

Predictor Effects Graphics Gallery

John Fox and Sanford Weisberg

October 29, 2018

Contents

1	Introduction	2
1.1	Effects and Predictor Effect Plots	2
1.2	General Outline for Constructing Predictor Effect Plots	7
1.3	How <code>predictorEffect</code> Chooses Conditioning Predictors	7
1.4	The <code>Effect()</code> Function	8
1.5	The <code>predictorEffects</code> Function	9
2	Optional Arguments for the <code>predictorEffect()</code> and <code>Effect()</code> Functions	10
2.1	Options for Predictors in the Conditioning Group: <code>xlevels</code>	10
2.2	Predictors in the Fixed Group	12
2.2.1	Factor Predictors	12
2.2.2	Numeric Predictors	12
2.3	Standard Errors and Confidence Intervals	12
2.4	Residuals	13
3	Arguments for Plotting Predictor Effects and Effects	13
3.1	<code>axes</code> Group: Specify Axis Characteristics	14
3.1.1	<code>x</code> : Horizontal Axis Specification	14
3.1.2	<code>y</code> : Specify the Vertical Axis, Linear Models	16
3.1.3	<code>y</code> : Specify the Vertical Axis, Generalized Linear Models	19
3.2	<code>lines</code> Group: Specifying Plotted Lines	22
3.2.1	<code>multiline</code> and <code>z.var</code> : Multiple Lines in a Plot	22
3.2.2	Line Color, Type, Width, Smoothness	27
3.3	<code>confint</code> : Specifying Confidence Interval Style and Inclusion	27
3.4	<code>lattice</code> : Specifying Standard lattice Package Arguments	29
3.4.1	Modifying the Key With <code>key.args</code>	29
3.4.2	<code>layout</code>	30
3.4.3	<code>array</code> of Multiple Plots	31
3.4.4	<code>strip</code>	31
3.5	<code>symbols</code> : Plotting symbols	32
4	Displaying Residuals in Predictor Effect and Effect Plots	33
5	Polytomous Categorical Responses	36
6	The Lattice Theme for the effects Package	41

Abstract

Predictor effect displays visualize the response surface of complex regression models by averaging and conditioning, producing a sequence of 2D line graphs, one graph or set of graphs for each predictor in the regression problem (Fox and Weisberg, 2019a,b). In this vignette, we give examples of effect plots produced by the **effects** package, and in the process systematically illustrate the optional arguments to functions in the package.

1 Introduction

Predictor effect plots (Fox and Weisberg, 2019b) provide graphical summaries for fitted regression models with a linear predictor, including linear models, generalized linear models, linear and generalized linear mixed models, and many others. These graphs provide an alternative to tables of fitted coefficients that can be much harder to interpret than effect plots. Predictor effect plots are implemented in R in the **effects** package, documented in Fox and Weisberg (2019a). This vignette provides many examples of variations on these graphical displays that can be obtained with the **effects** package. Many of the details, and more complete descriptions of the data sets used as examples, are provided in the references cited at the end of the vignette.

1.1 Effects and Predictor Effect Plots

We begin with an example of a multiple linear regression, using a data set in the **carData** package:

```
R> library("car") # also loads the carData package
R> Prestige$type <- factor(Prestige$type, levels=c("bc", "wc", "prof"))
R> lm1 <- lm(prestige ~ education + poly(women, 2) +
+           log(income)*type, data=Prestige)
```

The **Prestige** data set contains data, collected circa 1970, on 102 Canadian occupations. The model **lm1** is a linear model with response **prestige**, continuous predictors **income**, **education**, and **women**, and the factor predictor **type**, with three levels. Before fitting the model, we reordered the levels of the factor **type** as "bc" (blue-collar), "wc" (white-collar), "prof" (professional and managerial). The predictor **education** represents itself in the linear model, and so it is both a predictor and a regressor, as defined in Fox and Weisberg (2019a, Sec. 4.1). The predictor **income** is represented by the regressor $\log(\text{income})$. The variable **women**, a percentage between 0 and 100, is represented by regressors that define a polynomial of degree two using **poly()**'s default orthogonal polynomials. The variable **type** is a factor with three levels, so it is represented by two dummy variables defined by the default contrast-generating function in R, **contr.treatment()**. Finally the formula includes an interaction between **income** and **type** defined by multiplying each of the regressors that represent **income** by each of the regressors that represent **type**.

The usual numeric summary of the fit of **lm1** is a table of estimated coefficients, which we obtain via the **S()** function in the **car** package:¹

```
R> S(lm1)
```

```
Call: lm(formula = prestige ~ education + poly(women, 2) + log(income) * type,
data = Prestige)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-137.500	23.522	-5.85	8.2e-08
education	2.959	0.582	5.09	2.0e-06

¹Applied, to a regression model, **S()** is similar to, but somewhat more flexible than, the standard R **summary()** function.

poly(women, 2)1	28.339	10.190	2.78	0.0066
poly(women, 2)2	12.566	7.095	1.77	0.0800
log(income)	17.514	2.916	6.01	4.1e-08
typewc	0.969	39.495	0.02	0.9805
typeprof	74.276	30.736	2.42	0.0177
log(income):typewc	-0.466	4.620	-0.10	0.9199
log(income):typeprof	-7.698	3.451	-2.23	0.0282

Residual standard deviation: 6.2 on 89 degrees of freedom
(4 observations deleted due to missingness)

Multiple R-squared: 0.879

F-statistic: 81.1 on 8 and 89 DF, p-value: <2e-16

AIC	BIC
646.26	672.11

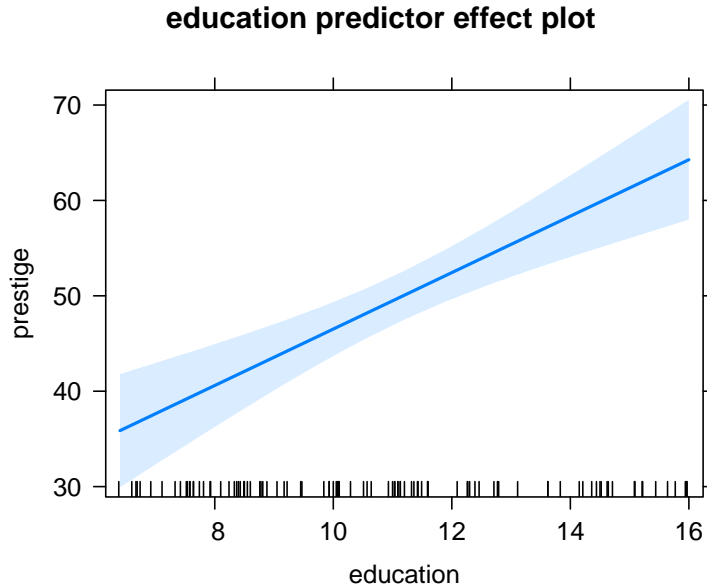
- Interpretation of this output is straightforward only for the predictor **education**, where an increase of **education** of one year holding other predictors fixed corresponds to an estimated expected increase in the response of 2.959 units.
- Even ignoring the interaction, the logarithm complicates the interpretation of the effect of **income**.²
- The predictor **women** is represented by two regressors, so the effect of **women** requires examining two coefficient estimates that are interpretable only by those knowledgeable about polynomial regression analysis, and even if raw polynomials were used, via `poly(women, 2, raw=TRUE)` in place of `poly(women, 2)`, interpretation of the effect of **women** is complicated.
- Understanding the coefficients for the main effect of **type** depends on the contrasts used to define the effect. The contrasts can be changed by the user, and the default contrasts in R are different from the default contrasts used by SAS or other programs, so the coefficients cannot be reliably interpreted without information not present in the summary shown.
- Finally, the interaction further complicates the interpretation of the effect of either **income** or **type** because the interaction coefficients need to be interpreted along with the main effect coefficients.

As a consequence, summarization of the effects of predictors using tables of coefficient estimates is often incomplete. Effects, and particularly plots of effects, can in many instances reveal the relationship of the response to the predictors more clearly. This conclusion is especially true for models with linear predictors that include interactions and multiple-coefficient terms such as regression splines and polynomials, as illustrated later in this vignette.

A predictor effect plot summarizes the role of a selected *focal* predictor in a fitted regression model. The `predictorEffect()` function is used to compute the appropriate summary of the regression, and then the `plot()` function may be used to graph the resulting object, as in the following example:

```
R> library("effects")
R> e1.lm1 <- predictorEffect("education", lm1)
R> plot(e1.lm1)
```

²An alternative here is to use logs-base-2 rather than natural logs, in which case the coefficient of log-income, again ignoring the interaction, would represent the effect of doubling **income** holding other predictors constant.



This graph visualizes the partial slope for `education`, that for each year increase in `education`, the fitted `prestige` increases by 2.959 points, when the other predictors are held fixed. The intercept of the line (which doesn't appear in the graph, which doesn't extend to `education = 0`) is determined by the choices for averaging over the fixed predictors, but for any choice of averaging method, the slope of the line would be the same. The shaded area is a pointwise confidence band for the fitted values, based on standard errors computed from the covariance matrix of the fitted regression coefficients. The rug-plot at the bottom of the graph shows the locations of the `education` values.

The information that is needed to draw the plot is computed by the `predictorEffect()` function. The minimal arguments for `predictorEffect()` are the quoted name of a predictor in the model followed by the fitted model object. The essential purpose of this function is to compute fitted values from the model with `education` varying and all other predictors fixed at some typical value (Fox and Weisberg, 2019a, Sec. 4.3). The command below displays the values of the regressors for which fitted values are computed, including a column of ones for the intercept:

```
R> e1.lm1$model.matrix
```

```
(Intercept) education poly(women, 2)1 poly(women, 2)2 log(income) typewc
1          1         6.4    0.000020998    -0.13496    8.8449 0.23469
2          1         8.8    0.000020998    -0.13496    8.8449 0.23469
3          1        11.0    0.000020998    -0.13496    8.8449 0.23469
4          1        14.0    0.000020998    -0.13496    8.8449 0.23469
5          1        16.0    0.000020998    -0.13496    8.8449 0.23469
typeprof log(income):typewc log(income):typeprof
1 0.31633          2.0758          2.7979
2 0.31633          2.0758          2.7979
3 0.31633          2.0758          2.7979
4 0.31633          2.0758          2.7979
5 0.31633          2.0758          2.7979
attr("assign")
[1] 0 1 2 2 3 4 4 5 5
attr("contrasts")
attr("contrasts")$type
[1] "contr.treatment"
```

Only five fitted values are computed by default, each for a different value of the focal predictor `education`. The remaining regressors have the same value for each fitted value. The value for

$\log(\text{income})$ is the logarithm of the sample mean `income`, the values for the regressors for `women` are computed at the average of `women` in the data, and the fixed values for the regressors for `type` effectively take a weighted average of the fitted values at the three levels of `type`, with weights proportional to the number of cases in each level of the factor.³ Differences in the fitted values are due to `education` alone because all the other predictors, and their corresponding regressors, are fixed. Thus the output gives the partial effect of `education` with all other predictors fixed.⁴

The computed fitted values can be viewed by printing the "eff" object returned by `predictorEffect()`, by summarizing the object, or by converting it to a data frame:

```
R> as.data.frame(e1.lm1)
```

	education	fit	se	lower	upper
1	6.4	35.864	2.9865	29.930	41.798
2	8.8	42.965	1.8275	39.334	46.596
3	11.0	49.474	1.2842	46.923	52.026
4	14.0	58.351	2.1500	54.079	62.623
5	16.0	64.268	3.1604	57.989	70.548

The values in the column `education` are the values the focal predictor. The remaining columns are the fitted values, their standard errors, and lower and upper end points of 95% confidence intervals for the fitted values. The *predictor effect plot* is simply a graph of the fitted values on the vertical axis versus the focal predictor on the horizontal axis. For a continuous focal predictor such as `education`, a line—in this case, a straight line—is drawn connecting the fitted values.

We turn next to the predictor effect plot for `income`. According to the regression model, the effect of `income` may depend on `type` due to the interaction between the two predictors, so simply averaging over `type` would be misleading. Rather, we should allow both `income` and `type` to vary, fixing the other predictors at their means or other typical values:

```
R> e2.lm1 <- predictorEffect("income", lm1)
R> as.data.frame(e2.lm1)
```

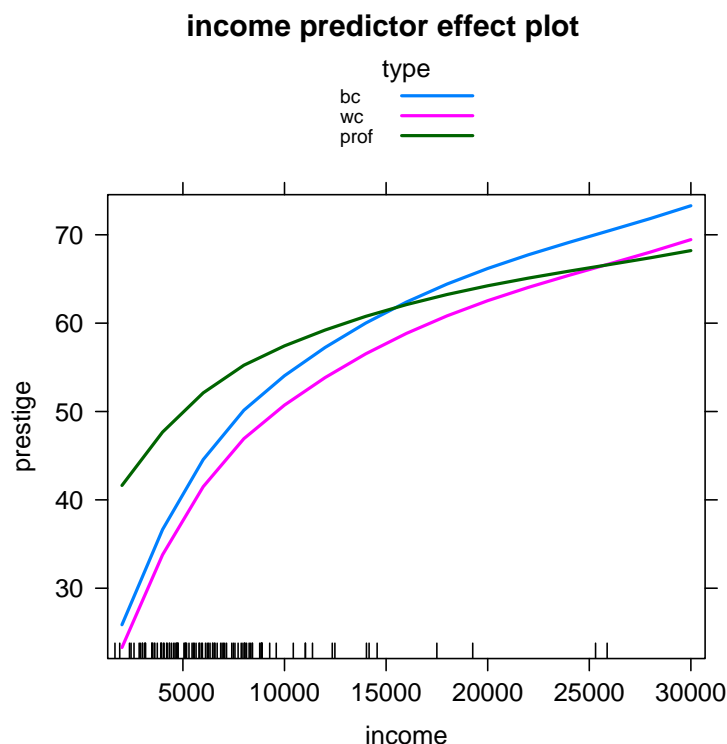
	income	type	fit	se	lower	upper
1	2000	bc	25.863	3.3037	19.299	32.428
2	8000	bc	50.142	2.3737	45.426	54.859
3	10000	bc	54.050	2.7996	48.487	59.613
4	20000	bc	66.190	4.4814	57.285	75.094
5	30000	bc	73.291	5.5708	62.222	84.360
6	2000	wc	23.290	4.5674	14.214	32.365
7	8000	wc	46.922	2.3106	42.331	51.513
8	10000	wc	50.726	3.0575	44.651	56.802
9	20000	wc	62.543	5.7716	51.075	74.011
10	30000	wc	69.455	7.4432	54.665	84.244
11	2000	prof	41.630	4.4812	32.726	50.534
12	8000	prof	55.237	2.3316	50.605	59.870
13	10000	prof	57.428	2.4552	52.549	62.306
14	20000	prof	64.231	3.6170	57.045	71.418
15	30000	prof	68.211	4.5680	59.135	77.288

The fitted values are now evaluated at 15 points, five selected values of `income` for each of the three levels of `type`. The predictor effect plot is

```
R> plot(e2.lm1, lines=list(multiline=TRUE))
```

³Although the treatment of the factor `type` sounds complicated, weighting averaging is achieved simply by computing the mean of each dummy regressor, which then corresponds to the proportion of cases coded 1 for that dummy regressor. This approach works more generally for other kinds of contrasts.

⁴Such effects are often called *marginal effects*. We prefer the term *partial effects* because it implies holding other predictors constant.

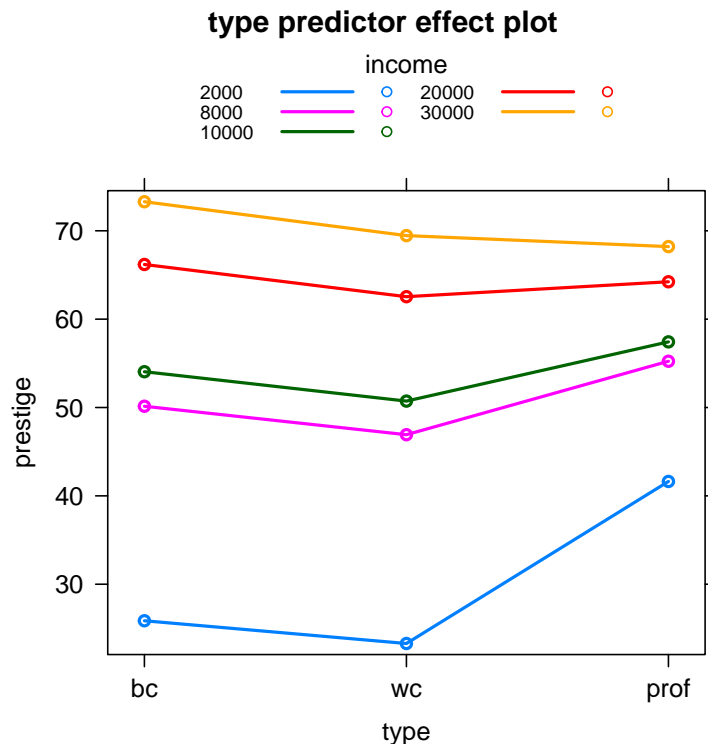


The focal predictor `income` is displayed on the horizontal axis. There is a separate line shown for the fitted values at each level of `type`. The lines are curved rather than straight because `income` appears in the model in log-scale but is displayed in the predictor effect plot in arithmetic (i.e., dollar) scale. The lines in the graph are not parallel because of the interaction between $\log(\text{income})$ and `type`. For `type = "prof"`, the fitted values of `prestige` are relatively high for lower values of `income`, and are relatively unaffected by increasing values of `income`.⁵

The predictor effect plot for `type` uses the same fitted values as the plot for `income`, but we now get five lines, one for each of the values of `income` selected by the `predictorEffect()` function:

```
R> plot(predictorEffect("type", lm1), lines=list(multiline=TRUE))
```

⁵The predictors `type` and `income` are related, however, with for example "prof" having relatively high incomes. As a consequence regions of the graph are essentially devoid of data and the fitted regression lines in these regions represent extrapolation beyond the data. We can effectively visualize this characteristic of the data by adding partial residuals to the predictor effect plot, a subject that we address in Section 4. We invite the reader to do this for the predictor `income` in model `e2.lm1`.



Here we both the `predictorEffect()` and `plot()` functions in the same command. Because the horizontal axis is now a factor, the fitted values are displayed explicitly as points, and the lines that join the points are merely a visual aid (representing what are sometimes called profiles of fitted values). Fitted `prestige` increases with `income` for all levels of `type`, but, as we found before, when `type = "prof"`, the fitted `prestige` is relatively high for lower `income`.

These initial examples use only default arguments for `predictorEffect()` and `plot()`, apart from the `multiline` argument to `plot()` to put all the fitted lines in the same graph. We explain how to customize predictor effect plots in subsequent sections of this vignette.

1.2 General Outline for Constructing Predictor Effect Plots

Using the `effects` package to draw plots usually follows this outline:

1. Fit a regression model with a linear predictor. The package supports models created by `lm()`, `glm()`, `lmer()` and `glmer()` in the `lme4` package, `lme()` in the `nlme` package, and many other regression-modeling functions (see `?Effect`).
2. The regression model created in the first step is then used as input to either `predictorEffect()`, to get the effects for one predictor, or `predictorEffects`, to get effects for one or more predictors. These functions do the averaging needed to get fitted values that will ultimately be plotted. There are many arguments for customizing the computation of the effects. The two predictor effect functions call the more basic `Effect()` function, and almost all of the material in this vignette applies to `Effect()` as well (see Section 1.4).
3. Use the generic `plot()` function to draw the graph or graphs based on the object created in Step 2.

1.3 How predictorEffect Chooses Conditioning Predictors

Suppose you select one focal predictor for which you want to draw a predictor effect plot. The `predictorEffect()` function divides the predictors in a formula into three groups:

1. The focal predictor.
2. The *conditioning group*, consisting of all predictors with at least one interaction in common with the focal predictor.
3. The *fixed group*, consisting of all other predictors, that is, those with no interactions in common with the focal predictor.

The predictors in the fixed group are all evaluated at a typical value, usually a mean, effectively averaging out the influence of these predictors on the fitted value. Fitted values are computed for all combinations of levels of the predictors in the conditioning group, with each continuous predictor in the conditioning group replaced by a few discrete values spanning the range of the predictor, for example, replacing years of `education` by a discrete variable with the values 8, 12, and 16 years.

Suppose that we fit a model with R formula

$$y \sim x1 + x2 + x3 + x4 + x2:x3 + x2:x4 \quad (1)$$

or, equivalently,

$$y \sim x1 + x2*x3 + x2*x4$$

There are four predictor effect plots for this model, one for each predictor selected in turn as the focal predictor:

Focal Predictor	Conditioning Group	Fixed Group
x1	none	x2, x3, x4
x2	x3, x4	x1
x3	x2	x1, x4
x4	x2	x1 x3

The predictor `x1` does not interact with any of the other predictors, so its conditioning set is empty and all the remaining predictors are averaged over. `x2` interacts with both `x3` and `x4`, `x3` interacts only with `x2` and `x4` interacts with `x2`.

1.4 The `Effect()` Function

Prior to late 2018 the primary function in **effects** for computing and displaying effects was the `Effect()` function.⁶ Whereas the `predictorEffect()` function automatically determines the conditioning group and the fixed group of predictors, the `Effect()` function puts that burden on the user. Each call to `predictorEffect()` is equivalent to a specific call the the `Effect()` function as follows. Suppose the `m` is the fitted model produced by the formula in 1; then

```
predictorEffect("x1", m) is equivalent to Effect("x1", m);
predictorEffect("x2", m) is equivalent to Effect(c("x2", "x3", "x4"), m);
predictorEffect("x3", m) is equivalent to Effect(c("x3", "x2"), m); and
predictorEffect("x4", m) is equivalent to Effect(c("x4", "x2"), m).
```

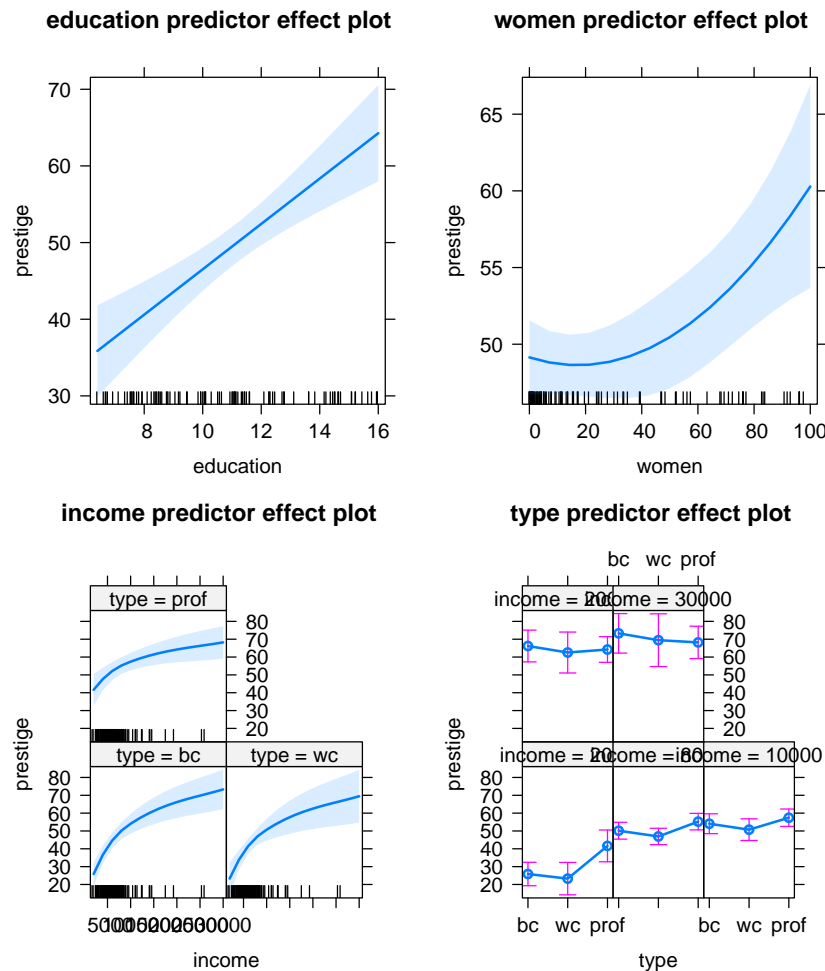
The `predictorEffect()` function determines the correct call to `Effect()` based on the choice of focal predictor and on the structure of main effects and interactions in the linear predictor for the model. It then uses the `Effect` function to do the computing. As a result most of the arguments to `predictorEffect()` are documented on the help page `help("Effect")` rather than the help page for `predictorEffect()`.

⁶The **effects** package also includes the older `allEffects()` function, which computes effects for each high-order term in a model with a linear predictor. As we explain in Fox and Weisberg (2019b), we prefer predictor effects to high-order term effects, and so, although its use is similar to `predictorEffects()`, we won't describe `allEffects()` in this vignette. There is also an older `effect()` function (with a lowercase "e"), which is a less flexible version of `Effect()`, which calls `Effect()` to perform computations, and which is retained simply for backwards compatability.

1.5 The predictorEffects Function

This function, whose name ends with an “s”—that is, the plural version of `predictorEffect()`—computes the values needed for one or more predictor effect plots, and by default for *all* of the predictors in the model. For example, the following command produces all of the predictor effect plots for the model we fit to the `Prestige` data:

```
R> eall.lm1 <- predictorEffects(lm1)
R> plot(eall.lm1)
```



The predictor effect plots for this model are displayed in an array of graphs. The plots for `income` and `type` have a separate panel for each level of the conditioning variable because the default argument `lines=list(multiline=FALSE)` was used. Confidence bounds are shown by default when `multiline=FALSE`.

The resulting object `eall.lm1` is a list with four elements, where `eall.lm1[[1]]` is the summary needed for the first predictor effect plot, `eall.lm1[[2]]` for the second plot, and so on. The following equivalent commands each draw an array of predictor effect plots:

```
R> plot(eall.lm1)
R> plot(predictorEffects(lm1))
R> plot(predictorEffects(lm1, ~ income + education + women + type))
```

If you want only the predictor effect plots for `type` and `education`, in that order, you could enter

```
R> plot(predictorEffects(lm1, ~ type + education))
```

Similarly, the commands

```
R> plot(predictorEffects(lm1, ~ women))
R> plot(predictorEffects(lm1)[[3]])
R> plot(predictorEffect("women", lm1))
```

all produce the same graph, the predictor effect plot for **women**.

Predictor effect plots in an array can be a useful shortcut for drawing many graphs quickly, but can lead to problems with the displayed graphs. For example, the horizontal axis labels for the plot for **income** are overprinted, and the labels at the top of the plots for **type** with conditioning variable **income** are larger than the available space. These problems can often be fixed using optional arguments described later in this vignette.

2 Optional Arguments for the `predictorEffect()` and `Effect()` Functions

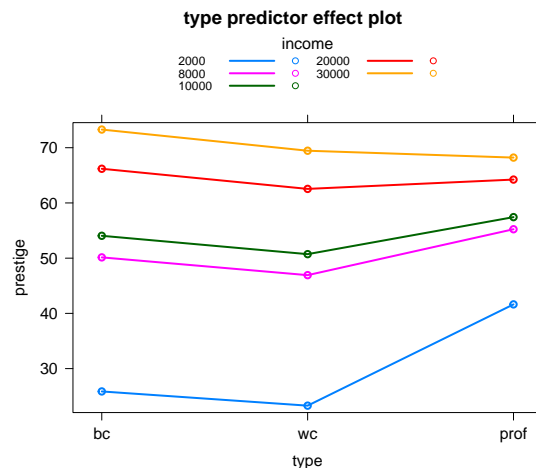
This section comprises a catalog of the arguments available to modify the behavior of the `predictorEffect()` and `Effect()` functions. The `help("Effect")` help page is somewhat more comprehensive, for example, explaining exceptions for use with `"svyglm"` objects or when plotting residuals.

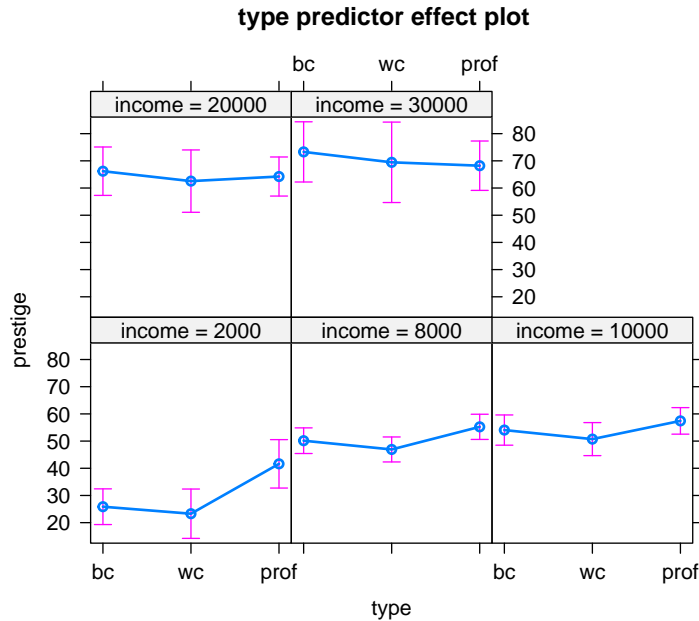
2.1 Options for Predictors in the Conditioning Group: `xlevels`

Numeric predictors in the conditioning group need to be discretized to draw a predictor effect plot. For example the predictor effect plot for **type** consists of a separate line, or a separate panel, for each discrete value of **income**:

```
R> e3.lm1 <- predictorEffect("type", lm1)
R> plot(e3.lm1, lines=list(multiline=TRUE))

R> plot(e3.lm1, lines=list(multiline=FALSE)) # the default
```

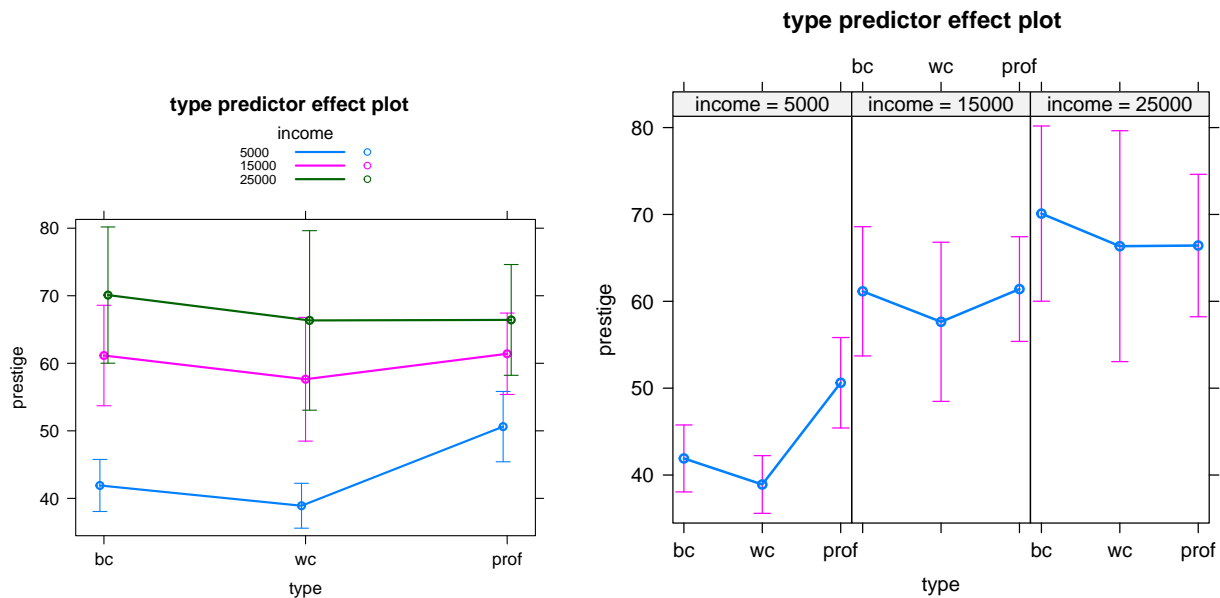




The numeric conditioning predictor `income` is evaluated by default at five equally spaced values `income`, and then rounded to “nice” numbers. In this example using three values of 5000, 15000, 25000 might produce a graph that’s easier to decode:

```
R> e3.lm1 <- predictorEffect("type", lm1,
+                             xlevels=list(income=c(5000, 15000, 25000)))
R> plot(e3.lm1, lines=list(multiline=TRUE),
+       confint=list(style="bars"))

