



**Statistics Netherlands**

MSP

MPP

*P.O.Box 4000  
2270 JM Voorburg  
The Netherlands*

---

## **Trading day adjustment for the consumption of Dutch households (methodological note)**

**Bastiaan Rooijackers**

**Mohamed El Bouchehati**

**24 January 2005**

# TRADING DAY ADJUSTMENT FOR THE CONSUMPTION OF DUTCH HOUSEHOLDS

*Summary:*

*Keywords:*

## 1. Introduction

The year-on-year volume changes of the monthly data on Dutch consumer spending are strongly influenced by differences in the trading day pattern. This makes it difficult to interpret the results correctly. To give a good insight in the underlying development of consumer expenditure it is necessary to adjust the figures for the trading day pattern.

This methodological note describes a study on trading day adjustment of the monthly data on domestic consumption of households with the software program Arima-X-12. This program is an improved and extended version of Arima-X11 which is widely used for seasonal and trading day adjustments. Arima-X12, developed by U.S. Census Bureau, is at the Statistics Netherlands linked with the user shell Vivaldi. An adjustment for trading days was first introduced in Dutch monthly consumer spending statistics in March 2004. Data are available from 1996 onwards.

The monthly domestic consumption is composed from multiple statistical resources, e.g. retail trade turnover statistics and use of gas, water and electricity. The monthly data on consumer spending are based on the concepts and definitions of the system of National Accounts (ESA). All data are adjusted to the National and Quarterly Accounts. The classification of goods and services used for compiling the short term statistics is similar to the one used in the annual National Accounts. However, publication on a monthly basis is at a less detailed level. In appendix A the list of different article groups and their aggregates is published.

The aggregates at two digit level are the basic level for composing trading day adjustments. Further research made clear that aggregates are best compiled according to the indirect method of adjustment; first adjust the original time series and then combine the adjusted series to an aggregate. The research was based on 18 two digit article group series. Separate regression models have been developed for all of the 18 series. For analysing purposes, the time series ran initially from January 1995 to December 2003. This is a relatively short period.

Once a year new analysis will lead to further improvement of the trading day adjustments, with more refined set-ups and more accurate adjusted figures. Each time a year is completed the results will be added to the original time series, but only after the data are adjusted to the results of the latest National accounts. The annual

National accounts provide new results for the consumption of households in the years t-1, t-2 en t-3.

Both the adjustment to the National accounts and the addition of a new year to the time span take place at the same moment. This is done for the users comfort, as revisions twice a year and shortly after one another can be very confusing. The new model will be fixed for one year. For the adjustment of the figures implemented in August 2004, the time series is updated with the data of the months of 2001, 2002 and 2003 corresponding with the National and Quarterly Accounts.

This methodological note describes the techniques and methods behind the composition of the regression models. With a view to the periodic revisions of the models, the most important model specifications are mentioned in Appendix B.

