Weighted Effect Coding for Observational Data with wec

by Rense Nieuwenhuis, Manfred te Grotenhuis, Ben Pelzer

Abstract Weighted effect coding refers to a specific coding matrix to include factor variables in generalised linear regression models. With weighted effect coding, the effect for each category represents the deviation of that category from the weighted mean (which corresponds to the sample mean). This technique has particularly attractive properties when analysing observational data, that commonly are unbalanced. The **wec** package is introduced, that provides functions to apply weighted effect coding to factor variables, and to interactions between (a.) a factor variable and a continuous variable and between (b.) two factor variables.

Introduction

Weighted effect coding is a type of dummy coding to facilitate the inclusion of categorical variables in generalised linear models (GLM). The resulting estimates for each category represent the deviation from the weighted mean. The weighted mean equals the arithmetic mean or sample mean, that is the sum of all scores divided by the number of observations As we will show, weighted effect coding has important advantages over traditional effect coding if unbalanced data are used (i.e. with categories holding different numbers of obwervations), which is common in the analysis of observational data. We describe weighted effect coding for categorical variables and their interactions with other variables. Basic weighted effect coding was first described in 1972 (Sweeney and Ulveling, 1972) and recently updated to include weighted effect interactions between categorical variables (Grotenhuis et al., 2017b,a). In this paper we develop the interaction between weighted effect coded categorical variables and continuous variables. All software is available in the **wec** package.

Treatment, effect, and weighted effect coding

Weighted effect coding is one of various ways to include categorical (i.e., nominal and ordinal) variables in generalised linear models. As this type of variables is not continuous, so-called dummy variables have to be created first which represent the categories of the categorical variable. In R, categorical variables are handled by factors, to which different contrasts can be assigned. For unordered factors, the default is dummy or treatment coding. With treatment coding, each category in the factor variable is tested against a preselected reference category. This coding can be specified with `contr.treatment`. Several alternatives are available, including orthogonal polynomials (default for ordered factors and set with `contr.poly`), helmert coding to compare each category to the mean of all subsequent categories (`contr.helmert`), and effect coding (`contr.sum`).

In effect coding (also known as deviation contrast or ANOVA coding), parameters represent the deviation of each category from the grand mean across all categories (i.e., the sum of arithmetic means in all categories divided by the number of categories). To achieve this, the sum of all parameters is constrained to 0. This implies that the possibly different numbers of observations in categories is not taken into account. In *weighted* effect coding, the parameters represent the deviation of each category from the *sample mean*, corresponding to a constraint in which the *weighted* sum of all parameters is equal to zero. The weights are equal to the number of observations per category.

