

# Package ‘interflex’

April 23, 2020

**Type** Package

**Title** Multiplicative Interaction Models Diagnostics and Visualization

**Version** 1.1.3

**Date** 2020-4-23

**Author** Jens Hainmueller, Jonathan Mummolo, Yiqing Xu and Ziyi Liu

**Maintainer** Yiqing Xu <[yiqingxu@stanford.edu](mailto:yiqingxu@stanford.edu)>

**Description** Performs diagnostic tests of multiplicative interaction models and plots non-linear marginal effects of a treatment on an outcome across different values of a moderator.

**URL** <http://yiqingxu.org/software/interaction/RGuide.html>

**License** MIT + file LICENSE

**Imports** Rcpp (>= 0.12.3), ggplot2 (>= 2.1.0), sandwich (>= 2.3-4),  
pcse (>= 1.9), lmtest (>= 0.9-34), Lmoments (>= 1.2-3),  
doParallel (>= 1.0.10), foreach (>= 1.4.3), mgcv (>= 1.8-16),  
lfe (>= 2.6-2291), gridExtra, grid, ggplotify, RColorBrewer,  
caret, gtable

**LinkingTo** Rcpp, RcppArmadillo

**NeedsCompilation** yes

**Depends** R (>= 3.6.0)

**LazyData** true

**Repository** CRAN

**Date/Publication** 2020-04-23 19:20:02 UTC

## R topics documented:

interflex-package . . . . .	2
inter.binning . . . . .	3
inter.gam . . . . .	8
inter.kernel . . . . .	9
inter.linear . . . . .	13
inter.raw . . . . .	17

inter.test . . . . .	19
interflex . . . . .	20
interflex-internal . . . . .	25
plot.interflex . . . . .	25
predict.interflex . . . . .	27
s1 . . . . .	30
s2 . . . . .	30
s3 . . . . .	30
s4 . . . . .	31
s5 . . . . .	31
<b>Index</b>	<b>32</b>

---

interflex-package	<i>Multiplicative Interaction Models Diagnostics and Visualization</i>
-------------------	--

---

## Description

Producing Flexible Marginal Effect Estimates with Multiplicative Interaction Models

## Details

This package performs diagnostics and visualizations of multiplicative interaction models. Besides conventional linear interaction models, it provides two additional estimation strategies—linear regression based on pre-specified bins and locally linear regressions based on Gaussian kernels—to flexibly estimate the conditional marginal effect of a treatment variable on an outcome variable across different values of a moderating variable. These approaches relax the linear interaction effect assumption and safeguard against excessive extrapolation.

## Author(s)

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

## References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.

## See Also

[interflex](#), [inter.raw](#), [inter.gam](#), [plot.interflex](#), and [predict.interflex](#)

inter.binning

*The Binning Estimator***Description**

Implementing the binning estimator, a generalization of the multiplicative interaction model and conducting various diagnostic tests

**Usage**

```
inter.binning(data, Y, D, X, treat.type = NULL, base = NULL,
  Z = NULL, FE = NULL, weights = NULL, full.moderate = TRUE,
  na.rm = FALSE, Xunif = FALSE, nbins = 3, cutoffs = NULL, CI = TRUE,
  vartype = "robust", nboots = 200, parallel = TRUE, cores = 4,
  cl = NULL, time = NULL, pairwise = TRUE, wald = TRUE, predict = FALSE,
  D.ref = NULL, figure = TRUE, order = NULL, subtitles = NULL,
  show.subtitles = NULL, Xdistr = "histogram", main = NULL,
  Ylabel = NULL, Dlabel = NULL, Xlabel = NULL, xlab = NULL,
  ylab = NULL, xlim = NULL, ylim = NULL, theme.bw = FALSE,
  show.grid = TRUE, cex.main = NULL, cex.sub = NULL, cex.lab = NULL, cex.axis = NULL,
  bin.labs = TRUE, interval = NULL, file = NULL, ncols = NULL, pool = FALSE,
  color = NULL, jitter = FALSE, legend.title = NULL)
```

**Arguments**

data	a data.frame that stores Y, D, X, and other variables.
Y	a string, name of the outcome variable.
D	a string, name of the treatment variable.
X	a string, name of the moderating variable.
treat.type	a string that specifies the type of the treatment variable. Can be one of the following: "discrete" and "continuous". If not specified, it will be automatically set based on the treatment variable. (If D is in numeric format and the number of multiple arms of D is larger than 5, treat.type will be set to "continuous". Otherwise, it will be set to "discrete".)
base	a string specifying the base group if <b>treat.type</b> is "discrete". If not specified, the base group will be selected based on numeric/character order of values in the treatment variable. For example, when treatment D can take values 0, 1, and 2, those observations with treatment 0 will be selected as the base group.
Z	a vector of variable names of control variables.
FE	a vector of variable names of fixed effects indicators.
weights	a string, name of the weighting variable.
full.moderate	a logical flag indicating whether to use fully moderated model. If it is set to TRUE, all interaction terms between moderator X and covariates Z(s) will be treated as extra covariates. The default is TRUE.

<code>na.rm</code>	a logical flag indicating whether to list-wise delete data. The algorithm will report error if missing data exist.
<code>Xunif</code>	a logical flag controlling whether to transform values of the moderator into their percentiles. The default is FALSE.
<code>nbins</code>	an integer that determines the number of bins of the moderator. The sample will be divided based on integral multiples of 100/nbins percentile. <b>nbins</b> will be automatically subtracted by 1 if more than one multiples of the percentile have the same value (for example, if the moderator has over 70% zeros, both the 33 and 66 percentiles are zero). The default is 3. Ignored when <b>cutoffs</b> is supplied.
<code>cutoffs</code>	a vector of numbers that determines how subgroups are divided based on the moderator X. When this option is supplied, <b>nbins</b> will be ignored. The smallest number of the first interval and the largest number of the last interval do not need to be specified. Numbers smaller than the minimum or larger than the maximum of X will be ignored. <b>nbins</b> equals to the length of this vector plus 1. Ignored if the treatment is dichotomous.
<code>CI</code>	a logical flag indicating whether the confidence intervals need to be shown.
<code>vartype</code>	a string that controls the variance-covariance estimator. Can be one of the following: "homoscedastic"; "robust", "cluster", "pcse" (panel corrected standard errors) and "bootstrap". The default is "robust".
<code>nboots</code>	an integer specifying the number of bootstrap runs when <b>vartype</b> is "bootstrap". The default is 200.
<code>parallel</code>	a logical flag indicating whether parallel computing will be used in bootstrapping.
<code>cores</code>	an integer indicating the number of cores to be used in parallel computing. If not specified, the algorithm will use the maximum number of logical cores of your computer (warning: this could prevent you from multi-tasking on your computer).
<code>cl</code>	a string specifying the name of clustering variable for clustered standard errors. <b>vartype</b> must be either "cluster" or "pcse". <b>vartype</b> will be automatically set to "cluster" if <b>cl</b> is supplied and <b>vartype</b> is not "pcse". "pcse" is not allowed when <b>FE</b> is specified.
<code>time</code>	a string specifying the name of time variable for panel corrected standard errors. Ignored when <b>vartype</b> is not "pcse".
<code>pairwise</code>	a logical flag indicating whether to switch on the <b>pairwise</b> option when estimating panel corrected standard errors. Ignored when <b>vartype</b> is not "pcse". The default is TRUE.
<code>wald</code>	a logical flag controlling whether to conduct a Wald test.
<code>predict</code>	a logical flag that specifies whether to estimate the expected value of Y conditional on X, D and the mean of all covariates. The default is FALSE. If it is set to TRUE, it will take a bit more time to estimate and save expected values of Y. Users can later use the command <code>predict(...)</code> to visualize the plots.
<code>D.ref</code>	a vector of numbers that specifies the reference values of D which the expected values of Y will be conditional on when <b>treat.type</b> is "continuous". If it is not specified, the 0.25, 0.5, and 0.75 quantiles of D will be chosen.

