New evidence on elementary index bias

Enikő Gábor† and Philip Vermeulen‡

†Central European University; Research Center Deutsche Bundesbank

‡Directorate General Research Monetary Policy Research Division European Central Bank

This version: December 05, 2014

Abstract

We provide evidence on the effect of elementary index choice on inflation measurement. Using scanner data for 15844 individual items from 42 product categories and 10 euro area countries, we compute product category level elementary price indexes using nine different elementary index formulas. Measured inflation outcomes of the different index formulas are compared with the Fisher Ideal index to quantify elementary index bias. Across product categories, mean levels of annual elementary index bias vary between -0.53 percentage points and 0.55 percentage points depending on the index, while the standard deviation is larger than 1 percentage point. National indexes based on aggregation of the elementary indexes remain biased. The average effect of elementary index bias on national inflation ranges from -0.45 to 0.45 percentage points depending on the index. The results show that elementary index bias is quantitatively more important than upper level substitution bias.

Key words: HICP, inflation measurement bias, elementary index, lower level substitution bias

JEL:E31, C43

We benefited from extensive discussions with Martin Eiglsperger. Thanks also to Aidan Meyler for advice. We received helpful comments from Jan de Haan. We also thank Jarko Pasanen who discussed a much earlier version of this paper. We also benefited from discussions with participants in the Eurosystem project on retail price analysis and participants at the Scanner Data Workshop 2014 at Statistics Austria. The views expressed in this paper are only those of the authors and should not be reported as presenting the views of the European Central Bank, Deutsche Bundesbank or the Eurosystem. Any remaining errors are solely to be attributed to the authors. Corresponding author: Philip Vermeulen: philip.vermeulen@ecb.europa.eu

Non-technical Summary

Inflation is one of the most important macroeconomic variables policymakers are interested in and is the subject of a vast economics literature. Much of this literature however simply uses the official inflation numbers in some type of analysis and is not concerned with how these inflation numbers are constructed. Inflation numbers are calculated through the aid of index numbers, i.e. by the aggregation of thousands of price observations. In other words, inflation is defined as the measured change of a price index number. This paper presents findings on how measured inflation outcomes differ when various index numbers are constructed.

Official consumer price indexes are usually constructed in two broad steps. First, for narrowly defined, relatively homogeneous products, also known as elementary aggregates, elementary price indexes are calculated. In a second step, these elementary price indexes are aggregated into a single consumer price index using expenditure weights. Elementary price indexes are therefore the building blocks of price index numbers. Their development over time measures the inflation of narrow product categories.

Official practices in elementary price index construction are still not uniform across countries (that is to say, different formulas can be used to aggregate the prices into one elementary price index), warranting further investigation in the consequences of different choices. Economic theory however suggests that combining price and quantity information of the items in superlative indexes is preferable if one wants to measure cost of living (which, it has to be mentioned, official price indexes not necessarily aim for). Due to the absence of expenditure information within the elementary aggregates, it is often the case that elementary price indexes are constructed using price information only.

Scanner data from retail stores, containing both prices and quantities, allows researchers to calculate how different index formulas at the elementary level perform. Using scanner data for 15844 individual items from 42 product categories and 10 euro area countries, this paper computes product category level elementary price indexes using nine different index formulas (Carli, Dutot, Jevons, Laspeyres, Paasche, Fisher Ideal, Lowe, Geometric Lowe and expenditure weighted Jevons). Measured inflation outcomes of the different elementary price index formulas are compared with the Fisher Ideal index. This difference is called elementary index bias as it measures the distance from a superlative index that, under certain theoretical conditions, measures cost of living.

The main findings are that across product categories, mean levels of annual elementary index bias vary between -0.53 percentage points and 0.55 percentage points depending on the index. The bias varies across product categories with a standard deviation larger than 1 percentage point for all of the indexes. A shift towards the use of cost of living indexes at the product category level would therefore have non-negligable effects on measured product

category inflation. Aggregation shows that it is often the case that the aggregate elementary index bias is larger than the upper level substitution bias. Although much attention has been given to consumers substituting between different product categories, it seems that more attention is needed for within product category substitution.

1 Introduction

Consumer price inflation measurement matters. In 1996, the Boskin commission pondered that even a small upward bias of the US CPI entailed billions of dollars of extra inflation-indexed social security outlays for the US government budget each year. Consumer price inflation is arguably one of the most important macro-economic variables, touching the life of each individual, whether it is trough inflation indexed social outlays, indexation of wages and prices, or monetary policy decisions. The simple fact that inflation numbers are *constructed* through price indexes, and not simply *observed*, combined with their enormous importance for economic decision making, implies that an investigation in how index formulas differ in their inflation assessment is warranted. The main contribution of this paper is an analysis of the sensitivity of measured inflation as a function of the index formula used to aggregate prices at the lowest level, i.e. the formula used to construct so called elementary price indexes.

Official consumer price indexes are usually constructed in two broad steps. First, for narrowly defined, relatively homogeneous products, also known as elementary aggregates, elementary price indexes are calculated. For the US CPI, there are currently 211 such elementary price indexes (called *basic indexes* by the Bureau of Labor Statistics (BLS)).[1] In a second step, these elementary price indexes are aggregated into a single consumer price index using expenditure weights, usually using some modified Laspeyres formula. Expenditure weights are commonly derived from household budget surveys.

The index formula used to construct elementary price indexes matters. The Boskin commission argued, based on a study by Moulton and Smedley (1995), that the US CPI suffered from an annual lower level substitution bias of 0.25 percentage points. The lower level substitution bias occurred because the arithmetic means used *within* the elementary price indexes implicitly assumed no substitution to changes in relative prices within elementary aggregates, say between coke and pepsi. This led to the Boskin commission recommendation that the BLS ought to move to geometric means to construct price indexes at the elementary aggregate level. Since the Boskin report, the BLS indeed moved towards the use of a (weighted) geometric means formula for elementary price indexes (away from arithmetic means of price ratios).

Importantly, official practices in elementary price index construction are still not uniform across countries, warranting further investigation in the consequences of different choices. Due to the absence of expenditure information within the elementary aggregates, it is often the case that elementary price indexes are constructed using *price information only.* In comparing N prices between two periods, using only price information, elementary price indexes can take the form of arithmetic means of price ratios (Carli index), ratios of arithmetic

[1] In countries in the European Union, similarly, elementary price indexes (referred to in EU regulation as *price indices for elementary aggregates*) form the basis of higher aggregate index construction. The number of elementary price indexes in the consumer price index varies across countries.

means (Dutot index) or geometric means of price ratios (Jevons index).[2] EU harmonized index of consumer prices (HICP) regulation allows countries to choose between geometric means and ratios of arithmetic means. For instance, the German HICP index uses arithmetic means for elementary indexes, the French HICP uses both arithmetic means and geometric means depending on the product.[3] An earlier literature, using the actual data underlying the consumer price index, has shown that the differences at the elementary aggregate level between the Dutot, Carli and Jevons index can be quite substantial (See Carruthers, et al, 1980, Dalén, 1994, Schultz, 1995 and Moulton and Smedley, 1995).

With the advent of scanner transactions data it is now possible[4] to use index formulas for elementary aggregates that combine both price and expenditure data. For instance, superlative index formula, such as the Fisher Ideal index or the Törnqvist index, can be constructed using scanner data that involves both prices and quantities. These two indexes approximate true cost of living indexes (Diewert, 1976). Most economists would agree that these indexes are preferable from a theory point of view, at least in cases where they are meant to measure cost of living.[5] At the level of the elementary aggregate, the basket changes as consumers react to relative price changes and switch consumption to very close substitutes, say shampoo from brand A to shampoo from brand B or, as in the example above, between coke and pepsi.[6] A Fisher Ideal index (or any other superlative index) takes into account this substitution.

The availability of scanner data[7] has shifted the former question, asked by the earlier literature, on the disparity between price-only indexes at the elementary level towards the question of how different price indexes at the elementary level are from cost of living indexes. Such a question can be asked for both price-only indexes and for price indexes that use expenditure weights, such as the Laspeyres index. Note that in reality however, at most instances, price indexes at the elementary level are still using price-only index formulas.[8]

Here we define *elementary index bias* (at the level of the elementary aggregate) as the

[2] The CPI manual of the ILO also mentions two other elementary indexes that use prices only: the harmonic average of price ratios and the geometric average of the Carli and harmonic formulas. Those two are not used in the HICP nor in the US CPI and are not considered further here.

[3] For the construction of elementary price indexes, the French HICP calculation uses geometric averages for heterogeneous products and ratios of arithmetic means for homogeneous products. Belgium and Ireland use ratios of arithmetic means and Spain, Italy, Austria and Portugal use geometric means. The Netherlands uses both geometric means and arithmetic means depending on the product. See the Eurostat website for more detail $http : //epp.eurostat.ec.europa.eu/cache/ITY_SDDS/EN/prc_hicp_esms.htm$

[4] Although possible, with only few exceptions, in practice this has not been taken up by Statistical Agencies. Field visits with price observation at the store are by far the most common method for price collection.

[5] This is also the viewpoint expressed by the CPI manual of the ILO. Notwithstanding this, note that most official price indexes, also the Harmonised Index of Consumer Prices (HICP) in the European Union, are not officially designed or meant to approximate true cost of living indexes.

[6] Note that such a switch could occur within a store or across stores.

[7] That is, for research purposes but not necessarily used in actual price index construction.

[8] Note though that the sampling method and the information that goes into the sampling of the prices used for the index can be seen as an implicit weighting.

difference between the price change measured using any particular elementary price index and a superlative price index. The elementary index bias provides an assessment of how far inflation gauged by an elementary index is from a true change in the cost of living. For the superlative price index we take the Fisher Ideal index. If substitution in the elementary price index is not taken into account in the index formula, this leads to lower level substitution bias. Diewert (1998) defines elementary substitution bias as the difference between a fixed base Laspeyres index and the corresponding Fisher Ideal index, both at elementary aggregate level. Here the term elementary index bias is thus a bit broader as it considers differences between any elementary index and the Fisher Ideal index.

Elementary index bias is potentially large. Clearly, the substitution elasticity between closely related products, such as two different brands of shampoo, is likely to be larger than say between shampoo and cars. Indeed, a small growing literature using scanner data compares inflation outcomes at the elementary aggregate level under different elementary index formulas. The findings of this literature suggest that *elementary index bias* can be quite large, or in other words that a different choice of index formula at the elementary level can lead to quite different inflation estimates (of elementary aggregates).

A number of findings from this literature are striking. Silver (1995) compares various price indices for the evolution of prices of TV sets in the UK in 1993 and finds large divergence between elementary price indexes that use only price information and superlative price indexes. E.g. from January to December 1993, a Dutot index drops by almost 14 percent, while a Törnqvist index rises by 1.3 percent. This result is due to large shifts in market shares of different television sets between those two months that is not taken into account by the Dutot index. Most of this literature however compares index formulas that use both prices and quantities, but differ in their substitution assumption. One finding there is that Laspeyres indexes at the elementary level tend to show higher rates of inflation compared to Fisher Ideal or Törnqvist indexes. A finding which is not surprising as the fixed quantity weights in the Laspeyres index do not take into account substitution at the elementary level.[9] Results on the elementary substitution bias, defined as the difference between a fixed basket index versus a superlative index, such as a Fisher Ideal or Törnqvist index, at the elementary aggregate level, can be found in a number of scanner data studies. Silver (1995) and Feenstra and Shapiro (2001) both compare a Laspeyres index with a Törnqvist index. Dalen (1997), Hawkes (1997), Reinsdorf (1999), Jain and Abello (2001), Ivancic, Diewert and Fox (2011), and Rotaru et al.(2011) all compare a Laspeyres index with a Fisher Ideal index.

