
User Defined Regressors with GenHol

Seasonal Adjustment With X-13ARIMA-SEATS

2019

Economic Statistical Methods Division

U.S. Census Bureau

Moving holiday effects

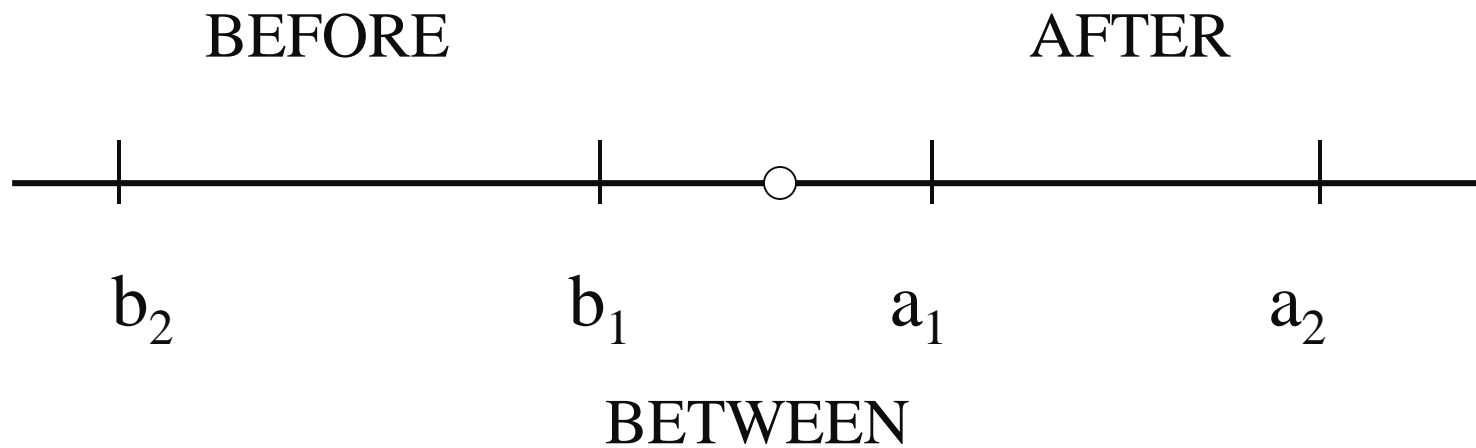
- Caused when dates for a given holiday change from year to year.
- X-13ARIMA-SEATS has built-in regressors to model moving holidays found to effect the U.S. economy:
 - Easter
 - Thanksgiving
 - Labor Day

Moving holiday effects

- How do we incorporate other effects, such as
 - Chinese New Year,
 - Ramadan,
 - Easter Monday,
 - Local festivals?
- Must generate user defined regressors to capture these effects

General schema

- For moving holidays, there may be effects
 - Before the holiday,
 - After the holiday,
 - Surrounding the holiday.
- Generate proportional regressors for intervals before and after the holiday where effect is likely to be strongest



“Before” holiday effect

- Define an interval that
 - starts b_1 days before and
 - ends b_2 days before the holiday
- Generate a regressor that is the proportion of the days within the interval for a given month or quarter.

Example

- Assume :
 - an interval 10 days long that starts 15 days before the holiday (Chinese New Year)
 - two of the days fall in January for a given year
- The regressor for that year will be:
 - 0.2 in January
 - 0.8 in February
 - 0.0 for the rest of the year

“After” holiday effect

- Define an interval that
 - starts a_1 days after and
 - ends a_2 days after the holiday
- Generate proportional regressor for the “after” effect, just as with the “before”, using this interval

“Surrounding” holiday effect

- Define an interval that
 - starts b_1+1 days before and
 - ends a_1-1 days after the holiday
- Generate proportional regressor as before.

Centering

- Done to keep
 - $\text{Total}(\text{OH}) \cong \text{Total}(\text{Orig})$
- Where
 - $\text{Total}(\text{Orig})$ is the yearly total of the original series
 - $\text{Total}(\text{OH})$ is the yearly total of the series with estimated holiday effects removed
- Also want to remove long term seasonal effects from the holiday effect

How to Center?