figure	a logical flag controlling whether to draw a marginal effect plot.
order	a vector of strings that determines the order of subplots when visualizing marginal effects. It should contain all kinds of treatment arms except for the base group.
subtitles	a vector of strings that determines the subtitles of subplots when <b>pool</b> is FALSE, or determines the label in the legend when <b>pool</b> is TRUE. It should have the same length as the number of categories of treatment arms minus one.
show.subtitles	a logical flag controlling whether to show subtitles.
Xdistr	a string indicating the way the distribution of the moderator will be plotted: "histogram" (or "hist"), "density", or "none". The default is "histogram".
main	a string that controls the title of the plot.
Ylabel	a string that controls the label of the outcome variable Y in the plot.
Dlabel	a string that controls the label of the treatment variable D in the plot.
Xlabel	a string that controls the label of the moderating variable X in the plot.
xlab	a string that specifies the label of the x-axis.
ylab	a string that specifies the label of the y-axis.
xlim	a two-element numeric vector that controls the range of the x-axis to be shown in the plot.
ylim	a two-element numeric vector that controls the range of the y-axis to be shown in the plot (with small adjustments to improve aesthetics).
theme.bw	a logical flag specifying whether to use a black-white theme.
show.grid	a logical flag indicating whether to show grid in the plot.
cex.main	a numeric value that controls the font size of the plot title.
cex.sub	a numeric value that controls the font size of subtitles.
cex.lab	a numeric value that controls the font size of the axis labels.
cex.axis	a numeric value that controls the font size of the axis numbers.
bin.labs	a logical flag indicating whether to show a label for each bin in a binning plot; only meaningful when <b>nbins</b> is 3 or 4. The default is TRUE.
interval	draw two vertical lines to demonstrate the interval used in replicated papers.
file	a string that specifies the filename in which the plot is saved.
ncols	an integral that controls the number of columns in the plot if <b>pool</b> is FALSE.
pool	a logical flag specifying whether to integrate marginal effects or expected values of Y for each treatment arm in one plot.
color	a vector of colors that determines the color of lines when drawing the pool plot.
jitter	a logical flag specifying whether to jitter error-bars to avoid overlapping. The default is FALSE.
legend.title	a string that specifies the title of the legend when <b>pool</b> is TRUE.

## Details

**inter.binning** implements the binning estimator. There are three steps to implement the estimator. First, we discretize the moderator  $X$  into several bins and create a dummy variable for each bin. The default is 3 bins, respectively corresponding to the three terciles. Second, we pick an evaluation point within each bin, where we want to estimate the conditional marginal effect of  $D$  on  $Y$ . Typically, we choose the median value of  $X$  in each bin. Third, we estimate a model that includes interactions between the bin dummies and the treatment indicator, the bin dummies and the moderator  $X$  minus the evaluation points, as well as the triple interactions. The last two terms are to capture the effect of  $D$  on  $Y$  within each bin.

The binning estimator has several key advantages over the standard multiplicative interaction model. First, the binning estimator is much more flexible as it jointly fits the interaction components of the standard model to each bin separately. The model does not impose the linear interaction effect (LIE) assumption. Instead, the conditional marginal effects can vary freely across the three bins and therefore can take on any non-linear or non-monotonic pattern that might describe the heterogeneity in the effect of  $D$  on  $Y$  across low, medium, or high levels of  $X$ .

Second, since the bins are constructed based on the support of  $X$ , the binning ensures that the conditional marginal effects are estimated at typical values of the moderator and do not rely on excessive extrapolation or interpolation.

Third, the binning estimator is easy to implement using any regression software and the standard errors for the conditional marginal effects are directly estimated by the regression so there are no need to compute linear combinations of coefficients to recover the conditional marginal effects.

Fourth, the binning estimator actually provides a generalization that nests the standard multiplicative interaction model as a special case. It can therefore serve as a formal test on the validity of a global LIE assumption imposed by the standard model. So in the special case when the standard multiplicative interaction model is correct and therefore the global LIE assumption holds, then—as the sample size grows—the marginal effect estimates from the binning estimator converge in probability on the unbiased marginal effect estimates from the standard multiplicative interaction model.

In the case of a discrete treatment, **inter.binning** displays at the bottom of the figure a stacked histogram that shows the distribution of the moderator  $X$ . In this histogram the total height of the stacked bars refers to the distribution of the moderator in the pooled sample and the red and gray shaded bars refer to the distribution of the moderator in the treatment and control groups, respectively. Adding such a histogram makes it easy to judge the degree to which there is common support in the data. In the case of a continuous treatment, it displays a histogram at the bottom that simply shows the distribution of  $X$  in the entire sample.

## Value

<code>est.lin</code>	stores the linear marginal effect estimates
<code>vcov.matrix</code>	stores the variance-covariance matrix of the series of point estimates.
<code>est.bin</code>	stores the binning estimates
<code>tests</code>	stores the results of various statistical tests. See below
<code>tests\$bin.size</code>	the percentage of observations within each bin
<code>tests\$X.LKurtosis</code>	the L-kurtosis measure of the moderator

tests\$correctOrder	whether the three binning estimates are in the correct order (i.e. monotonically increasing or decreasing) when there are 3 bins.
tests\$p.twosided	p values of pairwise t-tests of the binning estimates when there are 2 or 3 bins.
tests\$p.wald	p-value of a Wald test. The NULL hypothesis is that the linear interaction model and the binning model are statistically equivalent.
graph	stores the graphic output, if <b>pool</b> is TRUE, it is a ggplot object, otherwise it is a Grid object.
est.predict	stores the expected value of Y conditional on X, D and the mean of all covariates if <b>predict = TRUE</b> .

**Author(s)**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

**References**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.

**See Also**

[interflex](#)

**Examples**

```
library(interflex)
data(interflex)
inter.binning(Y = "Y", D = "D", X = "X", Z = "Z1",
  data = s1, nbins = 3, vartype = "homoscedastic",
  Ylabel = "Y", Dlabel = "Tr", Xlabel="X")

## specifying cutoffs
inter.binning(Y = "Y", D = "D", X = "X", Z = "Z1",
  data = s2, cutoffs = c(1,2,4,5))

## with fixed effects
s4[, "wgt"] <- 1
inter.binning(Y = "Y", D = "D", X = "X", Z = "Z1",
  weights = "wgt", FE = c("group", "year"), data = s4,
  cl = "group", vartype = "cluster")
```

---

`inter.gam`*Visualization of a Generalized Additive Model (GAM)*

---

**Description**

Estimating and plotting a GAM before estimating a multiplicative interaction model

**Usage**

```
inter.gam(data, Y, D, X, Z = NULL, weights = NULL, full.moderate = TRUE,  
          FE = NULL, SE = FALSE, k = 10, angle = c(30, 100, -30, -120),  
          Ylabel = NULL, Dlabel = NULL, Xlabel = NULL)
```

**Arguments**

<code>data</code>	a data.frame that stores Y, D, X, and other variables.
<code>Y</code>	a string, name of the outcome variable.
<code>D</code>	a string, name of the treatment variable.
<code>X</code>	a string, name of the moderating variable.
<code>Z</code>	a vector of variable names of control variables.
<code>full.moderate</code>	a logical flag indicating whether to use fully moderated model.
<code>weights</code>	a string, name of the weighting variable.
<code>FE</code>	a vector of variable names of fixed effects indicators.
<code>SE</code>	a logical variable indicating whether to show the surfaces one standard error above and below the predicted surface.
<code>k</code>	an integer indicating the dimension of the basis used to represent the smooth term. Smaller values of k produce smoother surfaces.
<code>angle</code>	a vector of numbers indicating viewing angles (allows 1 to 4 values).
<code>Ylabel</code>	a string that controls the label of the outcome variable Y in the plot.
<code>Dlabel</code>	a string that controls the label of the treatment variable D in the plot.
<code>Xlabel</code>	a string that controls the label of the moderating variable X in the plot.

