STATISTICS DENMARK

20 March 2018
Martin Nielsen
Prices and consumption

Abstract

This paper investigates the results of using different quality adjustment methods when calculating consumer price indexes for new cars based on Danish data. The consumer price index for cars is calculated with two hedonic methods and a more traditional method, which is the overall mean imputation method. The two hedonic methods that will be used, is the characteristics methods and the rolling time dummy method. The index is calculated for the period June 2013 to April 2017 on data from a tax register containing all sales of new cars to consumers in Denmark. The data set contains information on a wide set of characteristics for the cars including mileage, engine size, safety measures and car brand. The paper presents the resulting price indexes when using the different quality adjustment methods and give recommendations regarding the use of different methods based on the empirical findings.

Introduction

The paper presents the EU Harmonized Index Consumer Price (HICP) and the EU Harmonized Index Consumer Price Constant Taxes (HICP-CT) for new gasoline and diesel cars, calculated with different methods, for the period June 2013 to April 2017. Because the HICP is harmonized across the European countries, it is possible to use the HICP to compare the consumer prices between countries. The HICP-CT estimates how the prices would have developed for a given year if the taxes for a country not have changed in a given year. In other words, the HICP-CT shows how the index for a given year would be, if it was the tax rates from December the year before that given year that still was applicable.

The reminder of the paper is organized as follows. First the simple index calculated with the Overall Mean Imputation method (OMI) is shown and compared with the official Danish index for new cars. Then the two hedonic indexes is shown and compared with the Danish official index. Then the different calculated price indexes will be compared and lastly a concluding remark.

Background

Before going through the different price indexes, there are a couple of important general things.

First off all, the price indexes will be calculated for the period June 2013 to April 2017 and all the price indexes are only for gasoline and diesel cars. The reason for not including electric cars is because they not yet make out a large enough part of the car market. Another reason is that electric cars probably should be an individual price index, because there are so many differences between gasoline/diesel cars and electric cars.

The official price index for new cars, was for the time period investigated in this paper, calculated on a sample of approximately 30 prices for different cars. The sample was not divided into segments (more about segments later), which resulted in an overrepresentation of small cars. Because the registration laws in

Denmark affect small and large cars differently, the overrepresentation of small cars can therefore have had some unintended consequences.

From February 2018 the price index for new cars is divided up in six segments. The segments are a standard division of cars. The six segments are mini cars, small cars, medium cars, large cars, SUV cars and MPV cars. The division in the six segments, are also the segmentation that will be used in the price indexes calculated in this paper.

The data used for making the different price indexes in this paper, is data from a Danish tax register for the period June 2012 to April 2017. The Danish tax register contains information on all sold cars, busses, motorcycles etc. because there in Denmark is a registration tax for every vehicle. The car dealers report information needed to calculated the registration tax when they sell a car or another vehicle, and it is that information that is used in the calculation of the price indexes in this paper.

Before any actual calculation is done, the data is limited only to hold information for cars sold to private persons and then the data is "cleaned" - unrealistic observations are removed. The data hold the following information; registration month, registration year, consumer price, brand, model, variant, engine size, maximum speed, mileage, number of airbags, number of stars in the NCAP crash test weight, number of seats, number of doors, if the car have ABS, ESP, radio or tow bar and fuel type. From the different information in the data, it is possible to calculate what the selling price would have been, if it was sold by the rules applicable in December the year before, that price is called the CT price and is used when calculating the HICP-CT.

Overall mean imputation method

When calculating the HICP and HICP-CT with the Overall Mean Imputation (OMI) method, the average price is calculated for every car every month, where a car is defined on the basis of the brand, model and variant. Between every two months, all the cars there are in both months, is matched and divided into the six segments. When calculating the change in prices between two months, all the price changes for the cars within a segment, are equally weighted together to one price change for each segment. The total price change between two months, are the price change for each segment weighted together with turnover weights for the previous year for each segment.