The differences between treatment coding, effect coding, and weighted effect coding are illustrated in Figure 1, showing the mean wage score for 4 race categories in the USA. The grey circles represent the numbers of observations per category, with whites being the largest category. In treatment coding, the parameters for the Black, Hispanic and Asian populations reflect the mean wage differences from the mean wage in the white population that serves here as the reference category. The dotted double-headed arrow in Figure 1 represents the effect for Blacks based on treatment coding, with whites as the reference category. In effect coding, the reference is none of the four racial categories, but the grand mean. This mean is the sum of all four (arithmetic) mean wages divided by 4, and amounts to 49,762 and is shown as the dashed horizontal line in Figure 1. The effect for Blacks then is the difference between their mean wage score (37,664) and the grand mean, represented by the dashed double-headed arrow and amounts to (37,664 - 49,762 =) -12,096. Weighted effect coding accounts for the number of observations per category, and thus weighs the mean wages of all categories first by the different number of observations per category. Because whites outnumber all other other categories the weighted (sample) mean (= 52,320) is much larger than the unweighted (grand) mean, and is represented by the horizontal continuous line in Figure 1. As a consequence, the effect for
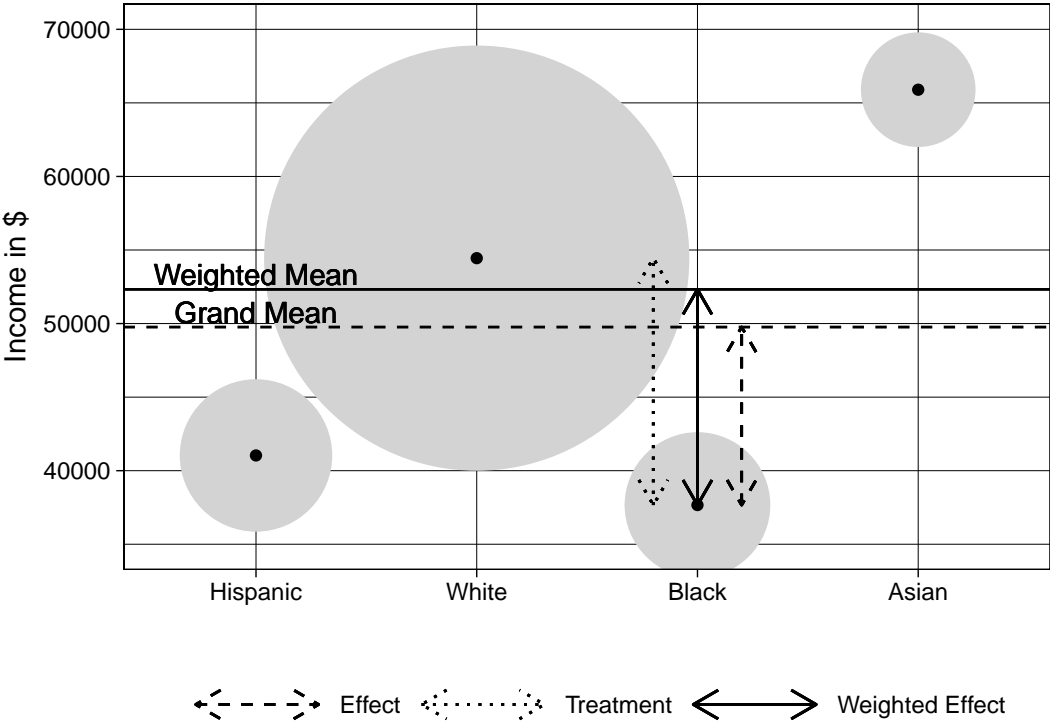
Figure 1: Illustration of treatment coding, effect coding, and weighted effect coding

Blacks now is much stronger ($37,664 - 52,320 = -14,654$) as represented by the (vertical) continuous double-headed arrow.

If the data are balanced, meaning that all categories have the same number of observations, the results for effect coding and weighted effect coding are identical. With unbalanced data, such as typically is the case in observational studies, weighted effect coding offers a number of interesting features that are quite different from those obtained by unweighted effect coding. First of all, in observational data the sample mean provides a natural point of reference. Secondly, the results of weighted effect coding are not sensitive to decisions on how observations were assigned to categories: when categories are split or combined, the grand mean is likely to shift as it depends on the means within categories. In weighted effect coding the sample mean of course remains unchanged. Therefore, combining or splitting *other* categories does not the change the effects of categories that were not combined or split. Finally, weighted effect coding allows for an interpretation that is complementary to treatment coding, and seems particularly relevant when comparing datasets from different populations (e.g., from different countries, or time-periods): the effects represent how deviant a specific category is from the sample mean, while accounting for differences in the composition between populations. Looking at Figure 1, this would allow for the finding that the Black population would have become more deviant over time in a situation where the whites grew in numbers (thus shifting the weighted mean upwards) while the wage gap between Blacks and whites remained constant (the dotted line, as would be estimated with treatment coding).

The coding matrix for weighted effect coding is shown in Table 1. In effect coding, the columns of the coding matrix would all have summed to 0. This can be seen in the first example of the next section. The coding matrix for *weighted* effect coding is based on the restriction that the columns multiplied by the proportions of the respective categories sum to 0. In other words, if the values in each cell of the

	Hispanic	Black	Asian
Hispanic	1	0	0
Black	0	1	0
Asian	0	0	1
White	$-(n_{hispanic}/n_{white})$	$-(n_{black}/n_{white})$	$-(n_{asian}/n_{white})$

Table 1: Coding Matrix Weighted Effect Coding

coding matrix in Table 1 are weighted by the relative number (or proportion) in each category, each column sums to 0.