**Details**

**inter.gam** provides a convenient way to visualize the relationships between the outcome, treatment, and moderator using GAM when both the treatment and moderator are continuous (based on [vis.gam](#) of the **mgecv** package).

**Author(s)**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu



## References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at SSRN: <http://bit.ly/HMX2019>.

## See Also

[interflex](#), [plot.interflex](#), and [predict.interflex](#)

## Examples

```
library(interflex)
data(interflex)
inter.gam(Y="Y", D="D", X="X", Z=c("Z1"), data=s2)
```

---

inter.kernel

*The Kernel Estimator*

---

## Description

Implementing the kernel estimator, based on a semi-parametric model that relaxes the linear interaction effect assumption of conventional multiplicative interaction models and safeguards against excessive extrapolation

## Usage

```
inter.kernel(data, Y, D, X, treat.type = NULL, base = NULL,
  Z = NULL, weights = NULL, FE = NULL, full.moderate = TRUE,
  na.rm = FALSE, Xunif = FALSE, CI = TRUE, conf.level = 0.95, cl = NULL,
  CV.method = NULL, kfold = 10, grid = 30, neval = 50, nboots = 200, parallel = TRUE,
  cores = 4, seed = 02139, bw = NULL, bw.adaptive = TRUE, quantile.eval = FALSE,
  metric = "MSPE", predict = FALSE, D.ref = NULL, figure = TRUE, order = NULL,
  subtitles = NULL, show.subtitles = NULL,Xdistr = "histogram", main = NULL,
  Ylabel = NULL, Dlabel = NULL, Xlabel = NULL, xlab = NULL, ylab = NULL,
  xlim = NULL, ylim = NULL, theme.bw = FALSE, interval = NULL, show.grid = TRUE,
  cex.main = NULL, cex.sub = NULL, cex.lab = NULL, cex.axis = NULL,
  file = NULL, ncols = NULL, pool = FALSE, color = NULL, legend.title = NULL,
  diff.values = NULL, percentile = FALSE)
```

## Arguments

data	a data.frame that stores Y, D, and X.
Y	a string, name of the outcome variable.
D	a string, name of the treatment variable.
X	a string, name of the moderating variable.

treat.type	a string that specifies the type of the treatment variable. Can be one of the following: "discrete" and "continuous". If not specified, it will be automatically set based on the treatment variable. (If D is in numeric format and the number of multiple arms of D is larger than 5, treat.type will be set to "continuous". Otherwise, it will be set to "discrete".)
base	a string specifying the base group if <b>treat.type</b> is "discrete". If not specified, the base group will be selected based on numeric/character order of values in the treatment variable. For example, when treatment D can take values 0, 1, and 2, those observations with treatment 0 will be selected as the base group.
Z	a vector of variable names of control variables.
weights	a string, name of the weighting variable.
FE	a vector of variable names of fixed effects indicators.
full.moderate	a logical flag indicating whether to use fully moderated model. If it is set to TRUE, all interaction terms between moderator X and covariates Z(s) will be treated as extra covariates. The default is TRUE.
na.rm	a logical flag indicating whether to list-wise delete missing data. The algorithm will report an error if missing data exist.
Xunif	a logical flag controlling whether to transform values of the moderator into their percentiles.
CI	a logical flag indicating whether the confidence intervals need to be shown. If <b>CI = FALSE</b> , the bootstrap procedure will not be conducted.
conf.level	a number that specifies the confidence interval. For example, 0.9 means a confidence interval of [5%, 95%]. The default is 0.95, i.e., [2.5%, 97.5%].
cl	a string specifying the name of clustering variable. If specified, the bootstrap procedure will be blocking on this variable. Ignored when <b>CI = FALSE</b> .
CV.method	a string specified the method of sampling in cross-validation. Can be one of the following: "simple", "cluster", and "stratify". The default is "simple".
kfold	an integer specified the number of folds in cross-validation, default to 10. If it is larger than the number of observations in the dataset, it will conduct a leave-one-out CV.
grid	either an integer indicating the number of candidates in a grid search to find the optimal bandwidth, or a vector of positive numbers of candidate bandwidths. Ignored if <b>bw</b> is specified.
neval	an integer specifying the number of evaluation points. A large number would produce smoother marginal effects estimates but the algorithm will take more time to run. The default is 50.
nboots	an integer specifying the number of bootstrap runs. The default is 200.
parallel	a logical flag indicating whether parallel computing will be used in bootstrapping and/or cross-validation. Ignored if <b>CI = FALSE</b> and <b>bw</b> is specified.
cores	an integer indicating the number of cores to be used in parallel computing. If not specified, the algorithm will use the maximum number of logical cores of your computer (warning: this could prevent you from multi-tasking on your computer).

seed	an integer that sets the seed in random number generation. Ignored if <b>CI</b> = FALSE and <b>bw</b> is specified.
bw	a number that sets the bandwidth in Gaussian kernel reweighting.
bw.adaptive	a logical flag indicating whether to use adaptive bandwidth. If it is set to TRUE, a small bandwidth will be assigned to the observations in the high-density region and vice versa, <b>bw</b> will indicate the bandwidth applied to the region of X with highest density of observations.
quantile.eval	a logical flag indicating whether to choose evaluation points based on quantiles rather than uniformly. If it is set to TRUE, denser evaluation points will be assigned to the region with high-density observations and vice versa. The default is FALSE.
metric	either "MSPE" (Mean Squared Prediction Error) or "MAPE" (Mean Absolute Prediction Error), criterion to be used in cross-validation to select bandwidth. The default is "MSPE".
predict	a logical flag that specifies whether to estimate the expected value of Y conditional on X, D and the mean of all covariates. The default is FALSE. If it is set to TRUE, it will take a bit more time to estimate and save expected values of Y. Users can later use the command predict(...) to visualize the plots.
D.ref	a vector of numbers that specifies the reference values of D which the expected values of Y will be conditional on when <b>treat.type</b> is "continuous". If it is not specified, the 0.25, 0.5, and 0.75 quantiles of D will be chosen.
figure	a logical flag controlling whether to draw a marginal effect plot.
order	a vector of strings that determines the order of treatment arms in the plot when visualizing marginal effects. It should contain all kinds of treatment arms except for the base group.
subtitles	a vector of strings that determines the subtitles of subplots when <b>pool</b> is FALSE, or determines the label in the legend when <b>pool</b> is TRUE. It should have the same length as the number of categories of treatment arms minus one.
show.subtitles	a logical flag controlling whether to show subtitles.
Xdistr	a string indicating the way the distribution of the moderator will be plotted: "histogram" (or "hist"), "density", or "none". The default is "histogram".
main	a string that controls the title of the plot.
Ylabel	a string that controls the label of the outcome variable Y in the plot.
Dlabel	a string that controls the label of the treatment variable D in the plot.
Xlabel	a string that controls the label of the moderating variable X in the plot.
xlab	a string that specifies the label of the x-axis.
ylab	a string that specifies the label of the y-axis.
xlim	a two-element numeric vector that controls the range of the x-axis to be shown in the plot.
ylim	a two-element numeric vector that controls the range of the y-axis to be shown in the plot (with small adjustments to improve aesthetics).
theme.bw	a logical flag specifying whether to use a black-white theme.