R> plot(e3.lm1,
+       lines=list(multiline=FALSE),
+       lattice=list(layout=c(3, 1)))
```



The argument `xlevels` is a list of sub-arguments that control how numeric predictors are discretized when used in the conditioning group. For example, `xlevels=list(x1=c(2, 4, 7), x2=6)` would

use the values 2, 4 and 7 for the levels of a predictor named `x1`, use six equally spaced values for the predictor `x2`, and use the default of five values for any other numeric predictors. Numeric predictors in the fixed group are not affected by the `xlevels` argument. We use the `layout` sub-argument of the `lattice` argument to arrange the panels of the second graph in 3 columns and 1 row (see Section 3.4.2).

The number of points at which the focal predictor is evaluated is also controlled by the `xlevels` argument. Increasing the number of points may make plotted lines smoother or more accurate. See `help("plot.eff")` for information on the `quantiles` argument that provides an alternative method of setting `xlevels` when partial residuals are displayed (as discussed in Section 4).

2.2 Predictors in the Fixed Group

2.2.1 Factor Predictors

Predictors in the fixed group are replaced by fixed “typical” values. Fitted values are then computed using these fixed values for the fixed group, varying the values in the conditioning group and for the focal predictor. The user can control how the fixed values are determined by specifying the `fixed.predictors` argument. This argument takes a list of sub-arguments that allow for controlling each predictor in the fixed group individually, with different rules for factors and numeric predictors.

For a fixed factor, imagine computing the fitted values evaluating the factor at each of its levels. The fitted value that is used is the weighed average of these within-level fitted values, with weights proportional to the number of observations at each level of the factor. This is an appropriate notion of “typical” if the data at hand are viewed as a random sample from a population, and so the sample fraction at each level estimates the population fraction.

A second approach is to average the level-specific fitted values with equal weights at each level. This may be appropriate, for example, in designed experiments in which the levels of a factor are assigned by an investigator. The latter method is invoked by setting `fixed.predictors=list(given.values="equal")`.

You can construct other weighting schemes for averaging over the levels of a factor, as described on the help page for the `Effect()` function.

2.2.2 Numeric Predictors

For numeric predictors in the fixed group, the default method of selecting a typical value is to apply the `mean()` function. The specification `fixed.predictors=list(typical=median)` would use the `median()` function; in general, `typical` can be any function that evaluates a vector and return a single number.

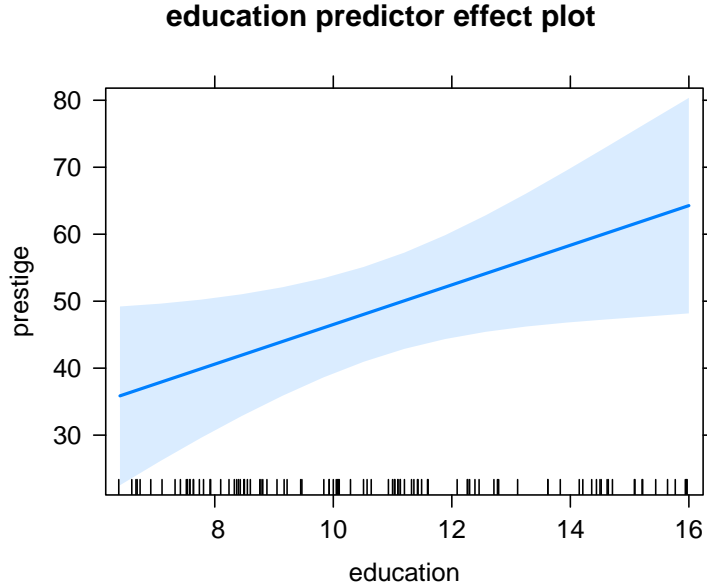
Other sub-arguments to `fixed.predictors` apply to the use of offsets, and to the `survey` package; see the help page for the `Effect()` function.

2.3 Standard Errors and Confidence Intervals

Standard errors and confidence intervals for fitted values are computed by default. The default corresponds to setting the argument `se=list(compute=TRUE, type="pointwise", level=.95)`. Setting `se=FALSE` omits standard errors, `type="scheffe"` uses wider Scheffé intervals that adjust for multiple testing, and `level=.8` produces 80% intervals.

Standard errors are based by default on the “usual” estimated covariance matrix of the estimated regression coefficients. You can replace the default coefficient covariance matrix with some other estimate, such as one obtained from the bootstrap or a sandwich coefficient covariance matrix estimator, by setting the argument `vcov` either to a function that returns a coefficient covariance matrix such as `hccm()` in the `car` package for linear models, or to a matrix of the correct size; for example:

```
R> e4.lm1 <- predictorEffect("education", lm1,
+                           se=list(type="scheffe", level=.99),
+                           vcov. = hccm)
R> plot(e4.lm1)
```



This plot displays 99% Scheffé intervals based on a robust coefficient covariance matrix computed by the sandwich method; see `help("hccm")`.

2.4 Residuals

The argument `residuals=TRUE` adds partial residuals to subsequent effect plots, a topic that we discuss in Section 4.

3 Arguments for Plotting Predictor Effects and Effects

The arguments described in Section 2 are for the `predictorEffect()` function or the `Effect()` function, and modify the computations that are performed, such as methods for averaging and fixing predictors, and for computing standard errors. Arguments to the `plot()` methods for the predictor effects and effects objects produced by the `predictorEffect()` and `Effect()` functions are described in this section, and these change the appearance of an effect plot or modify the quantities that are plotted. The help page `help("plot.eff")` describe these optional arguments in more details than the examples we provide here.⁷

In late 2018, we reorganized the `plot()` method for effect objects by combining arguments into five major groups of related sub-arguments, with the goal of simplifying the specification of effect plots. For example, the `lines` argument group is a list of sub-arguments for determining line type, color, and width, whether or not multiple lines should be drawn on the same graph, and whether plotted lines should be smoothed. The defaults for these sub-arguments are the choices we generally find the most useful, but they will not be the best choices in all circumstances. The cost of reorganizing the arguments in this manner is the necessity of specifying arguments as lists, some of whose elements are themselves lists, requiring the user to make sure that parentheses specifying the lists are properly balanced.

In addition to the five argument groups that we describe below, the `plot()` method for effect objects accepts the arguments `main` for the main title of a plot and `id` for identifying points in plots that includes residuals (as discussed in Section 4).

Finally, the `plot()` method has a number of “legacy” arguments shown in the help file. They have been kept so existing scripts using effects would not break, but they are all duplicated as sub-

⁷The `Effect()` function produces an S3 object of class "eff"; `predictorEffect()` produces an S3 object of class "predictoreff" inheriting from class "eff". The `plot()` method for class "predictoreffect" sets up a call to the method for class "eff", and the latter does most of the work of creating a predictor-effect plot.

arguments in the argument groups. The legacy arguments work, but they may not be supported forever, so learners should use the arguments and sub-arguments.

3.1 axes Group: Specify Axis Characteristics

The **axes** argument group has two major sub-arguments, **x** for the horizontal axis, **y** for the vertical axis, and two minor sub-arguments, the **grid** argument, which adds a background grid to the plot, and the **alternating** argument, for changing the placement of axis-tick labels in multi-panel plots.

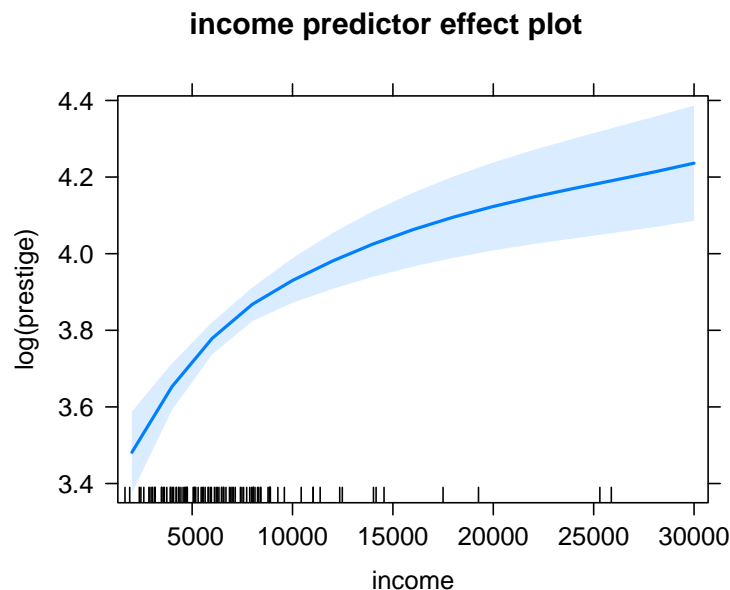
3.1.1 x: Horizontal Axis Specification

We introduce another linear model to serve as an example:

```
R> lm2 <- lm(log) ~ log(income) + education + type, Prestige)
```

The default predictor effect plot for `income` is

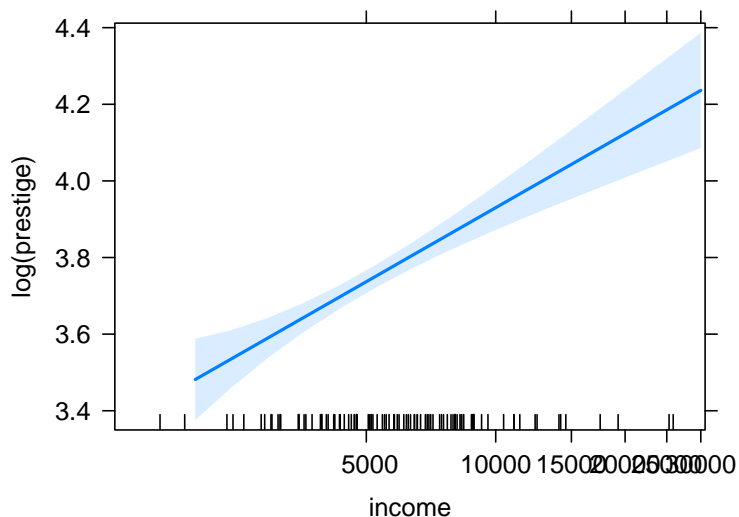
```
R> plot(predictorEffects(lm2, ~ income))
```



The plot is curved because the predictor `income` is represented by its logarithm in the model formula, but the default predictor effect plot uses the predictor `income`, not the regressor `log(income)`, on the horizontal axis. The `x` sub-argument can be used to transform the horizontal axis, for example to replace `income` by `log(income)`:

```
R> plot(predictorEffects(lm2, ~ income),  
+       axes=list(  
+         x=list(income=list(transform=list(trans=log, inverse=exp)))  
+       ))
```

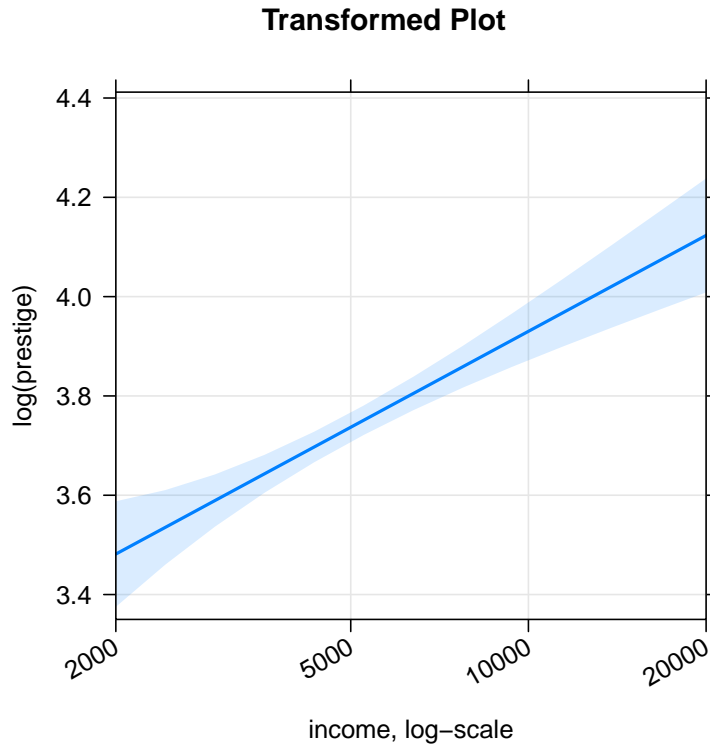
income predictor effect plot



The transformation changes the scale on the horizontal axis to log-scale, but leaves the tick labels in arithmetic scale, and the graph is now a straight line because of the change to log-scale. This plot has several undesirable features with regard to range on the horizontal axis and over-printing of tick marks; as we will explain, the user can correct these defects by additional arguments to `plot()`.

A more elaborate version of the graph illustrates all the sub-arguments to `x` in `lines`:

```
R> plot(predictorEffects(lm2, ~ income),
+       main="Transformed Plot",
+       axes=list(
+         grid=TRUE,
+         x=list(rotate=30,
+               rug=FALSE,
+               income=list(transform=list(trans=log, inverse=exp),
+                                     lab="income, log-scale",
+                                     ticks=list(at=c(1000, 2000, 5000, 10000, 20000)),
+                                     lim=c(2000, 20000))
+             )))
```



We use the top-level argument `main="Transformed Plot"` to set the title of the plot. The `axes` argument is a list with two sub-arguments, `grid` to turn on the background grid, and `x` to modify the horizontal axis.

The `x` sub-argument is itself a list with three elements. The sub-arguments `rotate` and `rug` set the rotation angle for the tick labels and suppress the rug plot, respectively. The additional sub-argument is a list called `income`, and its name is the same as the name of the focal predictor. If you were plotting many predictor effect plots you would supply one list named for each of the focal predictors. All of the sub-arguments for `income` are displayed in the example code above. The sub-argument `transform=list(trans=log, inverse=exp)` specifies how to transform the x -axis, using the `inverse` specification. The `ticks` and `lim` sub-arguments set the tick marks and range for the horizontal axis.

This is admittedly a *complex command*, but it allows you to fine-tune the graph to look the way you want. In specifying nested argument lists, you may have problems getting the parentheses in the right places. Be careful, indent your code to clarify the structure of the command, and patient!

3.1.2 y: Specify the Vertical Axis, Linear Models

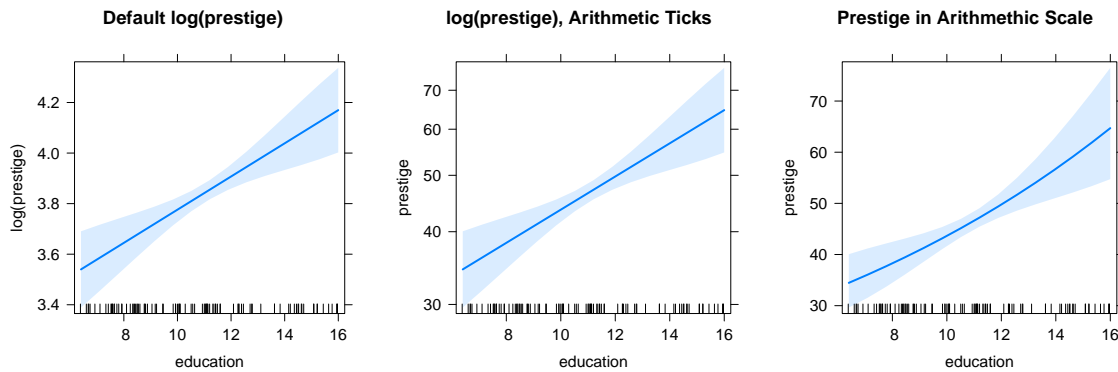
The model `lm2` has a transformed response `log(prestige)`, and “untransforming” the response to arithmetic scale may be desirable. This can be accomplished with the `y` sub-argument, which has two sub-arguments named `transform` and `type`. There are three options for drawing the predictor effect plot for a numeric response like `log(prestige)`:

```
R> # default:
R> plot(predictorEffects(lm2, ~ education),
+       main="Default log(prestige)")

R> # Change only tick-mark labels to arithmetic scale:
R> plot(predictorEffects(lm2, ~ education),
+       main="log(prestige), Arithmetic Ticks",
+       axes=list(y=list(transform=list(trans=log, inverse=exp),
+                             lab="prestige", type="rescale")))
```



```
R> # Replace log(prestige) by prestige:
R> plot(predictorEffects(lm2, ~ education),
+       main="Prestige in Arithmetic Scale",
+       axes=list(y=list(transform=exp, lab="prestige")))
```



The first plot is the default, with a log-response. In the second plot, the `transform` sub-argument specifies the transformation of the response and its inverse, and the sub-argument `type="rescale"` changes the tick marks on the vertical axis to arithmetic scale. In the third version, with `transform=exp, lab="prestige"`, the vertical axis now is in arithmetic scale, not log scale, although that may not be completely obvious in the example because $\log(x)$ is nearly linear over the range of 30 to 80 for `prestige` values in this data set: Look closely to see that the axis ticks marks in the second graph are unequally spaced, while those in the third graph are equally spaced and the plotted line is slightly curved. The help page `?plot.eff` provides a somewhat more detailed explanation of these options.

As a second example we will reconstruct Figure 7.10 in Fox and Weisberg (2019a, Sec. 7.2). In that section, we fit a linear mixed-effects model to data from the `Blackmore` data frame in the `carData` package. We transformed the response variable in the model, hours of `exercise`, using a transformation in a modified Box-Cox power family that allows zero or negative responses, explained briefly by Fox and Weisberg (2019a, Sec. 3.4) and more thoroughly by Hawkins and Weisberg (2017). The fitted model is:

```
R> library("lme4") # for lmer()