The reason for using average prices for the cars, are because there can be sold a lot of cars with relative large price differences within the same brand, model and variant. It is not possible to address the reason for the price differences within the cars, because information about extra equipment (air-condition, rearview camera and so on) not is available in the data. The use of average prices can raise problems about an upward or downward bias, if there is a trend against people choosing more or less extra equipment over time.

In figure 1 the HICP calculated with the OMI method and the official HICP is shown.

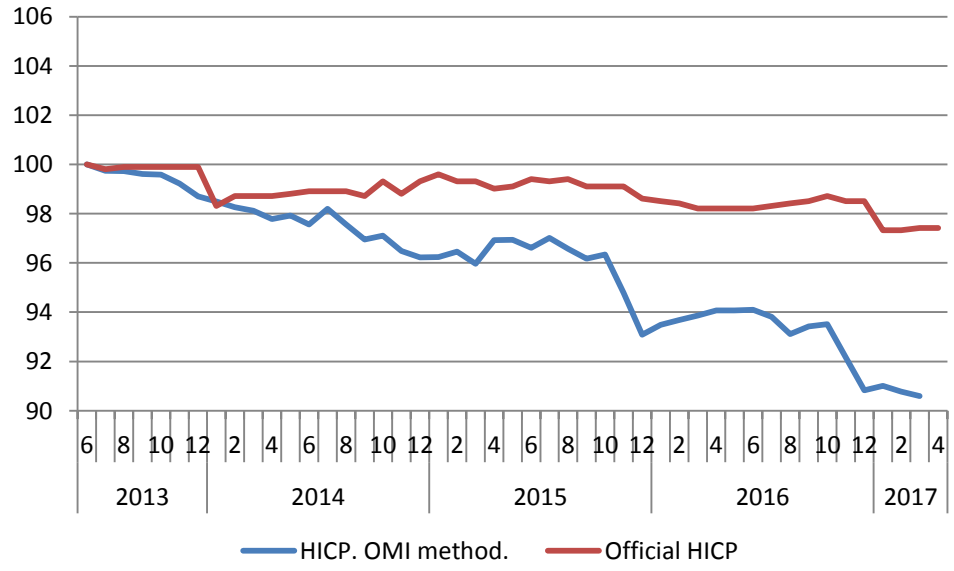
From figure 1 it is seen, that the index calculated with the OMI method at the end of the period is at least 6 index points under the official HICP. A difference of 6 index points is a relative big difference, especially because it only is over a

four year period. It is difficult to argue that the difference is due to a downward bias because of a trend with less extra equipment. Another possible problem with the OMI method, is that there can be a downward bias in the OMI index, because cars that is on the way out of the market, is in the index to the end, which can result in some clearance sale prices in the index. Another possible problem with the OMI index, are that price increases due to introduction of new models/variants, where the only change is a facelift, not is incorporated in the index, if there are any changes in the name of the car model or variant.

There off course also is the possibility that the OMI index gives the true index, and that the official index not have caught the full effect of the decrease of the registration taxes in the period. Because of the overrepresentation of small cars in the official index, it is possible that the decrease of the registration taxes not fully have been taking into account in the official index and therefore have overstated the price development.

The overall mean imputation method have a clear advantage compared to the official index, which is that the changes in the registration taxes in the end of 2015 and end of 2016 is shown clearly and at the time the rules are changed. It is also possible to see the effect of changes in the registration taxes in the official index, especially for the change in the end of 2016, but it takes a couple of months before the decrease of the registration taxes have been taking into account in the official index.

