Stat 412/512

Another Two-way ANOVA

Feb 9 2015

Charlotte Wickham stat512.cwick.co.nz

A two-way ANOVA

Sometimes only one factor is of interest, sometimes both are, sometimes the interaction is the primary interest.

The general approach is the same:

Start with the saturated model

Use F-tools to simplify

Then answer specific questions about means

Last time:

There is no evidence the treatment effect differs depending on the company (extra SS F-test on interaction term, p-value = 0.72).

comparing saturated model to the additive model

coming from additive model

There is moderate evidence that the pygmalion treatment changes the platoon's score (two sided p-value on t-test of treatment effect = 0.01).

It is estimated the pygmalion treatment adds 7.2 points to a platoon's score.

With 95% confidence, the pygmalion treatment adds between 1.8 and 12.6 points to a platoon's score

randomized

Note the **casual** language since this was an experiment. If it wasn't we would have written: "the pygmalion treatment platoon has on average a score 7.2 points higher than the control platoons"

Case1301: Fish Grazing

Influence on seaweed regeneration of certain grazers.

Scrape rocks clean, then exclude certain grazers.

Come back in 4 weeks and measure the % of rock covered in seaweed.

8 blocks covering different tidal conditions (e.g. just below high tide exposed to surf, mid tide exposed, ...).