- Remove calendar month means of the regressor if
 - the holiday can occur only in a few calendar months
 - Example: Chinese New Year
- Remove the overall mean of the regressor if
 - The holiday moves through all of the calendar months
 - Example: Ramadan

1991	1	0.0	0.0	0.0
1991	2	1.0	1.0	0.9
1991	3	0.0	0.0	0.1
1991	4	0.0	0.0	0.0
1991	5	0.0	0.0	0.0
1991	6	0.0	0.0	0.0
1991	7	0.0	0.0	0.0
1991	8	0.0	0.0	0.0
1991	9	0.0	0.0	0.0
1991	10	0.0	0.0	0.0
1991	11	0.0	0.0	0.0
1991	12	0.0	0.0	0.0
1992	1	0.7	0.0	0.0
1992	2	0.3	1.0	1.0
1992	3	0.0	0.0	0.0
1992	4	0.0	0.0	0.0
1992	5	0.0	0.0	0.0
1992	6	0.0	0.0	0.0
1992	7	0.0	0.0	0.0
1992	8	0.0	0.0	0.0
1992	9	0.0	0.0	0.0
1992	10	0.0	0.0	0.0
1992	11	0.0	0.0	0.0
1992	12	0.0	0.0	0.0
1993	1	1.0	1.0	0.4
1993	2	0.0	0.0	0.6
1993	3	0.0	0.0	0.0

Chinese New Year
Regressors,
Before Centering

1991	1	-0.5373	-0.2787	-0.0567
1991	2	0.5373	0.2787	0.0093
1991	3	0.0	0.0	0.0473
1991	4	0.0	0.0	0.0
1991	5	0.0	0.0	0.0
1991	6	0.0	0.0	0.0
1991	7	0.0	0.0	0.0
1991	8	0.0	0.0	0.0
1991	9	0.0	0.0	0.0
1991	10	0.0	0.0	0.0
1991	11	0.0	0.0	0.0
1991	12	0.0	0.0	0.0
1992	1	0.1627	-0.2787	-0.05667
1992	2	-0.1627	0.2787	0.1093
1992	3	0.0	0.0	-0.05267
1992	4	0.0	0.0	0.0
1992	5	0.0	0.0	0.0
1992	6	0.0	0.0	0.0
1992	7	0.0	0.0	0.0
1992	8	0.0	0.0	0.0
1992	9	0.0	0.0	0.0
1992	10	0.0	0.0	0.0
1992	11	0.0	0.0	0.0
1992	12	0.0	0.0	0.0
1993	1	0.4627	0.7213	0.3433
1993	2	-0.4627	-0.7213	-0.2907
1993	3	0.0	0.0	-0.05267

Chinese New Year
Regressors,
After Centering

Features of genhol

- Generates user-defined holiday regressors effects for intervals
 - before,
 - between,
 - and after holidays
- More than one holiday can be specified
 - A file of dates for each holiday is required.

Features of genhol

- Centers holiday regressors
 - Overall mean
 - Calendar mean
- Generates X-13ARIMA-SEATS commands that use the holiday regressors

genhol input files

- Before running genhol, an input file must be created
 - An ASCII file used to specify program options.
- Each file is made up of functional units called specifications (or “specs”)

General Input Syntax for genhol

```
spec1 {  
  argument = value  
  argument = "string"  
}  
  
spec2 {  
  argument = value  
  argument = "string"  
}
```

Two types of “specs”

- Global spec
 - Specify global options such as the output name
 - Always appears first
- Holiday specs
 - Separate holiday specs for each holiday

To run genhol

- Enter in Windows command prompt:

- ***genhol infile***

- Where *infile* is the genhol input file name

- Example:

- ***genhol cny.inp***

Sample global spec

```
global{  
  numhol = 3  #number of holidays defined in this input file  
  outfile = "chinahol.dat"  #output regressor file  
  outspec = "chinahol.reg"  #file for X-13 regression  
  commands  
  first = 1991  #holiday regressors start in this year  
  last = 2010  #and end in this year  
  firstmean = 1900  #holiday centering uses means of data  
  lastmean = 2100  #from years firstmean to lastmean  
  period = 4  }
```

Holiday spec(s)

- After the global spec
- Each holiday must have a separate spec
 - As many as specified by the numhol variable
 - Each holiday spec should be numbered (holiday1, holiday2)

Sample Holiday Specs

```
holiday1{
    name = cny
    begbefore = -10
    endbefore = -1
    infile = "cny.dat"
    center = calendar
}
holiday2{
    name = moon
    begbefore = -10
    endbefore = -1
    infile = "cny2.dat"
    center = calendar
}
holiday3{
    name = midfall
    begbefore = -10
    endbefore = -1
    infile = "cny3.dat"
    center = calendar
}
```

Three holiday regressors will be created in one file. Each has a 10-day holiday interval, starting ten days before the holiday and ending one day before the holiday. Each is centered only in the months the holiday can be in.