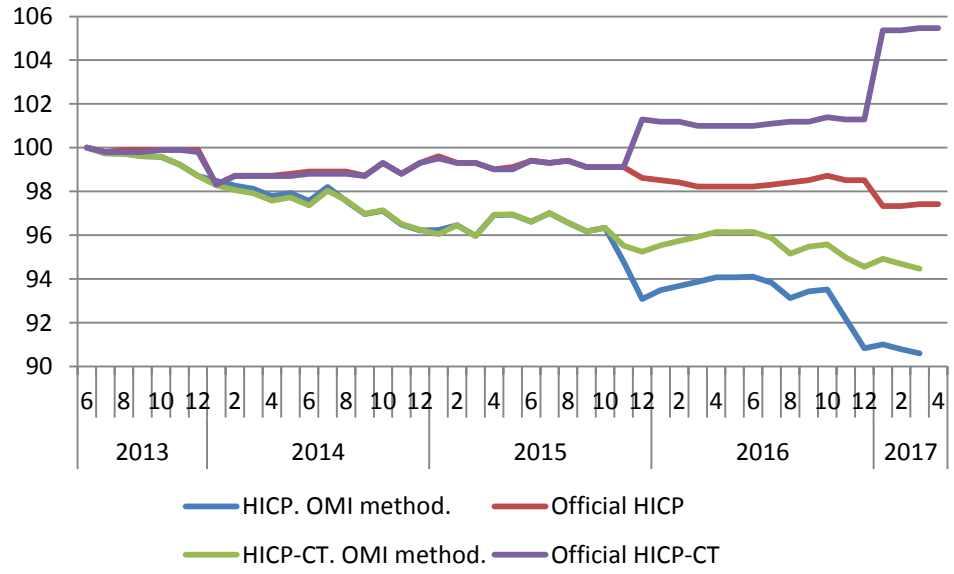
Figure 1: HICP index with overall mean imputation method and official HICP index (June 2013 = 100)



In figure 2 the HICP-CT are added to the two indexes from figure 1. From figure 2 it is seen that the connection between the official HICP and HICP-CT are pretty different compared to the connection between the OMI HICP and HICP-CT. When only looking at the connection between the HICP and HICP-CT, the OMI indexes, where changes in the registration taxes affect the HICP more than the HICP-CT, is probably more correct than the connection between the official HICP and HICP-CT, where the HICP-CT increases when there are changes in the registration taxes. The official HICP-CT is calculated on the basis of one price without taxes and vat equivalent to a medium sized car, which in itself not is ideal and the sample the official HICP index is calculated on had

some problems with an overrepresentation of smaller cars. Because the changes there have been in the registration taxes primarily affect medium and larger cars, it is likely that the official HICP index not fully have caught the effect of the changes in the registration taxes.

Figure 2: HICP and HICP-CT index with overall mean imputation method and official HICP and HICP-CT index (June 2013 = 100)



Hedonic index

There are used two different methods to calculate hedonic indexes. The two methods used are the characteristic method and the rolling time dummy method[1]. In both methods it is decided to use one year of data as basis for the models. At the moment the independent variables in the regression model is redefined every new year, where the dependent variable is the natural logarithm of the selling price. Which independent variable that are included in the model each year is decided by an automated procedure, more precise the SBC procedure. Every time the regression model is changed the different estimates and different test statistics are controlled. The yearly redefinitions of the model should not result in huge changes in the choice of explainable variables, unless there are some extraordinary changes in the market.

There have not been any extraordinary changes in the market in the time period analyzed in this paper, that can justify large changes in the choice of explanatory variables. The chosen explanatory variables in the models for the different year have also been really constant. The different explanatory variables in the regression models are listed in table 1.

