Plug-in TransReg

Version 0.2.11

September 27, 2017

Title Transformation of regression variables

Author Karsten Webel, Nina Gonschorreck

Maintainer Deutsche Bundesbank

Contact `JDemetra@bundesbank.de`

Depends JDemetra+ 2.2.0 or higher, Java 8 or higher

Repository `https://github.com/bbkrd/TransReg`

Bug reports `https://github.com/bbkrd/TransReg/issues` (or e-mail)

License Open source

Contents

1 Introduction

Calendar effects play an important role in the seasonal adjustment of many German macroeconomic stock and flow series. For example, they account for as much as 1 percentage point of the quarterly rate of change in real gross domestic product. Depending on the economic sector and/or the economic activity, calendar effects can be constant or vary over the months or quarters. For orders received by industry, for example, one additional working day leads to an average increase by 3.1 % in each month. For output in industry, however, the average rise is not constant over the year. More precisely, one additional working day in the months January to November leads, on average, to an increase in output by 3.5 %, whereas the average rise is significantly smaller in December (2.3 %), since many firms reduce production around Christmas. A detailed discussion of working-day and other calendar effects in German macroeconomic time series, including further examples, is given by Deutsche Bundesbank 2012.

To estimate calendar effects, regARIMA and related time series regression models have proved to be effective. In this regard, calendar variables have to centred in order to yield a meaningful estimate of the calendar component. As shown in figure 1 for orders received by industry, usage of calendar variables which have not been centred is likely to result in a mismatch between the levels of the original and calendar (and seasonally) adjusted time series. For this reason, many seasonal adjustment programs, such as X-12-ARIMA, X-13ARIMA-SEATS and TRAMO/SEATS, automatically centre built-in calendar variables and provide options for doing the same with user-defined calendar variables. In JDemetra+ (JD+), however, the latter option is not available as user-defined variables are assumed to be loaded in centred form.

The absence of a centring option in JD+ motivated work on this plug-in. In particular, we believe that centring user-defined calendar variables outside JD+ and then loading the centred version into the program is neither comfortable nor sustainable. In addition, creating a centring plug-in opens up the possibility for implementing further transformations of user-defined regression variables which are currently not available in JD+ (and the seasonal adjustment programs mentioned above). A prime example is splitting a calendar variable into two or more period-specific variables which is necessary for estimating period-specific calendar effects, as in the case of output in industry.

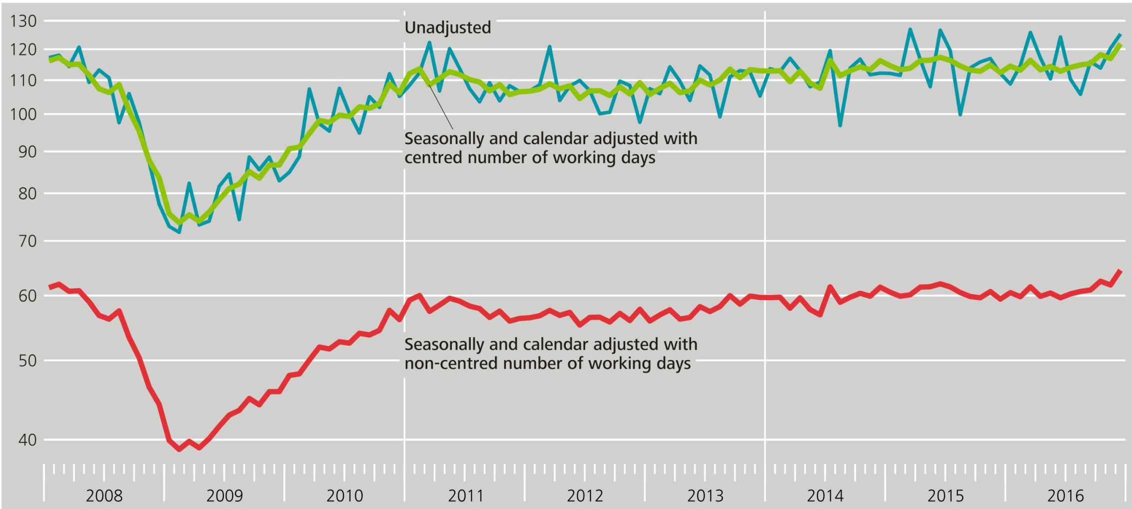
1.1 Overview

The TransReg plug-in allows the user to carry out grouping and centring of user-defined regression variables in JD+:

Orders received by industry

Figure 1

Value, 2010 = 100, log scale



Deutsche Bundesbank

- Grouping is the generation of a set of period-specific daughter regression variables from a mother regression variable. The user can (1) specify the maximum number of daughter regression variables and (2) assign the months or quarters to the daughter regression variables.