<code>interval</code>	draw two vertical lines to demonstrate the interval used in replicated papers.
<code>show.grid</code>	a logical flag indicating whether to show grid in the plot.
<code>cex.main</code>	a numeric value that controls the font size of the plot title.
<code>cex.sub</code>	a numeric value that controls the font size of subtitles.
<code>cex.lab</code>	a numeric value that controls the font size of the axis labels.
<code>cex.axis</code>	a numeric value that controls the font size of the axis numbers.
<code>file</code>	a string that specifies the filename in which the plot is saved.
<code>ncols</code>	an integer that controls the number of columns in the plot if <b>pool</b> is FALSE.
<code>pool</code>	a logical flag specifying whether to integrate marginal effects or expected values of Y for each treatment arm in one plot.
<code>color</code>	a vector of colors that determines the color of lines when drawing the pool plot.
<code>legend.title</code>	a string that specifies the title of the legend when <b>pool</b> is TRUE.
<code>diff.values</code>	a numeric vector contain 2 or 3 elements which are within the range of the moderator. The treatment effects at corresponding values of the moderator will be compared.
<code>percentile</code>	a logical flag indicating whether to take values of the moderator on a percentile scale.

## Details

**inter.kernel** implements a kernel smoothing estimator of the marginal effect, which is an application of semi-parametric smooth varying-coefficient models. This approach is more complicated than the simple binning estimator, but provides a generalization that allows researchers to flexibly estimate the functional form of the marginal effect of D on Y across the values of X by estimating a series of local effects with a Gaussian kernel re-weighting scheme.

The kernel regression nests the standard linear interaction model as a special case. However, in the kernel regression the conditional effect of D on Y does not have to fall on a linear line as required by the linear interaction effect (LIE) assumption, but can vary freely across the range of X. In addition, if covariates Z are included in the model, the coefficients of those covariates are also allowed to vary freely across the range of X resulting in a very flexible estimator that also helps to guard against misspecification bias with respect to the covariates.

**inter.kernel** produces a smooth marginal effect of D on Y with respect to X. It estimates a series of locally linear regression using kernel re-weighting based on the distance between each value of X (the moderator) and each evaluation point. If the bandwidth is not supplied, it is selected using a standard n-fold cross-validation procedure. Standard errors and confidence intervals can be computed using a bootstrap. It also adds a histogram (or density plot) at the bottom of the figure to help users assess the common support assumption based on the distribution of the moderator.

**inter.kernel** can accommodate linear fixed-effect models with multiple fixed-effect indicators. A fixed-effect model is estimated using a fast algorithm written in C++ at each evaluation point of the semi-parametric kernel estimation.

**inter.kernel** can estimate the expected value of Y conditional on the moderator X and treatment D when the option "predict" is set to TRUE.

**Value**

<code>bw</code>	specified or cross-validated bandwidth.
<code>est</code>	marginal effects estimates (and confidence intervals if <b>CI = TRUE</b> )
<code>vcov.matrix</code>	stores the variance-covariance matrix of the series of point estimates if <b>CI = TRUE</b> .
<code>CV.out</code>	results from cross-validation if <b>bw</b> is not specified.
<code>graph</code>	stores the graphic output, a ggplot object.
<code>est.predict</code>	stores the expected value of Y conditional on X, D and the mean of all covariates if <b>predict = TRUE</b> .
<code>t.test.diffs</code>	stores the difference between treatment effects at corresponding values specified in <b>diff.values</b> .

**Author(s)**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

**References**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.

**See Also**

[interflex](#)

**Examples**

```
library(interflex)
data(interflex)
inter.kernel(Y = "Y", D = "D", X = "X", Z = "Z1",
  data = s3, CI = FALSE, bw = 1.3, parallel = FALSE)
```

---

inter.linear

*The Linear Estimator*

---

**Description**

Implementing the linear interaction estimator

**Usage**

```
inter.linear(data, Y, D, X, treat.type = NULL, base = NULL,
  Z = NULL, FE = NULL, weights = NULL, full.moderate = TRUE,
  na.rm = FALSE, Xunif = FALSE, CI = TRUE, vartype = "robust",
  nboots= 200, parallel = TRUE, cores =4, cl = NULL, time = NULL,
  pairwise = TRUE, predict = FALSE, D.ref = NULL, figure = TRUE,
  order = NULL, subtitles = NULL, show.subtitles = NULL,Xdistr = "histogram",
  main = NULL, Ylabel = NULL, Dlabel = NULL, Xlabel = NULL,
  xlab = NULL, ylab = NULL, xlim = NULL, ylim = NULL, theme.bw = FALSE,
  show.grid = TRUE, cex.main = NULL, cex.sub = NULL, cex.lab = NULL,
  cex.axis = NULL, interval = NULL, file = NULL, ncols = NULL,
  pool = FALSE, color = NULL, legend.title = NULL, diff.values = NULL,
  percentile = FALSE)
```

**Arguments**

data	a data.frame that stores Y, D, X, and other variables.
Y	a string, name of the outcome variable.
D	a string, name of the treatment variable.
X	a string, name of the moderating variable.
treat.type	a string that specifies the type of the treatment variable. Can be one of the following: "discrete" and "continuous". If not specified, it will be automatically set based on the treatment variable. (If D is in numeric format and the number of multiple arms of D is larger than 5, treat.type will be set to "continuous". Otherwise, it will be set to "discrete".)
base	a string specifying the base group if <b>treat.type</b> is "discrete". If not specified, the base group will be selected based on numeric/character order of values in the treatment variable. For example, when treatment D can take values 0, 1, and 2, those observations with treatment 0 will be selected as the base group.
Z	a vector of variable names of control variables.
FE	a vector of variable names of fixed effects indicators.
weights	a string, name of the weighting variable.
full.moderate	a logical flag indicating whether to use fully moderated model. If it is set to TRUE, all interaction terms between moderator X and covariates Z(s) will be treated as extra covariates. The default is TRUE.
na.rm	a logical flag indicating whether to list-wise delete data. The algorithm will report error if missing data exist.
Xunif	a logical flag controlling whether to transform values of the moderator into their rank order. The default is FALSE.
CI	a logical flag indicating whether the confidence intervals need to be shown.
vartype	a string that controls the variance-covariance estimator. Can be one of the following: "homoscedastic"; "robust", "cluster", "pcse" (panel corrected standard errors) and "bootstrap". The default is "robust".

nboots	an integer specifying the number of bootstrap runs when <b>vartype</b> is "bootstrap". The default is 200.
parallel	a logical flag indicating whether parallel computing will be used in bootstrapping.
cores	an integer indicating the number of cores to be used in parallel computing. If not specified, the algorithm will use the maximum number of logical cores of your computer (warning: this could prevent you from multi-tasking on your computer).
cl	a string specifying the name of clustering variable for clustered standard errors. <b>vartype</b> must be either "cluster" or "pcse". <b>vartype</b> will be automatically set to "cluster" if <b>cl</b> is supplied and <b>vartype</b> is not "pcse". "pcse" is not allowed when <b>FE</b> is specified.
time	a string specifying the name of time variable for panel corrected standard errors. Ignored when <b>vartype</b> is not "pcse".
pairwise	a logical flag indicating whether to switch on the <b>pairwise</b> option when estimating panel corrected standard errors. Ignored when <b>vartype</b> is not "pcse". The default is TRUE.
predict	a logical flag that specifies whether to estimate the expected value of Y conditional on X, D and the mean of all covariates. The default is FALSE. If it is set to TRUE, it will take a bit more time to estimate and save expected values of Y. Users can later use the command plot(predict=TRUE,...) to visualize the plots.
D.ref	a vector of numbers that specifies the reference values of D which the expected values of Y will be conditional on when <b>treat.type</b> is "continuous". If it is not specified, the 0.25, 0.5, and 0.75 quantiles of D will be chosen.
figure	a logical flag controlling whether to draw a marginal effect plot.
order	a vector of strings that determines the order of subplots when visualizing marginal effects. It should contain all kinds of treatment arms except for the base group.
subtitles	a vector of strings that determines the subtitles of subplots when <b>pool</b> is FALSE, or determines the label in the legend when <b>pool</b> is TRUE. It should have the same length as the number of categories of treatment arms minus one.
show.subtitles	a logical flag controlling whether to show subtitles.
Xdistr	a string indicating the way the distribution of the moderator will be plotted: "histogram" (or "hist"), "density", or "none". The default is "histogram".
main	a string that controls the title of the plot.
Ylabel	a string that controls the label of the outcome variable Y in the plot.
Dlabel	a string that controls the label of the treatment variable D in the plot.
Xlabel	a string that controls the label of the moderating variable X in the plot.
xlab	a string that specifies the label of the x-axis.
ylab	a string that specifies the label of the y-axis.
xlim	a two-element numeric vector that controls the range of the x-axis to be shown in the plot.
ylim	a two-element numeric vector that controls the range of the y-axis to be shown in the plot (with small adjustments to improve aesthetics).

