Uncertainty visualization as a moral imperative

Matthew Kay
Assistant Professor
~~School of Information, University of Michigan~~
Computer Science and Communication, Northwestern University

Collective

What happens when we ignore uncertainty?

A mixed-design ANOVA with sex of face (male, female) as a within-subjects factor and self-rated attractiveness (low, average, high) and oral contraceptive use (true, false) as between-subjects factors revealed a main effect of sex of face, $F(1, 1276) = 1372$, $p < .001$, $\eta_p^2 = .52$. This was qualified by interactions between sex of face and SRA, $F(2, 1276) = 6.90$, $p = .001$, $\eta_p^2 = .011$, and between sex of face and oral contraceptive use, $F(1, 1276) = 5.02$, $p = .025$, $\eta_p^2 = .004$. The predicted interaction among sex of face, SRA and oral contraceptive use was not significant, $F(2, 1276) = 0.06$, $p = .94$, $\eta_p^2 < .001$. All other main effects and interactions were non-significant and irrelevant to our hypotheses, all $F \leq 0.94$, $p \geq .39$, $\eta_p^2 \leq .001$.

A mixed-design ANOVA with sex of face (male, female) as a within-subjects factor and self-rated attractiveness (low, average, high) and oral contraceptive use (true, false) as between-subjects factors revealed a main effect of sex of face, $F(1, 1276) = 1372$, $p < .001$, $\eta_p^2 = .52$. This was qualified by interactions between sex of face and SRA, $F(2, 1276) = 6.90$, $p = .001$, $\eta_p^2 = .011$, and between sex of face and oral contraceptive use, $F(1, 1276) = 5.02$, $p = .025$, $\eta_p^2 =$ [illegible]. The predicted interaction among sex of face, SRA and oral contraceptive use was not significant, $F(2, 1276) = 0.06$, $p = .94$, $\eta_p^2 < .001$. All other main effects and interactions were not significant and irrelevant to our hypotheses, all $F \leq$ [illegible], $p \geq .39$, $\eta_p^2 \leq .001$.

Alternatives...

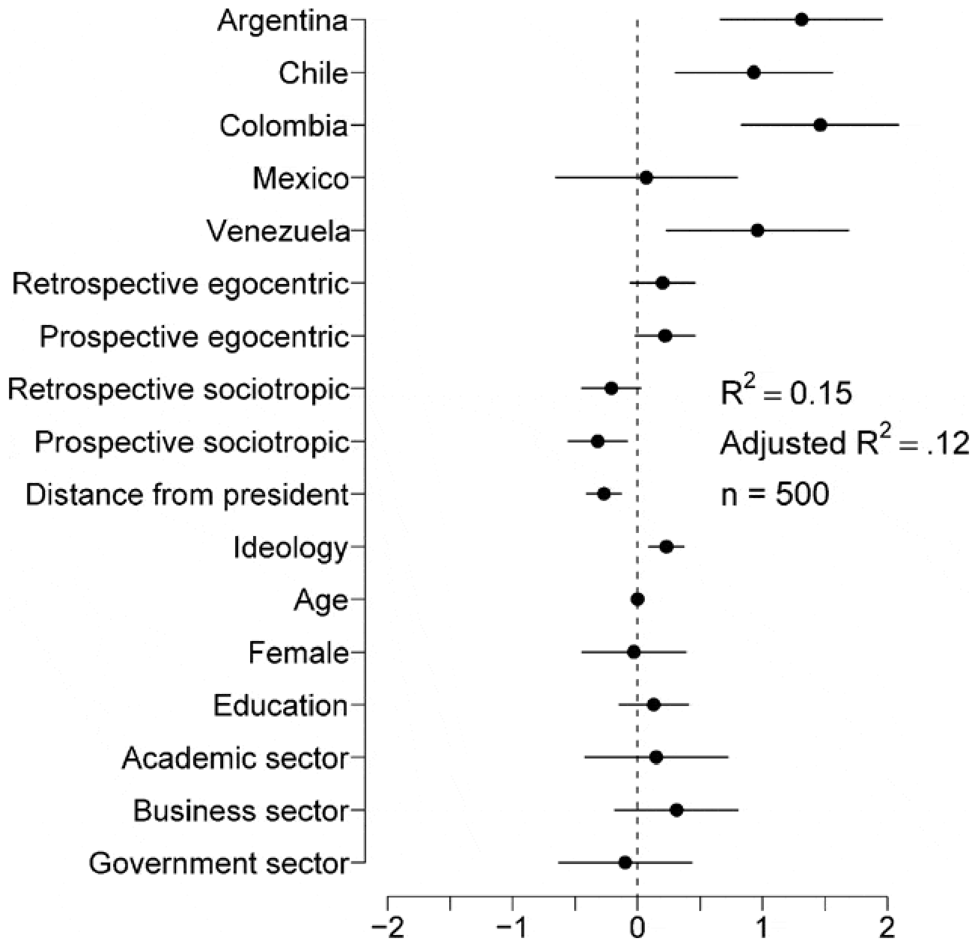
Variable	Coefficient (Standard Error)
Constant	.41 (.93)
Countries	
Argentina	1.31 (.33)**B,M
Chile	.93 (.32)**B,M
Colombia	1.46 (.32)**B,M
Mexico	.07 (.32)A,CH,CO,V
Venezuela	.96 (.37)**B,M
Threat	
Retrospective egocentric economic perceptions	.20 (.13)
Prospective egocentric economic perceptions	.22 (.12)#
Retrospective sociotropic economic perceptions	–.21 (.12)#
Prospective sociotropic economic perceptions	–.32 (.12)*
Ideological distance from president	–.27 (.07)**
Ideology	
Ideology	.23 (.07)**
Individual Differences	
Age	.00 (.01)
Female	–.03 (.21)
Education	.13 (.14)
Academic Sector	.15 (.29)
Business Sector	.31 (.25)
Government Sector	–.10 (.27)
R^2	.15
Adjusted R^2	.12
N	500

**p < .01, *p < .05, #p < .10 (twotailed)

Alternatives…

Variable	Coefficient (Standard Error)
Constant	.41 (.93)
Countries	
Argentina	1.31 (.33)^{**B,M}
Chile	.93 (.32)^{**B,M}
Colombia	1.46 (.32)^{**B,M}
Mexico	.07 (.32)^{A,CH,CO,V}
Venezuela	.96 (.37)^{**B,M}
Threat	
Retrospective egocentric economic perceptions	.20 (.13)
Prospective egocentric economic perceptions	.22 (.12)$^{\#}$
Retrospective sociotropic economic perceptions	–.21 (.12)$^{\#}$
Prospective sociotropic economic perceptions	–.32 (.12)*
Ideological distance from president	–.27 (.07)**
Ideology	
Ideology	.23 (.07)**
Individual Differences	
Age	.00 (.01)
Female	–.03 (.21)
Education	.13 (.14)
Academic Sector	.15 (.29)
Business Sector	.31 (.25)
Government Sector	–.10 (.27)
R^2	.15
Adjusted R^2	.12
N	500

$^{**}p < .01$, $^{*}p < .05$, $^{\#}p < .10$ (twotailed)

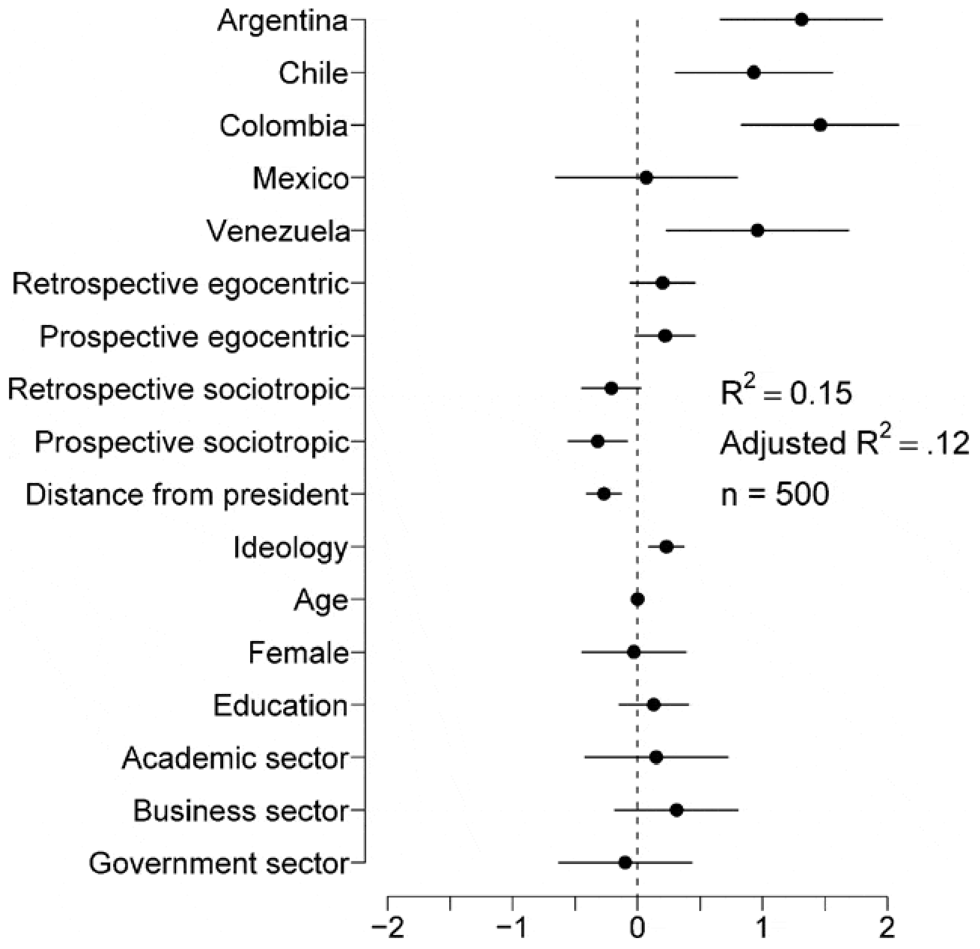


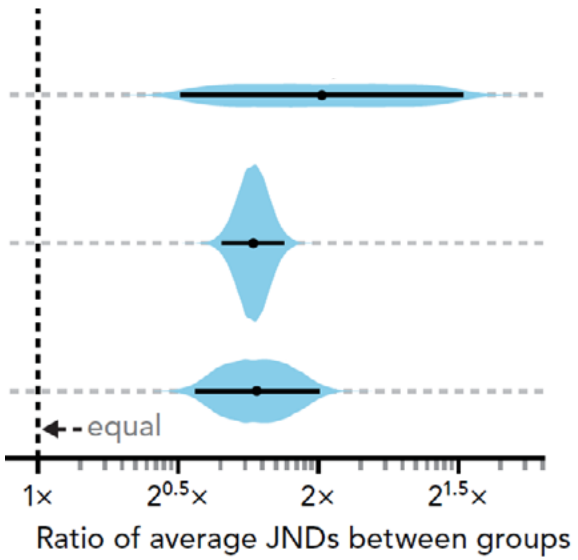
[Jonathan P Kastellec and Eduardo L Leoni. 2007. Using Graphs Instead of Tables in Political Science. Perspectives on politics 5, 4: 755–771]

Alternatives...

Variable	Coefficient (Standard Error)
Constant	.41 (.93)
Countries	
Argentina	1.31 (.33)$^{**\text{B,M}}$
Chile	.93 (.32)$^{**\text{B,M}}$
Colombia	1.46 (.32)$^{**\text{B,M}}$
Mexico	.07 (.32)$^{\text{A,CH,CO,V}}$
Venezuela	.96 (.37)$^{**\text{B,M}}$
Threat	
Retrospective egocentric economic perceptions	.20 (.13)
Prospective egocentric economic perceptions	.22 (.12)$^{\#}$
Retrospective sociotropic economic perceptions	–.21 (.12)$^{\#}$
Prospective sociotropic economic perceptions	–.32 (.12)*
Ideological distance from president	–.27 (.07)**
Ideology	
Ideology	.23 (.07)**
Individual Differences	
Age	.00 (.01)
Female	–.03 (.21)
Education	.13 (.14)
Academic Sector	.15 (.29)
Business Sector	.31 (.25)
Government Sector	–.10 (.27)
R^2	.15
Adjusted R^2	.12
N	500

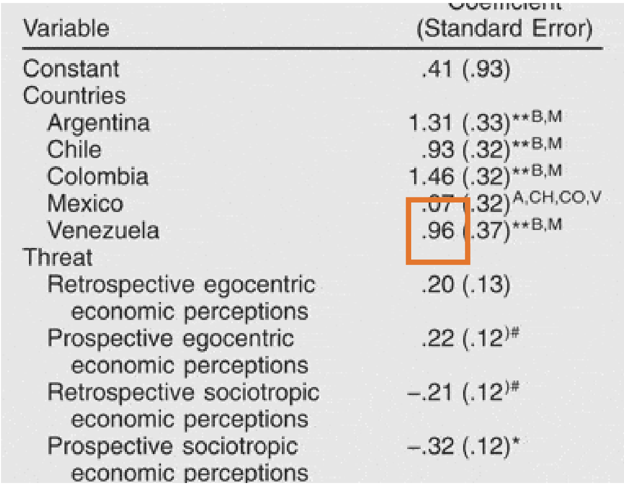
$^{**}p < .01$, $^{*}p < .05$, $^{\#}p < .10$ (twotailed)



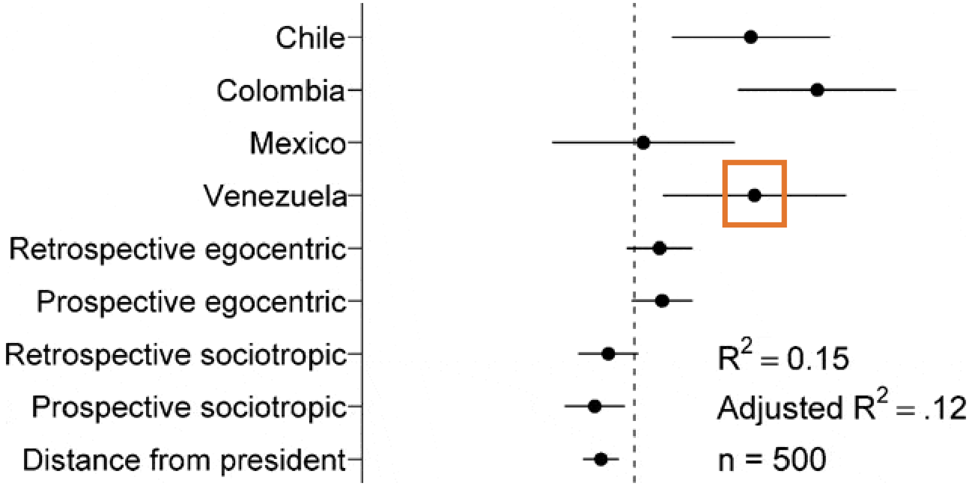


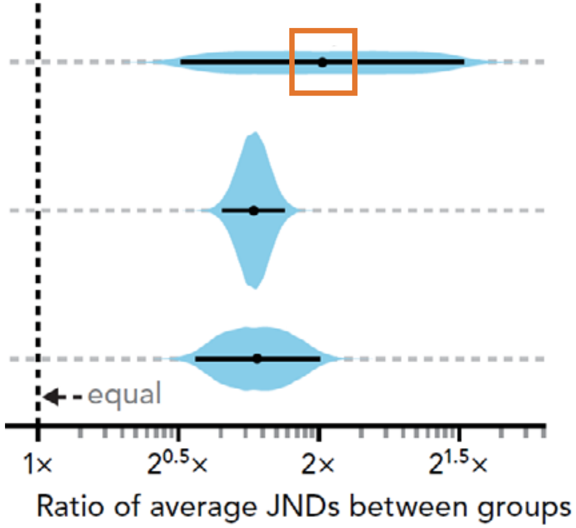
[Jonathan P Kastellec and Eduardo L Leoni. 2007. Using Graphs Instead of Tables in Political Science. Perspectives on politics 5, 4: 755–771]

How easy is it to ignore the uncertainty?



Variable	Coefficient (Standard Error)
Constant	.41 (.93)
Countries	
Argentina	1.31 (.33)**B,M
Chile	.93 (.32)**B,M
Colombia	1.46 (.32)**B,M
Mexico	.07 (.32)A,CH,CO,V
Venezuela	.96 (.37)**B,M
Threat	
Retrospective egocentric economic perceptions	.20 (.13)
Prospective egocentric economic perceptions	.22 (.12)#
Retrospective sociotropic economic perceptions	−.21 (.12)#
Prospective sociotropic economic perceptions	−.32 (.12)*





This contributes to dichotomania…

Dichotomania...

Predictions from 2016 presidential election

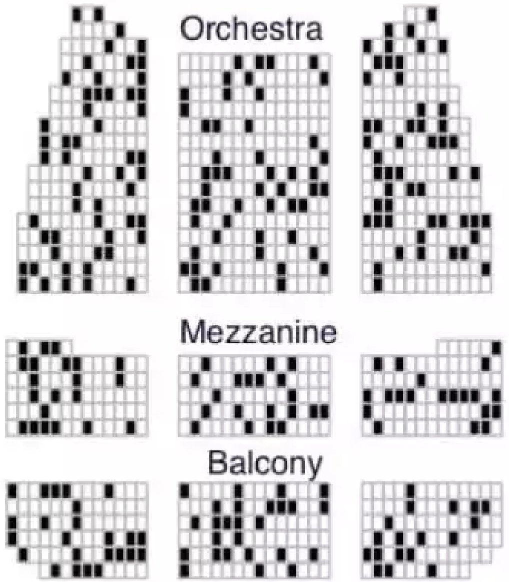
[Justin H. Gross, Washington Post, http://wapo.st/2fCYvDW]

FiveThirtyEight	NYT Upshot	HuffPo Pollster
28%	15%	2%

Predictions from 2016 presidential election

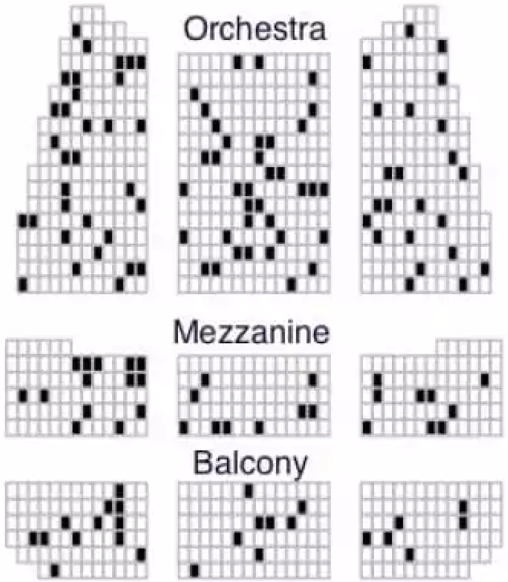
[Justin H. Gross, Washington Post, http://wapo.st/2fCYvDW]

FiveThirtyEight



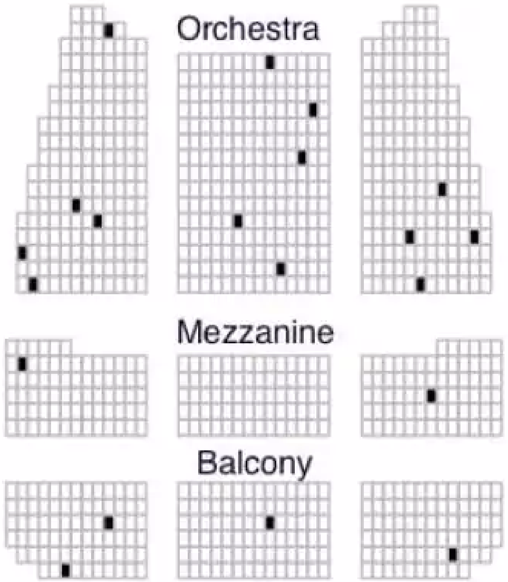
286 cases in 1,000

NYT Upshot



150 cases in 1,000

HuffPo Pollster



20 cases in 1,000

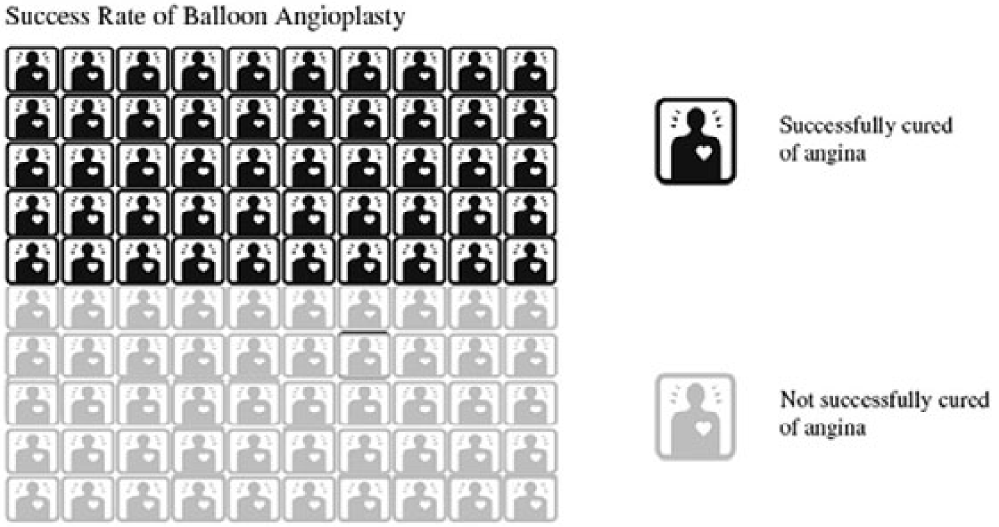
People are very good at ignoring uncertainty...

People are very good at ignoring uncertainty...

Especially when we provide bad uncertainty representations

Icon arrays in medical risk communication

[Figure from Fagerlin, Wang, Ubel. Reducing the influence of anecdotal reasoning on people's health care decisions: Is a picture worth a thousand statistics? Medical Decision Making 2005; 25:398–405]



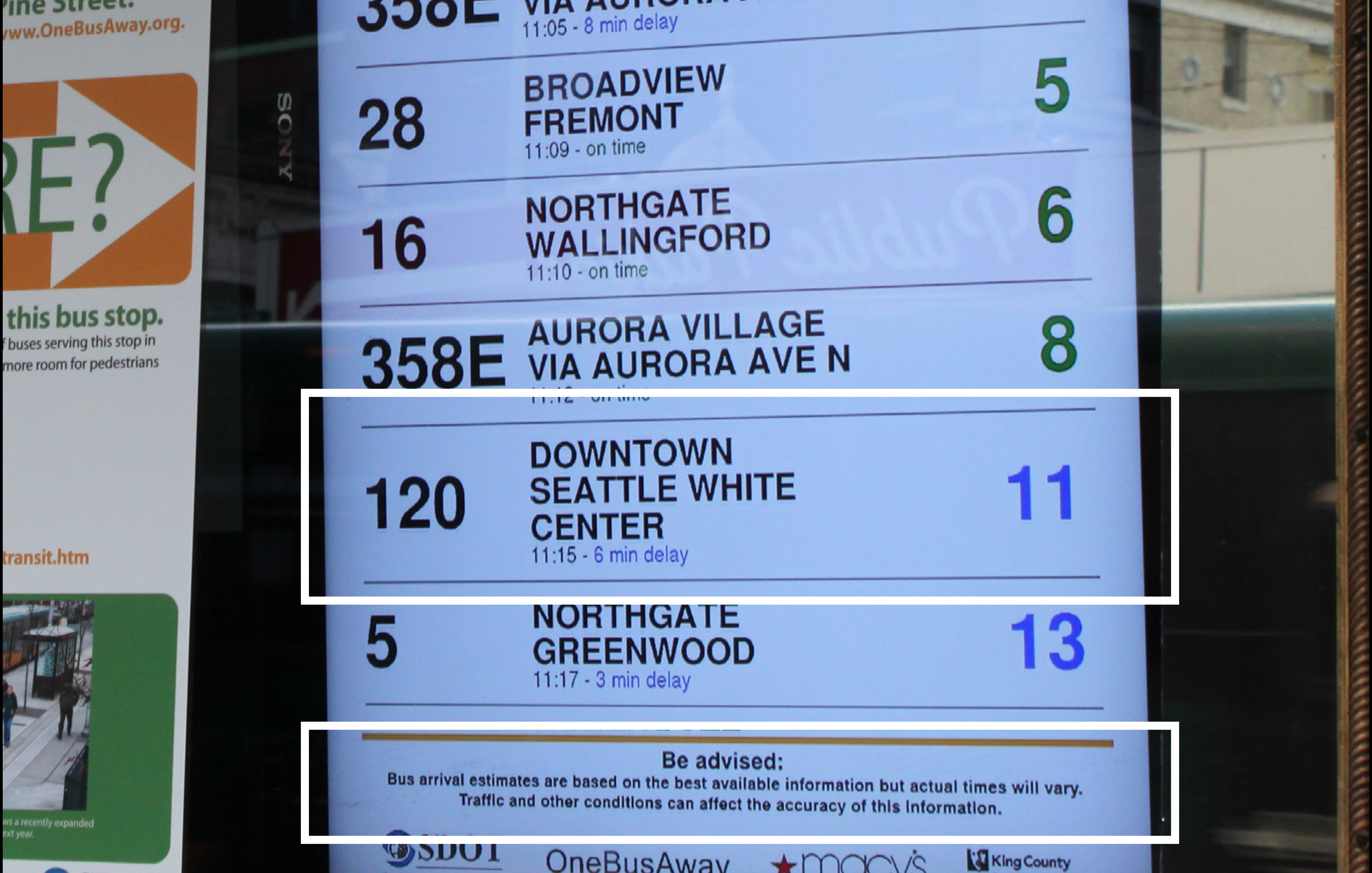
Frequency framing or discrete outcome visualization

What is an icon array for a continuous distribution?