- Centring is the subtraction of the sample mean. The user can (1) choose between the overall sample mean and period-specific sample means and (2) specify the time span to be used for calculating the chosen type of sample mean.

To illustrate the two transformations, we consider the monthly number of working days in Germany as reported in table 1.1. These figures account for both national and regional public holidays as the latter are counted as partial working days. This is a fundamental assumption of the working-day model which applies to a large number of German economic indicators and is described in more detail in Deutsche Bundesbank 2012.

1.2 Installation

Download the installation file `TransReg-0.2.11.nbm` from the plug-in's GitHub repository and save it to your local folder of choice. Choose

Tools ► Plugins

from the JD+ drop down menu, open the **Downloaded** sheet and press the Add Plugins button. Select the local copy of the installation file, click the Install button and accept the terms in the licence agreements. Once the installation is finished, TransReg documents

Table 1.1: Monthly number of working days in Germany.

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
2011	20.6	20.0	22.6	19.0	22.0	19.3	21.0	22.8	22.0	19.9	21.3	21.0
2012	21.6	20.6	22.0	19.0	20.0	20.3	22.0	22.8	20.0	21.9	21.3	17.0
2013	22.0	19.6	20.0	21.0	19.3	20.0	23.0	21.8	21.0	21.9	20.3	18.0
2014	21.6	20.0	20.6	20.0	20.0	19.3	23.0	20.8	22.0	21.9	20.0	19.0
2015	20.6	19.6	22.0	20.0	18.0	21.3	23.0	21.0	22.0	22.0	21.0	20.0
2016	19.6	20.6	21.0	21.0	19.3	22.0	21.0	22.8	22.0	19.9	21.3	21.0
2017	21.6	19.6	23.0	18.0	21.0	20.3	21.0	22.8	21.0	20.0	21.3	19.0
2018	22.0	19.6	21.0	20.0	19.3	21.0	22.0	22.8	20.0	21.9	21.3	17.0
2019	22.0	20.0	20.6	20.0	21.0	18.3	23.0	21.8	21.0	21.9	20.3	18.0
2020	21.6	19.6	22.0	20.0	19.0	20.3	23.0	21.0	22.0	22.0	21.0	20.0

and a global option menu are available. Note that the plug-in is automatically activated and, therefore, a restart of JD+ is not necessary.
The remainder of this manual is organised as follows:

- Section 2 describes the basic structure of the TransReg document as well as data input and storage.
- Section 3 explains how the available transformations can be carried out in the specification panel of the TransReg document.
- Finally, section 4 addresses the global options available for customising the plug-in according to the user's preferences.

2 The TransReg document

The TransReg document is the main device of the plug-in as it enables input and transformation of regression variables as well as storage of transformed regression variables. A new document can be created and opened simultaneously by selecting

Tools ► TransReg

from the JD+ drop down menu. As a sequential alternative, a new TransReg document can be created by choosing New after right-clicking on

⊞ Utilities ► TransReg

in the **Workspace** window. Then the created document can be opened by double-clicking on it. Either way, the default name of a new TransReg document is **TransReg-Number**, where **Number** is a counter of the TransReg documents which are available in the current workspace. For example, a new document created in a workspace which does not contain any other TransReg document will be named **TransReg-1**. Any default name can be changed by right-clicking on

⊞ Utilities ► TransReg ► TransReg-Number

in the **Workspace** window and selecting Rename. For convenience, we subsequently use the name **MyTransReg** to refer to the TransReg document currently in use. Similarly, we will use **MyRegVar** as the default name of a user-defined regression variable.

2.1 General structure

The TransReg document consists of three elements: the top banner, the data panel and the specification panel.

- The top banner contains the data input field labelled *Drop data here*, in line with familiar JD+ standards.
- The data panel is essentially a table with a maximum of eight columns. The information reported in these columns is summarised in table 2.1. By default, all columns are displayed but the user can customise the appearance of the data panel by right-clicking on any spot in the light grey title row and selecting the preferred columns, except for variable which is always displayed.
- The specification panel provides functionalities for transforming regression variables loaded into the data panel. A detailed description of the available transformations is given in section 3. The appearance of this panel has been adopted from the specifications used for seasonal adjustment.

Table 2.1: Data panel content of the TransReg document.

Column	*Information*	*Displayed for*
Variable	Name of regression variable. Transformed variables are ordered in a tree-type structure to facilitate identification of dependencies.	All variables.
Level	Type of regression variable: original for untransformed variables, group identifier for grouped variables, centring identifier for centred variables.	All variables.
Frequency	Periodicity of regression variable: monthly or quarterly.	All variables.
Period	Observation span.	All variables.
Pre-test	Test if regression variable has already been centred (see section 3.2.2).	Original and grouped variables.
Timestamp	Time of data generation.	All variables.
Mean calculation span	Span chosen for the calculation of sample mean as given by centring identifier (see level).	Centred variables.
Sample mean	Value of the sample mean as given by centring identifier (see level).	Centred variables.