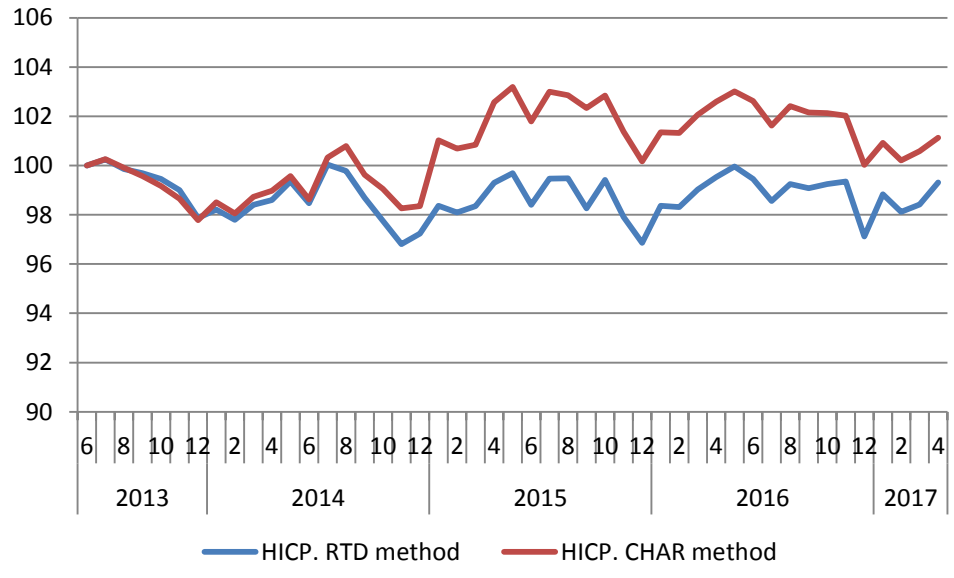
[1] For more about hedonic methods see eg. "Triplett, J. (2004), "Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products", *OECD Science, Technology and Industry Working Papers*, No. 2004/09, OECD Publishing, Paris. " and "de Haan, J. (2015, May). Rolling year time dummy indexes and the choice of splicing method. In *Room Document at the 14th meeting of the Ottawa Group, May* (Vol. 22)."

Table 1: Description of variables in the regression models

Variable	Variable description
Ny_model	Dummy variable that takes the value 1 if the car model is from the same or the previous year as the registration year.
Slagvolumen	Engine capacity
Maksimumhastighed	Maximum speed
KMPERLITER	Mileage. Affect the price directly because the registration tax depends on the mileage.
Sele	Number of seatbelt alarms. The registration tax is decreased with an amount for every alarm up to 3 alarms.
INTBA	Number of integrated child seats. Decreases the registration tax directly.
Seks_pladser	Dummy variable that takes value 1 if the car have six seats or more.
Airbag	Number of airbags. The registration tax is affected by the number of airbags.
Traek	Towbar
NCAPtest	Dummy variable that takes the value 1 if the car have received 5 stars in the international crash test. The registration tax is decreased if the car receives the maximum number stars in the crash test.
Budget	Takes the value 1 if the brand of the car is perceived as a cheap brand.
Praemium	Takes the value 1 if the brand of the car is perceived as an expensive brand.
Benzin	Takes the value 1 if it is a gasoline car.

Besides the variables in table 1 there is also five dummy variables for the six different segments and time dummy variables for the months. How many of the time dummy variables that are included in the model differs a little from year to year.

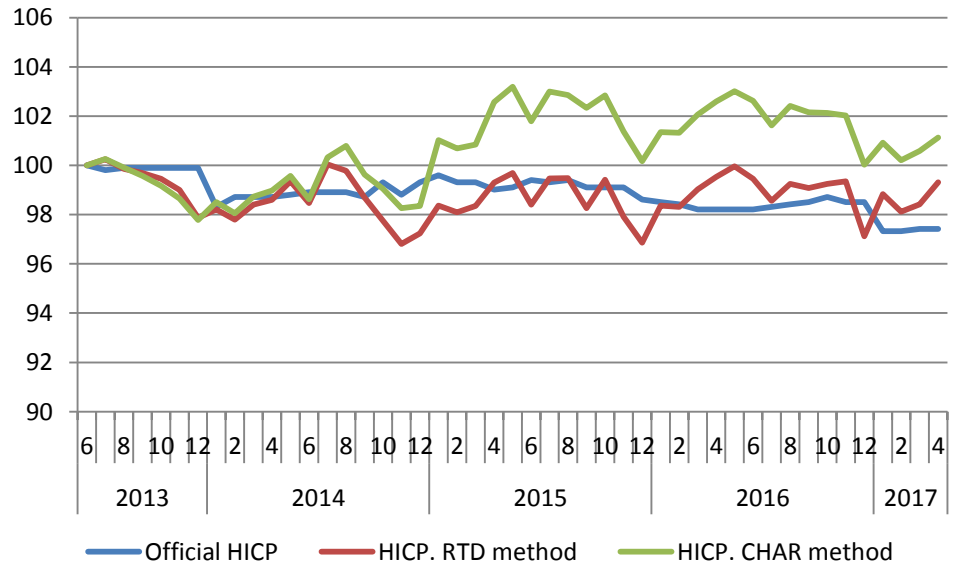
Figure 3: HICP indexes with the characteristic method (CHAR) and rolling time dummy method (RTD). (June 2013 = 100)