## Composition of setups

### General approach

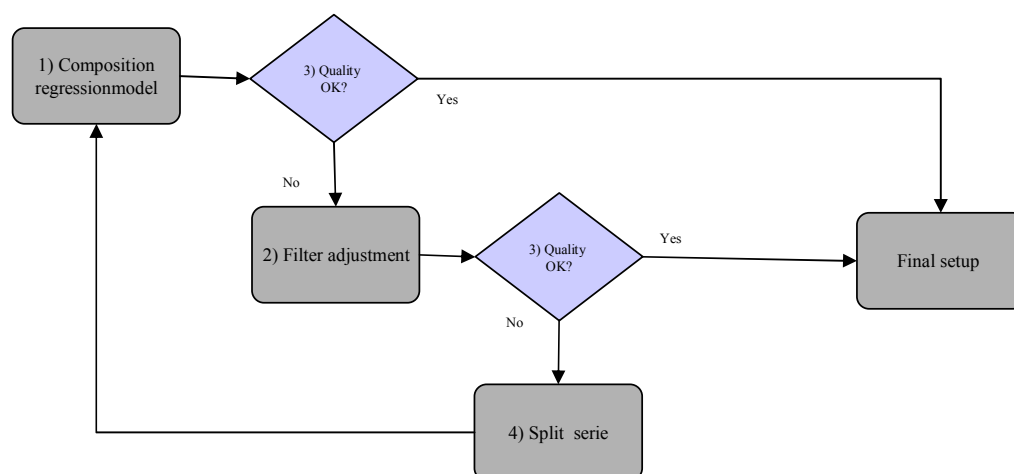


Figure 1: composition of setup<sup>1</sup>

In X12-Arima, the regression models are implemented in so called setups. Each of the setups is connected to a specific original time series. The adjusted series are the result of running the individual setups with their corresponding original time series. Figure 1 explains in general how a final setup is made. In the first stage, extensive tests are performed to get the optimal regression model. The resulting model is judged by quality criteria formulated beforehand. If the criteria are not met, the second stage will be fine-tuning the model by adjusting the seasonal and trend filters. In the next paragraphs, the stages as well as the quality checks will be explained in detail.

### *1) Composition of regression models*

#### **- Decomposition**

With Arima-X-12 two main types of decomposition can be performed: additive or multiplicative. The type of decomposition is selected according to (one of) the following 3 methods:

- 1). Automatic: Arima-X-12 performs an automatic analysis and makes a decision between the two types of transformation

---

<sup>1</sup> Michiel Jansen –Seasonal and working day adjustment for the Industry Production Index – Methodological Note

2). Statistical: compare the outcomes for the F-tests for seasonality with an additive and a multiplicative run and choose the decomposition with the highest F-value.

3). Judge visually based on the graphic of the time series: if the trend of the original series is upward or downward and/or if the peaks increase or decrease in time, choose for multiplicative decomposition. Choose for additive decomposition if the trend is smooth and constant over time.

For most of the series, all three methods lead to the same conclusion and the type of decomposition is unambiguously selected. For series with adverse conclusions the choice is based on the second method.

#### **- ARIMA-model**

For all setups, the program is instructed to select the Arima-model automatically. Standard definitions are:

- method=best: the estimated model with the lowest within-sample forecast error
- mode=forecast: the model will be used to produce a year of forecasts.

#### **- Regression variables**

Arima-X-12 distinguishes regression variables into two main groups: trading day and user-defined. Both groups are judged through the Chi-squared Test to determine whether a group of regression variables has a significant added value to the adjustment. This is expressed in a P-value. Under the condition of a 95-percent confidence interval, the P-value may not exceed 0,05. Judgement of the individual regression variables is based on the t-values. Regression variables with t-values ranging from -1,96 to 1,96 lie within the 95-percent confidence interval and are considered as sufficiently stable to be included in the model. A regression model can be composed by a maximum of 5 different regression variables. Research has been done on the following effects:

- 1) Trading day effects: The contribution of different weekdays to the consumption.
- 2) Leap year effects: The contribution of an extra day in February to the consumption.
- 3) Easter effects (user defined): The contribution to the consumption from the 10 days before Easter. These days variously fall in March and/or April.
- 4) Holiday effects (user defined): The negative contribution of national holidays (e.g. New Years Day, Easter, Christmas) to the consumption.

Regression variables are selected according to a fixed procedure. Starting point is a standard set which contains regression variables for trading day effects, Easter effects, holiday effects and leap year effects. The regression variables will be judged on their t- and P-values. If necessary, a new set is composed by replacing, deleting or adding regression variables. The new set of variables will again be judged on their t- and P-values. This process is continued until the 'ideal' set of regression variables is found for the series concerned.

The following comments are relevant for the selection of regression variables:

- The 6 day effect regression variables are judged by their collective P-value. If the value exceeds 0,05, the article group is left out of the analysis as the meaning of the analysis is to compute the trading day effects.
- The standard set of regression variables contains a leap year variable. The regression variable is removed from the original set if the t-value is not significant.
- The sign of t-values and parameter estimates should be interpreted as follows. A negative t-value or parameter estimate will increase the original volume of the consumption and a positive sign will decrease the volume of the consumption.

- Outliers and level shifts

The default critical value for tracing outliers is 3.0. For some series a higher or lower critical value is chosen, ranging from 2.0 to 3.8. Deviations from the default value were made if the resulting adjustment improved.

