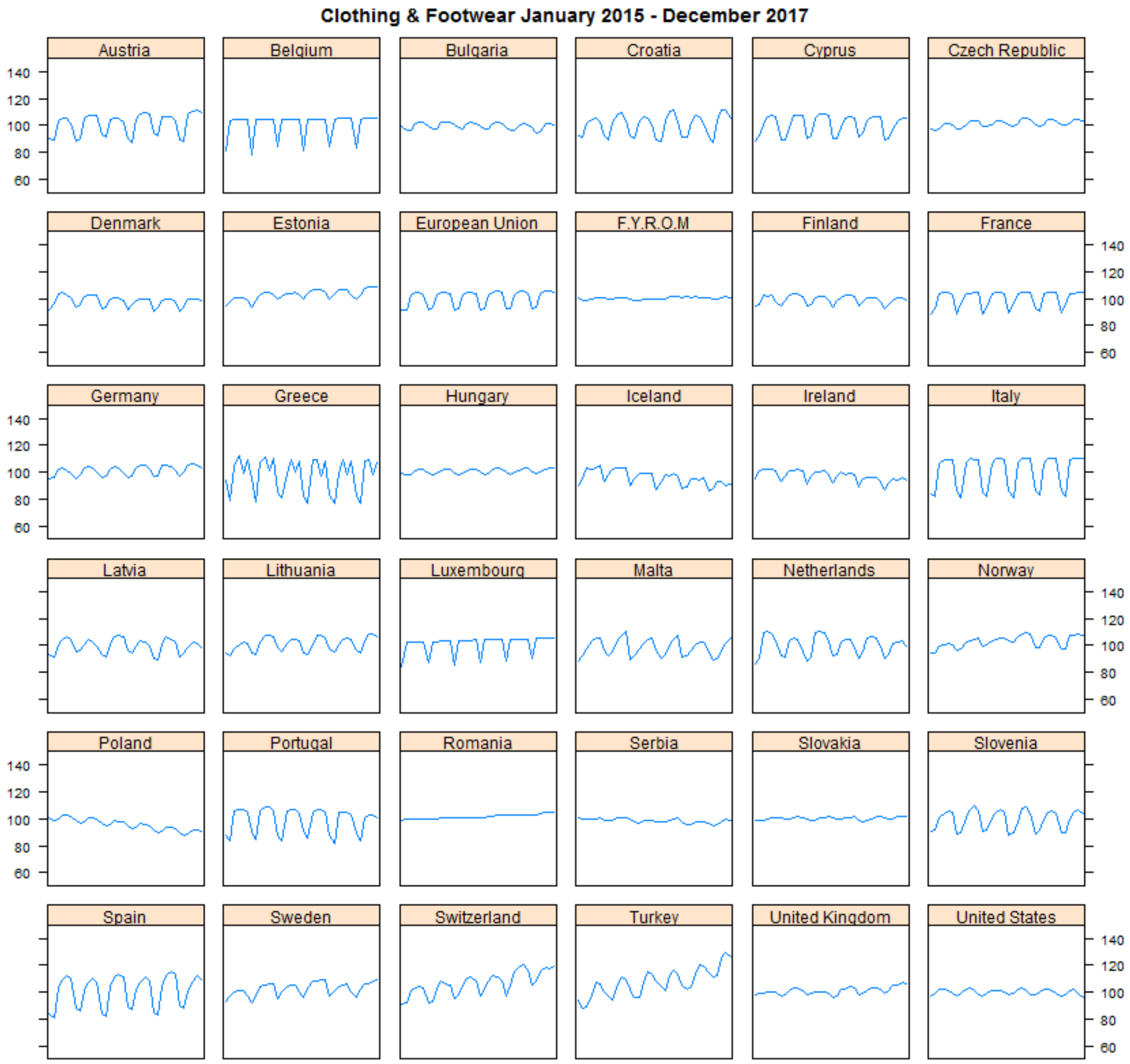


Figure 8: Comparison of trends in the Index for Clothing and Footwear across NSIs

3.2 Alternative Quality Adjustment Methodology

We will take various approaches to improving our clothing and footwear indices (e.g. clearer instructions to pricers when choosing replacements, narrower product descriptions, seasonal items) but in this paper we concentrate on the issue of deciding whether a replacement is comparable or non-comparable. It seems clear that we need to broaden our criteria for deciding a replacement is comparable. The challenge is to set these boundaries clearly, both to pricers and to the CSO staff who monitor and can overrule the pricers' decisions. Our approach is to simulate different criteria on our historic data and examine the resulting indices.

