Title

tssmooth hwinters — Holt–Winters nonseasonal smoothing

Description

`tssmooth hwinters` is used in smoothing or forecasting a series that can be modeled as a linear trend in which the intercept and the coefficient on time vary over time.

Quick start

Create `smooth` using Holt–Winters nonseasonal smoothing over `y` with `tsset` data

```
tssmooth hwinters smooth=y
```

Same as above, but forecast 10 periods out of sample

```
tssmooth hwinters smooth=y, forecast(10)
```

Same as above, but use 111 and 112 as the initial values for the recursion

```
tssmooth hwinters smooth=y, forecast(10) s0(111 112)
```

Same as above, but use 0.5 and 0.3 as the smoothing parameters

```
tssmooth hwinters smooth=y, forecast(10) s0(111 112) parms(.5 .3)
```

Note: The above commands can also be used to apply the smoother separately to each panel of a panel dataset when a *panelvar* has been specified using `tsset` or `xtset`.

Menu

Statistics > Time series > Smoothers/univariate forecasters > Holt–Winters nonseasonal smoothing

Syntax

```
tssmooth hwinters [type] newvar = exp [if] [in] [, options]
```

options	Description
Main	
`replace`	replace *newvar* if it already exists
`parms(`$\#_\alpha$ $\#_\beta$`)`	use $\#_\alpha$ and $\#_\beta$ as smoothing parameters
`samp0(#)`	use # observations to obtain initial values for recursion
`s0(`$\#_{\text{cons}}$ $\#_{\text{lt}}$`)`	use $\#_{\text{cons}}$ and $\#_{\text{lt}}$ as initial values for recursion
`forecast(#)`	use # periods for the out-of-sample forecast
Options	
`diff`	alternative initial-value specification; see *Options*
Maximization	
maximize_options	control the maximization process; seldom used
`from(`$\#_\alpha$ $\#_\beta$`)`	use $\#_\alpha$ and $\#_\beta$ as starting values for the parameters

You must `tsset` your data before using `tssmooth hwinters`; see **[TS] tsset**.

exp may contain time-series operators; see **[U] 11.4.4 Time-series varlists**.

`collect` is allowed; see **[U] 11.1.10 Prefix commands**.

Options

Main

`replace` replaces *newvar* if it already exists.

`parms(`$\#_\alpha$ $\#_\beta$`)`, $0 \le \#_\alpha \le 1$ and $0 \le \#_\beta \le 1$, specifies the parameters. If `parms()` is not specified, the values are chosen by an iterative process to minimize the in-sample sum-of-squared prediction errors.

If you experience difficulty converging (many iterations and "not concave" messages), try using `from()` to provide better starting values.

`samp0(#)` and `s0(`$\#_{\text{cons}}$ $\#_{\text{lt}}$`)` specify how the initial values $\#_{\text{cons}}$ and $\#_{\text{lt}}$ for the recursion are obtained.

By default, initial values are obtained by fitting a linear regression with a time trend using the first half of the observations in the dataset.

`samp0(#)` specifies that the first # observations be used in that regression.

`s0(`$\#_{\text{cons}}$ $\#_{\text{lt}}$`)` specifies that $\#_{\text{cons}}$ and $\#_{\text{lt}}$ be used as initial values.

`forecast(#)` specifies the number of periods for the out-of-sample prediction; $0 \le \# \le 500$. The default is `forecast(0)`, which is equivalent to not performing an out-of-sample forecast.

Options

`diff` specifies that the linear term is obtained by averaging the first difference of exp_t and the intercept is obtained as the difference of *exp* in the first observation and the mean of D.exp_t.

If the `diff` option is not specified, a linear regression of exp_t on a constant and t is fit.

Maximization

maximize_options controls the process for solving for the optimal α and β when `parms()` is not specified.

maximize_options: `nodifficult`, `technique(`*algorithm_spec*`)`, `iterate(`#`)`, [`no`]`log`, `trace`, `gradient`, `showstep`, `hessian`, `showtolerance`, `tolerance(`#`)`, `ltolerance(`#`)`, `nrtolerance(`#`)`, and `nonrtolerance`; see [R] **Maximize**. These options are seldom used.

