WIFO

ÖSTERREICHISCHES INSTITUT
FÜR WIRTSCHAFTSFORSCHUNG

WORKING PAPERS

Direct Versus Indirect Approach in Seasonal Adjustment

Marcus Scheiblecker

460/2014

WIFO
ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG
AUSTRIAN INSTITUTE OF ECONOMIC RESEARCH

Direct Versus Indirect Approach in Seasonal Adjustment

Marcus Scheiblecker

WIFO Working Papers, No. 460

January 2014

Abstract

With seasonal adjustment one has to decide whether to seasonal adjust an aggregate like GDP directly or to sum up its seasonally adjusted components. This choice is usually driven by subjective motives or practical convenience. In the case of seasonal adjustment with chain-linked data one might feel forced to use the direct approach as components do not even add up to aggregates before the adjustment. This paper presents a guide for practitioners, which recommends a more objective way of decision-making, based on several indicators. It proposes some of these criteria which can facilitate the decision between the direct and the indirect approach. For the case of chain-linked series, where the indirect approach seems not to be feasible because components are not adding up to an aggregate, the paper presents a method how the indirect approach of seasonal adjustment nevertheless can be applied. Finally it deals with a possible balancing process between the results of the direct and the indirect approach and a practical application example is given.

E-mail address: Marcus.Scheiblecker@wifo.ac.at

2014/017/W/0

Medieninhaber (Verleger), Hersteller: Österreichisches Institut für Wirtschaftsforschung • 1030 Wien, Arsenal, Objekt 20 • Tel. (43 1) 798 26 01-0 • Fax (43 1) 798 93 86 • http://www.wifo.ac.at/ • Verlags- und Herstellungsort: Wien
Die Working Papers geben nicht notwendigerweise die Meinung des WIFO wieder
Kostenloser Download: http://www.wifo.ac.at/wwa/pubid/47115

Direct versus indirect approach in seasonal adjustment

Marcus Scheiblecker

Austrian Institute of Economic Research (WIFO)

Arsenal Objekt 20

1030 Vienna, Austria

e-mail: marcus.scheiblecker@wifo.ac.at

January 2014

Abstract

By doing seasonal adjustment one has to decide frequently whether to seasonal adjust an aggregate like GDP directly or to sum up its seasonally adjusted components. This choice is usually driven by subjective motives or practical convenience. In the case of seasonal adjustment with chain-linked data one might feel forced to use the direct approach as components do not even add up to aggregates before the adjustment.

This paper presents a guide for practitioners, which recommends a more objective way of decision-making, based on several indicators. Chapter 1 proposes some of these criteria which can alleviate the decision between the direct and the indirect approach. For the case of chain-linked series, where the indirect approach seems not to be feasible because components are not adding up to an aggregate, the second chapter presents a method how the indirect approach of seasonal adjustment nevertheless can be applied. The third chapter deals with a possible balancing process between the results of the direct and the indirect approach. In the forth chapter a practical application example is given. Conclusions are drawn in the fifth chapter.

JEL codes: C22, C80

Keywords: Seasonal adjustment, direct indirect method, chain-linking

Direct versus indirect approach in seasonal adjustment

1. Criteria for choosing the direct or indirect approach

Usually seasonal adjustment methods (or at least the preprocessing in order to detect outliers, calendar effects or to extend the series length) are based on time series models. In contrast to other econometric models, more often several models at the same time show satisfying properties which makes it difficult to discriminate between them.

