

mfcurve: Visualizing Results From Multifactorial Designs

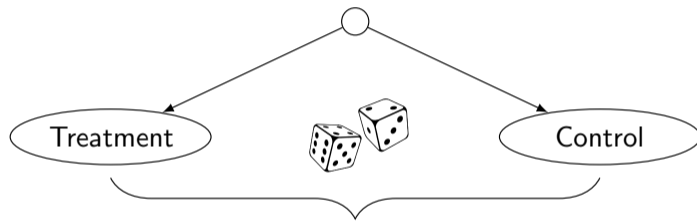
Daniel Krähler
LMU Munich

2023 German Stata Conference
Berlin | 16 June



The Classical Experimental Research Design

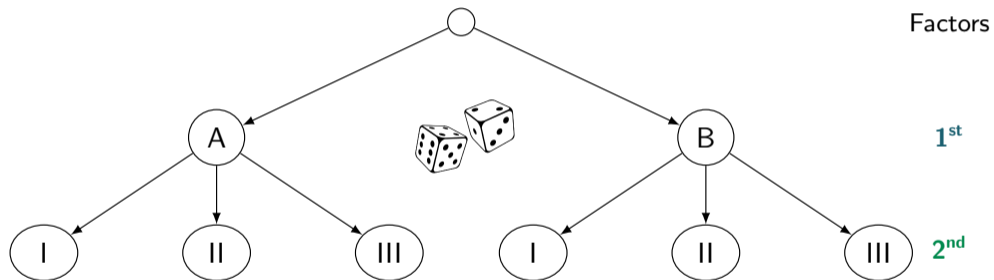
Estimand: Main effect of one particular treatment



$$ATE = E[Y^1] - E[Y^0]$$

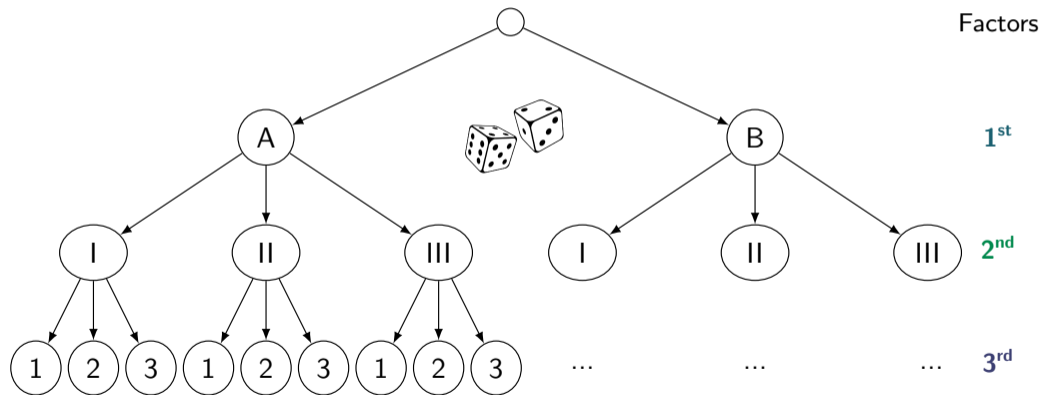
Multifactorial Research Designs

Estimand: Joint effect of several treatments



Multifactorial Research Designs

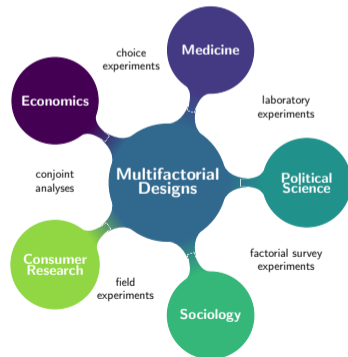
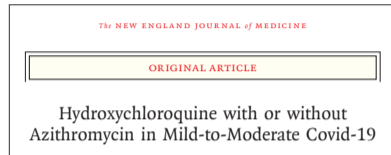
Estimand: Joint effect of several treatments



Multifactorial Research Designs Are...

