

Statistical modelling

05. Linear models

Léo Belzile, HEC Montréal

2024

Interaction

An **interaction** occurs if some explanatory variables, when coupled together, have different impacts than the superposition of each.

If X_j and X_k interact, the marginal effect of $E(Y | \mathbf{X})$ with respect to X_j is a function of X_k or vice-versa.

We will restrict attention to the cases where one or more of the explanatories is a categorical variable (factor).

Insurance premium

Smokers who have a BMI of 30 and above pay a hefty premium, but there is also seemingly a linear increase in the amount charged with BMI. We see no such behaviour for non-smokers.

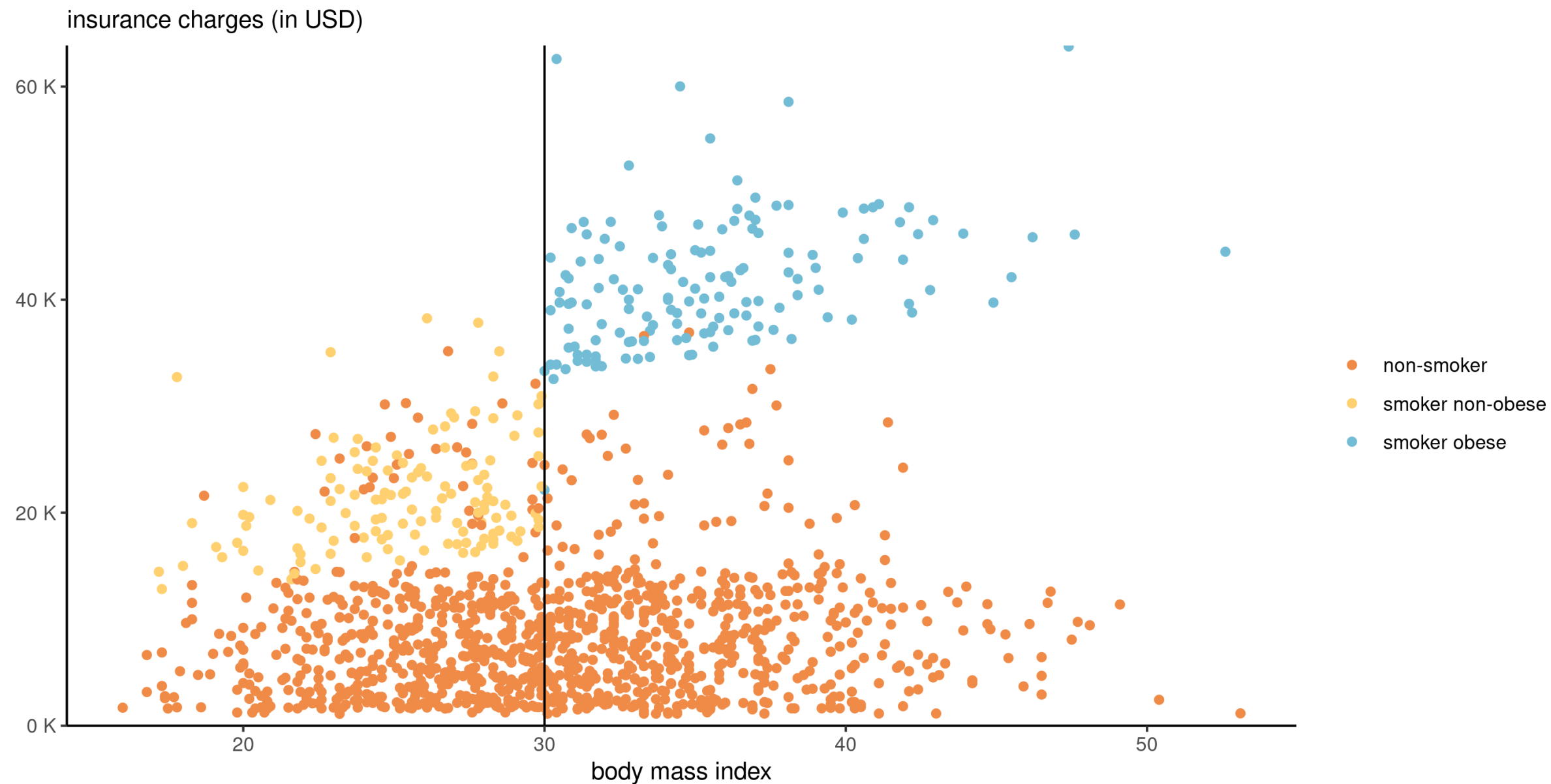


Figure 1: Graph of [insurance](#) charges against body mass index, colored by smoking status.

Toy example 1 – continuous vs categorical

We consider a toy model for the *interaction* data. The base model, without interaction, is

$$E(\text{intention} \mid \cdot) = \beta_0 + \beta_1 \text{sex} + \beta_2 \text{fixation},$$

where $\text{sex}=1$ for women and $\text{sex}=0$ for men.

The effect of fixation in this model is the same regardless of sex.

