Uncertainty propagation for predictor effects and bias correction in generalized linear (mixed) models

Bicko Cygu, Ben Bolker, Jonathan Dushoff

November 23, 2021

1 Introduction

In many applications, predictions of expected responses, or response probabilities, are often of interest. Predictions provide a great way to summarize what the regression model is telling us and are very useful for interpreting and visualizing model estimates using graphs. For example, in logistic regression models, the coefficient estimates are usually not easy to interpret, and we may want to explore the "effects" of covariates on predicted probabilities, or predictions of responses at various levels of covariates – *predictor effects*. For example, as a way to plan, public health officials may want to know the "effects" of household income, wealth index, etc, on the predicted probability of having improved water services among slum dwellers.
Both simple and generalized linear (mixed) models (GL(M)Ms) can examine very complex relationships, including nonlinear relationships between response and predictors, interactions between predictors (and via splines for example), and nonlinear transformations via link functions due to their flexibility. This flexibility comes at a cost, for example, complex multivariate models may risk misinterpretation, and miscalculation of quantities of interest. Also, coefficient estimates of models involving nonlinear link functions or interactions lose their direct interpretation [Leeper, 2017], meaning that interpretation of derived quantities from these estimates requires some understanding of the specified model.

When visually presented, predictor effects provide a unified and intuitive way of describing relationships from a fitted model, especially complex models involving interaction terms or some kind of transformations on the predictors whose estimates are usually, but not always, a subject to less clarity of interpretation. Further, generating predictor effects together with the associated confidence intervals for regression models has a number of challenges. In particular:

1. choice of representative values of *focal* predictor(s) and appropriate "model center" for *non-focal* predictors especially in multivariate models

2. propagation of uncertainty ([how] can we incorporate uncertainty in nonlinear components of (G)L(M)Ms? Should we exclude variation in non-focal parameters?)

3. bias in the expected mean prediction induced by the nonlinear transformation of the response variable (especially in GL(M)Ms)

Most common way of dealing with the first challenge is taking unique levels of the focal predictor if discrete or taking appropriately sized quantiles (or bins) if continuous, and then calculating the predictions while holding non-focal predictors at their typical values (e.g., averages). This generates – *predictor effect* [Fox and Hong, 2009], *marginal predictions* [Leeper et al., 2017] and *estimated marginal means* [Lenth and Lenth, 2018]. In this article, we refer to this quantity as *predictor effect* since it should, for example, tell us what we would expect the presponse to be at a particular value or level of the predictor. Formerly, predictor effects computes the expected outcome by meaningfully holding the non-focal predictors constant (or averaged in some meaningful way) while varying the focal predictor, with the goal that the outcome expected prediction represents how the model responds to the changes in the focal predictor.
An alternative to averaging the non-focal predictors, is the population-based approach which involves computing the prediction over the population of the non-focal predictors and then averaging across the values of the focal predictor. Dealing with second challenge is discussed in Section 4
When dealing with nonlinear link functions, the correct predictions for example, are even much harder to estimate. One approach is to make predictions on the transformed scale (linear predictor scale), and then back-transform to the original scale. However, the back-transformation may either result in

biased predictions or requires some approximation. In particular, bias in expected mean prediction induced by nonlinear transformation of the response variable can lead inaccurate predictions.

The main purpose of this article is to discuss and implement various approaches for computing predictor effects and provide an alternative method for computing the associated confidence intervals. We further explore approaches for correcting bias in predictions for GL(M)Ms involving nonlinear link functions.

The outline of this article is as follows: Section 2 provides formal definition various quantities of interest, Section 3 provides statistical background of estimation of model predictions, Section 4 describes mathematical and computation implementation of the proposed method for uncertainty propagation, Section 5 describes methods for bias correction in GL(M)M together with the computational implementation, and lastly, in Section 6 we consider some simulation examples.

2 Quantities of interest

Several quantities of interest may be derived from regression models. The first one is the coefficient estimates. Others are *predictors*, *model matrix*, *predicted values* and *marginal effects*. Predictors are model input variables and is associated with one or more *variables* (in polynomials, splines and models with interaction). Model matrix refers to the design matrix whose rows include all combination of variables appearing in the interaction terms, along with the "typical" values of the focal and non-focal predictors.

To illustrate some of the concepts, consider results from, hypothetical simulated example, regression of household size as function of household wealth index and age of household head, and the second model involving these predictors plus the interaction between them.