2.2 Data input

To load a user-defined regression variable into the plug-in, either create and open an empty new TransReg document or open an existing document which may already contain some data. Then, drag and drop the regression variable from your data provider of choice into the data input field of the top banner of the TransReg document. Note that missing values are allowed and treated as being not a number (NaN). If the document already contains one or more regression variables with the same name (**MyRegVar**), then the name of the regression variable added most recently will be changed automatically to

MyRegVar ► Number,

where **Number** is a counter which works in exactly the same way as the counter of the TransReg documents.

This data input procedure also works for loading multiple regression variables simultaneously. For example, dragging and dropping an Excel sheet with k user-defined regression variables into the data input field will input the k variables in one step.

If a regression variable which has been loaded into a TransReg document is changed or updated in the external data source, it can be refreshed in the TransReg document by selecting Refresh after right-clicking on

⊞ Utilities ► TransReg ► MyTransReg

in the **Workspace** window. In this case, all specified transformations will be recalculated automatically for the refreshed regression variable using the saved options.

2.3 Specification panel

The specification panel shows up automatically at the right edge of the TransReg document as soon as a regression variable is selected in the data panel. It allows the user to carry out the transformations described in section 3. In this regard, it should be noted that the basic principle of the plug-in is simultaneous rather than sequential performance of multiple transformations. Therefore, the following rules apply:

- For untransformed regression variables, the user has to specify all transformations which should be performed in a single specification panel. Once the [Calculate] button in the bottom line is pressed, the transformed regression variables obtained at the intermediate steps and the final stage are saved to the data panel in a tree-type structure.
- For transformed regression variables, the user cannot specify any further transformation. Still, the saved options used for obtaining the transformed variable are displayed in the specification panel.

As an example, assume a user-defined mother regression variable should be grouped into two daughter regression variables, each of which should be centred around period-specific sample means. Then, the two transformations (grouping, centring) have to be specified jointly since grouping the mother regression variable in a first step and centring the saved daughter regression variables in a second step will not work according the plug-in's basic principle.

2.4 Data storage

In general, TransReg documents can be saved easily as part of the current workspace. In this case, however, the transformed regression variables are not available in JD+ for other purposes, such as seasonal adjustment. To remedy this situation, the transformed regression variables which should be generally available in JD+ could be copied to a variable list accessible under

⊞ Utilities ► Variables

in the **Workspace** window. Note that this approach also enables exchange of transformed regression variables with users who did not install the TransReg plug-in to their JD+ environment.

To copy data from a TransReg document to a variables list, double-click on the transformed regression variable in the TransReg document which opens a new data window. Select [Chart] or [Grid] and proceed as follows:

- For Chart, click on any spot in the time series of the transformed regression variable and drag and drop it into the opened variables list.
- For Grid, click on the upper left corner of the table of the transformed regression variable and drag and drop it into the opened variables list.

3 Transformations

Let the regression variable under consideration be denoted by

$$\mathbf{X} = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^\top . \tag{3.1}$$

Adopting the matrix notation used in table 1.1 to represent the monthly number of working days in Germany, equation (3.1) can be rewritten loosely as

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,\tau} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n_1,1} & x_{n_2,2} & \cdots & x_{n_\tau,\tau} \end{pmatrix}, \tag{3.2}$$

which will be used for convenience in the sequel. In equation (3.2), n_j is the number of observations in the j-th period and τ is the number of observations per year (that is $\tau = 4$ for quarterly data and $\tau = 12$ for monthly data), giving $n = \sum_{j=1}^{\tau} n_j$. In most cases, the regression variable will cover a span of complete years, leading to $n_1 = n_2 = \cdots = n_\tau$. However, this does not constitute a necessary condition and, therefore, the plug-in allows for some flexibility here.

3.1 Grouping

Grouping is the process of creating a set of period-specific daughter regression variables from a mother regression variable. The maximum number of groups allowed in this process, g_{max}, can be specified in the global option menu, bearing in mind that it is restricted naturally to $g_{max} \leq \tau$ (see section 4). For any daughter regression variable, the columns which should not be taken into account are treated as NaN's. Thus, separation of the figures for period τ from $\mathbf{X}$ as given in equation (3.2) creates the two daughter variables

$$\mathbf{X}_1 = \begin{pmatrix} x_{1,1} & \cdots & x_{1,\tau-1} & . \\ x_{2,1} & \cdots & x_{2,\tau-1} & . \\ \vdots & \ddots & \vdots & \vdots \\ x_{n_1,1} & \cdots & x_{n_{\tau-1},\tau-1} & . \end{pmatrix} \quad \text{and} \quad \mathbf{X}_2 = \begin{pmatrix} . & \cdots & . & x_{1,\tau} \\ . & \cdots & . & x_{2,\tau} \\ \vdots & \ddots & \vdots & \vdots \\ . & \cdots & . & x_{n_\tau,\tau} \end{pmatrix},$$