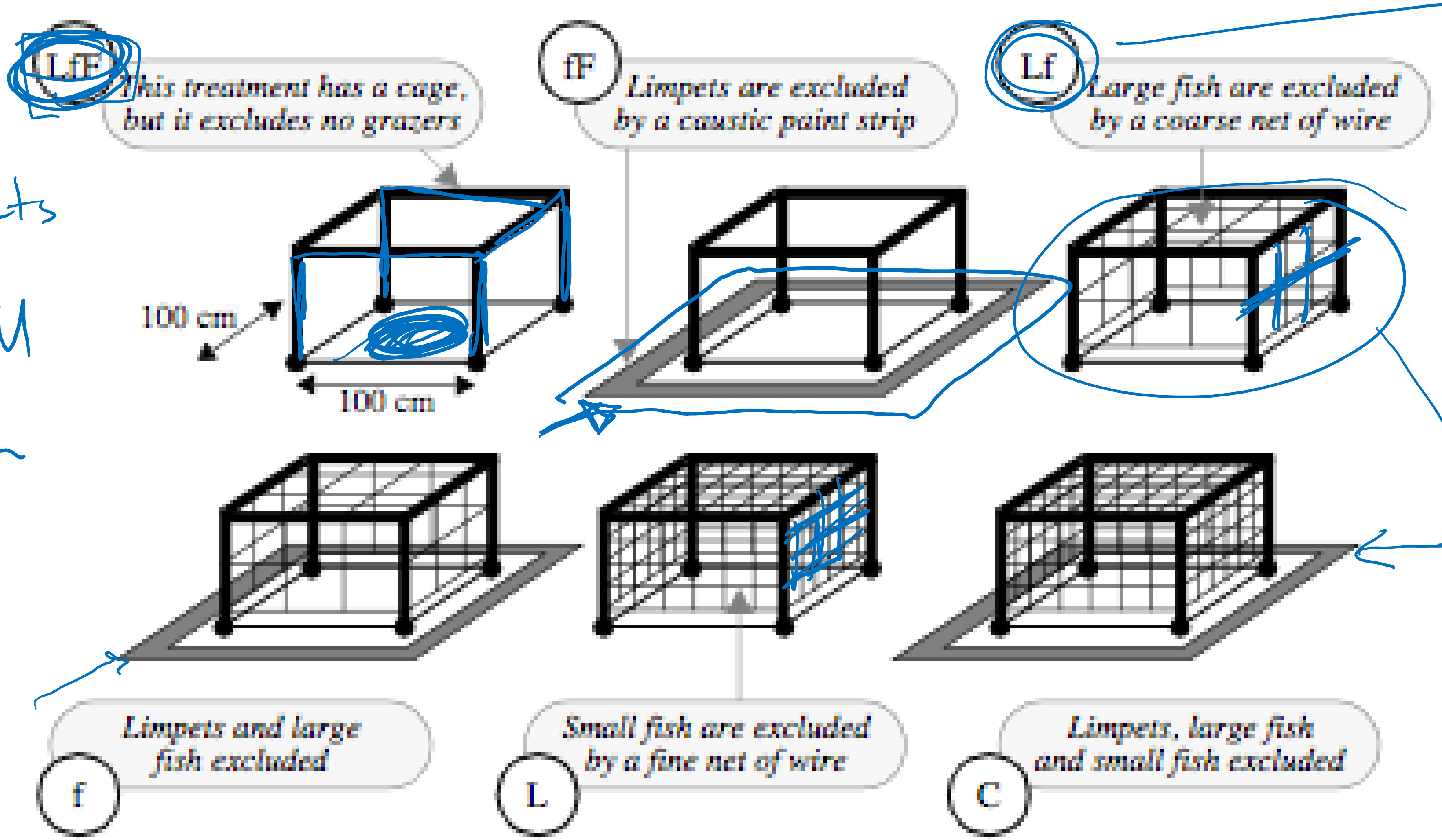
Display 13.1 p. 376

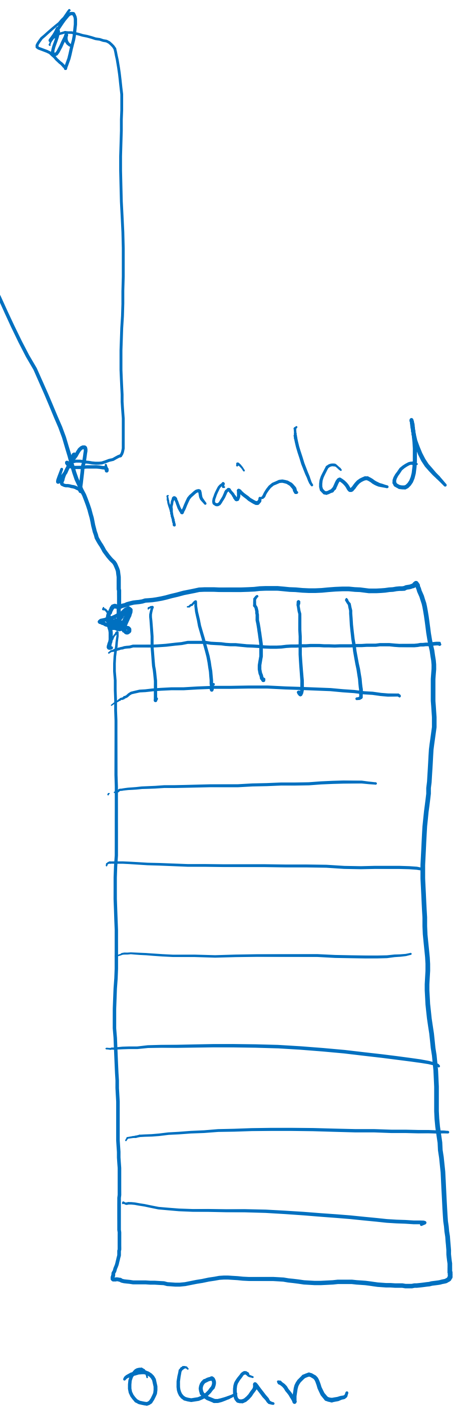
Six treatments excluding three kinds of intertidal grazers from regenerating seaweed on the Oregon coast

L limpets

f small fish

F big fish





Each block divided into twelve plots, treatments randomly assigned to plot, 2 plots per treatment within each block

Your turn

What are the two factors?

Different tidal zones, not in the experimenters control, BLOCKS

Cages, experimenters control, TREAT

How many levels do they each have?

8

6

experimental technique, to reduce variation.