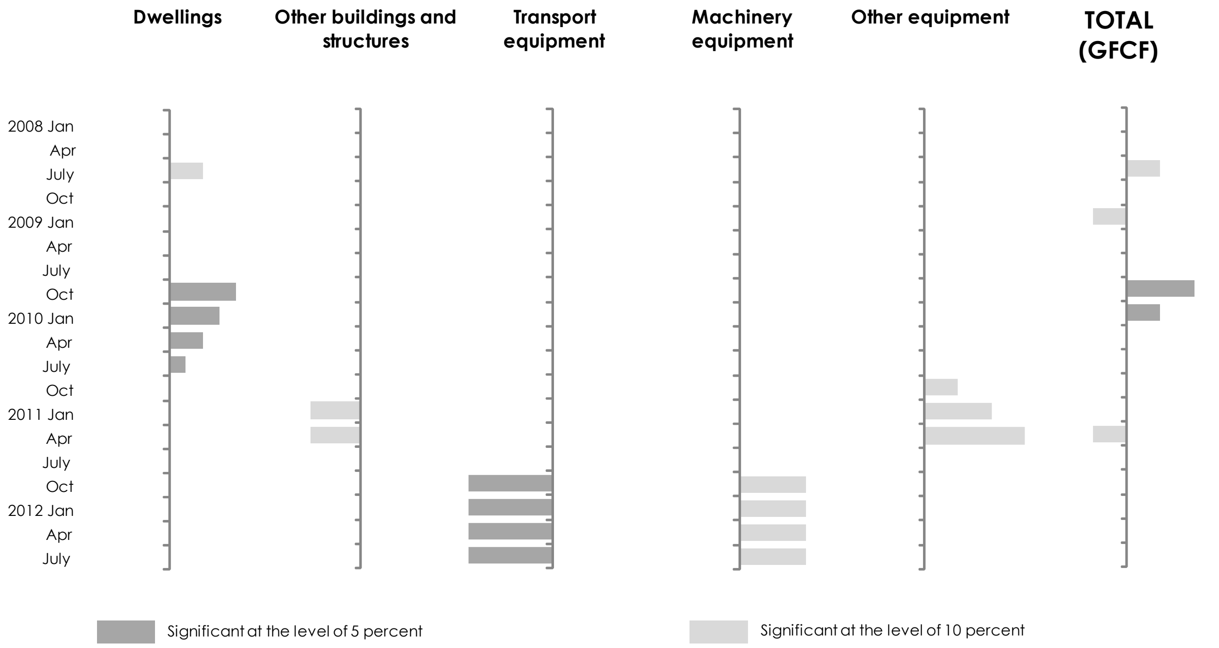
In this case, information which is not included in the model can be used as a guide for discrimination. Such external information can be a prior of sign and size of seasonality, calendar effects or outliers. Clearly, these priors can be formed better for components than for the aggregate. For example the sign of the Easter effect should be negative for manufacturing and positive for the tourism industry. For an aggregate representing the total economy, the total remaining effect can hardly be evaluated, however.

A further example is the detection of outliers. Data compilers in national accounts usually know in which quarter and which component, due to a change in the legislation (e.g. a change in the tax rate), a level-shift is to be expected. Sometimes even some knowledge about the size of this effect exists. This information can be used for discriminating between models. The one that reproduces the assumed effect in the best way should be chosen.

If such prior information is used for discriminating between time series models , the indirect method should bring the best results in the majority of cases. But there are also examples where the direct adjustment theoretically gives better results. For instance, if a large enterprise has been reclassified from one industry to another, a negative level shift in one industry should be reflected with a positive sign in another one. As it is unlikely that an automatic procedure used for the detection of such outliers will estimate them of the same size (with different signs) or type, the direct approach would be the better choice here. If so, no outlier will be found, meaning that the reclassification balances out completely in the correct period.

Calendar effects and outliers found in component series may help to discriminate between competing models for modeling the aggregate directly, too. In case that one should be selected that gives the most consistent picture of outliers and calendar effects for the component series. For this purpose it would be helpful if the software used for seasonal adjustment would provide a good overview over detected outliers in an easy comparable form as shown in *Figure 1*. Here, for an illustration, total gross fixed capital formation as well as its components are used. It can be seen that the statistically non-significant additive outlier found in the aggregate (TOTAL GFCF) in the third quarter of 2008 is consistent with the one found in the component series “dwellings”. Also the significant transitory outliers that are starting in the fourth quarter of 2009 are approximately the same in both series. The outliers found in “other buildings and structures” and “other equipment” in 2011 seem to balance in the aggregate what makes sense from a logical point of view. The same goes for the level shifts found in “transport equipment” and “machinery equipment" starting in the last quarter of 2011. Here probably a reclassification from one component series to the other took place whose effects cancel out each other in the aggregate.