Table 1: Example of simple linear model output

	Dependent variable:	
	hhsize	
	(1)	(2)
age	0.987^{***} (0.035)	0.996^{***} (0.032)
wealthindex	2.137^{***} (0.035)	1.993^{***} (0.032)
age:wealthindex		0.477^{***} (0.031)
Constant	5.063^{***} (0.036)	5.039^{***} (0.033)
Observations	1,000	1,000
Note:		*p<0.1; **p<0.05; ***p<0.01

In simple linear models with no interaction terms, the default out for the coefficient estimates in Table 1 model (1) is simple and directly interpretable as the expected change in household size for a unit change in age. In particular, a unit change in age is associated with a household that is 0.987 bigger. This is the *unconditional marginal effect* [Leeper, 2017] and it is constant across all the observations and levels of all other predictors. As a result, the interpretation of interaction models differs in an important way from linear-additive regression models. To see this more clearly, compare the marginal effect of x_1 in the following two models:

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_1 \tag{1}$$
$$y_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon_2 \tag{2}$$

The marginal effect of x_1 in Equation 1 is $\frac{\partial y_1}{\partial x_1} = \beta_1$. On the other hand, the marginal effect of x_1 in Equation 2 is given by $\frac{\partial y_2}{\partial x_1} = \beta_1 + \beta_{12} x_2$. In other words, if there are no interactions, the marginal effect of x_1 on y_1 is constant, while, if there are interactions in the model, the marginal effect of a change in x_1 on y_1 depends on the value of the *conditioning* predictor x_2.
Figure 1 shows the comparison of predictor effect and unconditional marginal effect of age on household size, based on model estimates in Table 1. Since the model has no interactions, the relationship between the predicted household size and age is linear, hence the marginal effect of age is the slope ($\frac{\Delta \text{hhsize}}{\Delta \text{age}}$) of the predictor effect line and can be calculated irrespective of the values of wealth index.

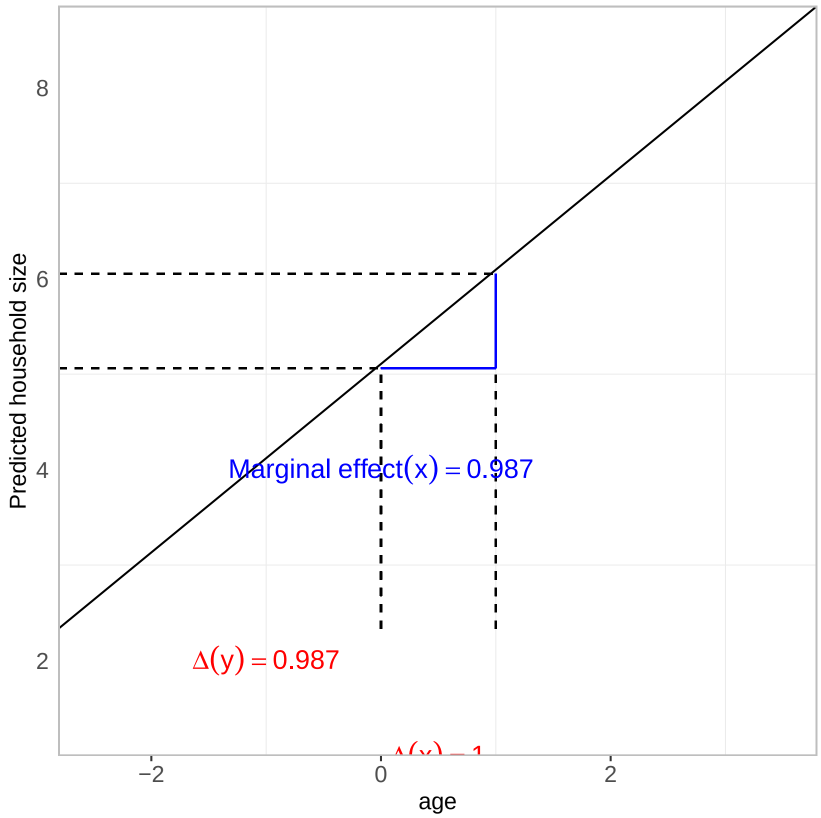

Figure 1: Variable effect plot and marginal effect of age of household head on household size. For a linear model with no interaction, the marginal effect is the slope of the predictor effect line.

Predictor effect, on the other hand, is the expected household size for a particular age, holding wealth index at its mean. In particular, the purpose and goal of a predictor effect seems fairly straightforward; for specified values of (a) focal predictor(s), we want to give a point estimate and confidence intervals for the prediction of the model for a "typical" (= random sample over the multivariate distribution of non-focal parameters) individual with those values of the predictors.
On the other hand, when model specification involves other kinds of terms such as variable interactions, log and power transformations, etc., the coefficient estimates cannot easily and directly communicate the relationship between the outcome and the independent variable of interest because of

the multiple coefficients of the same variable. For instance, in model (2), the interaction between age and wealth index is added, hence the coefficient estimate for age (0.996) can only be regarded as unconditional marginal effect when effect of wealth index (and thus `age:wealthindex`) is zero, which is never the case. In other words, for models involving interactions terms, looking at a single estimated coefficient in isolation and failing to take into account the estimates involving the interaction terms may lead to inaccurate interpretations [Brambor et al., 2006, Leeper, 2017].