What is an icon array for a continuous distribution?

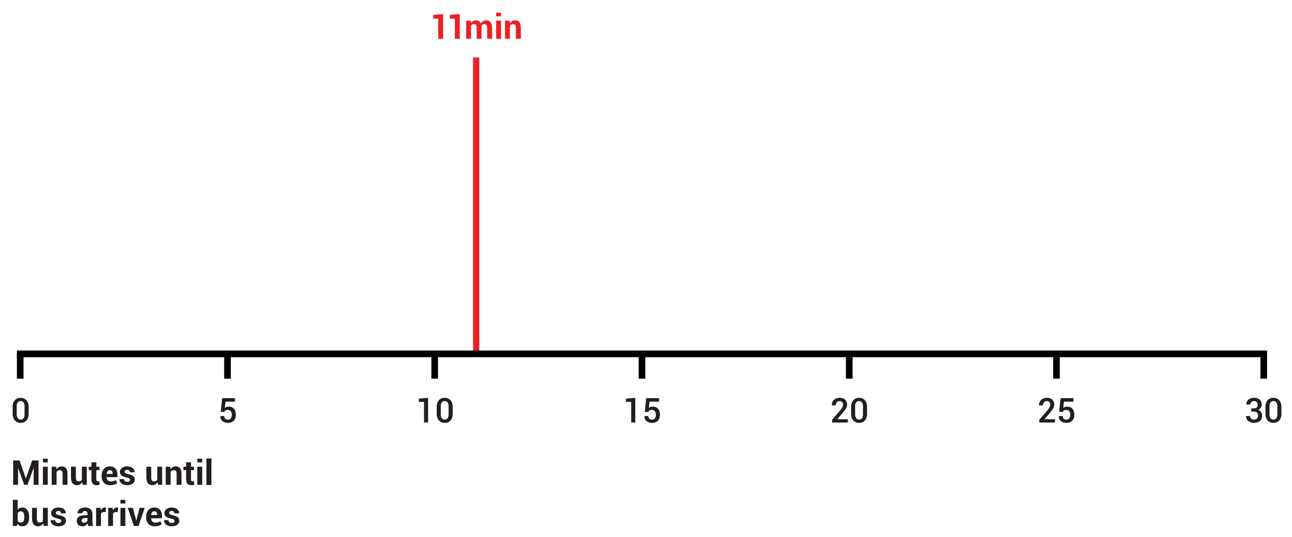
An example scenario...

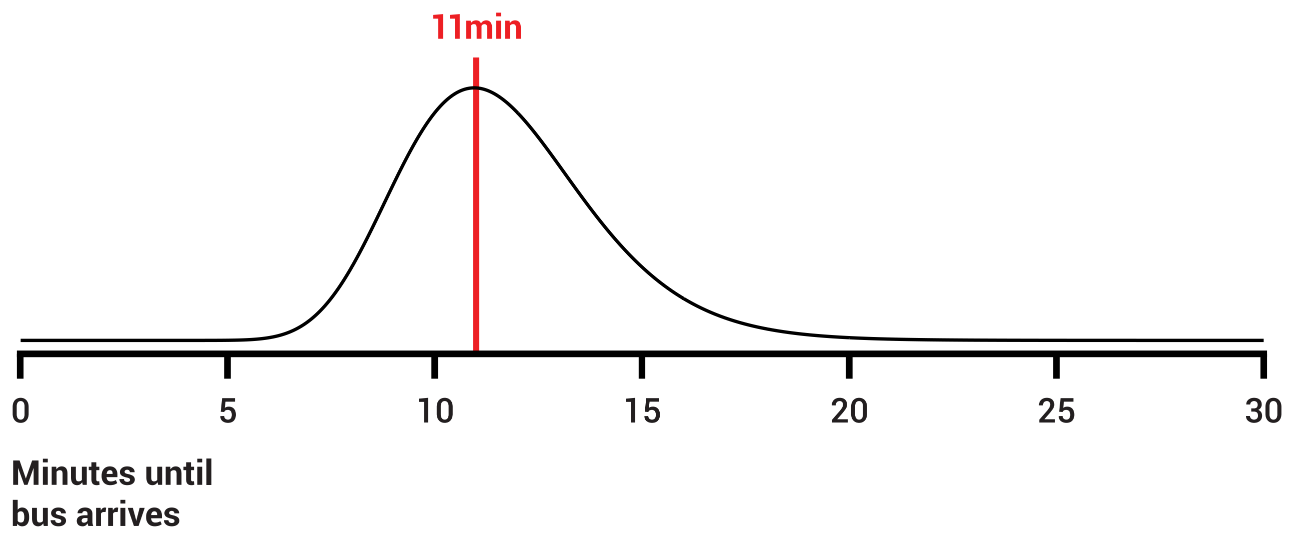
11:05 - 8 min delay
28
BROADVIEW
FREMONT
11:09 - on time
5
16
NORTHGATE
WALLINGFORD
11:10 - on time
6
358E
AURORA VILLAGE
VIA AURORA AVE N
8
120
DOWNTOWN
SEATTLE WHITE
CENTER
11:15 - 6 min delay
11
5
NORTHGATE
GREENWOOD
11:17 - 3 min delay
13
Be advised:
Bus arrival estimates are based on the best available information but actual times will vary.
Traffic and other conditions can affect the accuracy of this information.
SDOT
OneBusAway
King County
this bus stop.
transit.htm

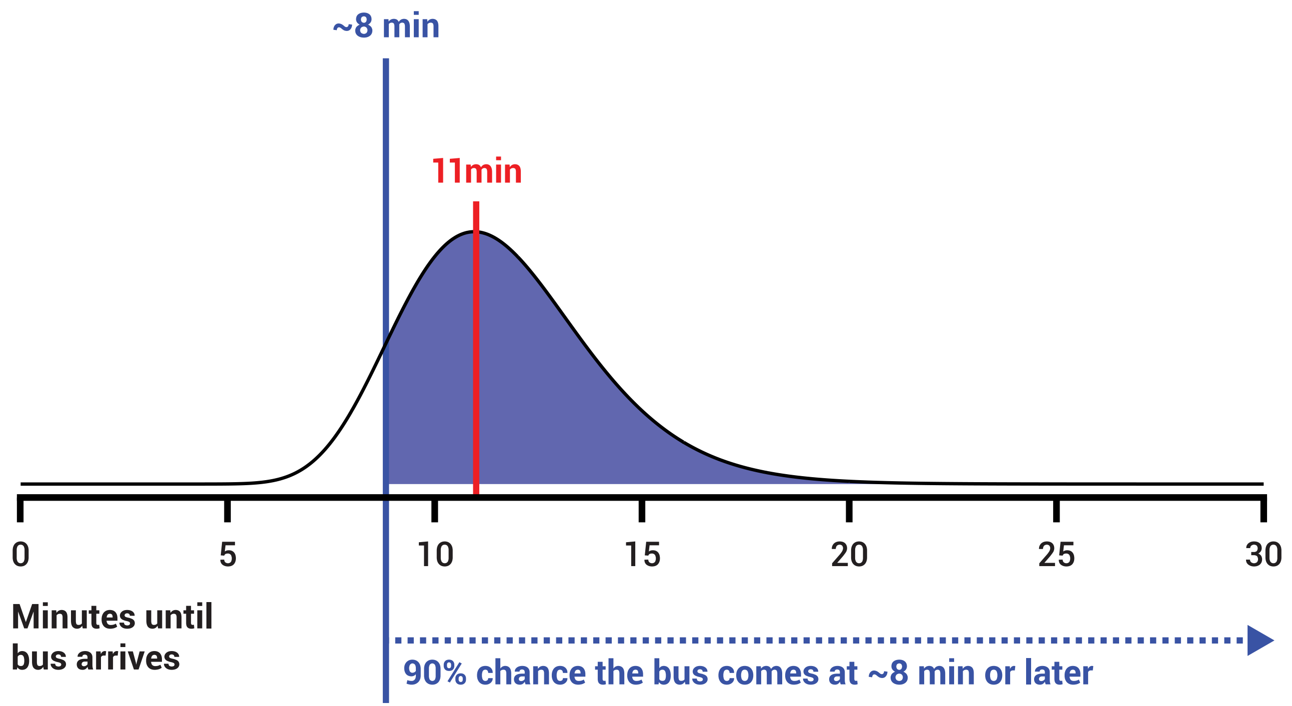


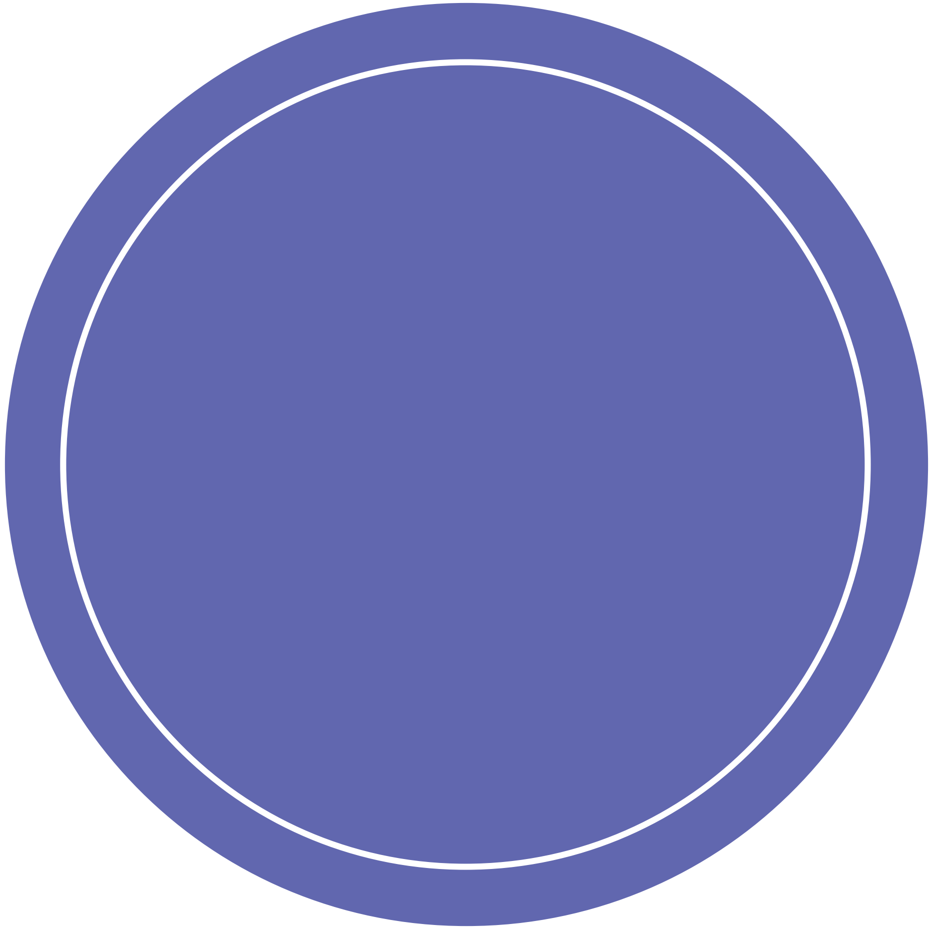
www.OneBusAway.org.
RE?
this bus stop.
f buses serving this stop in
more room for pedestrians
transit.htm
SONY
11:05 - 8 min delay
28
BROADVIEW
FREMONT
11:09 - on time
5
16
NORTHGATE
WALLINGFORD
11:10 - on time
6
358E
AURORA VILLAGE
VIA AURORA AVE N
8
120
DOWNTOWN
SEATTLE WHITE
CENTER
11:15 - 6 min delay
11
5
NORTHGATE
GREENWOOD
11:17 - 3 min delay
13
Be advised:
Bus arrival estimates are based on the best available information but actual times will vary.
Traffic and other conditions can affect the accuracy of this information.
SDOT
OneBusAway
macy's
King County

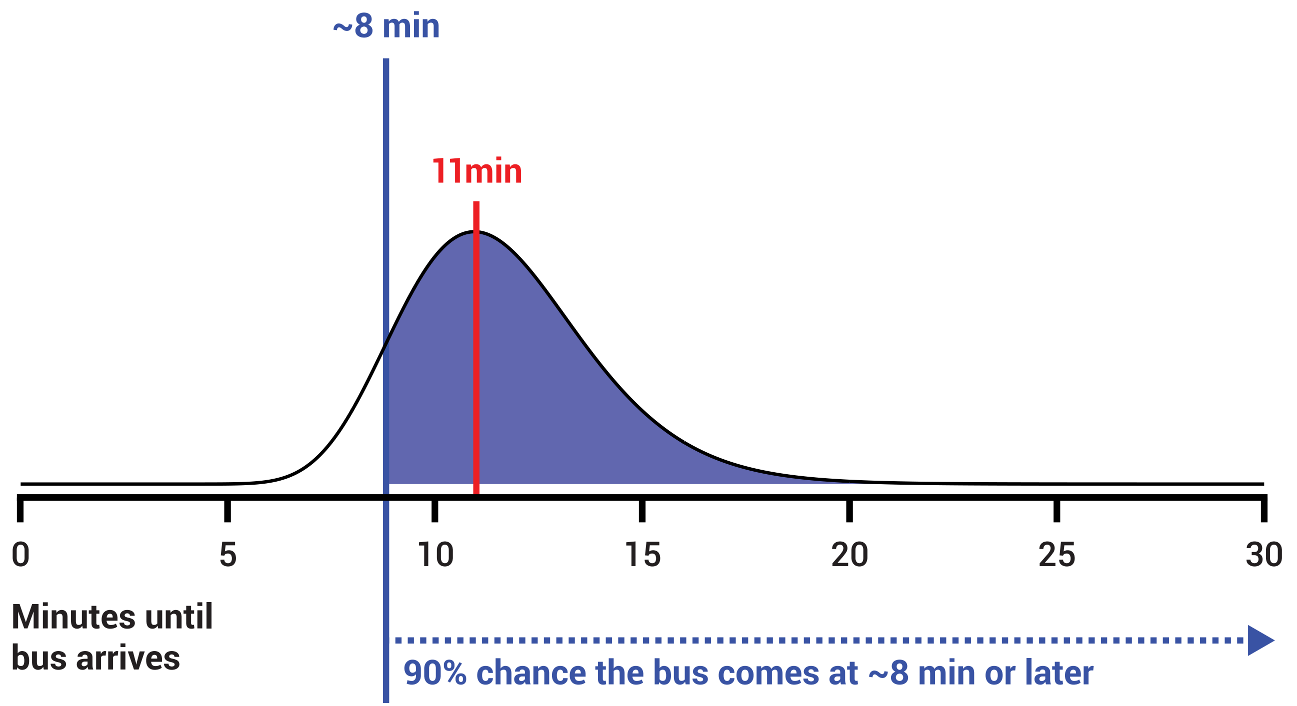
Do I have time to get a coffee?

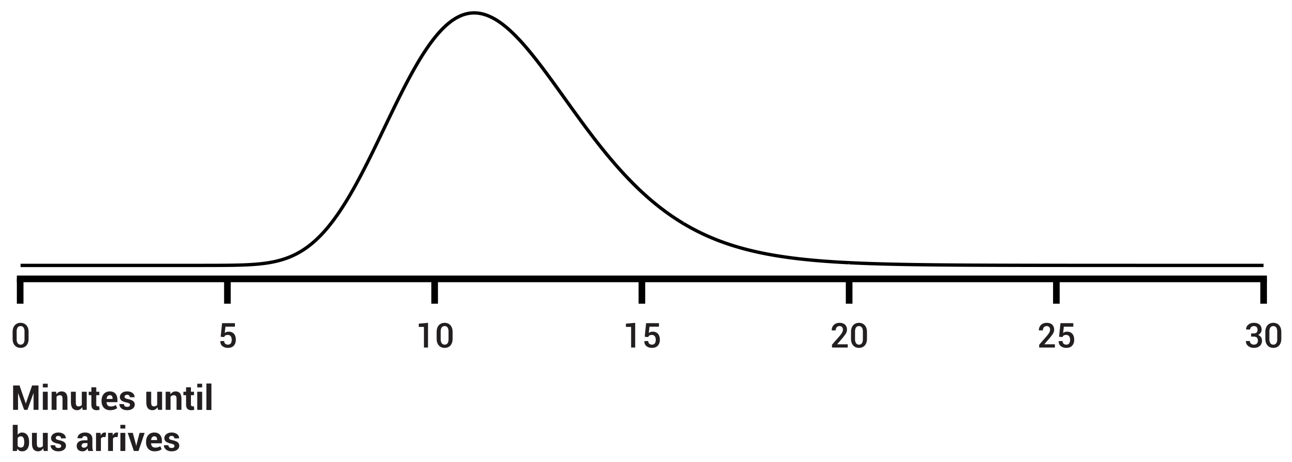

11min
0
5
10
15
20
25
30
Minutes until
bus arrives

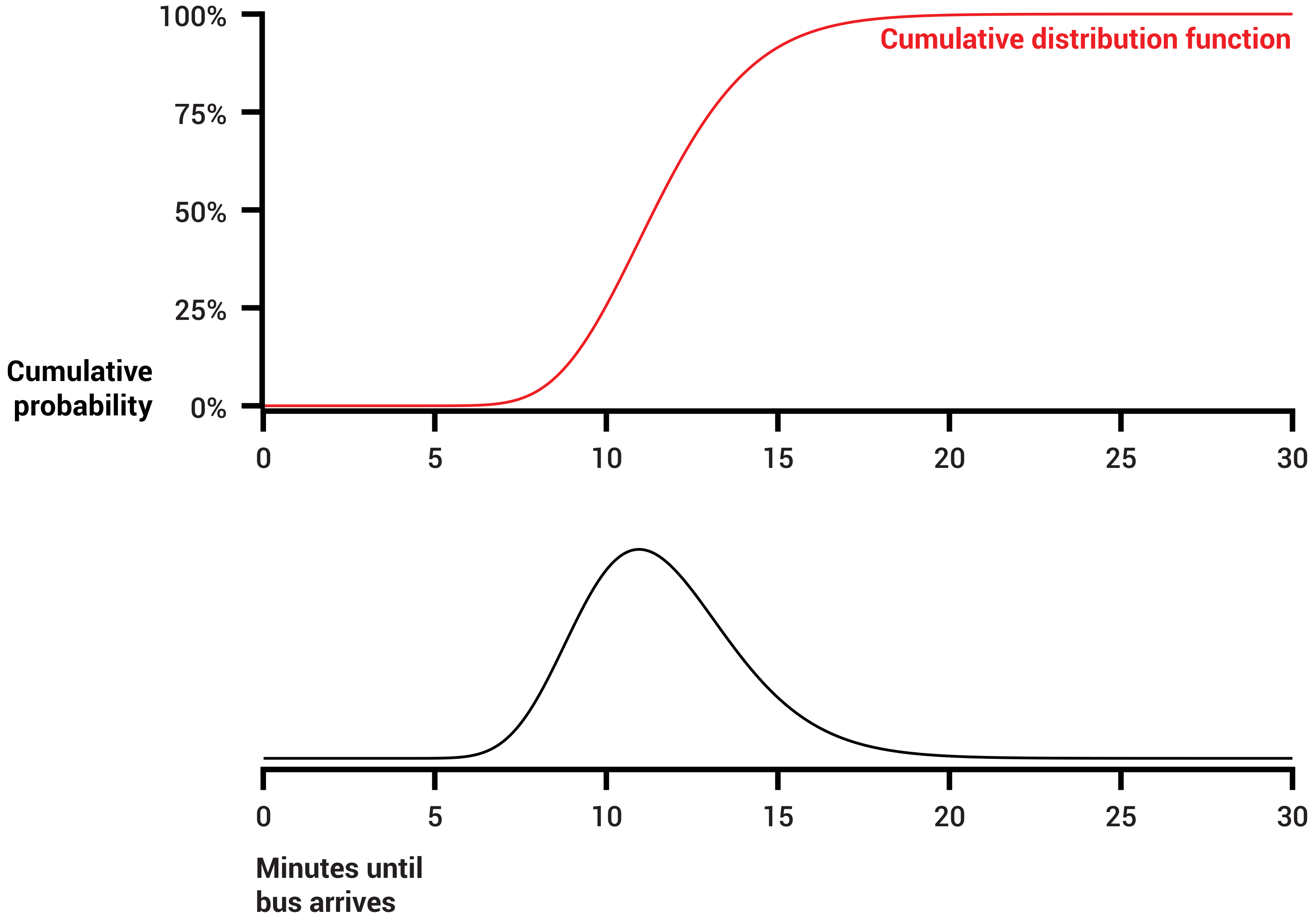

11min
0
5
10
15
20
25
30
Minutes until
bus arrives


~8 min
11min
0
5
10
15
20
25
30
Minutes until
bus arrives
90% chance the bus comes at ~8 min or later