In order to add a different slope for men and women, we can create a new variable equal to the product $\text{fixation} \times \text{sex}$ and add it to the model,

$$E(\text{intention} \mid \cdot) = \beta_0 + \beta_1 \text{sex} + \beta_2 \text{fixation} + \beta_3 \text{fixation} \cdot \text{sex}.$$

Is there an interaction?

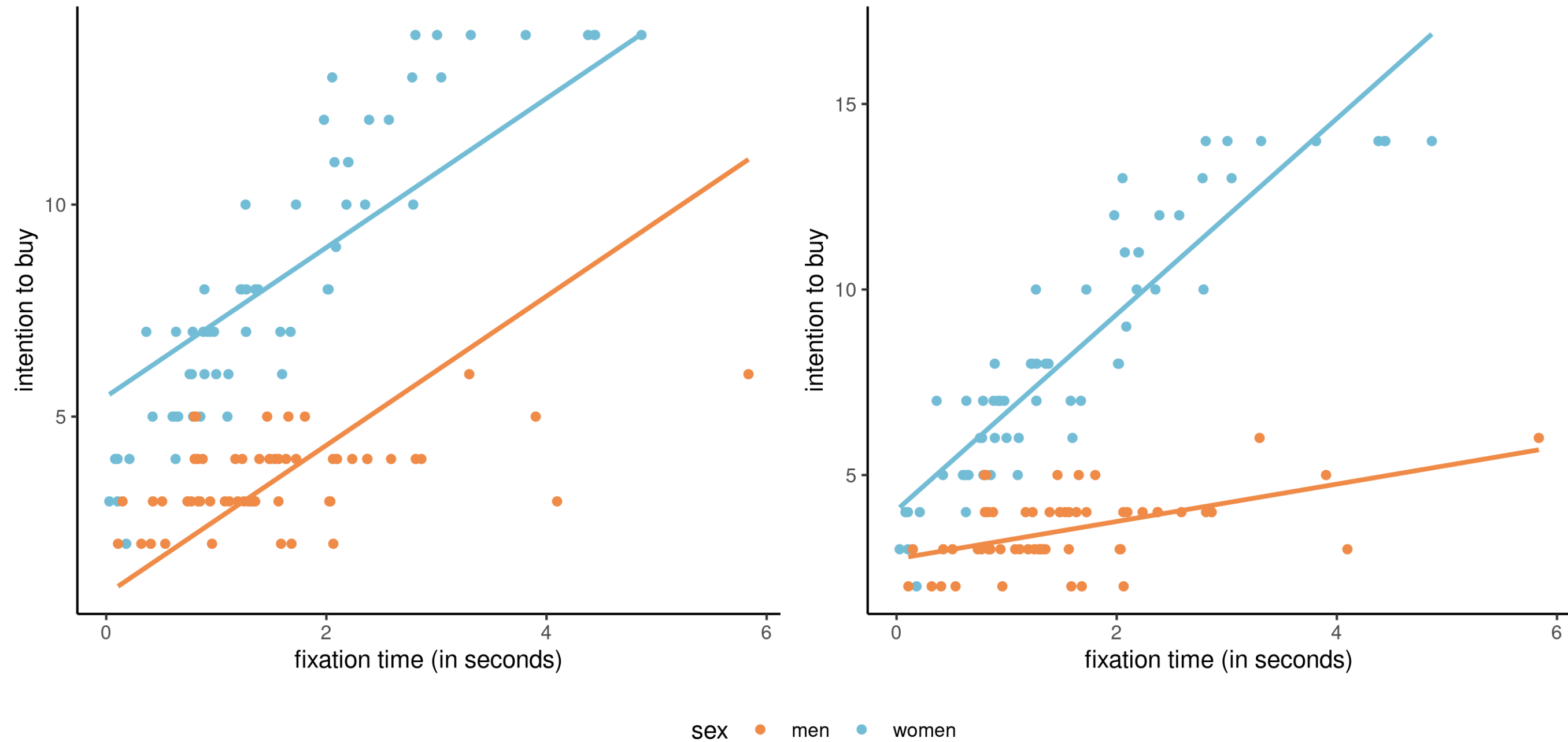


Figure 2: Scatterplots and fitted lines for a model with a single continuous and binary explanatory, without (left) and with (right) an interaction term.

Parameter interpretation

Depending on the value of the binary variable **sex**, we get

$$E(\mathbf{intention} \mid \cdot) = \begin{cases} (\beta_0 + \beta_1) + (\beta_2 + \beta_3)\mathbf{fixation}, & \mathbf{sex} = 1 \text{ (women)}, \\ \beta_0 + \beta_2\mathbf{fixation}, & \mathbf{sex} = 0 \text{ (men)}. \end{cases}$$

The interpretation of the coefficients in the model is as usual with the treatment contrast parametrization:

- β_0 is the average buying intention when the fixation time is zero for men,
- β_1 is the difference in intercept for women vs men,
- β_2 is the unit increase in intention to buy per second of fixation for men,
- β_3 is the difference in slope for women vs men.