Examples

This article introduces the **wec** package and provides functions to obtain coding matrices based on weighted effect coding. The examples in this article are based on the PUMS data.frame, which has data on wages, education, and race in the United States in 2013. It is a subset of 10,000 randomly sampled observations, all aged 25 or over and with a wage larger than zero, originating from the PUMS 2013 dataset (Census, 2013). Because the calculation of weighted effect coded variables involves numbers of observations, it is important to first remove any relevant missing values (i.e., list-wise deletion), before defining the weighted effect coded variables.

```
> library(wec)
> data(PUMS)
```

We first demonstrate the use of standard effect coding, which is built into base R, to estimate the effects of race on wages. To ensure that the original race variable remains unaltered, we create a new variable 'race.effect'. This is a factor variable with 4 categories ('Hispanic', 'Black', 'Asian', and 'White'). Effect coding is applied using the contr.sum() function. 'White' is selected as the omitted category by default. Then, we use this new variable in a simple, OLS regression model. This is shown below:

```
> PUMS$race.effect <- factor(PUMS$race)
> contrasts(PUMS$race.effect) <- contr.sum(4)
> contrasts(PUMS$race.effect)

         [,1] [,2] [,3]
Hispanic    1    0    0
Black       0    1    0
Asian       0    0    1
White      -1   -1   -1

> m.effect <- lm(wage ~ race.effect, data=PUMS)
> summary(m.effect)$coefficients
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    49762        954    52.2  0.0e+00
race.effect1   -8724       1649    -5.3  1.3e-07
race.effect2  -12096       1702    -7.1  1.3e-12
race.effect3   16135       2042     7.9  3.0e-15
```

The results of regressing wages on the effect coded race variable (only the fixed effects are shown above) indicate that the grand mean of wages is 49,762. In Figure 1 this grand mean was shown as the horizontal, dashed line. This is the grand (unweighted) mean of the average (arithmetic) wages among Hispanics, Blacks, Asians, and white Americans. The mean wage among Blacks ('race.effect2', refer to the coding matrix to see which category received which label) is, on average, 12,096 dollar lower than this grand mean. This was shown as the dashed double-headed arrow in Figure 1. The wages of Asians ('race.effect3'), on the other hand, are on average 16,135 dollar higher than the grand mean.

We already saw in Figure 1 that not only the average wages vary across races, but also that the number of Hispanics, Blacks, Asians, and whites are substantially different. As these observational data are so unbalanced, the grand mean is not necessarily the most appropriate point of reference. Instead, the sample (arithmetic) mean may be preferred as a point of reference. To compare and test the deviations of all four mean wages from the sample mean, weighted effect coding has to be used:

```
> PUMS$race.wec <- factor(PUMS$race)
> contrasts(PUMS$race.wec) <- contr.wec(PUMS$race.wec, "White")
> contrasts(PUMS$race.wec)

         Hispanic Black  Asian
Hispanic     1.00  0.00  0.000
Black        0.00  1.00  0.000
Asian        0.00  0.00  1.000
White       -0.12 -0.11 -0.069

> m.wec <- lm(wage ~ race.wec, data=PUMS)
```

```
> summary(m.wec)$coefficients
               Estimate Std. Error t value Pr(>|t|)
(Intercept)       52320        587    89.1  0.0e+00
race.wecHispanic -11282       1810    -6.2  4.8e-10
race.wecBlack    -14654       1905    -7.7  1.6e-14
race.wecAsian     13577       2484     5.5  4.7e-08
```

The example above again creates a new variable ('race.wec') and uses the new contr.wec() function to assign a *weighted* effect coding matrix. Unlike many other functions for contrasts in R, contr.wec() requires that not only the omitted category is specified, but also the specification of the factor variable for which the coding matrix is computed. The reason for this is that, as seen in Table 1, to calculate a weighted effect coded matrix, information on the number of observations within each category is required. The coding matrix now shows a (negative) weight (which is the ratio between the number of observations in category x and the omitted category 'Whites') for the omitted category, which was -1 in the case of effect coding.