The simplest approach would be to let all replacements be considered comparable. The graph below shows that the result of this would have led to a 4% increase in price from April 2016 to February 2018 in the elementary index for ladies' dresses. The actual index for ladies' dresses had a 17% decrease during this period.

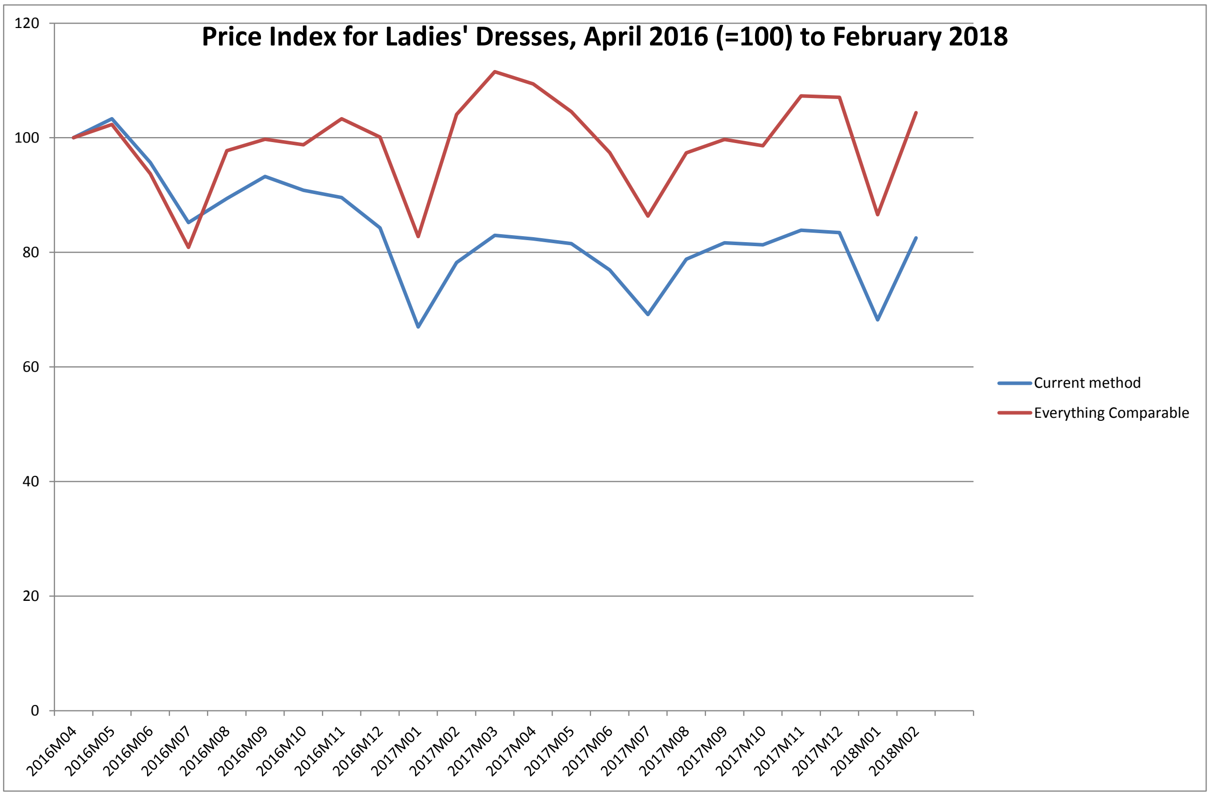


Figure 9: Price Index for Ladies' Dresses from April 2016 to February 2018

The effect is less dramatic for an item less subject to changes in fashion. Letting all replacements be considered comparable for men's vests and white cotton t-shirts results in a 4% increase in price from April 2016 to February 2018. The actual elementary index for men's vests and white cotton t-shirts had a 1% decrease during this period.