where single dots denote NaN's, see also example 3.1 below.
To carry out grouping in a TransReg document, the checkbox **Groups by period** has to be marked first in the ⊞ Grouping section of the specification panel. Afterwards, click into the white box next to ⊞ Groups to open a new menu where the groups can be allocated to

the periods using the respective drop down menus. To save and close the allocation menu, click on the red X-button in the upper right corner. Confirm your choice by clicking on Calculate in the bottom line of the specification panel and the daughter regression variables appear in the data panel under the following default names:

MyRegVar ► Group1,
MyRegVar ► Group2,
...
MyRegVar ► GroupG,

where $G \leq g_{max}$ is the number of groups actually assigned. Any default name can be changed by selecting Rename after right-clicking on the respective grouped regression variable in the data panel.

Example 3.1 (Splitting off workings days in December) *The monthly number of working days in Germany as given in table 1.1 is to be separated into two daughter variables: one should cover the months from January through November, whereas the other should carry working days only for December. Assigning the months January through November to group 1 and December to group 2 yields the desired daughter variables*

$$\mathbf{X}_1 = \begin{pmatrix} 20.6 & 20.0 & 22.6 & 19.0 & 22.0 & 19.3 & 21.0 & 22.8 & 22.0 & 19.9 & 21.3 & . \\ 21.6 & 20.6 & 22.0 & 19.0 & 20.0 & 20.3 & 22.0 & 22.8 & 20.0 & 21.9 & 21.3 & . \\ 22.0 & 19.6 & 20.0 & 21.0 & 19.3 & 20.0 & 23.0 & 21.8 & 21.0 & 21.9 & 20.3 & . \\ 21.6 & 20.0 & 20.6 & 20.0 & 20.0 & 19.3 & 23.0 & 20.8 & 22.0 & 21.9 & 20.0 & . \\ 20.6 & 19.6 & 22.0 & 20.0 & 18.0 & 21.3 & 23.0 & 21.0 & 22.0 & 22.0 & 21.0 & . \\ 19.6 & 20.6 & 21.0 & 21.0 & 19.3 & 22.0 & 21.0 & 22.8 & 22.0 & 19.9 & 21.3 & . \\ 21.6 & 19.6 & 23.0 & 18.0 & 21.0 & 20.3 & 21.0 & 22.8 & 21.0 & 20.0 & 21.3 & . \\ 22.0 & 19.6 & 21.0 & 20.0 & 19.3 & 21.0 & 22.0 & 22.8 & 20.0 & 21.9 & 21.3 & . \\ 22.0 & 20.0 & 20.6 & 20.0 & 21.0 & 18.3 & 23.0 & 21.8 & 21.0 & 21.9 & 20.3 & . \\ 21.6 & 19.6 & 22.0 & 20.0 & 19.0 & 20.3 & 23.0 & 21.0 & 22.0 & 22.0 & 21.0 & . \end{pmatrix}$$

and

$$\mathbf{X}_2 = \begin{pmatrix} . & . & . & . & . & . & . & . & . & . & . & 21.0 \\ . & . & . & . & . & . & . & . & . & . & . & 17.0 \\ . & . & . & . & . & . & . & . & . & . & . & 18.0 \\ . & . & . & . & . & . & . & . & . & . & . & 19.0 \\ . & . & . & . & . & . & . & . & . & . & . & 20.0 \\ . & . & . & . & . & . & . & . & . & . & . & 21.0 \\ . & . & . & . & . & . & . & . & . & . & . & 19.0 \\ . & . & . & . & . & . & . & . & . & . & . & 17.0 \\ . & . & . & . & . & . & . & . & . & . & . & 18.0 \\ . & . & . & . & . & . & . & . & . & . & . & 20.0 \end{pmatrix}.$$

3.2 Centring

A variable $\mathbf{X}$ is called centred if $\mathbb{E}(\mathbf{X}) = 0$. Therefore a centred version of $\mathbf{X}$ is always obtained in theory by the transformation $\mathbf{X} - \mathbb{E}(\mathbf{X})$. In practice, however, the theoretical expectation of $\mathbf{X}$ is usually unknown and thus estimated by a sample mean.

- Section 3.2.1 briefly describes the two sample means implemented in the plug-in.
- Section 3.2.2 provides details on the centring pre-test which is performed automatically for any non-centred regression variable.
- Section 3.2.3 outlines the options available for centring and how they can be put into practice in the specification panel.

3.2.1 Sample means

The plug-in implements the global and period-specific sample means.