The studies in this literature have in common that they are based on a small sample of products or a small geographical area. The estimates of lower level substitution bias are

[9] An exception is de Haan and Opperdoes (1997) which find almost no difference between a Laspeyres and a Fisher Ideal index. However, they only use a rather small scanner dataset of only six items in the Dutch coffee market and argue that there is very little price variation in this particular market.

therefore mostly limited to biases that occur at the elementary level for particular products. For instance, Silver (1995) compares different price indexes for TV sets in the UK. Hawkes (1999) and Reinsdorf (1999) compare different elementary indexes using the same dataset, i.e. scanner data for roasted bean coffee and instant coffee in Chicago and Washington DC in the US. Dalen (1997) uses Swedish scanner data for four different item groups (fats, detergents, breakfast cereals and frozen fish). Jain and Abello (2001) uses Australian scanner data from one Australian city to estimate differences between different indexes for 19 products. Rotaru et al. (2011) use US scanner data from 6 cities and 5 products. Ivancic, Diewert and Fox (2011) use one of the largest datasets: Australian scanner data of 19 major supermarket item categories and over 8000 individual products.

So a number of important questions remain. How generalisable are the results of this literature? Are lower level substitution biases a general problem for all products? How wide is the variation of elementary index bias across products? Or in other words, what is the standard deviation of this bias? Do lower level substitution biases show up in aggregate inflation numbers? More precisely, if for some elementary aggregate the bias is positive and for others it is negative, is it possible that higher aggregates show less bias simply through averaging of a set of elementary aggregates? This last question is particularly relevant for the price information only elementary indexes. The direction of the bias for price indexes that use price information only is theoretically ambiguous and can therefore differ across product categories, in contrast to say a Laspeyres index which is expected to have a positive bias for every product category.

These questions are taken up in this paper. It expands on this literature in a number of dimensions. It uses a large dataset covering 42 product categories in 10 countries. It uses detailed item level price and expenditure information obtained from scanner data provided by AC Nielsen. Each individual item belongs to one of 42 narrowly defined product categories. Within each country, for each of the product categories, elementary price indexes using different formulas are constructed. This rich dataset allows to make a comparison of elementary indexes that use prices only (Carli, Dutot and Jevons) versus elementary indexes that use both price and expenditure data (Laspeyres, Paasche, expenditure weighted Jevons, Lowe, geometric lowe and Fisher Ideal index). The elementary index bias is defined by comparing these indexes with a Fisher Ideal index. The effects on measured inflation outcomes are presented and discussed first at the elementary level itself (i.e. the product category level). Thereafter, elementary indexes from the first step are aggregated to the national level using both Laspeyres and Lowe aggregation. Measured inflation outcomes at the national level are presented and discussed. Finally, euro area level aggregated inflation outcomes are presented. This allows to investigate the important question whether differences at the elementary level ”wash out” upon aggregation or not.

Ultimately, the insights of this study are important to understand the impact of elementary level index choice on inflation measurement. With scanner data, statistical agencies could construct expenditure weights at the lowest level of aggregation and therefore compute elementary price indexes that use both price and quantity information. One might therefore imagine a future in which the traditional unweighted-price only elementary indices, Dutot and Jevons, are gradually replaced by expenditure weighted indices. The results of this paper show significant effects on measured inflation, at the elementary, at the national level of aggregation, and at the euro area level as well.

The remainder of the paper is structured as follows. In Section 2 the dataset is described. Section 3 describes the methodology employed, the various unweighted and weighted indices at the elementary level, national and euro area aggregate indexes. Section 4 shows results and section 5 concludes.

2 The data

This paper uses a large dataset provided by the marketing research firm AC Nielsen. The raw data (available to AC Nielsen, but not to us), used to construct this dataset, consists of retail scanner transaction data (price and quantity) of a large set of individual stock keeping units over the period 2009-2010 at the individual store level for a large set of retailers, for a set of 13 countries.[10] In the parlance of retailers, a stock keeping unit is a uniquely identifiable product, of a particular brand, product name, package size, package form (e.g glass or plastic), and content. An example of a stock keeping unit is a six pack 'Heineken Light' consisting of 6 0,25 liter glass bottles.).

One of the features of raw scanner data is that literally thousands of transactions of the same stock keeping unit are occurring in the same period (at different stores). It is clear that price index theory breaks down if every single transaction in every single store is considered separately. Some aggregation, over time and space, needs to be done. This is also the conclusion of Ivancic, Diewert and Fox (2011) who provide a more detailed analysis on the effects of time aggregation with scanner data. Diewert (1995) argues in favor of using unit values as prices at the very lowest level of aggregation: *"To summarize: at the individual outlet level, we recommend using the unit value and the total quantity transacted to form price and quantity estimates of the homogenous commodity for the two periods under consideration*". On the spacial dimension, Diewert (1995) argues further: *Thus the lowest level aggregates would normally be shop specific unit values. However if individual outlet data on transactions were not available or were considered to be too detailed, then unit values for a homogeneous commodity over all outlets in a market area might form the lowest level of*

[10]The exact number and identity of retailers is only known to AC Nielsen, but AC Nielsen attempts to cast a net as wide as possible as this data is generally sold to the large consumer goods manufactures for marketing research purposes.

aggregation.

To make the data available to us, AC Nielsen aggregated the scanner data of each individual stock keeping unit over a homogeneous set of retailers within the country. That is, for each individual stock keeping unit the monthly unit price and quantity is available at the *store type level* (not at the individual store level). That is, AC Nielsen first sorts the individual stores into homogeneous store types. As the retailing landscape differs according to country, each country has a different set of homogeneous store categories (e.g. in our dataset Germany has 14 store types and France has 8). Within each country, for each store type-stock keeping unit pair we have available a monthly unit price and quantity. Monthly unit prices are calculated as total sales (in euros) of the stock keeping unit divided by total volume.[11] We use the term "item" to mean a "*country*"-"*store type*"-"*stock keeping unit*" triple. Our unit prices and quantities sold are therefore at the item level. An example of an item is a 1 liter plastic bottle regular Coca Cola in Germany sold at gas stations.

Each stock keeping unit belongs to exactly one of 42 product categories. The 42 product categories cover a wide range of grocery products that can be considered to be typically found in the shopping cart of a consumer. The list of the 42 product categories is: 100 percent orange juice, diapers, ground coffee, instant coffee, all purpose cleaner, automatic dishwasher detergent, baby food, beer, bouillon, cat food, cereal, chewing gum, condoms, carbonated soft drinks, deodorant, dog food, dry pasta, fabric softener, frozen fish, ice cream, strawberry jam, laundry detergent, margarine, refrigerated milk, uht milk, olive oil, paper towels, panty liner, frozen peas, rice, shampoo, shaving preparation, sugar, tinned peas, tinned tuna, toilet tissue, toothpaste, vodka, sparkling water, still water, wet soups, whiskey.

The data made available to us does obviously not cover all stock keeping units sold in the country. The selection was done to get "*representative*" products. The following procedure was followed. First, for each product category, within each country, (say for example ice cream in Germany), all possible brands that sell stock keeping units in that product category in that country are identified. Of these, four brands are selected. The four brands singled out are those that have the highest market share in terms of sales. Within each country-product category pair two brands are taken that have the highest market share in that product category at the European level and two brands are selected that have the highest market share in that product category nationally. If the brands in that product category with the highest market share at the national level also happen to have the highest European level market share, brands further down the chain of national market share were identified so that the number of brands selected remains four in total for each country-product category pair.

[11]The initial dataset contained unit prices and volumes at a slightly different frequency than monthly (mostly 4 weekly, but also a few 5 weekly and some other frequencies). We had available the monthly frequency data which was constructed by linearly interpolating the initial dataset. For example, imagine unit prices p_1 and p_2 for the two consecutive 4 weekly periods: 6 April 2009 to 3 May 2009 and 4 May 2009 to 31 May 2009. The unit price for May 2009 was constructed as $(3 * p_1 + 28 * p_2)/31$.

After the brands are decided on, individual stock keeping units are chosen. For each brand, three stock keeping units are picked. The selection criterium being the three most sold stock keeping units of that particular brand in that particular product category in that country.

From the dataset, a balanced panel of 15844 individual items was extracted so that for each item unit price and quantity data are available for each month in the period January 2009-December 2010. The following ten countries had a sufficient amount of data for the purpose of this paper: Austria, Belgium, Germany, Spain, France, Greece, Italy, Ireland, Netherlands and Portugal. The average number of items per country, in the balanced dataset, is thus 1584.

Table 1: Number of product categories and items per country

Country	product categories	stock keeping units	store types	items
AT	41	371	9	1925
BE	38	338	4	925
DE	40	383	14	2350
ES	16	153	15	816
FR	32	333	8	2002
GR	34	347	5	1680
IE	33	306	4	891
IT	34	346	20	2952
NL	33	242	12	779
PT	37	347	9	1524

One of the requirements to construct the index numbers, all with identically the same data, is that the price and quantity information of each item is available for the entire period January 2009-December 2010. This criterion implies that not all product categories are available for all countries. Table 1 shows the number of product categories, stock keeping units, store types and the number of items for each country in the balanced dataset.

3 The price indexes

This section describes the different indexes at the elementary level and at the aggregate level used in the paper. The theoretical discussion is held brief. More extensive theoretical discussions on elementary indexes can be found in Chapter 20 of the Consumer price index manual of the International Labour Organisation (ILO), Diewert (1995) and Triplett (1998).

3.1 Elementary indexes

The Carli, Dutot and Jevons indexes are elementary indexes at the product category level that only use price information. They are defined as follows. Let there be N_i prices of individual items for a given product category i observed in base month t_0 and in month t.

Let p_{int} be the price of the n-th item in product category i in month t. The Carli index, Dutot index and Jevons index for elementary aggregate i are defined as follows:

Carli index:

$$CARLI_{it} = (1/N_{it}) \sum_{n}^{N_i} \frac{p_{int}}{p_{int_0}} \tag{1}$$

Dutot Index ;

$$DUTOT_{it} = \frac{\sum_n^{N_i} p_{int}}{\sum_n^{N_i} p_{int_0}} = \sum_{n}^{N_i} (\frac{p_{int}}{p_{int_0}})(\frac{p_{int_0}}{\sum_n^{N_i} p_{int_0}}) \tag{2}$$

Jevons Index:

$$JEVONS_{it} = \prod_{n}^{N_i} (\frac{p_{int}}{p_{int_0}})^{1/N_i} \tag{3}$$

In the Carli index, the price ratios of every item carry the same weight and are averaged arithmetically. The Dutot index can be seen as an arithmetically weigthed average of price ratios where the higher priced items carry the higher weight. The Jevons index is an equally weighted geometric average of price ratios. It is quite well known that a geometric mean always is below an arithmetic mean, so that the Jevons index will always be below the Carli index. Of those three indexes, the Carli index is the only one that does not pass the time reversal test (Diewert 1995).