3 Statistical background

To get an intuition of how conditioning on the mean values of the non-focal predictors work, suppose we are interested in predictor effects of a particular predictor (hence forth referred as *focal* predictor otherwise *non-focal*), x_f, from the set of predictors. To keep it simple, assume that the model has no interaction terms. Then the idea is to *anchor* the values of non-focal predictors to some particular values. For example fixing the values of *non-focal* predictor(s) at some typical values – typically determined by averaging (for now) in some meaningful way, for example, arithmetic mean for continuous and average over the levels of the factors for categorical non-focal predictors. The easiest way to achieve this is by constructing $\mathbf{X}^\star$ by averaging the columns of non-focal predictors in model matrix $\mathbf{X}$, and together with appropriately chosen values of focal predictor(s).
Consider a simple linear model with linear predictor $\eta = \mathbf{X}\boldsymbol{\beta}$ and let $g(\boldsymbol{\mu}) = \boldsymbol{\eta}$ be an identity link function (in the case of simple linear model), where $\boldsymbol{\mu}$ is the expected value of response variable y. Let $\hat{\boldsymbol{\beta}}$ be the estimate of $\boldsymbol{\beta}$, together with the estimated covariance matrix $\Sigma = V(\hat{\boldsymbol{\beta}})$ of $\hat{\boldsymbol{\beta}}$. Let $\mathbf{X}^*$ be the model matrix, inheriting most of its key properties, for example transformations on predictors and interactions from the model matrix, $\mathbf{X}$. Then the prediction $\hat{\boldsymbol{\eta}}^\star = \mathbf{X}^\star\hat{\boldsymbol{\beta}}$ is the predictor effect for the focal predictor in question.
An alternative, save for later, formulation of predictor effect involves, expressing the linear predictor as the sum of the focal and non-focal predictor linear predictor. In particular, $\eta^\star(x_f, \bar{x}_{\{n\}}) = \beta_f x_f + \sum \beta_{\{n\}}\bar{x}^\star_{\{n\}}$, where $\bar{x}^\star_{\{n\}}$ are the appropriately averaged entries of non-focal predictors and x_f is a vector of values of the focal predictors for a particular observation.

3.1 Higher order interactions

Higher order terms such as interactions, splines, polynomials, etc., can be between the non-focal predictors or focal and non-focal predictor(s). In the former case, we treat the interactions as just another column in the variable space (of the model matrix). In the later case, the non-focal variables in the model matrix are averaged as before. However, for the interacting focal predictors, a combination of each unique levels (or quantiles) are first generated and then the interaction terms are generated by multiplying these combinations. For example, consider the model 3 below:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{23} x_2 x_3 + \epsilon \tag{3}$$

For the first case, if we consider x_1 as the focal predictor, the interaction $(x_2 x_3)$ is between the non-focal predictors, x_2 and x_3, the linear predictor is

$$\eta^\star(x_{1i}, \bar{x}^\star_{\{n\}}) = \beta_0 \mathbf{1} + \beta_1 x_{1i} + \beta_2 \bar{x}_2 + \beta_3 \bar{x}_3 + \beta_{23} \bar{x}_2 \bar{x}_3$$

where x_{1i} are the carefully chosen levels of the focal predictor. On the other hand, for the second case, if we consider x_2 as the focal predictor, then the interaction $(x_2 x_3)$ is between focal and non-focal predictor. In this case, the linear predictor is given by

$$\eta^\star(x_{2ij}, \bar{x}^\star_{\{n\}}) = \beta_0 \mathbf{1} + \beta_1 \bar{x}_1 + \beta_2 x_{2ij} + \beta_3 x_{3j} + \beta_{23} x_{2ij} x_{3j}$$

In general, our formulation, even for more complicated interactions, follow these two basic principles – interaction between non-focal predictors and interaction between focal and non-focal predictors.

4 Uncertainty propagation

What about the confidence intervals (CI)? The limits of the confidence intervals are points, not mean values. In principle, every observation/value of focal predictors has a different CI. The traditional way to compute variances for predictions is $\sigma^2 = \mathrm{Diag}(\mathbf{X}^\star \Sigma \mathbf{X}^{\star\top})$ [Lenth and Lenth, 2018, Fox and Hong, 2009], so that the confidence intervals are $\eta \pm q\sigma$, where q is an appropriate quantile of Normal or t distribution. This approach incorporates all the uncertainties – including the uncertainty due to non-focal predictors. But what if we are only interested in the uncertainty as a result of the focal predictor, so that the confidence intervals are $\eta \pm q\sigma_f$, i.e., *isolated* confidence intervals?