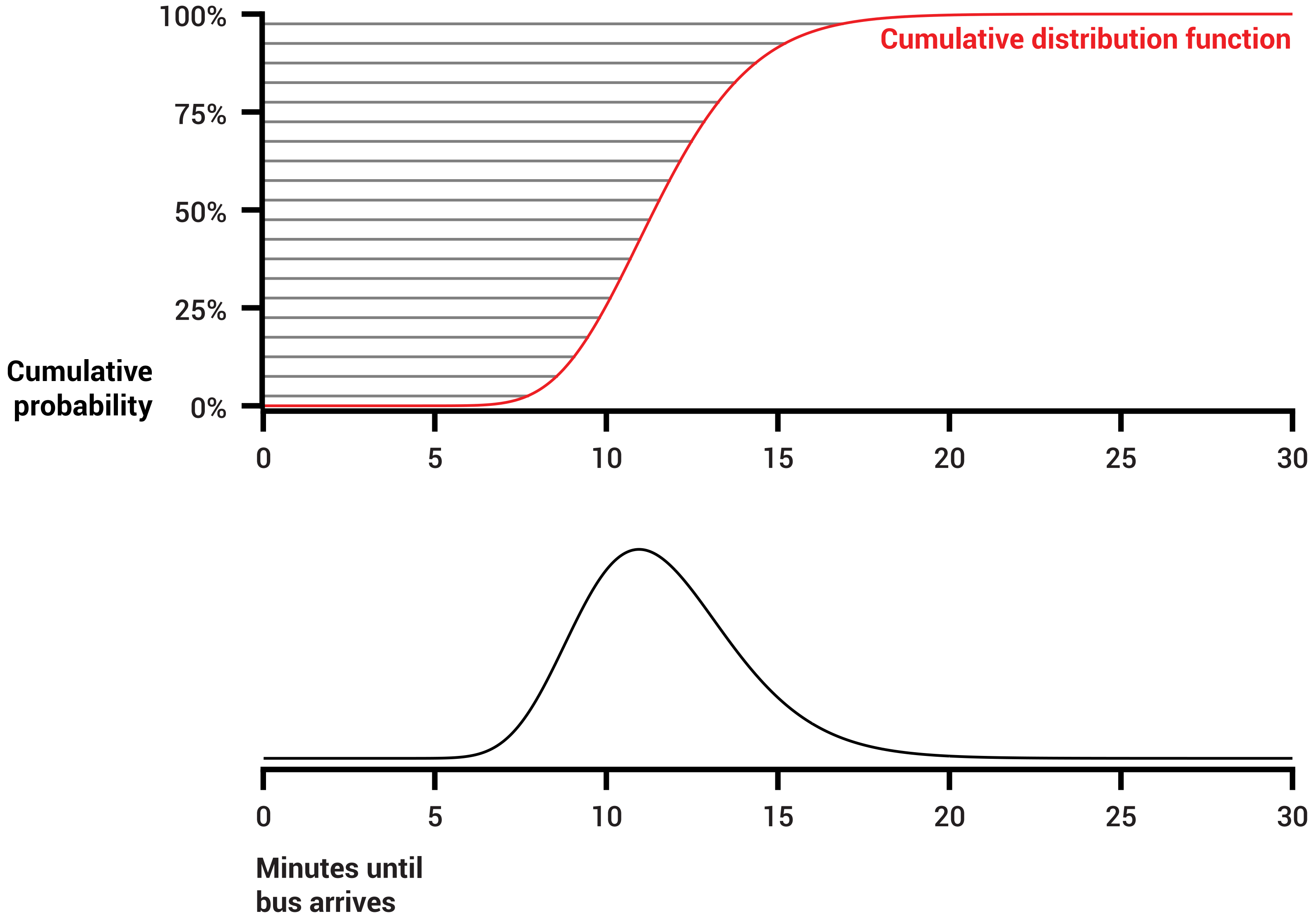


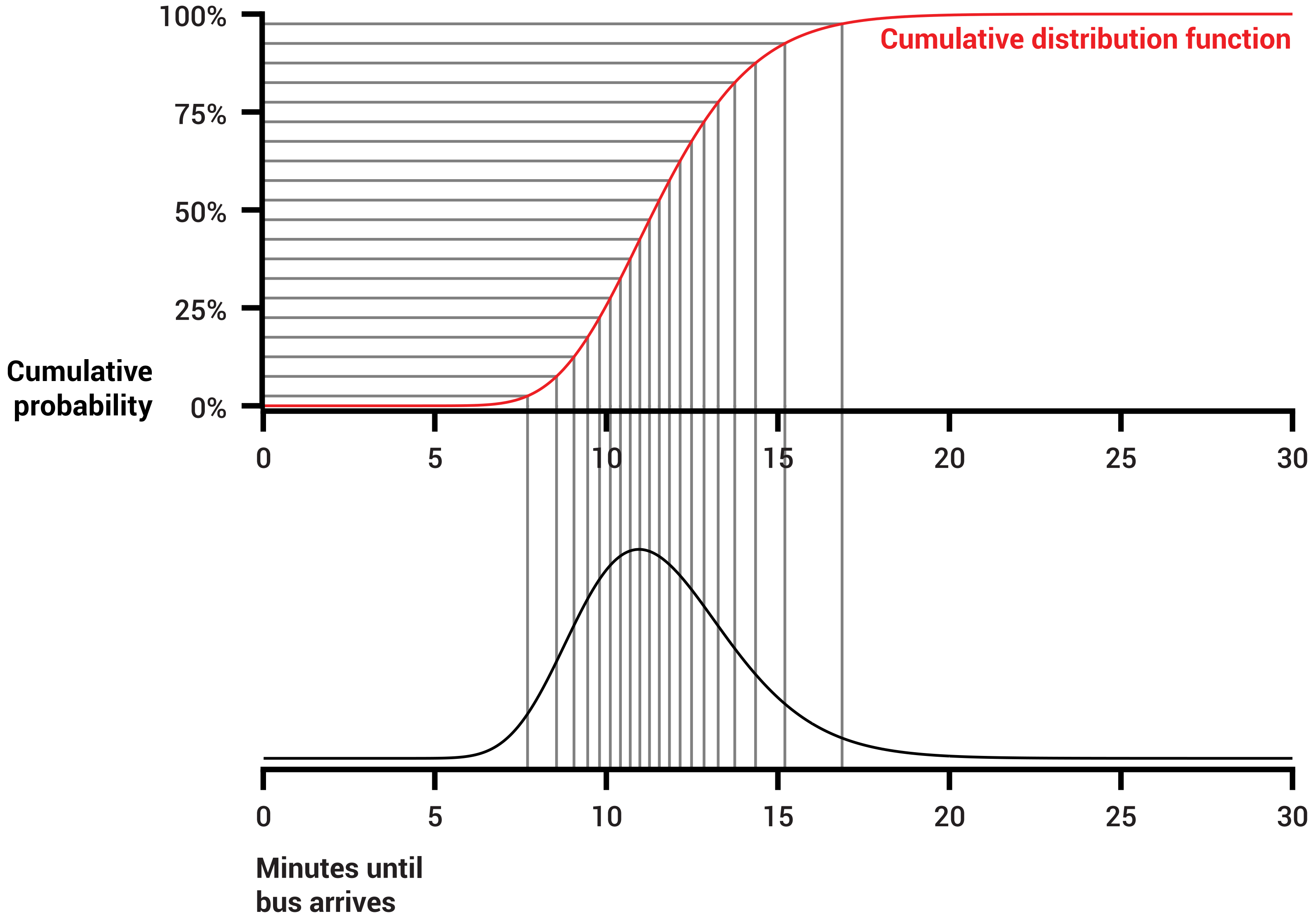

~8 min
11min
0
5
10
15
20
25
30
Minutes until
bus arrives
90% chance the bus comes at ~8 min or later

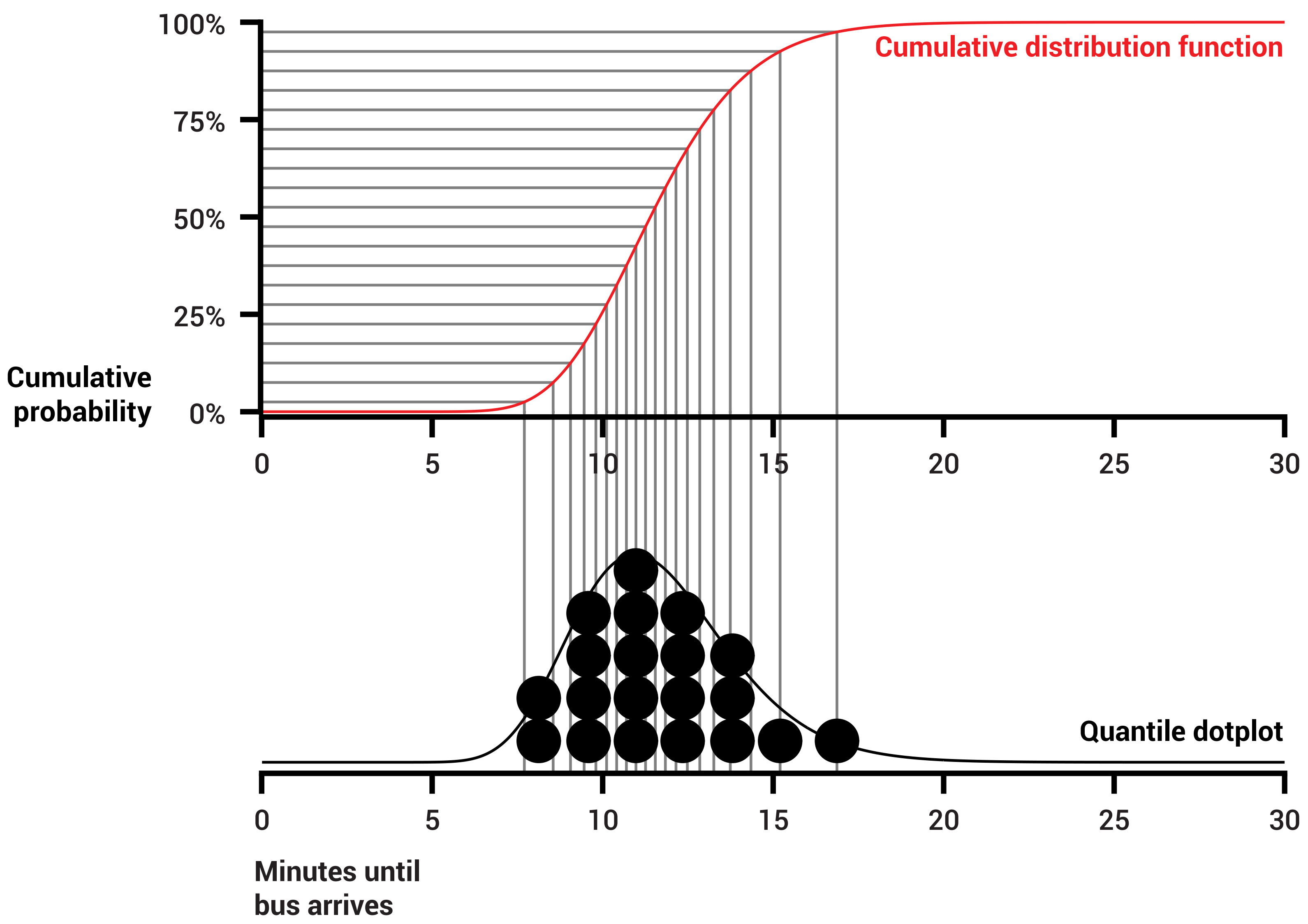

0
5
10
15
20
25
30
Minutes until
bus arrives

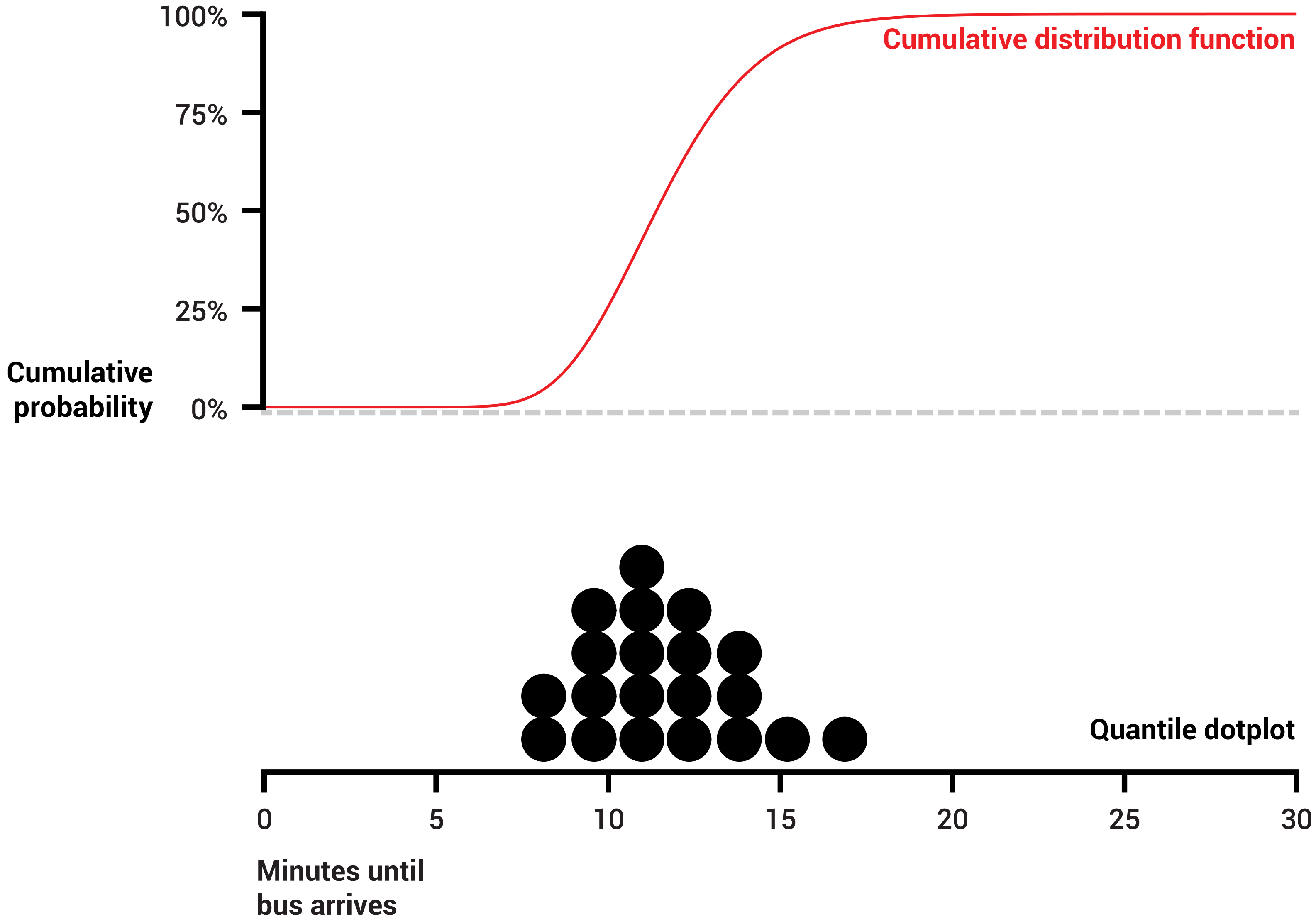


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
0
5
10
15
20
25
30
Minutes until bus arrives

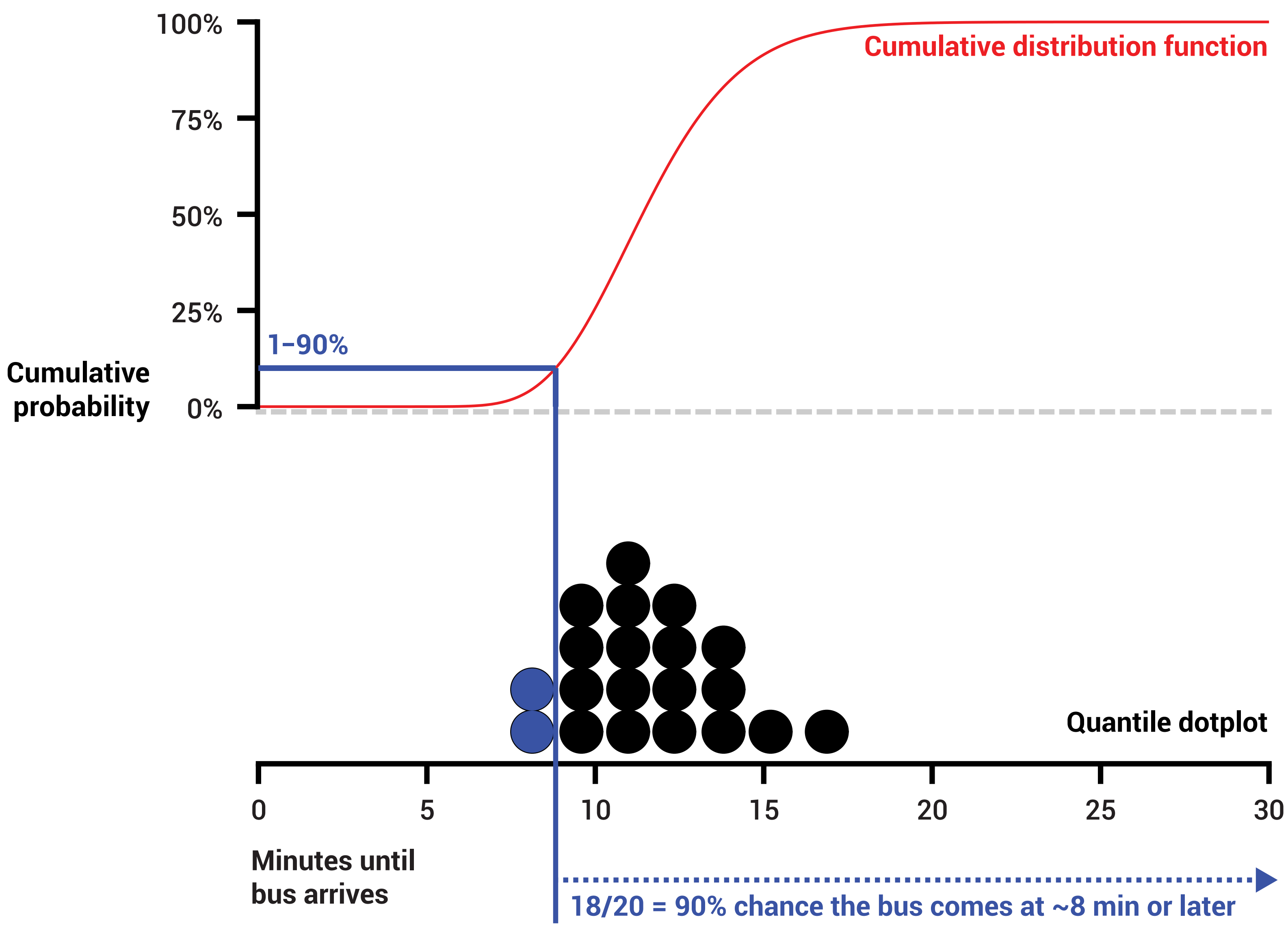

100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
0
5
10
15
20
25
30
Minutes until bus arrives

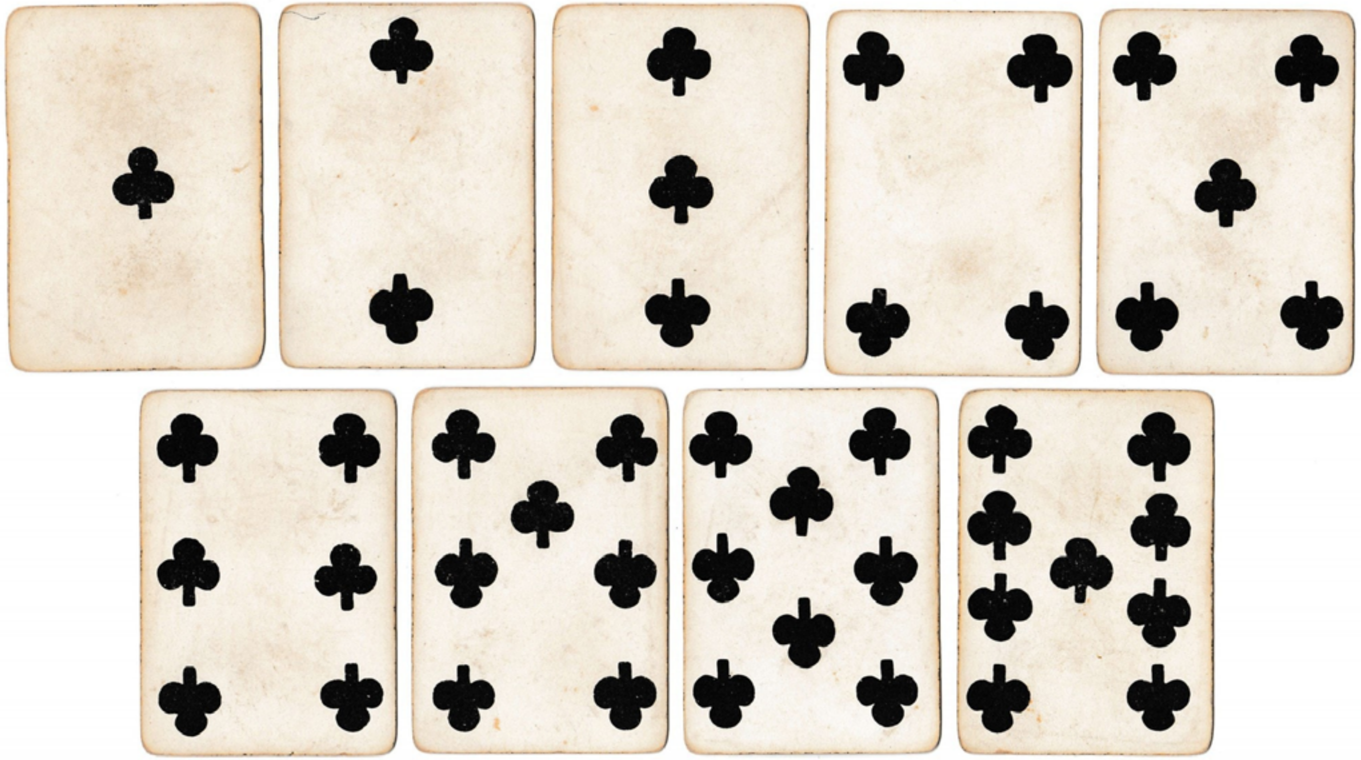

100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
0
5
10
15
20
25
30
Minutes until bus arrives

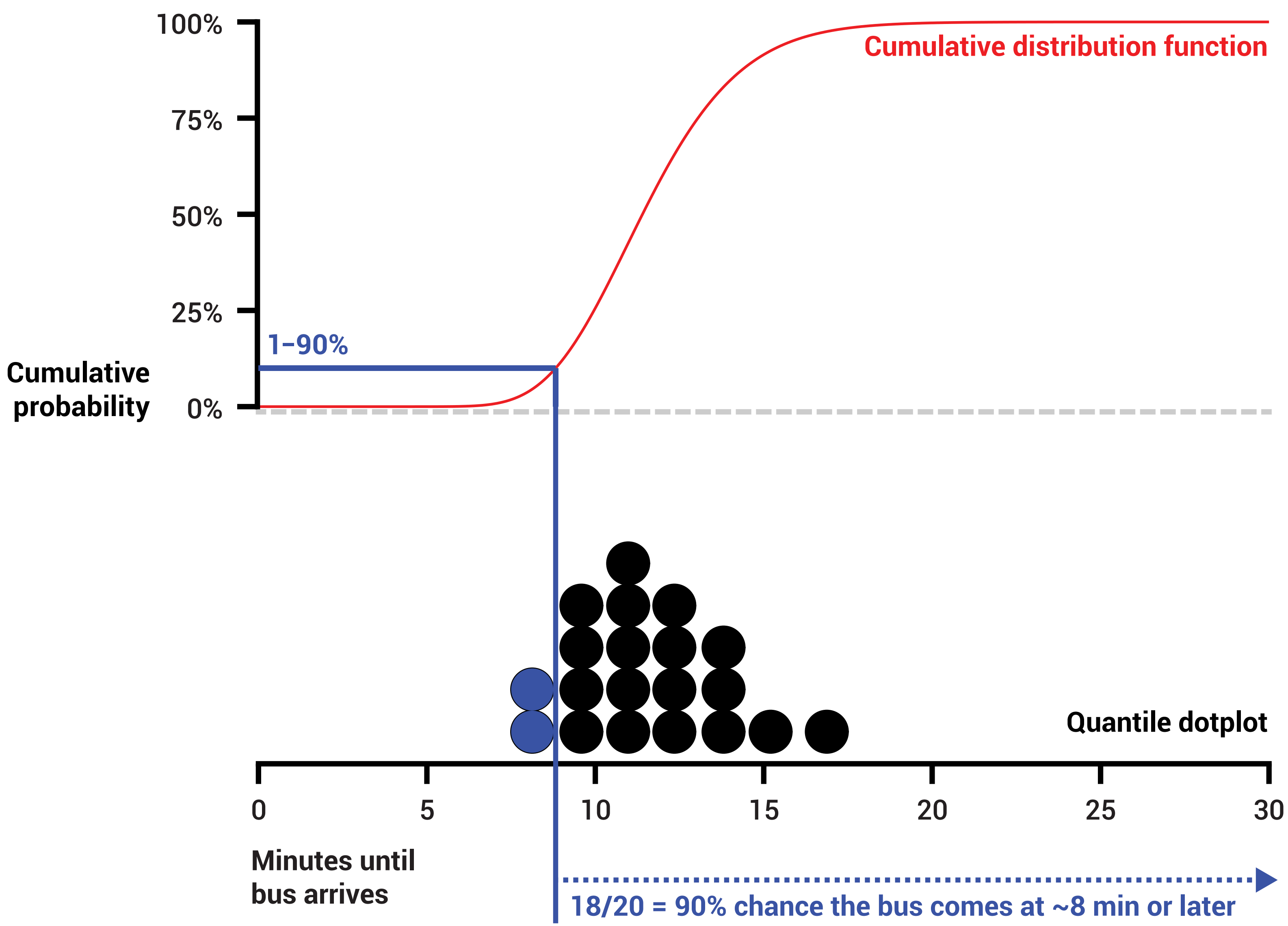

100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
0
5
10
15
20
25
30
Quantile dotplot
0
5
10
15
20
25
30
Minutes until bus arrives



100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
Quantile dotplot
0
5
10
15
20
25
30
Minutes until bus arrives


100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
1−90%
Quantile dotplot
0
5
10
15
20
25
30
Minutes until bus arrives
18/20 = 90% chance the bus comes at ~8 min or later