Infile is the file containing the holidays' dates.

Another Example of a Holiday Spec

```
holiday1{  
  name = IdulFitri  
  infile = "IFHolInd.dat"  
  begbefore = -42  
  endbefore = -21  
  begafter = 1  
  endafter = 7  
  center = mean  
}
```

Creates a 3-part regressor for the holiday: from 42 days before to 21 days before; from 20 days before to the day of; and from the day after to 7 days after.

The holiday is centered using the overall mean.

Date file for holiday

- ASCII file
- One date per line
- Each date consists of
 - Month
 - Day
 - Four digit year
- Separate file for each holiday

Example : cny.dat (Chinese New Year)

2	15	1991
2	4	1992
1	23	1993
2	10	1994
1	31	1995
2	19	1996
2	7	1997
1	28	1998
2	16	1999

To generate specific holiday regressors

- “before” holiday effect
 - Must specify **begbefore** and **endbefore**
- “after” holiday effect
 - Must specify **begafter** and **endafter**

“Between” Holiday Effects

- Must specify endbefore and begafter
 - There must be at least two observations in this window
 - If you only want a “between” effect, specify only endbefore and begafter

X-13A-S regression commands generated by genhol

```
regression{  
  user=(  
    Beforecny      Betweencny      Aftercny  
    Beforemoon     Betweenmoon     Aftermoon  
    Beforemidfall  Betweenmidfall  Aftermidfall  
  )  
  file="chinahol.dat"  
  format="datevalue"  
  start=1991.1  
  usertype=holiday
```

Holiday regressors for stock series

- Findley (2009) develops a method of generating moving holiday regressors for inventory series
 - Uses the view of stock series as accumulations of flow series to construct moving holiday regressors
- Genhol can generate moving holiday regressors for inventory series
 - Need to specify the stock day