Currently, commonly used R packages for constructing predictions do not exclude the uncertainties resulting from the non-focal predictors when computing the CIs. A non-trivial way to exclude uncertainties associated with non-focal predictors in some of these packages is to provide a user defined variance-covariance matrix with the covariances of non-focal terms set to 0 – *zeroing-out* variance-covariance matrix. This only works when the input predictors are *centered* prior to model fitting, in case of numerical predictors, and even much complicated when the predictors are categorical. We first describe this variance-covariance based approach and then discuss our proposed method which is based on *model centering* and does not require input predictors to be scaled prior to model fitting.

4.1 Variance-covariance

The computation of $\hat{\eta}^\star$ remains the same as described above. However, to compute σ, Σ is modified by *zeroing-out* (the variance-covariance of all non-focal predictors are set to zero) variances of non-focal terms. Although this is the simplest approach, it requires centering of continuous, i.e., $x_c = x - \bar{x}$, predictors prior to model fitting and proper way to average categorical predictors.

4.2 Centered model matrix

Consider centered model matrix $\mathbf{X}_c^\star = \{\mathbf{X}_f^\star, \mathbf{X}_{\{n\}}^\star - \bar{\mathbf{X}}_{\{n\}}^\star\}$. It follows that the non-focal terms in $\mathbf{X}_c^\star$ are all zero in simple models without interactions but are isolated (narrower) around the model center. Consequently the uncertainty due to non-focal predictors are isolated in the computation of $\sigma^2 = \mathrm{Diag}(\mathbf{X}_c^\star \Sigma \mathbf{X}_c^{\star\top})$. In addition, the computation of $\mathbf{X}_c^\star$ impacts only on the intercepts and the non-focal terms, i.e., the slopes and variance of the focal predictors are not affected. This means that we can still generate isolated CIs without necessarily centering the predictors prior to model fitting.

5 Bias correction

In many applications, it usually important to report the estimates that reflect the expected values of the untransformed response. However, when dealing with nonlinear link functions, it is even harder to generate correct predictions that reflect the untransformed response due to the bias in the expected mean induced by the nonlinear transformation of the response variable. In such cases, bias correction is needed when back-transforming the predictions to the original scales. Most common approach for bias-adjustment is second-order Taylor approximation [Lenth and Lenth, 2018, Duursma and Robinson, 2003]. Here, we describe and implement a different approach, *population-based* approach for bias correction.

5.1 Population-based approach for bias correction

The most precise (although not necessarily accurate!) way to predict is to condition on a value F of the focal predictor and make predictions for all observations (members of the population) for which $x_f = F$ (or in a small range around F ...). A key point is that the nonlinear transformation involved in these computations is always *one-dimensional*; all of the multivariate computations required are at the stage of collapsing the multidimensional set of predictors for some subset of the population (e.g. all individuals with $x_f = F$) to a one-dimensional distribution of $\eta^\star$.
Once we have got our vector of $\eta^\star$ (which is essentially a set of samples from the distribution over $\eta^\star(x_f, x_{\{n\}})$ for the conditional set), we may want a mean and confidence intervals on the mean on the response (data) scale, i.e. after back-transforming. In other words we can use the observations themselves then we just compute the individual values of $g^{-1}(\eta^\star)$ and compute their mean.
In population-based approach, $\eta^\star$ is constructed as a function of properly constituted levels of focal predictors, x_f, and entire population of the non-focal predictors, $x_{\{n\}}$, as opposed to averaged non-focal predictors, $\bar{x}_{\{n\}}$, described in the previous section. More specifically:

- compute linear predictor associated with the non-focal predictors, $\eta_{\{n\}} = \sum \beta_{\{n\}} x_{\{n\}}$
- compute linear predictor associated with the focal predictors, $\eta_{jf} = \sum \beta_f x_{jf}$

- for every value of the focal predictor, η_{jf}:
  - $\eta_j^\star = \eta_{jf} + \eta_{\{n\}}$
  - $\hat{y}_j = \text{mean } g^{-1}\left(\eta_j^\star\right)$

We make similar adjustments to compute the variances of the predictions at every level of the focal predictor:

$$\sigma_{jf}^2 = \text{Diag}(\mathbf{X}_{jc}^\star \Sigma \mathbf{X}_{jc}^{\star\top}) \tag{4}$$

where $\mathbf{X}_{jc}^\star = \{\mathbf{X}_{jf}^\star, \mathbf{X}_{\{n\}}^\star - \bar{\mathbf{X}}_{\{n\}}^\star\}$ and

$$\text{CI}_j = \text{mean } \eta_j^\star \pm q\sigma_{jf} \tag{5}$$

6 Simulation examples

In this section, we illustrate predictor effects (together with the associated CI) in the context of linear models involving continuous and categorical predictors with or without interactions. We compare the expected to the known simulation “truth”.