In the regression analysis, the intercept now represents the sample mean and all other effects represent deviations from that sample mean. This corresponds to the continuous double-headed arrow and line in Figure 1. For instance, Blacks earn on average 14,654 dollars less compared to the sample mean. To be able to test how much whites' mean wage differs from the sample mean, the omitted category must be changed and subsequently a new variable is to be used in an updated regression analysis:

```
> PUMS$race.wec.b <- PUMS$race.wec
> contrasts(PUMS$race.wec.b) <- contr.wec(PUMS$race.wec, "Black")
> m.wec.b <- lm(wage ~ race.wec.b, data=PUMS)
> summary(m.wec.b)$coefficients

                  Estimate Std. Error t value Pr(>|t|)
(Intercept)          52320        587    89.1  0.0e+00
race.wec.bHispanic  -11282       1810    -6.2  4.8e-10
race.wec.bAsian      13577       2484     5.5  4.7e-08
race.wec.bWhite       2128        325     6.5  6.5e-11
```

Here, the omitted category was changed to Blacks. Note that the intercept as well as the estimates for Hispanics and for Asians did not change. This is unlike treatment coding, where each estimate represents the deviation from the omitted category (in treatment coding: the reference category). The new estimate shows that whites earn 2,128 dollar more than the mean wage in the sample. In the remainder of this article we use 'White' as the omitted category by default, but in all analyses the omitted category can be changed.

Next, we control the results for respondents' level of education using a continuous variable (which is mean centred to keep the intercept at 52,320).

```
> m.wec2 <- lm(wage ~ race.wec + education.int, data=PUMS)
> summary(m.wec2)$coefficients
               Estimate Std. Error t value Pr(>|t|)
(Intercept)       52320        560    93.4  0.0e+00
race.wecHispanic  -4955       1738    -2.9  4.4e-03
race.wecBlack    -11276       1820    -6.2  6.0e-10
race.wecAsian      5151       2385     2.2  3.1e-02
education.int      9048        287    31.5  7.9e-208
```

The results show that one additional point of education is associated with an increase in wages of 9,048. This represents the average increase of wages due to education in the sample while controlling for race. The estimates for the categories of race again represent the deviation from the sample mean *controlled* for education. When more control variables are added, the weighted effect coded estimates still represent the deviation from the sample mean, but now controlled for all other variables as well. Comparing these estimates to those from the model without the control for education suggests that educational differences partially account for racial wage differences. In the next section, we discuss how weighted effect coded factor variables can be combined with interactions, to test whether the wage returns of educational attainment vary across race.

Interactions

Weighted effect coding can also be used in generalised linear models with interaction effects. The weighted effect coded interactions represent the additional effects over and above the main effects

obtained from the model without these interactions. This was recently shown for an interaction between two weighted effect coded categorical variables (Grotenhuis et al., 2017a). In this paper we address the novel interaction between weighted effect coded categorical variables and a continuous variable. In the previous section a positive effect (9,048) of education on wage was found. The question is whether this effect is equally strong for all four racial categories.

With treatment coding, an interaction would represent how much the effect of (for instance) education for one category differs from the educational effect in another category that was chosen as reference. With effect coding, the interaction terms represent how much the effect of (for instance) education for a specific category differs from the unweighted main effect (which here happens to be 8,405). Because the data are unbalanced, weighted effect coding is considered here an appropriate parameterisation. In weighted effect coded interactions the point of reference is the main effect in he same model but without the interactions.

In our case this educational main effect on wage is 9,048 (see Figure 2), which we calculated in the example above. Let's assume we already know, as will be confirmed in later examples, that the estimates for the effect of education on wages among whites is 9,461, among Hispanics 5,782, among Asians 12,623, and finally among Blacks the effect is 5,755. The weighted effect coded interactions then are, respectively, $9,461 - 9,048 = 413$; $5,782 - 9,048 = -3,266$; $12,623 - 9,048 = 3,575$; and $5,755 - 9,048 = -3,293$. These estimates represent how much the education effect for each group differs from the main effect of education in the sample.

With weighted effect coded interactions, one can obtain these estimates simultaneously with the mean effect of education. To do so, a coding matrix has to be calculated. This coding matrix is based on the restriction that if the above-mentioned effects are multiplied by the sum of squares of education within each category, the sum of these multiplications is zero. This is the weighted effect coded restriction for interactions.