The hedonic HICP index calculated with the rolling time dummy method and the characteristic method is illustrated in figure 3. It is very positive that the two hedonic indexes are very identical besides two month in 2014. In the middle of 2014 the characteristic index increases a bit where the rolling time dummy index decreases a bit. In the end of 2014 both indexes increases but the characteristic index increases considerably more than the rolling time dummy index. Whether the rolling time dummy index or the characteristic index gives the best picture of the true price changes is very difficult to conclude, but there is nothing in the registration rules or anything else there can explain why the index should increase more than two index points at the end of 2014, why the rolling time dummy index is assumed to be more plausible than the characteristic index.

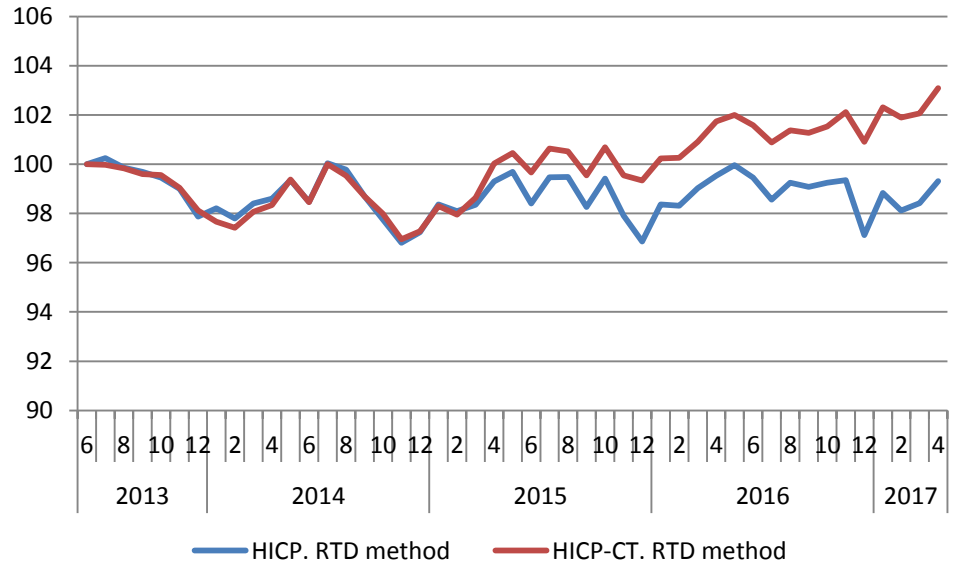
Is the official HICP index compared with the two hedonic indexes, which is done in figure 4, it is seen that the trend in the official index is close to the trend in the rolling time dummy index.

Figure 4: HICP indexes with the characteristic method (CHAR), rolling time dummy method (RTD) and the official HICP index. (June 2013 = 100)



In the following the focus will be on the rolling time dummy index within the hedonic indexes. The rolling time dummy HICP and HICP-CT index is illustrated in figure 5 (In appendix A the same can be seen for the characteristic method), where it clearly is seen that the HICP-CT index almost have the same fluctuations as the HICP index, except when there are changes in the registration taxes. There are small fluctuations in the HICP-CT index in the entire time period, but for the time period June 2012 to April 2017 there have not been a year without at least small changes in the registration taxes.