6.1 Continuous predictors

Consider simulation example described in Section 2, with slight modifications to incorporate interaction between predictors. The first case, is the model with no interaction:

$$\begin{aligned}
\text{hh size} &= \beta_0 + \beta_{\text{A}}\text{Age}_i + \beta_{\text{W}}\text{Wealthindex}_i + \epsilon_i \\
\text{Age}_i &\sim \text{Normal}(0.2, 1) \\
\text{Wealthindex}_i &\sim \text{Normal}(0, 1) \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
\beta_0 &= 5 \\
\beta_{\text{A}} &= 0.1 \\
\beta_{\text{W}} &= 2 \\
i &= 1, \cdots, 100
\end{aligned} \tag{6}$$

Secondly, we add interaction between the non-focal predictors (`wealth index` and `household expenditure`) but not with the focal predictor `age`:

$$
\begin{aligned}
\text{hh size} &= \beta_0 + \beta_{\text{A}}\text{Age}_i + \beta_{\text{W}}\text{Wealthindex}_i + \beta_{\text{E}}\text{Expenditure}_i \\
&\quad + \beta_{\text{WE}}\text{Wealthindex}_i * \text{Expenditure}_i + \epsilon_i \\
\text{Age}_i &\sim \text{Normal}(0.2, 1) \\
\text{Wealthindex}_i &\sim \text{Normal}(0, 1) \\
\text{Expenditure}_i &\sim \text{Normal}(0, 1) \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
\beta_0 &= 5 \\
\beta_{\text{A}} &= 0.1 \\
\beta_{\text{W}} &= 2 \\
\beta_{\text{E}} &= 1.5 \\
\beta_{\text{WE}} &= 1 \\
i &= 1, \cdots, 100
\end{aligned}
\tag{7}
$$

And lastly, the model with interaction terms between the focal predictor (`age`) and one of the non-focal predictor (`wealth index`):

$$
\begin{aligned}
\text{hh size} &= \beta_0 + \beta_{\text{A}}\text{Age}_i + \beta_{\text{W}}\text{Wealthindex}_i + \beta_{\text{E}}\text{Expenditure}_i \\
&\quad + \beta_{\text{AW}}\text{Age}_i * \text{Wealthindex}_i + \epsilon_i \\
\text{Age}_i &\sim \text{Normal}(0.2, 1) \\
\text{Wealthindex}_i &\sim \text{Normal}(0, 1) \\
\text{Expenditure}_i &\sim \text{Normal}(0, 1) \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
\beta_0 &= 5 \\
\beta_{\text{A}} &= 0.1 \\
\beta_{\text{W}} &= 2 \\
\beta_{\text{E}} &= 1.5 \\
\beta_{\text{AW}} &= 1 \\
i &= 1, \cdots, 100
\end{aligned}
\tag{8}
$$

We first fitted all the three models and in each case, compared three different predictor effects together with the corresponding CIs – 1) uncertainty due to non-focal predictors included (everything), 2) uncertainty due to non-focal predictors removed using variance-covariance matrix (isolated (vcov)), and 3) uncertainty due to non-focal predictors removed using centered model matrix (isolated (mm)). The results are shown in Figure 2.

Figure 2a represents predictor effects for the model with no interaction, Figure 2b shows predictor effects for the model with interaction between non-focal predictors, and Figure 2c shows predictor effects for the model with interaction between focal and non-focal predictors. For perfect predictions, we expect the horizontal yellow and grey dotted lines to overlay each other and they should intersect with the black solid line and the vertical dotted grey line at the same point (or very close), i.e., model center. In other words, at the model center (mean of the focal predictor), generally, we expect the model to predict the average of the response variable and narrower CIs. If we properly anchor the model at its center (or any other appropriate value, assuming no interactions), we would expect the variance at the anchor to be zero, i.e., the CIs bands crosses at that point (see isolated (mm) and isolated (vcov) in Figure 2). Although the overall predictor effects trend lines are different in each of the simulation, the expected household size for a household head with average age is not only the same in all the three simulations but also very close to the "truth".