The sum of squares (SS) of the continuous variable x (education) for level j of the categorical variable (race) is calculated as:

$$SS_j = \sum_{i=1}^{I} \left(x_{ij} - \overline{x}_j \right)^2 \tag{1}$$

where, for the example, x_{ij} denotes the education of a person i in race j, I denotes the total number of people in race j and $\overline{x}j$ denotes the mean of education for people in race j.

To impose this restriction we replaced the weights in Table 1 by the ratio between two sums of squares to obtain a new coding matrix (see Table 2) (Lammers, 1991). The denominator of this ratio is the sum of squares of education among the omitted category. If we multiply this coding matrix with the mean centred education variable, then we get three interaction variables, and the estimates for these variables reflect the correct deviations from the main education effect together with the correct statistical tests. To have the intercept unchanged, we finally mean centred the new interaction variables within each category of race.

In previous approaches to interactions with weighted effect coding (West et al., 1996; Aguinis, 2004), it was not possible to have the effects of the first order model unchanged. This is because a restriction to the coding matrix was used based on the number of observations rather than on the sum of squares used here.

An attractive interpretation of interaction terms is provided: as the (multiplicative) interaction terms are orthogonal to the *main* effects of each category, these main effects remain unchanged upon adding the interaction terms to the model. The interaction terms represent, and test the significance of, the additional effect to the main effects.

The logic of interactions between weighted effect coded dummies and a continuous variable is demonstrated in Figure 2. The dashed blue and red lines represent the effects of education for Blacks and whites, respectively (Hispanics and Asians not shown here). The dashed black line represents the effect of education that is the average of the effects among the four racial categories. This is the effect of education one would estimate if effect coding was used to estimate the interaction, and the differences in slopes between this reference and each racial categories would be the interaction parameters. However, the observations of whites influence the height of the average effect of education in the sample to a larger extent than the Blacks, due to their larger sum of squares. Therefore, the weighted effect of education, shown as the continuous line, is a more useful reference. The sum of squares are represented in Figure 2 by grey squares, and are distinct from the grey circles representing frequencies in Figure 1. The sum of squares pertain to the complete regression slope, and therefore the position of the grey squares was chosen arbitrarily at the center of the x-axis.

Finally, we briefly address the interaction between two weighted effect coded categorical variables. Unlike dummy coding and effect coding, the interaction variables are not simply the multiplication of the two weighted effect coded variables. Instead, partial weights are assigned to the interaction
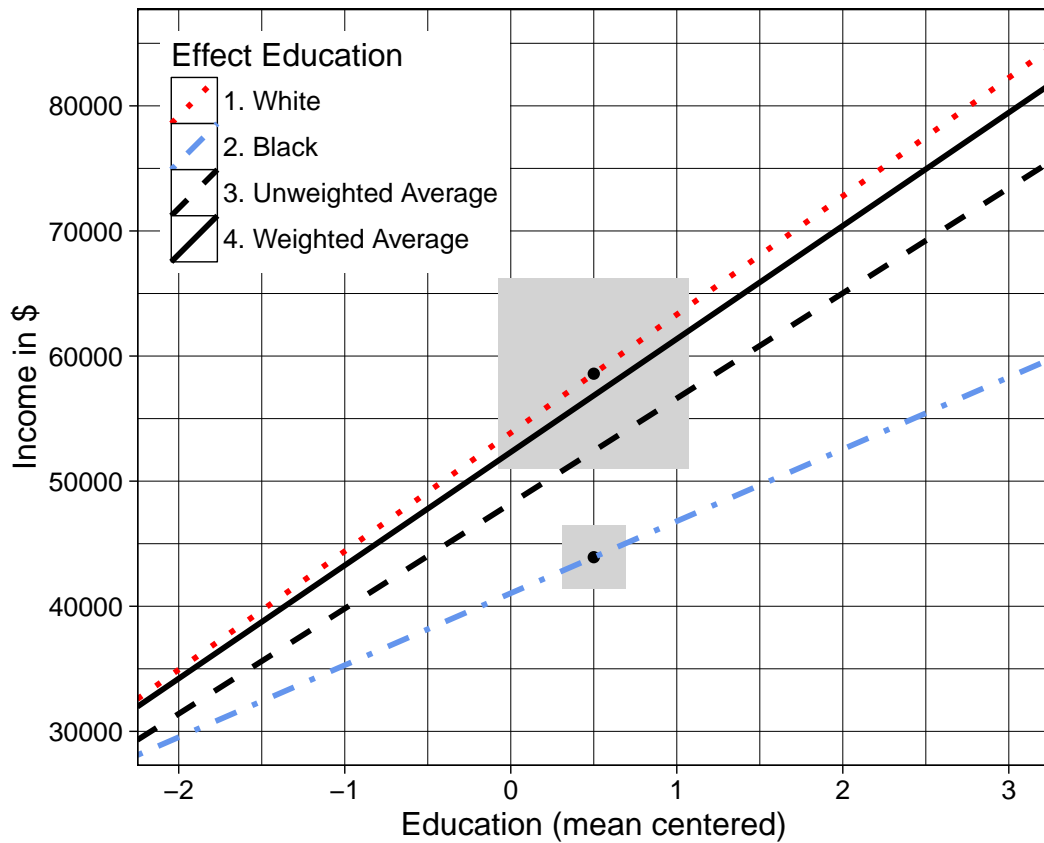
Figure 2: Illustration of interacting weighted effect coding and continuous variable.
Note: Grey squares represent sum of squares (position on x-axis was chosen arbitrarily).