... epistemologically useful,

... versatile & wide-spread,

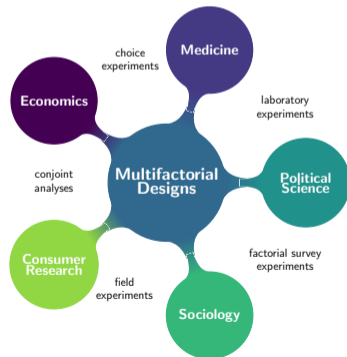
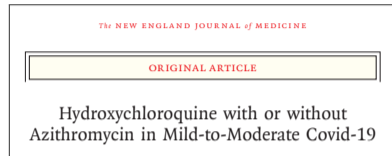


Multifactorial Research Designs Are...

... epistemologically useful,

... versatile & wide-spread,

... **but difficult to visualize**, due to the spiraling number of treatment conditions
→ Potential remedy: `mfcurve`



Command `mfcurve`: Basics

```
mfcurve depvar, factors(indepvar) [options]
```

Input

- Outcome, i.e. dependent variable
- Factors, i.e. independent variables

Command `mfcurve`: Basics

```
mfcurve depvar, factors(indepvar) [options]
```

Input

- Outcome, i.e. dependent variable
- Factors, i.e. independent variables

Example: Average hourly wage across groups defined by race, region, and union membership (descriptive)

Command `mfcurve`: Basics

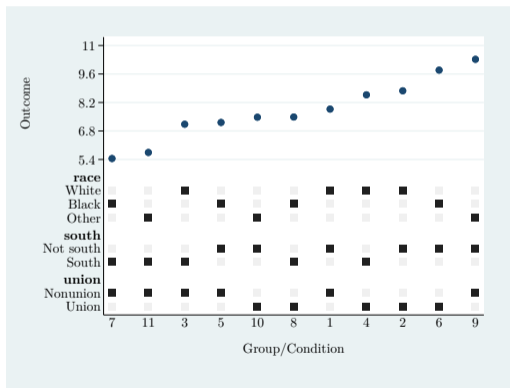
```
mfcurve depvar, factors(indepvar) [options]
```

Input

- Outcome, i.e. dependent variable
- Factors, i.e. independent variables

Example: Average hourly wage across groups defined by race, region, and union membership (descriptive)

```
sysuse nlsw88, clear  
mfcurve wage, ///  
    factors(race south union)
```



► mechanics of `mfcurve`

Command `mfcurve`: Options

Options	Description
groupvar (<i>varname</i>)	specify group identifier
test (<i>mean zero</i>)	perform significance tests
level (#)	set confidence level
show (<i>show_options</i>)	add elements to the plot (see below)
boxplot	use boxplots instead of point estimates ▶ demo
style... (<i>marker_options</i>)	customize graph elements ▶ demo
<i>twoway_options</i>	twoway options, other than <code>by()</code>

<i>show_options</i>	Description
mean	add a horizontal mean line
sig	highlight significant estimates
ci_regular	add CIs, using solid lines
ci_gradient	add CIs, using color gradients ▶ demo
groupsize	add case numbers to the x-axis ▶ demo

Command `mfcurve`: Application

Working Paper:

Care to Share? Experimental Evidence on Code Sharing Behavior in the Social Sciences



Command `mfcurve`: Application

Working Paper:

Care to Share? Experimental Evidence on Code Sharing Behavior in the Social Sciences

- ① What determines researchers' willingness to share analysis code upon request?
- ✉ Field experiment including more than 1,200 researchers across the social sciences
- 💡 Experimental variation of the code request's wording



Command `mfcurve`: Application

Working Paper:

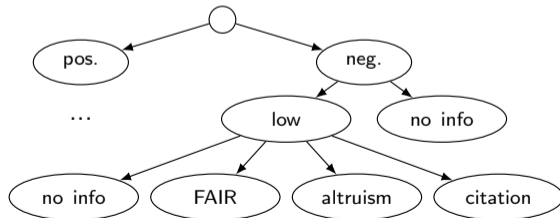
Care to Share? Experimental Evidence on Code Sharing Behavior in the Social Sciences

- 🔗 What determines researchers' willingness to share analysis code upon request?
- ✉ Field experiment including more than 1,200 researchers across the social sciences
- 💡 Experimental variation of the code request's wording