Testing for an interaction

Whether the interaction is significant boils down to testing $\mathcal{H}_0 : \beta_3 = 0$.

```

1 data(interaction, package = "hecstatmod")
2 # To specify an interaction use :
3 mod <- lm(intention ~ sex + fixation + sex:fixation,
4           data = interaction)
5 # A shortcut is sex*fixation, which expands to the above
6 summary(mod)$coefficients
7 ##              Estimate Std. Error t value Pr(>|t|)
8 ## (Intercept)      2.7         0.28     9.7 1.0e-16
9 ## sex              1.3         0.38     3.5 7.7e-04
10 ## fixation        0.5         0.15     3.3 1.3e-03
11 ## sex:fixation    2.1         0.20    10.7 5.6e-19

```

The model with the interaction is significantly better, meaning that the effect of fixation time on intention to buy varies according to sex.

Marginality principle

All lower interaction terms should be included if an interaction is present.

For example, we would **not** remove **fixation** while keeping the interaction term **fixation*sex**, even if we fail to reject $\mathcal{H}_0 : \beta_2 = 0$ because otherwise

$$E(\mathbf{intention} \mid \cdot) = \begin{cases} (\beta_0 + \beta_1) + \beta_3 \mathbf{fixation}, & \mathbf{sex} = 1 \text{ (women)}, \\ \beta_0, & \mathbf{sex} = 0 \text{ (men)}; \end{cases}$$

this implies that intention to buy is constant for men, regardless of the fixation time.

As the choice of baseline is arbitrary, changing the dummy (0 for women, 1 for men), would yield a different model and so potentially different inferences.

Example 2 - categorical vs categorical

Consider a linear model with factors A and B and their interactions.

This is a **two-way ANOVA model**, in which each subgroup (a_i, b_j) has a different mean μ_{ij} .

e.g., if A has $n_a = 3$ levels and B has $n_b = 2$ levels.

	B	b_1	b_2	row mean
A				
a_1		μ_{11}	μ_{12}	$\mu_{1.}$
a_2		μ_{21}	μ_{22}	$\mu_{2.}$
a_3		μ_{31}	μ_{32}	$\mu_{3.}$
column mean		$\mu_{.1}$	$\mu_{.2}$	μ

- Sample estimates are obtained by replacing μ_{ij} by subgroup sample means.
- Row averages $\mu_{i.}$, column averages $\mu_{.j}$ and the global average μ are **equiweighted** combinations of the cell means μ_{ij} .

Example 1 - perception of debt

Sharma, Tully, and Cryder (2021) conducted a 2 by 2 between-subjects comparison (two-way ANOVA) varying

- the type of debt (whether the money was advertised as *credit* or *loan*) and
- the type of purchase the latter would be used for (*discretionary* spending or *need* for necessary purchases).

```
1 data(STC21_SS5, package = "hecedsm")
2 mod1 <- lm(likelihood ~ purchase * debtttype, data = STC21_SS5)
```

The response is the average of the likelihood and interest in the product, both measured using a 9 point Likert scale from 1 to 9.

Example 2 - Perceptions of cultural appropriation by ideology

We consider a three-way ANOVA from Lin et al. (2024). Their Study 4 focused on cultural appropriation for a soul food recipe, an Afro-American tradition, that appears in a cookbook from Chef Dax.

The authors manipulated the ethnicity of Chef Dax, who was either black (or not), and the description of the way he obtained the recipes (by peeking without permission in kitchens, by asking permission or no mention for control).

```
1 data(LKUK24_S4, package = "hecedsm")
2 mod2 <- lm(appropriation ~ politideo * chefdax * brandaction,
3           data = LKUK24_S4)
```

Authors postulated that the perception of appropriation would vary by political ideology (liberal or conservative). The study results in a 3 by 2 by 2 three-way ANOVA.

Interaction plot

Plot the averages (with confidence intervals) as a function of the explanatories. If lines are parallel, then there is no interaction.