Input file to generate stock holiday regressors

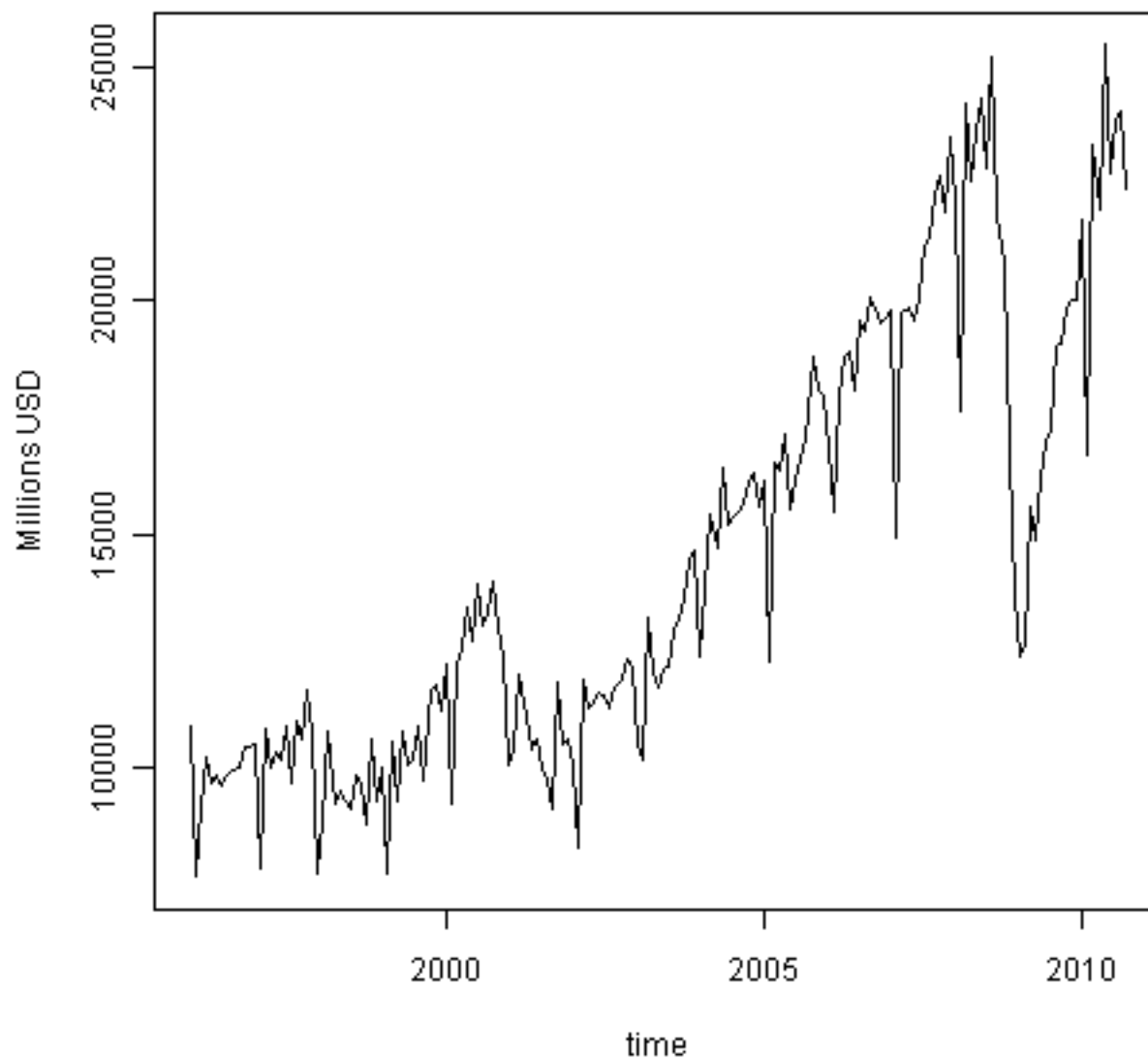
```
global{
    numhol = 1
    outfile = "cnystock1b1a.dat"
    outspec = "cnystock1b1a.reg"
    first = 1981
    last = 2020
    stockday = 31
}
holiday1{
    name = cnystock
    begbefore = -7
    endbefore = -1
    begafter = 7
    infile = "cny.dat"
    center = calendar
}
```

1981	1	-0.0580952380952382	-0.2457142857142857
1981	2	0.0	0.0
1981	3	0.0	0.0
1981	4	0.0	0.0
1981	5	0.0	0.0
1981	6	0.0	0.0
1981	7	0.0	0.0
1981	8	0.0	0.0
1981	9	0.0	0.0
1981	10	0.0	0.0
1981	11	0.0	0.0
1981	12	0.0	0.0
1982	1	0.5133333333333332	0.7542857142857143
1982	2	0.0	0.0
1982	3	0.0	0.0
1982	4	0.0	0.0
1982	5	0.0	0.0
1982	6	0.0	0.0
1982	7	0.0	0.0
1982	8	0.0	0.0
1982	9	0.0	0.0
1982	10	0.0	0.0
1982	11	0.0	0.0
1982	12	0.0	0.0
1983	1	-0.48666666666666668	-0.2457142857142857
1983	2	0.0	0.0
1983	3	0.0	0.0

Chinese New Year Regressors for Stock Series

Case Study

Taiwan exports



Exports from Taiwan Series

- Three holiday models
 - No holiday regressors
 - Chinese new year regressors
 - Chinese new year, Midfall festival, and Dragon Boat festival regressors
- Everything else about the regARIMA model will be the same (trading day, outliers, ARIMA model)

What do we need to run genhol?

- Date file for holidays
- Input files for genhol
 - Need to give some thought to the intervals that define the before, during and after regressors

My choices

Holiday	“Before” interval	“Between” interval	“After” interval
Chinese New Year	Two weeks before the holiday	The week that starts on the holiday	The week that starts one week after the holiday
Midfall Festival	One week before the holiday	The week that starts on the holiday	The week that starts one week after the holiday
Dragon Boat Festival	One week before the holiday	The week that starts on the holiday	The week that starts one week after the holiday

Genhol input for Chinese New Year only

```
global{
  numhol = 1
  outfile = "taiwan2.dat"
  outspec = "taiwan2.reg"
  first = 1981
  last = 2020
}
holiday1{
  name = cny
  begbefore = -14
  endbefore = -1
  begafter = 7
  endafter = 13
  infile = "cny.dat"
  center = calendar
}
```