Questions of interest

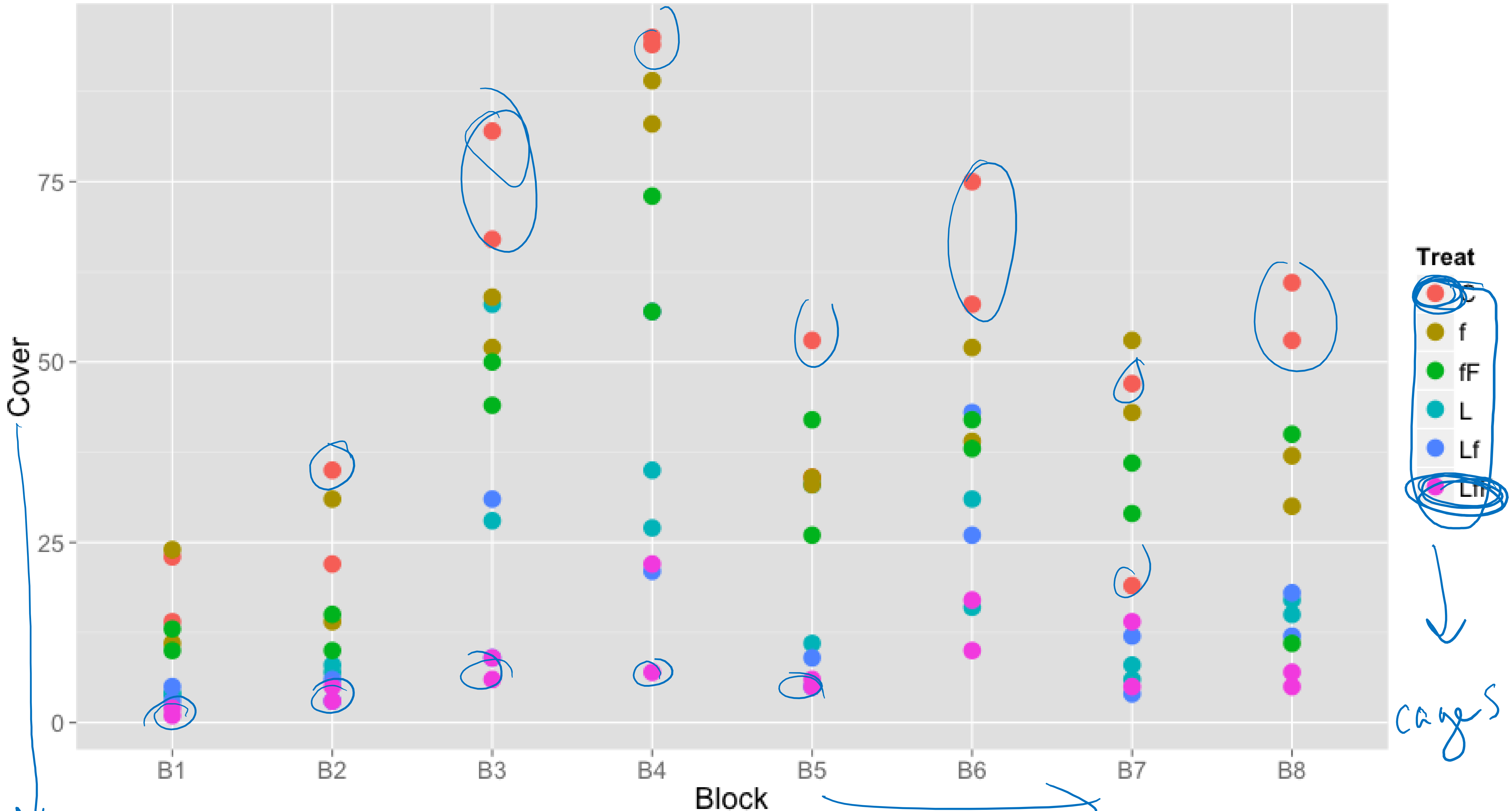
Which grazer consumes the most seaweed?

Do the different grazers impact each other?

Are grazing effects similar in all microhabitats?