Figure 1: Outliers in time series of total gross fixed capital formation and its components



Source: WIFO calculations.

Apart from this selection criteria there are others which can be helpful for deciding whether to adjust directly or indirectly. A further guide may be the testing for residual seasonality of the aggregated components . From a theoretical point of view this should clearly not be the case but in practice it cannot be excluded, nevertheless. This test can be refined by not only testing the adjusted components and the aggregate for residual seasonality but also the difference between the summed components and the aggregate. Sometimes this difference shows a statistically significant residual seasonality (see below). In this case the model selection procedure for all components and the aggregate should be started again, with a more careful look at the remaining seasonality in residuals and their squares.

A further criterion which can be used for discrimination between time series models is based on their theoretic properties when summed up. According to theory the following relation should hold:

$$\text{ARMA}[p_1,q_1] + \text{ARMA}[p_2,q_2] = \text{ARMA}[p,q]$$

with $p \leq p_1+p_2$

and $q = \max(p_1+q_2,\ p_2+q_1)$

It could be checked whether these conditions are fulfilled or not. If not, a reexamination of the time series models should be made in order to drop the ones that harm this theoretic relation and to choose other models.

2. Indirect seasonal adjustment with chain-linked series

The problem of whether to adjust aggregates directly or indirectly seems to arise only for nominal series and real units (heads, hours, pieces etc.) as it is possible to construct sums of them. In the case of chain-linked series this is not possible. Such chain-linked series are constructed in the process of price adjustment with no fixed base year. Here the advantage of updating the base period every year in order to avoid a possible substitution bias comes at the cost that only time series of two years length can be constructed or interpreted meaningfully. In order to get longer time series of absolute price adjusted values, growth rates are derived from these fragments which are used for constructing index series. Such series typically start with an index number of 100 or 1. In many countries these index series are converted into currency units by using a certain year as the reference period in order to meet user demands. Despite this, these series still represent just growth rates meaning that – apart from the values in the reference year and the one following - they do not sum up over components or countries. Due to their non-additivity behavior compilers of seasonal adjusted figures feel compelled to use the direct approach. This has the disadvantage that a possibly existing inconsistency between the aggregate and its components - due to the stochastic nature of the process of modeling - cannot be discovered and removed.

One approach followed e.g. by Eurostat is just to take the seasonally adjusted components (countries) and – as the original series are in most cases chain-linked by the annual overlap method – to dechain them in order to get something like unchained quarters at previous year's average prices but seasonally adjusted. As additivity holds at this level, aggregates (like the GDP of the euro area or the EU) can be constructed easily and be rechained, thereafter. Of course this aggregate has to be rechecked for a possible remaining seasonality.

Here a further procedure is proposed which allows to follow the indirect seasonal adjustment method in the environment of chain-linked series. This procedure is simply based on adjusting the component series as well as the non-additivity term stemming from the unadjusted chain-linked values. Afterwards these series can be summed up and should logically represent the indirect approach.

In brief, the steps of this procedure are:

a) Calculate the non-additivity from chained original series.
b) Adjust all component series, the aggregate and the non-additivity component for seasonal and calendar variations[1].
c) Add up the components and the adjusted non-additivity in order to get the indirectly adjusted aggregate.
d) Recheck for remaining seasonality in the indirect aggregate.

---

[1] The difficulty here is that no external information about sign and size of calendar effects or possible outliers included in the non-additivity component is available.

3. Balancing the indirectly and the directly adjusted aggregate

For econometricians it is clear that, due to the stochastic nature of isolating seasonal and calendar components in time series, adjusted data of components do not add up completely to the independently adjusted aggregate. However, from a pure accounting view this inconsistency may not be acceptable. In this case some kind of balancing process between the sum of the components and its aggregate has to be done. Consequently, it has to be decided which single series represents the sum of the components as well as the aggregate. From a logical point of view it seems to be clear that this series should be calculated as a weighted sum with ω=[1,0] of both with the restriction that $\omega_C + \omega_A$=1.