In contrast, weighted indexes can be constructed when expenditure data is available.[12] Index number theory as it relates to cost of living suggests a superlative index such as the Fisher Ideal index. The Fisher Ideal price index is thus preferred by economic theory. As stated in Diewert (1995) : *"Thus if price and quantity information is available at the elementary level, it seems preferable to use the Fisher ideal price index to aggregate the basic level price quotes rather than the Laspeyres, Paasche or geometric indexes..."*

Fisher Ideal index:

$$FISH_{it} = \sqrt{LASP_{it} * PAAS_{it}} \tag{4}$$

with

Laspeyres index:

$$LASP_{it} = \frac{\sum_n^{N_i} p_{int} * q_{int_0}}{\sum_n^{N_i} p_{int_0} * q_{int_0}} = \sum_{n}^{N_i} \left(\frac{p_{int}}{p_{int_0}}\right) * \left(\frac{p_{int_0} * q_{int_0}}{\sum_n^{N_i} p_{int_0} * q_{int_0}}\right) \tag{5}$$

[12]See also chapter nine of the Consumer Price Index Manual of the International Labour Organisation (ILO) for a discussion of weighted indexes at the elementary level.

Paasche index:

$$PAAS_{it} = \frac{\sum_n^{N_i} p_{int} * q_{int}}{\sum_n^{N_i} p_{int_0} * q_{int}} = \left[\sum_n^{N_i} \left(\frac{p_{int_0}}{p_{int}} \right) * \left(\frac{p_{int} * q_{int}}{\sum_n^{N_i} p_{int} * q_{int}} \right) \right]^{-1} \tag{6}$$

The Fisher Ideal index uses quantities in both period t_o and t and allows for substitution effects. It can be contrasted to the Laspeyres index, which keeps weights fixed at the base period. In a Laspeyres index the base period of the quantities and prices are identical.

For higher levels of aggregation however, official statistics rarely follow an exact Laspeyres formula. Modified Laspeyres or so-called Lowe indexes, where the base for quantities and prices differ, are more common. Practice in the EU is to take a base month for the price (December) but a base year for the quantities. The idea is that seasonality might make monthly quantities unstable, so that quantities aggregated over a whole year might be preferable.

A modified version of the Laspeyres index is the following Lowe index:

Lowe index:

$$LOWE_{it} = \frac{\sum_n^{N_i} p_{int} * q^s_{int_0}}{\sum_n^{N_i} p_{int_0} * q^s_{int_0}} = \sum_n^{N_i} \left(\frac{p_{int}}{p_{int_0}} \right) * \left(\frac{p_{int_0} * q^s_{ink}}{\sum_n^{N_i} p_{int_0} * q^s_{ink}} \right) \tag{7}$$

with

$$q^s_{ink} = \sum_{j=0}^{11} q_{in(t_0-j)} \tag{8}$$

where q^s_{ink} are the 12 month quantities of individual item n (i.e. the sum of quantities of months $t_0, t_0 - 1, ..., t_0 - 11$).

The Laspeyres and Lowe indexes use formulas that employ a weighted arithmetic mean of price ratios. They will only differ in measured inflation to the extent that the annual based quantity weights from the Lowe index are dissimilar from the monthly based quantity weights of the Laspeyres index. They can also be compared to a Carli index, which is an unweighted arithmetic mean of price ratios or the Dutot index which implicitly uses the prices as weights. So Laspeyres, Lowe, Carli and Dutot indexes when applied to the elementary level can all be seen to belong to the same family of averages of price ratios, all using different weights to take arithmetic averages of price ratios.

Alternatively, price ratios can be geometrically averaged. Geometric averaging allows for substitution between items. The Jevons index defined above uses equal weights across all items. It is used in official practice if weights of individual items are not available. Alternatively, items can be weighted differently, such as in the expenditure weighted Jevons index and the geometric Lowe index. These last two are defined as follows.

expenditure weighted Jevons index:

$$JEVEW_{it} = \prod_{n}^{N_i} \left[\frac{p_{int}}{p_{int_0}}\right]^{\left(\frac{p_{int_0} * q_{int_0}}{\sum_n^{N_i} p_{int_0} * q_{int_0}}\right)} \tag{9}$$

geometric Lowe index

$$GLOWE_{it} = \prod_{n}^{N_i} \left[\frac{p_{int}}{p_{int_0}}\right]^{\left(\frac{p_{int_0} * q^s_{ink}}{\sum_n^{N_i} p_{int_0} * q^s_{ink}}\right)} \tag{10}$$

The expenditure weighted Jevons index is an exact measure of the cost of living for a Cobb-Douglas utility function. Note that for such a utility function, the expenditure shares are constant over time. The geometric Lowe index has a similar form as the expenditure weighted Jevons index. The expenditure shares of the geometric Lowe index are constructed by taking annual quantities at base prices.

This gives in total eight different elementary index formulas at the product category level that are compared below with a Fisher Ideal index. Three indexes use only price information: Carli, Dutot and Jevons. Five indexes use price and quantity information: Laspeyres, Paasche, Lowe, expenditure weighted Jevons and geometric Lowe.

The following well known results from index number theory should be expected to be found back in the empirical exercise (see Diewert (1995) and references therein). The Carli index is always above the Dutot and Jevons index. The Laspeyres index is expected to be above the Fisher Ideal index as it does not allow for substitution effects as prices change. The Paasche index is expected to be below the Fisher Ideal index. The Laspeyres and the Lowe index are expected to approximate each other if products are not very seasonal, so that the base month weights of the Laspeyres approximate the annual weights of the Lowe index. The same holds for the expenditure weighted Jevons and geometric Lowe index.

3.2 Aggregate indexes

The elementary indexes defined above aggregate individual items within a product category within each country. The data covers on average 34 product categories per country. These clearly do not cover the whole universe of product categories that go into a national price index. Nevertheless, the large set of product categories available allow us to investigate if upon aggregating over such a large set of product categories, biases at the elementary level remain apparent in national aggregates.

Aggregating the elementary indexes at the national level is done using two aggregation formulas, Laspeyres and Lowe. Define I_{it} to be the elementary index of product category i (within country c).[13] Let K be the number of product categories. Laspeyres is the textbook

[13]In the country level indexes that follow, all the elementary indexes I_{it} and the prices p_{int_0} and so on,

fixed weight price index formula:

$$LASP_{ct}^{a} = \sum_{i}^{K} (I_{it}) * \left(\frac{\sum_{i}^{N_i} p_{int_0} * q_{int_0}}{\sum_{i}^{K} \sum_{n}^{N_i} p_{int_0} * q_{int_0}} \right) \tag{11}$$

As indicated above, for higher levels of aggregation official statistics rarely follow an exact Laspeyres formula. Modified Laspeyres, i.e. Lowe indexes, are more common. A Lowe aggregation of elementary indexes is defined as:

$$LOWE_{ct}^{a} = \sum_{i}^{K} (I_{it}) * \left(\frac{\sum_{n}^{N_i} p_{int_0} * q_{ink}^{s}}{\sum_{i}^{K} \sum_{n}^{N_i} p_{int_0} * q_{ink}^{s}} \right) \tag{12}$$

with

$$q_{ink}^{s} = \sum_{j=0}^{11} q_{in(t_0-j)} \tag{13}$$

where q_{ink}^{s} are the 12 month quantities of individual item n (i.e. the sum of quantities of months $t_0, t_0 - 1, ..., t_0 - 11$).

Alternatively, a direct Fisher Ideal index can be calculated. Note that the direct Fisher Ideal index is not consistent in aggregation (Diewert 1978). Therefore, the Fisher Ideal index is built up directly from the item data, and not as a Fisher Ideal index from elementary indexes that have the Fisher Ideal index form. At the country level, the following direct Fisher Ideal index is constructed:

$$FISH_{ct} = \sqrt{LASP_{ct} * PAAS_{ct}} \tag{14}$$

with Laspeyres and Paasche country level indexes:

$$LASP_{ct} = \frac{\sum_{i}^{K} \sum_{n}^{N_i} p_{int} * q_{int_0}}{\sum_{i}^{K} \sum_{n}^{N_i} p_{int_0} * q_{int_0}} = \sum_{i}^{K} \sum_{n}^{N_i} \left(\frac{p_{int}}{p_{int_0}} \right) * \left(\frac{p_{int_0} * q_{int_0}}{\sum_{i}^{K} \sum_{n}^{N_i} p_{int_0} * q_{int_0}} \right) \tag{15}$$

and

Paasche aggregate index

$$PAAS_{ct} = \frac{\sum_{i}^{K} \sum_{n}^{N_i} p_{int} * q_{int}}{\sum_{i}^{K} \sum_{n}^{N_i} p_{int_0} * q_{int}} = \left[\sum_{i}^{K} \sum_{n}^{N_i} \left(\frac{p_{int_0}}{p_{int}} \right) * \left(\frac{p_{int} * q_{int}}{\sum_{i}^{K} \sum_{n}^{N_i} p_{int} * q_{int}} \right) \right]^{-1} \tag{16}$$

Aggregating the national level to the euro area level is done similarly as above using three aggregation formulas, Laspeyres, Lowe, and a direct Fisher Ideal index. The exact formulas used at the euro area level are in the Appendix.

should have a country subscript c. To simplify notation the c is dropped.

4 Results

4.1 Elementary index bias at the product category level

The elementary index bias is defined as the difference between the elementary index and the Fisher Ideal index, both calculated at the product category level. Each elementary index is calculated for a total of 338 product category observations (i.e. an average of 33.8 product categories per country). Table 2 shows summary statistics of the elementary index bias at the product category level. It shows the average, the standard deviation, minimum and maximum, and a number of percentiles of the distribution of the elementary index bias across product categories for eight indexes (Carli, Dutot, Jevons, Laspeyres, Lowe, Paasche, expenditure weighted Jevons and geometric Lowe index).

Table 2: Elementary index bias for different elementary index formulas (percentage points)

	CARLI	DUTOT	JEVONS	LASP	LOWE	PAAS	JEVEW	GLOWE
mean	0.49	-0.05	0.00	0.55	0.23	-0.53	0.27	-0.05
sd	3.12	3.58	3.01	1.29	1.08	1.22	1.03	1.07
min	-19.74	-27.53	-21.59	-5.09	-4.54	-10.47	-5.57	-6.87
p5	-3.36	-5.03	-4.07	-0.20	-0.89	-2.76	-0.43	-1.42
p25	-0.77	-1.25	-1.00	0.06	-0.07	-0.61	-0.02	-0.23
p50	0.42	0.14	0.08	0.20	0.10	-0.20	0.09	-0.01
p75	1.58	1.54	1.37	0.62	0.34	-0.06	0.34	0.17
p95	5.05	5.30	4.04	2.85	1.97	0.20	1.99	1.33
max	24.03	9.98	13.54	11.69	7.69	5.36	10.39	5.99

Note: Results for each column are based on 338 product category index observations.
Elementary index bias is defined as elementary index minus Fisher Ideal index.
Each elementary index is the direct comparison index at December 2010, base December 2009.

Consider first the elementary indexes that only use price information, the Carli, Dutot and Jevons index. The average elementary index bias is largest for the Carli index at 0.49 percentage points. On average, the Carli index overestimates price inflation. Note that it is forbidden in EU HICP regulation, and also not any longer used in the US, which seems to be for good reason. The average elementary index bias for the Dutot index is low, at -0.05 and even absent at 0.00 percentage points for the Jevons index. Both the Dutot and Jevons index are used at the product category level in actual practice of statistical agencies for official price index measurement. The low average elementary index bias seems to support their official use.