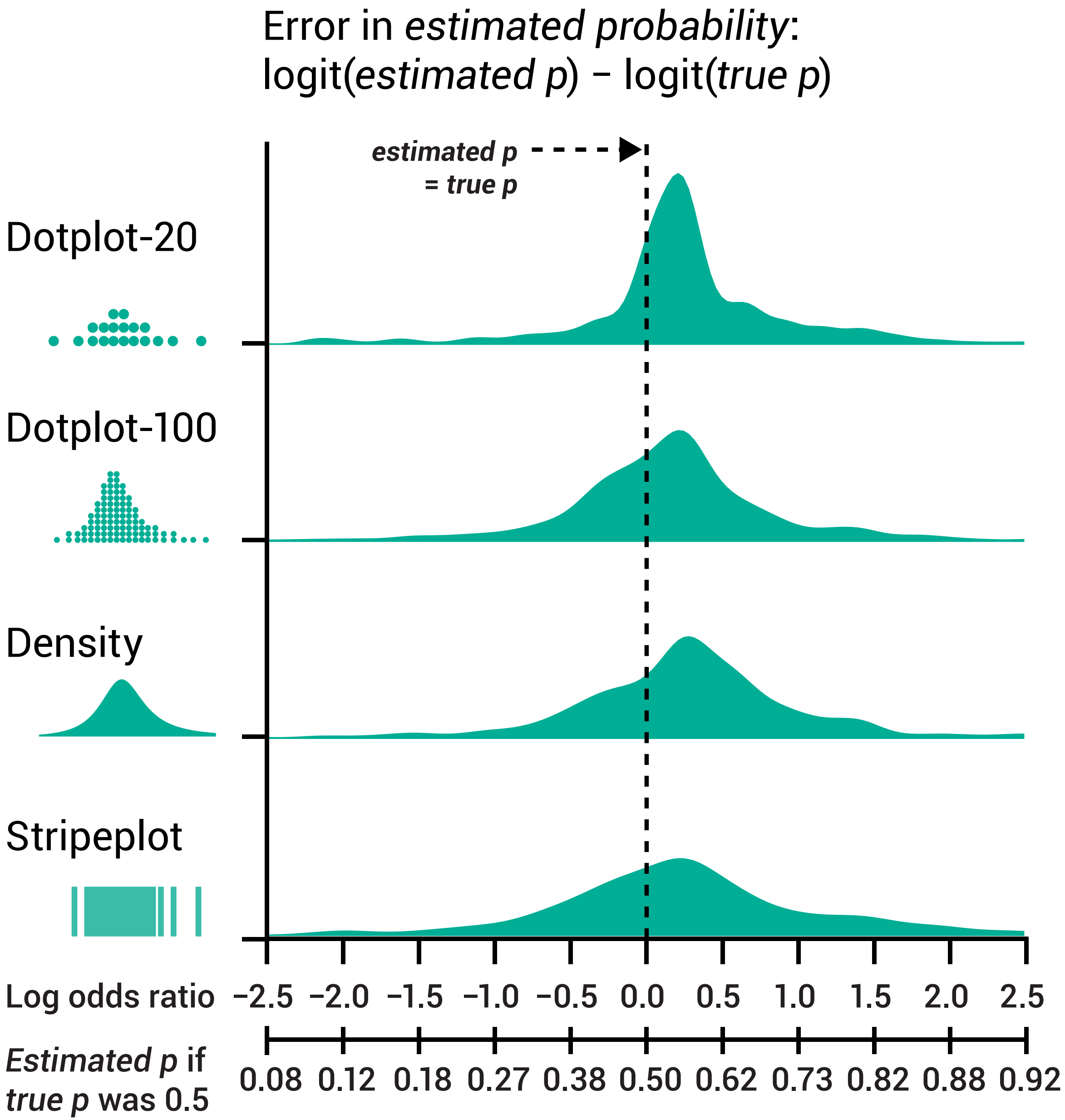



100%
75%
50%
25%
0%
Cumulative probability
Cumulative distribution function
1−90%
Quantile dotplot
0
5
10
15
20
25
30
Minutes until bus arrives
18/20 = 90% chance the bus comes at ~8 min or later

Quantile dotplots

[**Kay**, Kola, Hullman, Munson. When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. CHI 2016]

Better estimates (perceptually)



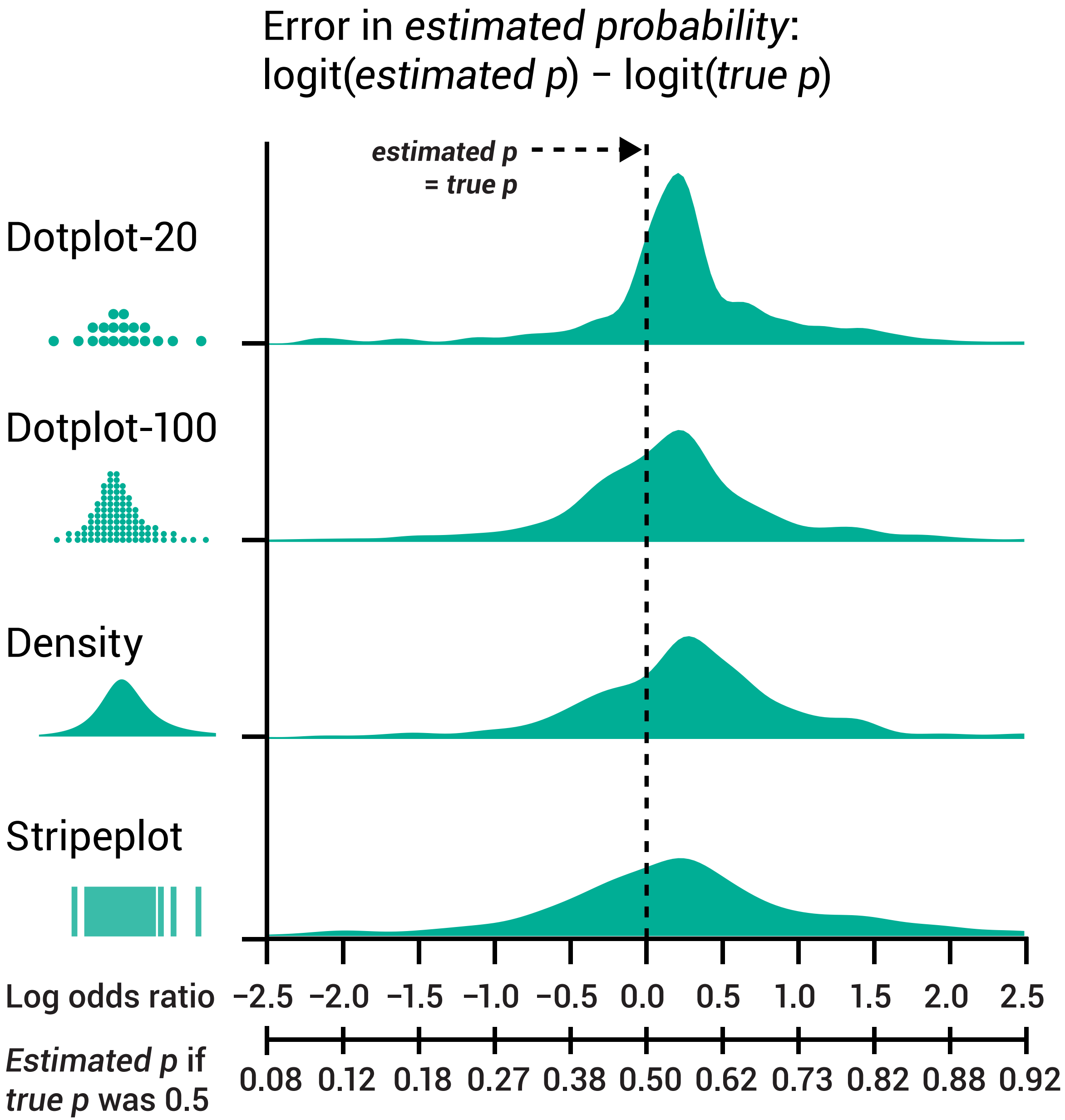
Quantile dotplots

[**Kay**, Kola, Hullman, Munson. When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. CHI 2016]

Better estimates
(perceptually)

↓?

better decisions



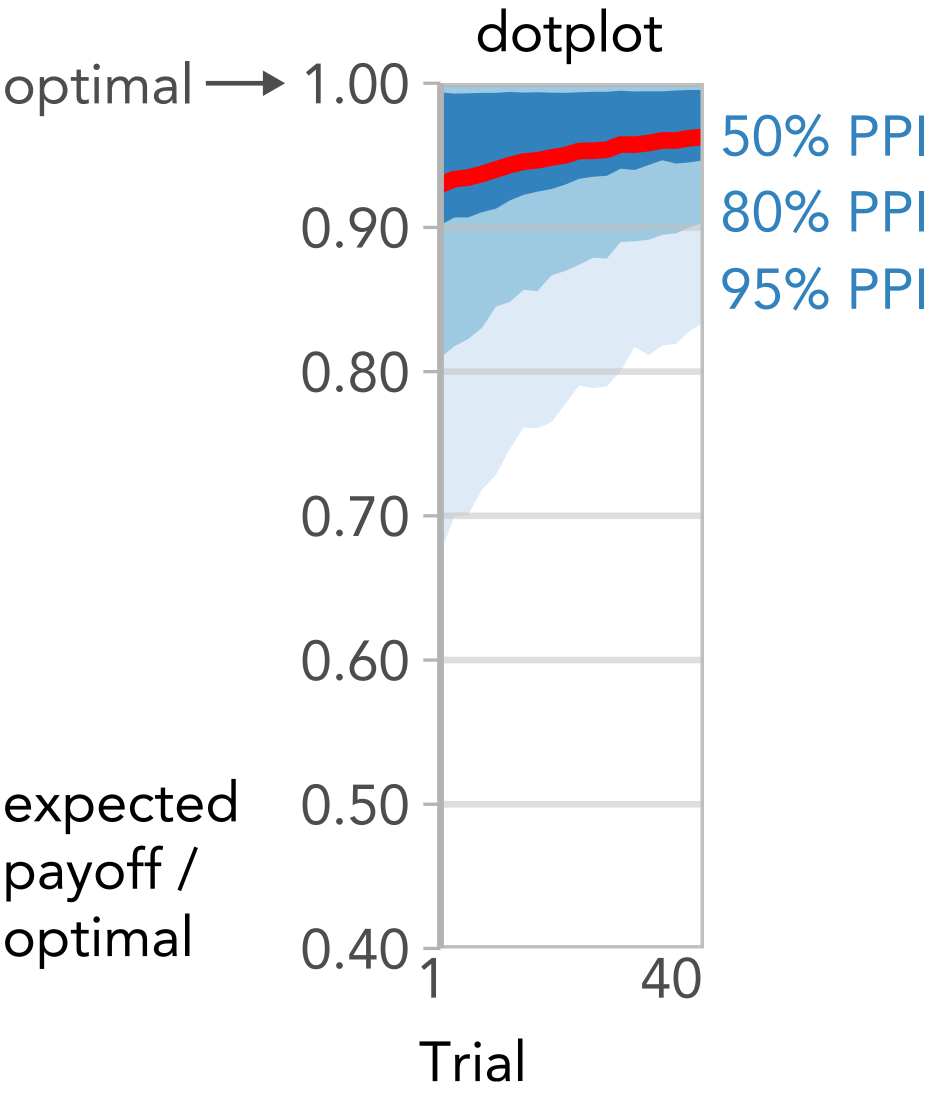
Quantile dotplots

[Fernandes, Munson, Hullman, **Kay**. Uncertainty Displays Using Quantile Dotplots or CDFs Improve Transit Decision-Making. CHI 2018. **Honorable Mention**]

Better estimates
(perceptually)

↓

better decisions
(in this case)



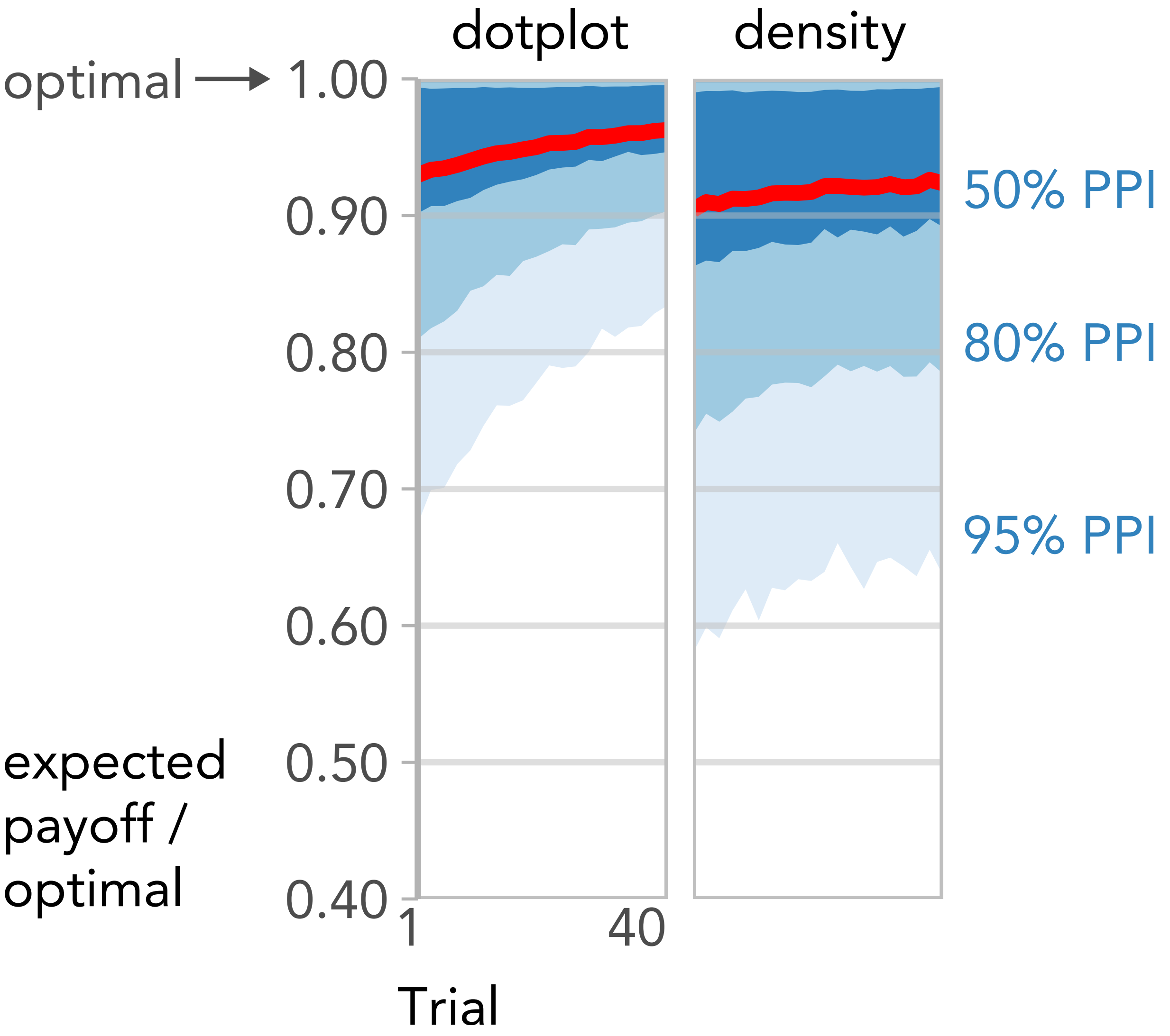
Quantile dotplots

[Fernandes, Munson, Hullman, **Kay**. Uncertainty Displays Using Quantile Dotplots or CDFs Improve Transit Decision-Making. CHI 2018. **Honorable Mention**]

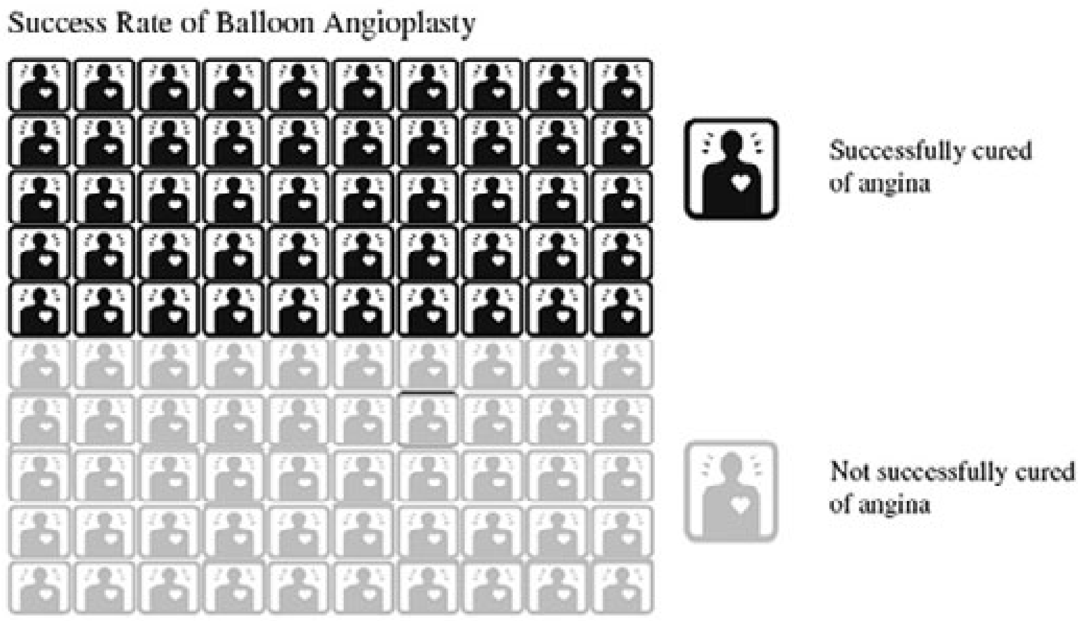
Better estimates
(perceptually)

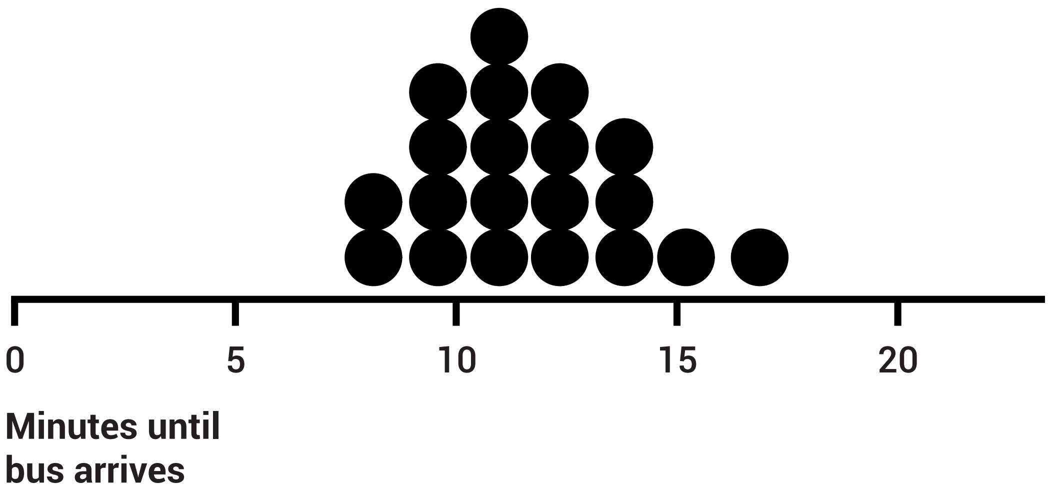
↓

better decisions
(in this case)



Discrete outcome / frequency framing





Other discrete outcome uncertainty visualizations…

Predictions from 2016 presidential election

[Justin H. Gross, Washington Post, http://wapo.st/2fCYvDW]

FiveThirtyEight	NYT Upshot	HuffPo Pollster
28%	15%	2%

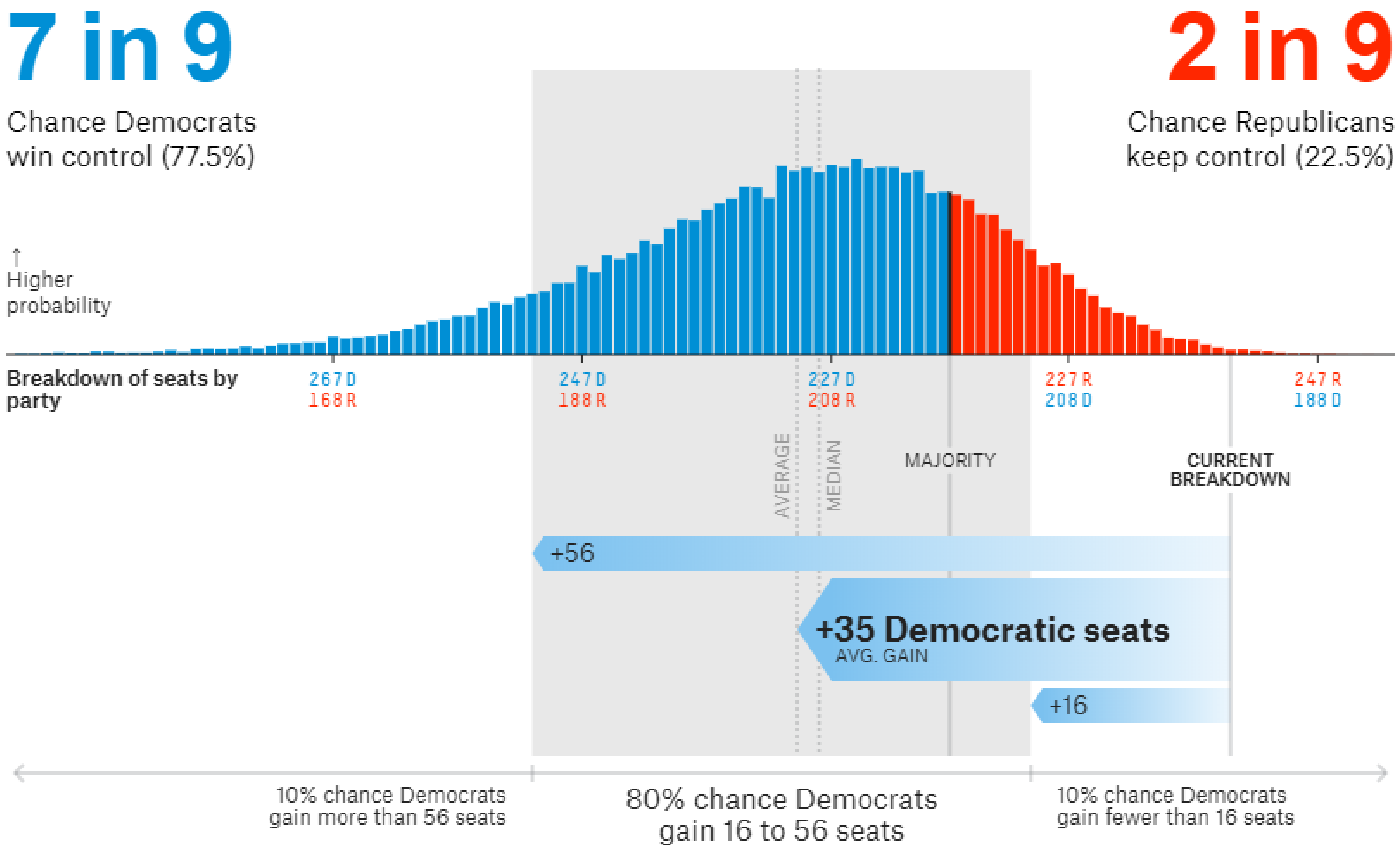
Predictions from 2016 presidential election

[Justin H. Gross, Washington Post, http://wapo.st/2fCYvDW]