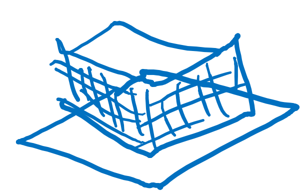
tidal zones

shape = Treat, geom = "jitter"



% rock covered in seaweed

tidal zones

C =  = no animals

Strategy

Start with saturated model

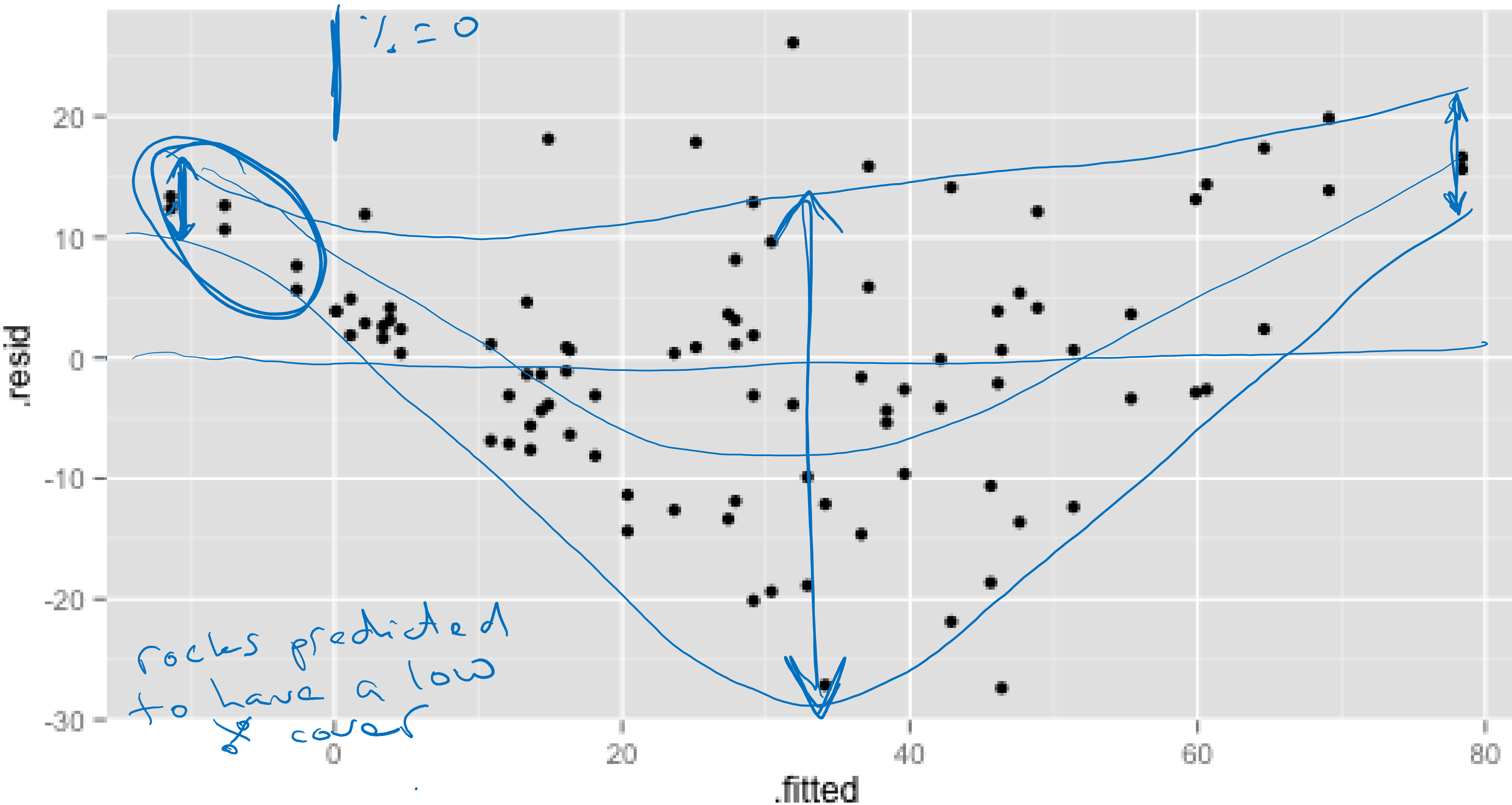
Check fit

Is a simpler model adequate?