Notably, for all of these three indexes, there is a large variation in the bias across product categories. For instance, the standard deviation for the Dutot index is 3.58 percentage points. It is smaller but still substantial at 3.01 for the Jevons index and 3.12 for the Carli index. The low average bias but large standard deviation for the Dutot and Jevons index implies that an unweighted average of elementary indexes across a number of product categories will be rather close to an unweighted average of Fisher Ideal indexes, however for any particular individual product category, the bias might be large. Choosing between the Dutot or Jevons index, Jevons seems to be the better choice (i.e. has the lowest standard deviation with a zero mean bias). The substantial standard deviation of the bias across product categories implies that annual inflation numbers at the product category level, when measured using elementary indexes such as the Dutot or Jevons index, have to be considered with considerable caution.

Looking in more detail to the indexes that use both price and quantity information, the Laspeyres index has the largest average positive bias at 0.55 percentage points. The Paasche index has the (in absolute terms) largest average negative bias, i.e. -0.53 percentage points. Both results are as one should expect, the Laspeyres index, which uses base period expenditure weights, does not take into account substitution and has a built in upward bias. The Paasche index uses current period expenditure weights and has a built in downward bias (i.e. it takes already into account the shift towards a new basket as relative prices change). The geometric Lowe index has a small average bias, at -0.05 percentage points. Both the Laspeyres and the geometric Lowe index have lower standard deviation than the Dutot and Jevons index, at 1.29 and 1.07 percentage points respectively. The expenditure weighted Jevons index has both a lower bias and standard deviation than the Laspeyres index, 0.27 and 1.03 percentage points respectively. The Lowe index has a lower bias and a lower standard deviation than the Laspeyres index. Of all indexes, the expenditure weighted Jevons index has the lowest standard deviation.

These results can be summarized as follows. For any elementary index, the standard deviation of the bias is large, and using both price and quantity information reduces the standard deviation of the bias considerably compared to the indexes that only use price information. Roughly, the standard deviation is more than halved. Nevertheless, the standard deviation of the bias, for all the indexes that use both price and quantity information, remains still considerable and is in all cases above 1 percentage point.

The average bias and its large variation, for all the indexes considered here, is consistent with the large variety of the magnitudes of the elementary index bias found in the previous literature (using a more limited number of product categories and index comparisons). The large difference in TV price inflation when using the Dutot index versus a Törnqvist index found by Silver (1995) was already mentioned above. But also Laspeyres indexes show large varying biases at the elementary level. For instance, comparing the Laspeyres with the Fisher Ideal index for coffee scanner data in Chicago and Washington, Hawkes (1997) finds differences in measured annual inflation rates as large as 9.3 percentage points for roasted coffee in Chicago versus 1.6 percentage points for instant coffee in Chicago. For Washington, the numbers are respectively 1 and 1.9 percentage points (see Table 2 in Hawkes, 1997). So, even for as closely related products as roasted and instant coffee, biases show a large variation.

Dalen (1997) shows a difference in measured annual inflation between the Laspeyres index and the Fisher Ideal index of 0.50 percentage points for detergents, 0.20 percentage points for breakfast cereal and 1.3 percentage points for frozen fish.

Diewert (1998) summarizes the scanner data literature on elementary index bias as follows: *"Some estimates of elementary index bias (on an annual basis) that emerged from these studies were: 1.1 percentage points for television sets in the United Kingdom; 4.5 percentage points for coffee in the United States; 1.5 percentage points for ketchup, toilet tissue, milk and tune in the United States; 1 percentage points for fats, detergents, breakfast cereals and frozen fish in Sweden; 1 percentage point for coffee in the Netherlands and 3 percentage points for coffee in the United States... It is unclear to what extent these large estimates can be generalized to other commodities."* The results in this paper, based on 338 product category estimates per index, show that indeed these estimates can be generalized to other commodities, at least to the large set of grocery products considered here. This sheds new light on the importance of elementary index bias and its considerable variation. It also shows that having weights available one can drastically reduce this variation.

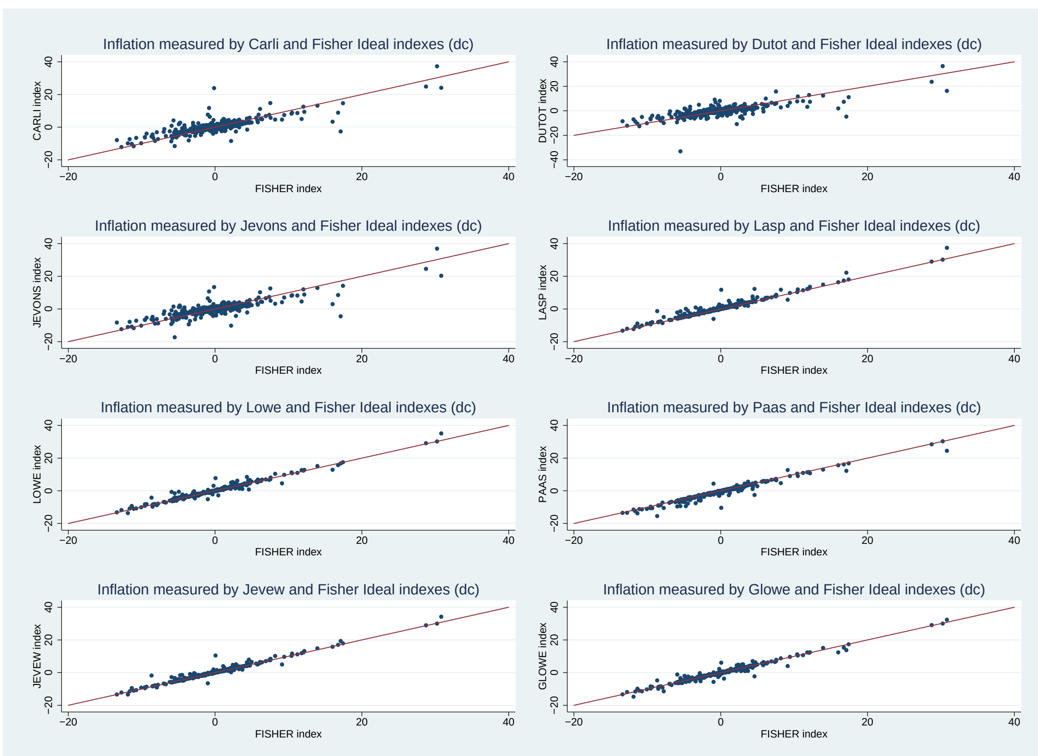

Figure 1: *Inflation at product category level*

Another way of showing that the variation across product categories is substantial is demonstrated in Figure 1. Figure 1 depicts the measured inflation rates for the different elementary indexes (y-axis) compared to the measured inflation rate by the Fisher Ideal index (x-axis). Each dot represents the two measured inflation rates of one product category

Table 3: Elementary Index Bias for different elementary index formulas: country results (mean and standard deviation) (percentage points)

	AT	BE	DE	ES	FR	GR	IE	IT	NL	PT	Total
CARLI	1.78	0.92	0.87	2.22	-0.02	1.34	-2.27	0.63	0.40	-0.50	0.49
	2.15	2.50	1.62	6.00	1.11	3.17	5.11	0.95	3.21	2.36	3.12
DUTOT	1.19	0.85	0.71	0.73	-0.42	0.70	-3.08	0.37	-0.50	-1.11	-0.05
	2.34	2.46	2.25	2.59	1.56	2.79	5.97	1.10	6.35	2.75	3.58
JEVONS	1.28	0.51	0.55	1.28	-0.20	0.90	-2.90	0.28	-0.42	-1.04	0.00
	2.11	2.38	1.57	3.57	1.13	2.99	5.46	0.93	3.53	2.32	3.01
LASP	0.88	0.16	0.40	0.19	0.19	0.70	1.23	0.36	0.71	0.46	0.55
	1.40	1.02	1.01	0.73	0.24	1.49	1.76	0.35	2.16	0.90	1.29
LOWE	0.21	0.21	-0.02	-0.15	0.12	0.50	0.49	0.04	0.41	0.30	0.23
	1.00	1.06	0.98	0.75	0.20	1.04	1.55	0.21	1.71	1.12	1.08
PAAS	-0.85	-0.15	-0.39	-0.19	-0.19	-0.68	-1.19	-0.36	-0.66	-0.45	-0.53
	1.32	1.05	1.00	0.73	0.24	1.40	1.69	0.35	1.96	0.88	1.22
JEVEW	0.49	-0.04	0.20	-0.06	0.11	0.43	0.67	0.22	0.30	0.26	0.27
	0.77	1.04	0.98	0.54	0.20	1.28	1.16	0.28	1.88	0.77	1.03
GLOWE	-0.16	0.00	-0.21	-0.36	0.04	0.21	-0.06	-0.09	-0.08	0.08	-0.05
	1.30	1.04	1.11	0.79	0.20	0.94	1.41	0.21	1.59	1.07	1.07

Note: Results for each column are based on the number of product category observations shown in table 1.
Elementary index bias is defined as elementary index minus Fisher Ideal index.
Each elementary index is the direct comparison index at December 2010, base December 2009.
Last column gives mean and standard deviation of all country results pooled.

in one particular country. Dots above the 45 degree line indicate that the elementary index has a positive bias, dots below the 45 degree line indicate a negative bias. As expected for the Laspeyres index, most of the dots are above the 45 degree line, indicating positive bias for most product categories, the opposite is true for the Paasche index. The larger standard deviation of the bias for the price only indexes, Carli, Dutot and Jevons is also clear as the dots are scattered more widely around the 45 degree line. Note that the level of the bias does not seem related to the level of inflation measured by the Fisher Ideal index. Inflation rates, as measured by this index, at the product category level, have a wide range from about -20 percent to +40 percent. The deviations from the 45 degree line however do not seem to increase for higher levels of inflation.

The results when pooling all countries together masks some interesting facts that are revealed when looking at the country results. Table 3 shows the mean and standard deviation of the elementary index bias at the product category level for each of the different indexes per country. Note that the number of product categories varies slightly across countries, ranging from 16 in Spain to 41 in Austria. First, one readily observes that the Laspeyres index has consistently a positive average elementary index bias, i.e. in all ten countries. The Paasche index has a consistent average negative bias for all the ten countries. As a matter of fact, the Laspeyres and Paasche indexes are the *only* two indexes with an average bias that has a consistent sign across the ten countries. The average bias of the Laspeyres index ranges from 0.16 percentage points in Belgium to 1.23 percentage points in Ireland. Also, the standard deviation of the Laspeyres bias varies quite a lot, from a low 0.24 percentage points in France to a high 2.16 percentage points in the Netherlands. The index that comes closest to the Laspeyres index is the modified one, i.e. the Lowe index, which has a mean bias that is positive in 8 countries and negative in 2 countries. To the extent that the base month weights of the Laspeyres index are similar to the yearly weights of the Lowe index this result is to be expected. The other indexes have positive or negative bias depending on the country. For instance, the Dutot index has a mean bias that is positive in 6 countries and negative in 4 countries. The same is true for the Jevons index. The Carli index has a mean bias that is positive in 7 countries and negative in 3 countries. The geometric Lowe index has a mean bias that is positive in 4 countries and negative in 6 countries. The country results underscore the variability of the elementary index bias, not only across countries but also within countries.