FiveThirtyEight	NYT Upshot	HuffPo Pollster
28%	15%	2%

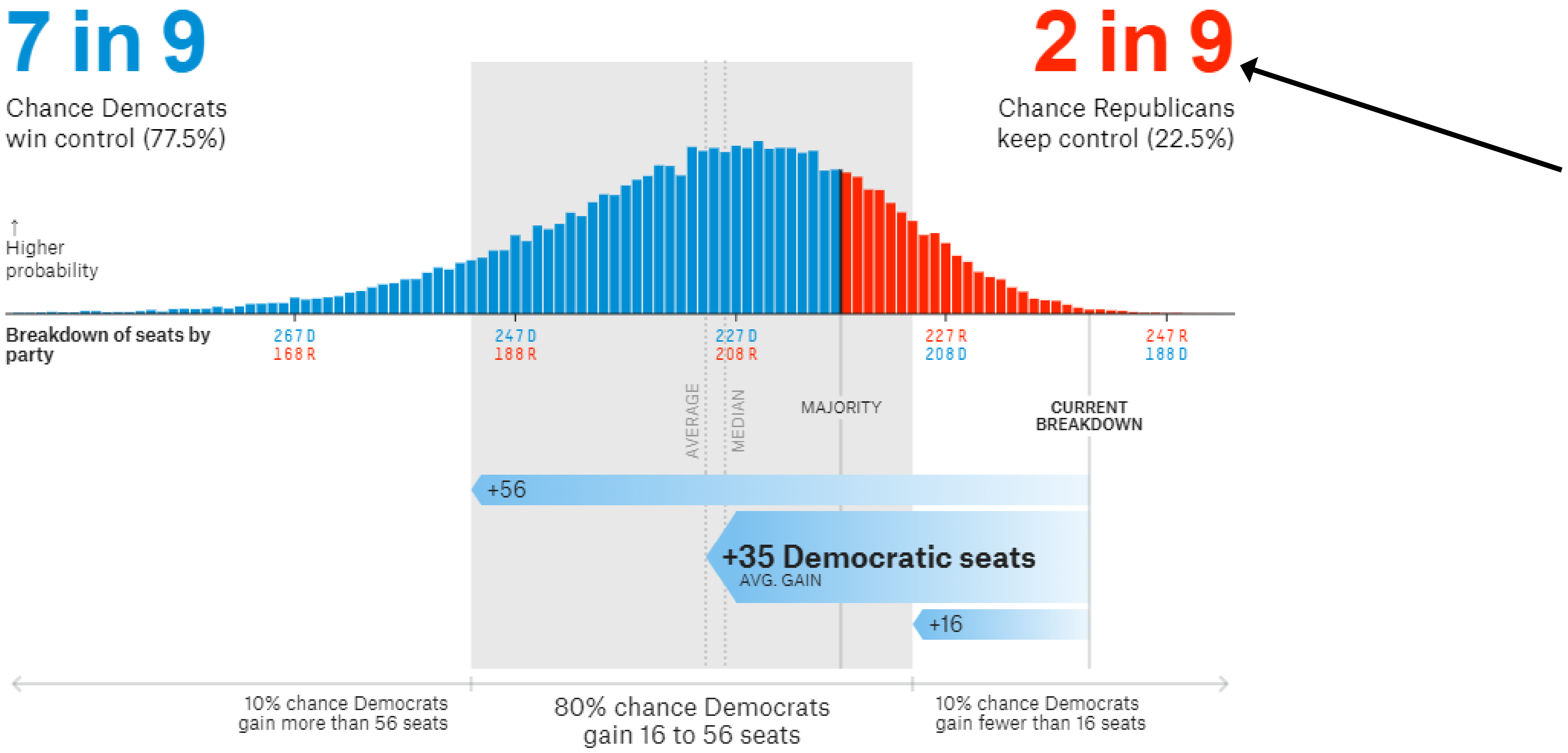
FiveThirtyEight's 2018 House forecast

[https://projects.fivethirtyeight.com/2018-midterm-election-forecast/house/]



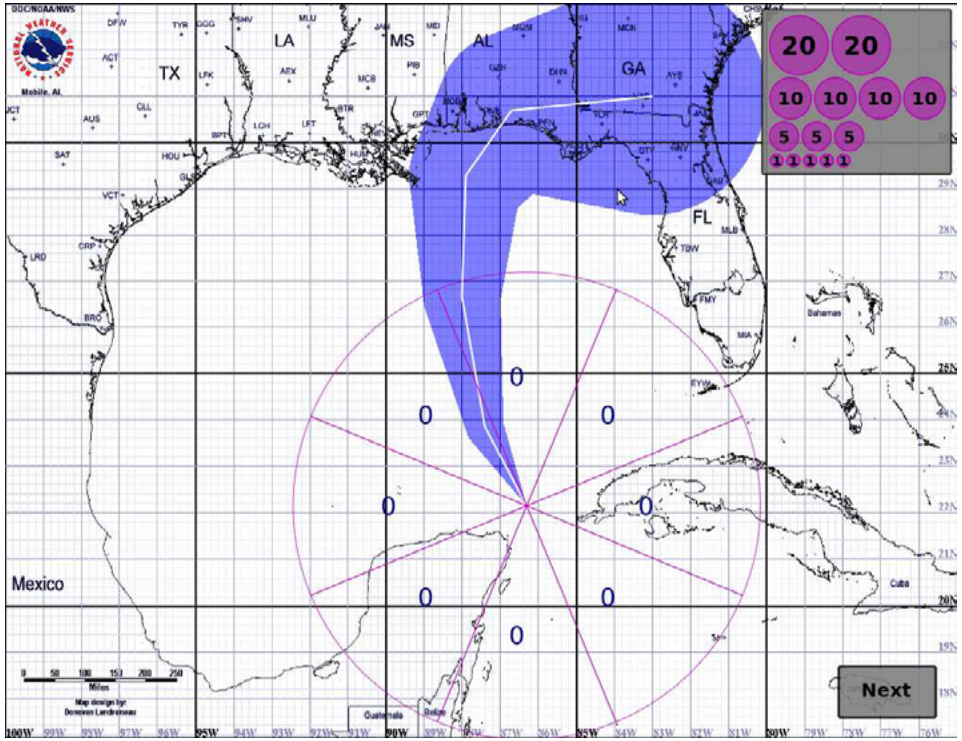
FiveThirtyEight's 2018 House forecast

[https://projects.fivethirtyeight.com/2018-midterm-election-forecast/house/]



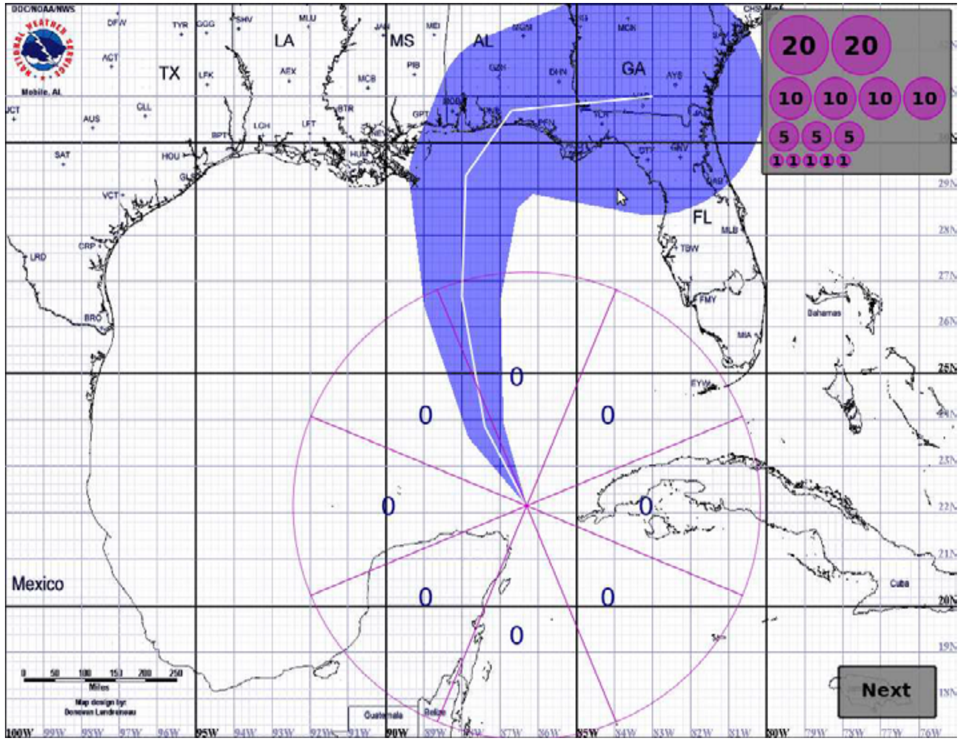
Hurricane error cones

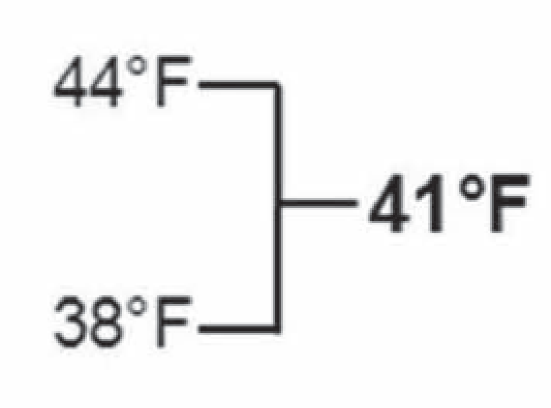
[Cox, House, Lindell. Visualizing Uncertainty in Predicted Hurricane Tracks. International Journal for Uncertainty Quantification, 3(2), 143–156, 2013]



Deterministic construal errors

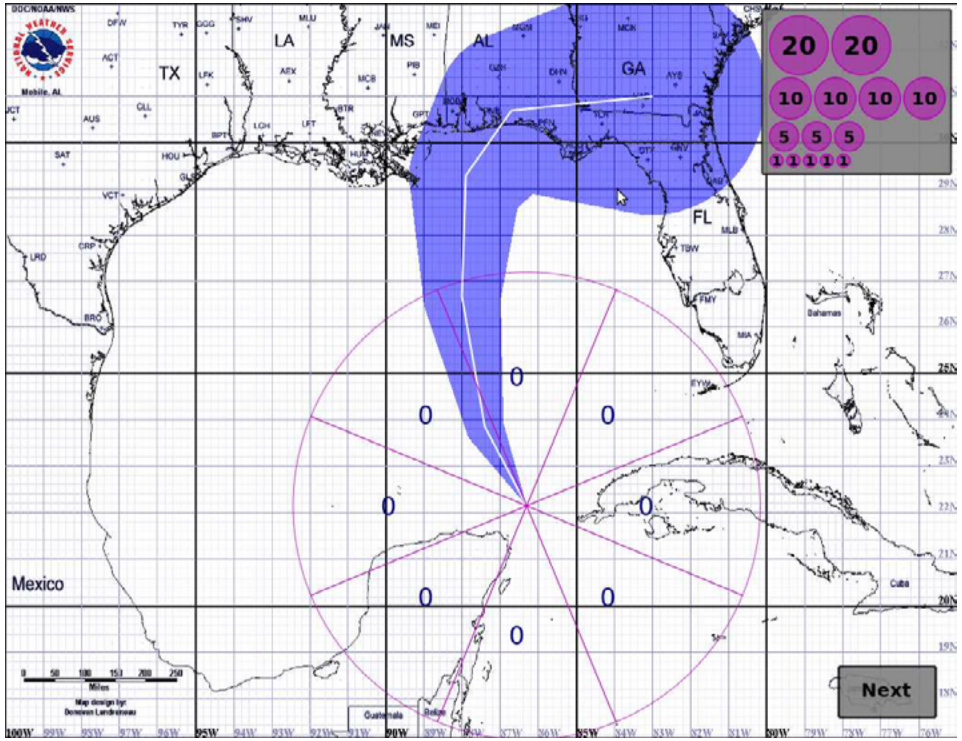
[Joslyn & LeClerc. Decisions With Uncertainty: The Glass Half Full. Current Directions in Psych. Science, 22(4), 2013]

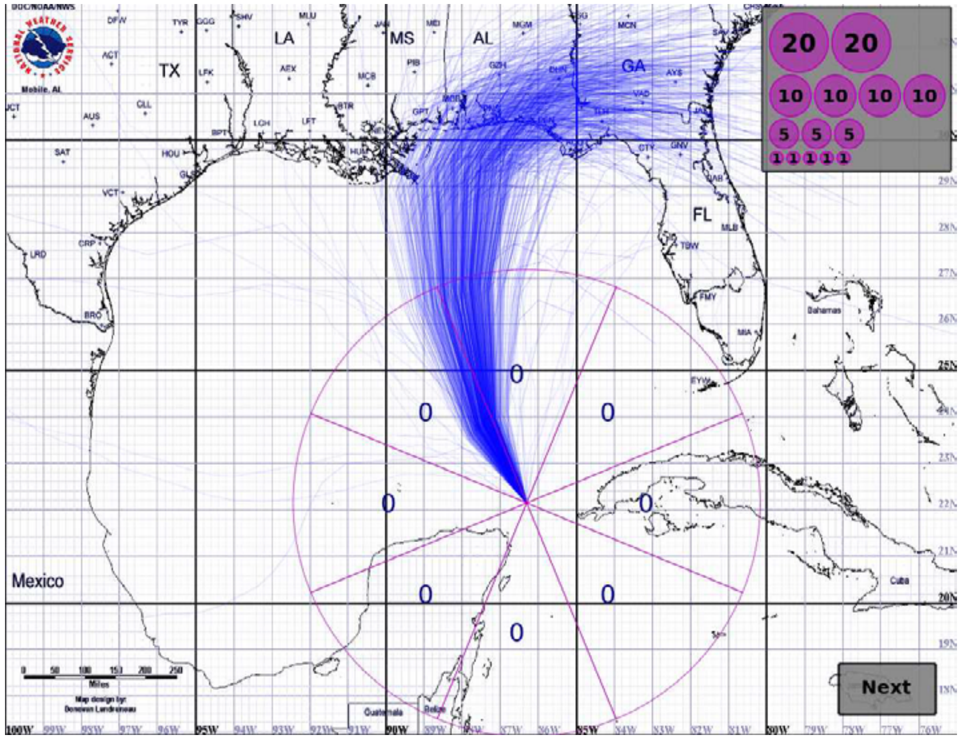




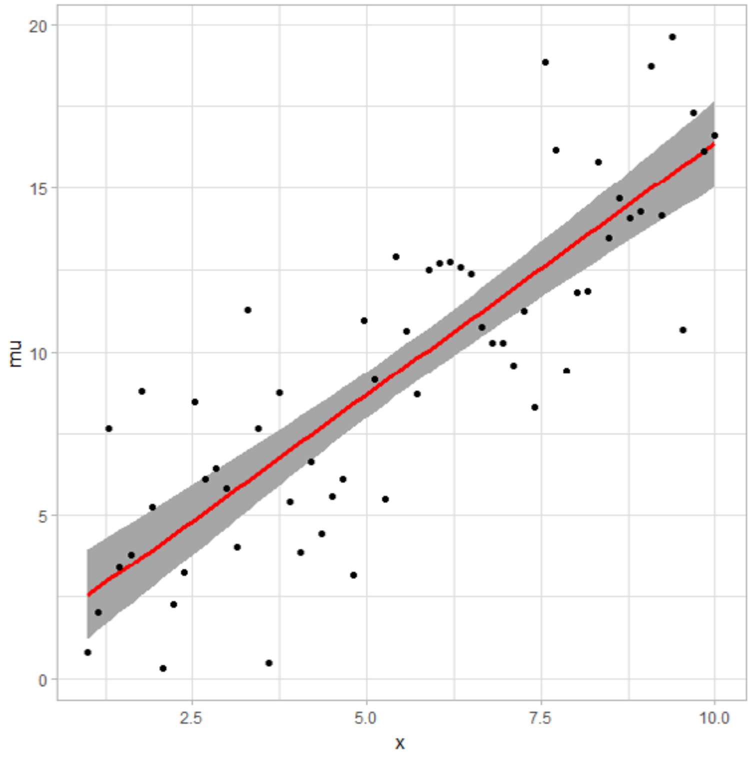
Hurricane error cones

[Cox, House, Lindell. Visualizing Uncertainty in Predicted Hurricane Tracks. International Journal for Uncertainty Quantification, 3(2), 143–156, 2013]

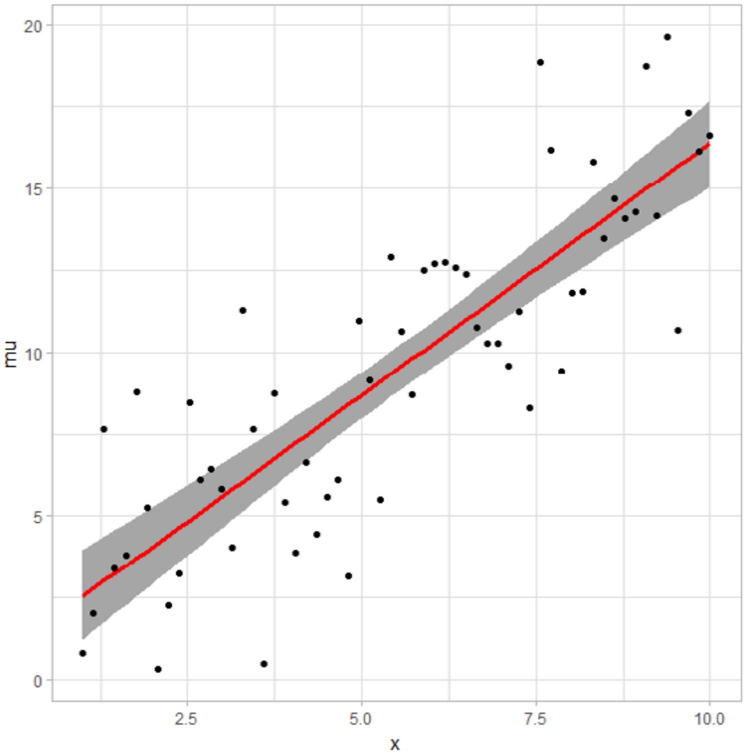


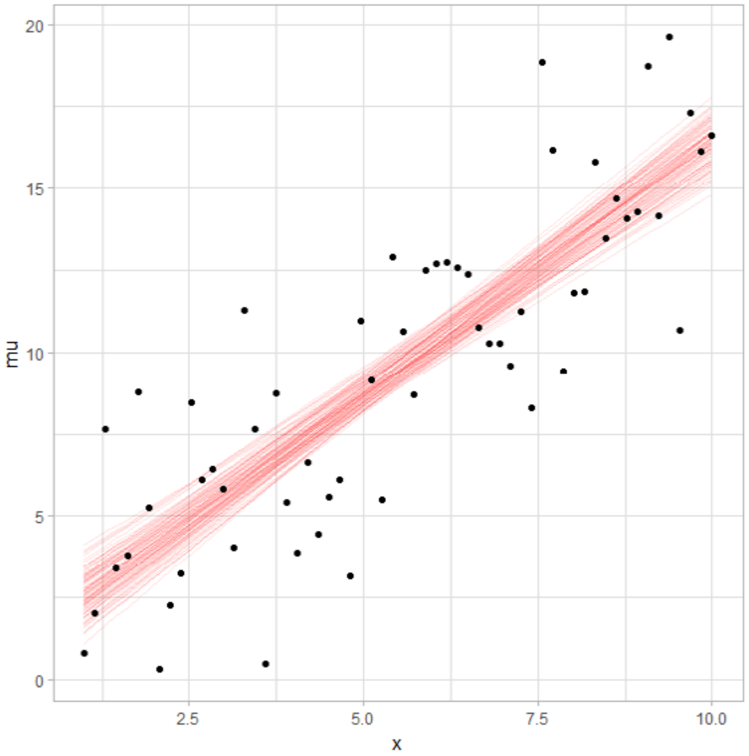


Fit line uncertainty



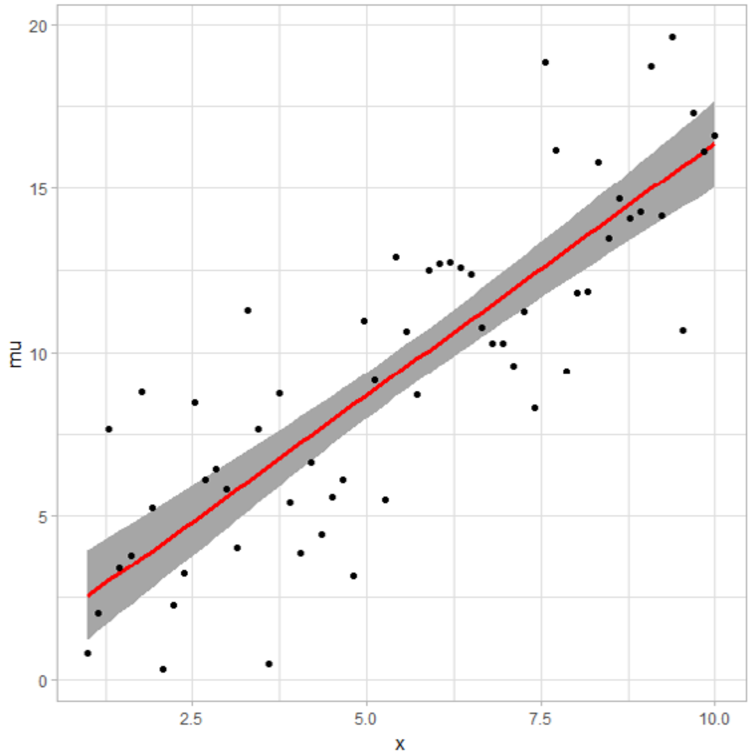
Fit line uncertainty

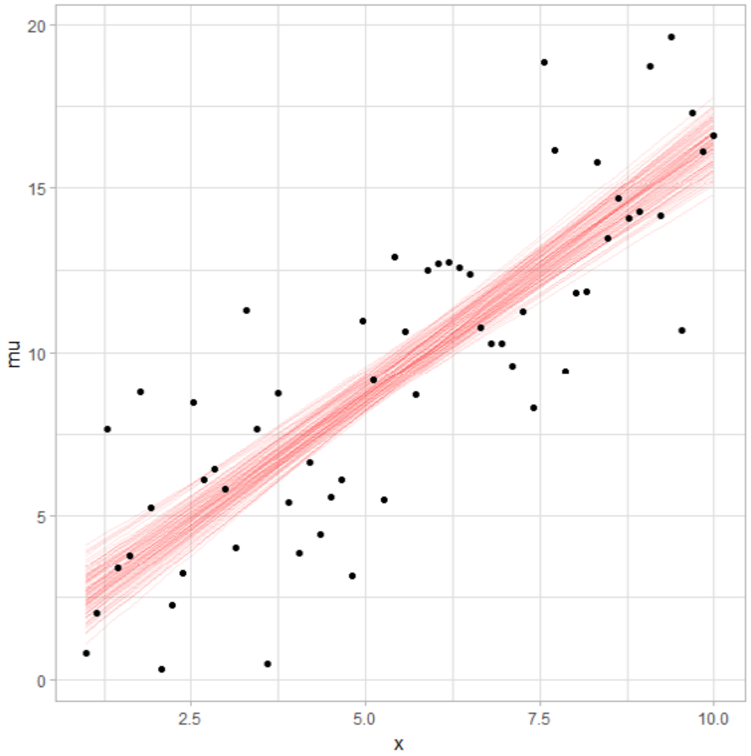


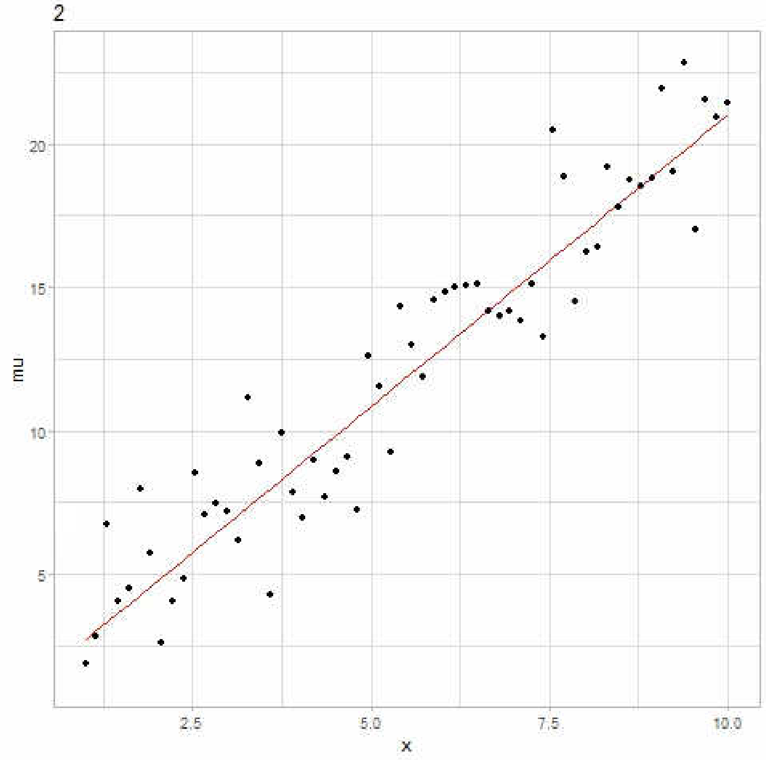


Fit line uncertainty

Hypothetical outcome plots (HOPs)







[Hullman, Resnick, Adar. Hypothetical Outcome Plots Outperform Error Bars and Violin Plots for Inferences about Reliability of Variable Ordering. PloS One, 10(11). 2015]

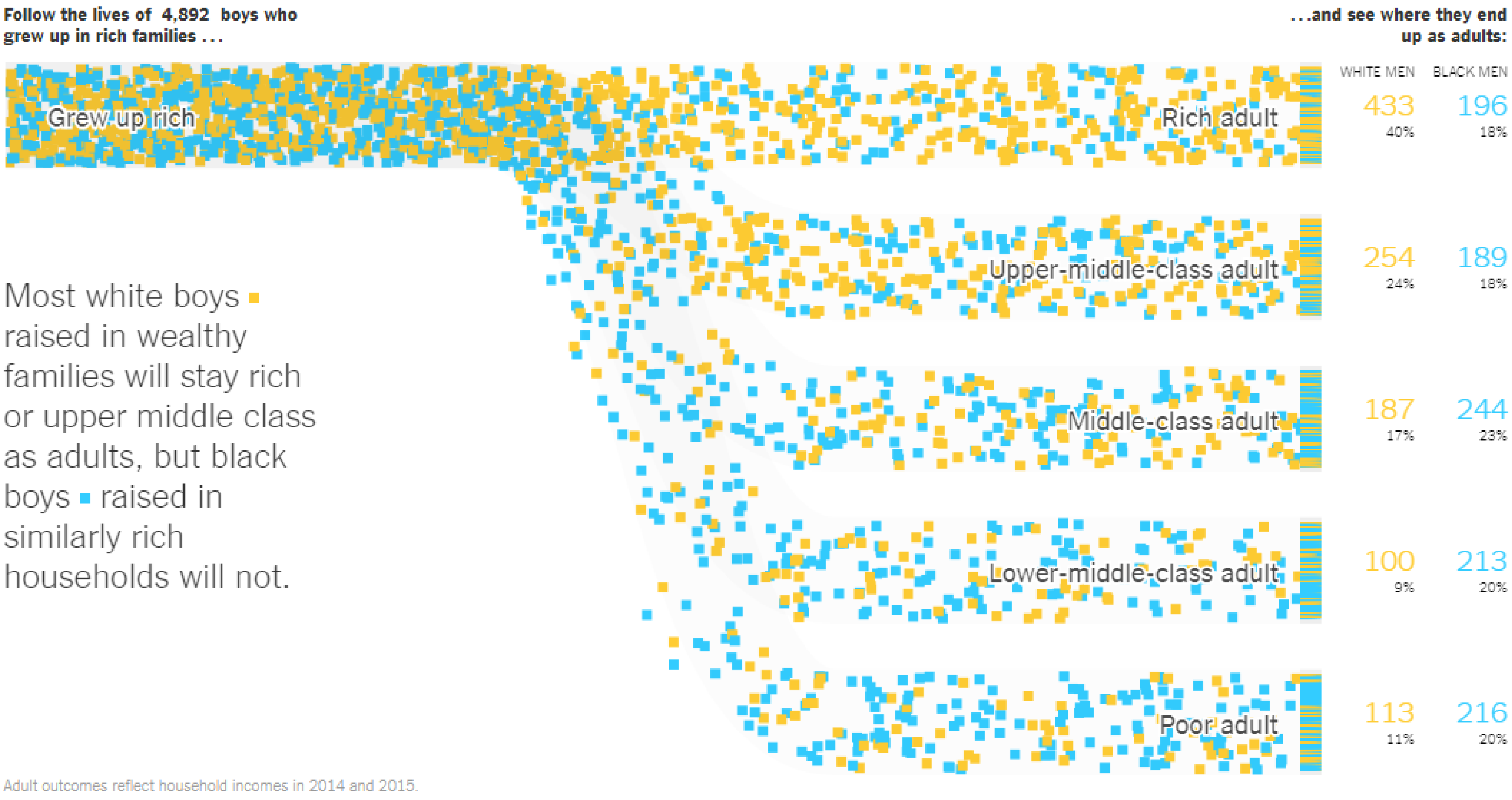
[Kale, Nguyen, **Kay**, Hullman. Hypothetical Outcome Plots Help Untrained Observers Judge Trends in Ambiguous Data. IEEE TVCG (Proc. InfoVis), 2018]

Animation helps people experience uncertainty

This can be very powerful...