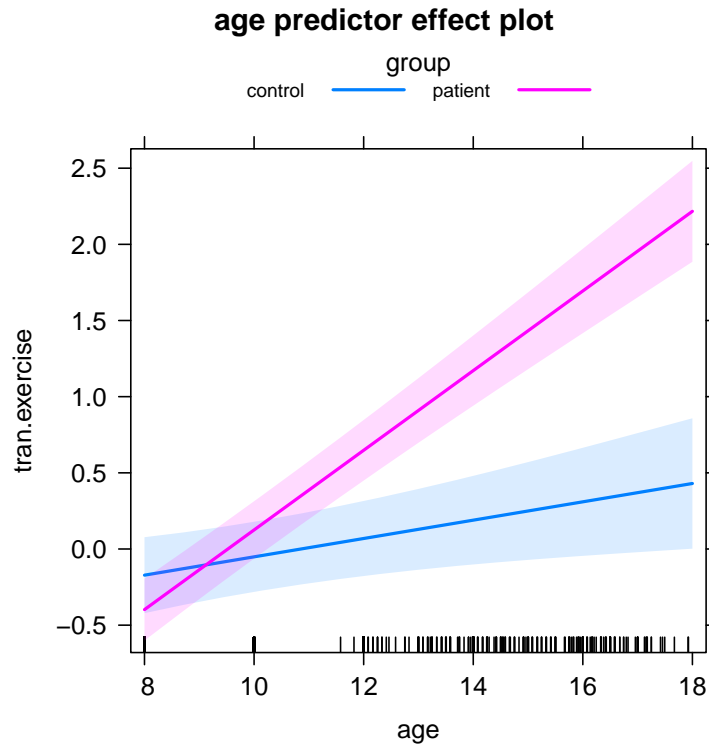
Loading required package: Matrix

R> Blackmore$tran.exercise <- bcnPower(Blackmore$exercise,
+                                     lambda=0.25, gamma=0.1)
R> mm1 <- lmer(tran.exercise ~ I(age - 8)*group +
+             (I(age - 8) | subject), data=Blackmore)
```

This model, with numeric predictor `age` and factor predictor `group`, is a linear mixed model with random intercepts and slopes for `age` that vary by `subject`. The response variable is a transformation of `exercise` similar to the fourth root with adjustment for zero values; see `help("bcnPower")`.

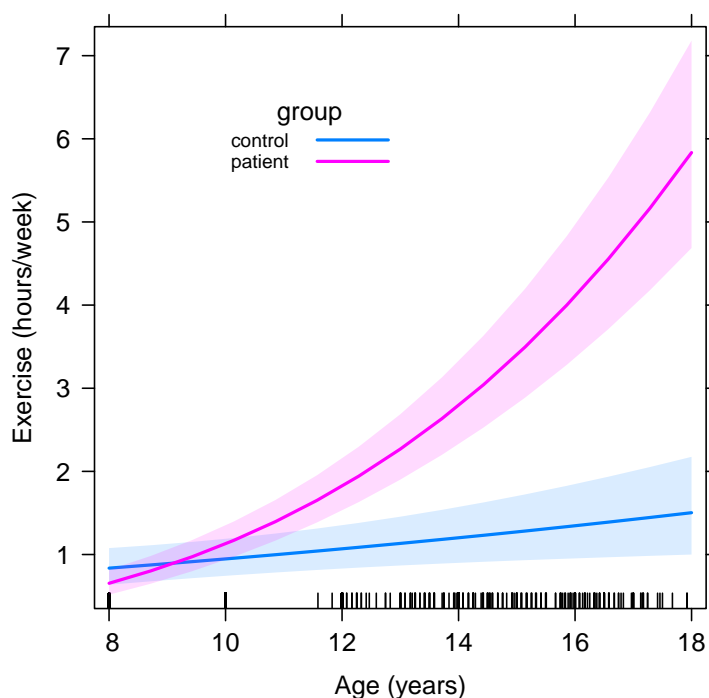
The predictor effect plot for the fixed effect of `age` is

```
R> e1.mm1 <- predictorEffect("age", mm1)
R> plot(e1.mm1, lines=list(multiline=TRUE), confint=list(style="auto"))
```



The plot clearly shows the difference between the "control" and "patient" groups, with the fitted response for the latter having a larger slope. The graph is hard to decode, however, because the vertical axis is more or less in the scale of the fourth-root of hours of exercise, so untransforming the response may produce a more informative plot. Because the `bcnPower()` transformation is complex, the `car` package includes the function `bcnPowerInverse()` to reverse the transformation:

```
R> f.trans <- function(x) bcnPower(x, lambda=0.25, gamma=0.1)
R> f.inverse <- function(x) bcnPowerInverse(x, lambda=0.25, gamma=0.1)
R> plot(e1.mm1, lines=list(multiline=TRUE),
+       confint=list(style="auto"),
+       axes=list(y=list(transform=list(trans=f.trans,
+                                       inverse=f.inverse),
+                                       type="response")),
+       lattice=list(key.args=list(x=.20, y=.75, corner=c(0, 0),
+                                       padding.text=1.25)),
+       xlab="Age (years)", ylab="Exercise (hours/week)", main=""
+ )
```



The response scale is now in hours per week, and we see that hours of exercise increase more quickly on average in the patient group for older subjects. We use several additional arguments in this plot to match Fox and Weisberg (2019a, Fig. 7.10), including moving the key inside of the graph (see Section 3.4.1), changing the axis labels, and removing the main title to the plot using standard arguments for **lattice** graphics.⁸

3.1.3 y: Specify the Vertical Axis, Generalized Linear Models

Transforming the vertical axis for generalized linear models also uses the **y** sub-argument to the **axes** argument. You do not need to specify the **transform** argument because **plot()** obtains the right functions from the regression model's **family** component. The **type** sub-argument has the same three possible values as for linear models, but their interpretation is somewhat different:

1. Effect plots in **type="link"** or linear predictor scale, in which the horizontal axis of each plot is a predictor and the vertical axis is in the scale of the linear predictor. For logistic regression, the vertical axis is the log-odds (logit) scale. For Poisson regression with log-link the vertical axis is the log-mean (log-count) scale.
2. Predictor effect plots in **type="response"** or mean scale are obtained by “untransforming” the *y* axis using the inverse of the link function. For the log-link, this corresponds to transforming the *y* axis and plotting $\exp(y)$. For logistic regression, $y = \log(p/(1 - p))$ and, solving for p , $p = \exp(y)/[1 + \exp(y)] = 1/[1 + \exp(-y)]$, so the plot in mean scale uses $1/[1 + \exp(-y)]$ on the vertical axis.
3. We also a third option, **type="rescale"**, which plots in linear predictor (e.g., logit) scale, but labels the tick marks on the vertical axis in mean (e.g., probability) scale. This third option, which retains the linear structure of the model but labels the vertical axis on the usually more familiar mean scale, is the default.

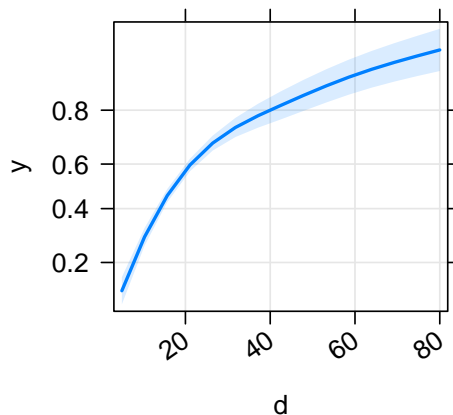
⁸The code shown for this graph in Fox and Weisberg (2019a) uses “legacy” arguments, and is therefore somewhat different from the code given here. Both commands produce the same plot.

We use the `Blowdown` data from the `alr4` package to provide examples. These data model the probability of blowdown, a tree being uprooted as the result of a major straight-line wind storm in the Boundary Waters Canoe Area Wilderness in 1999, as a function of the diameter `d` of the tree, the local severity `s` of the storm, and the species `spp` of the tree. We fit a main-effects model and then display all three predictor effect plots:

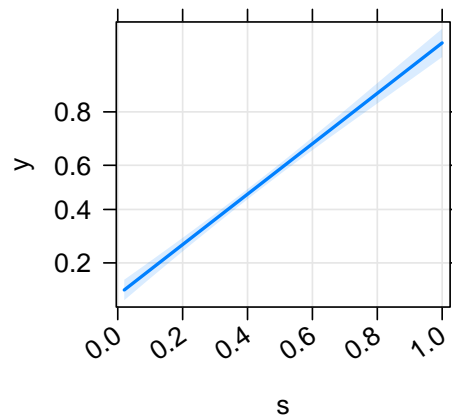
```
R> library("alr4") # for the data
R> gm1 <- glm(y ~ log(d) + s + spp, family=binomial, data=Blowdown)

R> plot(predictorEffects(gm1),
+       axes=list(grid=TRUE, x=list(rug=FALSE, rotate=35)))
```

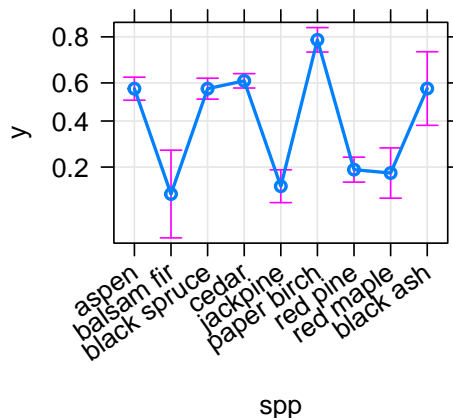
d predictor effect plot



s predictor effect plot



spp predictor effect plot



The `rug=FALSE` sub-argument to `x` suppresses the rug plot that appears by default at the bottom of graphs for numeric predictors, and the `grid` sub-argument to `axes` adds background grids. The `rotate` sub-argument prints the horizontal tick labels at an angle to avoid overprinting.

Interpretation of GLM predictor effect plots in link scale is similar to predictor effect plots for linear models, and all the modifications previously described can be used for these plots. Because the default is `type="rescale"`, the vertical axis is in linear predictor scale, but the vertical axis labels are in probability scale, so the tick-marks are not equally spaced.

The next three graphs show the possible values of `type`:

```
R> e1.gm1 <- predictorEffect("spp", gm1)
R> plot(e1.gm1, main="type='rescale'",
```

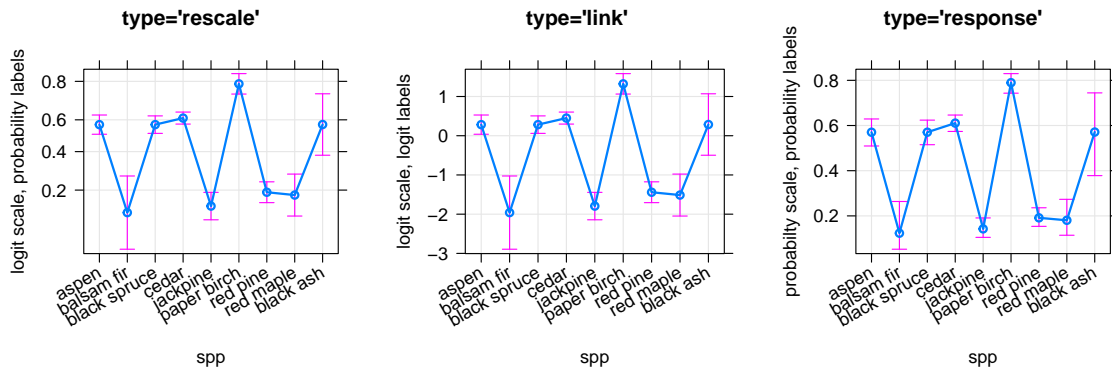
```

+       axes=list(y=list(type="rescale",
+                         lab="logit scale, probability labels"),
+                 x=list(rotate=30),
+                 grid=TRUE))

R> plot(e1.gm1, main="type='link'",
+       axes=list(y=list(type="link",
+                         lab="logit scale, logit labels"),
+                 x=list(rotate=30),
+                 grid=TRUE))

R> plot(e1.gm1, main="type='response'",
+       axes=list(y=list(type="response", grid=TRUE,
+                         lab="probability scale, probability labels"),
+                 x=list(rotate=30),
+                 grid=TRUE))

```



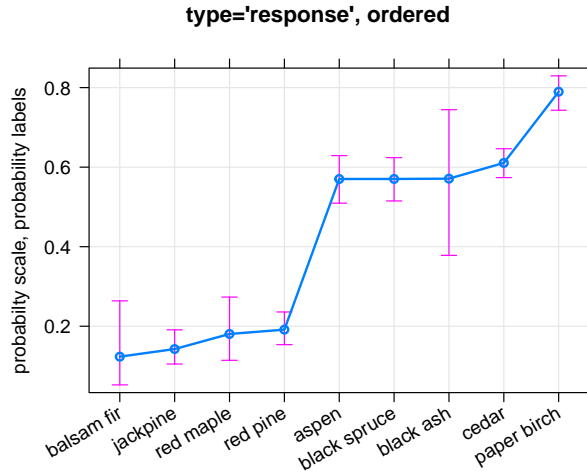
The first two graphs show the same plot, but in the first the tick-marks on the vertical axis are unequally spaced and are in probability scale, while in the second the tick-marks are equally spaced and are in log-odds scale. In the third graph, the vertical axis has been transformed to probability scale, and the corresponding tick-marks are now equally spaced.

The plot for species would be more helpful if the levels of the factor were ordered according to the estimated log-odds of blowdown. First, we need to recover the fitted values in link scale, which are log-odds of blowdown for a logistic model. The fitted log-odds are stored in `as.data.frame(e1.gm1)$fit` using the `e1.gm1` object previously computed.

```

R> or <- order(as.data.frame(e1.gm1)$fit) # order smallest to largest
R> Blowdown$spp1 <- factor(Blowdown$spp,
+                          levels=levels(Blowdown$spp)[or])
R> gm2 <- update(gm1, ~ . - spp + spp1)
R> plot(predictorEffects(gm2, ~ spp1), main="type='response', ordered",
+       axes=list(y=list(type="response",
+                         lab="probability scale, probability labels"),
+                 x=list(rotate=30, spp=list(lab="Species")),
+                 grid=TRUE))

```



In this order, the separation of species into two groups of lower and higher probability species is reasonably clear, with paper birch more susceptible to blowdown than the other species and possibly in a group by itself.

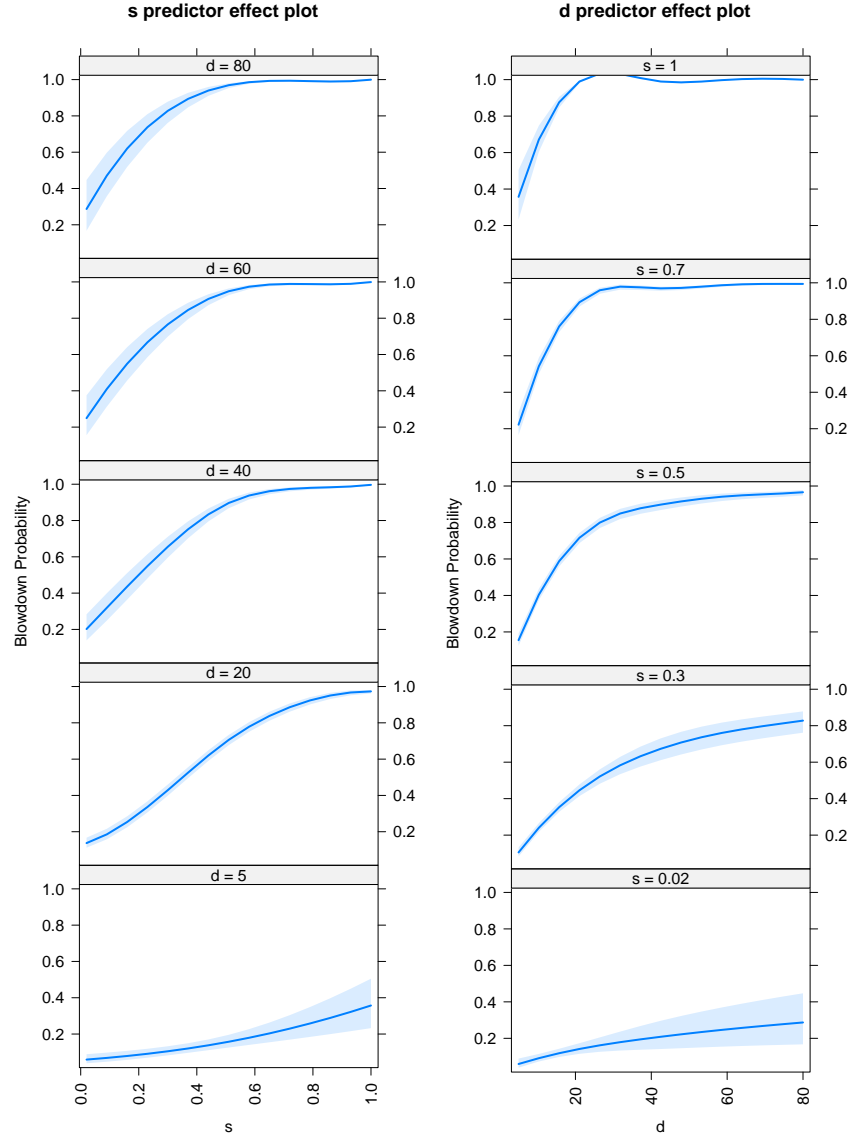
3.2 lines Group: Specifying Plotted Lines

The **lines** argument group allows the user to specify the color, type, thickness, and smoothness of lines. This can be useful, for example, if the colors used by **effects** by default are for some reason unacceptable, such as for publications in which only black or gray-scale lines are permitted. The most common use of this argument group is to allow more than one line to be plotted on the same graph or panel using the **multiline** sub-argument.

3.2.1 multiline and z.var: Multiple Lines in a Plot

Default predictor effect plots with conditioning variables generate a separate plot for each level of the conditioning variable(s). For an example, we add the **log(d):s** interaction to the model **gm1**, and plot the predictor effect plots for **s** and for **d**:

```
R> gm3 <- update(gm2, ~ . + s:log(d)) # add an interaction
R> plot(predictorEffects(gm3, ~ s + d),
+       axes=list(x=list(rug=FALSE, rotate=90),
+                 y=list(type="response", lab="Blowdown Probability")),
+       lattice=list(layout=c(1, 5)))
```



As explained in the preceding section, setting the sub-argument `type="response"` for the `y` axis, plots the response on the probability scale; setting `layout=c(1, 5)` arranges each predictor effect plot in 1 column of 5 rows (see the description of the `lattice` argument in Section 3.4). The predictor effect plot for `s` conditions on the level of `d`, and displays the plot of the fitted values for `y` versus `s` in a separate panel for each value of `d`. Similarly, the predictor effect plot for `d` displays a separate panel for each conditioning level of `s`. Confidence bands are displayed by default around each fitted line. These two graphs are based on the same fitted values, with the values of the interacting predictors `s` and `d` varying, and an fixing the factor predictor `spp` to its distribution in the data, as described in Section 2.2.1. Concentrating on the graph for the focal predictor `d`, when `s` is small the probability of blowdown is estimated to be in the range of about 0.53 to 0.57 for any value of `d`, but for larger values of `s`, the probability of blowdown increases rapidly with `d`. Similar comments can be made concerning the predictor effect plot for `s`.

Setting `multiline=TRUE` superimposes the lines for all the groups in a single graph. In the example below we will reduce the number of levels of the conditioning variable for each predictor effect plot to three each to produce simpler graphs, although this is not required:⁹

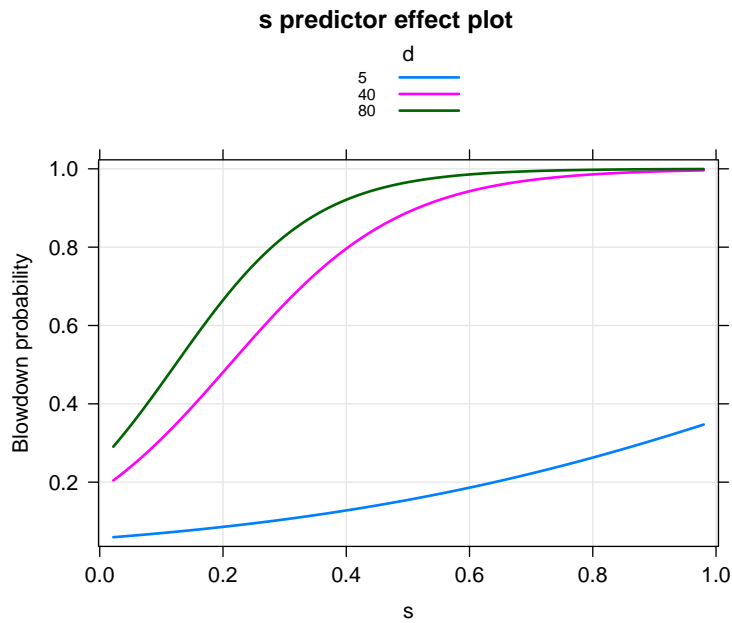
⁹We draw the plots separately because we don't want to reduce the number of values for the *focal* predictor in each plot to just three values, and indeed increase the number of values for the focal predictor to 50 from the default of five to obtain smoother plotted curves; using only three values doesn't adequate support drawing smooth curves

```

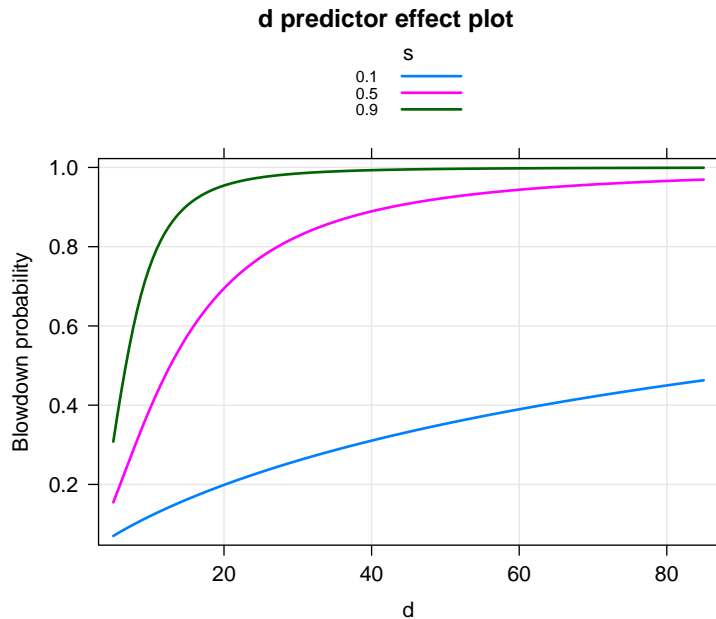
R> plot(predictorEffects(gm3, ~ s,
+                       xlevels=list(s=50, d=c(5, 40, 80))),
+       axes=list(grid=TRUE,
+                 x=list(rug=FALSE),
+                 y=list(type="response", lab="Blowdown probability")),
+       lines=list(multiline=TRUE))

R> plot(predictorEffects(gm3, ~ d,
+                       xlevels=list(s=c(0.1, 0.5, 0.9), d=50)),
+       axes=list(grid=TRUE,
+                 x=list(rug=FALSE),
+                 y=list(type="response", lab="Blowdown probability")),
+       lines=list(multiline=TRUE))

```



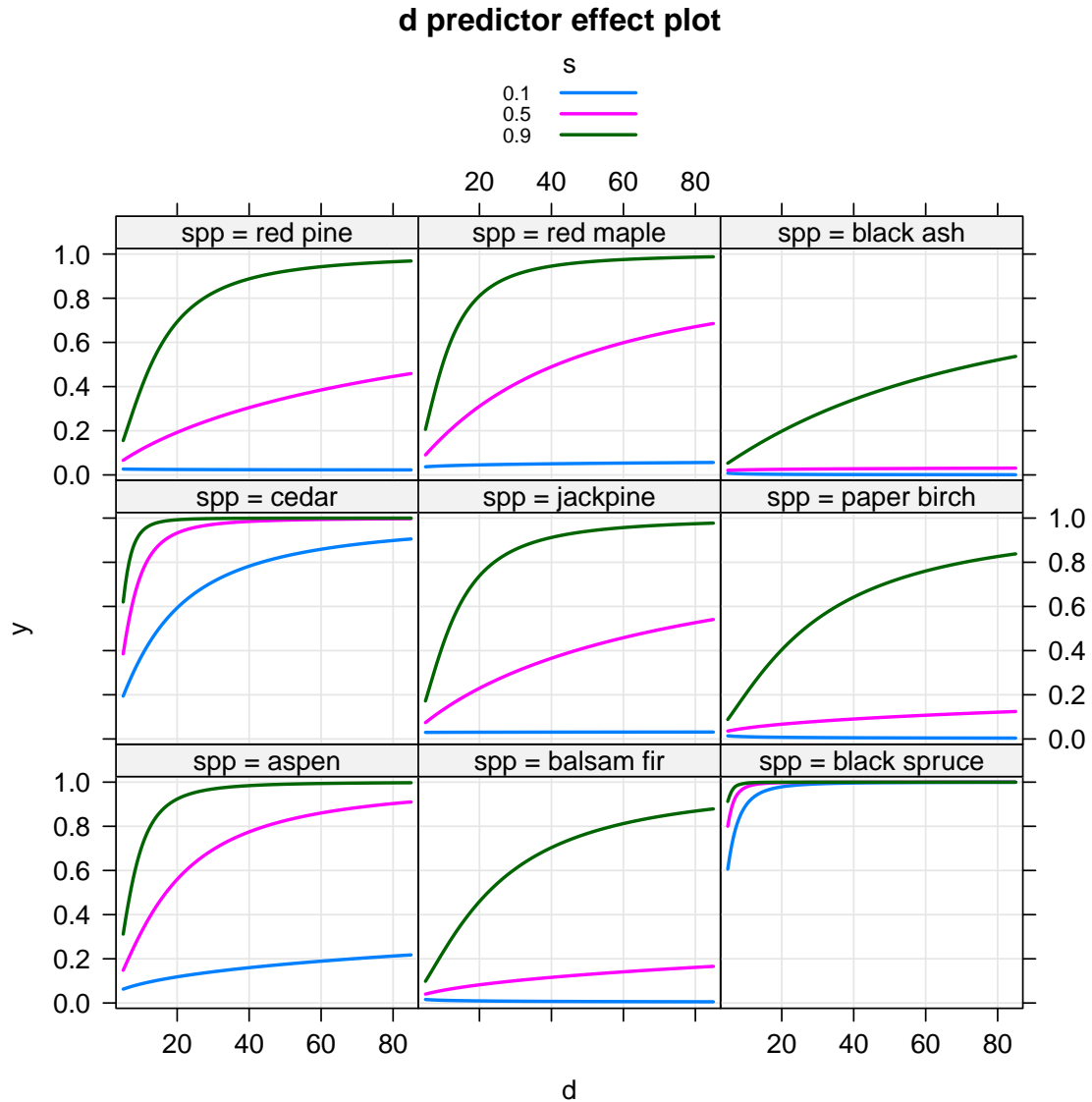
connecting the plotted points for the values of the focal predictor: Try it! If you look closely at the separate panels produced in the non-multiline version of the predictor effect plots, shown previously, you'll see inaccuracies in the plotted curves at the high end of the response probabilities produced by evaluating the probabilities at only five values of the focal predictor in each plot.



In each graph, we kept, more or less, the lowest, middle, and highest values of the conditional predictor for the interaction. We also added a grid to each graph. Multiline plots by default omit confidence bands or intervals, but these can be included using the `confint` argument (see Section 3.3). By default different values of the conditioning predictor are distinguished by color, and a key is provided. The placement and appearance of the key are controlled with a sub-argument in the `lattice` group (discussed in Section 3.4).

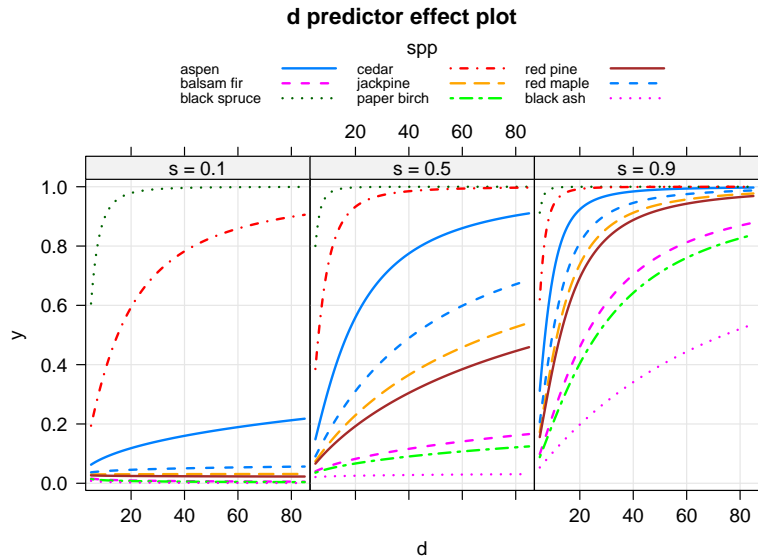
When the conditioning group includes two or more (and certainly when it includes three or more) predictors, multiline plots are almost always helpful because otherwise the resulting array of panels becomes too complicated. Suppose that we add the `spp:log(d)` interaction to the model. The predictor effect plot for `d` now includes both `s` and `spp` in the conditioning set because `d` interacts with both:

```
R> gm4 <- update(gm3, ~ . + spp:log(d))
R> plot(predictorEffects(gm4, ~ d,
+                       xlevels=list(s=c(0.1, 0.5, 0.9), d=50)),
+       axes=list(grid=TRUE,
+                 y=list(type="response"),
+                 x=list(rug=FALSE)),
+       lines=list(multiline=TRUE))
```



This plot now displays the lines for all conditioning values of *s* within the panel for each level of the conditioning factor *spp*. Compare this graph to the much more confusing plot in which different lines are drawn for the nine levels of the conditioning factor *spp*:

```
R> plot(predictorEffects(gm4, ~ d, xlevels=list(s=c(0.1, 0.5, 0.9), d=50)),
+       rug=FALSE,
+       axes=list(grid=TRUE, y=list(type="response")),
+       lines=list(multiline=TRUE, z.var="spp", lty=1:9),
+       lattice=list(layout=c(3, 1)))
```



The `z.var` sub-argument for `lines` selects the predictor that determines the lines within a panel and the remaining predictors (here just `s`) distinguish the panels. The default that the program chooses is usually, but not always, appropriate. We also used the `lattice` argument to set the array of graphs to have 3 columns and 1 row, and different the lines by type (`lty`, see Section 3.2.2 immediately below) as well as by color.

3.2.2 Line Color, Type, Width, Smoothness

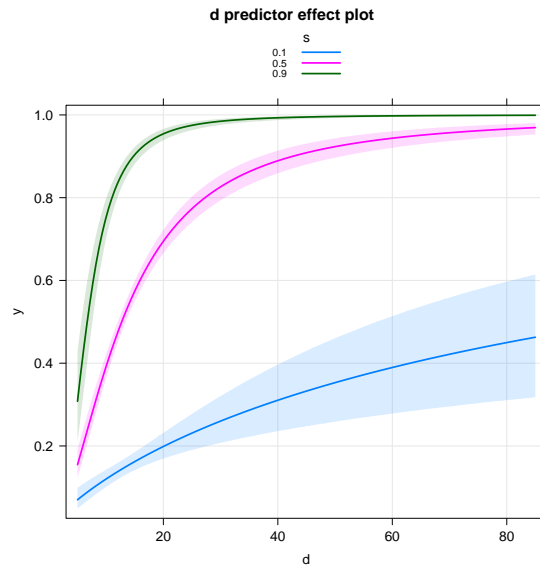
Different lines in the same plot are differentiated by default using color. This can be modified by the sub-arguments `lty`, `lwd` and `col` to set line types, widths, and colors, respectively. For example, in the last graph shown you can get all black lines of different line types using `lines=list(multiline=TRUE, col="black", lty=1:9)`, or using a gray scale, `lines=list(multiline=TRUE, col=gray((1:9)/10))`.

The `plot()` method for effect objects uses smoothing splines to interpolate between plotted points. Smoothing can be turned off with `splines=FALSE` in the `lines` argument, but we rarely expect this to be a good idea, although it may be necessary to increase the number of values at which the focal predictor is evaluated to support a reasonable spline smooth, as we discovered in some of the examples above.

3.3 confint: Specifying Confidence Interval Style and Inclusion

The `confint` argument group controls the inclusion and appearance of confidence intervals and regions. This argument takes three sub-arguments: The `style` sub-argument is either `"bars"`, for confidence bars typically around the adjusted fitted value for a factor, `"bands"` for shaded confidence bands typically for numeric focal predictors, `"auto"` to let the program automatically choose between `"bars"` and `"bands"`, `"lines"` to draw only the edges of confidence bands with no shading, or `"none"` to suppress confidence intervals. The default is `"auto"` when `multiline=FALSE` and `"none"` when `multiline=TRUE`. Setting `confint="auto"` produces bars for factors and bands for numeric predictors. For example

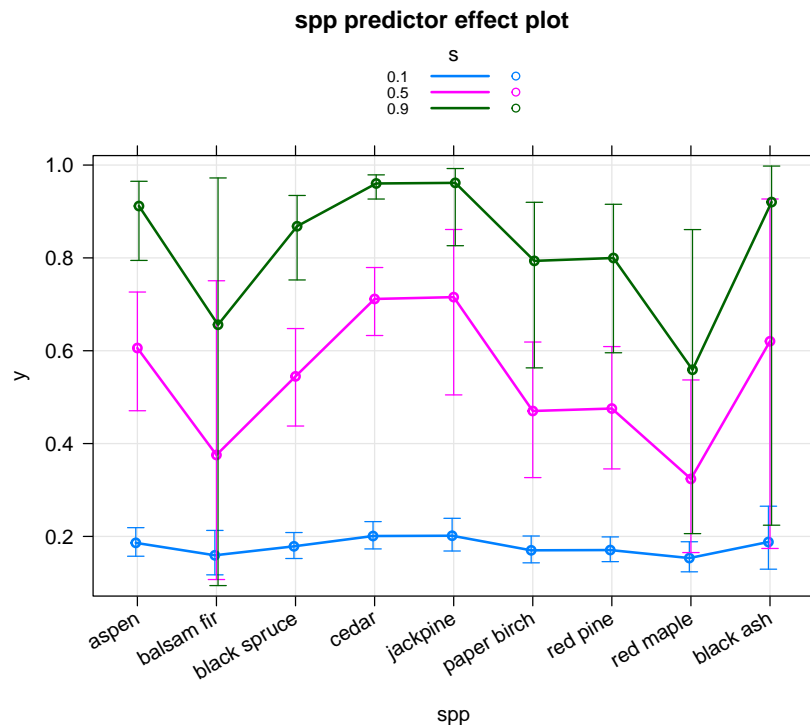
```
R> plot(predictorEffects(gm3, ~ d, xlevels=list(s=c(0.1, 0.5, 0.9), d=50)),
+       axes=list(grid=TRUE,
+                 x=list(rug=FALSE),
+                 y=list(type="response")),
+       lines=list(multiline=TRUE),
+       confint=list(style="auto"))
```



In this example the confidence bands are well separated, so including them in a multiline graph isn't problematic; in other cases, overlapping confidence bands produce an artistic but uninterpretable mess.

With a factor focal predictor, we get:

```
R> gm5 <- update(gm2, ~ . + spp:s)
R> plot(predictorEffects(gm5, ~ spp, xlevels=list(s=c(0.1, 0.5, 0.9))),
+       axes=list(grid=TRUE,
+                 y=list(type="response"),
+                 x=list(rug=FALSE, rotate=30)),
+       lines=list(multiline=TRUE),
+       confint=list(style="auto"))
```



The error bars for the various levels of **s** are slightly staggered to reduce over-plotting.

Two additional arguments, **col** and **alpha**, control respectively the color of confidence bars and regions and the transparency of confidence regions. Users are unlikely to need these options. Finally, the type of confidence interval shown, either pointwise or Scheffé corrected for multiple comparisons, is controlled by the **se** argument to the **predictorEffect()** or **Effect()** function (see Section 2.3).

3.4 lattice: Specifying Standard lattice Package Arguments

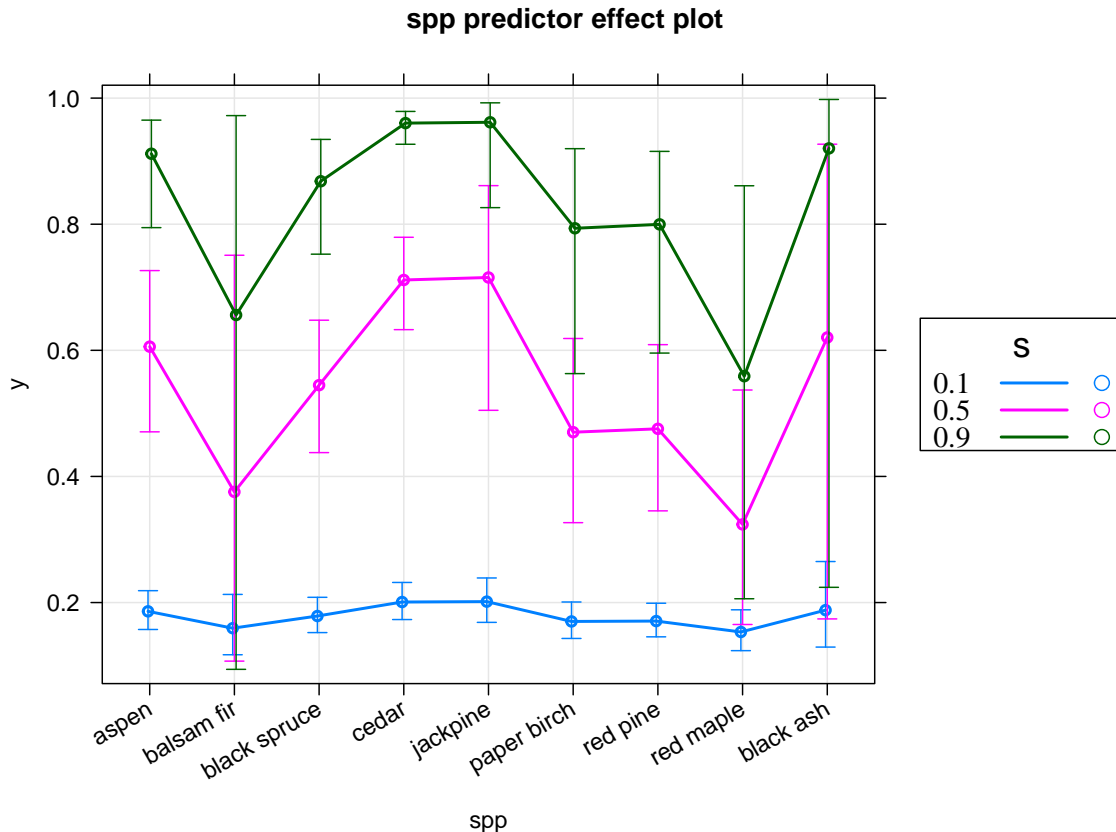
The **plot()** methods defined in the **effects** package use functions in the **lattice** package (Sarkar, 2008), such as **xyplot()**, to draw effect plots, which often comprise rectangular arrays of panels. In particular, the **plot()** method for the "eff" objects returned by the **Effect()** function are "trellis" objects, which can be manipulated in the normal manner.

The **lattice** group of arguments to the **plot()** method for effect objects may be used to specify various standard arguments for **lattice** graphics functions such as **xyplot()**. In particular, you can control the number of rows and columns when panels are displayed in an array, modify the key (legend) for the graph, and specify the contents of the "strip" displayed in the shaded region of text above each panel in a **lattice** array. In addition, the **array** sub-argument, for advanced users, controls the layout of multiple predictor effect plots produced by the **predictorEffects()** function.

3.4.1 Modifying the Key With key.args

A user can modify the placement and appearance of the key with sub-arguments of the **key.args** argument. For example:

```
R> plot(predictorEffects(gm5, ~ spp, xlevels=list(s=c(0.1, 0.5, 0.9))),
+       rug=FALSE,
+       axes=list(grid=TRUE,
+                 y=list(type="response"),
+                 x=list(rotate=30)),
+       lines=list(multiline=TRUE),
+       confint=list(style="auto"),
+       lattice=list(key.args=list(space="right",
+                                   columns=1,
+                                   border=TRUE,
+                                   fontfamily="serif",
+                                   cex=1.25,
+                                   cex.title=1.5)))
```

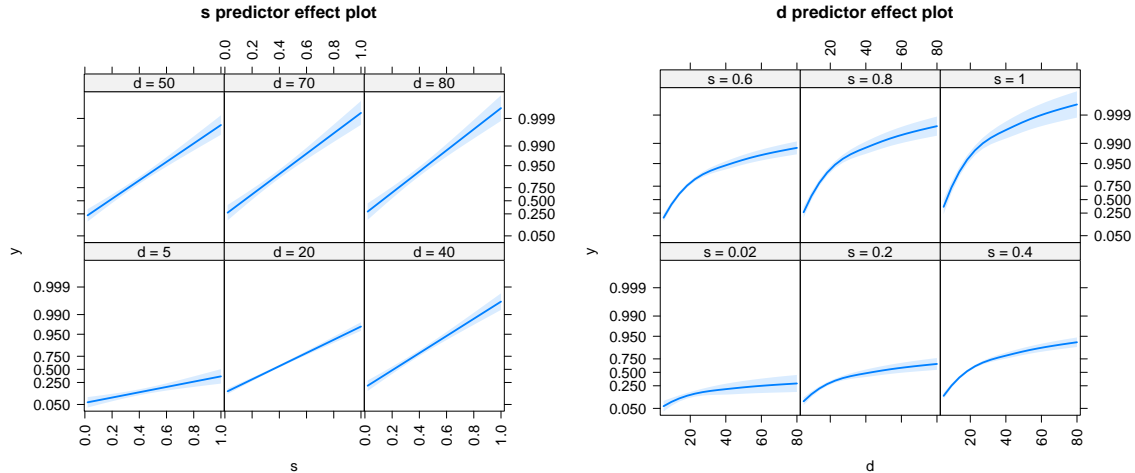


The sub-argument `space="right"` moves the key to the right of the graph, overriding the default `space="top"`. Alternatively the key can be placed inside the graph using the `x`, `y`, and `corner` sub-arguments, as illustrated in the graph on page 19. The choices for `fontfamily` are `c("sans", "serif")` and affect only the key; the rest of the plot uses `"sans"`. The sub-arguments `cex` and `cex.title` control the relative sizes of the key entries and the key title, respectively. Finally any argument documented in `help("xyplot")` in the `key` section can be set with this argument. If you use the default `space="top"` for placement of the key, you may wish to adjust the number of columns in the key, particularly if the level names are long.

3.4.2 layout

The `layout` sub-argument to the `lattice` argument allows a user to select a layout of multiple panels in a lattice graph; for example:

```
R> plot(predictorEffects(gm3, ~ s + d, xlevels=list(s=6, d=6)),
+       axes=list(x=list(rug=FALSE, rotate=90),
+                 y=list(ticks=list(at=c(.999, .99, .95, .75, .5, .25, .05)))),
+       lattice=list(layout=c(3, 2)))
```



Here, the `layout` sub-argument specifies an array of 3 columns and 2 rows for each of the predictor effect plots.

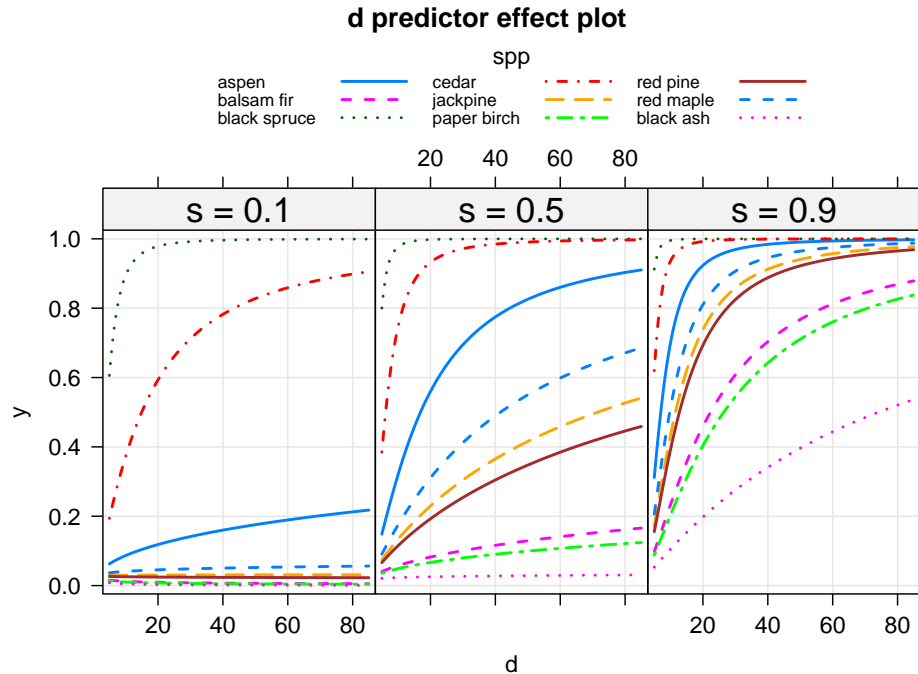
3.4.3 array of Multiple Plots

If you create several predictor effect objects with the `predictorEffects()` function, the `plot()` method for the resulting object uses the **lattice** array to divide the plotting region into sub-plots so that the predictor effect plots, each potentially with several panels, can be drawn without overlapping. The user can also supply this argument directly to create a custom array of predictor effect plots; see `help("plot.eff")` for the syntax.

3.4.4 strip

Lattice graphics with more than one panel typically provide a title at the top of each panel in an area called the *strip*. The default strip title contains the name of the conditioning predictor and the value to which it is set in the panel; if there is more than one conditioning predictor, then all of their names and corresponding values are shown. For example:

```
R> plot(predictorEffects(gm4, ~ d, xlevels=list(s=c(0.1, 0.5, 0.9), d=50)),
+       axes=list(grid=TRUE,
+                 x=list(rug=FALSE),
+                 y=list(type="response")),
+       lines=list(multiline=TRUE, z.var="spp", lty=1:9),
+       lattice=list(layout=c(3, 1),
+                     strip=list(factor.names=TRUE,
+                               values=TRUE,
+                               cex=1.5)))
```

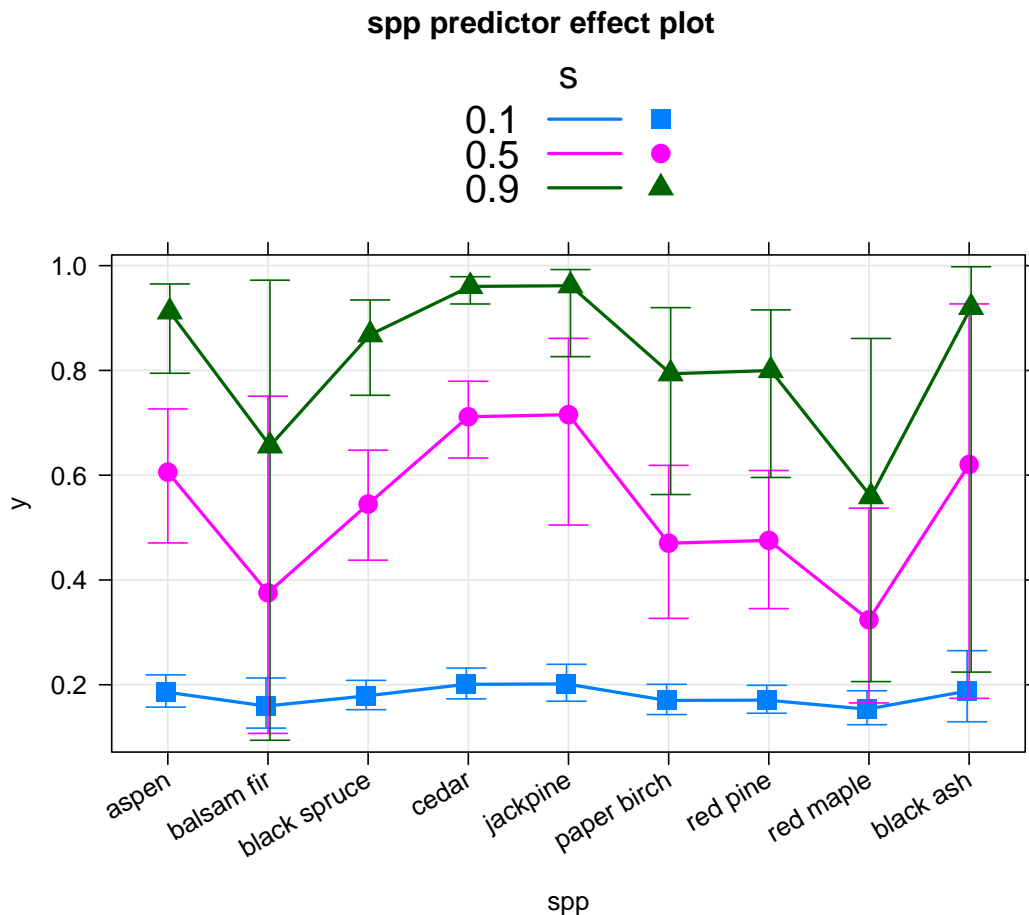


Setting `factor.names=FALSE` (the default is `TRUE`) displays only the value, and not the name, of the conditioning predictor in each strip; usually, this is desirable only if the name is too long to fit, in which case you could also rename the predictor. Setting `values=FALSE` replaces the conditioning value with a line in the strip that represents the value: the line is at the left of the strip for the smallest conditioning value, at the right for the largest value, and in a proportional intermediate position in between the two extremes. The most generally useful sub-argument is `cex`, which allows you to reduce or expand the relative size of the text in the strip, in this case increasing the size to 150% of standard size.

3.5 symbols: Plotting symbols

Symbols are used to represent adjusted fitted values (e.g., adjusted means) when the focal predictor is a factor. You can control the symbols used and their relative size:

```
R> gm5 <- update(gm2, ~ . + spp:s)
R> plot(predictorEffects(gm5, ~ spp, xlevels=list(s=c(0.1, 0.5, 0.9))),
+       symbols=list(pch=15:19, cex=1.5),
+       axes=list(grid=TRUE,
+                 y=list(type="response"),
+                 x=list(rotate=30)),
+       lines=list(multiline=TRUE),
+       confint=list(style="auto"),
+       lattice=list(key.args=list(cex=1.5, cex.title=1.5)))
```

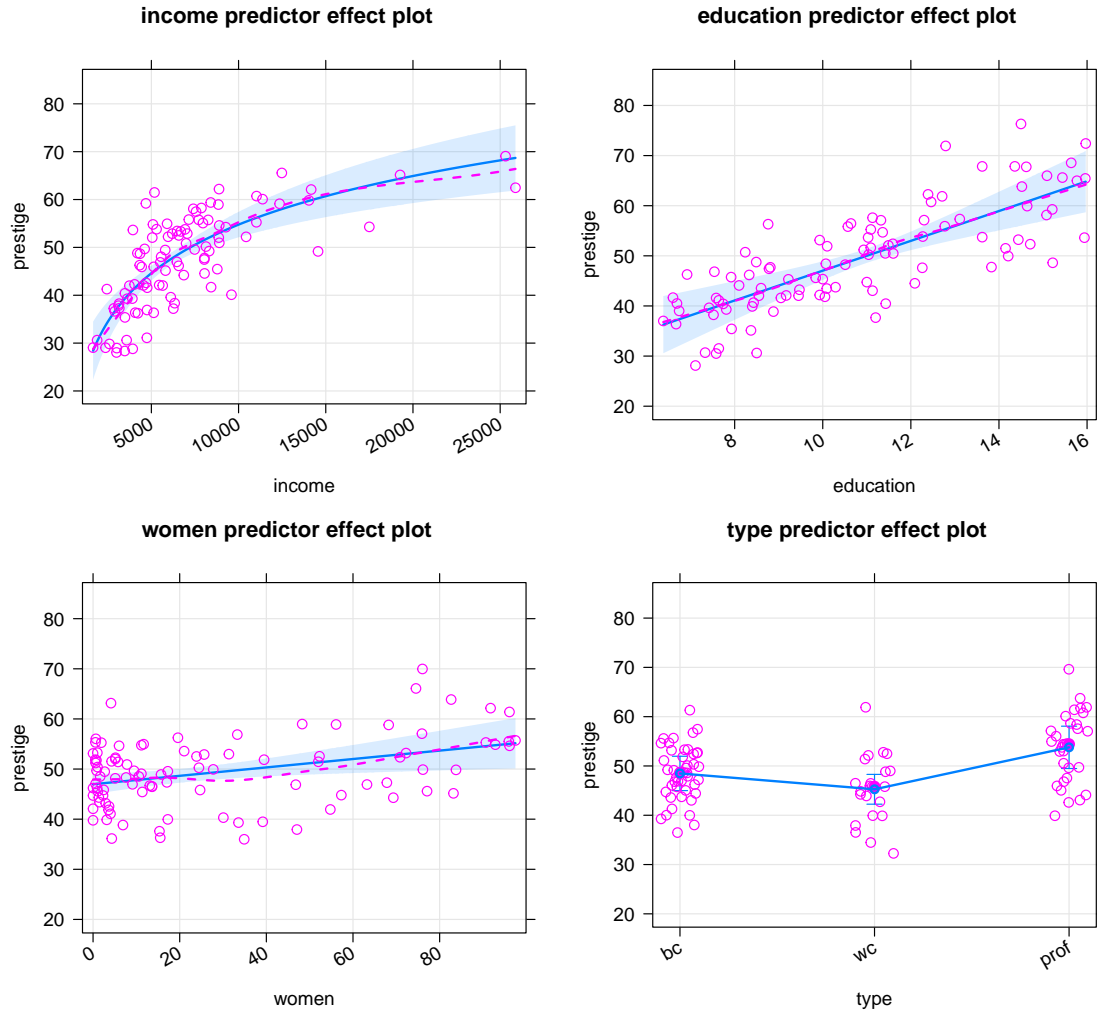
We use the `pch` sub-argument to set the symbol number for plotted symbols; you can enter the commands `plot(1:25, pch=1:25)` and `lines(1:25, lty=2, type="h")` to see the 25 plotting symbols in R. The sub-argument `pch` can also be a character vector, such as `letters[1:10]`. In this example, we set `cex=1.5` to increase the symbol size by the factor 1.5. Because only one value is given, it is recycled and used for all of the symbols. We need to change the size of the symbols in the key separately, as we do here via the `key.args` sub-argument to the `lattice` argument (see Section 3.4.1).

4 Displaying Residuals in Predictor Effect and Effect Plots

Fox and Weisberg (2019b) introduce methodology for adding partial residuals to a predictor effect or effect plot. This can be desirable to display variation in data around a fitted partial regression surface or to diagnose possible lack of fit, as the resulting plots are similar to traditional component-plus-residual plots (Fox and Weisberg, 2019a, Sec. 8.4).

The predictor effect plot for a focal predictor that does not interact with other predictors is equivalent to a standard component-plus-residual plot; for example:

```
R> lm5 <- lm(prestige ~ log(income) + education + women + type,
+           Prestige)
R> plot(predictorEffects(lm5, residuals=TRUE),
+       axes=list(grid=TRUE,
+                 x=list(rotate=30)),
+       partial.residuals=list(smooth=TRUE,
+                             lty="dashed"))
```



The partial residuals to be plotted are computed using the `residuals` argument to the `predictorEffect()`, `predictorEffects()`, or `Effect()` function. For the numeric predictors `income`, `education` and `women`, the plotted points are each equal to a point on the fitted (blue) line, representing the partial fit, plus the corresponding residual. For `income`, the fitted partial-regression line is curved because of the log transformation of the predictor, but the partial-regression function is a straight line for the other two numeric predictors. The dashed line, in the same magenta color as the plotted points on the graph, is a nonparametric-regression smooth of the points. If the model adequately represents the data, then the dashed line should approximately match the fitted partial-regression line from the model. For the factor `type`, the points have been jittered horizontally to separate them visually, because the only possible horizontal coordinates are at the three distinct factor levels. Smooths are not fit to factors and instead the conditional means of the partial residuals are plotted as magenta dots; in the current model, the magenta dots and the blue dots representing the fitted adjusted means of the response at the levels of `name` necessarily match.

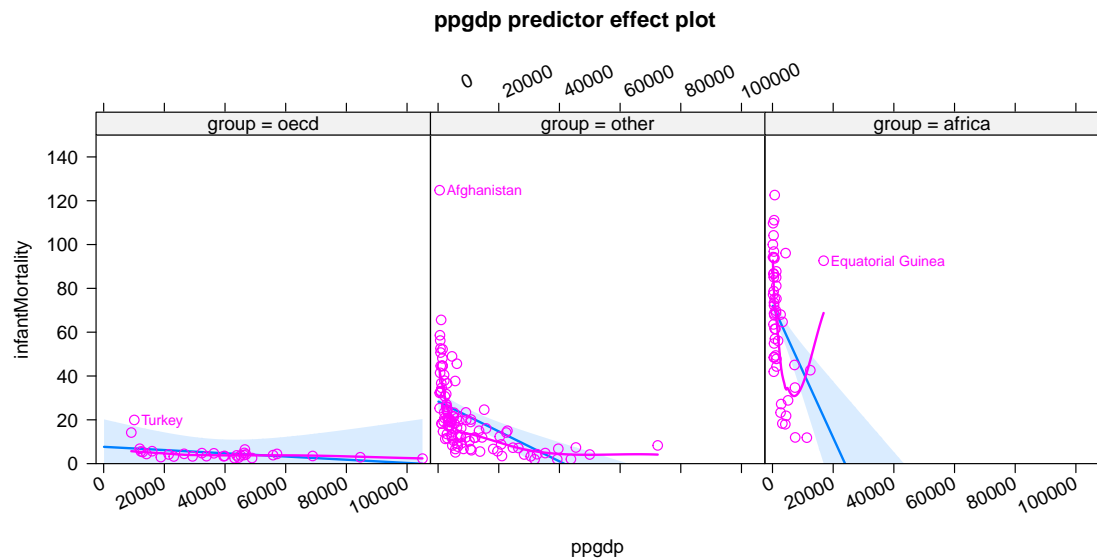
The `plot()` method for effect objects has a `partial.residuals` argument, with several sub-arguments that control how partial residuals are displayed. In the command above, we used the sub-argument `smooth=TRUE` to add the smoother (which is the default when residuals are included in the effect object), and `lty="dashed"` to change the line type for the smooth from the default solid line to a dashed line. All the `smooth` sub-arguments are described in `help("plot.eff")`.

For a second example, we turn to a linear model with an interaction, modelling `infantMortality` rate as a function of `ppgdp`, per person GDP, and country `group` in the UN data set in the `carData` package:

```

R> options(scipen=10) # suppress scientific notation
R> lm6 <- lm(infantMortality ~ group*ppgdp, data=UN)
R> plot(predictorEffects(lm6, ~ ppgdp, partial.residuals=TRUE),
+       axes=list(x=list(rotate=25),
+                   y=list(lim=c(0, 150))),
+       id=list(n=1),
+       lattice=list(layout=c(3, 1)))

```



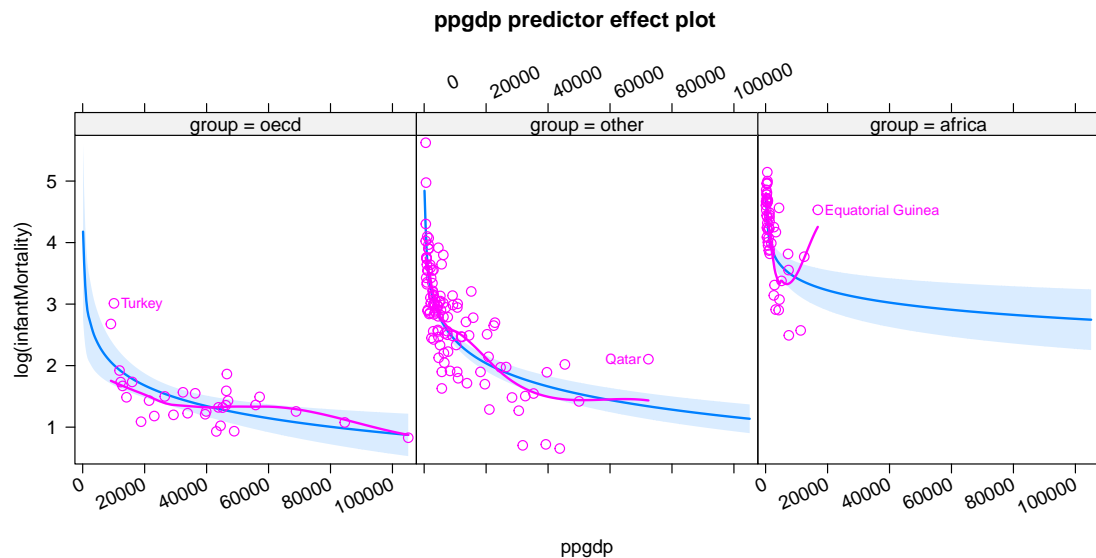
The predictor effect plot for `ppgdp` conditions on the factor `group` because of the interaction between these two predictors. Several problems are apparent in this plot: The `id` argument is used to identify the most unusual point in each panel, as described in detail in `help("plot.eff")`. Turkey has higher than predicted infant mortality for the "oecd" group, Equatorial Guinea is clearly unusual for the "africa" group, and Afghanistan, in the "other" group, has infant mortality much higher than predicted. In addition, the smooths through the points do not match the fitted lines in the "other" and "africa" groups. We use the command `options(scipen=10)` to suppress annoying scientific notation in the axis labels, and instead rotate these labels so that they fit without over-plotting.

Log-transforming both the predictor `ppgdp` and the response `infantMortality` produces a better fit to the data:

```

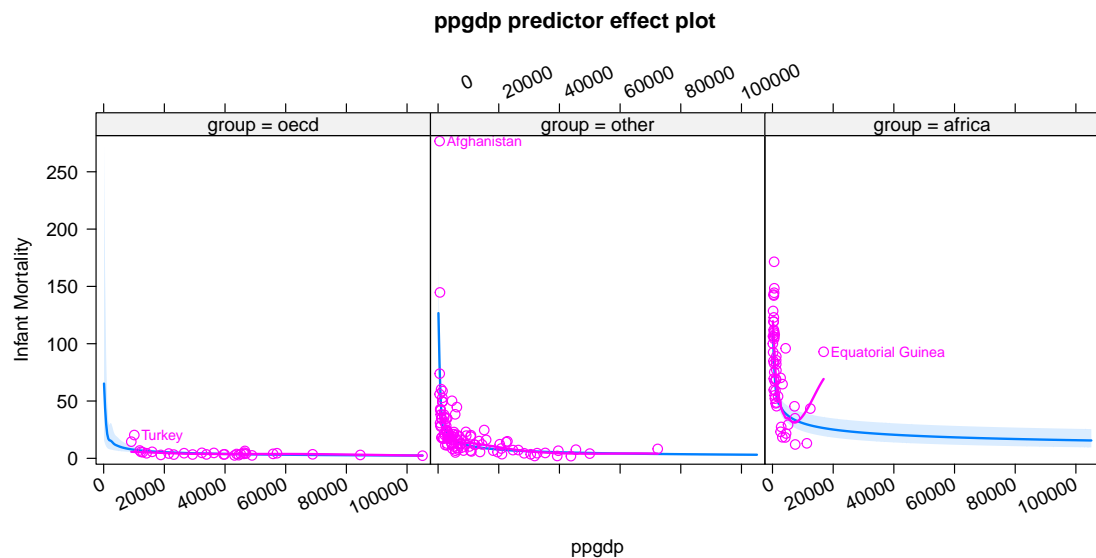
R> lm7 <- lm(log(infantMortality) ~ group*log(ppgdp), dat=UN)
R> plot(predictorEffects(lm7, ~ ppgdp, partial.residuals=TRUE),
+       axes=list(x=list(rotate=25),
+                   y=list(lim=c(0, 150))),
+       id=list(n=1),
+       lattice=list(layout=c(3, 1)))

```



Equatorial Guinea is still anomalous, however. Rescaling the vertical axis to arithmetic scale produces a slightly different, but possibly useful, picture:

```
R> plot(predictorEffects(lm7, ~ ppgdp, partial.residuals=TRUE),
+       axes=list(x=list(rotate=25),
+                 y=list(transform=list(trans=log, inverse=exp),
+                                   type="response",
+                                   lab="Infant Mortality"))),
+       id=list(n=1),
+       lattice=list(layout=c(3, 1)))
```



Partial residual plots can be added to linear or linear mixed models or to generalized linear or generalized linear mixed models in the default link scale.

5 Polytomous Categorical Responses

The **effects** package produces special graphs for ordered and unordered polytomous categorical response variables. In an ordinal regression the response is an ordered categorical variable. For

example, in a study of labor force participation the response could be not working, working part time, or working full time. The proportional-odds model (Fox and Weisberg, 2019a, Sec. 6.9) estimates the probability of a response in each of these three categories given a linear combination of regressors defined by a set of predictors, assuming a logit link function.

We illustrate the proportional-odds model with the `Womenlf` data set, for young married Canadian women, in the `carData` package and the `polr()` function in the `MASS` package:

```
R> library("MASS") # for the the polr() function
R> Womenlf$partic <- factor(Womenlf$partic,
+       levels=c("not.work", "parttime", "fulltime")) # order response levels
R> or1 <- polr(partic ~ log(hincome) + children, data=Womenlf)
R> S(or1)
```

Re-fitting to get Hessian

```
Call: polr(formula = partic ~ log(hincome) + children, data = Womenlf)
```

```
Coefficients:
              Estimate Std. Error z value      Pr(>|z|)
log(hincome)    -0.666     0.233   -2.86      0.0042
childrenpresent -1.948     0.287   -6.80 0.00000000001
```

```
Intercepts (Thresholds):
              Estimate Std. Error z value Pr(>|z|)
not.work|parttime  -2.747     0.654   -4.20 0.000027
parttime|fulltime  -1.837     0.640   -2.87  0.0041
```

Residual Deviance: 441.12

```
logLik      df      AIC      BIC
-220.56      4  449.12  463.40
```

The response variable `partic` had its levels in alphabetical order, which does not correspond to their natural ordering. We therefore start by reordering the levels to increase from "`not.work`", to "`parttime`" work, to "`fulltime`" work. The printed summary is relatively complex, and is explained in (Fox and Weisberg, 2019a, Sec. 6.9).

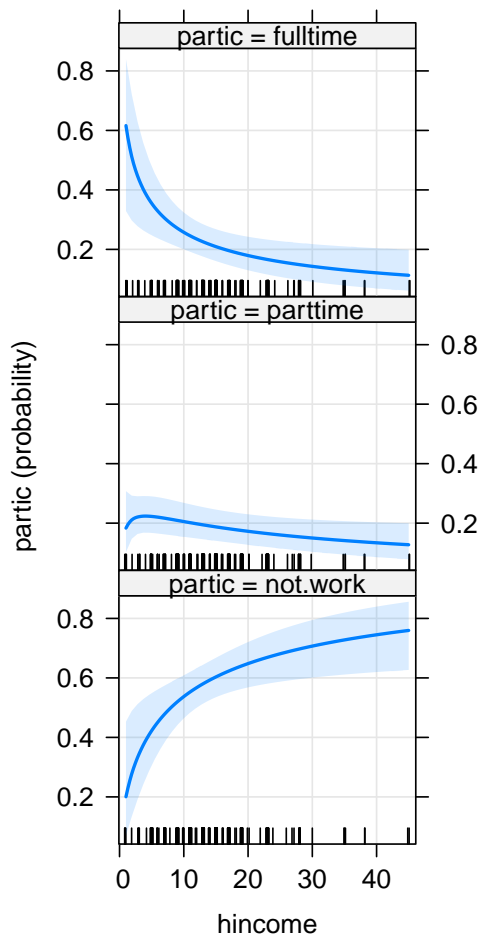
Predictor effect plots greatly simplify interpretation of the fitted model:

```
R> plot(predictorEffects(or1, xlevels=50),
+       axes=list(grid=TRUE),
+       lattice=list(key.args=list(columns=1)))
```

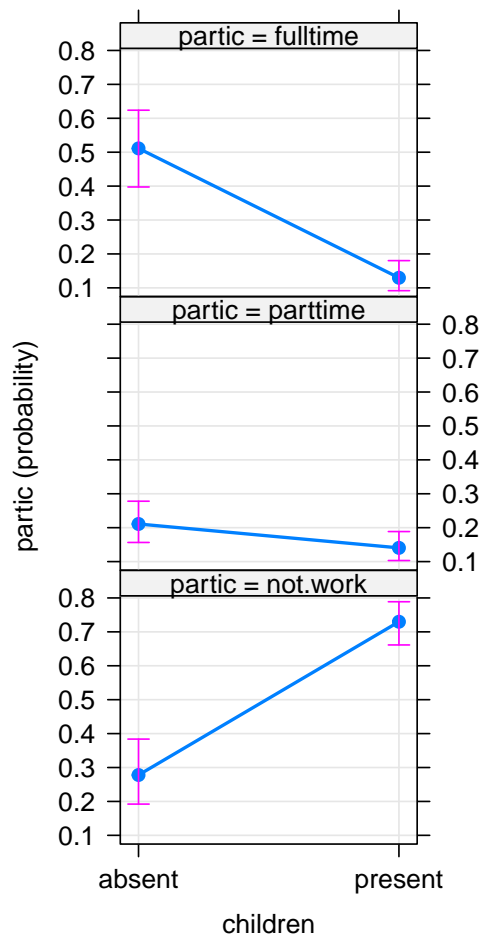
Re-fitting to get Hessian

Re-fitting to get Hessian

hincome predictor effect plot



children predictor effect plot



Unlike predictor effect plots for generalized linear models, the default scaling for the vertical axis is the probability scale, equivalent to `axes=list(y=list(type="response"))` for a GLM, and the alternative is `axes=list(y=list(type="logit"))`, which is analogous to `type="link"` for a GLM.¹⁰ Confidence bands are present by default, unless turned off with the argument `confint=list(style="none")`. We specify the argument `xlevels=50` to `predictorEffects()` to draw smoother lines for the `hincome` predictor; the default number of values to plot is five. The plot for `hincome` suggests high probability of full time work if husband's income is low, with the probability of full time work sharply decreasing to about \$15,000 and then leveling off at about .1 to .2. The probability of not working rapidly increases with husband's income, while the probability of working part time is fairly flat. A similar pattern is apparent for children present in the home, with full time work much less prevalent and not working much more prevalent when children are present than when they are absent.

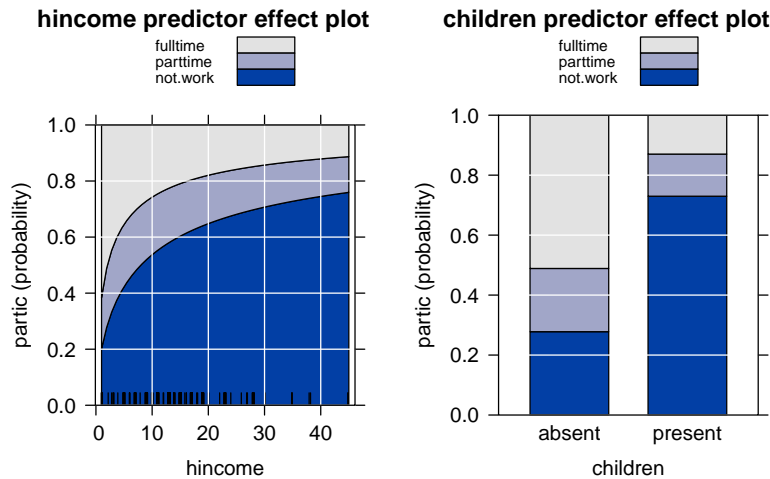
Stacked area plots are sometimes more useful for examining polytomous response models; for example:

```
R> plot(predictorEffects(or1, xlevels=50),
+       axes=list(grid=TRUE, y=list(style="stacked")),
+       lattice=list(key.args=list(columns=1)))
```

¹⁰The logits plotted, however, correspond to the individual-level probabilities and are not the adjacent-category logits in the definition of the proportional-odds model.

Re-fitting to get Hessian

Re-fitting to get Hessian



For each fixed value on the horizontal axis, the vertical axis “stacks” the probabilities in the three response categories. For example, with children absent from the household and `hincome` set to its mean, about 30% of women did not work outside the home, about 20% worked part time, and the remaining 50% worked full time. The `xlevels` argument was again used to evaluate `hincome` over a fine grid to get smoother curves in the plot.

Some models produced by the functions `clm()`, `clm2()`, and `clmm()` in the **ordinal** package can be used with the **effects** package. To work with model objects produced by these functions, you must also load the **MASS** package

The **effects** package can produce similar graphs for the more general multinomial logit model, in which the polytomous categorical response has unordered categories (see Fox and Weisberg, 2019a, Sec. 6.7). The details of the model, its parameters, and its assumptions are different from the proportional-odds model and from other ordered-response models, but predictor effect plots for these models are similar.

As an example, we use the BEPS data set in the **carData** package, consisting of about 1,500 observations from the 1997-2001 British Election Panel Study. The response variable, `vote`, is party choice, one of "Liberal Democrat", "Labour", or "Conservative". There are numerous predictors of `vote` in the data set, and we fit the model

```
R> library("nnet")
R> mr1 <- multinom(vote ~ age + gender + economic.cond.national +
+                   economic.cond.household + Blair + Hague + Kennedy +
+                   Europe*political.knowledge, data=BEPS)

# weights: 36 (22 variable)
initial value 1675.383740
iter 10 value 1240.047788
iter 20 value 1163.199642
iter 30 value 1116.519687
final value 1116.519666
converged
```

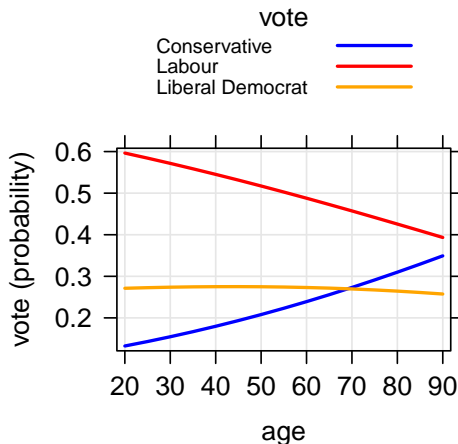
There are nine predictors, seven of which are scales with values between 0 and 5 concerning respondents' attitudes; these predictors enter the model as main effects. The remaining two predictors are scales between 0 and 3 for `political.knowledge` and between 1 and 11 for `Europe` (attitude toward European integration of the UK in the European Union, with high values representing

“Euroscepticism”—i.e., a negative attitude toward Europe); these predictors enter the model with a two-factor interaction.

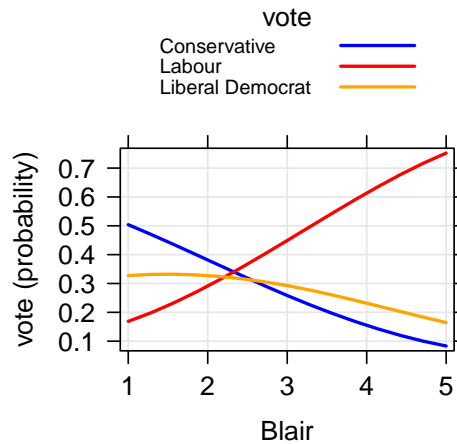
Drawing all nine predictor effect plots simultaneously is not productive is not a good idea because the plots won’t fit reasonably in a single display. We therefore draw only a few of these plots at a time.

```
R> plot(predictorEffects(mr1, ~ age + Blair + Hague + Kennedy),
+       axes=list(grid=TRUE, x=list(rug=FALSE)),
+       lattice=list(key.args=list(columns=1)),
+       lines=list(multiline=TRUE, col=c("blue", "red", "orange")))
```

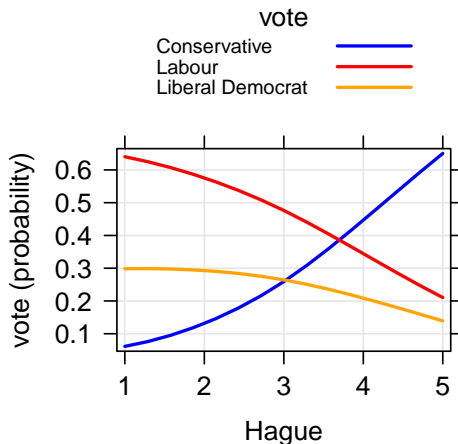
age predictor effect plot



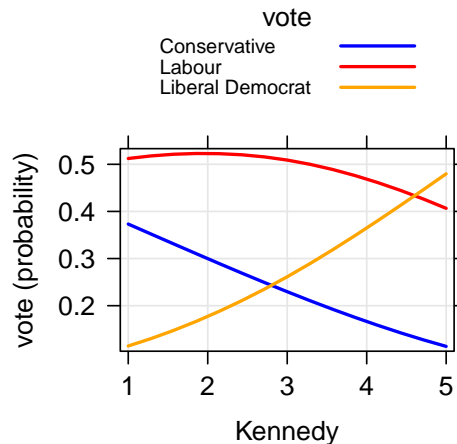
Blair predictor effect plot



Hague predictor effect plot



Kennedy predictor effect plot



We use optional arguments to get a multiline plot, with a grid and no rug plot, and to modify the key. The color specification for the lines represents the traditional colors of the three parties. Interpreting these plots is challenging: The probability of preferring Labour decreases with age, increases with attitude toward the Labour leader Blair, strongly decreases with attitude toward the Conservative leader Hague, and is relatively unaffected by attitude toward the Liberal Democrat leader Kennedy. In general a positive attitude toward a party leader increases the probability of favoring that leader’s party, as one would expect. Of course, the causal direction of these relationships is unclear.

We next turn to the interaction between `Europe` and `political.knowledge`, this time using a stacked area display. We also use the `xlevels` argument to control the number of panels in each effect plot:

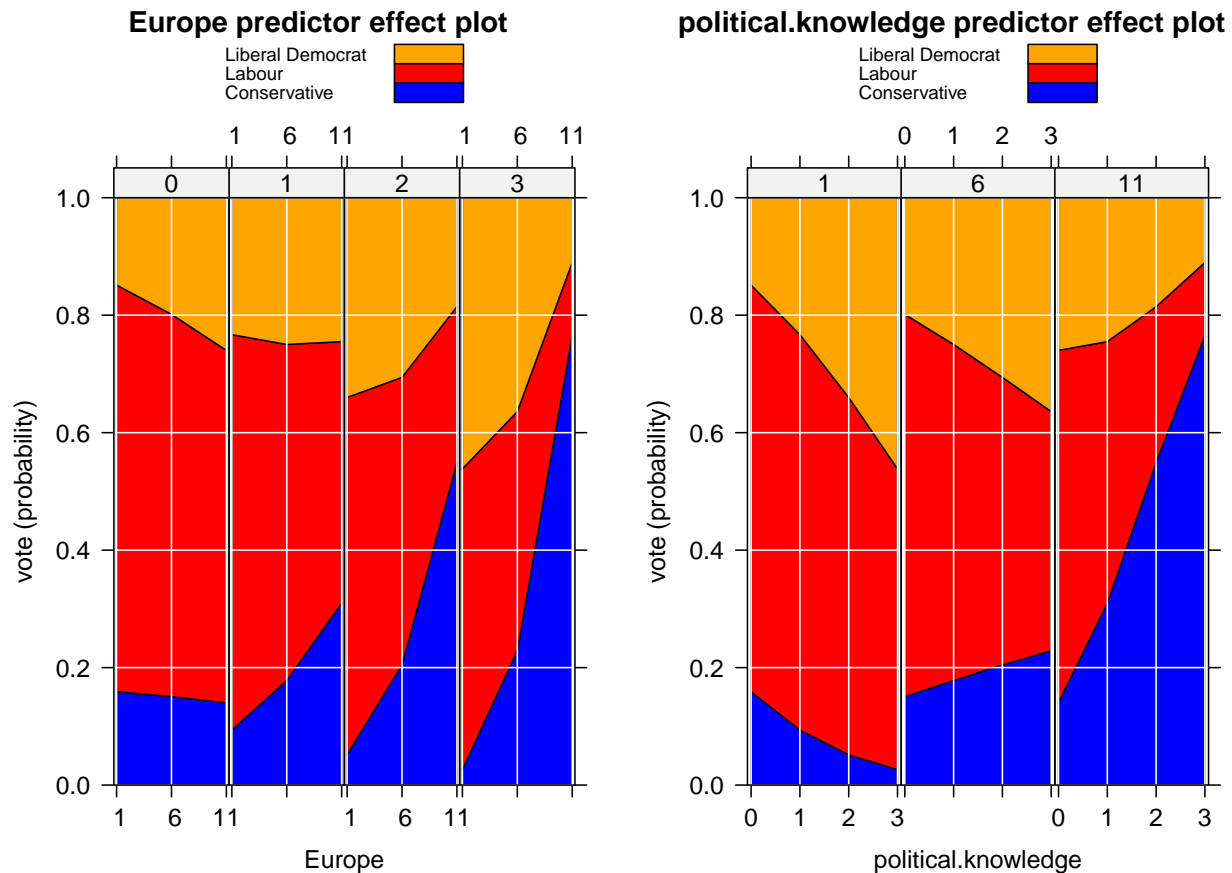
```
R> plot(predictorEffects(mr1, ~ Europe + political.knowledge,
```



```

+               xlevels=list(political.knowledge=0:3, Europe=c(1, 6, 11))),
+   axes=list(grid=TRUE,
+             x=list(rug=FALSE,
+                   political.knowledge=list(ticks=list(at=0:3)),
+                   Europe=list(ticks=list(at=c(1, 6, 11)))),
+             y=list(style="stacked")),
+   lines=list(col=c("blue", "red", "orange")),
+   lattice=list(key.args=list(columns=1),
+               strip=list(factor.names=FALSE)))

```



The `lines` argument is used to specify the colors for the stacked areas representing the parties. Both effect plots are of the same fitted values, in the first graph with `Europe` varying and conditioning on `political.knowledge`, and in the second with `political.knowledge` varying and conditioning on `Europe`. Setting `strip=list(factor.names=FALSE)` suppresses the names of the conditioning predictor in each effect plot; these names are too long for the strips at the tops of the panels. Concentrating on the first graph, we see, for example, that preference for the Conservative Party increases with `Europe` (i.e., with Euroscepticism) for respondents with high political knowledge, but not for those with low political knowledge. More generally, voters with high political knowledge are more likely to align their votes with the positions of the parties (Eurosceptic for the CONservatives, pro-Europe for Labour and the Liberal Democrats) than are voters with low political knowledge.

6 The Lattice Theme for the effects Package

The `effects` package uses the `xypplot()` and `barchart()` functions in the standard `lattice` package (Sarkar, 2008) to draw graphs. The `lattice` package has many options for customizing the appearance

of graphs that are collected into a *lattice theme*. We have created a custom theme for use with the **effects** package that automatically replaced the default theme, *unless the **lattice** package has been previously loaded*. You can load the **effects** package theme directly by the command

```
R> effectsTheme()
```

You can also customize the **effects** package theme; see `help("effectsTheme")`. Finally, because `plot()` methods in the **effects** package return lattice objects, these objects can be edited and manipulated in the normal manner, for example by functions in the **latticeExtra** package ([Sarkar and Andrews, 2016](#)).

7 Tests with Predictor Effects

Not written. This will make connections to Anova and emmeans.

References

- Fox, J. and S. Weisberg (2019a). *An R Companion to Applied Regression* (Third ed.). Sage.
- Fox, J. and S. Weisberg (2019b). Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software*.
- Hawkins, D. M. and S. Weisberg (2017). Combining the Box-Cox power and generalised log transformations to accommodate negative responses in linear and mixed-effects linear models. *South African Statistics Journal* 51, pp. 317–328.
- Sarkar, D. (2008). *Lattice: Multivariate Data Visualization With R*. Springer Science & Business Media.
- Sarkar, D. and F. Andrews (2016). *latticeExtra: Extra Graphical Utilities Based on Lattice*. R package version 0.6-28.