## *2) Selection of filters*

- Seasonal filter

The seasonal filter estimates the seasonal components of a series. Default adjustment for the filter is Moving Seasonality Ratio (MSR), which means the programme will chose the final seasonal filter automatically. The seasonal filter can also be manually adjusted. For monthly series from 1995, three possible filters are available: 3x3, 3x5 and 3x9. For one series it was necessary to choose a separate filter for each month.

- Henderson filter

The Henderson filter is a trend filter. It estimates the trend-cycle component or the long-term movement of series. The programme selects automatically the optimal Henderson-filter. Most of the series have a 13-term Hendersonfilter. Two series have a 23-term Hendersonfilter.

## *3) Quality check*

#### - Quality Measures

The quality of an adjustment is expressed in eleven M-measures, which are automatically generated by the programme. The M-measures indicate different statistical diagnostics. Their value can range from 0,0 to 3,0, and ideally they lie next to zero. The main M-measures are:

M2: The relative contribution of the irregular component to the stationary portion of the variance (from Table F 2.F).

M7: The amount of moving seasonality present relative to the amount of stable seasonality (from Table F 2.I).

M10: The size of the fluctuations in the seasonal component throughout the last three years

M11: The average linear movement in the seasonal component throughout the last three years.

The overall acceptability of an adjustment is measured in Q-measures:

Q: Overall index of the acceptability of the seasonal adjustment

Q2: Q statistic computed without the M2 Quality Measure statistic

The Q-measures can range from 0,0 to 3,0, with 0,0 as best performance. Both for the individual M- and Q-measures applies that values above 1,0 are better avoided. Especially M7 should not exceed 1,0. Generally, the quality of an adjustment is considered as reliable if the values of the mentioned measures do not exceed 0,7.

#### - Visual check

The quality of adjustments is also visually checked by using graphs. Especially the following tables were closely inspected:

A1: Original time series data

B1xA8 : series adjusted for calendar effects (i.e. trading day and holidays), but including outliers

D11: Final seasonally adjusted data

D12: Final trend cycle

## **Practicalities**

### **- Graphs**

Graphs of all 12 series are available from the author. For all series two graphs are presented. The upper one contains the original time series and the series adjusted for calendar effects, but including outliers. The lower graph includes the final seasonally adjusted data (also adjusted for calendar effects), with replacement of outliers and the final trend cycle. For most series two years of estimations are shown in the upper graph. For technical reasons, no estimations are shown for series without an automatically chosen ARIMA-model. The graphs can be useful for analyses on behalf of future updates of the setups.

### **- Output files**

All generated tables, model specifications and statistical analyses can be shown in the output files of the ARIMA-X-12 programme. The files can be useful for analyses on behalf of future updates of the setups.

### **- User-defined regression files**

One type of user-defined regression variable is defined in the models: holidays. The user-defined regression variable is attached to the models by a separate file. The regression variable is updated to December 2008.

### **- Publication table**

One table in particular is used for monthly publication:

- B1\*A8: original series (A1) adjusted for trading day patterns.

## **Publication**

Publication of the consumption figures happens at 1-digit level in year on year volume changes. As all original time series are based on fixed prices 1995, it is not possible to simply add the time series up to a 1-digit aggregate. As of 1997 (t+2) this will give a statistical discrepancy between the total of the 2-digit time series and the 1-digit aggregates. The reason for this is that the aggregates are derived from index figures.

To avoid the problem of statistical differences, we use for every year the fixed prices of t-1. The changes of the 2-digit time series are both, based on fixed prices 1995 and fixed prices t-1, equal. This means that it is allowed to use the mutations based on fixed prices 1995 on volume data of t-1 with fixed prices of t-1. Then the 2-digit data can be added up to a correct 1-digit aggregate to compile the volume change of the aggregate.