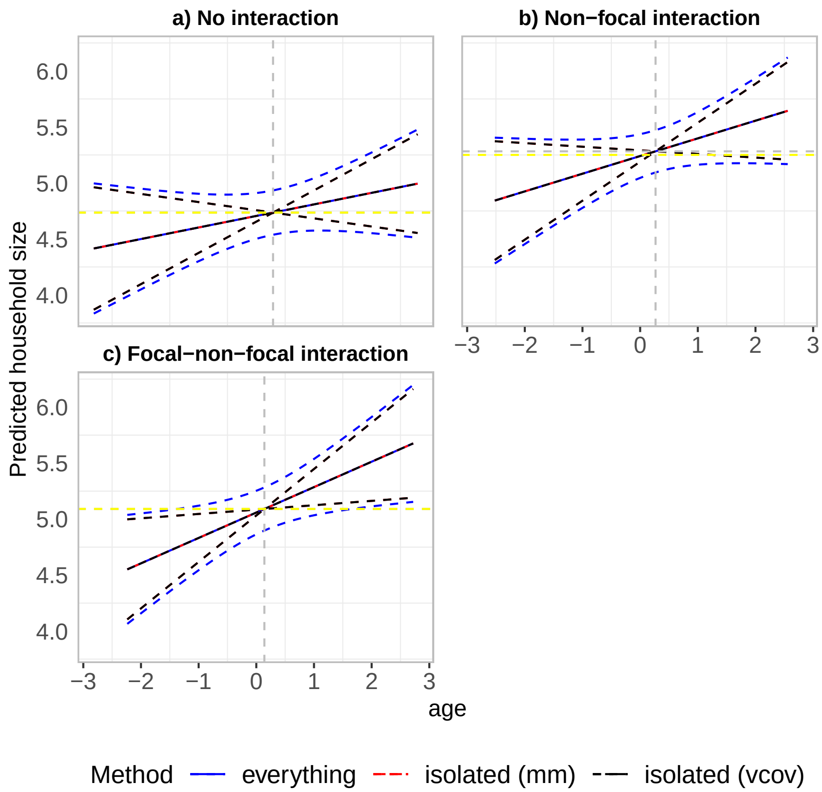

Figure 2: A comparison of predictor effects and their corresponding CIs. Figure a) shows model with no interaction, Figure b) shows model with non-focal predictors interacting with each other, and Figure c) shows model with focal and non-focal predictors interacting. The dotted yellow and grey horizontal lines are the expected and observed/true means, respectively. The vertical grey line is the mean of the focal predictor (model center). The dotted black, red (non-focal uncertainty excluded) and the solid blue lines (all uncertainty included) are the predictor effects together with their corresponding CIs. For properly generated predictor effects or simple models with no interactions, we expect the yellow and the grey horizontal dotted lines to overlay each other and intersect with the black, blue, red solid trend lines, and the vertical dotted grey line at the same point (model center).

6.2 Polynomial interaction

We consider more complex interactions in which the focal predictor is modelled as a cubic polynomial. Specifically:

$$
\begin{aligned}
\text{hh size} &= \beta_0 + \beta_{\mathrm{A}_1}\mathrm{Age}_i + \beta_{\mathrm{A}_2}\mathrm{Age}_i^2 + \beta_{\mathrm{A}_3}\mathrm{Age}_i^3 + \beta_{\mathrm{W}}\mathrm{Wealthindex}_i + \epsilon_i \\
\mathrm{Age}_i &\sim \mathrm{Normal}(0,1) \\
\mathrm{Wealthindex}_i &\sim \mathrm{Normal}(0,1) \\
\epsilon_i &\sim \mathrm{Normal}(0,10) \\
\beta_0 &= 20 \\
\beta_{\mathrm{A}_1} &= 0.1 \\
\beta_{\mathrm{A}_2} &= 0.8 \\
\beta_{\mathrm{A}_3} &= 0.3 \\
\beta_{\mathrm{W}} &= 0.8 \\
i &= 1, \cdots, 100
\end{aligned} \tag{9}
$$

In Figure 3 we compare the non-isolated (everything) and the isolated predictor effects. In the first case (Figure 3a), the focal predictor, `age`, has a higher order interaction in the form of cubic polynomial while in the second case (Figure 3b), the focal predictor, `wealthindex`, is simple but the non-focal predictor is a cubic polynomial. In Figure 3a, the average prediction and the "truth" are very similar but in this case, cubic polynomial, the model center is not a point in the predictor space. On the hand, in Figure 3b, the average prediction is higher than the observed marginal, "truth" but the model center is a point in the predictor space.