- The global sample mean of $\mathbf{X}$ is given by

$$\bar{\mathbf{X}}_g = \frac{1}{n} \sum_{j=1}^{\tau} \sum_{i=1}^{n_j} x_{i,j}. \tag{3.3}$$

- The vector of period-specific sample means is defined as $\bar{\mathbf{X}}_{ps} = \begin{pmatrix} \bar{\mathbf{X}}_1 & \cdots & \bar{\mathbf{X}}_\tau \end{pmatrix}^\top$, where

$$\bar{\mathbf{X}}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{i,j} \tag{3.4}$$

  for each $j \in \{1, \ldots, \tau\}$.

Combining equations (3.3) and (3.4), it follows immediately that

$$\bar{\mathbf{X}}_g = \sum_{j=1}^{\tau} w_j \bar{\mathbf{X}}_j, \tag{3.5}$$

where $w_j = n_j/n$. Thus the global sample mean of $\mathbf{X}$ always equals the weighted sum of period-specific sample means of $\mathbf{X}$ where the weight of each period is given by the period's share to the total number of observations.

3.2.2 Pre-test

The centring pre-test is performed automatically for any regression variable which is labelled either as original or as grouped in order to avoid centring of regression variables that have been centred already. The basic idea is to check whether the global sample mean and/or the vector of period-specific sample means of the variable in question is sufficiently close to zero. To this end, let $\varepsilon_{low} > 0$ and $\varepsilon_{upp} > 0$ with $\varepsilon_{low} < \varepsilon_{upp}$ be a lower and an upper threshold, respectively, which account for different levels of sufficiency and can be customised in the global options (see section 4). The pre-test then consists of two steps:

Table 3.1: Possible comments of the centring pre-test.

Pre-test column	*Displayed if*
Centred (seasonal means)	$\|\| \bar{\mathbf{X}}_{ps} \|\|_2 \leq \varepsilon_{low}$.
Probably centred (seasonal means)	$\varepsilon_{low} < \|\| \bar{\mathbf{X}}_{ps} \|\|_2 \leq \varepsilon_{upp}$.
Centred (global mean)	$\|\| \bar{\mathbf{X}}_{ps} \|\|_2 > \varepsilon_{upp}$ and $\| \bar{\mathbf{X}}_g \| \leq \varepsilon_{low}$.
Probably centred (global mean)	$\|\| \bar{\mathbf{X}}_{ps} \|\|_2 > \varepsilon_{upp}$ and $\varepsilon_{low} < \| \bar{\mathbf{X}}_g \| \leq \varepsilon_{upp}$.
Not centred	$\|\| \bar{\mathbf{X}}_{ps} \|\|_2 > \varepsilon_{upp}$ and $\| \bar{\mathbf{X}}_g \| > \varepsilon_{upp}$.

(1) It is checked whether $\mathbf{X}$ has been centred already around its period-specific sample means. This is assumed to be true if $|| \bar{\mathbf{X}}_{ps} ||_2 \leq \varepsilon_{low}$ or at least $|| \bar{\mathbf{X}}_{ps} ||_2 \leq \varepsilon_{upp}$, where $|| \cdot ||_2$ denotes the Euclidean norm. If so, the global sample mean of $\mathbf{X}$ is also close to zero according to equation (3.5) and, therefore, not checked separately.

(2) If $|| \bar{\mathbf{X}}_{ps} ||_2 > \varepsilon_{upp}$ holds in step (1), it is checked whether $\mathbf{X}$ has been centred already around its global sample mean. This is assumed to be true if $| \bar{\mathbf{X}}_g | \leq \varepsilon_{low}$ or at least $| \bar{\mathbf{X}}_g | \leq \varepsilon_{upp}$.

Depending on the outcome of the pre-test, a verbal comment is displayed in the respective column of the data panel. The possible comments are summarised in table 3.1. Note that any indication of centring will generate a warning message and the regression variable in question is not centred.

3.2.3 Calculations

For any original regression variable $\mathbf{X}$, the transformed variable centred around the global sample mean of $\mathbf{X}$ is obtained as:

$$\mathbf{X}_g^{Cen} = \begin{pmatrix} x_{1,1} - \bar{\mathbf{X}}_g & x_{1,2} - \bar{\mathbf{X}}_g & \cdots & x_{1,\tau} - \bar{\mathbf{X}}_g \\ x_{2,1} - \bar{\mathbf{X}}_g & x_{2,2} - \bar{\mathbf{X}}_g & \cdots & x_{2,\tau} - \bar{\mathbf{X}}_g \\ \vdots & \vdots & \ddots & \vdots \\ x_{n_1,1} - \bar{\mathbf{X}}_g & x_{n_2,2} - \bar{\mathbf{X}}_g & \cdots & x_{n_\tau,\tau} - \bar{\mathbf{X}}_g \end{pmatrix},$$