4.2 National indexes as aggregates of elementary indexes

National official consumer price indexes are generally produced in two steps. First, for narrowly defined product categories, elementary indexes are produced. These are thereafter aggregated into a national index. This latter step typically takes place using some modified Laspeyres formula, such as the Lowe index. National official price indexes are therefore composites of product category elementary indexes.

In this section, we provide an answer to the following question: How does a Laspeyres or Lowe aggregate of elementary indexes compare to a direct Fisher Ideal index?[14] Note that the aim of this section is *not* to obtain estimates of biases in national official price indexes. The dataset used here (a large sample of grocery products) versus the samples used at the national level in offical statistics (samples of prices of the entire consumer basket) are obviously very different. Rather, by showing results of aggregates of elementary indexes, the aim is to infer something meaningful on the potential bias in aggregate indexes that stem from lower level elementary index choice. How do the different elementary index formulas perform in this respect? Does it matter that one uses a Dutot or a Jevons index at the product category level if one is only interested in national inflation? Do elementary indexes

[14] The formula for the direct Fisher Ideal index is given in section 3.

that use price and quantity information lead to less variable bias in the aggregate index? These questions are definitely not settled as official practices in elementary index choice still differ across countries. The answers to these questions are relevant for the construction of official statistics and especially the choice of elementary index in practice.

In this section, both the textbook Laspeyres aggregation and the Lowe aggregation (more common in practice) at the national level are considered. The expenditure weights are directly derived from the scanner dataset.[15] This leads to one national index for each different elementary index formula. Each national index calculated here is the aggregate of around 34 product category elementary indexes (the number varies slightly across countries and is given in Table 1). The results are compared with a direct Fisher Ideal index.

In the national index, constructed with the Laspeyres or Lowe aggregation, two sources of bias, relative to the Fisher Ideal index, can be distinguished. First, the elementary index choice (being different from a Fisher Ideal index) and second, the fact that the aggregation uses the fixed weights Laspeyres (or Lowe formula). With respect to the first source of bias, when using elementary indexes at the product category level that are different from a Fisher Ideal index, each product category index has an elementary index bias, the result having been discussed above. The Laspeyres (or Lowe) aggregation of the elementary indexes, each with their own bias, leads to an *aggregate elementary index bias* for the national index. With respect to the second source of bias, a Laspeyres (or Lowe) aggregation of the product category indexes leads to an upper level substitution bias, as substitution between different product categories (say between ice cream and strawberry jam or rice and pasta) is not taken into account.

The two biases for the Laspeyres index can be computed as follows. Denote by $LASP^a_{ct}$ the Laspeyres aggregation of the elementary indexes (see equation 11), denote by $LASP^F_{ct}$ the Laspeyres aggregation of the elementary indexes which use the Fisher Ideal index formula and denote by $FISH_{ct}$ the direct Fisher Ideal index at the country level (see equation 14).

The aggregate elementary index bias is defined as $LASP^a_{ct} - LASP^F_{ct}$. This bias captures the effect of using a different index formula than the Fisher Ideal index at the elementary level. The upper level substitution bias is defined as $LASP^F_{ct} - FISH_{ct}$. This bias captures the fact of using Laspeyres aggregation of product categories to obtain the national index.

The difference between the Laspeyres aggregation of elementary indexes and the direct Fisher Ideal index is equal to the sum of the two biases.

$$LASP^a_{ct} - FISH_{ct} = (LASP^a_{ct} - LASP^F_{ct}) + (LASP^F_{ct} - FISH_{ct}) \tag{17}$$

The two biases are defined similarly for the Lowe aggregation. When price indexes are used at the elementary level that do not account for substitution, aggregate elementary

[15] Tables 10 and 11 in the Appendix present the expenditure weights used in the aggregation.

index bias has also been called lower level substitution bias. This terminology was used by the Boskin commission to denote the bias that was caused by using the Carli formula at the elementary level in the US CPI. This led the BLS to switch to the geometric means formula in 1999. The term aggregate elementary index bias is used here to denote any bias caused by not using the Fisher Ideal index at the elementary level, so it also applies for elementary indexes that take substitution into account, such as those that use geometric averages.

Table 4 shows the results for the upper level substitution bias and for the aggregate elementary index bias. Results are shown both for the Laspeyres aggregation and for the Lowe aggregation. For each different elementary index formula a national Laspeyres (and Lowe) aggregate index was constructed, leading to ten country indexes per elementary index formula. This allows us to also construct a mean and standard deviation for the two biases.

The results for the Laspeyres and Lowe aggregation are very similar. The mean estimate (based on 10 country results) of upper level substitution bias is 0.20 percentage points for the Laspeyres aggregation and 0.25 percentage points for the Lowe aggregation. These levels are similar to the upper level substitution bias of 0.15 quoted for the US CPI by the Boskin commission (Boskin et al., 1998).

The Laspeyres and Lowe aggregate indexes are weighted arithmetic means of elementary indexes, so that the aggregate elementary index bias is also a weighted mean of elementary index biases over the different product categories. The fact that aggregate elementary index bias derives from averaging is an important point. The high estimates of the elementary index bias for individual product categories of the earlier scanner data studies, and also found back here, are individually worrisome, that is when considering product category inflation estimates. However, averaging the biases generally leads to less extreme results for the aggregate elementary index bias.

The mean estimate (across countries) of the aggregate elementary index bias differs depending on the elementary index. This should not be surprising, as the unweighted mean elementary index bias at the product category level considered in the section above also differed. Considering the Laspeyres aggregation, it ranges from -0.44 percentage points for the Paasche index at the elementary level to 0.44 percentage points for the Laspeyres index at the elementary level. Interestingly, the mean estimate for the aggregate elementary index bias using the Carli index, at 0.34 percentage points found here, is similar to the lower level substitution bias mentioned by the Boskin commission of 0.25 which led to the abolishment of the Carli index by the BLS.

Note also that the standard deviation of the aggregate elementary index bias when using price only elementary indexes (Carli, Dutot, Jevons) are still rather large and are in the range from 1.19 percentage points to around 1.49 percentage points. The standard deviation is much smaller (smaller by a factor five !) when using elementary indexes that use both price and quantity information. This finding strongly supports the use of scanner data as it can reduce

Table 4: Upper level substitution bias and aggregate elementary index bias: summary table (percentage points)

Elementary Index	mean	sd	min	p50	max
Laspeyres aggregation of elementary indexes					
Upper level substitution bias					
	0.20	0.25	0.01	0.10	0.66
Aggregate elementary index bias					
CARLI	0.34	1.19	-1.84	0.48	2.31
DUTOT	-0.13	1.49	-3.00	-0.01	2.55
JEVONS	-0.06	1.22	-2.42	0.15	1.78
LASP	0.45	0.30	0.15	0.34	0.98
LOWE	0.22	0.22	-0.03	0.11	0.59
PAAS	-0.44	0.29	-0.96	-0.34	-0.15
JEVEW	0.26	0.24	-0.00	0.16	0.65
GLOWE	0.03	0.17	-0.16	-0.01	0.40
Lowe aggregation of elementary indexes					
Upper level substitution bias					
	0.25	0.51	-0.26	0.15	1.56
Aggregate elementary index bias					
CARLI	0.34	1.18	-1.81	0.51	2.21
DUTOT	-0.13	1.46	-2.90	0.08	2.44
JEVONS	-0.07	1.22	-2.38	0.18	1.68
LASP	0.47	0.32	0.15	0.35	1.06
LOWE	0.23	0.23	-0.04	0.12	0.60
PAAS	-0.45	0.30	-1.04	-0.35	-0.14
JEVEW	0.27	0.26	-0.01	0.18	0.69
GLOWE	0.03	0.18	-0.18	-0.00	0.41

Note: Results for each row are based on 10 country level indexes. Each elementary index is the direct comparison index at December 2010, base December 2009. National index is the direct comparison index at December 2010, base December 2009. Aggregate elementary index bias is defined as the difference between the Laspeyres (or Lowe) aggregation of the elementary indexes and the Laspeyres (or Lowe) aggregation of Fisher Ideal elementary indexes. Upper level substitution bias is defined as the difference between the Laspeyres (or Lowe) aggregation of Fisher Ideal elementary indexes and the direct Fisher Ideal index.

Table 5: Upper level substitution bias and aggregate elementary index bias: national results (percentage points)

	AT	BE	DE	ES	FR	GR	IE	IT	NL	PT	Total
Laspeyres aggregation of elementary indexes											
Upper level substitution bias											
	0.12	0.12	0.10	0.09	0.01	0.66	0.66	0.09	0.06	0.05	0.20
Aggregate elementary index bias											
CARLI	1.24	2.31	0.46	0.99	0.11	1.02	-0.51	0.50	-1.84	-0.91	0.34
DUTOT	0.85	2.55	-0.19	0.41	-0.22	0.66	-0.98	0.16	-3.00	-1.52	-0.13
JEVONS	0.86	1.78	0.10	0.57	-0.02	0.71	-0.95	0.20	-2.42	-1.40	-0.06
LASP	0.76	0.16	0.33	0.15	0.21	0.82	0.98	0.36	0.47	0.29	0.45
LOWE	0.41	0.10	0.01	-0.03	0.12	0.59	0.47	0.02	0.42	0.10	0.22
PAAS	-0.75	-0.16	-0.32	-0.15	-0.21	-0.78	-0.96	-0.35	-0.45	-0.28	-0.44
JEVEW	0.48	-0.00	0.16	0.01	0.13	0.64	0.65	0.23	0.16	0.12	0.26
GLOWE	0.13	-0.07	-0.14	-0.16	0.05	0.40	0.15	-0.11	0.09	-0.06	0.03
Lowe aggregation of elementary indexes											
Upper level substitution bias											
	0.19	0.38	-0.26	-0.09	-0.04	0.12	1.56	-0.04	0.22	0.47	0.25
Aggregate elementary index bias											
CARLI	1.22	2.21	0.52	1.01	0.18	1.07	-0.62	0.50	-1.81	-0.91	0.34
DUTOT	0.81	2.44	-0.03	0.35	-0.19	0.71	-1.19	0.19	-2.90	-1.53	-0.13
JEVONS	0.84	1.68	0.17	0.55	0.02	0.77	-1.11	0.19	-2.38	-1.41	-0.07
LASP	0.73	0.17	0.34	0.15	0.22	0.86	1.06	0.36	0.49	0.28	0.47
LOWE	0.41	0.11	-0.01	-0.04	0.12	0.60	0.50	0.04	0.44	0.12	0.23
PAAS	-0.72	-0.17	-0.34	-0.14	-0.22	-0.82	-1.04	-0.36	-0.47	-0.27	-0.45
JEVEW	0.46	0.00	0.17	-0.01	0.14	0.68	0.69	0.23	0.19	0.11	0.27
GLOWE	0.14	-0.07	-0.17	-0.18	0.04	0.41	0.15	-0.10	0.12	-0.04	0.03

the variation of the bias considerably. For the Laspeyres, Paasche, Lowe, geometric Lowe and expenditure weighted Jevons index the standard deviation is in the range of 0.17 to 0.30 percentage points. The geometric Lowe index seems to perform quite well, it has almost zero average aggregate elementary index bias (0.03 percentage points), with the lowest standard deviation (0.17 percentage points).