	Hispanic	Black	Asian
Hispanic	1	0	0
Black	0	1	0
Asian	0	0	1
White	$-(SS_{hispanic}/SS_{white})$	$-(SS_{black}/SS_{white})$	$-(SS_{asian}/SS_{white})$

Table 2: Coding Matrix for interaction Factor with Weighted Effect Coding and Continuous variable

		Main Effects			Interaction Effects		
	Degree	Hispanic	Black	Asian	Degree × Hispanic	Degree × Black	Degree × Asian
HS & H	$-(n_d/n_{hs})$	1	0	0	$-(n_{d,h}/n_{hs,h})$	0	0
HS & B	$-(n_d/n_{hs})$	0	1	0	0	$-(n_{d,b}/n_{hs,b})$	0
HS & A	$-(n_d/n_{hs})$	0	0	1	0	0	$-(n_{d,a}/n_{hs,a})$
HS & W	$-(n_d/n_{hs})$	$-(n_h/n_w)$	$-(n_b/n_w)$	$-(n_a/n_w)$	$(n_{d,h}/n_{hs,w})$	$(n_{d,b}/n_{hs,w})$	$(n_{d,a}/n_{hs,w})$
D & H	1	1	0	0	1	0	0
D & B	1	0	1	0	0	1	0
D & A	1	0	0	1	0	0	1
D & W	1	$-(n_h/n_w)$	$-(n_b/n_w)$	$-(n_a/n_w)$	$-(n_{d,h}/n_{d,w})$	$-(n_{d,b}/n_{d,w})$	$-(n_{d,a}/n_{d,w})$

Table 3: Coding Matrix for interaction two factor variables with Weighted Effect Coding

variables to obtain main effects that equal the effects from the model without these interactions (see Table 3 for the weights, for in-depth matrix information about how to create these partial weights please visit http://ru.nl/sociology/mt/wec/downloads/). The orthogonal interaction effects in our example denote the extra wage over and above the mean wages found in the model without these interactions, no matter whether the data are unbalanced or not. In case the data are completely balanced, the estimates from weighted effect coding are equal to those from effect coding, but they can

be quite different in effect size and associated t-values when the data are unbalanced.

Examples of interactions

To demonstrate interactions that include weighted effect coded factor variables, we continue our previous example. For these interactions, the functions in the **wec** package deviate a little from standard R conventions. This is a direct result of how weighted effect coding works. With many forms of dummy coding, interaction variables can be created by simply multiplying the values of the two variables that make up the interaction. This is not true for weighted effect coding, as the coding matrix for the interaction is a function of the numbers of observations of the two variables that interact. So, instead of multiplying two variables in the specification of the regression model in typical R-fashion, a new, third, variable is created prior to specifying the regression model and then added. Here, we refer to these additional variables as the 'interaction' variable.