see also example 3.2 below. Similarly, the transformed variable centred around the period-specific sample means of $\mathbf{X}$ is given by:

$$\mathbf{X}_{ps}^{Cen} = \begin{pmatrix} x_{1,1} - \bar{\mathbf{X}}_1 & x_{1,2} - \bar{\mathbf{X}}_2 & \cdots & x_{1,\tau} - \bar{\mathbf{X}}_\tau \\ x_{2,1} - \bar{\mathbf{X}}_1 & x_{2,2} - \bar{\mathbf{X}}_2 & \cdots & x_{2,\tau} - \bar{\mathbf{X}}_\tau \\ \vdots & \vdots & \ddots & \vdots \\ x_{n_1,1} - \bar{\mathbf{X}}_1 & x_{n_2,2} - \bar{\mathbf{X}}_2 & \cdots & x_{n_\tau,\tau} - \bar{\mathbf{X}}_\tau \end{pmatrix},$$

see also example 3.3 below. For any grouped regression variable, these transformations are carried out only in the "active" periods and the NaN's introduced in the "inactive"

Table 3.2: Sample means for centring (default in boldface).

Sample mean	*Action*
None	No centring.
Global	Centring around global sample mean as defined in equation (3.3).
Seasonal	Centring around period-specific sample means as defined in equation (3.4).

Table 3.3: Calculation span for centring (default in boldface).

Type	*Start*	*End*
All	Start of regression variable	End of regression variable
From	User-defined date	End of regression variable
To	Start of regression variable	User-defined date
Between	User-defined date	User-defined date
Last	User-defined number of periods	End of regression variable
First	Start of regression variable	User-defined number of periods
Excluding	User-defined number of periods	User-defined number of periods

periods are replaced with zeros, see example 3.4 below. Note that the latter replacement is also applied to missing values in both original and grouped regression variables.
To carry out centring in a TransReg document, go to the ⊞ Centring section of the specification panel and use the drop down menu **Sample mean** in order to specify the type of sample mean. The implemented options are summarised in table 3.2. In case of centring around a global sample mean or period-specific sample means, the span used for calculating the selected sample mean can be further specified. To this end, click on ⊞ Calculation span, use the drop down menu **Type** to select a span category and, if necessary, provide the requested information. The implemented span categories are summarised in table 3.3. Eventually, confirm your choice by clicking on [Calculate] in the bottom line of the specification panel and the centred regression variable appears in the data panel under the following default name:

MyRegVar ► Centred.

If a set of G grouped regression variables has been created prior to centring, the default names of the centred daughter regression variables are:

MyRegVar ► Group1 ► Centred,
MyRegVar ► Group2 ► Centred,
...
MyRegVar ► GroupG ► Centred.

Example 3.2 (Centring around global sample mean) *Using the span from January 2011 to December 2020, the global sample mean of the monthly number of working days is (rounded to one decimal place) given by*

$$\bar{\mathbf{X}}_g = 20.8.$$

Accordingly, the monthly number of working days centred around this average reads

$$\mathbf{X}_g^{Cen} = \begin{pmatrix} -0.2 & -0.8 & 1.8 & -1.8 & 1.2 & -1.5 & 0.2 & 2.0 & 1.2 & -0.9 & 0.5 & 0.2 \\ 0.8 & -0.2 & 1.2 & -1.8 & -0.8 & -0.5 & 1.2 & 2.0 & -0.8 & 1.1 & 0.5 & -3.8 \\ 1.2 & -1.2 & -0.8 & 0.2 & -1.5 & -0.8 & 2.2 & 1.0 & 0.2 & 1.1 & -0.5 & -2.8 \\ 0.8 & -0.8 & -0.2 & -0.8 & -0.8 & -1.5 & 2.2 & 0.0 & 1.2 & 1.1 & -0.8 & -1.8 \\ -0.2 & -1.2 & 1.2 & -0.8 & -2.8 & 0.5 & 2.2 & 0.2 & 1.2 & 1.2 & 0.2 & -0.8 \\ -1.2 & -0.2 & 0.2 & 0.2 & -1.5 & 1.2 & 0.2 & 2.0 & 1.2 & -0.9 & 0.5 & 0.2 \\ 0.8 & -1.2 & 2.2 & -2.8 & 0.2 & -0.5 & 0.2 & 2.0 & 0.2 & -0.8 & 0.5 & -1.8 \\ 1.2 & -1.2 & 0.2 & -0.8 & -1.5 & 0.2 & 1.2 & 2.0 & -0.8 & 1.1 & 0.5 & -3.8 \\ 1.2 & -0.8 & -0.2 & -0.8 & 0.2 & -2.5 & 2.2 & 1.0 & 0.2 & 1.1 & -0.5 & -2.8 \\ 0.8 & -1.2 & 1.2 & -0.8 & -1.8 & -0.5 & 2.2 & 0.2 & 1.2 & 1.2 & 0.2 & -0.8 \end{pmatrix}.$$

What jumps immediately to the eye is the fact that the centred number of working days is always negative in February and positive or at least zero in July and August. This might be disadvantageous in some practical applications. Thus the use of period-specific sample means seems to be preferable in these cases.