Table 5 shows the results of the upper level substitution bias and aggregate elementary index bias for each country individually. The upper level substitution bias is notably high for Greece and Ireland at 0.66 percentage points (compared to the average 0.20 percentage points). In the 2009-2010 period consumers in these countries were severely hit by the financial crisis which likely led to larger levels of substitution. The aggregate elementary index bias can be quite large, in particular for the price only indexes Carli, Dutot and Jevons. For instance, the bias for the Dutot index is 2.55 in Belgium and -3.00 in the Netherlands. The indexes that use both price and quantity information generally have less extreme biases. Further, it is often the case that the aggregate elementary index bias is larger than the upper level substitution bias. This reconfirms that the choice of elementary index has indeed an important effect on measured inflation.

4.3 Results at the euro area level

Aggregate elementary index bias at the national level can still be quite large. Again, one should expect that aggregating over a set of countries (say at the euro area level) the bias should be smaller. In this section, the Laspeyres and Lowe aggregation of the national indexes into a euro area level index is considered. Here as well, this aggregation is compared with a direct Fisher Ideal index.

The two different biases, upper level substitution bias and aggregate elementary index bias are now defined at the euro area level in a similar way as at the national level. Denote by $LASP_{et}^{a}$ the Laspeyres aggregation of the national indexes (see equation 20), denote by $LASP_{et}^{F}$ the Laspeyres aggregation of the national indexes that use the Fisher Ideal index formula at the elementary level and denote by $FISH_{et}$ the direct Fisher Ideal index at the euro area level (see equation 23). The aggregate elementary index bias at the euro area level is defined as $LASP_{et}^{a} - LASP_{et}^{F}$. The upper level substitution bias is defined as $LASP_{et}^{F} - FISH_{et}$. The biases are similarly defined for the Lowe aggregation.

Table 6 shows the results for the upper level substitution bias and for the aggregate elementary index bias. Results are again shown both for the Laspeyres aggregation and for the Lowe aggregation. The upper level substitution bias is 0.11 percentage points for the Laspeyres aggregation and 0.01 percentage points for the Lowe aggregation. The aggregate elementary bias for the Laspeyres aggregation ranges from -0.33 percentage points, when using the Paasche elementary index, to 0.34 percentage points, when using the Laspeyres

Table 6: Upper level substitution bias and aggregate elementary index bias: euro area (percentage points)

	Euro Area
Laspeyres aggregation of national and elementary indexes	
Upper level substitution bias	
	0.11
Aggregate elementary index bias	
CARLI	0.31
DUTOT	-0.13
JEVONS	0.01
LASP	0.34
LOWE	0.12
PAAS	-0.33
JEVEW	0.20
GLOWE	-0.02
Lowe aggregation of national and elementary indexes	
Upper level substitution bias	
	0.01
Aggregate elementary index bias	
CARLI	0.33
DUTOT	-0.10
JEVONS	0.03
LASP	0.35
LOWE	0.12
PAAS	-0.29
JEVEW	0.20
GLOWE	-0.02

elementary index. The results for the Lowe aggregation are similar.

Indeed, the large variation of the elementary index bias does not necessarily lead to large aggregate elementary index bias at higher levels of aggregation. Nevertheless, biases of the order of magnitude found in Table 6 are still large enough to matter.

5 Conclusion

This paper has provided estimates of elementary index bias at the narrow product category level in 10 countries. Of course, the level of the biases is specific to the products in this dataset and the particular period considered. Nevertheless, their general characteristics seem to be confirming what has been found in the earlier literature on more limited datasets. We also have new and striking results on the remaining effect of elementary index bias upon aggregation of multiple elementary indexes to the national or even euro area level. Some general conclusion can be drawn.

With respect to the results at the product category level, the elementary index bias, independently of the index formula, has been shown to be quite variable, across countries, as well as across product categories. A shift towards the use of cost of living indexes at the product category level, from currently Dutot or Jevons indexes in official practices, would therefore have non-negligible effects on measured product category inflation, potentially of multiple percentage points. A comparison of price only indexes, such as Dutot, Jevons and Carli, with indexes that use both price and quantity information shows that weighting reduces the variability of the bias quite substantially. From the viewpoint of variability, price only elementary indexes are certainly inferior to indexes that use both price and quantity information. This seems to support the application of scanner data, which combines price and quantity information, above price only sampling methods. The relatively large elementary index biases of the Laspeyres and Paasche indexes indicate that lower level substitution matters.

With respect to the results at the national level, aggregation shows that it is often the case that the aggregate elementary index bias is larger than the upper level substitution bias. Although much attention has been given to consumers substituting between different product categories, it seems that more attention is needed for within product category substitution.

As Silver (1995) states, *micro-indices are the building blocks of a CPI.* Good measurement of the building blocks is important to get reliable aggregate indices. The findings in this paper show that measurement of price change at the lowest level of aggregation is sensitive to the elementary index choice. One question that remains is the variability and level of the bias for services. For services no scanner data is available, however, these form a substantial part of the consumer price index. Obtaining jointly price and quantity information of services to perform a similar exercise will certainly not be easy. However, taking into account the importance of elementary indexes and the variability of the biases considered here, it seems that further analysis for a wider set of products and services is certainly warranted and could lead to new insights into the relevance of these biases in our official price indexes.

References

[1] Boskin, Michael J., Dulberger Ellen R., Gordon Robert J., Griliches, Zvi, and Dale W. Jorgenson, 1998, “Consumer prices, the consumer price index, and the Cost of living”, *Journal of Economic Perspectives, Vol 12, N1* 3-26.

[2] Carruthers A.G., Sellwood D.J.and P.W. Ward, 1980,“Recent developments in the retail price index”, *The Statistician, vol 29 n1* 1-32.

[3] Dalén, Jorgen, 1994,“Sensitivity analyses for harmonising european consumer price indices”, *Paper presented at the International Conference on Price Indices, 1994, Ottawa, Statistics Canada.*

[4] Dalén, Jorgen, 1997,“Experiments with Swedisch Scanner data”, *Paper presented at the International Conference on Price Indices, 1997, Voorburg, The Netherlands.*

[5] Dalton Kenneth V., Greenlees, John S., Steward, Kenneth J., 1998, “Incorporating a geometric mean formula into the CPI”, *Monthly Labor Review* 3-7.

[6] Diewert, W.E., 1976, “Exact and superlative index numbers”, *Journal of Econometrics, 4*, 115-145.

[7] Diewert, W.E., 1978, “Superlative index numbers and consistency in aggregation”, *Econometrica, 46*, 883-900.

[8] Diewert, W.E., 1995, “Axiomatic and Economic Approaches to Elementary Price Indexes”, NBER Working Paper No. 5104

[9] Diewert, W.E., 1998, “Index Number Issues in the consumer price index”, *Journal of Economic Perspectives 12,no1*, 47-58

[10] Haan, Jan de, Opperdoes Eddy, 1997, “Estimation of the Coffee Price index Using scanner data: the choice of the Micro index”, *Paper presented at the International Conference on Price Indices, 1997, Voorburg, The Netherlands*

[11] Hawkes, William J., 1997, “Reconciliation of consumer price index trends with corresponding trends in average prices for quasi-homogeneous goods using scanner data,” *Paper presented at the International Conference on Price Indices, 1997, Voorburg, The Netherlands*

[12] Ivancic, Lorraine, W.E. Diewert and K.J. Fox, 2011, “Scanner data, time aggregation and the construction of price indexes”, *Journal of Econometrics, 16 24-35.*

[13] ILO, Consumer price index manual: theory and practice.

[14] Jain Malti, Ruel Abello, 2001, "Construction of price indexes and exploration of biases using scanner data", *Room document at the Sixth Meeting of the International Working Group on Price Indices, Canberra, Australia 2-6 April, 2001.*

[15] Moulton Brent R., Karin E. Smedley, 1995, "A comparison of estimators for elementary aggregates of the CPI", Bureau of Labor Statistics, June 1995.

[16] Reinsdorf, Marshall B., 1999 "Using Scanner Data to Construct CPI Basic Component Indexes", *Journal of Business and Economics Statistics, April 152-160.*

[17] Richardson David H., 2000,"Scanner Indexes for the CPI", *Paper presented at the Sixth Meeting of the International Working Group on Price Indices, Canberra, Australia, 2-6 April 2001.*

[18] Rotaru Christian, McNaughton Shaun and Soriano Franklin, 2011,"Revisiting the estimation of substitution biases and the construction of spatial price indexes using scanner data", *Paper presented at Economic Measurement Group Workshop Decembber 2011*

[19] Schultz Bohdan, 1995,"Choice of price index formulare at the micro-aggregation level: The Canadian Empirical evidence", Version updated in June 1995, Paper presented at the first meeting of the Ottawa Group on Price indices, October 31-November 2, 1994.

[20] Silver, Mick, 1995, "Elementary Aggregates, micro-indies and scanner data: some issues in the compilation of consumer price indices" *Review of Income and Wealth, December 427-438.*

[21] Triplett, Jack, 1998, "Elementary indexes for a consumer price index", Paper presented at the fourth meeting of the Ottawa Group on Price indices.

Appendix A: Indices at the euro area level

Definitions of the symbols used

t_0 : time t-zero

t : time t

N_i : the number of Stock Keeping Units in product category i.

p_{int} : the price of the n-th Stock Keeping unit in product category i in month t

p_{int_0}: the price of the n-th Stock Keeping unit in product category i in month t-zero

q_{int_0}: the quantity of the n-th Stock Keeping unit in product category i in month t-zero

q_{int} : the quantity of the n-th Stock Keeping unit in product category i in month t

Direct comparison Indices between base month t_0 (i.e. December 2009) and month t (i.e. December 2010)

Let

C: the number of countries

e: indicates euro area

I_{ct}: the direct comparison index at the national level (Laspeyres or Lowe aggregation).

Note that, in principle, for the euro area indexes below, one should write p_{int_0} , q_{int_0} , q^s_{ink} and so on, with a country subscript "c", i.e. as p_{cint_0}, q_{cint_0} , q^s_{cink}. However, to safe on notation, the country subscript "c" is understood to be there implicitly in the formulas below.

Euro area Lowe aggregation index

$$LOWE^a_{et} = \sum_c^C (I_{ct}) * \left(\frac{\sum_i^K \sum_n^{N_i} p_{int_0} * q^s_{ink}}{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int_0} * q^s_{ink}} \right) \tag{18}$$

with

$$q^s_{ink} = \sum_{j=0}^{11} q_{in(t_0-j)} \tag{19}$$

where q^s_{ink} are the 12 month quantities of individual product n (i.e the sum of quantities of months $t_0, t_0 - 1, ...t_0 - 11$

Euro area Laspeyres aggregation index

$$LASP^a_{et} = \sum_c^C (I_{ct}) * \left(\frac{\sum_i^K \sum_n^{N_i} p_{int_0} * q_{int_0}}{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int_0} * q_{int_0}} \right) \tag{20}$$

euro area Laspeyres aggregate Index

$$LASP_{et} = \frac{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int} * q_{int_0}}{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int_0} * q_{int_0}} = \sum_c^C \sum_i^K \sum_n^{N_i} \left(\frac{p_{int}}{p_{int_0}} \right) * \left(\frac{p_{int_0} * q_{int_0}}{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int_0} * q_{int_0}} \right) \tag{21}$$

euro area Paasche aggregate index

$$PAASCH_{et} = \frac{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int} * q_{int}}{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int_0} * q_{int}} = \left[\sum_c^C \sum_i^K \sum_n^{N_i} \left(\frac{p_{int_0}}{p_{int}} \right) * \left(\frac{p_{int} * q_{int}}{\sum_c^C \sum_i^K \sum_n^{N_i} p_{int} * q_{int}} \right) \right]^{-1} \tag{22}$$

euro area Fisher (ideal) aggregate index:

$$FISH_{et} = \sqrt{LASP_{et} * PAASCH_{et}} \tag{23}$$

Appendix B: List of product categories and COICOP classification

The product categories available in the dataset are quite narrowly defined. To get an idea of the coverage, the list below shows a matching between corresponding categories according to the COICOP classification and the respective product categories available in the dataset.