If there is no external information about how to weight both series (the sum of the components and the aggregate itself) together, a pragmatic approach would be to take the average of both with $\omega_C = \omega_A$ =0.5. A further possibility would be to use subjective weights based on external information. This external information can be derived from detected outliers as outlined in chapter 1. If the aggregate logically reflects all outliers found in the component series, a simple average could be used. The less the aggregate reflects them, the more weight could be given to the component series. A parametric approach could be to link the weights to some numerical test statistics that are usually produced by the used software as additional information in the modeling process. For instance the F-test of the models used for adjusting the component series can be confronted with the one from the aggregate. Of course interpretations of the F-test ratios of the individual components have to consider the magnitude of the series so that small components cannot dominate the process.

The next step after fixing the common series by weighing is the balancing procedure. It has to be determined by how much each component series has to be adjusted. Before doing this, one has to decide whether the seasonally adjusted non-additivity component which carries some balancing adjustment load, too, should also be submitted to the balancing procedure. There are no theoretic arguments for including or excluding this component in the balancing process but, as it can be expected that this component is rather small, the decision should be arbitrary.

The difference to be balanced over all components can be distributed in several ways. One option is to do this proportionally according to their size. A more refined version could be to take account for the reliabilities of the models. The idea behind is that the higher the uncertainty of the model, used for representing the series, the more it should carry of the adjustment load; and the more reliable the less its values should be altered. Again the weighing criterion could be an F-test for instance.

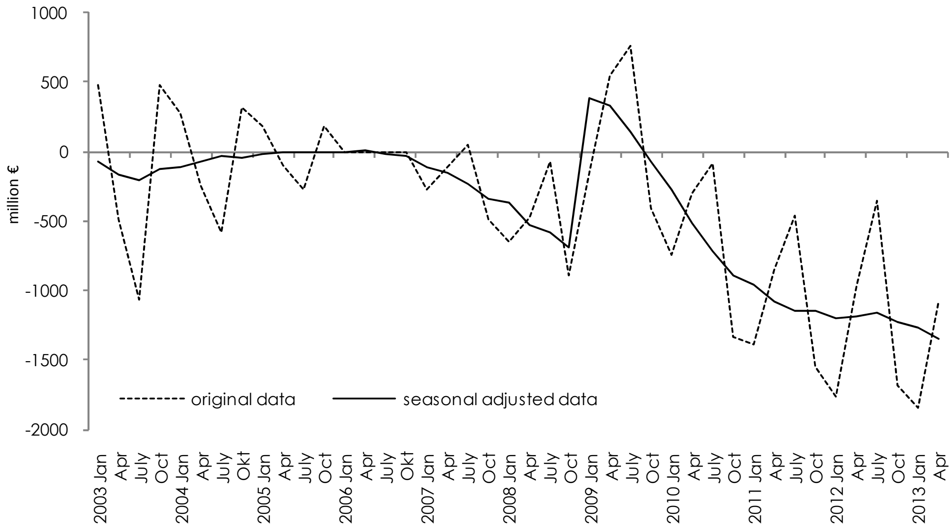
There also exist automatic balancing procedures like the multivariate Denton procedure or the RAS technique. The first tries to minimize the original variations of the data series over time by respecting their cross dimensional restriction of the sum over components. The latter is a computational iterative process which does not explicitly consider the time series behavior of the data and may not be as good as the first in this context.

4. A numerical example of the direct/indirect approach with chain-linked data

In order to demonstrate the proposed approach for a joint use of the direct and indirect approach with a following balancing procedure, euro area investment data for several components was used here. The data set consists of unadjusted chain-linked investment data for six categories and total gross fixed capital investment as their aggregate between the first quarter of 2003 and the second quarter of 2013 and includes 42 observation points for each series[2].

Figure 2 shows the difference between the sum of the unadjusted chain-linked components "residentials", "non-residentials", "vehicles", "other machinery", "cultivated assets", "intangibles" and the aggregate "total gross fixed capital formation". This non-additivity component shows a clear seasonal pattern and values around zero for the reference year 2005 and 2006.