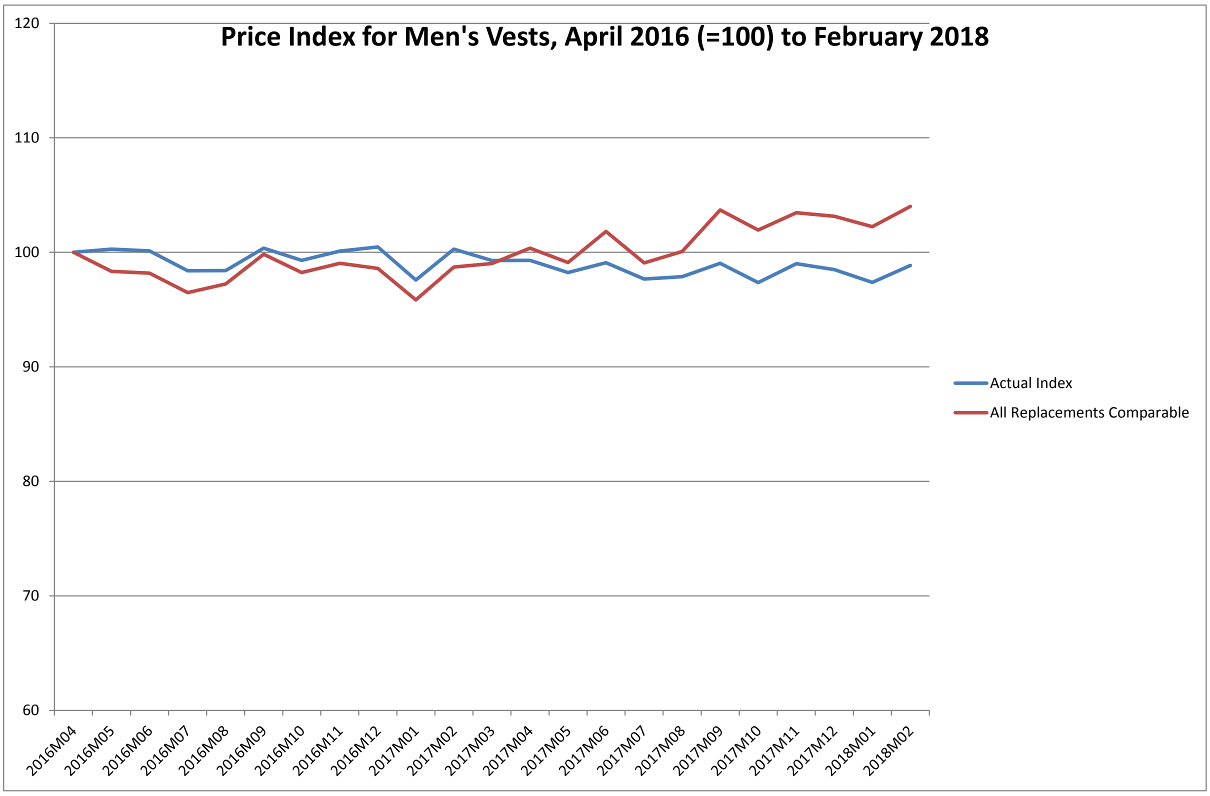


Figure 10: Price Index for Men's Vests from April 2016 to February 2018

The Office of National Statistics (ONS) in Britain has made the vast majority of their clothing replacements comparable since 2010. In 2013 only 1.7% of clothing replacements were considered non-comparable. (Johnson 2015:206). However, clothing and footwear items in the Irish index have broader descriptions than their ONS equivalents. This may mean that there is more risk in the Irish case that pricers can choose replacements that are truly non-comparable and should be identified as such.

We therefore need to be more precise about what will constitute a comparable clothing and footwear replacement, particularly for fashion items. We will do a range of simulations to gain insight into what are the most appropriate rules for deciding whether a replacement is comparable or non-comparable. We will also be able to give our pricers clear instructions on how they should best choose replacements.

It is clear that the approach taken to the question of comparability has a strong effect on the clothing and footwear index, with different approaches resulting in decreasing, flat, or increasing indices. We should therefore look to benchmark the index to other data. The Household Budget Survey (HBS) for Ireland estimated that from the 2009/10 survey to the 2015/16 survey the average weekly household expenditure on clothing and footwear decreased by 16%. The clothing and footwear price index decreased by 14% over the same period. This implies the volume of clothing and footwear bought per household in 2015/16 was 2% less than in 2009/10.

	2009/10	2015/16	Change
CPI (Base Dec 2016=100)	118.4	101.8	-14%
HBS spending (€)	40.11	33.65	-16%

Table 5: CPI and HBS weekly household spending for Clothing and Footwear, 2009/10 and 2015/16

The HBS data suggests that the downward trend of the clothing and footwear index could be correct. Comparing our index with data collected for the Irish Purchasing Power Parities survey, on the other hand, suggests that this downward trend could be incorrect.

The evidence does not point all in one direction, but the evidence suggesting that the index is behaving in implausible ways seems strong. Behaviour that seems implausible includes:

- Sale recovery consistently setting index lower than it was in month before sale
- Index falling faster for fashion items than for non-fashion items
- Index falling faster for ladies' clothing and footwear than for men's or children's.