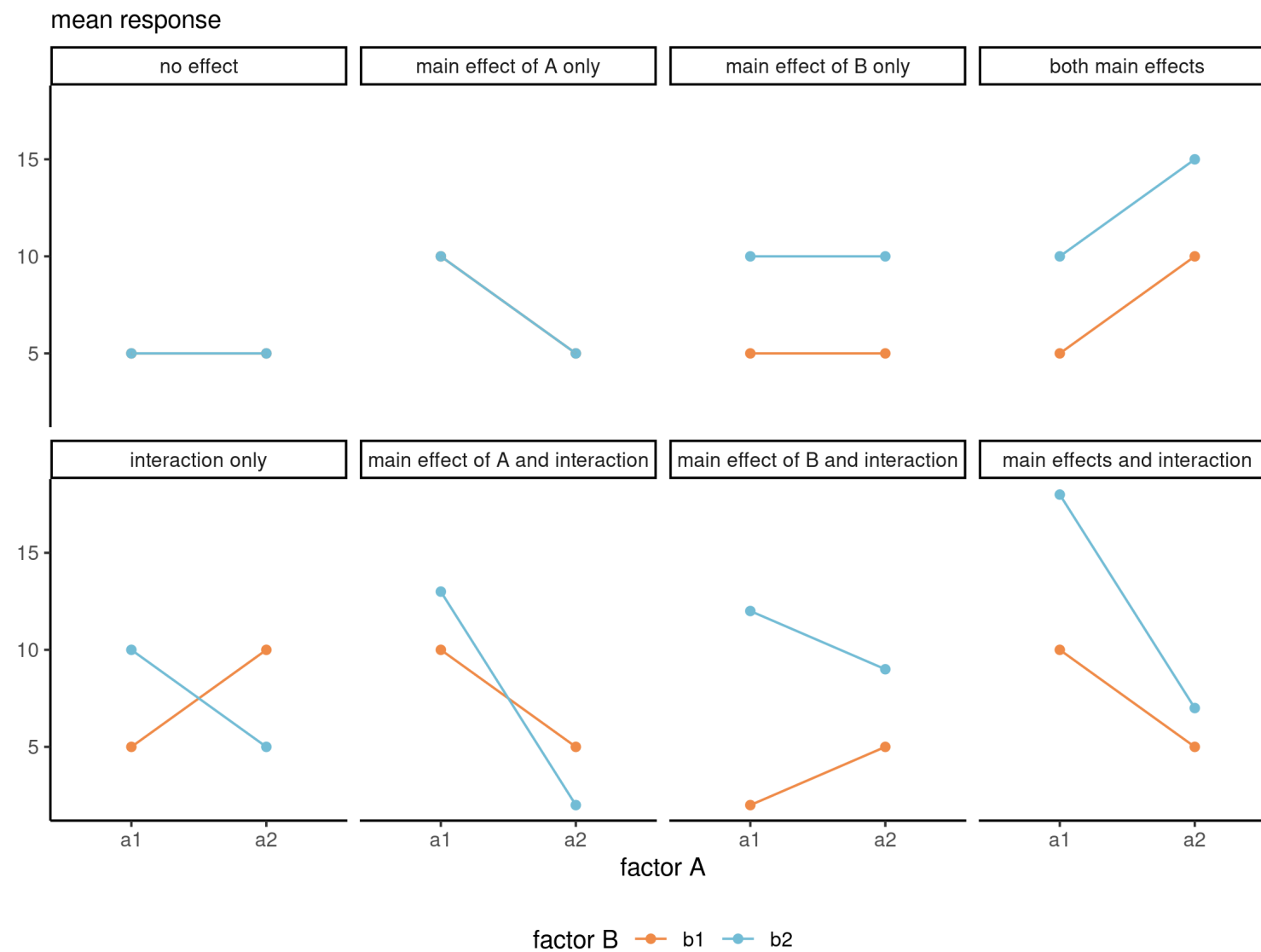


Figure 3: Interaction plots for 2 by 2 designs. Illustration adapted from Figure 10.2 of Crump, Navarro and Suzuki (2019) by Matthew Crump (CC BY-SA 4.0 license)

Cell means for 2 by 2 designs

Refer to the previous slide for the scenarios.