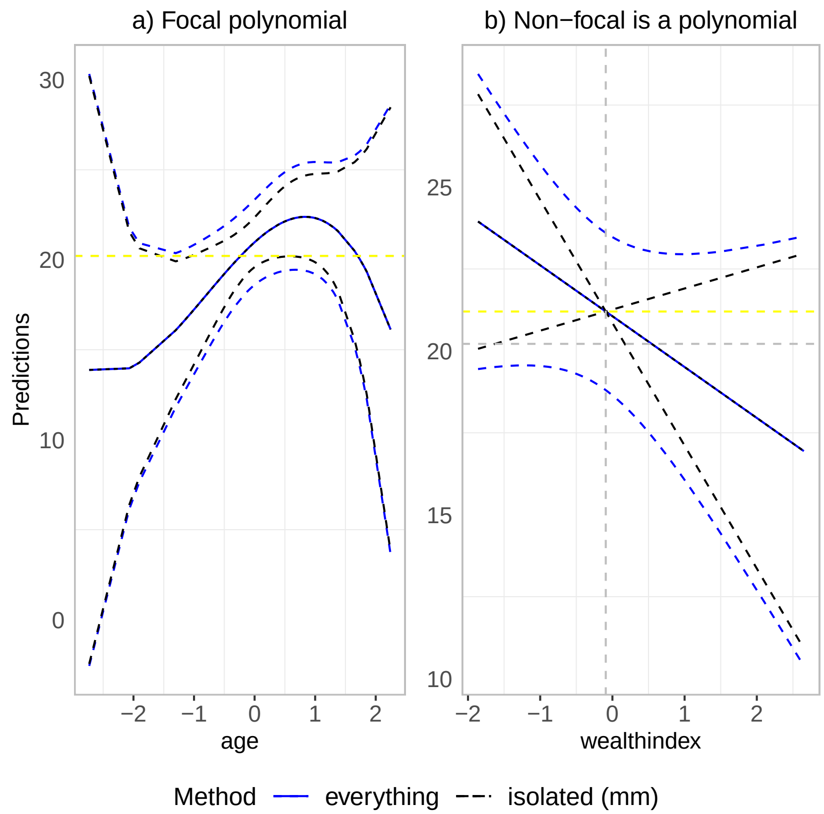

Figure 3: A comparison of predictor effects and their corresponding CIs. In Figure a) the focal predictor has a higher cubic polynomial interaction while in Figure b) the focal predictor has no interaction but the non-focal predictor is a cubic polynomial. The dotted yellow and grey horizontal lines are the expected and observed/true means, respectively. The vertical grey line is the mean of the focal predictor (model center). The dotted black (non-focal uncertainty excluded) and blue lines (all uncertainty included) are the predictor effects together with their corresponding CIs.

Even in simple interaction model in Figure 2b and Figure 3b, we start to observe bias in the predicted and observed means. To illustrate the bias correction, we regenerate the predictor effects in Figure 3b but incorporate bias adjustment.

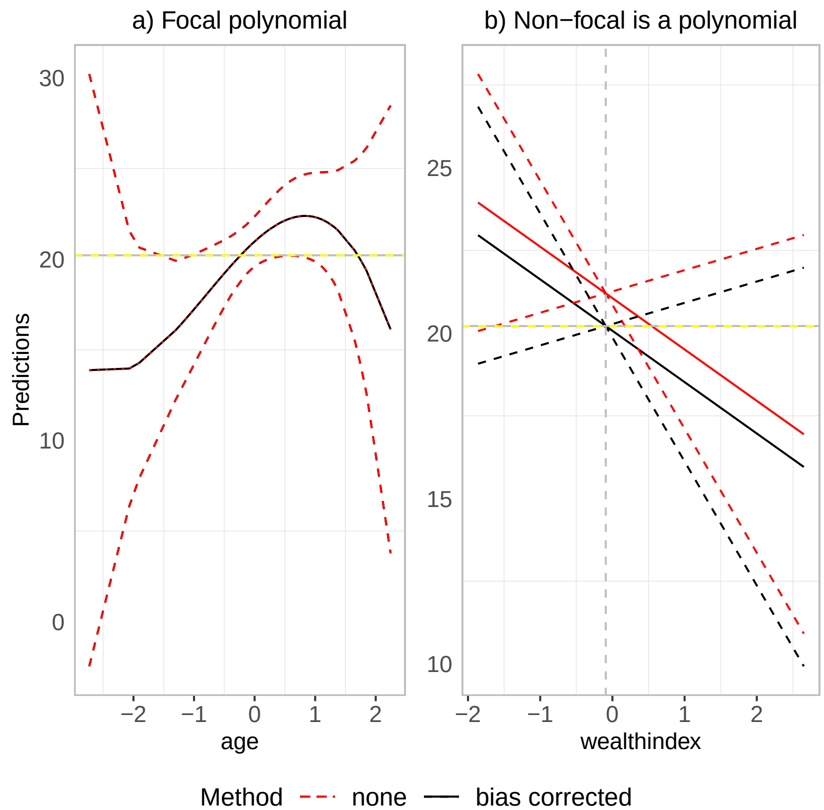

Figure 4: A comparison of bias corrected and uncorrected predictor effects. The red solid and dotted lines are not bias corrected (none), the black ones are corrected for bias, the horizontal grey and yellow lines are the observed and predicted average household size, respectively, while the vertical dotted line is the average wealth index. In the absence of bias, the two approaches give similar estimates (Figure a). On the other hand, in Figure b, without correcting for the bias, the predictions are slightly higher than the corrected ones which matches the observed relatively well.