<code>theme.bw</code>	a logical flag specifying whether to use a black-white theme.
<code>show.grid</code>	a logical flag indicating whether to show grid in the plot.
<code>cex.main</code>	a numeric value that controls the font size of the plot title.
<code>cex.sub</code>	a numeric value that controls the font size of subtitles.
<code>cex.lab</code>	a numeric value that controls the font size of the axis labels.
<code>cex.axis</code>	a numeric value that controls the font size of the axis numbers.
<code>interval</code>	draw two vertical lines to demonstrate the interval used in replicated papers.
<code>file</code>	a string that specifies the filename in which the plot is saved.
<code>ncols</code>	an integer that controls the number of columns in the plot if <b>pool</b> is FALSE.
<code>pool</code>	a logical flag specifying whether to integrate marginal effects or expected values of Y for each treatment arm in one plot.
<code>color</code>	a vector of colors that determines the color of lines when drawing the pool plot.
<code>legend.title</code>	a string that specifies the title of the legend when <b>pool</b> is TRUE.
<code>diff.values</code>	a numeric vector contain 2 or 3 elements which are within the range of the moderator. The treatment effects at corresponding values of the moderator will be compared.
<code>percentile</code>	a logical flag indicating whether to take values of the moderator on a percentile scale.

## Details

**inter.linear** implements the classic linear estimator.

## Value

<code>est.lin</code>	stores the linear marginal effect estimates.
<code>vcov.matrix</code>	stores the variance-covariance matrix of the series of point estimates.
<code>tests</code>	stores the results of various statistical tests. See below
<code>tests\$X.LKurtosis</code>	the L-kurtosis measure of the moderator
<code>graph</code>	stores the graphic output, if <b>pool</b> is TRUE, it is a ggplot object, otherwise it is a Grid object.
<code>est.predict</code>	stores the expected value of Y conditional on X, D and the mean of all covariates if <b>predict = TRUE</b> .
<code>t.test.diffs</code>	stores the difference between treatment effects at corresponding values specified in <b>diff.values</b> .

## Author(s)

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

## References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.



**See Also**[interflex](#)**Examples**

```
library(interflex)
data(interflex)
inter.linear(Y = "Y", D = "D", X = "X", Z = "Z1",
  data = s1, vartype = "homoscedastic",
  Ylabel = "Y", Dlabel = "Tr", Xlabel="X")

inter.linear(Y = "Y", D = "D", X = "X", Z = c("Z1","Z2"),
  data = s5, base = 'A')
```

inter.raw

*Plotting Raw Data***Description**

Plotting raw data before estimating a multiplicative interaction model

**Usage**

```
inter.raw(data, Y, D, X, treat.type=NULL, subtitles=NULL,
  order=NULL, Z=NULL, FE=NULL, weights = NULL, full.moderate = TRUE,
  nbins = 3, cutoffs = NULL, span = NULL, pos = NULL,
  main = NULL, Ylabel = NULL, Dlabel = NULL,
  Xlabel = NULL, theme.bw = FALSE, show.grid = TRUE, cex.main = NULL,
  cex.lab = NULL, cex.axis = NULL, ncols = NULL)
```

**Arguments**

data	a data.frame that stores Y, D, and X.
Y	a string, name of the outcome variable.
D	a string, name of the treatment variable.
X	a string, name of the moderating variable.
treat.type	a string that specifies the type of the treatment variable. Can be one of the following: "discrete" and "continuous". If not specified, it will be automatically set based on the treatment variable. (If D is in numeric format and the number of multiple arms of D is larger than 5, treat.type will be set to "continuous". Otherwise, it will be set to "discrete".)
subtitles	a vector of strings that determines the subtitles of subplots when <b>treat.type</b> is "discrete". It should have the same length as the number of categories of treatment arms.

order	a vector of strings that determines the order of treatment arms in the plot when <b>treat.type</b> is "discrete". It should contain all kinds of treatment arms.
Z	a vector of variable names of control variables.
FE	a vector of variable names of fixed effects indicators.
weights	a string, name of the weighting variable. It will influence both the linear and lowess fits.
full.moderate	a logical flag indicating whether to use fully moderated model.
nbins	an integer that determines the number of bins of the moderator (and hence the number of subplots) when the treatment is continuous. The sample will be divided based on integral multiples of 100/nbins percentile. <b>nbins</b> will be automatically subtracted by 1 if more than one multiples of the percentile have the same value (for example, if the moderator has over 70% zeros, both the 33 and 66 percentiles are zero). The default is 3. Ignored when <b>cutoffs</b> is supplied.
cutoffs	a vector of numbers that determines how subgroups are divided based on the moderator X. When this option is supplied, <b>nbins</b> will be ignored. The smallest number of the first interval and the largest number of the last interval do not need to be specified. Numbers smaller than the minimum or larger than the maximum of X will be ignored. <b>nbins</b> equals the length of this vector plus 1. Ignored if the treatment is dichotomous.
span	hyper-parameter for the lowess fit. Large numbers produce smoother lines (from <b>ggplot2</b> ).
pos	the interval of Y to be shown.
main	a string that controls the title of the plot.
Ylabel	a string that controls the label of the outcome variable Y in the plot.
Dlabel	a string that controls the label of the treatment variable D in the plot.
Xlabel	a string that controls the label of the moderating variable X in the plot.
theme.bw	a logical flag specifying whether to use a black-white theme.
show.grid	a logical flag indicating whether to show grid in the plot.
cex.main	a numeric value that controls the font size of the plot title.
cex.lab	a numeric value that controls the font size of the axis labels.
cex.axis	a numeric value that controls the font size of the axis numbers.
ncols	a integral value that controls the number of columns when drawing marginal effect plot, default to the number of categories of treatment.

## Details

**inter.raw** provides a simple visual diagnostic, a scatter-plot, to help researchers to detect potential problems with the linear interaction effect assumption and the lack of common support.

If the treatment D is binary, it plots outcome Y against the moderator X separately for the sample of treatment group observations (D=1) and the sample of control group observations (D=0). In each sample, it superimposes a linear regression line as well as lowess fit in each group.

To check whether there is sufficient common support in the data, we can simply compare the distribution of X in both groups and examine the range of X values for which there are a sufficient

number of data points in both groups. The box plots near the center of the figures display quantiles of the moderator at each level of the treatment. The dot in the center denotes the median, the end points of the thick bars denote the 25th and 75th percentiles, and the end points of the thin bars denote the 5th and 95th percentiles.

If the treatment and moderator are continuous, we adopt a simple binning approach. Specifically, **inter.raw** will split the sample into several subgroups based on the moderator. By default, it will split the sample into three roughly equal sized subgroups: low X (first tercile), medium X (second tercile), and high X (third tercile). For each of the subgroups, it plots Y against D while again adding the estimates for both the linear fit and the lowess fit.

If FE or Z are specified, the same plots will be constructed after residualizing with respect to FE and Z.

### Value

graph stores the graphic output, a **ggplot2** object.