	b1	b2
<i>a1</i>	5	5
<i>a2</i>	5	5

	b1	b2
<i>a1</i>	10	10
<i>a2</i>	5	5

	b1	b2
<i>a1</i>	5	10
<i>a2</i>	5	10

	b1	b2
<i>a1</i>	5	10
<i>a2</i>	10	15

	b1	b2
<i>a1</i>	5	10
<i>a2</i>	10	5

	b1	b2
<i>a1</i>	10	13
<i>a2</i>	5	2

	b1	b2
<i>a1</i>	2	12
<i>a2</i>	5	9

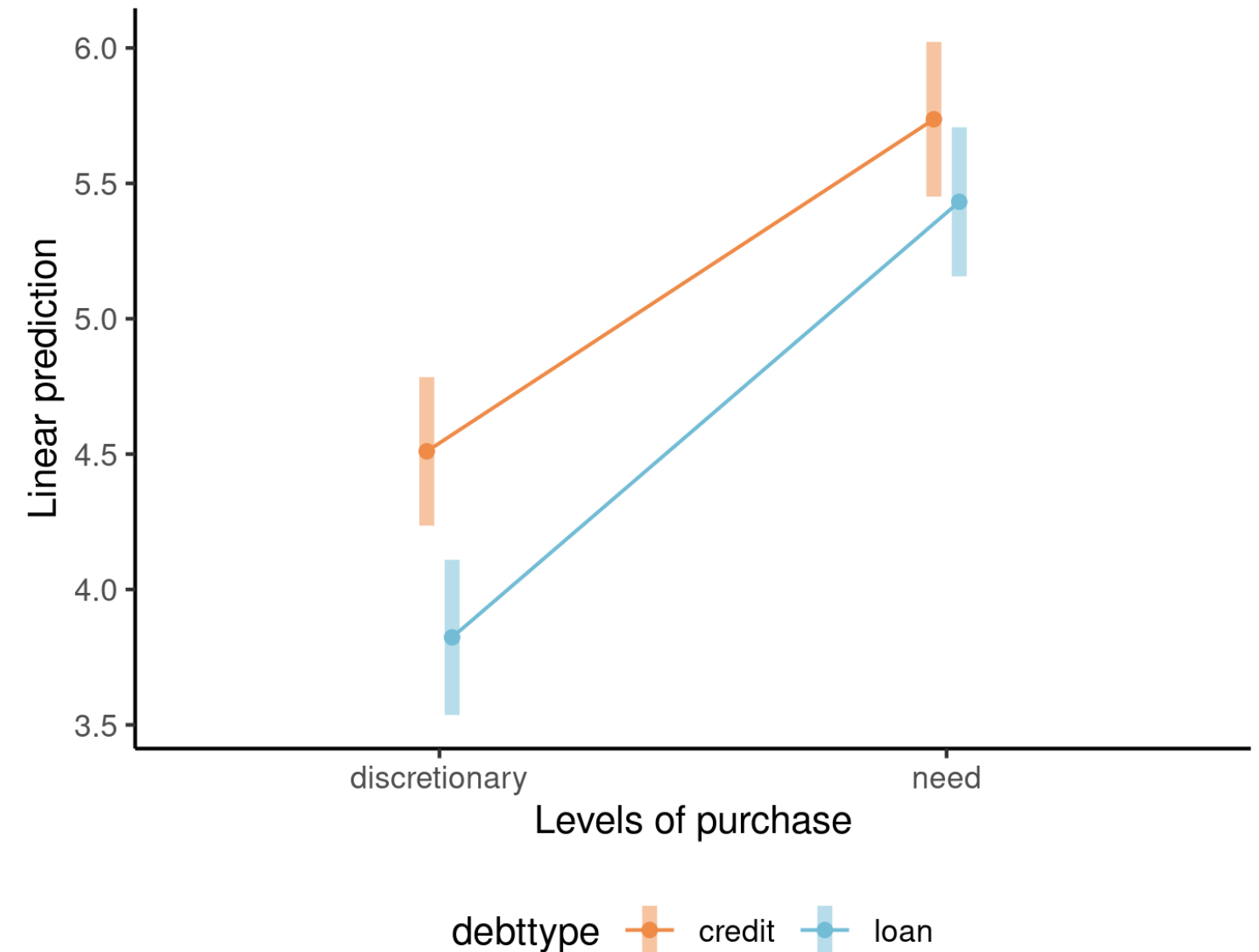
	b1	b2
<i>a1</i>	10	18
<i>a2</i>	5	7

Table 1: Patterns for means for each of the possible kinds of general outcomes in a 2 by 2 design.

Interaction plot for Example 1

Supplementary study 5 of Sharma, Tully, and Cryder (2021) consists of a 2×2 between-subject (i.e., no repeated measure per individual) ANOVA with factors

- debt type (*debttype*), either “loan” or “credit”
- purchase type, either *discretionary* or not (*need*)