6.3 Generalized linear mixed effect model

Here we consider generalized linear models with random effects components. In the simplest case, we start with a single grouping factor as the random effect component, i.e., random intercept model. This way, we can investigate

the contribution of random effects overall predicted probability.
Suppose, in the simulation described in Section 3, the observations are recorded at least once per household and we are interested in the status improved water services (improved or unimproved). In particular, let H be the number of households indexed by the grouping factor, and h[i] be the household of the ith observation such that

$$
\begin{aligned}
\text{logit}(\text{status} = 1) &= \eta \\
\eta &= \beta_0 + \alpha_{0,\text{h}[i]} + \beta_{\text{A}}\text{Age}_i + \epsilon_i \\
\alpha_{0,\text{h}} &\sim \text{Normal}(0, 5), \quad h = 1, \cdots, 10 \\
\text{Age}_i &\sim \text{Normal}(0, 3) \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
\beta_0 &= 5 \\
\beta_{\text{A}} &= 0.8 \\
i &= 1, \cdots, 500
\end{aligned}
\tag{10}
$$

We compare predictor effects generated without any bias correction to those corrected for the bias, and then compare the average predicted probability to the observed. The results are shown in Figure 5. The uncorrected approach overestimated the predicted probabilities as opposed to the bias corrected estimates which closely match the observations and the observed average probability.

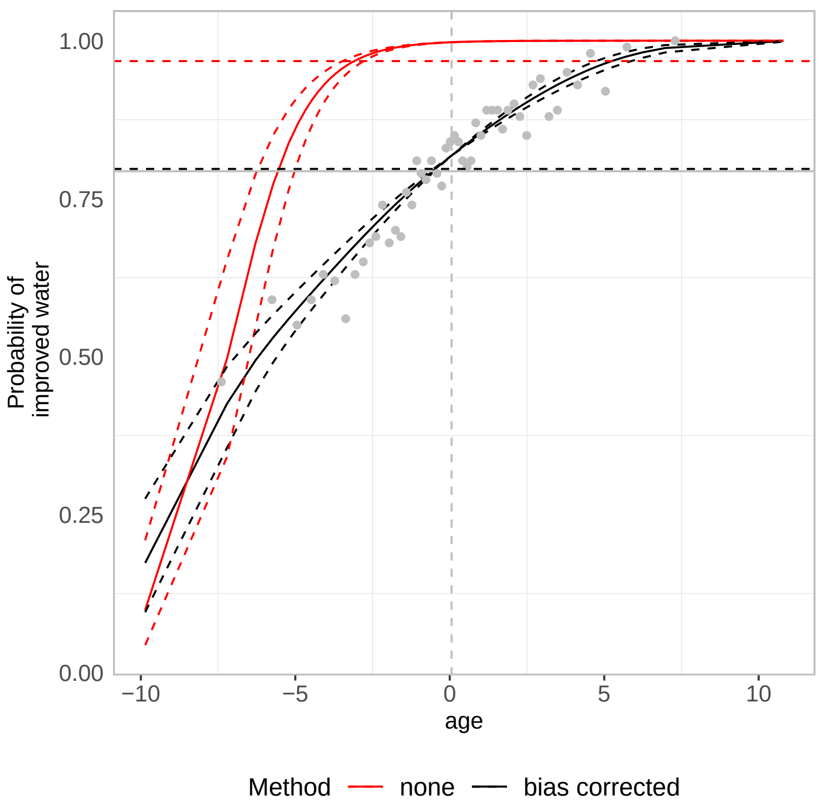

Figure 5: A compare of bias corrected and uncorrected predictor effects. The red lines are the bias uncorrected estimates while the black ones are the bias corrected. The horizontal red and black lines are average predicted probability of improved water. The horizontal grey (overlaid by the black) dotted line is the observed average probability of improved water. The vertical dotted grey line is the average age while the grey points are the binned observations.

References

Thomas Brambor, William Roberts Clark, and Matt Golder. Understanding interaction models: Improving empirical analyses. *Political analysis*, 14 (1):63–82, 2006.

RA Duursma and AP Robinson. Bias in the mean tree model as a consequence of jensen's inequality. *Forest Ecology and Management*, 186(1-3): 373–380, 2003.

John Fox and Jangman Hong. Effect displays in r for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software*, 32(1):1–24, 2009.

Thomas J Leeper. Interpreting regression results using average marginal effects with r's margins. *Reference manual*, 32, 2017.

Thomas J Leeper, Jeffrey Arnold, Vincent Arel-Bundock, and Maintainer Thomas J Leeper. Package 'margins'. *accessed December*, 5:2019, 2017.

Russell Lenth and Maintainer Russell Lenth. Package 'lsmeans'. *The American Statistician*, 34(4):216–221, 2018.