### Author(s)

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer), Ziyi Liu

### References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at SSRN: <http://bit.ly/HMX2019>.

### See Also

[inter.gam](#), [plot.interflex](#), and [predict.interflex](#)

### Examples

```
library(interflex)
data(interflex)
inter.raw(Y = "Y", D = "D", X = "X", data = s1,
  Ylabel = "Outcome", Dlabel = "Treatment", Xlabel="Moderator")
```

---

inter.test

*Conduct t-test of Treatment Effects*

---

### Description

Compare treatment effects at specific values of the moderator

### Usage

```
inter.test(x, diff.values, percentile=FALSE, k=15)
```

**Arguments**

x	an <b>interflex</b> object.
diff.values	a numeric vector contain 2 or 3 elements which are within the range of the moderator. The treatment effects at corresponding values of the moderator will be compared.
percentile	a logical flag indicating whether to take values of the moderator on a percentile scale.
k	an integer specifying the dimension(s) of the bases used to represent the smooth term, default to 15.

**Details**

**inter.test** compare treatment effects at specific values of the moderator using marginal effects and vcov matrix derived from linear/kernel estimation. Based on GAM model(relies on **mgcv** package), users can approximate the treatment effect and its variance using smooth functions without re-estimating the model, hence saving time.

**Author(s)**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

**References**

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.

**See Also**

[interflex](#) and [predict.interflex](#)

**Examples**

```
library(interflex)
data(interflex)
out <- inter.linear(data = s1, Y = "Y", D = "D", X = "X", Z = "Z1")
inter.test(out,diff.values=c(0.25,0.5,0.75),percentile=TRUE)
```

---

interflex

---

*Flexible Interactive Models*


---

**Description**

Producing flexible marginal effect estimates with linear, binning, and kernel estimators

**Usage**

```
interflex(data,Y, D, X, estimator, treat.type = NULL,
  base = NULL, Z = NULL, weights = NULL,
  FE = NULL,cl = NULL, full.moderate = TRUE,
  na.rm = FALSE, Xunif = FALSE,
  CI = TRUE, conf.level = 0.95, nbins = 3, cutoffs = NULL,
  vartype = "robust", time = NULL,
  pairwise = TRUE, wald = TRUE, bin.labs = TRUE,
  CV.method = NULL, kfold = 10, grid = 30,
  neval = 50, nboots = 200, parallel = TRUE,
  cores = 4, seed = 02139, bw = NULL, bw.adaptive = TRUE,
  quantile.eval = FALSE, metric = "MSPE", figure = TRUE,
  order = NULL, subtitles = NULL,
  show.subtitles = NULL, Xdistr = "histogram",
  main = NULL, Ylabel = NULL, Dlabel = NULL, Xlabel = NULL,
  xlab = NULL,ylab = NULL, xlim = NULL, ylim = NULL,
  theme.bw = FALSE, interval = NULL, show.grid = TRUE,
  cex.main = NULL, cex.sub = NULL, cex.lab = NULL, cex.axis = NULL,
  file = NULL, ncols = NULL, pool = FALSE, color = NULL,
  legend.title = NULL, jitter = FALSE,
  predict = FALSE, D.ref = NULL, diff.values = NULL,
  percentile = FALSE)
```

**Arguments**

<code>data</code>	a data.frame that stores Y, D, and X.
<code>Y</code>	a string, name of the outcome variable.
<code>D</code>	a string, name of the treatment variable.
<code>X</code>	a string, name of the moderating variable.
<code>estimator</code>	a string that specifies the estimator. Can be one of the following: "linear", "binning" and "kernel".
<code>treat.type</code>	a string that specifies the type of the treatment variable. Can be one of the following: "discrete" and "continuous". If not specified, it will be automatically set based on the treatment variable. (If D is in numeric format and the number of multiple arms of D is larger than 5, treat.type will be set to "continuous". Otherwise, it will be set to "discrete".)
<code>base</code>	a string specifying the base group if <b>treat.type</b> is "discrete". If not specified, the base group will be selected based on numeric/character order of values in the treatment variable. For example, when treatment D can take values 0, 1, and 2, those observations with treatment 0 will be selected as the base group.
<code>Z</code>	a vector of variable names of control variables.
<code>weights</code>	a string, name of the weighting variable.
<code>FE</code>	a vector of variable names of fixed effects indicators.
<code>cl</code>	a string specifying the name of clustering variable. If specified, the bootstrap and cross-validation procedure will be blocking on this variable. Ignored when

	<b>CI = FALSE</b> . When <b>vartype</b> is "pcse", it specifies the name of unit variable for panel corrected standard errors.
full.moderate	a logical flag indicating whether to use fully moderated model. If it is set to TRUE, all interaction terms between moderator X and covariates Z(s) will be treated as extra covariates. The default is TRUE.
na.rm	a logical flag indicating whether to list-wise delete missing data. The algorithm will report an error if missing data exist.
Xunif	a logical flag controlling whether to transform values of the moderator into their rank order.
CI	a logical flag indicating whether the confidence intervals need to be shown. If <b>CI = FALSE</b> , the bootstrap procedure will not be conducted.
conf.level	a number that specifies the confidence interval. For example, 0.9 means a confidence interval of [5%, 95%]. The default is 0.95, i.e., [2.5%, 97.5%].
nbins	an integer that determines the number of bins of the moderator. The sample will be divided based on integral multiples of 100/nbins percentile. <b>nbins</b> will be automatically subtracted by 1 if more than one multiples of the percentile have the same value (for example, if the moderator has over 70% zeros, both the 33 and 66 percentiles are zero). The default is 3. Ignored when <b>cutoffs</b> is supplied. When nbins are set to 1, it only conducts a linear fit.
cutoffs	a vector of numbers that determines how subgroups are divided based on the moderator X. When this option is supplied, <b>nbins</b> will be ignored. The smallest number of the first interval and the largest number of the last interval do not need to be specified. Numbers smaller than the minimum or larger than the maximum of X will be ignored. <b>nbins</b> equals to the length of this vector plus 1. Ignored if the treatment is dichotomous.
vartype	a string that controls the variance-covariance estimator. Can be one of the following: "homoscedastic"; "robust", "cluster", "pcse" (panel corrected standard errors) and "bootstrap". The default is "robust".
time	a string specifying the name of time variable for panel corrected standard errors. Ignored when <b>vartype</b> is not "pcse".
pairwise	a logical flag indicating whether to switch on the <b>pairwise</b> option when estimating panel corrected standard errors. Ignored when <b>vartype</b> is not "pcse". The default is TRUE.
wald	a logical flag controlling whether to conduct a Wald test.
bin.labs	a logical flag indicating whether to show a label for each bin in a binning plot; only meaningful when <b>nbins</b> is 3 or 4. The default is TRUE.
CV.method	a string specified the method of sampling in cross-validation. Can be one of the following: "simple", "cluster", and "stratify".The default is "simple".
kfold	an integer specified the number of folds in cross-validation, default to 10. If it is larger than the number of observations in the dataset, it will conduct a leave-one-out CV.
grid	either an integer indicating the number of candidates in a grid search to find the optimal bandwidth, or a vector of positive numbers of candidate bandwidths. Ignored if <b>bw</b> is specified.