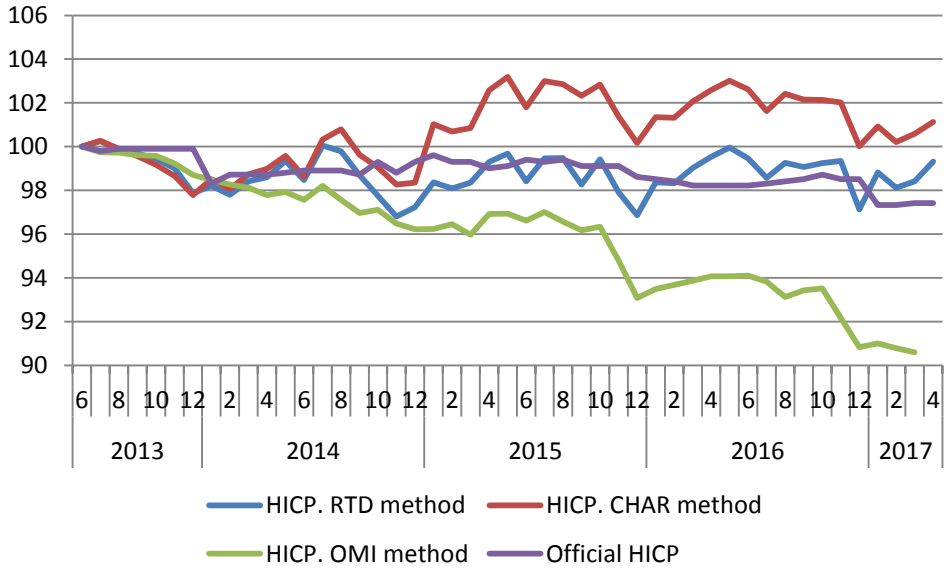
Figure 5: HICP and HICP-CT indexes with rolling time dummy method (RTD) (June 2013 = 100)



Comparison of trends

When comparing the trends the focus is only on the HICP indexes. In figure 6 the four HICP indexes calculated with different methods are illustrated. The index calculated with the characteristic method is the only index that gives an increasing trend. The OMI index gives a really negative trend compared to the other indexes. When commenting on the OMI index it is important to remember that there is a risk that the OMI index have a severe downward bias. The rolling time dummy index have a trend that is more or less identical with the official index, not that the official index necessarily is the truth. As earlier commented the official index have a problem with overrepresentation of small cars, which have the effect that the decrease of the registration rules not fully have been integrated in the index, which leads to an possible upward bias in the official index. On the other hand the rolling time dummy index indicates that the manufactures of the cars, just have raised the prices after the decrease of the registration taxes. The OMI index actually also show small increases in prices shortly after the decrease of the registration taxes.

Figure 6: HICP indexes with different methods (June 2013 = 100)



Focusing on the rolling time dummy index and the OMI index in figure 6, it is clear that the two indexes mainly have identical fluctuations, one clear difference is, that the increases in the rolling time dummy index are larger compared to the increases in the OMI index.

Because the OMI index is so far from the other indexes it would be interesting to make an index, where a number of cars is chosen in the beginning and then every new year redefine the chosen cars. Besides the redefinition of the cars every new year, there should throughout the year be a manual substitution of cars leaving the market. An index with manually substitutions would be interesting because it could remove the problem with clearance sale prices in the OMI index.

Conclusion

Because of the possible severe downward bias in the OMI index, the rolling time dummy index is at the moment assumed to be the best index of the ones tested in this paper. The rolling time dummy index is assumed to be better than the characteristic index, because there not are any large fluctuations, where there not are changes in the registration taxes or other things in the market that can explain large fluctuations.

Even though the official index more or less have the same trend as the rolling time dummy index, the rolling time dummy index is seen as a better index, because the sample for the rolling time dummy index almost covers the total amount of sold cars in the period and it also make an quality correction.

Appendix

A

HICP and HICP-CT indexes with characteristic method (CHAR) (June 2013 = 100)