Interaction variables for interacting weighted effect coded factor variables are produced by the wec.interact() function. The first variable entered ('x1') must be a weighted effect coded factor variable. The second ('x2') can either be a continuous variable or another weighted effect coded factor variable. By default, this function returns an object containing one column for each of the interaction variables required. However, by specifying output.contrasts = TRUE, the coding matrix (see Table 2) is returned:

```
> wec.interact(PUMS$race.wec, PUMS$education.int, output.contrasts = TRUE)

   [,1]   [,2]   [,3]
1  1.0   0.000  0.000
2  0.0   1.000  0.000
3  0.0   0.000  1.000
4 -0.1  -0.098 -0.066
```

The example above shows the coding matrix for interacting the (weighted effect coded) race variable with the continuous education variable. The omitted category is, again, 'Whites' (category 4), and the coding matrix shows the ratio of sum of squares as was defined in Table 2. To include this in the regression analysis, a new factor variable is created:

```
> PUMS$race.educint <- wec.interact(PUMS$race.wec, PUMS$education.int)
> m.wec.educ <- lm(wage ~ race.wec + education.int + race.educint, data=PUMS)
> summary(m.wec.educ)$coefficients
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	52320	559	93.5	0.0e+00
race.wecHispanic	-4955	1736	-2.9	4.3e-03
race.wecBlack	-11276	1817	-6.2	5.7e-10
race.wecAsian	5151	2381	2.2	3.1e-02
education.int	9048	287	31.6	2.3e-208
race.educintinteractHispanic	-3266	977	-3.3	8.3e-04
race.educintinteractBlack	-3293	990	-3.3	8.8e-04
race.educintinteractAsian	3575	1217	2.9	3.3e-03

The wec.interact function is now called without the output.contrasts = TRUE option. The first specification is the factor variable and the second term is the continuous variable. The results are stored in a new variable. This new interaction variable is entered into the regression model *in addition to* the variables with the main effects for race and education.

The results show that the returns to education, in terms of wages, for Hispanics and Blacks are lower than the average returns to education in the sample, and the returns to education are higher among Asians than it is in the sample as a whole. Note that without additional control variables, all effects for race and for education, as well as the estimate for the intercept, remained unchanged compared to previous examples after including the (weighted effect coded) interaction variable.

Note that if one wants to estimate the effects and standard errors for the omitted category, in this case 'Whites', not only the contrasts for the categorical variable need to be changed (as demonstrated above), but also the interaction variable needs to be updated.

Below, we specify the interaction between the race variable with a factor variable differentiating respondents who have a high school diploma and those who have a higher degree. Of course, both are weighted effect coded:

```
> PUMS$education.wec <- PUMS$education.cat
> contrasts(PUMS$education.wec) <- contr.wec(PUMS$education.cat, "High school")
```

```
> PUMS$race.educat <- wec.interact(PUMS$race.wec, PUMS$education.wec)

> m.wec.educwec <- lm(wage ~ race.wec + education.wec + race.educat, data=PUMS)
> summary(m.wec.educwec)$coefficients