Answer questions of interest about means

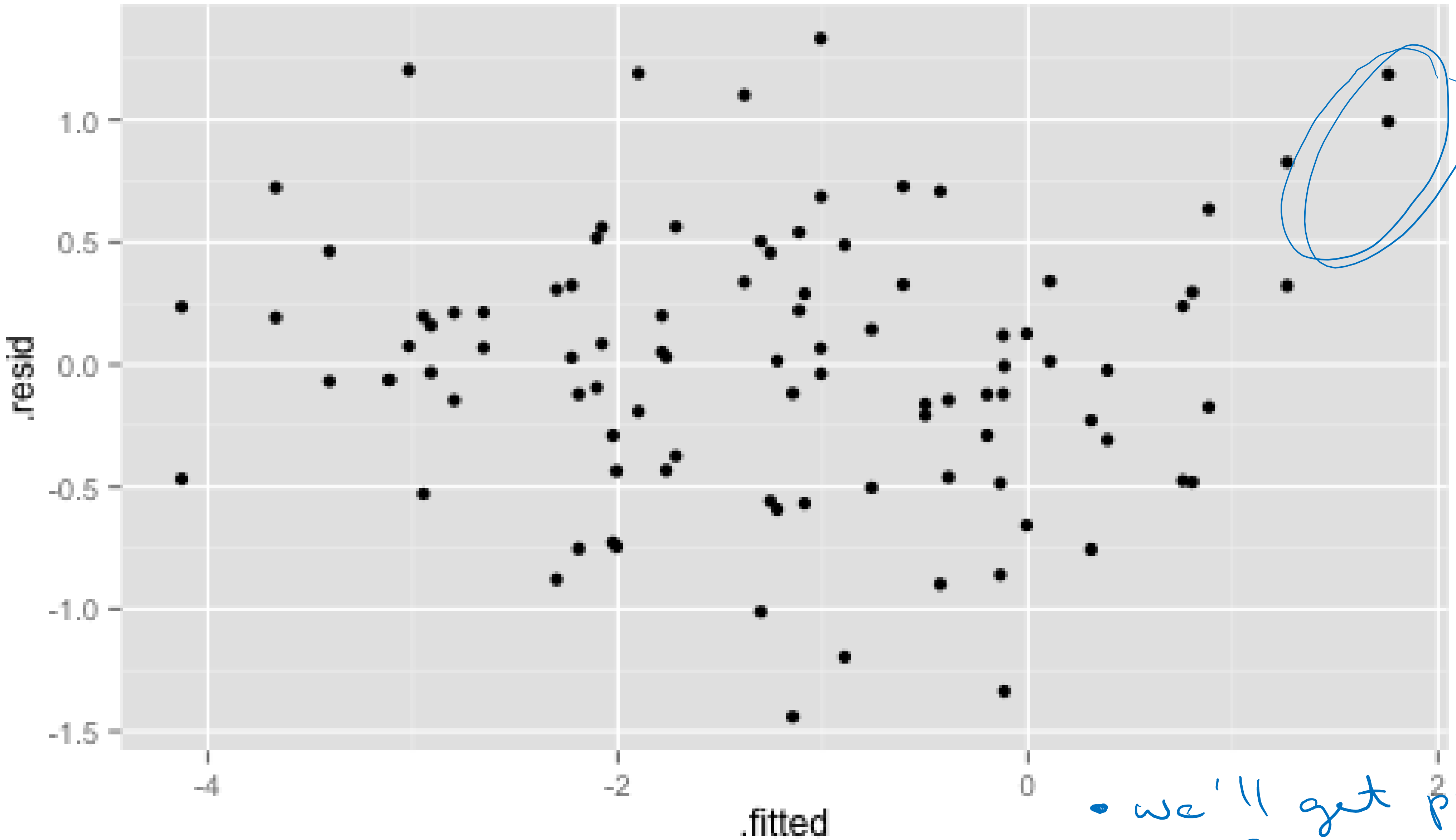
Residuals of saturated model

μ{ Cover | Block, Treat} =
BLOCK + TREAT + BLOCK:TREAT



Transform and try again

μ{ log(Cover/(100-Cover)) | Block, Treat} =
BLOCK + TREAT + BLOCK:TREAT



• we'll get predicted values for cover between 0% & 100%.