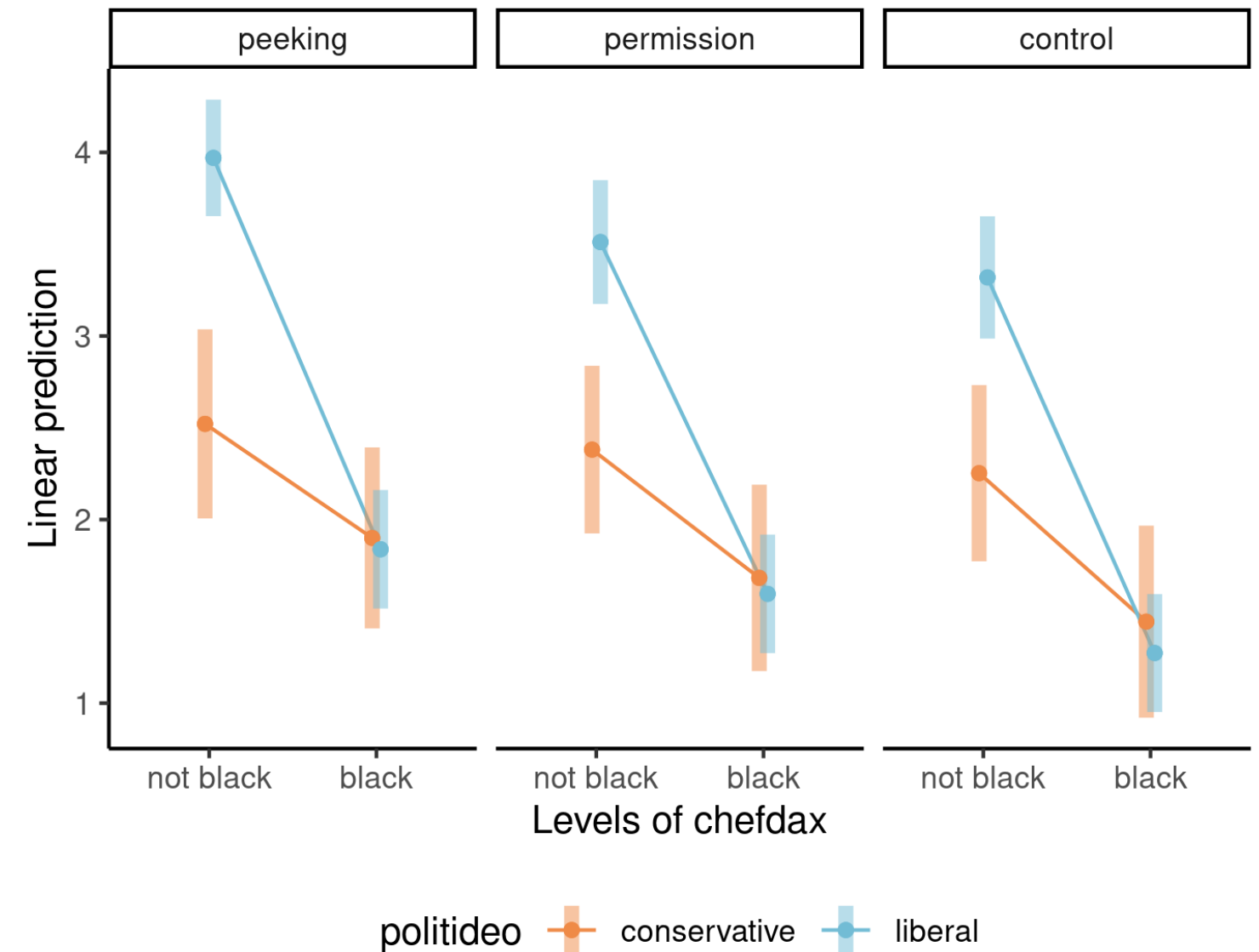


No evidence of interaction.

Example 2 - cultural appropriation

Study 4 of Lin et al. (2024) uses a $3 \times 2 \times 2$ between-subject ANOVA with factors

- **chefdax**: ethnicity, either **black** or **not black**
- **politideo**: political ideology, either **conservative** or **liberal**
- **brandaction**: how recipes were obtained by Chef Dax, either peeking, with permission or no mention (control).



Evidence of an interaction between **chefdax** (ethnicity) and **ideology**.

Analysis of variance table

The analysis of variance table compares models with (or without) A , B , AB .

If the sample is **balanced** (same number of observations per subgroup), we can uniquely decompose the variance as

$$\underset{\text{total}}{SS_c} = \underset{\text{factor } A}{SS_A} + \underset{\text{factor } B}{SS_B} + \underset{\text{interaction}}{SS_{AB}} + \underset{\text{residual}}{SS_e}$$

- a model without AB would have residual sum of squares of $SS_{AB} + SS_e$, but $(n_a - 1)(n_b - 1)$ fewer parameters.
- a model without A (i.e., only factor B) would have $n_a - 1$ fewer parameter, and residual sum of squares of $SS_A + SS_{AB} + SS_e$, versus $SS_{AB} + SS_e$ for the alternative.
- the difference in sum of squares is SS_A for A , SS_{AB} for the interaction $A \times B$, etc.

Test statistic for ANOVA

Consider removal of A in a balanced design. If the alternative model has $n_a n_b$ parameters for the mean, and we impose $n_a - 1$ linear restrictions under the null hypothesis to the model estimated based on n independent observations, the test statistic is

$$F = \frac{SS_A / (n_a - 1)}{SS_e / (n - n_a n_b)}$$

- The numerator is the difference in sum of squares, denoted SS_A , from models fitted under \mathcal{H}_0 and \mathcal{H}_a , divided by degrees of freedom $n_a - 1$ (number of additional parameters).
- The denominator is equal to S^2 , the unbiased estimator of the variance.

Analysis of variance table

term	degrees of freedom	mean square	F
A	$n_a - 1$	$MS_A = SS_A / (n_a - 1)$	MS_A / MS_{res}
B	$n_b - 1$	$MS_B = SS_B / (n_b - 1)$	MS_B / MS_{res}
AB	$(n_a - 1)(n_b - 1)$	$MS_{AB} = SS_{AB} / \{(n_a - 1)(n_b - 1)\}$	$MS_{AB} / MS_{\text{res}}$
residuals	$n - n_a n_b$	$MS_{\text{res}} = SS_e / (n - ab)$	
total	$n - 1$		

Read the table backward (starting with the interaction).

Comparing nested models with unbalanced data

Depending on the decomposition, we are comparing different nested models!

Table 2: Sum of square decompositions in ANOVA tables. Comparison of sum of squares between null, versus alternative model.

	type 1	type 2	type 3
<i>A</i>	intercept vs <i>A</i>	<i>B</i> vs (<i>A</i> , <i>B</i>)	(<i>B</i> , <i>AB</i>) vs (<i>A</i> , <i>B</i> , <i>AB</i>)
<i>B</i>	<i>A</i> vs (<i>A</i> , <i>B</i>)	<i>A</i> vs (<i>A</i> , <i>B</i>)	(<i>A</i> , <i>AB</i>) vs (<i>A</i> , <i>B</i> , <i>AB</i>)
<i>AB</i>	(<i>A</i> , <i>B</i>) vs (<i>A</i> , <i>B</i> , <i>AB</i>)	(<i>A</i> , <i>B</i>) vs (<i>A</i> , <i>B</i> , <i>AB</i>)	(<i>A</i> , <i>B</i>) vs (<i>A</i> , <i>B</i> , <i>AB</i>)

Use the type 2 sum of square by default (default with `car::Anova`). Type 1 decomposition (`anova`) is sequential, while type 3 does not respect marginality principles.