neval	an integer specifying the number of evaluation points. A large number would produce smoother marginal effects estimates but the algorithm will take more time to run. The default is 50.
nboots	an integer specifying the number of bootstrap runs. The default is 200.
parallel	a logical flag indicating whether parallel computing will be used in bootstrapping and/or cross-validation. Ignored if <b>CI</b> = FALSE and <b>bw</b> is specified.
cores	an integer indicating the number of cores to be used in parallel computing. If not specified, the algorithm will use the maximum number of logical cores of your computer (warning: this could prevent you from multi-tasking on your computer).
seed	an integer that sets the seed in random number generation. Ignored if <b>CI</b> = FALSE and <b>bw</b> is specified.
bw	a number that sets the bandwidth in Gaussian kernel reweighting.
bw.adaptive	a logical flag indicating whether to use adaptive bandwidth. If it is set to TRUE, a small bandwidth will be assigned to the observations in the high-density region and vice versa, <b>bw</b> will indicate the bandwidth applied to the region of X with highest density of observations.
quantile.eval	a logical flag indicating whether to choose evaluation points based on quantiles rather than uniformly. If it is set to TRUE, denser evaluation points will be assigned to the region with high-density observations and vice versa. The default is FALSE.
metric	either "MSPE" (Mean Squared Prediction Error) or "MAPE" (Mean Absolute Prediction Error), criterion to be used in cross-validation to select bandwidth. The default is "MSPE".
figure	a logical flag controlling whether to draw a marginal effect plot.
order	a vector of strings that determines the order of treatment arms in the plot when visualizing marginal effects. It should contain all kinds of treatment arms except for the base group.
subtitles	a vector of strings that determines the subtitles of subplots when <b>pool</b> is FALSE, or determines the label in the legend when <b>pool</b> is TRUE. It should have the same length as the number of categories of treatment arms minus one.
show.subtitles	a logical flag controlling whether to show subtitles.
Xdistr	a string indicating the way the distribution of the moderator will be plotted: "histogram" (or "hist"), "density", or "none". The default is "histogram".
main	a string that controls the title of the plot.
Ylabel	a string that controls the label of the outcome variable Y in the plot.
Dlabel	a string that controls the label of the treatment variable D in the plot.
Xlabel	a string that controls the label of the moderating variable X in the plot.
xlab	a string that specifies the label of the x-axis.
ylab	a string that specifies the label of the y-axis.
xlim	a two-element numeric vector that controls the range of the x-axis to be shown in the plot.

<code>ylim</code>	a two-element numeric vector that controls the range of the y-axis to be shown in the plot (with small adjustments to improve aesthetics).
<code>theme.bw</code>	a logical flag specifying whether to use a black-white theme.
<code>interval</code>	draw two vertical lines to demonstrate the interval used in replicated papers.
<code>show.grid</code>	a logical flag indicating whether to show grid in the plot.
<code>cex.main</code>	a numeric value that controls the font size of the plot title.
<code>cex.sub</code>	a numeric value that controls the font size of subtitles.
<code>cex.lab</code>	a numeric value that controls the font size of the axis labels.
<code>cex.axis</code>	a numeric value that controls the font size of the axis numbers.
<code>file</code>	a string that specifies the filename in which the plot is saved.
<code>ncols</code>	an integer that controls the number of columns in the plot if <b>pool</b> is FALSE.
<code>pool</code>	a logical flag specifying whether to draw marginal effects or expected values of Y for each treatment arm in a single plot. The default is FALSE.
<code>color</code>	a vector of colors that determines the color of lines when drawing the pool plot.
<code>legend.title</code>	a string that specifies the title of the legend when <b>pool</b> is TRUE.
<code>jitter</code>	a logical flag specifying whether to jitter error-bars to avoid overlapping. The default is FALSE.
<code>predict</code>	a logical flag that specifies whether to estimate the expected value of Y conditional on X, D and the mean of all covariates. The default is FALSE. If it is set to TRUE, it will take a bit more time to estimate and save expected values of Y. Users can later use the command <code>plot(predict=TRUE,...)</code> to visualize the plots.
<code>D.ref</code>	a vector of numbers that specifies the reference values of D which the expected values of Y will be conditional on when <b>treat.type</b> is "continuous". If it is not specified, the 0.25, 0.5, and 0.75 quantiles of D will be chosen.
<code>diff.values</code>	a numeric vector contain 2 or 3 elements which within the range of the moderator. The treatment effects at corresponding values of the moderator will be compared.
<code>percentile</code>	a logical flag indicating whether to take values of the moderator on a percentile scale.

## Details

**interflex** is an umbrella function that allow users to call `inter.kernel` and `inter.binning` more conveniently.

## Author(s)

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

## References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.



**See Also**

[inter.raw](#), [inter.gam](#), [plot.interflex](#), and [predict.interflex](#)

**Examples**

```
library(interflex)
data(interflex)
s5.binning <- interflex(estimator = 'binning', Y = "Y", D = "D", X = "X",
Z = c("Z1", "Z2"), data = s5, CI = FALSE)
```

---

interflex-internal      *Internal Interflex Functions*

---

**Description**

Internal Interflex functions

**Details**

These are not to be called by the user, or in some cases are just waiting for proper documentation to be written :).

---

plot.interflex      *Plotting Marginal Effect Estimates*

---

**Description**

Plotting marginal effect estimates after either the binning estimator or the kernel estimator is applied

**Usage**

```
## S3 method for class 'interflex'
plot(x, to.show = NULL, order = NULL, subtitles = NULL,
      show.subtitles = NULL, CI = NULL, diff.values = NULL,Xdistr = "histogram",
      main = NULL, Ylabel = NULL, Dlabel = NULL, Xlabel = NULL, xlab = NULL,
      ylab = NULL, xlim = NULL, ylim = NULL, theme.bw = FALSE, show.grid = TRUE,
      cex.main = NULL, cex.sub = NULL, cex.lab = NULL, cex.axis = NULL,
      bin.labs = TRUE, interval = NULL, file = NULL, ncols = NULL, pool = FALSE,
      legend.title = NULL, color = NULL, jitter = FALSE, ...)
```

**Arguments**

<code>x</code>	an <b>interflex</b> object.
<code>to.show</code>	a vector of strings that determines which treatment arms will be displayed when visualizing marginal effects. It should be a subset of all treatment arms except for the base group.
<code>order</code>	a vector of strings that determines the order of treatment arms in the plot when visualizing marginal effects. It should contain all kinds of treatment arms except for the base group or all elements in "to.show".
<code>subtitles</code>	a vector of strings that determines the subtitles of subplots. Its length should be equal to the number of categories of treatment arms minus one.
<code>show.subtitles</code>	a logical flag controlling whether to show subtitles. Default to TRUE when the number of treatment arms is larger than 2.
<code>CI</code>	a logical flag indicating whether the confidence intervals need to be shown.
<code>diff.values</code>	a two-element numeric vector. The treatment effects at corresponding values of the moderator will be marked on the plot.
<code>Xdistr</code>	a string indicating the way the distribution of the moderator will be plotted: "histogram" (or "hist"), "density", or "none". The default is "histogram".
<code>main</code>	a string that controls the title of the plot.
<code>Ylabel</code>	a string that controls the label of the outcome variable Y in the plot.
<code>Dlabel</code>	a string that controls the label of the treatment variable D in the plot.
<code>Xlabel</code>	a string that controls the label of the moderating variable X in the plot.
<code>xlab</code>	a string that specifies the label of the x-axis.
<code>ylab</code>	a string that specifies the label of the y-axis.
<code>xlim</code>	a two-element numeric vector that controls the range of the x-axis to be shown in the plot.
<code>ylim</code>	a two-element numeric vector that controls the range of the y-axis to be shown in the plot (with small adjustments to improve aesthetics).
<code>theme.bw</code>	a logical flag specifying whether to use a black-white theme.
<code>show.grid</code>	a logical flag indicating whether to show grid in the plot.
<code>cex.main</code>	a numeric value that controls the font size of the plot title.
<code>cex.sub</code>	a numeric value that controls the font size of subtitles.
<code>cex.lab</code>	a numeric value that controls the font size of the axis labels.
<code>cex.axis</code>	a numeric value that controls the font size of the axis numbers.
<code>bin.labs</code>	a logical flag indicating whether to show a label for each bin in a binning plot; only meaningful when <b>nbins</b> is 3 or 4. The default is TRUE.
<code>interval</code>	draw two vertical lines to demonstrate the interval used in replicated papers.
<code>file</code>	a string that specifies the filename in which the plot is saved.
<code>ncols</code>	an integer that controls the number of columns in the plot if <b>pool</b> is FALSE.
<code>pool</code>	a logical flag specifying whether to integrate marginal effects for each treatment arm in one plot.

legend.title	a string that specifies the title of the legend when <b>pool</b> is TRUE.
color	a vector of colors that determines the color of lines when <b>pool</b> is TRUE.
jitter	a logical flag specifying whether to jitter error-bars to avoid overlapping. The default is FALSE.
...	other argv.