log(Cover/(100-Cover)) : log(recovery ratio)

Saturated model

Full model: μ{ log(Cover/(100-Cover)) | Block, Treat} = BLOCK + TREAT + BLOCK:TREAT

BLOCK + TREAT

Extra SS F-test

Display 13.10 **p. 385**

Analysis of variance for the log of the seaweed regeneration ratio; non-additive model

	Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p-value
1	Between Groups	188.4622	47	4.0098	13.2407	<0.0001
2	Blocks	76.2386	7	10.8912	35.9634	<0.0001
3	Treatments	96.9932	5	19.3986	64.0554	<0.0001
4	Interactions	15.2304	35	0.4352	1.4369	0.1209
	Within Groups	14.5364	48	0.3028		
	Total	202.9986	95			

R-squared = 92.84% **adj. R-squared = 85.83%** **Estimated SD = 0.5503**

no evidence for interactions

response ~ (Blocks - 1):Treat

In R:

1. μ{ *log(C/(1- C))* | blocks, treat} = μ
2. μ{ *log(C/(1- C))* | blocks, treat} = TREAT + BLOCKS x TREAT
3. μ{ *log(C/(1- C))* | blocks, treat} = BLOCKS + BLOCKS x TREAT
4. μ{ *log(C/(1- C))* | blocks, treat} = BLOCKS + TREAT

Additive model

Full model: μ{ log(Cover/(100-Cover)) | Block, Treat} = BLOCK + TREAT

Display 13.11 **p. 387**

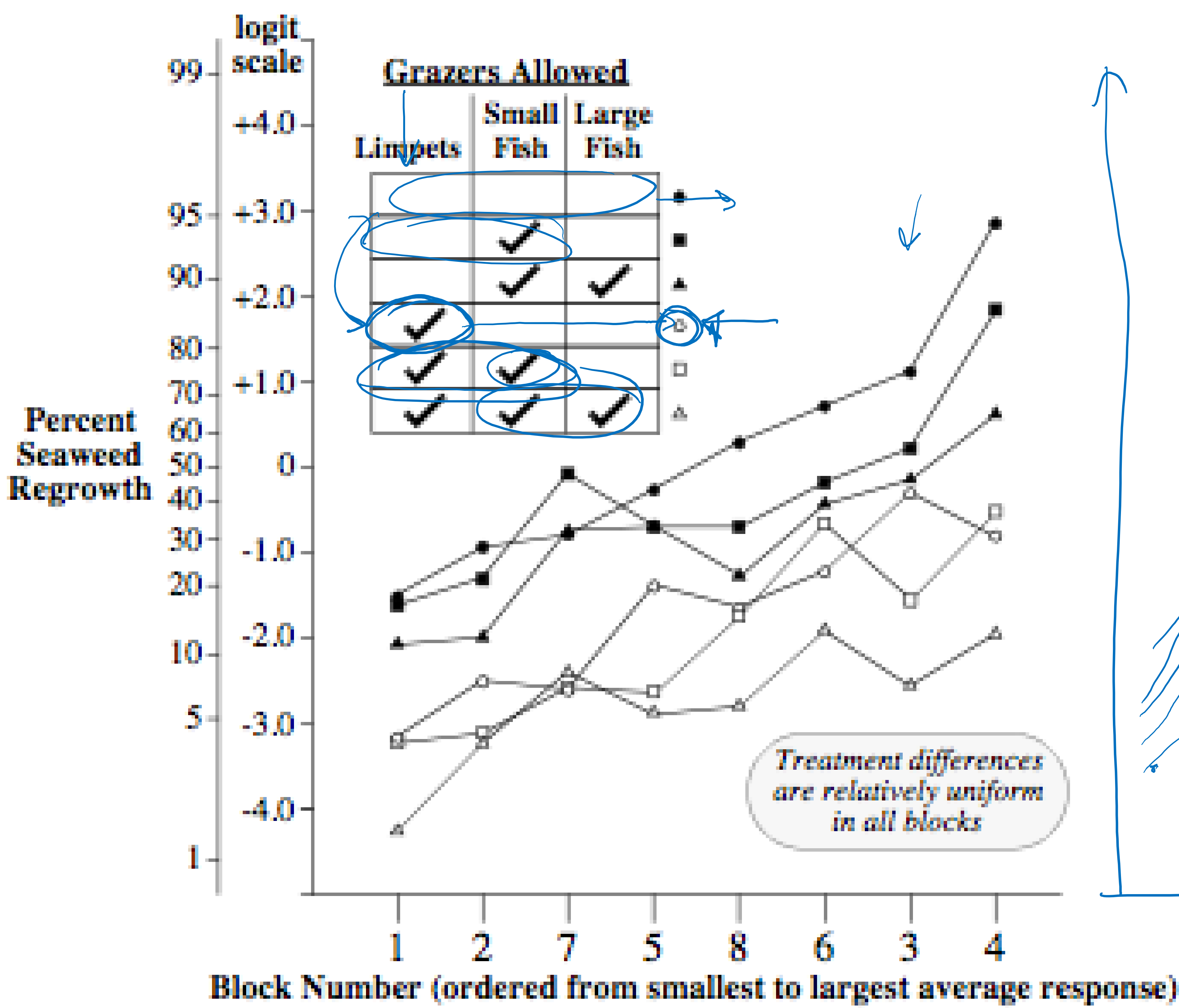
Analysis of variance for the log of the seaweed regeneration ratio; additive model