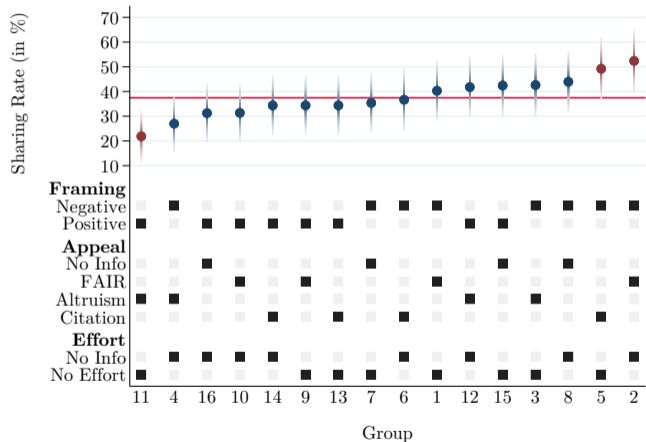
Framing (2 levels)

Effort (2 levels)

Appeal (4 levels)

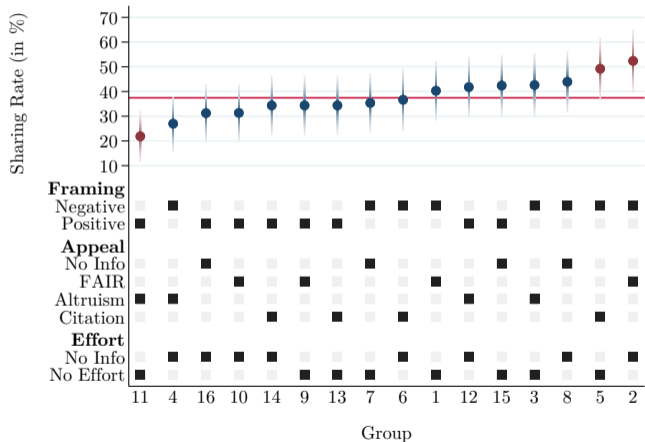


Command mfcurve: Application



► mfcurve vs. specification curves

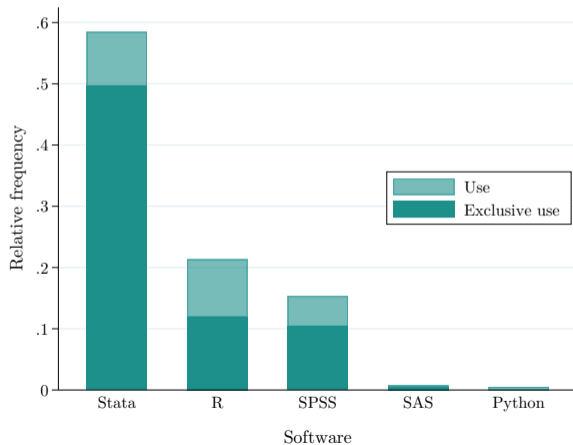
Command mfcurve: Application



- + Transparent reporting of all results
- + Ranking of interventions by effectiveness
- ~ Readability
- ~ Main effects
- ~ Interaction effects

By The Way

Software usage among authors who shared their code ($n = 385$)



Summary

- Multifactorial research designs are popular across disciplines
- They are notoriously difficult to visualize
- `mfcurve` may provide a solution to handle multidimensionality
- May also be used for simple n -dimensional description

- Installation from GitLab:

```
net install mfcurve, from("https://tinyurl.com/mfcurve")
```



Thanks for your time!

Your comments and suggestions are appreciated.

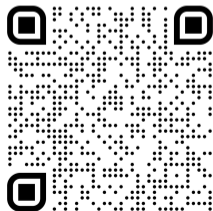


@dkraehmer@sciences.social



daniel.kraehmer@soziologie.uni-muenchen.de

Project Homepage:



Mechanics of `mfcurve` I

Preparatory work:

race	south	union
Black	Not South	Union
Black	Not South	Nonunion
Black	South	Union
Black	South	Nonunion
White	Not South	Union
White	Not South	Nonunion
...

Mechanics of `mfcurve` I

Preparatory work:

- Define distinct groups in the n -dimensional space defined by variables in `factor(...)`

race	south	union	group
Black	Not South	Union	1
Black	Not South	Nonunion	2
Black	South	Union	3
Black	South	Nonunion	4
White	Not South	Union	5
White	Not South	Nonunion	6
...

Mechanics of `mfcurve` I

Preparatory work:

- Define distinct groups in the n -dimensional space defined by variables in `factor(...)`
- Calculate mean outcome by group

race	south	union	group	wage
Black	Not South	Union	1	9.79
Black	Not South	Nonunion	2	7.22
Black	South	Union	3	7.49
Black	South	Nonunion	4	5.45
White	Not South	Union	5	8.77
White	Not South	Nonunion	6	7.87
...