Figure 2: Non-additivity in millions of chain-linked euros (original and seasonally adjusted data)



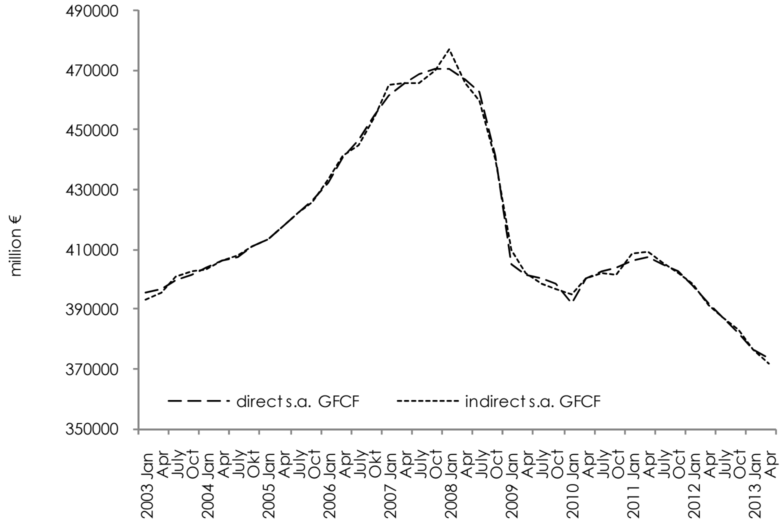
Source: WIFO calculations.

In a first step this non-additivity component was adjusted for seasonal and possible existing calendar variations. As for the directly adjusted aggregate it was difficult to form priors about sign and size of outliers and calendar effects which could help discriminating between competing models. So effects were taken as suggested by the TRAMO-SEATS version, included in the software package Demetra+. An Airline Model was automatically chosen, showing a significant seasonality but no calendar effects. Furthermore, a negative level shift in the first quarter 2009 was detected. The output is given in *Figure 2* as the smoother line.

[2] The data was downloaded from the Eurostat data base on 12. November 2013. At request the data is available from the author.

The next step was to add this adjusted non-additivity component to the sum of the other components which should be the indirectly adjusted aggregate. In *Figure 3* both the directly as well as the indirectly seasonally adjusted GFCF of the euro area are shown.

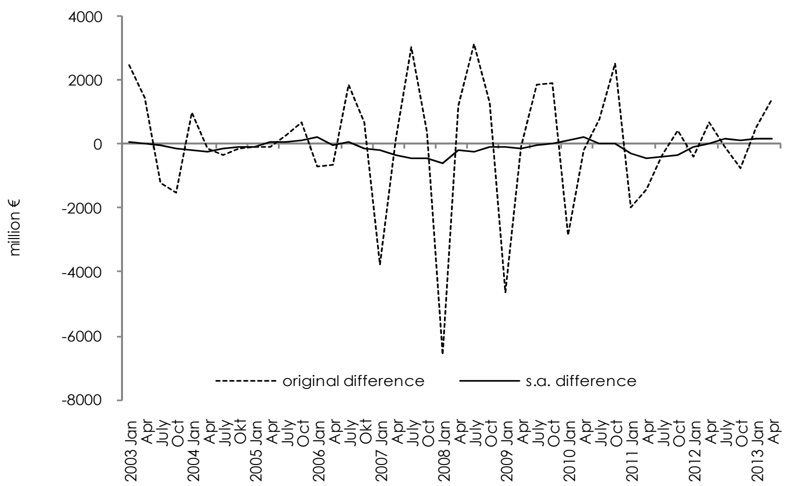
Figure 3: Direct and indirect seasonally adjusted gross fixed capital formation of the euro area



Source: WIFO calculations.