## Details

**plot.interflex** draws a margining effect plot based on a **interflex** object after either **inter.binning** or **inter.kernel** is applied. It allows users to flexibly change the look of a plot without re-estimating the model, hence saving time.

## Value

graph stores the graphic output, a **ggplot2** object.

## Author(s)

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

## References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.

## See Also

[interflex](#) and [predict.interflex](#)

## Examples

```
library(interflex)
data(interflex)
out <- inter.binning(data = s1, Y = "Y", D = "D", X = "X", Z = "Z1", CI = FALSE)
plot(out)
```

---

predict.interflex      *Plotting Marginal Effect Estimates*

---

## Description

Plotting expected outcomes given fixed values of the treatment and moderator after either the linear, binning or the kernel estimator is applied.

**Usage**

```
## S3 method for class 'interflex'
predict(object, to.show = NULL, order = NULL, subtitles = NULL,
  show.subtitles = NULL, Xdistr = "histogram", CI = NULL, pool = FALSE, main = NULL,
  Ylabel = NULL, Xlabel = NULL, xlab = NULL, ylab = NULL, xlim = NULL, ylim = NULL,
  theme.bw = FALSE, show.grid = TRUE, cex.main = NULL, cex.sub = NULL,
  cex.lab = NULL, cex.axis = NULL, color = NULL, file = NULL, interval = NULL,
  legend.title = NULL, ncols = NULL, ...)
```

**Arguments**

object	an <b>interflex</b> object ( <b>predict==TRUE</b> ).
to.show	a vector of strings that determines which treatment arms will be displayed when visualizing expected values. It should be a subset of all treatment arms including the base group.
order	a vector of strings that determines the order of treatment arms in the plot when visualizing expected values. It should contain all treatment arms or all elements in "to.show".
subtitles	a vector of strings that determines the subtitles of subplots when <b>pool</b> is FALSE, or determines the label in the legend when <b>pool</b> is TRUE.
show.subtitles	a logical flag controlling whether to show subtitles. Default to TRUE when the number of treatment arms is larger than 2.
Xdistr	a string indicating the way the distribution of the moderator will be plotted: "histogram" (or "hist"), "density", or "none". The default is "histogram".
CI	a logical flag indicating whether the confidence intervals need to be shown.
pool	a logical flag specifying whether to integrate expected values of Y for each treatment arm in one plot.
main	a string that controls the title of the plot.
Ylabel	a string that controls the label of the outcome variable Y in the plot.
Xlabel	a string that controls the label of the moderating variable X in the plot.
xlab	a string that specifies the label of the x-axis.
ylab	a string that specifies the label of the y-axis.
xlim	a two-element numeric vector that controls the range of the x-axis to be shown in the plot.
ylim	a two-element numeric vector that controls the range of the y-axis to be shown in the plot (with small adjustments to improve aesthetics).
theme.bw	a logical flag specifying whether to use a black-white theme.
show.grid	a logical flag indicating whether to show grid in the plot.
cex.main	a numeric value that controls the font size of the plot title.
cex.sub	a numeric value that controls the font size of subtitles.
cex.lab	a numeric value that controls the font size of the axis labels.
cex.axis	a numeric value that controls the font size of the axis numbers.

<code>color</code>	a vector of colors that determines the color of lines when <b>pool</b> is TRUE.
<code>file</code>	a string that specifies the filename in which the plot is saved.
<code>interval</code>	draw two vertical lines to demonstrate the interval used in replicated papers.
<code>legend.title</code>	a string that specifies the title of the legend when <b>pool</b> is TRUE.
<code>ncols</code>	na integral that controls the number of columns in the plot if <b>pool</b> is FALSE.
<code>...</code>	other argv.

## Details

**predict.interflex** visualize expected outcomes given fixed values of the treatment and moderator after either the linear, binning or the kernel estimator is applied. It allows users to flexibly change the look of a plot without re-estimating the model, hence saving time.

## Value

`graph` stores the graphic output, a **ggplot2** object.

## Author(s)

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu (Maintainer); Ziyi Liu

## References

Jens Hainmueller; Jonathan Mummolo; Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis*, Vol. 27, Iss. 2, April 2019, pp. 163–192. Available at: <http://bit.ly/HMX2019>.

## See Also

[interflex](#) and [plot.interflex](#)

## Examples

```
library(interflex)
data(interflex)
out <- inter.binning(data = s1, Y = "Y", D = "D", X = "X", Z = "Z1", predict = TRUE)
predict(out)
```

---

s1	<i>s1</i>
----	-----------

---

**Description**

Simulated dataset 1: dichotomous treatment with linear marginal effects.

**Usage**

s1

**Format**

dataframe

---

s2	<i>s2</i>
----	-----------

---

**Description**

Simulated dataset 2: continuous treatment with linear marginal effects.

**Usage**

s2

**Format**

dataframe

---

s3	<i>s3</i>
----	-----------

---

**Description**

Simulated dataset 3: dichotomous treatment with nonlinear marginal effects.

**Usage**

s3

**Format**

dataframe

---

s4

s4

---

**Description**

Simulated dataset 4: dichotomous treatment with fixed effects.

**Usage**

s4

**Format**

dataframe

---

s5

s5

---

**Description**

Simulated dataset 5: discrete treatments with nonlinear marginal effects.

**Usage**

s5

**Format**

dataframe

# Index

## \*Topic **datasets**

- s1, [30](#)
- s2, [30](#)
- s3, [30](#)
- s4, [31](#)
- s5, [31](#)

## \*Topic **graphics**

- inter.binning, [3](#)
- inter.gam, [8](#)
- inter.kernel, [9](#)
- inter.linear, [13](#)
- inter.raw, [17](#)
- inter.test, [19](#)
- plot.interflex, [25](#)
- predict.interflex, [27](#)

## \*Topic **ts**

- interflex-internal, [25](#)

- coefs.new (interflex-internal), [25](#)
- crossvalidate.new (interflex-internal),  
[25](#)

- fastplm (interflex-internal), [25](#)
- fastplm\_predict (interflex-internal), [25](#)

- inter.binning, [3](#)
- inter.gam, [2](#), [8](#), [19](#), [25](#)
- inter.kernel, [9](#)
- inter.linear, [13](#)
- inter.plot.pool (interflex-internal), [25](#)
- inter.raw, [2](#), [17](#), [25](#)
- inter.test, [19](#)
- interflex, [2](#), [7](#), [9](#), [13](#), [17](#), [20](#), [20](#), [27](#), [29](#)
- interflex-internal, [25](#)
- interflex-package, [2](#)
- interflex.default (interflex-internal),  
[25](#)
- interflex\_fastplm (interflex-internal),  
[25](#)

- interflex\_fastplm\_predict  
(interflex-internal), [25](#)

- plot.interflex, [2](#), [9](#), [19](#), [25](#), [25](#), [29](#)
- predict.interflex, [2](#), [9](#), [19](#), [20](#), [25](#), [27](#), [27](#)

- s1, [30](#)
- s2, [30](#)
- s3, [30](#)
- s4, [31](#)
- s5, [31](#)

- vcovCluster (interflex-internal), [25](#)
- vis.gam, [8](#)