Mean model for Example 1

The mean model with an interaction can be written using the treatment contrast parametrization as

$$\begin{aligned} E(\text{likelihood} \mid \cdot) = & \beta_0 + \beta_1 \mathbf{1}_{\text{purchase}=\text{need}} + \beta_2 \mathbf{1}_{\text{debttype}=\text{loan}} \\ & + \beta_3 \mathbf{1}_{\text{purchase}=\text{need}} \mathbf{1}_{\text{debttype}=\text{loan}}. \end{aligned}$$

We can retrieve the average of each subgroup and deduce the interpretation of coefficients:

- $\mu_1 = \beta_0$ for `purchase=discretionary` and `debttype=credit`
- $\mu_2 = \beta_0 + \beta_1$ for `purchase=need` and `debttype=credit`
- $\mu_3 = \beta_0 + \beta_2$ for `purchase=discretionary` and `debttype=loan`
- $\mu_4 = \beta_0 + \beta_1 + \beta_2 + \beta_3$ for `purchase=need` and `debttype=loan`

Thus, β_3 represents the difference in means $\mu_4 - \mu_2 - \mu_3 + \mu_1$.

Fitting ANOVA and extracting group means

```
1 # Analysing Supplementary Study 5
2 # of Sharma, Tully, and Cryder (2021)
3 data(STC21_SS5, package = "hecedsm")
4 # Use 'aov' to fit models with categorical variables
5 # Equivalent to 'lm' with sum-to-zero contrasts
6 aov_mod1 <- aov(likelihood ~ purchase*debttype,
7               data = STC21_SS5)
8 # Check counts per subcategory (data are unbalanced)
9 xtabs(~purchase + debttype, data = STC21_SS5)
10 ##                debttype
11 ## purchase      credit loan
12 ## discretionary  392   359
13 ## need          361   389
14 # Compute overall/rows/columns/cells means
15 mean_groups <- model.tables(aov_mod1, type = "means")
```

Model comparison

```

1 car::Anova(mod1, type = 2)
2 ## Anova Table (Type II tests)
3 ##
4 ## Response: likelihood
5 ##
6 ##           Sum Sq   Df F value  Pr(>F)
7 ## purchase         752    1  98.21 < 2e-16 ***
8 ## debttype          92    1  12.04 0.00054 ***
9 ## purchase:debttype  14    1   1.79 0.18171
10 ## Residuals      11467 1497
11 ## ---
12 ## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

In the analysis of variance table, we focus exclusively on the line `purchase:debttype`.

The F statistic is 1.79; using the Fisher (1, 1497) distribution as null distribution, we obtain a p -value of 0.18 so there is no evidence the effect of purchase depends on debt type.

Main effects and marginalization

Main effects are comparisons between row or column averages

Obtained by *marginalization*, i.e., averaging over the other dimension.

Main effects are misleading if there is an interaction.

For example, the main effects of A are:

A	row mean
a_1	$\mu_{1.}$
a_2	$\mu_{2.}$
a_3	$\mu_{3.}$

Simple effects

When there are interactions, the effect of A depends on the value of B .

Simple effects are comparisons between cell averages within a given row or column.

A	b_1
a_1	μ_{11}
a_2	μ_{21}
a_3	μ_{31}

Example of main effects

Since the interaction is not significant, we can extract the main effect of `debttype` and compute the pairwise difference.

```

1 # Pairwise comparisons within levels of purchase
2 # Using the main effects
3 emmeans::emmeans(mod1,
4                   # what variable to keep (so average over "purchase")
5                   specs = "debttype",
6                   contr = "pairwise")
7 ## $emmeans
8 ##   debttype emmean    SE    df lower.CL upper.CL
9 ##   credit      5.1 0.101 1497     4.9     5.3
10 ##   loan        4.6 0.101 1497     4.4     4.8
11 ##
12 ## Results are averaged over the levels of: purchase
13 ## Confidence level used: 0.95
14 ##
15 ## $contrasts
16 ##   contrast      estimate    SE    df t.ratio p.value
17 ##   credit - loan        0.5 0.143 1497   3.500 0.0005
18 ##
19 ## Results are averaged over the levels of: purchase

```

ANOVA table for Example 2

For the K -way ANOVA, we always start with estimating the full model with all K -way interaction (provided there are enough data to estimate the latter, which implies there are repetitions).

Table 3: Analysis of variance table (type II decomposition) for the data from Study 4 of Lin et al. (2024).

term	sum of squares	df	stat	p-value
politideo	48.49	1	21.35	<0.001
chefdax	473.72	1	208.61	<0.001
brandaction	34.24	2	7.54	<0.001
politideo:chefdax	65.00	1	28.63	<0.001
politideo:brandaction	1.56	2	0.34	0.71
chefdax:brandaction	0.62	2	0.14	0.87
politideo:chefdax:brandaction	0.66	2	0.15	0.86
Residuals	1587.33	699		

There is no three-way interaction and a single two-way interaction between political ideology and the race of Chef Dax. We cannot interpret the p -value for the main effect of `brandaction`, but we could look at the marginal means.

Dimension reduction and simple effects

Collapse to a 2 by 2 two-way ANOVA, averaging over `brandaction`.