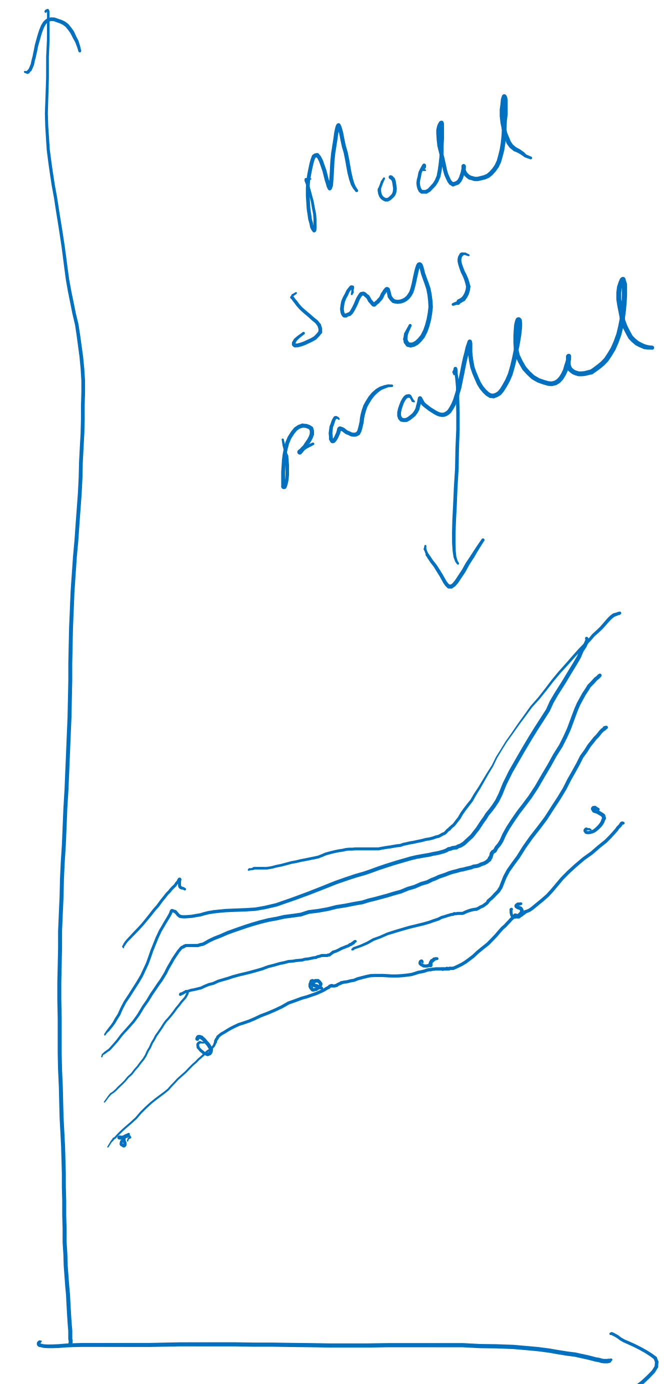
Source of Variation	Sum of Squares	df	Mean Square	F-Statistic	p-value
1 Model	173.2318	12	14.4360	40.2520	<0.0001
2 Blocks	76.2386	7	10.8912	30.3684	<0.0001
3 Treatments	96.9932	5	19.3986	54.0900	<0.0001
Residual	29.7668	83	0.35864		
Total	202.9986	95			