`from(`$\#_\alpha$ $\#_\beta$`)`, $0 < \#_\alpha < 1$ and $0 < \#_\beta < 1$, specifies starting values from which the optimal values of α and β will be obtained. If `from()` is not specified, `from(.5 .5)` is used.

Remarks and examples

stata.com

The Holt–Winters method forecasts series of the form

$$\widehat{x}_{t+1} = a_t + b_t t$$

where $\widehat{x}_t$ is the forecast of the original series x_t, a_t is a mean that drifts over time, and b_t is a coefficient on time that also drifts. In fact, as Gardner (1985) has noted, the Holt–Winters method produces optimal forecasts for an ARIMA(0,2,2) model and some local linear models. See [TS] **arima** and the references in that entry for ARIMA models, and see Harvey (1989) for a discussion of the local linear model and its relationship to the Holt–Winters method. Abraham and Ledolter (1983), Bowerman, O'Connell, and Koehler (2005), and Montgomery, Johnson, and Gardiner (1990) all provide good introductions to the Holt–Winters method. Chatfield (2001, 2004) provides helpful discussions of how this method relates to modern time-series analysis.

The Holt–Winters method can be viewed as an extension of double-exponential smoothing with two parameters, which may be explicitly set or chosen to minimize the in-sample sum-of-squared forecast errors. In the latter case, as discussed in *Methods and formulas*, the smoothing parameters are chosen to minimize the in-sample sum-of-squared forecast errors plus a penalty term that helps to achieve convergence when one of the parameters is too close to the boundary.

Given the series x_t, the smoothing parameters α and β, and the starting values a_0 and b_0, the updating equations are

$$a_t = \alpha x_t + (1-\alpha)\left(a_{t-1} + b_{t-1}\right)$$

$$b_t = \beta\left(a_t - a_{t-1}\right) + (1-\beta)\, b_{t-1}$$

After computing the series of constant and linear terms, a_t and b_t, respectively, the τ-step-ahead prediction of x_t is given by

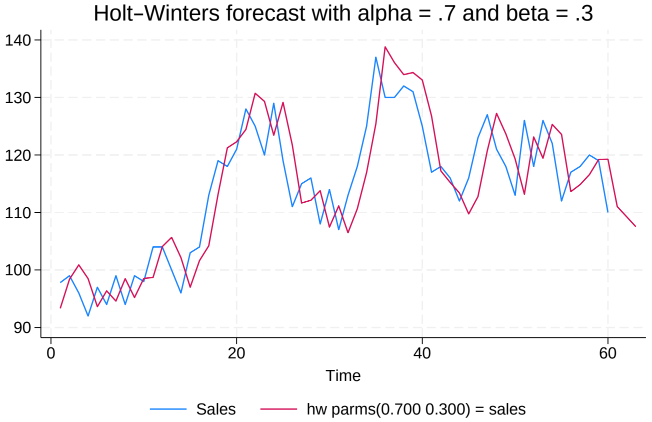
$$\widehat{x}_{t+\tau} = a_t + b_t \tau$$

▷ Example 1: Smoothing a series for specified parameters

Below we show how to use `tssmooth hwinters` with specified smoothing parameters. This example also shows that the Holt–Winters method can closely follow a series in which both the mean and the time coefficient drift over time.

Suppose that we have data on the monthly sales of a book and that we want to forecast this series with the Holt–Winters method.

```
. use https://www.stata-press.com/data/r18/bsales
. tssmooth hwinters hw1=sales, parms(.7 .3) forecast(3)
Specified weights:
                   alpha = 0.7000
                    beta = 0.3000
sum-of-squared residuals = 2301.046
 root mean squared error = 6.192799
. tsline sales hw1, title("Holt-Winters forecast with alpha = .7 and beta = .3")
```



The graph indicates that the forecasts are for linearly decreasing sales. Given a_T and b_T, the out-of-sample predictions are linear functions of time. In this example, the slope appears to be too steep, probably because our choice of α and β.