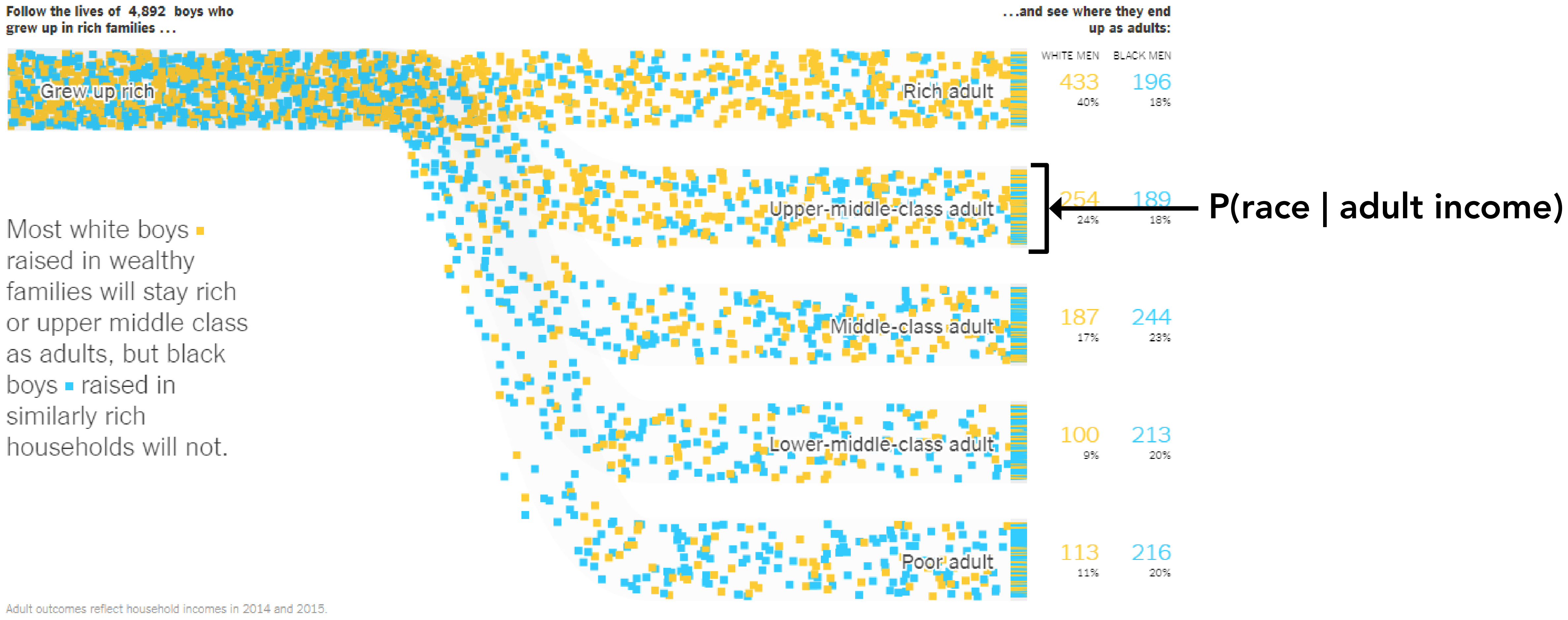
Income of black boys from wealthy families

[Badger, Miller, Pearce, Quealy. Extensive Data Shows Punishing Reach of Racism for Black Boys, NYT Upshot, 2018, https://nyti.ms/2GGpFZw]



Income of black boys from wealthy families

[Badger, Miller, Pearce, Quealy. Extensive Data Shows Punishing Reach of Racism for Black Boys, NYT Upshot, 2018, https://nyti.ms/2GGpFZw]

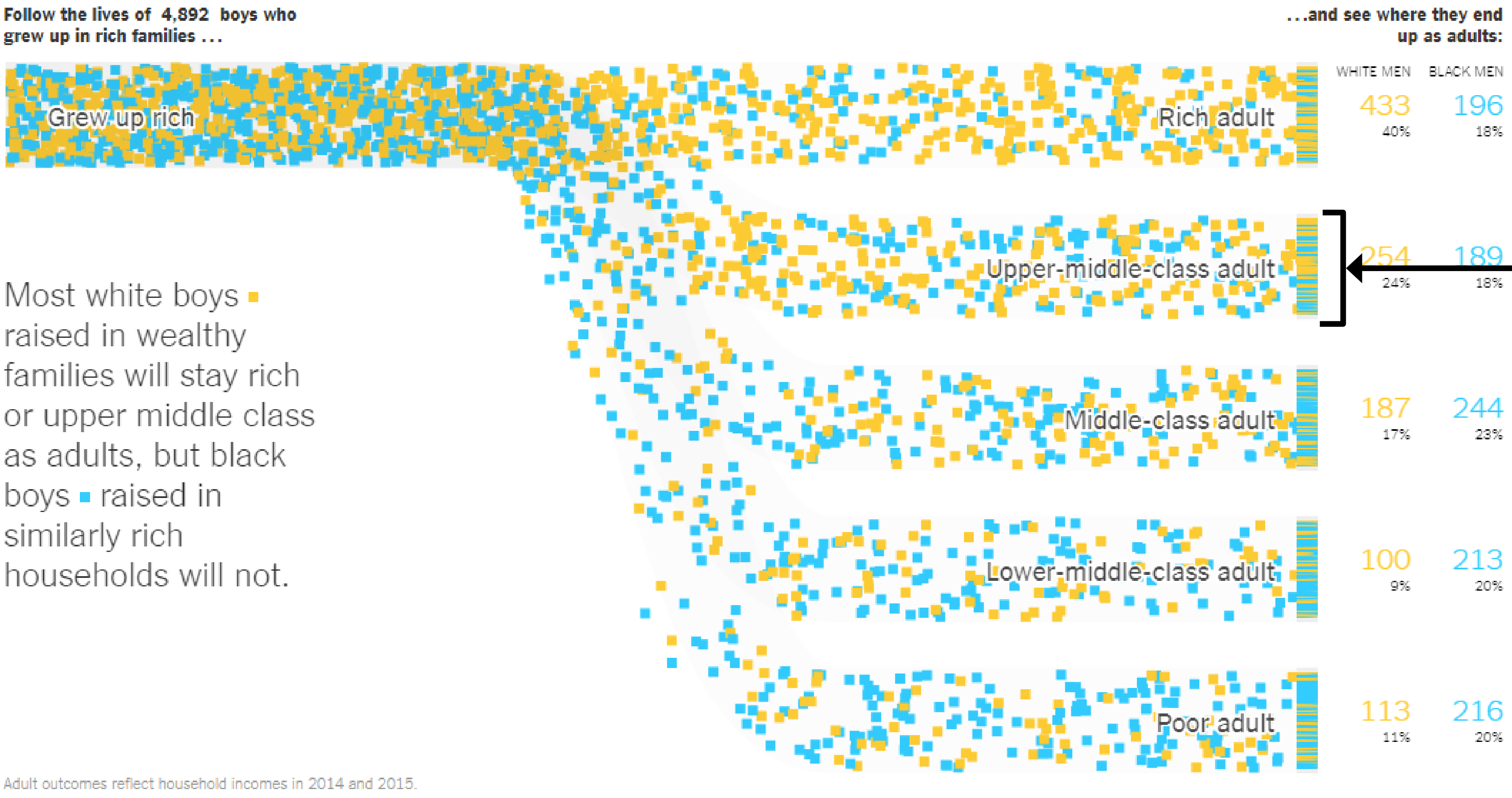


Income of black boys from wealthy families

[Badger, Miller, Pearce, Quealy. Extensive Data Shows Punishing Reach of Racism for Black Boys, NYT Upshot, 2018, https://nyti.ms/2GGpFZw]

I want:

P(adult income | race)



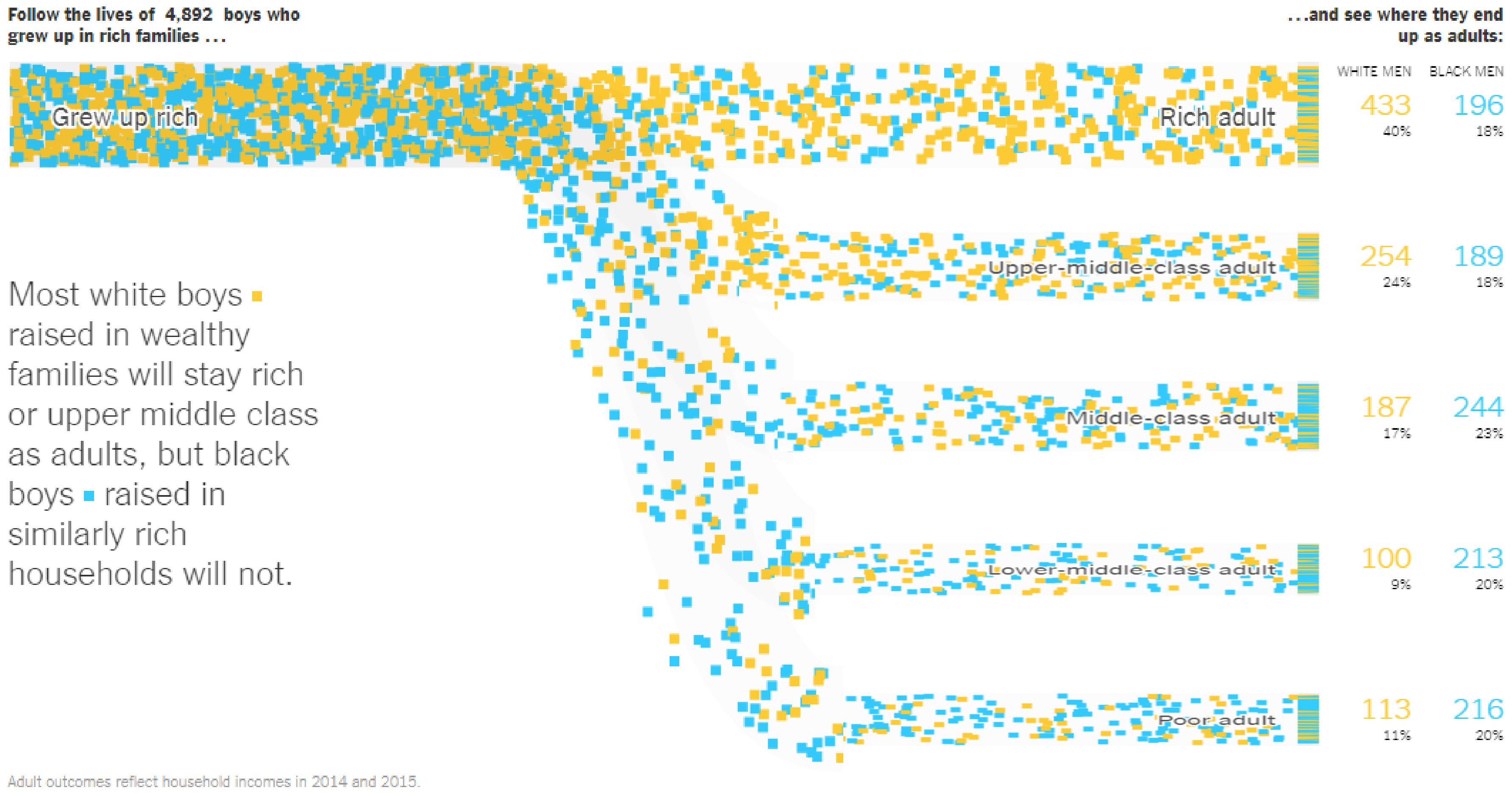
P(race | adult income)

Income of black boys from wealthy families

[Badger, Miller, Pearce, Quealy. Extensive Data Shows Punishing Reach of Racism for Black Boys, NYT Upshot, 2018, https://nyti.ms/2GGpFZw]

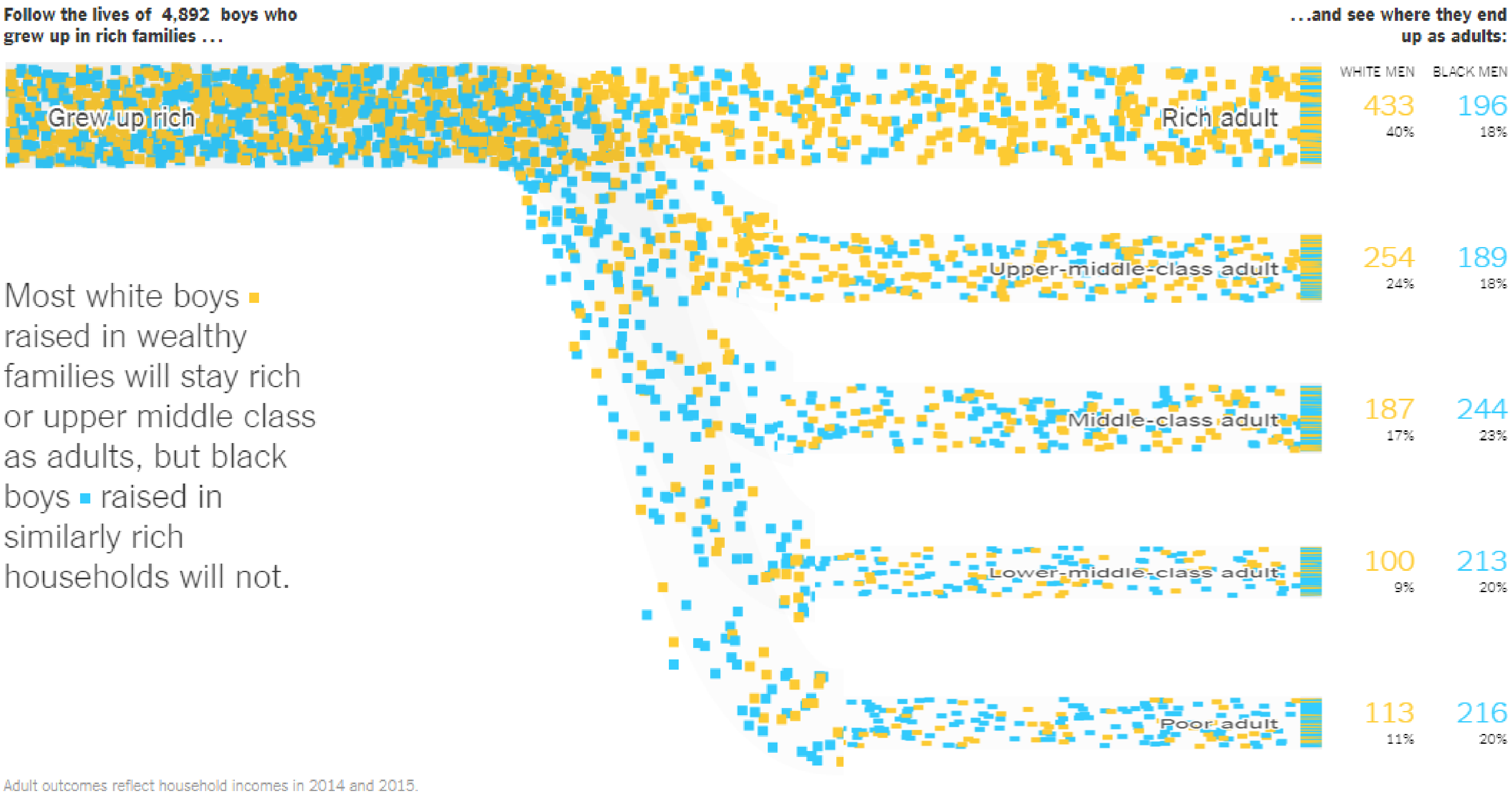
I want:

P(adult income | race)



Income of black boys from wealthy families

[Badger, Miller, Pearce, Quealy. Extensive Data Shows Punishing Reach of Racism for Black Boys, NYT Upshot, 2018, https://nyti.ms/2GGpFZw]

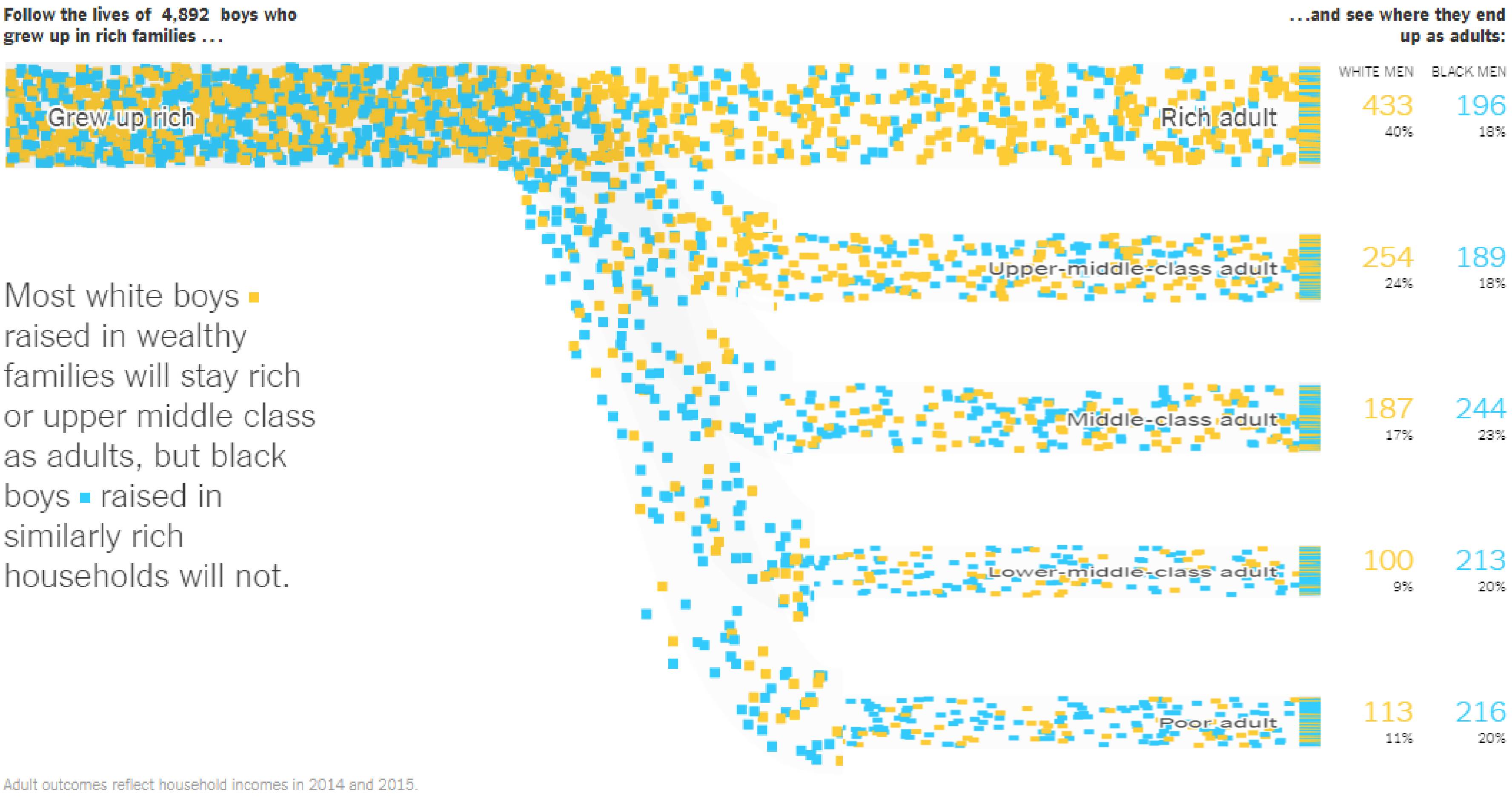


I want:

P(adult income | race)

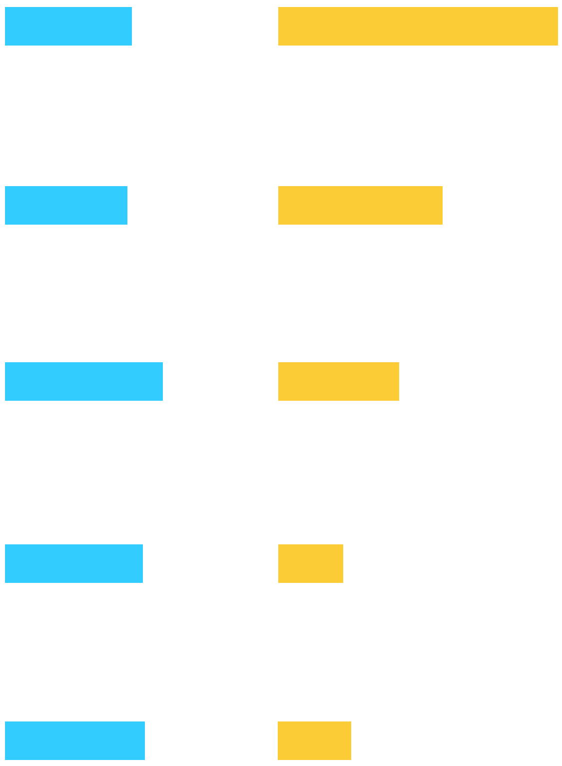
Income of black boys from wealthy families

[Badger, Miller, Pearce, Quealy. Extensive Data Shows Punishing Reach of Racism for Black Boys, NYT Upshot, 2018, https://nyti.ms/2GGpFZw]



I want:

P(adult income | race)



Building effective, complex, correct uncertainty visualizations is **hard**

Building effective, complex, correct uncertainty visualization is hard

Prototyping animation takes time, is brittle

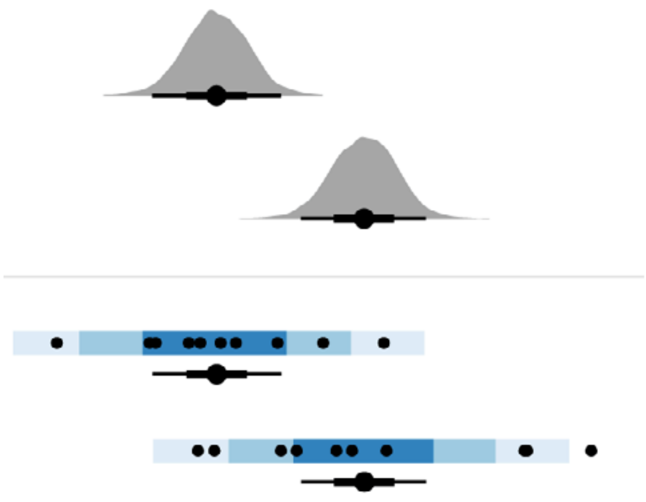
Specifications involve conditional probabilities (a pain!)