R-squared = 85.34% **adj. R-squared =** 83.22% **Estimated SD =** 0.5989

1. μ{ *log(C/(1- C))* | blocks, treat} = μ
2. μ{ *log(C/(1- C))* | blocks, treat} = TREAT
3. μ{ *log(C/(1- C))* | blocks, treat} = BLOCKS

Averages of the log of the seaweed regeneration ratio versus block number, with code for treatment





Model says parallel

Estimating effects

not of the treatments, but of the animals

Two approaches:

Using averages over cell, rows and columns.

HARD, and only relevant for balanced data

Using indicator variables and multiple regression.

A regression approach

f, Ff, Lf, LFf, ⊔, L

Set up indicators:

sml = 1, small fish are present if f, Ff, Lf, LFf

big = 1, large fish are present if Ff, LFf

limp =1, limpets are present if L, Lf, LFf

Equivalent to the additive model (TREAT + BLOCK):

baseline?

BLOCK + sml + big + limp + sml x limp + big x limp

sml x big : can't estimate, since big fish always present with little fish.

```
Analysis of Variance Table

Model 1: log(Cover/(100 - Cover)) ~ Block + L + f + F
Model 2: log(Cover/(100 - Cover)) ~ Block + L + f + F + L:F + L:f
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     85 29.996
2     83 29.767  2   0.22928 0.3197 0.7273
```

no evidence for animal interactions

```
Call:
lm(formula = log(Cover/(100 - Cover)) ~ Block + L + f + F, data = case1301)

Residuals:
     Min       1Q   Median       3Q      Max
-1.47682 -0.40585  0.03001  0.33617  1.30143

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -1.2545     0.2011  -6.238 1.66e-08 ***
BlockB2       0.4600     0.2425   1.897  0.06127 .
BlockB3       2.1046     0.2425   8.678 2.42e-13 ***
BlockB4       2.9807     0.2425  12.291  < 2e-16 ***
BlockB5       1.2160     0.2425   5.014 2.87e-06 ***
BlockB6       2.0251     0.2425   8.350 1.11e-12 ***
BlockB7       1.1085     0.2425   4.571 1.64e-05 ***
BlockB8       1.3300     0.2425   5.484 4.19e-07 ***
L            -1.8288     0.1213 -15.082  < 2e-16 ***
f            -0.3933     0.1485  -2.648  0.00965 **
F            -0.6140     0.1485  -4.135 8.31e-05 ***
```

estimates of effects

There is no evidence that the grazing effects differ depending on microhabitat (extra SS F-test on interaction between grazers and blocks, p-value = 0.12).

There is no evidence that the different grazers impact each other (extra SS F-test on interactions between limpets and fish, p-value = 0.72).

Allowing limpets access to plots caused significant changes in the regeneration of seaweed (two sided p-value < 0.00001 from a t-test on the effect of limpets). It is estimated that the median regeneration ratio when limpets were present is estimated to be only 0.161 times as large as the median regeneration time when they are excluded (95% CI: 0.126 to 0.205).

exp(-1.82) = 0.161

... two more, one for small fish, one for big fish