Our aim is to find a clear set of rules which pricers and CSO staff can implement, and that result in an index that behaves in a logically consistent manner.

4 Conclusion

The establishment of a dedicated CPI development unit has allowed us to focus more closely on specific methodological issues, of which quality adjustment is one of the most important. It was vital to not simply add extra resources to the CPI but to separate the functions of production and development. Otherwise, the most immediate issues would still tend to dominate available resources. The dedicated development unit (currently the two authors of this paper) is able to step back from the most immediate issues and take a longer term view of the index, improving methodology and developing new products. Our intention is to expand the unit to become a development unit for all price surveys in the CSO. The unit will work with academia and the CSO methodology unit, as well as coordinating closely with the production units for price surveys.

We aim to implement the improved quality adjustment methods for computers and printers and for clothing and footwear in the index from 2019 onwards. A further challenge will be to explain the changes we are making to our users. We will outline the changes on the CSO website, on a new page dedicated to methodological changes to the Irish CPI (Available at: http://www.cso.ie/en/methods/prices/consumerpriceindex/).

We will also communicate directly with our users. We need to strike a balance between 1) making it clear to those using the data that there has been an important methodological change and 2) retaining user confidence in the overall index.

The development unit will improve the consistency and coherence of the various price indices produced by the CSO. It will develop new products that respond to users' needs. It will take advantage of new opportunities such as web-scraping. But most importantly, it will enable us to give methodological issues such as quality adjustment the attention they need.

The authors wish to thank Nicolas Laloue and Loreline Court of Ensai who began this analysis of quality adjustment in the Irish CPI, and Vytas Vaiciulis of CSO for assistance on the viability of web-scraping.

References

Central Statistics Office (2016). *Consumer Price Index - Introduction of Updated Series*. Cork: Central Statistics Office. Available at: http://www.cso.ie/en/media/csoie/methods/consumerpriceindex/CPI_-_introduction_to_series_2016.pdf [Accessed 4 Mar. 2018].

Cohen, R. A. (2006). Introducing the GLMSELECT procedure for model selection. In *Proceedings of the Thirty-First Annual SAS Users Group International Conference*.

Johnson, P. (2015). *UK Consumer Price Statistics: A Review*. Newport: Office for National Statistics. Available at: https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/12/images-ukconsumerpricestatisticsarevie_tcm97-44345.pdf [Accessed 1 Apr. 2018].

Triplett, J. (2004), "Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products", *OECD Science, Technology and Industry Working Papers*, No. 2004/09, OECD Publishing, Paris, http://dx.doi.org/10.1787/643587187107.

Wells, J. and Restieaux, A. (2014). *Review of Hedonic Quality Adjustment in UK Consumer Price Statistics and Internationally*. Newport: Office for National Statistics. Available at: https://www.ons.gov.uk/ons/guide-method/user-guidance/prices/cpi-and-rpi/review-of-hedonic-quality-adjustment-in-uk-consumer-price-statistics-and-internationally.pdf [Accessed 1 Apr. 2018].

Appendix

A. Sample Regression Output for Computers

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	4.856504794	0.17552972	27.67	<.0001
CPU_Speed	0.263010623	0.06043003	4.35	<.0001
SSD	0.001021391	0.00025869	3.95	0.0001
HDD	0.000144987	0.00007435	1.95	0.0537
eMMC	-0.003279647	0.00196684	-1.67	0.0983
RAM	0.055285319	0.01253935	4.41	<.0001
Real_Pixel_Density	0.002426720	0.00060592	4.01	0.0001
Graphics_Memory	0.054893516	0.02335620	2.35	0.0206
Battery_Hours	0.024016830	0.01066715	2.25	0.0264
Touch	0.186977922	0.08595860	2.18	0.0318
MacOS	0.497308522	0.09943387	5.00	<.0001

B. Sample Regression Output for Computer Accessories (Printers)

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	4.106549000	0.16810951	24.43	<.0001
Max_Yield_Black	0.000189501	0.00002782	6.81	<.0001
Resolution_Width	0.000190906	0.00011384	1.68	0.0973
PPM_Black	0.009624295	0.00575279	1.67	0.0981
Scanner_res_Length	0.000106814	0.00004341	2.46	0.0159
Fax	0.475444140	0.10905099	4.36	<.0001
Portable	1.473236177	0.32136912	4.58	<.0001
Inkjet	-0.345928018	0.15237925	-2.27	0.0258