◁

▷ Example 2: Choosing the initial values

The graph in the previous example illustrates that the starting values for the linear and constant series can affect the in-sample fit of the predicted series for the first few observations. The previous example used the default method for obtaining the initial values for the recursion. The output below illustrates that, for some problems, the differenced-based initial values provide a better in-sample fit for the first few observations. However, the differenced-based initial values do not always outperform the regression-based initial values. Furthermore, as shown in the output below, for series of reasonable length, the predictions produced are nearly identical.

```
. tssmooth hwinters hw2=sales, parms(.7 .3) forecast(3) diff
Specified weights:
                   alpha = 0.7000
                    beta = 0.3000
sum-of-squared residuals = 2261.173
 root mean squared error = 6.13891
. list hw1 hw2 if _n<6 | _n>57

      +----------------------+
      |      hw1         hw2 |
      |----------------------|
  1.  | 93.31973    97.80807 |
  2.  | 98.40002    98.11447 |
  3.  | 100.8845     99.2267 |
  4.  | 98.50404    96.78276 |
  5.  | 93.62408     92.2452 |
      |----------------------|
 58.  | 116.5771    116.5771 |
 59.  | 119.2146    119.2146 |
 60.  | 119.2608    119.2608 |
 61.  | 111.0299    111.0299 |
 62.  | 109.2815    109.2815 |
      |----------------------|
 63.  | 107.5331    107.5331 |
      +----------------------+
```

When the smoothing parameters are chosen to minimize the in-sample sum-of-squared forecast errors, changing the initial values can affect the choice of the optimal α and β. When changing the initial values results in different optimal values for α and β, the predictions will also differ.

◁

When the Holt–Winters model fits the data well, finding the optimal smoothing parameters generally proceeds well. When the model fits poorly, finding the α and β that minimize the in-sample sum-of-squared forecast errors can be difficult.

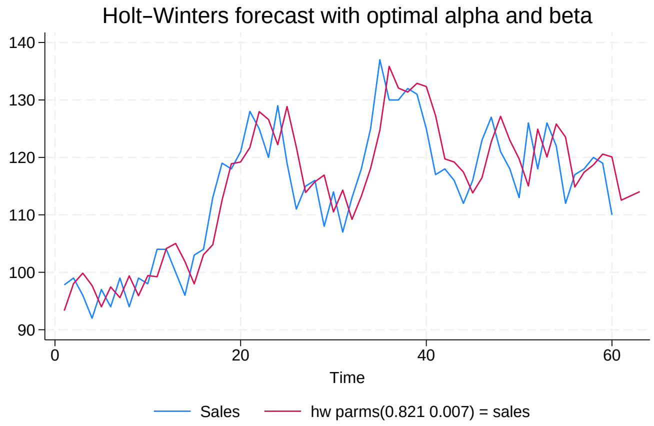
▷ Example 3: Forecasting with optimal parameters

In this example, we forecast the book sales data using the α and β that minimize the in-sample squared forecast errors.

```
. tssmooth hwinters hw3=sales, forecast(3)
computing optimal weights
Iteration 0:  Penalized RSS = -2632.2073  (not concave)
Iteration 1:  Penalized RSS = -1982.8431
Iteration 2:  Penalized RSS = -1976.4236
Iteration 3:  Penalized RSS = -1975.9172
Iteration 4:  Penalized RSS = -1975.9036
Iteration 5:  Penalized RSS = -1975.9036
Optimal weights:
                             alpha = 0.8209
                              beta = 0.0067
penalized sum-of-squared residuals = 1975.904
          sum-of-squared residuals = 1975.904
           root mean squared error = 5.738617
```

The following graph contains the data and the forecast using the optimal α and β. Comparing this graph with the one above illustrates how different choices of α and β can lead to very different forecasts. Instead of linearly decreasing sales, the new forecast is for linearly increasing sales.

```
. tsline sales hw3, title("Holt-Winters forecast with optimal alpha and beta")
```



◁

Stored results

tssmooth hwinters stores the following in r():

Scalars

r(N)	number of observations	r(N_pre)	number of observations used in calculating starting values
r(alpha)	α smoothing parameter		
r(beta)	β smoothing parameter	r(s2_0)	initial value for linear term
r(rss)	sum-of-squared errors	r(s1_0)	initial value for constant term
r(prss)	penalized sum-of-squared errors, if parms() not specified	r(linear)	final value of linear term
		r(constant)	final value of constant term
r(rmse)	root mean squared error		

Macros

r(method)	smoothing method	r(timevar)	time variables specified in tsset
r(exp)	expression specified	r(panelvar)	panel variables specified in tsset

Methods and formulas

A truncated description of the specified Holt–Winters filter is used to label the new variable. See [D] **label** for more information on labels.