Mechanics of `mfcurve` I

Preparatory work:

- Define distinct groups in the n -dimensional space defined by variables in `factor(...)`
- Calculate mean outcome by group
- Rank groups by mean outcome

race	south	union	group	wage	rank
Black	Not South	Union	1	9.79	1
Black	Not South	Nonunion	2	7.22	6
Black	South	Union	3	7.49	4
Black	South	Nonunion	4	5.45	5
White	Not South	Union	5	8.77	2
White	Not South	Nonunion	6	7.87	3
...

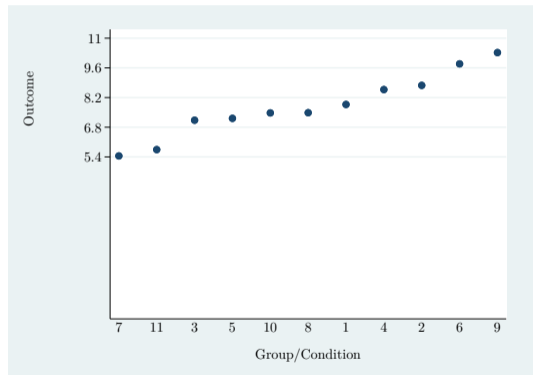
Mechanics of `mfcurve` II

- Keep only one observation per group to increase efficiency (Stata Tip 19)
- Generate indicator variables, signaling each level's presence/absence

<code>race</code>	<code>race_d_white</code>	<code>race_d_black</code>	<code>race_d_other</code>	<code>south</code>	<code>union</code>	<code>wage</code>
Black	0	1	0	...	Union	9.79
Black	0	1	0	...	Nonunion	7.22
Black	0	1	0	...	Union	7.49
Black	0	1	0	...	Nonunion	5.45
White	1	0	0	...	Union	8.77
White	1	0	0	...	Nonunion	7.87
...

Mechanics of `mfcurve` III

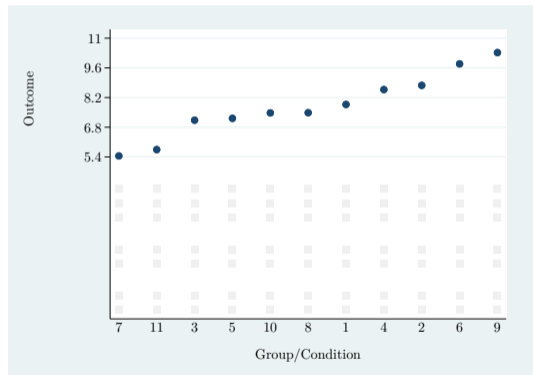
Overlay plots:



scatter outcome rank (using *group* labels!)

Mechanics of `mfcurve` III

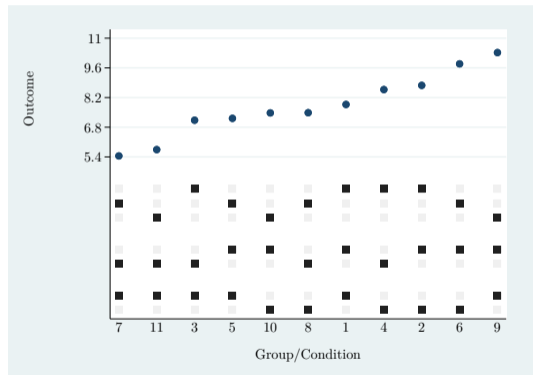
Overlay plots:



scatter outcome rank (using *group* labels!)
+ indicators

Mechanics of `mfcurve` III

Overlay plots:



scatter outcome rank (using *group* labels!)