```

1 # Marginal means for political ideology/Chef Dax
2 # Compute simple effects, by political ideology
3 (emm2 <- emmeans(mod2,
4     specs = "chefdax",
5     by = "politideo"))
6 ## politideo = conservative:
7 ##   chefdax   emmean    SE  df lower.CL upper.CL
8 ## not black   2.4 0.142 699     2.1     2.7
9 ## black       1.7 0.149 699     1.4     2.0
10 ##
11 ## politideo = liberal:
12 ##   chefdax   emmean    SE  df lower.CL upper.CL
13 ## not black   3.6 0.097 699     3.4     3.8
14 ## black       1.6 0.095 699     1.4     1.8
15 ##
16 ## Results are averaged over the levels of: brandaction
17 ## Confidence level used: 0.95

```

Pairwise contrasts between ethnicity, by political ideology

```

1 emm2 |> contrast(method = "pairwise")
2 ## politideo = conservative:
3 ## contrast          estimate      SE   df t.ratio p.value
4 ## not black - black    0.71 0.206 699   3.400 0.0006
5 ##
6 ## politideo = liberal:
7 ## contrast          estimate      SE   df t.ratio p.value
8 ## not black - black    2.03 0.135 699  15.000 <.0001
9 ##
10 ## Results are averaged over the levels of: brandaction

```

We see that the liberals are much more likely to view Chef Dax cookbook as an instance of cultural appropriation if he is not black; there is limited evidence of any difference between conservatives and liberal when Chef Dax is black.

We can also compare marginal effects for ideology separately for `black` or `not black`. Both differences are statistically significant, but the differences (and thus evidence of an effect) is much stronger for left-leaning respondents.

Comparisons of main effects of brandaction

We expect participants will view peeking less favorably than if Chef Dax asked for permission to publish the recipes. It's tricky to know the effect of the control, as we are not bringing the point to the attention of participants in this instance.

```

1 # Marginal mean for brandaction
2 emm_brand <- emmeans(mod2, specs = c("brandaction"))
3 emm_brand
4 ## brandaction emmean SE df lower.CL upper.CL
5 ## peeking 2.56 0.108 699 2.35 2.77
6 ## permission 2.29 0.105 699 2.09 2.50
7 ## control 2.07 0.108 699 1.86 2.28
8 ##
9 ## Results are averaged over the levels of: politideo, chefdax
10 ## Confidence level used: 0.95
11 # Joint F test for the main effect of brandaction
12 emm_brand |> pairs() |> joint_tests()
13 ## model term df1 df2 F.ratio p.value
14 ## contrast 2 699 5.100 0.0064

```

A joint F -test, obtained by collapsing everything to a one-way ANOVA, shows that there are indeed differences.

Contrasts

We can view the multiway ANOVA as a one-way analysis of variance with $n_a \times n_b \times \dots$ levels, corresponding to each sub-group.

Marginal effects, interactions, and simple effects correspond to particular **linear contrasts**.

```
1 mod <- lm(likelihood ~ group,  
2         data = STC21_SS5 |>  
3         dplyr::mutate(group = interaction(debttype, purchase)))  
4 emmeans(mod, specs = "group") |>  
5   contrast(method = list(main_pairwise = c(1,-1,1,-1)/2))  
6 ## contrast      estimate      SE    df t.ratio p.value  
7 ## main_pairwise      0.5 0.143 1497   3.500 0.0005
```

Recap 1

- Interactions occur when the effect of a variable depend on another: we typically model this by adding the product of the two (one or more being categorical with dummy indicators).
- **Interaction plots**, showing group averages, are useful conceptually to look for interactions, but formal tests are needed.
- Tests of statistical significance consider removal of interactions based on fitting the complete model (when possible).
- The marginality principle implies we keep all lower effects: use type II effects.

Recap 2

- Analysis of variance models are simply linear regression models with categorical explanatories.
- The models with all interactions correspond to each subgroup having a specific average.
- We compare of main effects (if interactions are not present) or simple effects (when they are), or more general contrasts, based on the full model.
- A multiway ANOVA can always be cast as a one-way ANOVA.

References

- Lin, Jason D, Nicole You Jeung Kim, Esther Uduehi, and Anat Keinan. 2024. "Culture for Sale: Unpacking Consumer Perceptions of Cultural Appropriation." *Journal of Consumer Research*. <https://doi.org/10.1093/jcr/ucad076>.
- Sharma, Eesha, Stephanie Tully, and Cynthia Cryder. 2021. "Psychological Ownership of (Borrowed) Money." *Journal of Marketing Research* 58 (3): 497–514. <https://doi.org/10.1177/0022243721993816>.