Example 3.3 (Centring around month-specific sample means) *Using again the span from January 2011 to December 2020, the month-specific sample means of the number of working days are (again rounded to one decimal place) given by*

$$\bar{\mathbf{X}}_{ps} = \begin{pmatrix} 21.3 & 19.9 & 21.5 & 19.8 & 19.9 & 20.2 & 22.2 & 22.0 & 21.3 & 21.3 & 20.9 & 19.0 \end{pmatrix}^\top .$$

Thus the monthly number of working days centred around these averages is

$$\mathbf{X}_{ps}^{Cen} = \begin{pmatrix} -0.7 & 0.1 & 1.1 & -0.8 & 2.1 & -0.9 & -1.2 & 0.8 & 0.7 & -1.4 & 0.4 & 2.0 \\ 0.3 & 0.7 & 0.5 & -0.8 & 0.1 & 0.1 & -0.2 & 0.8 & -1.3 & 0.6 & 0.4 & -2.0 \\ 0.7 & -0.3 & -1.5 & 1.2 & -0.6 & -0.2 & 0.8 & -0.2 & -0.3 & 0.6 & -0.6 & -1.0 \\ 0.3 & 0.1 & -0.9 & 0.2 & 0.1 & -0.9 & 0.8 & -1.2 & 0.7 & 0.6 & -0.9 & 0.0 \\ -0.7 & -0.3 & 0.5 & 0.2 & -1.9 & 1.1 & 0.8 & -1.0 & 0.7 & 0.7 & 0.1 & 1.0 \\ -1.7 & 0.7 & -0.5 & 1.2 & -0.6 & 1.8 & -1.2 & 0.8 & 0.7 & -1.4 & 0.4 & 2.0 \\ 0.3 & -0.3 & 1.5 & -1.8 & 1.1 & 0.1 & -1.2 & 0.8 & -0.3 & -1.3 & 0.4 & 0.0 \\ 0.7 & -0.3 & -0.5 & 0.2 & -0.6 & 0.8 & -0.2 & 0.8 & -1.3 & 0.6 & 0.4 & -2.0 \\ 0.7 & 0.1 & -0.9 & 0.2 & 1.1 & -1.9 & 0.8 & -0.2 & -0.3 & 0.6 & -0.6 & -1.0 \\ 0.3 & -0.3 & 0.5 & 0.2 & -0.9 & 0.1 & 0.8 & -1.0 & 0.7 & 0.7 & 0.1 & 1.0 \end{pmatrix} .$$

As opposed to the use of the global sample mean in example 3.2, this centred number of working days has positive and negative values in each month.

To demonstrate the impact of the span used for calculating the sample means, we now shorten this span to the period from January 2015 to December 2020 and note that the month-specific sample means change visibly, especially in May and June, to

$$\bar{\mathbf{X}}_{ps} = \begin{pmatrix} 21.2 & 19.8 & 21.6 & 19.8 & 19.6 & 20.5 & 22.2 & 22.0 & 21.3 & 21.3 & 21.0 & 19.2 \end{pmatrix}^\top .$$

As a result, the centred monthly number of working days changes visibly as well. It now reads

$$\mathbf{X}_{ps}^{Cen} = \begin{pmatrix} -0.6 & 0.2 & 1.2 & -0.8 & 2.4 & -1.2 & -1.2 & 0.8 & 0.7 & -1.4 & 0.3 & 1.8 \\ 0.4 & 0.8 & 0.4 & -0.8 & 0.4 & -0.2 & -0.2 & 0.8 & -1.3 & 0.6 & 0.3 & -2.2 \\ 0.8 & -0.2 & -1.6 & 1.2 & -0.3 & -0.5 & 0.8 & -0.2 & -0.3 & 0.6 & -0.7 & -1.2 \\ 0.4 & 0.2 & -1.0 & 0.2 & 0.5 & -1.2 & 0.8 & -1.2 & 0.7 & 0.6 & -1.0 & -0.2 \\ -0.6 & -0.2 & 0.4 & 0.2 & -1.6 & 0.8 & 0.8 & -1.0 & 0.7 & 0.7 & 0.0 & 0.8 \\ -1.6 & 0.8 & -0.6 & 1.2 & -0.9 & 1.5 & -1.2 & 0.8 & 0.7 & -1.4 & 0.3 & 1.8 \\ 0.4 & -0.2 & 1.4 & -1.8 & 1.4 & -0.2 & -1.2 & 0.8 & -0.3 & -1.3 & 0.3 & -0.2 \\ 0.8 & -0.2 & -0.6 & 0.2 & -0.3 & 0.5 & -0.2 & 0.8 & -1.3 & 0.6 & 0.3 & -2.2 \\ 0.8 & 0.2 & -1.0 & 0.2 & 1.4 & -1.6 & 0.8 & -0.2 & -0.3 & 0.6 & -0.7 & -1.2 \\ 0.4 & -0.2 & 0.4 & 0.2 & -0.6 & -0.2 & 0.8 & -1.0 & 0.7 & 0.7 & 0.0 & 0.8 \end{pmatrix} ,$$