- 0111 Bread and Cereals : rice, cereal, dry pasta
- 0113 Fish and Seafood : tinned tuna, frozen fish
- 0114 Milk, cheese and eggs : refrigerated milk, UHT milk
- 0115 Oils and Fats : margarine, olive oil
- 0117 Vegetables : tinned peas, peas frozen
- 0118 Sugar, jam, honey, chocolate and confectionery : jam strawberry, ice cream, sugar, chewing gum
- 0119 Food products, n.e.c: wet soups, baby food, bouillon
- 0121 Coffee, tea and cocoa: Ground coffee, Instant coffee
- 0122 Mineral waters, soft drinks, fruit and vegetable juices: sparkling water, still water, 100 percent orange juice, carbonated soft drinks
- 0211 Spirits: vodka ,whiskey

- 0213 Beer: beer
- 0561 Non-durable household goods: all purpose cleaner, automatic dishwasher detergent, fabric softenen, laundry detergent
- 0612 Other medical products: condoms
- 0934 Pets and related products: dog food, cat food
- 1213 Other appliances, articles and products for personal care: deodorant, paper towels, shampoo, shave preps, toilet tissue, toothpaste, pantyliner, Diapers

Appendix C: Elementary index bias at the product category level: individual product results

Table 7: Elementary index bias: product category results

	CARLI	DUTOT	JEVONS	LASP	LOWE	PAAS	JEVEW	GLOWE
100pc_juice_rev								
mean	-1.13	-1.68	-1.42	0.04	-0.35	-0.02	-0.19	-0.57
sd	2.38	2.76	2.41	1.55	1.93	1.58	1.71	2.05
Diapers								
mean	0.19	0.40	-0.31	0.10	0.11	-0.10	0.04	0.03
sd	1.57	1.66	1.95	0.25	0.26	0.25	0.20	0.26
Ground_coffee								
mean	-1.85	-2.86	-2.08	0.38	0.36	-0.38	0.25	0.22
sd	2.15	3.00	2.28	0.45	0.36	0.45	0.36	0.30
Instant_coffee								
mean	0.22	0.30	0.03	0.53	0.35	-0.52	0.31	0.13
sd	1.74	2.38	1.84	0.74	0.54	0.73	0.56	0.45
apc								
mean	1.04	-0.20	0.08	1.83	0.26	-1.74	0.82	-0.61
sd	2.02	3.70	2.47	2.82	2.40	2.63	1.34	3.05
auto_dish_det								
mean	3.69	2.72	2.84	0.03	0.11	0.02	-0.21	-0.12
sd	4.05	3.55	3.47	2.43	2.58	2.48	2.43	2.57
baby_food								
mean	0.23	-0.17	-0.03	0.05	-0.04	-0.05	-0.09	-0.17
sd	1.28	2.01	1.52	0.10	0.24	0.10	0.28	0.43
beer								
mean	0.53	0.18	0.08	0.51	0.43	-0.50	0.36	0.28
sd	2.90	3.54	2.73	0.79	0.73	0.77	0.65	0.61
bouillon								
mean	-0.05	-0.57	-0.23	0.09	0.20	-0.09	-0.02	0.09
sd	1.42	2.34	1.45	0.26	0.41	0.26	0.30	0.31
cat_food								
mean	0.21	0.42	0.03	0.54	0.63	-0.53	0.39	0.48
sd	1.36	1.48	1.51	1.00	1.06	0.97	0.85	0.89
cereal								
mean	-0.41	-0.94	-0.75	0.22	0.04	-0.22	0.11	-0.07
sd	1.72	1.89	1.96	0.20	0.30	0.20	0.17	0.33
chewing_gum								
mean	-0.30	-0.44	-0.49	0.06	-0.04	-0.06	-0.05	-0.15
sd	0.84	0.93	0.90	0.19	0.22	0.19	0.13	0.20
condoms								
mean	0.27	0.43	-0.23	0.30	0.28	-0.30	0.11	0.10
sd	2.70	2.79	2.47	0.43	0.54	0.42	0.23	0.41
csd								
mean	0.39	-0.20	-0.14	0.41	0.27	-0.40	0.21	0.10
sd	2.52	2.74	2.45	0.52	0.57	0.51	0.41	0.54
deodorant								
mean	-1.83	-2.48	-2.28	0.88	0.05	-0.85	0.34	-0.50
sd	6.85	7.35	7.35	1.63	0.30	1.56	0.73	1.13
dog_food								
mean	0.68	0.69	0.30	0.64	0.57	-0.63	0.35	0.29
sd	2.12	3.55	1.78	1.06	1.20	1.02	0.82	0.98
dry_pasta								
mean	-1.39	-1.88	-1.83	0.30	-0.41	-0.30	0.04	-0.66
sd	5.10	5.59	5.04	0.27	1.12	0.27	0.28	1.23
fabric_soft								
mean	1.12	0.79	0.33	1.32	0.27	-1.26	0.63	-0.42
sd	3.11	2.79	2.96	2.03	2.05	1.98	1.43	1.79
frozen_fish								
mean	1.63	1.46	1.00	1.89	0.36	-1.83	1.38	-0.16
sd	2.87	2.80	2.60	1.79	1.23	1.71	1.29	1.36
ice_cream								
mean	0.96	0.75	0.75	0.40	0.39	-0.40	0.28	0.25
sd	2.38	3.32	2.41	0.53	0.77	0.52	0.44	0.76
jam_strawberry								
mean	-0.05	-0.42	-0.23	0.34	-0.04	-0.33	0.18	-0.19
sd	3.13	3.68	3.15	0.57	0.49	0.56	0.44	0.59
laundry_detergent								
mean	-0.28	-0.35	-0.57	0.68	0.20	-0.67	0.53	0.02
sd	0.56	0.97	0.50	0.29	0.11	0.29	0.28	0.20
margarine								

Table 7 – *Continued from previous page*

	CARLI	DUTOT	JEVONS	LASP	LOWE	PAAS	JEVEW	GLOWE
mean	0.12	0.09	-0.12	0.13	0.08	-0.13	-0.07	-0.13
sd	1.10	1.46	1.23	0.22	0.11	0.21	0.39	0.27
milk_refr								
mean	0.13	0.05	0.05	0.05	-0.02	-0.05	0.00	-0.07
sd	1.02	1.25	1.01	0.10	0.13	0.10	0.09	0.15
milk_uht								
mean	1.13	-2.28	0.15	0.82	0.65	-0.80	0.44	0.27
sd	5.23	10.87	6.31	1.35	1.36	1.30	0.78	0.79
olive_oil								
mean	0.79	0.49	0.46	1.40	0.54	-1.33	1.16	0.31
sd	1.76	2.19	1.56	2.40	1.55	2.22	2.24	1.42
pantyliner_rev								
mean	2.19	1.24	1.30	0.58	0.37	-0.57	0.17	-0.07
sd	4.35	2.62	3.76	0.65	0.49	0.64	0.42	0.44
paper_towels								
mean	-1.62	-3.51	-2.57	1.55	0.97	-1.39	1.09	0.45
sd	3.93	4.99	4.41	4.15	3.06	3.73	3.85	2.63
peas_frozen								
mean	0.52	-0.13	-0.11	0.53	-0.32	-0.53	0.33	-0.69
sd	2.39	1.83	2.41	0.70	1.58	0.69	0.54	1.85
rice								
mean	0.90	1.25	0.51	0.37	-0.10	-0.36	0.15	-0.30
sd	2.81	2.03	3.10	0.68	0.36	0.67	0.53	0.39
shampoo								
mean	-0.10	-1.60	-1.01	1.41	0.59	-1.36	0.54	-0.35
sd	2.88	5.10	3.82	2.15	1.44	2.05	1.12	1.43
shave_preps								
mean	6.07	3.08	3.46	0.46	0.24	-0.45	-0.08	-0.21
sd	8.26	3.05	5.13	0.78	0.57	0.77	0.51	0.85
sugar								
mean	0.44	-0.14	0.05	0.05	-0.04	-0.05	-0.09	-0.21
sd	0.77	3.59	1.09	0.25	0.40	0.25	0.21	0.29
tinned_peas								
mean	0.29	-0.39	0.08	0.17	0.11	-0.17	0.04	-0.03
sd	1.84	1.19	1.84	0.16	0.23	0.16	0.08	0.18
tinned_tuna_rev								
mean	0.36	-0.19	-0.20	0.46	-0.05	-0.46	0.22	-0.33
sd	2.66	3.26	2.68	0.68	0.72	0.67	0.86	1.07
toilet_tissue								
mean	0.37	-0.05	-0.14	0.61	0.73	-0.60	0.38	0.44
sd	2.32	2.89	2.64	0.77	0.73	0.75	0.75	0.62
toothpaste								
mean	1.54	1.37	1.16	0.46	0.58	-0.46	0.10	0.20
sd	1.97	1.59	1.96	0.40	1.00	0.39	0.15	0.72
vodka								
mean	1.82	1.23	1.47	0.19	0.11	-0.19	0.11	0.02
sd	2.59	2.29	2.43	0.27	0.39	0.27	0.26	0.37
water_sparkling								
mean	0.52	0.71	0.24	0.28	0.14	-0.28	0.16	-0.00
sd	1.43	1.31	1.59	0.28	0.24	0.28	0.18	0.34
water_still								
mean	1.47	1.14	1.11	0.46	0.13	-0.45	0.20	-0.10
sd	1.32	2.27	1.41	0.46	0.73	0.46	0.32	0.87
wet_soups								
mean	1.00	1.46	0.51	1.17	0.55	-1.14	0.73	0.10
sd	2.75	4.10	2.95	1.52	1.23	1.47	1.15	0.80
whiskey								
mean	0.31	-0.03	0.07	0.28	0.03	-0.28	0.21	-0.05
sd	2.87	3.65	2.86	0.37	0.32	0.37	0.32	0.33

Appendix D: Measured inflation outcomes

Table 8: Measured inflation of national aggregates for different elementary indexes (percentage points)