An untruncated description of the specified Holt–Winters filter is saved in the characteristic named `tssmooth` for the new variable. See [P] **char** for more information on characteristics.

Given the series, x_t; the smoothing parameters, α and β; and the starting values, a_0 and b_0, the updating equations are

$$a_t = \alpha x_t + (1-\alpha)\,(a_{t-1} + b_{t-1})$$

$$b_t = \beta\,(a_t - a_{t-1}) + (1-\beta)\,b_{t-1}$$

By default, the initial values are found by fitting a linear regression with a time trend. The time variable in this regression is normalized to equal one in the first period included in the sample. By default, one-half of the data is used in this regression, but this sample can be changed using `samp0()`. a_0 is then set to the estimate of the constant, and b_0 is set to the estimate of the coefficient on the time trend. Specifying the `diff` option sets b_0 to the mean of D.x and a_0 to $x_1 - b_0$. `s0()` can also be used to specify the initial values directly.

Sometimes, one or both of the optimal parameters may lie on the boundary of $[\,0,1\,]$. To keep the estimates inside $[\,0,1\,]$, `tssmooth hwinters` parameterizes the objective function in terms of their inverse logits, that is, in terms of $\exp(\alpha)/\{1+\exp(\alpha)\}$ and $\exp(\beta)/\{1+\exp(\beta)\}$. When one of these parameters is actually on the boundary, this can complicate the optimization. For this reason, `tssmooth hwinters` optimizes a penalized sum-of-squared forecast errors. Let $\widehat{x}_t(\widetilde{\alpha},\widetilde{\beta})$ be the forecast for the series x_t, given the choices of $\widetilde{\alpha}$ and $\widetilde{\beta}$. Then the in-sample penalized sum-of-squared prediction errors is

$$P = \sum_{t=1}^{T} \left[\{x_t - \widehat{x}_t(\widetilde{\alpha},\widetilde{\beta})\}^2 + I_{|f(\widetilde{\alpha})|>12)}(|f(\widetilde{\alpha})| - 12)^2 + I_{|f(\widetilde{\beta})|>12)}(|f(\widetilde{\beta})| - 12)^2 \right]$$

where $f(x) = \ln\{x(1-x)\}$. The penalty term is zero unless one of the parameters is close to the boundary. When one of the parameters is close to the boundary, the penalty term will help to obtain convergence.

Acknowledgment

We thank Nicholas J. Cox of the Department of Geography at Durham University, UK, who is coeditor of the *Stata Journal* and author of *Speaking Stata Graphics* for his helpful comments.

References

Abraham, B., and J. Ledolter. 1983. *Statistical Methods for Forecasting*. New York: Wiley.

Bowerman, B. L., R. T. O'Connell, and A. B. Koehler. 2005. *Forecasting, Time Series, and Regression: An Applied Approach*. 4th ed. Pacific Grove, CA: Brooks/Cole.

Chatfield, C. 2001. *Time-Series Forecasting*. London: Chapman and Hall/CRC.

——. 2004. *The Analysis of Time Series: An Introduction*. 6th ed. Boca Raton, FL: Chapman and Hall/CRC.

Chatfield, C., and M. Yar. 1988. Holt-Winters forecasting: Some practical issues. *Statistician* 37: 129–140. https://doi.org/10.2307/2348687.

Gardner, E. S., Jr. 1985. Exponential smoothing: The state of the art. *Journal of Forecasting* 4: 1–28. https://doi.org/10.1002/for.3980040103.

Harvey, A. C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Holt, C. C. 2004. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting* 20: 5–10. https://doi.org/10.1016/j.ijforecast.2003.09.015.

Montgomery, D. C., L. A. Johnson, and J. S. Gardiner. 1990. *Forecasting and Time Series Analysis*. 2nd ed. New York: McGraw–Hill.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. *Management Science* 6: 324–342. https://doi.org/10.1287/mnsc.6.3.324.

Also see

[TS] **tsset** — Declare data to be time-series data

[TS] **tssmooth** — Smooth and forecast univariate time-series data

For suggested citations, see the FAQ on citing Stata documentation.