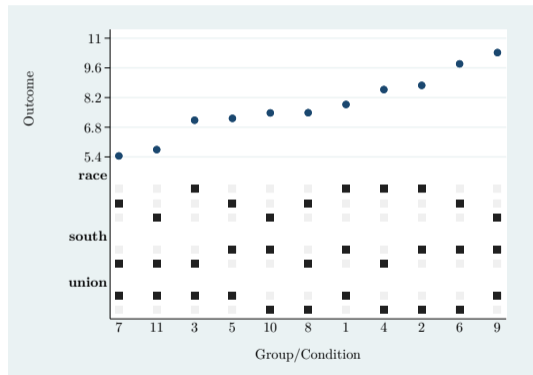
+ indicators

+ active indicators

(based on dummies == 1)

Mechanics of `mfcurve` III

Overlay plots:

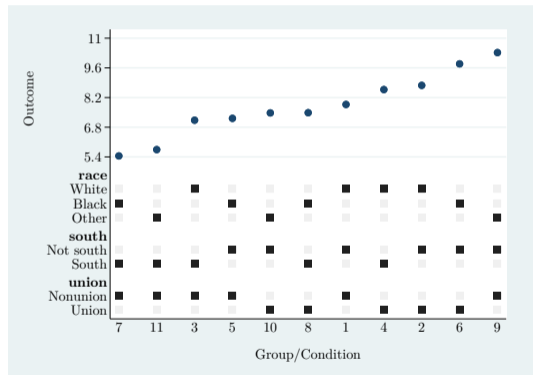


scatter outcome rank (using *group* labels!)

- + indicators
- + active indicators
(based on dummies == 1)
- + ylabels based on *variable* labels

Mechanics of `mfcurve` III

Overlay plots:

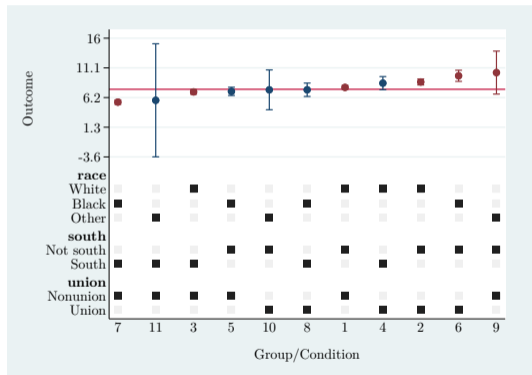


scatter outcome rank (using *group* labels!)

- + indicators
- + active indicators
(based on dummies == 1)
- + ylabels based on *variable* labels
- + ylabels based on *value* labels

Mechanics of `mfcurve` III

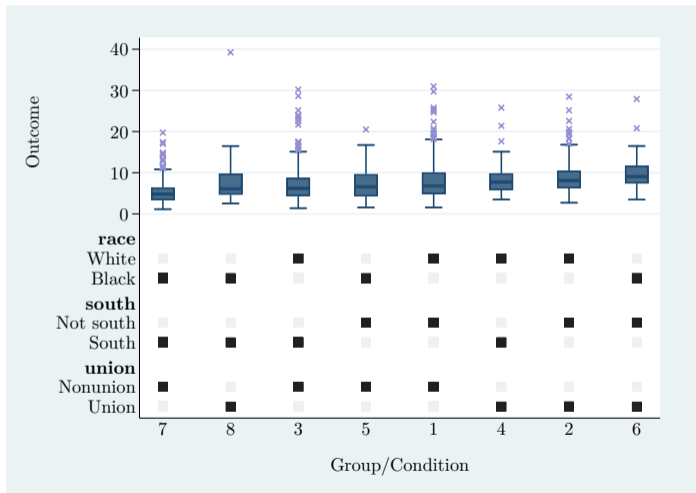
Overlay plots:



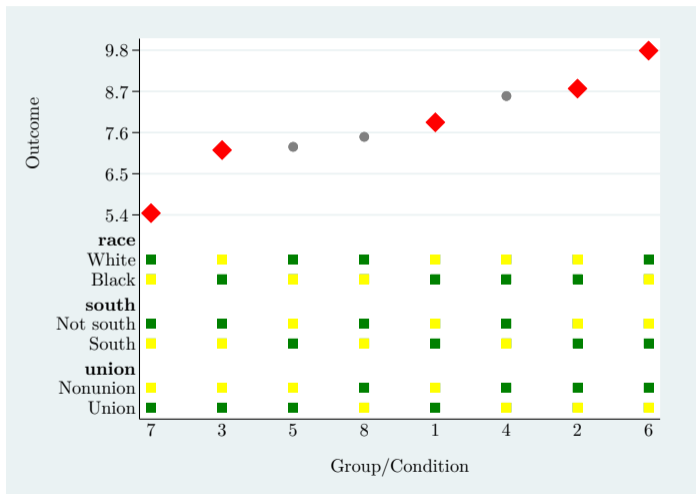
scatter outcome rank (using *group* labels!)