## Appendix A

List of article groups and their aggregates (publication level in bold)

|                  |   |
|------------------|---|
| <b>1000-2000</b> | <b>Food, beverages and tobacco</b>        |
| <b>3000</b>      | <b>Durable consumer goods</b>             |
| 3100             | Textiles and clothing                     |
| 3200             | Leather goods and footwear                |
| 3300             | Home furnishing                           |
| 3400             | Household appliances                      |
| 3500             | Household articles                        |
| 3600             | Vehicles                                  |
| 3700             | Other durable consumer goods              |
| <b>4000</b>      | <b>Other goods</b>                        |
| 4100             | Fuel                                      |
| 4200             | Energy                                    |
| 4300             | Other goods n.e.c.                        |
| <b>1000-4000</b> | <b>Goods</b>                              |
| <b>5000</b>      | <b>Services</b>                           |
| 5100             | Housing                                   |
| 5200             | Hotels and restaurants services           |
| 5300             | Recreational and cultural services        |
| 5400             | Transport and communication services      |
| 5500             | Medical services and welfare              |
| 5600             | Financial and business services           |
| 5700             | Other services                            |
| <b>1000-5000</b> | <b>Domestic consumption of households</b> |

## Appendix B

Model specifications

Explanation of abbreviations;

DC=type of decomposition (additive / multiplicative)

TD = trading day regression

SF = seasonal filter

LY = leap year regression

HF = Henderson trend filter

F = Holiday regression

MSR = moving seasonality ratio

E = Easter (10days)

| SBI       | start  | DC | SF  | HF | Arimamodel      | TD | LY | Holi-<br>day | Levelshift<br>Temp.ch | Outliers   | M-values-<br>M2/M7/M10/M11/Q1/Q2 |
|-----------|--------|----|-----|----|-----------------|----|----|--------------|-----------------------|--|----------------------------------|
| 1000-2000 | 1995.1 | M  | 5   | 13 | (0,1,2)(0,1,2)  | x  | x  | E            |                       | 98.1/ 00.4   | 0.0/0.1/0.5/0.5/0.4/0.4          |
| 3100      | 1995.1 | M  | MSR | 23 | (2,1,2)(0,1,1)  | x  |    | E            |                       | 96.5/99.2/00.3/01.4/01.10/97.8/99.9/00.6/01.6/01.9 | 0.2/0.2/0.5/0.5/0.6/0.6          |
| 3200      | 1995.1 | M  | 5   | 13 | (2,1,2)(0,1,1)  |    |    | F            |                       | 01.5/01.9/96.5/97.4/03.5                           | 0.0/0.1/0.3/0.3/0.3/0.3          |
| 3300      | 1995.1 | M  | 5   | 13 | (2,1,2)(0,1,1)  | x  |    |              |                       | 96.5/98.7/00.12/01.2/97.8                          | 0.0/0.3/0.5/0.4/0.4/0.4          |
| 3400      | 1995.1 | M  | 5   |    | (0,1,1)(0,1,1)  | x  | x  |              |                       | 01.4/99.12   | 0.1/0.1/0.4/0.3                  |
| 3500      | 1995.1 | M  | 5   | 23 | (0,1,1)(0,1,1)  | x  | x  |              |                       | 03.2   | 0.1/0.1/0.4/0.4/0.7/0.8          |
| 3600      | 1995.1 | M  | 9   | 23 | (0,1,2)(0,1,1)  | x  |    |              |                       | 00.12/02.12  | 0.1/0.2/0.3/0.3/0.5/0.5          |
| 3700      | 1995.1 | M  | 9   | 13 | (0,1,1)(0,1,1)  | x  |    |              |                       | 00.12  | 0.2/0.2/0.4/0.4/0.4/0.4          |
| 4100      | 1995.1 | M  | 9   | 13 | (0,1,2)(0,1,1)  | x  | x  |              |                       | 97.7/95.12/96.12/00.4/97.12/98.11                  | 0.4/0.2/0.3/0.3/0.4/0.4          |
| 4300      | 1995.1 | M  | 5   | 13 | (2,1,2)(0,1,1)  | x  | x  |              |                       | 97.8/95.7/95.12                                    | 0.0/0.1/0.2/0.2/0.2/0.3          |
| 5200      | 1995.1 | M  | 5   | 13 | (0,1,2)(0,1,1)  | x  |    | F            |                       | 96.8/98.5/00.3/00.7/01.9/03.2                      | 0.0/0.1/0.4/0.4/0.3/0.4          |
| 5400      | 1995.1 | M  | MSR | 9  | (2,1,2)((0,1,1) |    |    |              | 01.10                 |  | 0.0/0.3/0.5/0.5/0.2/0.2          |