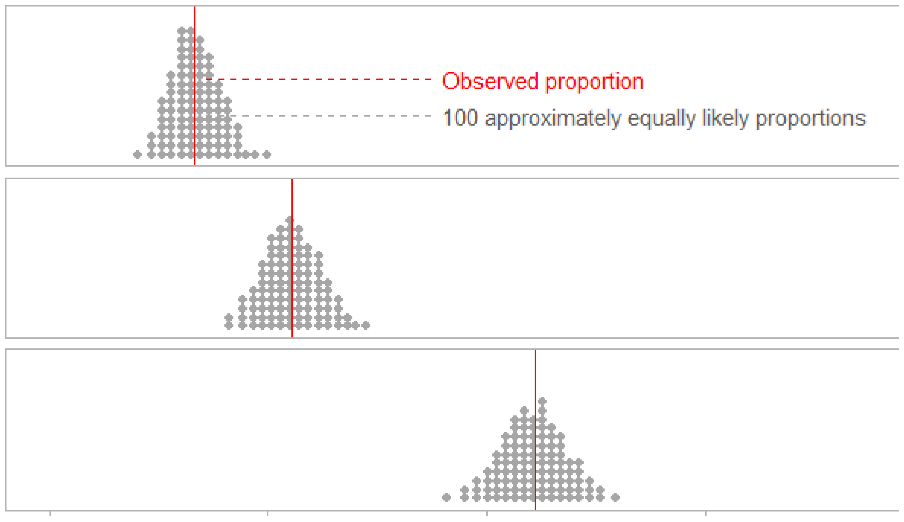
Need to be able to navigate the design space

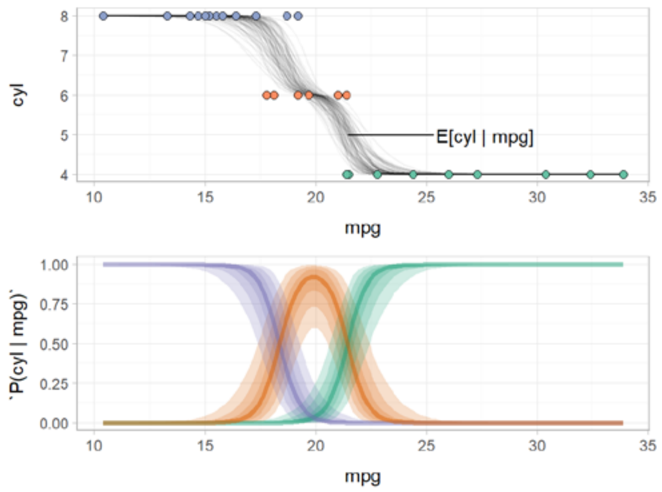
Building effective, complex, correct uncertainty visualization is hard

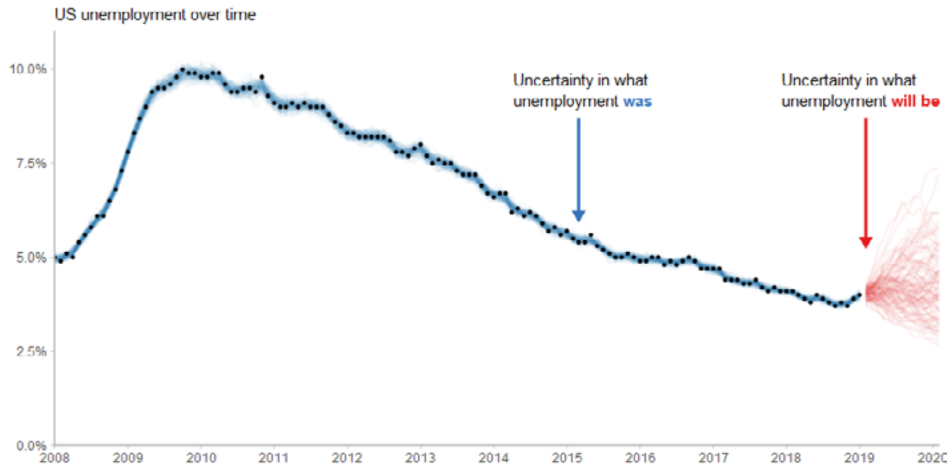
Tackling this on two fronts:

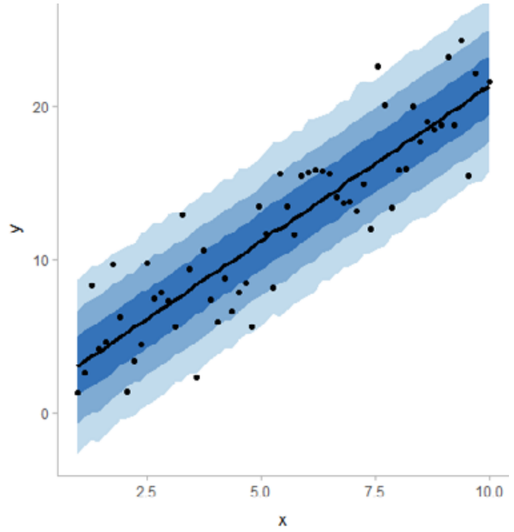
1. R packages for handling output from Bayesian models and visualizing uncertainty: tidybayes, ggdist





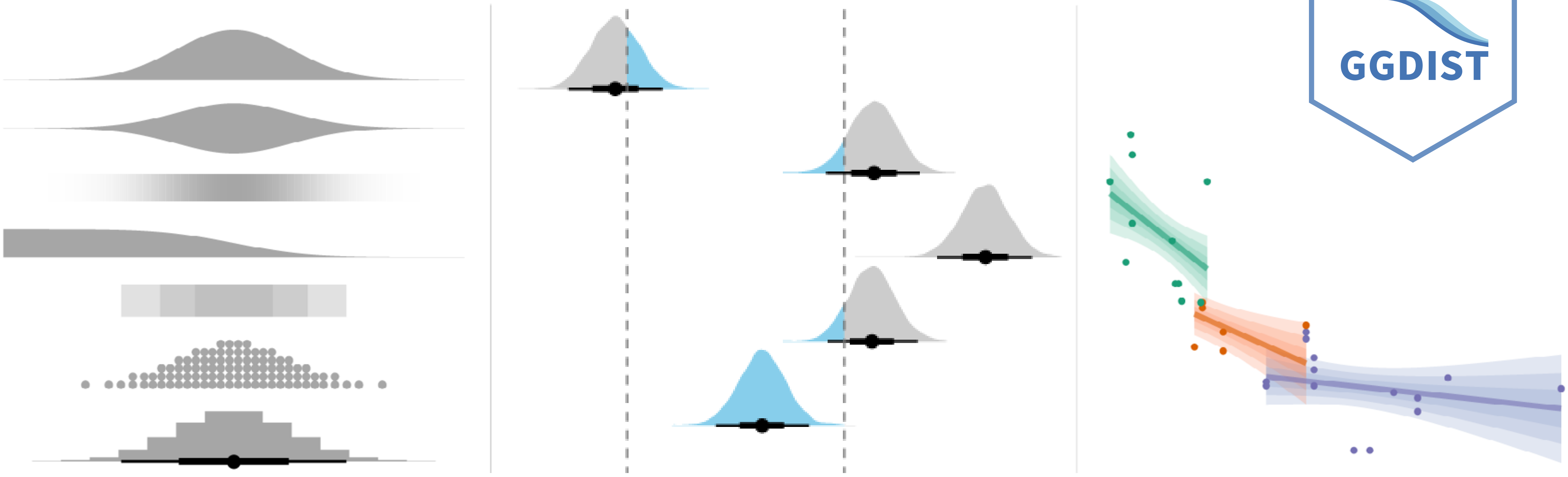






http://mjskay.github.io/tidybayes/

https://github.com/mjskay/uncertainty-examples



GGDIST

http://mjskay.github.io/ggdist/

Building effective, complex uncertainty visualization is hard

Tackling this on two fronts:

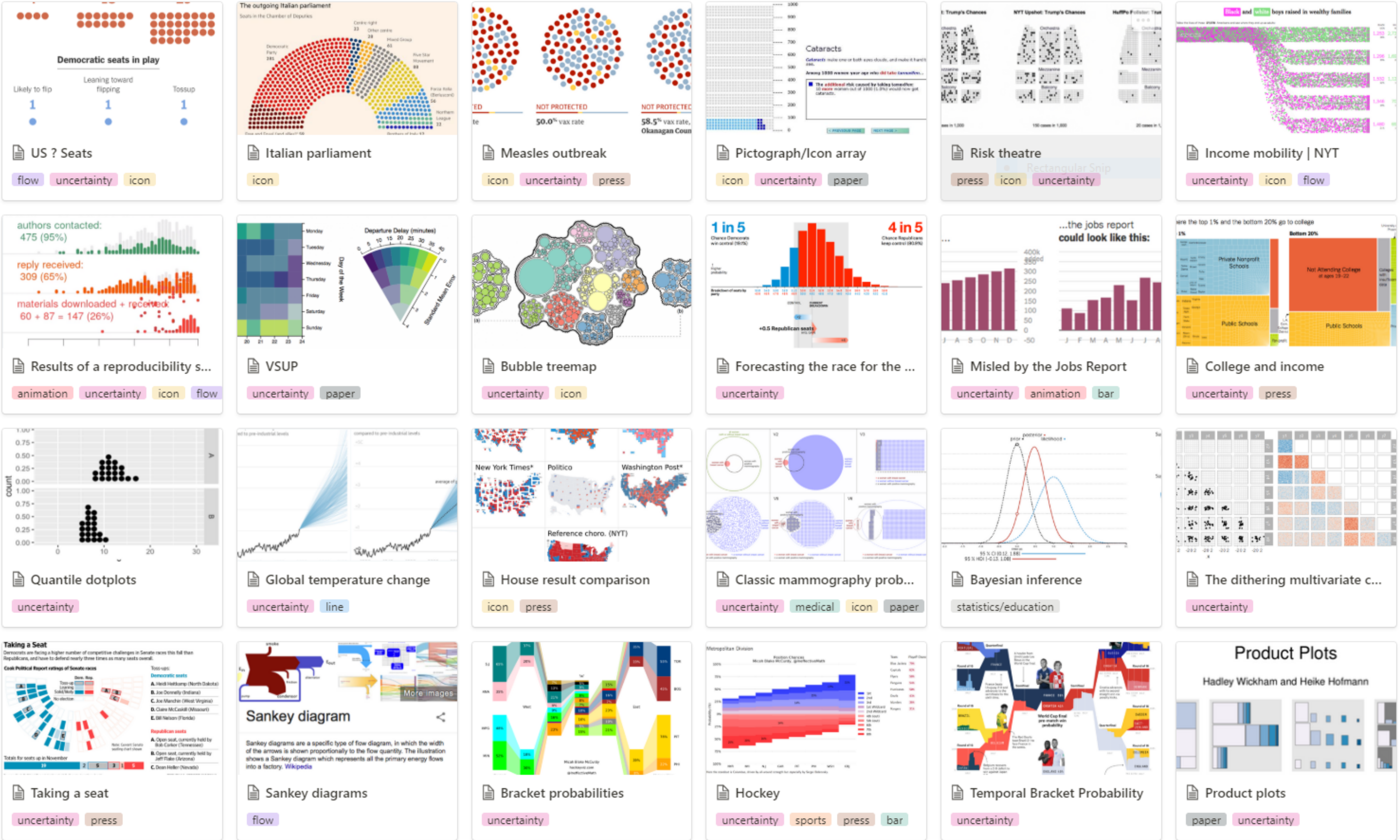
1. R packages for handling output from Bayesian models and visualizing uncertainty: tidybayes, ggdist

Building effective, complex uncertainty visualization is hard

Tackling this on two fronts:

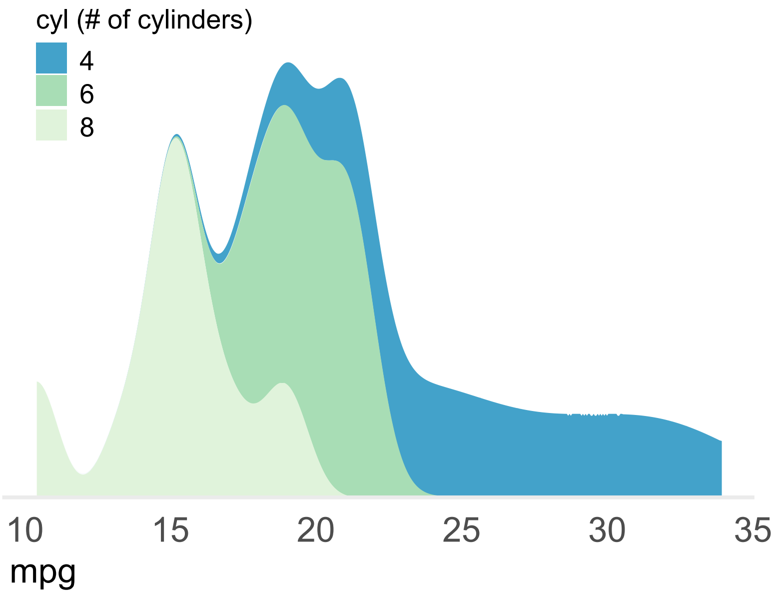
1. R packages for handling output from Bayesian models and visualizing uncertainty: tidybayes, ggdist

2. A probabilistic grammar of graphics for uncertainty visualization specification [Xiaoying Pu]



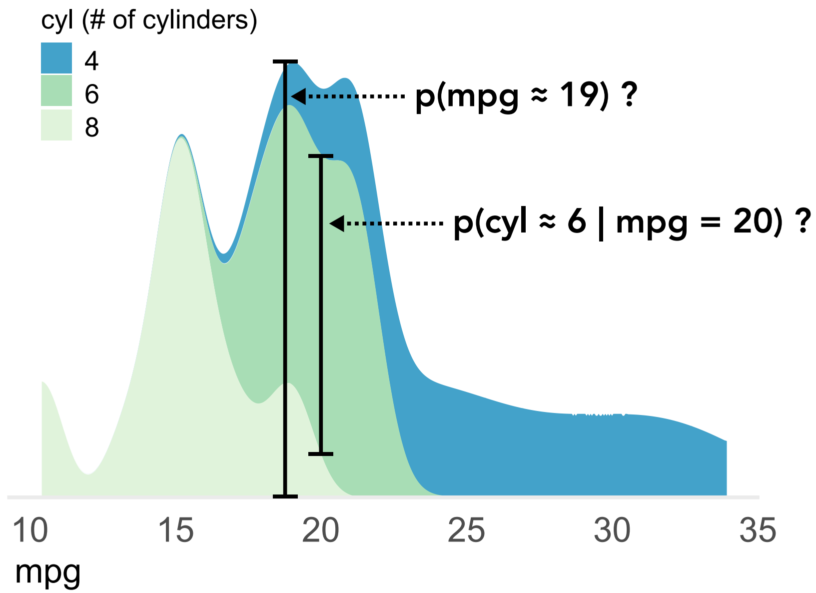
http://mucollective.co/uncertainty_collection/ [collected by Xiaoying Pu, Puhe Liang]

Grammar of graphics



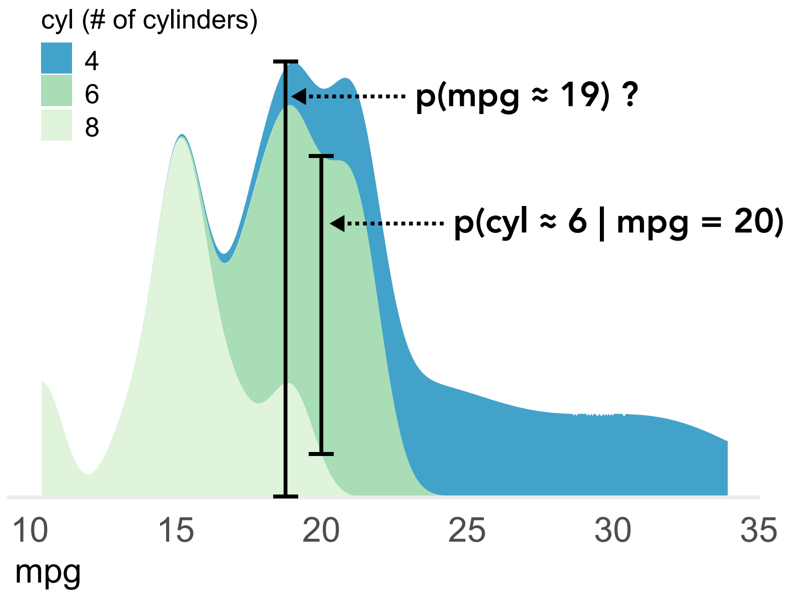

```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    fill = cyl
  ),
  position = "stack"
)
```

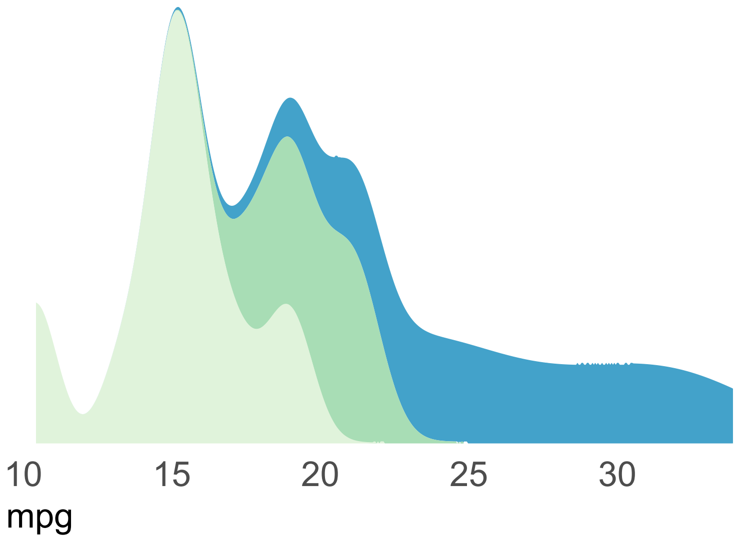
Grammar of graphics




```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    fill = cyl
  ),
  position = "stack"
)
```

Grammar of graphics

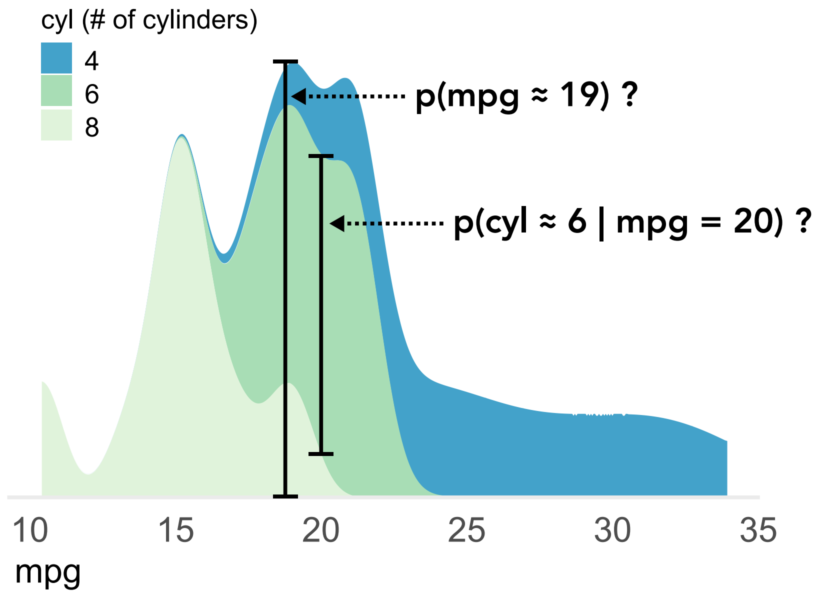


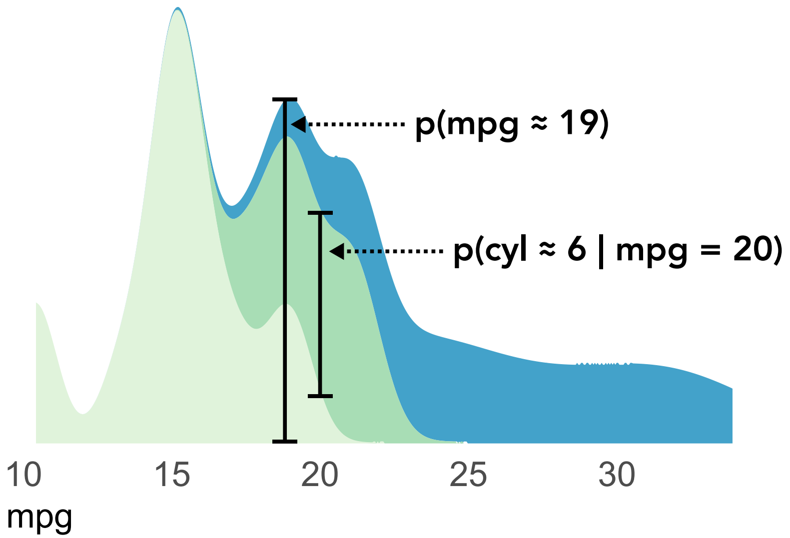



```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    fill = cyl
  ),
  position = "stack"
)
```

```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    y = stat(density * n),
    fill = cyl
  ),
  position = "stack"
)
```

Grammar of graphics



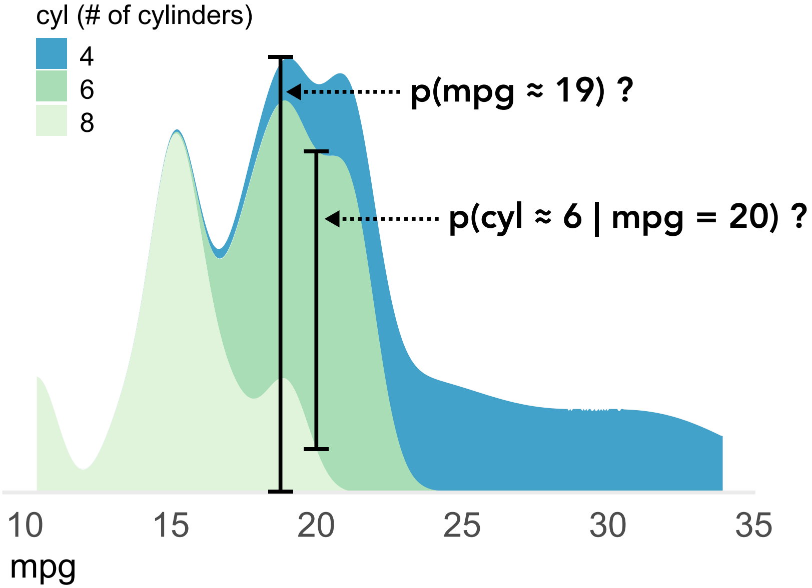


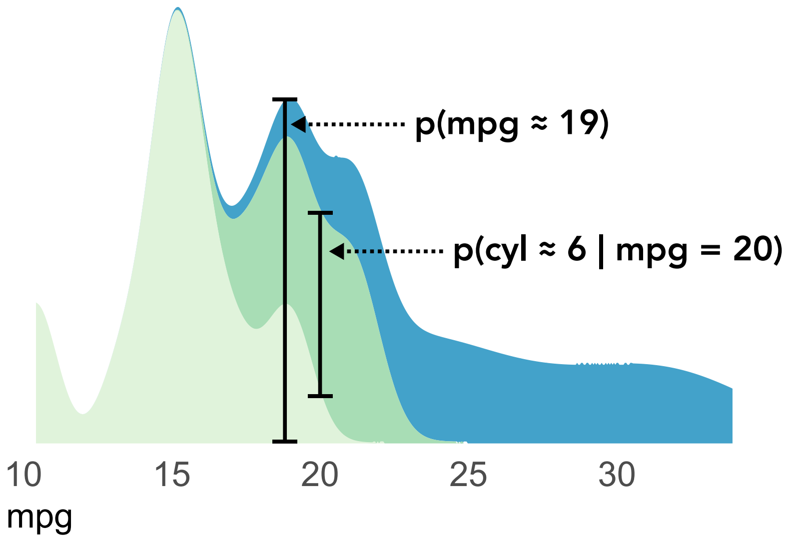

```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    fill = cyl
  ),
  position = "stack"
)
```