- + indicators
- + active indicators
(based on dummies == 1)
- + ylabels based on *variable* labels
- + ylabels based on *value* labels
- + custom graph types
(rcap, rspike, etc.)

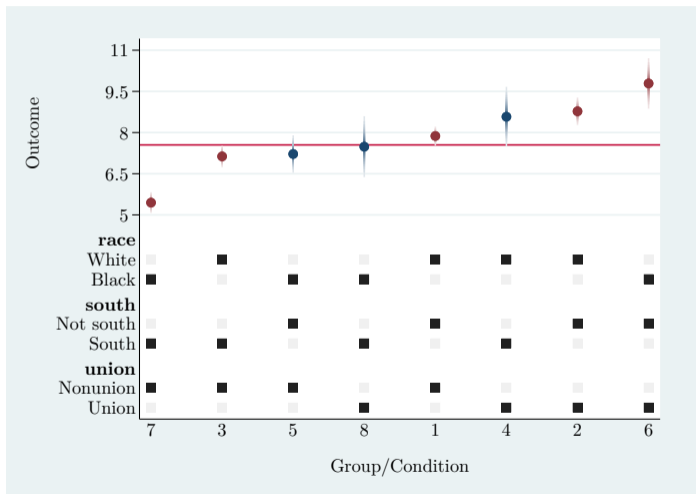
Demo: boxplot



Demo: style... (*marker_options*)

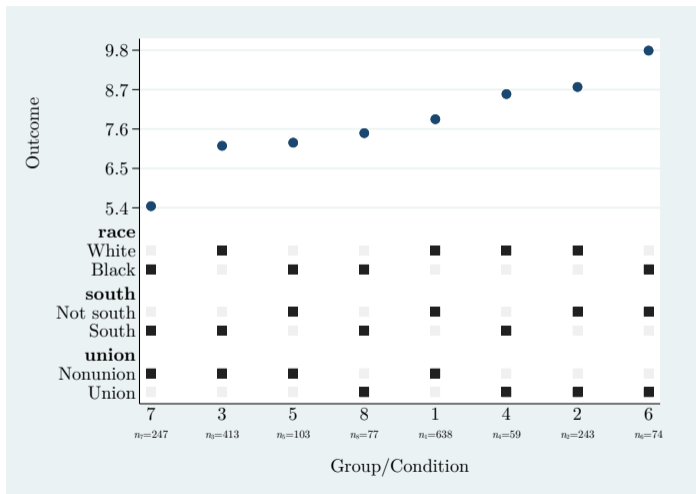


Demo: show(mean ci_gradient)

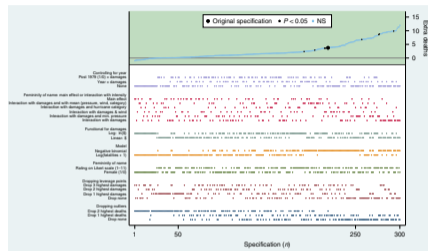
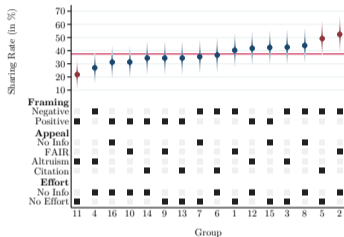


◀ back

Demo: show (groupsize)



Comparison: mfcurve vs. specification curves



Graphical differences

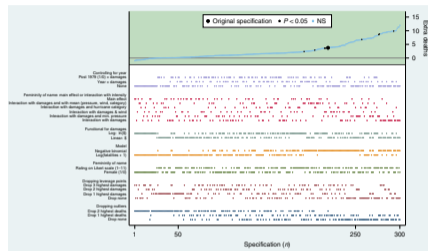
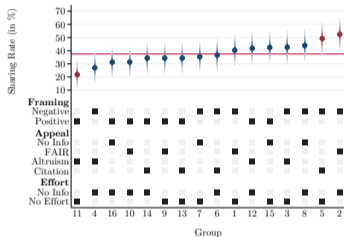
upper panel:
lower y-axis:
x-axis:

mean values
levels
treatment specifications

multivariate regression coefficients
model ingredients
model specifications

◀ back

Comparison: `mfcurve` vs. specification curves



Conceptual differences

analytical units:
applications:
computation:

disjunct subsamples
inferential & descriptive
parsimonious

(overlapping) models
inferential
intensive

◀ back