exhibiting even some sign changes compared to the centred version calculated above.

Example 3.4 (Grouping and centring around period-specific sample means) *In example 3.1, the monthly number of working days has been split into two daughter regression variables: one for the months January through November and one for December. Using the month-specific sample means calculated from the entire observation span in example 3.3, the centred versions of the grouped regression variables are given by*

$$\mathbf{X}_{1,ps}^{Cen} = \begin{pmatrix} -0.7 & 0.1 & 1.1 & -0.8 & 2.1 & -0.9 & -1.2 & 0.8 & 0.7 & -1.4 & 0.4 & 0.0 \\ 0.3 & 0.7 & 0.5 & -0.8 & 0.1 & 0.1 & -0.2 & 0.8 & -1.3 & 0.6 & 0.4 & 0.0 \\ 0.7 & -0.3 & -1.5 & 1.2 & -0.6 & -0.2 & 0.8 & -0.2 & -0.3 & 0.6 & -0.6 & 0.0 \\ 0.3 & 0.1 & -0.9 & 0.2 & 0.1 & -0.9 & 0.8 & -1.2 & 0.7 & 0.6 & -0.9 & 0.0 \\ -0.7 & -0.3 & 0.5 & 0.2 & -1.9 & 1.1 & 0.8 & -1.0 & 0.7 & 0.7 & 0.1 & 0.0 \\ -1.7 & 0.7 & -0.5 & 1.2 & -0.6 & 1.8 & -1.2 & 0.8 & 0.7 & -1.4 & 0.4 & 0.0 \\ 0.3 & -0.3 & 1.5 & -1.8 & 1.1 & 0.1 & -1.2 & 0.8 & -0.3 & -1.3 & 0.4 & 0.0 \\ 0.7 & -0.3 & -0.5 & 0.2 & -0.6 & 0.8 & -0.2 & 0.8 & -1.3 & 0.6 & 0.4 & 0.0 \\ 0.7 & 0.1 & -0.9 & 0.2 & 1.1 & -1.9 & 0.8 & -0.2 & -0.3 & 0.6 & -0.6 & 0.0 \\ 0.3 & -0.3 & 0.5 & 0.2 & -0.9 & 0.1 & 0.8 & -1.0 & 0.7 & 0.7 & 0.1 & 0.0 \end{pmatrix}$$

and

$$\mathbf{X}_{2,ps}^{Cen} = \begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 2.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -2.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -1.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 2.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -2.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & -1.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \end{pmatrix}.$$

Note that the process of centring replaces automatically the NaN's in each grouped monthly number of working days with zeros.

4 Global options

The global option menu enables some degree of customisation of the plug-in. It can be assessed by selecting

Tools ► Options

from the JD+ drop down menu and clicking on the **TransReg** sheet of the **Demetra** window. The implemented global options are summarised in table 4.1.

Table 4.1: TransReg global options.

Option	*Variable*	*Valid choices*	*Defaults*
Maximum number of groups	The maximum number of period-specific daughter regression variables, g_{max}, to be allowed for grouping a mother regression variable. It applies to monthly and quarterly data alike. If $g_{max} > 4$ and a quarterly regression variable is considered, then $g_{max} = 4$ is automatically used in this particular case.	$g_{max} \in \{2, \ldots, 12\}$	$g_{max} = 2$
Thresholds for pre-test	The exponents k_{upp} and k_{low} to be used for calculating the upper and lower thresholds according to $$\varepsilon_{upp} = 10^{-k_{upp}}$$ and $$\varepsilon_{low} = 10^{-k_{low}}.$$ Any choice where $k_{upp} \geq k_{low}$ will generate a warning message and cannot be saved.	$k_{upp} \in \{0, \ldots, 99\}$ $k_{low} \in \{1, \ldots, 100\}$	$k_{upp} = 4$ $k_{low} = 12$

Bibliography

Deutsche Bundesbank (2012). “Calendar effects on economic activity”. In: *Monthly Report* 64.12, pp. 51–60.