Both lines show the same level but different variations for certain periods. The remaining difference between them should logically stem just from the process of seasonal adjustment and not from the non-additivity of the original chained data. This difference is given in *Figure 4.*

Figure 4: Original and seasonally adjusted difference between direct and indirect SA of gross fixed capital formation



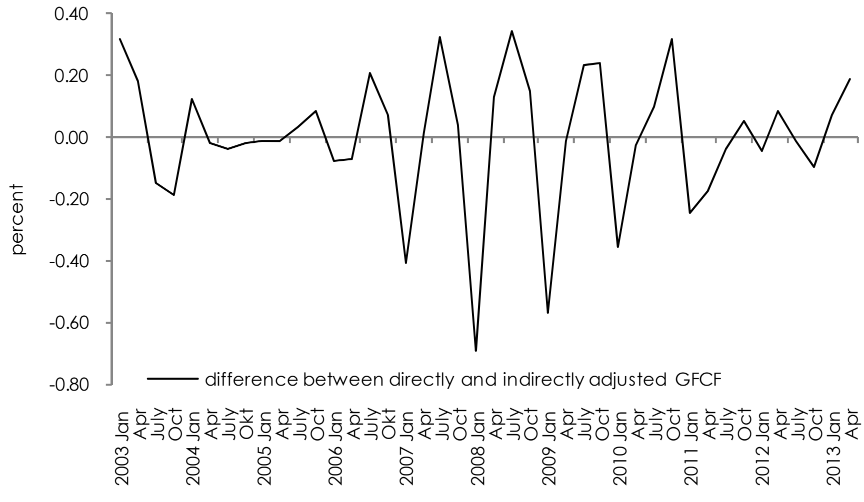
Source: WIFO calculations.

It is interesting that neither the directly seasonally adjusted aggregate nor the indirectly sum of adjusted components showed remaining seasonal variations as their difference did. *Figure* 4 shows seasonal fluctuations even by visual inspection what was confirmed by statistical tests. In this figure the seasonally adjusted series is given, too. If this difference is spread over the different components, series are theoretically recontaminated with seasonal variations. As this difference is quite small, it can be expected that this recontamination would not lead to a statistically significant season in the series. In order to check this, one common series for seasonally adjusted GFCF was derived by simply taking the arithmetic mean of both series and recheck for included seasonal variations. Taking the average means that only half of this difference goes to each series which reduces further the problem of a significant recontamination. Not surprisingly, the statistical tests for seasonality showed that there was by far no statistical significant one included in the common series derived by averaging.

In the last step, the part of the difference which was added to the indirectly adjusted series had to be distributed over its component series. As mentioned in section 3, different approaches can be used like e.g. the multivariate Denton method or the proportional method with and without special weights stemming from information criteria. In order to keep this exercise simple this difference was proportionally distributed over the components' weight in the indirectly adjusted GFCF series. In this case the relative distortion in the component series was immediately clear. It was the same as the total difference compared to the indirectly adjusted series (minus the non-additivity if it was chosen that it should not carry an adjustment load). These relative differences can be seen in *Figure 5*, with the largest in the first quarter of 2008 accounting for nearly 0.7% of GFCF in this period.

After distributing these differences over all components an indirectly adjusted series of GFCF was obtained which was fully consistent with the directly adjusted one, apart from the non-additivity term stemming from chain-linking of the original data.

Figure 5: Difference between directly and indirectly adjusted GFCF in % of average



Source: WIFO calculations.

A last check whether the component series still showed no remaining seasonal fluctuations was necessary, as distributed differences included such. In the present example no significant fluctuations were found.

5. Conclusions

This paper proposes a guide on whether to perform seasonal adjustment of aggregates directly or indirectly by summing up its adjusted components. Three criteria are recommended: a) the empirical informational content of outliers that are found in the adjustment process based on theoretical considerations, b) the consistency of the number, types, occurrence and size of outliers between the individually adjusted components and its directly adjusted aggregate and c) the theoretical properties of summed ARIMA processes.

For opening the possibility to adjust even non-additive components like chain-linked series, additionally to the existing dechaining-rechaining method, a further approach is proposed. This approach is based on adjusting the difference between the sum of the components and the aggregate - the so-called non-additivity term - as a separate component. Doing so, the sum of all adjusted components including this term should represent the indirectly adjusted aggregate. All differences between the indirectly and directly adjusted time series can solely be attributed to the stochastic nature of the seasonal adjustment process and therefore be balanced, if necessary.

In a practical example using euro area data on gross fixed capital formation and its components the procedure is demonstrated as well as a possibly following balancing process between the adjusted components and their aggregate. This should give a fully consistent data set of adjusted series, apart from the non-additivity term.