	AT	BE	DE	ES	FR	GR	IE	IT	NL	PT	Total
					direct Fischer Ideal index						
	-0.51	-0.38	1.51	-0.13	-0.18	1.96	-3.48	-1.02	0.96	1.42	0.02
					Laspeyres aggregation of elementary indexes						
CARLI	0.85	2.05	2.07	0.94	-0.05	3.64	-3.33	-0.44	-0.82	0.56	0.55
DUTOT	0.46	2.29	1.42	0.37	-0.38	3.29	-3.79	-0.77	-1.97	-0.05	0.09
JEVONS	0.47	1.52	1.71	0.52	-0.19	3.34	-3.76	-0.74	-1.40	0.07	0.15
LASP	0.37	-0.10	1.94	0.10	0.04	3.45	-1.83	-0.58	1.49	1.76	0.66
LOWE	0.02	-0.16	1.62	-0.08	-0.04	3.21	-2.35	-0.91	1.44	1.58	0.43
PAAS	-1.14	-0.42	1.29	-0.20	-0.37	1.84	-3.77	-1.28	0.57	1.19	-0.23
JEVEW	0.10	-0.27	1.77	-0.04	-0.03	3.26	-2.17	-0.70	1.18	1.60	0.47
GLOWE	-0.26	-0.33	1.47	-0.21	-0.12	3.02	-2.67	-1.04	1.11	1.41	0.24
FISH	-0.39	-0.26	1.61	-0.05	-0.16	2.62	-2.82	-0.93	1.02	1.47	0.21
					Lowe aggregation of elementary indexes						
CARLI	0.90	2.22	1.77	0.78	-0.04	3.15	-2.54	-0.56	-0.63	0.99	0.60
DUTOT	0.49	2.45	1.22	0.13	-0.40	2.79	-3.12	-0.88	-1.72	0.37	0.13
JEVONS	0.52	1.68	1.41	0.33	-0.19	2.85	-3.04	-0.87	-1.20	0.49	0.20
LASP	0.41	0.17	1.59	-0.08	0.01	2.94	-0.86	-0.70	1.67	2.18	0.73
LOWE	0.09	0.11	1.23	-0.26	-0.09	2.68	-1.42	-1.02	1.62	2.02	0.50
PAAS	-1.04	-0.16	0.91	-0.37	-0.43	1.26	-2.96	-1.42	0.71	1.62	-0.19
JEVEW	0.14	0.01	1.42	-0.23	-0.07	2.76	-1.23	-0.83	1.37	2.01	0.53
GLOWE	-0.19	-0.06	1.07	-0.41	-0.17	2.49	-1.77	-1.16	1.30	1.86	0.30
FISH	-0.32	0.01	1.25	-0.23	-0.21	2.08	-1.92	-1.06	1.18	1.90	0.27

Table 9: Measured inflation at euro area level for different elementary indexes (percentage points)

	Euro Area
direct Fischer Ideal index	
	0.09
Laspeyres aggregation of elementary indexes	
CARLI	0.50
DUTOT	0.06
JEVONS	0.21
LASP	0.54
LOWE	0.32
PAAS	-0.13
JEVEW	0.40
GLOWE	0.18
FISH	0.20
Lowe aggregation of elementary indexes	
CARLI	0.43
DUTOT	0.00
JEVONS	0.13
LASP	0.45
LOWE	0.22
PAAS	-0.19
JEVEW	0.30
GLOWE	0.08
FISH	0.10

Appendix E: Laspeyres and Lowe weights of product categories in national indexes

Table 10: Laspeyres weights used for product category aggregation

product	AT	BE	DE	ES	FR	GR	IE	IT	NL	PT	Mean
100pc_juice_rev	2.48		1.98		2.78	4.26	1.84	2.26	3.31	0.21	2.39
Diapers	0.06	0.12			0.21			1.88	0.47		0.55
Ground_coffee	4.64	5.74	13.15		5.33	3.70	0.45	10.30	13.34	5.32	6.88
Instant_coffee	1.69	0.56	2.78		2.37	7.85	2.31	0.48	1.41	2.51	2.44
apc	0.71	0.23	0.69	4.12		0.57			0.62	0.07	1.00
auto_dish_det	1.27	0.21	0.48	2.75	0.54			0.47	0.86	0.45	0.88
baby_food	0.13	0.65	1.14		0.49	1.68	0.72	8.05		0.97	1.73
beer	16.95	20.38	22.48		6.79	4.69	18.82	4.18	28.03	11.66	14.89
bouillon	1.80	0.80	0.75	8.63	0.87	0.40	0.33	2.79	0.34	2.15	1.89
cat_food	2.38	2.61	0.93			0.64	0.62		0.91	1.80	1.41
cereal	0.42	1.25	0.43	3.96	0.65	2.59	4.52	1.81	0.59	2.35	1.86
chewing_gum	0.56	0.80	1.87	18.63	1.97	0.70	2.62	4.59	0.61	0.92	3.33
condoms	0.77	0.04	0.28	1.59	0.34	0.65		0.32	0.16	0.26	0.49
csd	7.85	11.58	9.71		9.95	9.76	9.63	10.91	14.56	5.48	9.94
deodorant	1.14	1.16	0.89	7.16	0.97	1.05	0.51	0.99		1.80	1.74
dog_food	1.20	1.55	0.62			0.42	1.36		1.39	0.83	1.06
dry_pasta	0.72	0.90	0.53		0.70	1.72	0.26		0.71	0.65	0.77
fabric_soft	0.82	1.10	1.57	1.97	0.87	0.90	0.64	0.87		1.15	1.10
frozen_fish	2.12	0.76	2.31		0.42		1.33	1.50	3.68	2.20	1.79
ice_cream	0.57	2.01	0.84		0.63	0.18	0.86	0.19	0.92	0.51	0.75
jam_strawberry	0.60	0.64	0.81	2.24	0.48	0.32	0.27	0.17	0.77		0.70
laundry_detergent			0.88					1.32		1.59	1.27
margarine	4.51	4.91	6.98	8.46	4.60	5.90			1.90	4.17	5.18
milk_refr	5.35	0.69	2.32			13.54	13.39	6.64	4.51	1.41	5.98
milk_uht	1.04	4.48	1.04		10.78	0.89		10.53	3.65	15.37	5.97
olive_oil	1.34	1.30	0.71		2.67	6.07	0.49	4.52	0.81	4.97	2.54
pantyliner_rev	0.60	0.49	0.87			0.70	0.16			1.88	0.78
paper_towels	0.63		0.58	1.98		1.28	0.08	2.67	0.60	0.74	1.07
peas_frozen	0.72	0.18					0.50	1.66	0.06	0.78	0.65
rice	1.81	0.95	0.84		1.40	2.91	0.71	0.93	1.00	4.70	1.69
shampoo	1.05	0.51	2.03	8.46		1.82	0.25	1.51	0.31	2.58	2.06
shave_preps	0.21	0.47	0.23	1.82		0.34		0.28	0.54		0.55
sugar	3.73	4.85	2.20		2.16		1.24	2.61	2.84		2.81
tinned_peas	0.09	0.31	0.18		0.61		0.60	0.67	0.59		0.44
tinned_tuna_rev	2.88	0.84	0.47		1.73	3.21	0.40	3.07	1.51	3.64	1.97
toilet_tissue	8.35	1.44	1.22	4.49	4.48	3.67	2.36	5.08	3.41	1.51	3.60
toothpaste	3.26	0.95	1.80	6.12	1.30	1.21		1.62	3.11	1.73	2.35
vodka	2.30	2.40	2.54		3.03	2.19	11.62	0.47		0.69	3.15
water_sparkling	5.64	4.83	3.82		6.51	1.10	0.85			3.20	3.71
water_still	3.19	10.92	2.71		8.90	4.09	4.47			3.07	5.34
wet_soups	2.30		0.48	17.62	1.81		1.22	0.65	2.48	1.00	3.45
whiskey	2.11	6.39	3.82		13.67	9.00	14.57	4.01		5.69	7.41

Table 11: Lowe weights used for product category aggregation

product	AT	BE	DE	ES	FR	GR	IE	IT	NL	PT	Mean
100pc_juice_rev	2.28		2.05		2.69	4.53	2.11	3.11	2.98	0.16	2.49
Diapers	0.09	0.11			0.22			1.98	0.44		0.57
Ground_coffee	4.71	6.15	11.32		5.30	3.07	0.47	7.61	14.24	4.89	6.42
Instant_coffee	1.69	0.57	2.60		2.38	9.02	2.48	0.48	1.45	2.31	2.55
apc	0.64	0.24	0.71	5.22		0.62			0.68	0.07	1.17
auto_dish_det	1.31	0.18	0.41	2.98	0.52			0.55	0.94	0.57	0.93
baby_food	0.15	0.63	1.15		0.57	1.51	0.77	8.54		1.13	1.81
beer	18.82	21.06	23.76		9.01	7.26	19.12	5.71	26.96	13.57	16.14
bouillon	1.59	0.57	0.63	7.38	0.66	0.33	0.27	2.37	0.28	1.94	1.60
cat_food	2.41	2.30	0.87			0.59	0.77		0.86	1.82	1.37
cereal	0.47	1.30	0.44	4.40	0.64	3.17	5.49	2.26	0.68	2.77	2.16
chewing_gum	0.70	0.99	1.99	17.89	2.14	0.66	2.87	4.74	0.80	0.93	3.37
condoms	0.76	0.04	0.30	1.70	0.30	0.63		0.34	0.17	0.26	0.50
csd	6.35	11.97	9.26		8.71	8.35	11.93	9.81	14.18	5.09	9.52
deodorant	1.24	1.31	0.88	8.36	1.09	1.19	0.68	1.27		1.77	1.98
dog_food	1.20	1.53	0.58			0.41	1.42		1.36	0.82	1.05
dry_pasta	0.85	0.88	0.56		0.74	1.68	0.28		0.81	0.61	0.80
fabric_soft	1.11	1.24	2.09	3.01	0.79	1.09	0.81	0.99		1.76	1.43
frozen_fish	1.70	0.93	3.05		0.50		1.52	1.69	4.79	2.26	2.05
ice_cream	1.84	2.32	1.16		1.31	1.48	1.88	0.86	0.95	1.10	1.43
jam_strawberry	0.62	0.81	0.82	2.32	0.52	0.35	0.34	0.22	0.65		0.74
laundry_detergent			1.04					1.67		2.31	1.67
margarine	3.64	4.95	6.53	8.53	4.35	5.23			1.77	3.82	4.85
milk_refr	5.49	0.74	2.49			13.78	12.94	6.83	4.35	1.37	6.00
milk_uht	1.29	3.76	1.09		10.46	0.79		10.48	3.21	14.64	5.71
olive_oil	1.24	1.34	0.69		2.58	6.27	0.45	4.87	0.81	3.49	2.41
pantyliner_rev	0.58	0.63	0.90			0.65	0.21			1.88	0.81
paper_towels	0.88		0.57	1.98		1.19	0.40	2.85	1.00	0.65	1.19
peas_frozen	0.57	0.14					0.72	1.42	0.07	0.79	0.62
rice	1.65	0.91	0.81		1.50	2.38	0.74	0.94	1.00	3.68	1.51
shampoo	1.10	0.49	2.02	10.27		1.82	0.34	1.59	0.36	3.01	2.34
shave_preps	0.21	0.43	0.21	1.93		0.29		0.31	0.48		0.55
sugar	3.33	4.61	2.50		2.00		1.13	2.07	2.48		2.59
tinned_peas	0.08	0.30	0.16		0.63		0.56	0.74	0.57		0.44
tinned_tuna_rev	3.41	1.22	0.38		3.54	3.13	0.62	3.76	1.46	5.09	2.51
toilet_tissue	7.14	1.66	1.23	4.16	4.83	3.60	2.72	5.77	4.44	1.38	3.69
toothpaste	3.43	1.07	1.73	6.64	1.34	1.10		1.87	2.92	1.81	2.44
vodka	1.91	1.79	2.50		2.29	1.94	10.10	0.43		0.69	2.71
water_sparkling	7.10	5.03	4.08		6.18	1.00	1.03			3.06	3.93
water_still	3.80	11.33	3.07		10.12	4.81	6.12			4.22	6.21
wet_soups	1.49		0.41	13.22	1.27		1.09	0.69	1.88	0.95	2.62
whiskey	1.12	4.46	2.97		10.81	6.08	7.64	1.20		3.32	4.70