Genhol Input for All Holidays

```
global{
  numhol = 3
  outfile =
  "taiwan2HolAll2.dat"
  outspec =
  "taiwan2HolAll2.reg"
  first = 1981
  last = 2020
  usergroup = yes
}
holiday1{
  name = cny
  begbefore = -14
  endbefore = -1
  begafter = 7
  endafter = 13
  infile = "cny.dat"
  center = calendar
}
```

```
holiday2{
  name = dragon
  begbefore = -7
  endbefore = -1
  begafter = 7
  endafter = 13
  infile = "dragon.dat"
  center = calendar
}
holiday3{
  name = midfall
  begbefore = -7
  endbefore = -1
  begafter = 7
  endafter = 13
  infile = "midfall.dat"
  center = calendar
}
```

taiwan2HolAll2.reg

```
regression{
  user=(
    Beforecny      Betweencny      Aftercny
    Beforedragon   Betweenragon   Afterdragon
    Beforemidfall  Betweenmidfall  Aftermidfall
  )
  file="taiwanholall2.dat"
  format="datevalue"
  start=1981.1
  usertype=(
    holiday        holiday        holiday
    holiday2       holiday2       holiday2
    holiday3       holiday3       holiday3
  )
}
```

AICC results for Exports

Holiday Regressors	AICC
None	2723.28613
Chinese New Year Only	2655.04785
Chinese New Year, Midfall, Dragon Boat	2647.19727

However

- When we look at the output for the model with regressors for all the holidays, we see
 - The Dragon Boat group is insignificant;
 - T-statistics for the “after holiday” regressors are insignificant
- Can we improve the model by
 - Dropping the Dragon Boat Festival regressors
 - Not including the “after holiday”

Genhol Input for 2 Holidays (no after)

```
global{
    numhol = 2
    outfile =
    "taiwan2Hol2c.dat"
    outspec =
    "taiwan2Hol2c.reg"
    first = 1981
    last = 2020
    usergroup = yes
}
holiday1{
    name = cny
    begbefore = -14
    endbefore = -1
    begafter = 7
    infile = "cny.dat"
    center = calendar
}

holiday2{
    name = midfall
    begbefore = -7
    endbefore = -1
    begafter = 7
    infile = "midfall.dat"
    center = calendar
}
```

AICC results for Exports

Holiday Regressors	AICC
None	2723.28613
Chinese New Year Only	2655.04785
Chinese New Year, Midfall, Dragon Boat	2647.19727
Chinese New Year, Midfall	2643.24365
Chinese New Year, Midfall (no after regressors)	2638.2146

For More Information

- “Modeling Lunar Calendar Holiday Effects in Taiwan,” by Jin-Lung Lin and Tian-Syh Liu

<http://www.census.gov/srd/www/sapaper/sapaper.html>

Genhol in R seasonal

- The genhol program was rewritten in R as part of the seasonal package
- Run with the genhol function, one regressor at a time
- Arguments:
 - x: a date vector, dates of the holiday
 - center: “calendar”, “mean”, or “none” (default)
 - start: number of days before (negative number) or after (positive number) the holiday to start the effect
 - end: number of days before (negative number) or after (positive number) the holiday to end the effect
 - frequency
- Run data(holiday) to load date vectors of Easter (easter), Chinese New Year (cny), and Diwali (diwali)

Chinese New Year with genhol() function

- To create the regressors that genhol would produce with

```
holiday1{ name = cny begbefore = -14  
          endbefore = -1 begafter = 7  
          endafter = 13 ... center = calendar  
}
```

use:

```
data(holiday)
```

```
before.cny <- genhol(cny,start=-14,end=-1,center="calendar")
```

```
during.cny<-genhol(cny,start=0,end=6,center="calendar")
```

```
after.cny<-genhol(cny,start=7,end=13,center="calendar")
```

Running genhol()'s Chinese New Year

To run all three regressors:

```
m <- seas(out, xreg=cbind(before.cny,during.cny,after.cny),  
regression.usertype="holiday", regression.aictest=c("td","user") )
```

To run only the during effect:

```
m <- seas(out, xreg=during.cny, regression.usertype="holiday",  
regression.aictest=c("td","user") )
```