                          Estimate Std. Error t value Pr(>|t|)
(Intercept)                  52320        569    92.0  0.0e+00
race.wecHispanic             -6645       1764    -3.8  1.7e-04
race.wecBlack               -11738       1849    -6.3  2.3e-10
race.wecAsian                 7528       2419     3.1  1.9e-03
education.wecDegree          14343        572    25.1 1.6e-134
race.educatx1Hispanic:x2Degree -7674     2441    -3.1  1.7e-03
race.educatx1Black:x2Degree   -6682       2252    -3.0  3.0e-03
race.educatx1Asian:x2Degree    4022       1536     2.6  8.8e-03
```

We created a new categorical variable 'education.wec' and assigned a coding matrix based on weighted effect coding, with 'High school' as the omitted category. The results show that respondents with a degree on average earn 14,343 dollar more than the sample average (52,320). Hispanics benefit 7,674 dollar less from having a degree compared to the average benefit of a degree, while Asians benefit 4,022 dollar more. All in all, the results are very similar to those in the previous model with the continuous variable for education. It should be noted that in the model with interactions between weighted effect coded factor variables, the intercept again shows the same value, representing the average wage in the sample. Just like with the previous examples, the omitted estimates and standard errors (for instance the income effect of Hispanics without a degree) can be obtained by changing the omitted categories in the weighted effect coded factor variables, and by re-calculating the interaction variable(s).

Conclusion

This article discussed benefits and applications of weighted effect coding. It covered weighted effect coding as such, interactions between two weighted effect coded variables, and interactions with a weighted effect coded variable and an continuous variable. The **wec** package to apply these techniques in R was introduced. The examples shown in this article were based on OLS regression, but weighted effect coding (also) applies to all generalised linear models.

The benefits of using weighted effect coding are apparent when analysing observational data that, unlike experimental data, typically do not have an equal number of observations across groups or categories. When this is the case, the grand mean is not necessarily the appropriate point of reference. Consequently estimates of effects and standard errors based on weighted effect coding are not sensitive to how *other* observations are categorised.

With weighted effect coding, compared to treatment coding, no arbitrary reference category has to be selected. Instead, the sample mean serves as a point of reference. With treatment coding, selecting as a reference a category with a small number of observations and a deviant score can lead to significant results while this reference category has little contribution to the overall sample mean.

When weighted effect coded variables are used in interactions, the main effects remain unchanged after the introduction of the interaction terms. In previous, related, approaches this was not possible (West et al., 1996; Aguinis, 2004). This allows for the straightforward interpretation that the interaction terms represent how much the effect is weaker / stronger in each category. That is, when interacting with treatment coded categorical variables, the so-called 'main' effect refers to the reference category, whereas with weighted effect coding the unconditional main (/mean) effect is shown. As such, it can be used to test the assumption that estimated effects do not vary across groups.

It should be noted that the R-square of regression models does not depend on which type of dummy coding is selected. This means that the predicted values based on models using treatment coding, effect coding, or weighted effect coding, will be exactly the same. Yet, as each type of dummy coding selects a different point of reference, the interpretation of the estimates differs and a different statistical test is performed.

To conclude, the **wec** package contributes functionality to apply weighted effect coding to factor variables and interactions between (a.) a factor variable and a continuous variable and between (b.) two factor variables. These techniques are particularly relevant with unbalanced data, as is often the case when analysing observational data.

Bibliography

H. Aguinis. *Regression Analysis for Categorical Moderators*. The Guilford Press, New York / London, 2004. [p5, 8]

Census. *Public Use Microdata Sample (PUMS)*. prepared by the U.S. Census Bureau, United States of America, 2013. doi: 10.3886/ICPSR33802.v1. [p3]

M. Grotenhuis, B. Pelzer, R. Eisinga, R. Nieuwenhuis, A. Schmidt-Catran, and R. Konig. A novel method for modelling interaction between categorical variables. *International Journal of Public Health*, 62(3):427–431, 2017a. doi: 10.1007/s00038-016-0901-1. [p1, 5]

M. Grotenhuis, B. Pelzer, R. Eisinga, R. Nieuwenhuis, A. Schmidt-Catran, and R. Konig. When size matters: advantages of weighted effect coding in observational studies. *International Journal of Public Health*, 62(1):263–167, 2017b. doi: 10.1007/s00038-016-0902-0. [p1]

J. Lammers. *Grootschalig veldonderzoek. Inleiding in de data-analyse (An introduction to the analysis of observational data)*. Coutinho, Muiderberg, 1991. [p5]

R. E. Sweeney and E. F. Ulveling. A transformation for simplifying the interpretation of coefficients of binary variables in regression analysis. *The American Statistician*, 1972. doi: 10.1080/00031305.1972. 10478949. [p1]

S. G. West, L. S. Aiken, and J. L. Krull. Experimental Personality Designs: Analyzing Categorical by Continuous Variable Interactions. *Journal of personality*, 64(1):1–48, Mar. 1996. doi: 10.1111/j.1467-6494.1996.tb00813.x. [p5, 8]

Rense Nieuwenhuis
Swedish Institute for Social Research (SOFI), Stockholm University
Affiliate of The Linnaeus Center on Social Policy and Family Dynamics in Europe (SPaDE), Stockholm University
106 91 Stockholm
Sweden
rense.nieuwenhuis@sofi.su.se

Manfred te Grotenhuis
Radboud University
Box 9104, 6500 HE, Nijmegen
The Netherlands
m.tegrotenhuis@maw.ru.nl

Ben Pelzer
Radboud University
Box 9104, 6500 HE, Nijmegen
The Netherlands
b.pelzer@maw.ru.nl