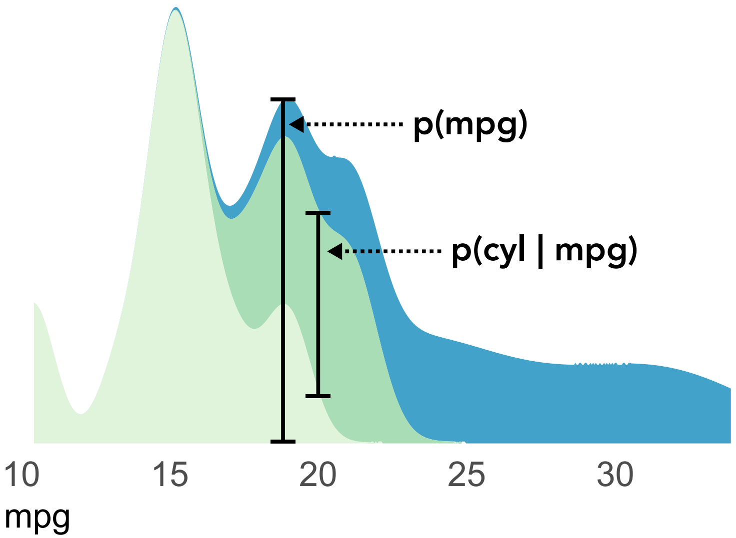
```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    y = stat(density * n),
    fill = cyl
  ),
  position = "stack"
)
```

Probabilistic grammar of graphics (PGoG)

[Pu, **Kay**. A Probabilistic Grammar of Graphics. CHI 2020, **Honorable Mention**]



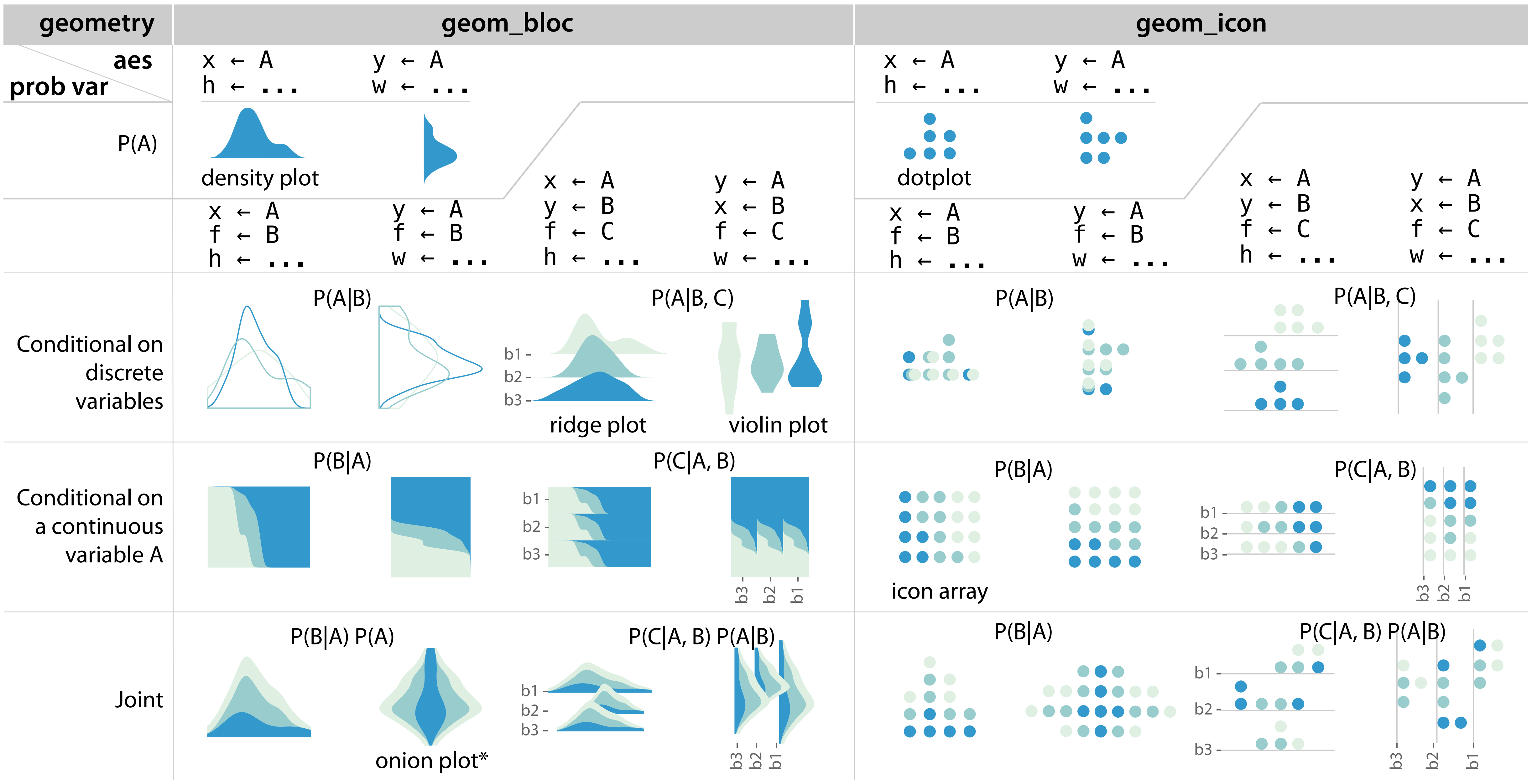





```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    fill = cyl
  ),
  position = "stack"
)
```

```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    y = stat(density * n),
    fill = cyl
  ),
  position = "stack"
)
```

```
ggplot(mtcars) +
geom_density(
  aes(
    x = mpg,
    height = P(cyl|mpg)*P(mpg),
    fill = cyl
  )
)
```



geometry
geom_bloc
geom_icon
aes
prob var
x ← A
h ← ...
y ← A
w ← ...
P(A)
density plot
dotplot
x ← A
f ← B
h ← ...
y ← A
f ← B
w ← ...
x ← A
y ← B
f ← C
h ← ...
y ← A
x ← B
f ← C
w ← ...
Conditional on discrete variables
P(A|B)
P(A|B, C)
b1
b2
b3
ridge plot
violin plot
Conditional on a continuous variable A
P(B|A)
P(C|A, B)
icon array
Joint
P(B|A) P(A)
P(C|A, B) P(A|B)
onion plot*

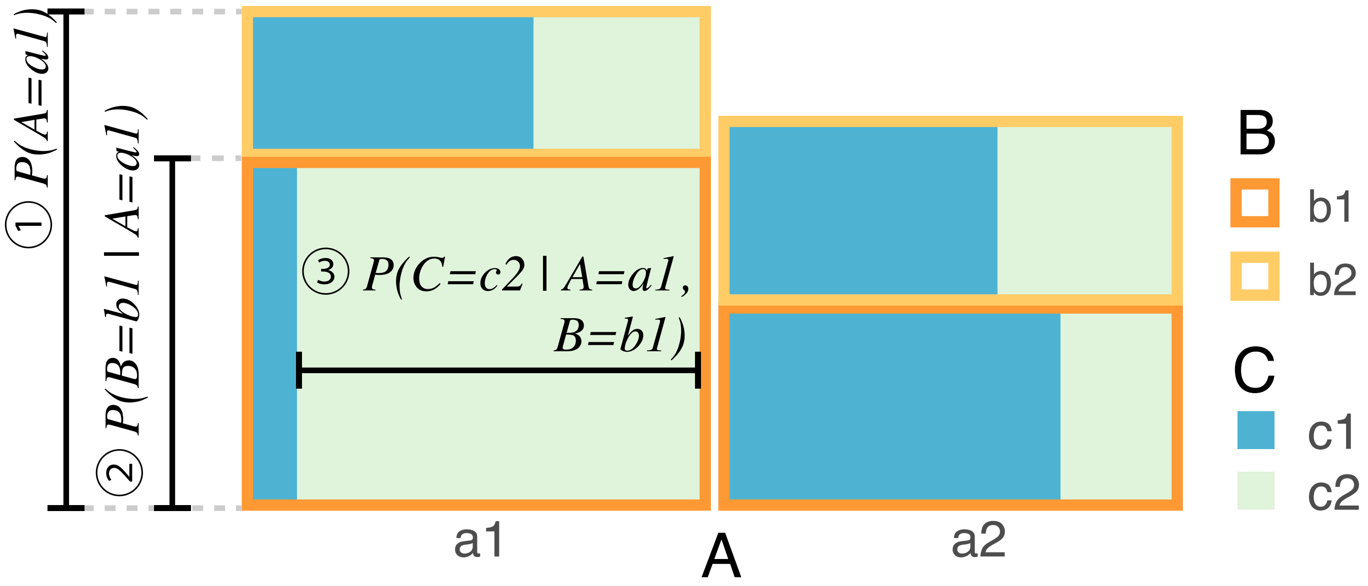
PGoG specification and validation

[Pu, **Kay**. A Probabilistic Grammar of Graphics. CHI 2020, **Honorable Mention**]

PGoG specification

```
geom_bloc:
 height ← P(A)
          P(B|A)
  width ← P(C|A,B)
      x ← A
  color ← B
   fill ← C
```

Resulting area plot



PGoG specification and validation

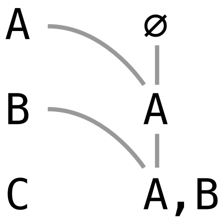
[Pu, **Kay**. A Probabilistic Grammar of Graphics. CHI 2020, **Honorable Mention**]

PGoG specification

```
geom_bloc:
 height ← P(A)
          P(B|A)
  width ← P(C|A,B)
      x ← A
  color ← B
   fill ← C
```

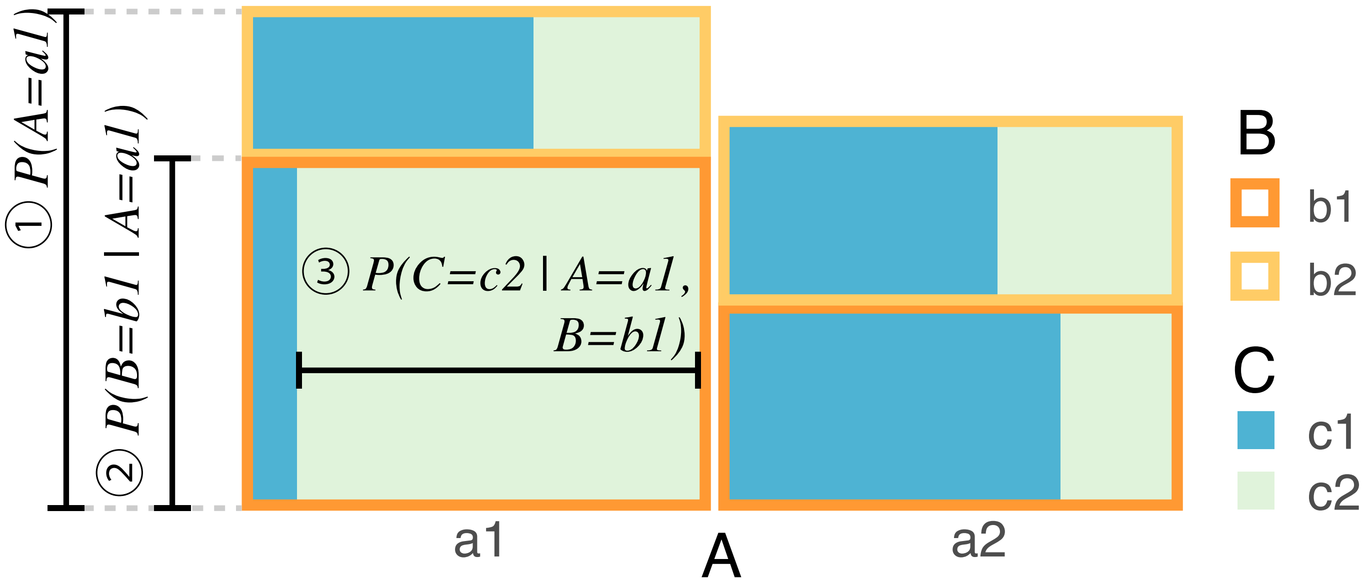
Validation

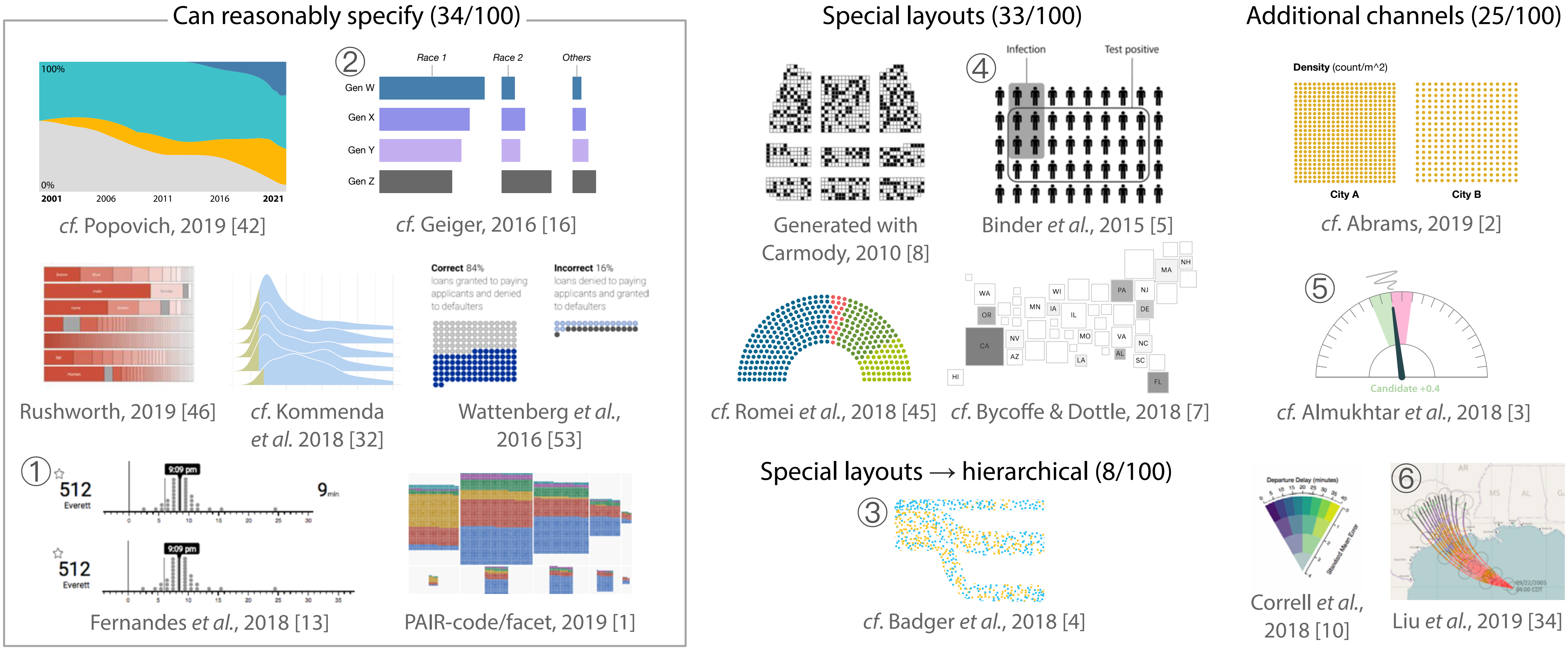
$$\begin{aligned} & P(A, B, C) \\ &= P(A) && \text{...①} \\ &\times P(B|A) && \text{...②} \\ &\times P(C|A,B) && \text{...③} \end{aligned}$$



(chain)

Resulting area plot





Can reasonably specify (34/100)
100%
0%
2001
2006
2011
2016
2021
cf. Popovich, 2019 [42]
Race 1
Race 2
Others
Gen W
Gen X
Gen Y
Gen Z
cf. Geiger, 2016 [16]
Rushworth, 2019 [46]
cf. Kommenda et al. 2018 [32]
Correct 84%
loans granted to paying applicants and denied to defaulters
Incorrect 16%
loans denied to paying applicants and granted to defaulters
Wattenberg et al., 2016 [53]
512
Everett
9:09 pm
9min
Fernandes et al., 2018 [13]
PAIR-code/facet, 2019 [1]
Special layouts (33/100)
Generated with Carmody, 2010 [8]
Infection
Test positive
Binder et al., 2015 [5]
cf. Romei et al., 2018 [45]
cf. Bycoffe & Dottle, 2018 [7]
Special layouts → hierarchical (8/100)
cf. Badger et al., 2018 [4]
Additional channels (25/100)
Density (count/m^2)
City A
City B
cf. Abrams, 2019 [2]
Candidate +0.4
cf. Almukhtar et al., 2018 [3]
Departure Delay (minutes)
Standard Mean Error
Correll et al., 2018 [10]
Liu et al., 2019 [34]

Future of PGoG

[NSF Award #1910431]

Expanding the space of supported visualizations

Animation, time

Specifications directly from probabilistic programs

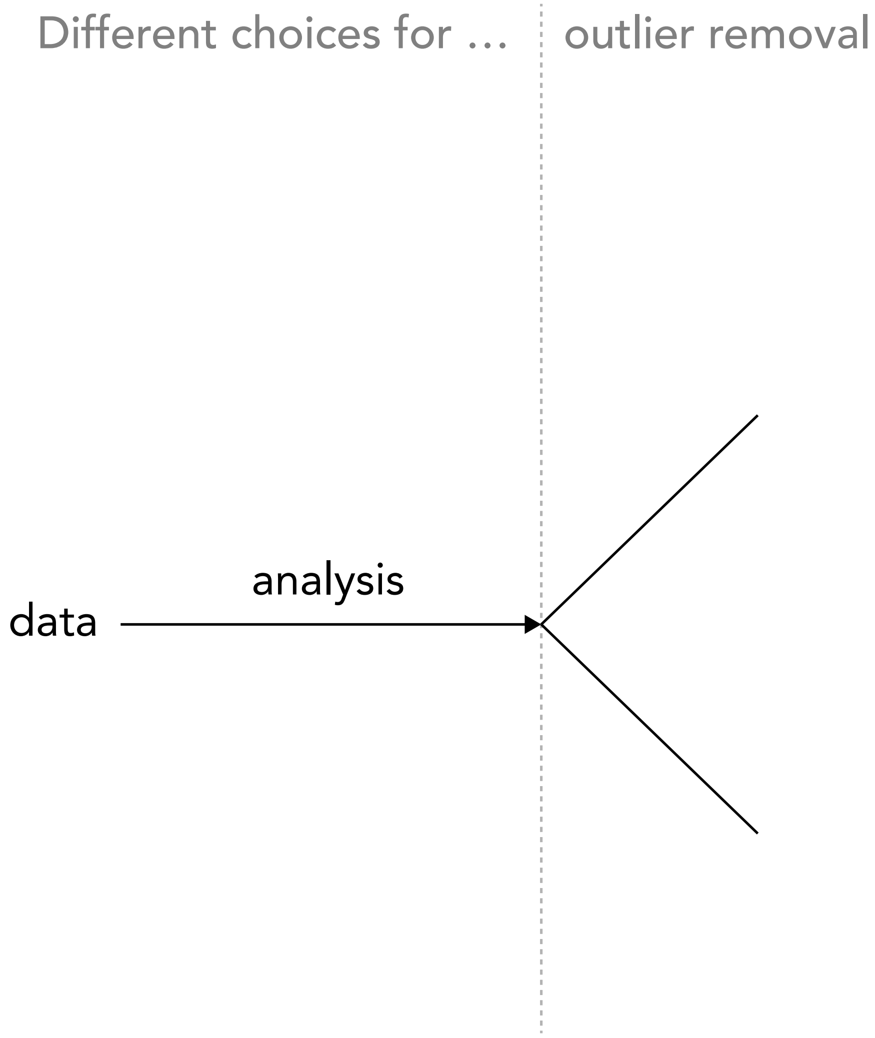
Changes in ggdist

Let’s step back from
strictly probabilistic uncertainty

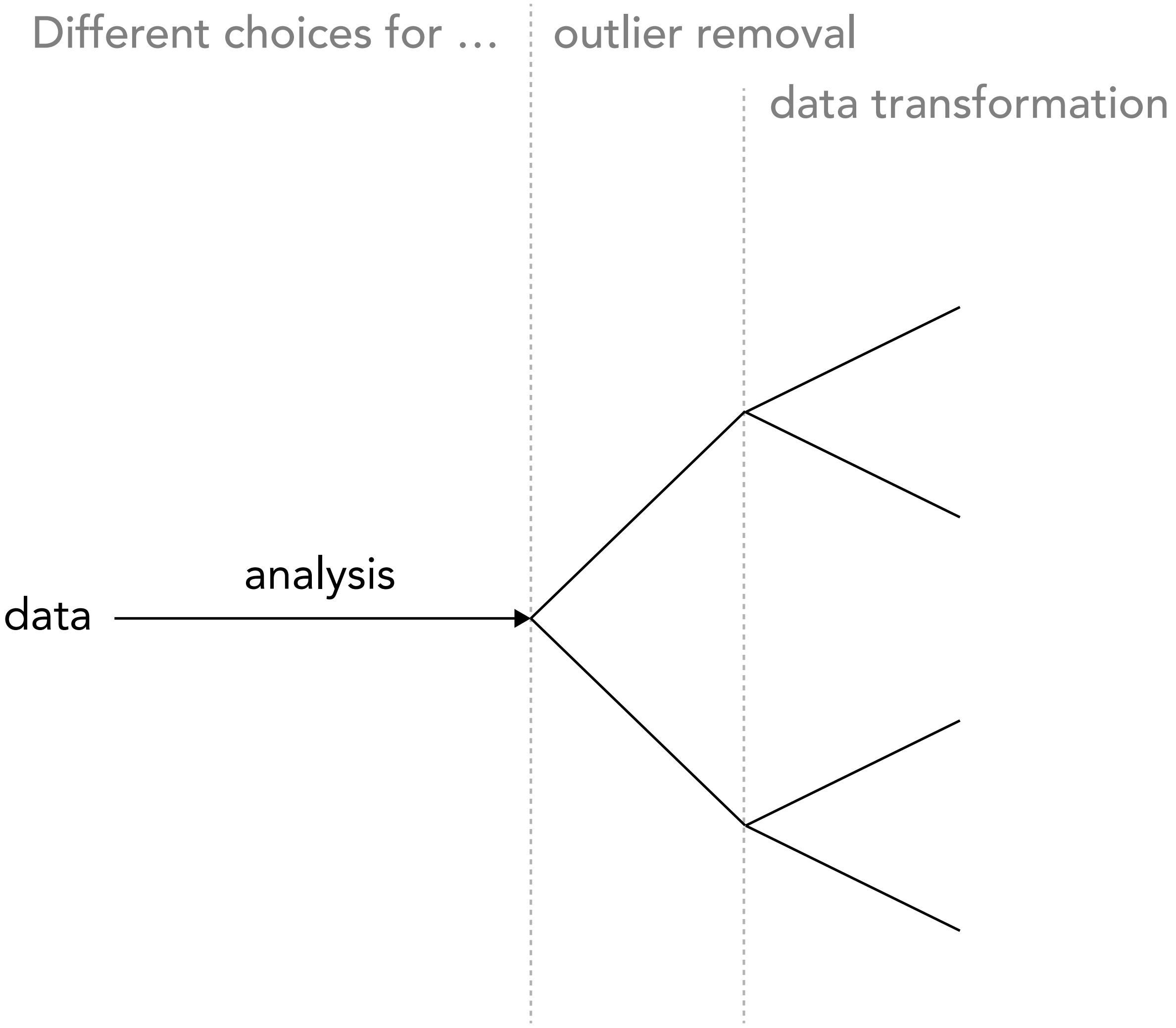
$$\text{data} \xrightarrow{\text{analysis}} p < 0.05$$

$$\text{data} \xrightarrow{\text{analysis}}$$

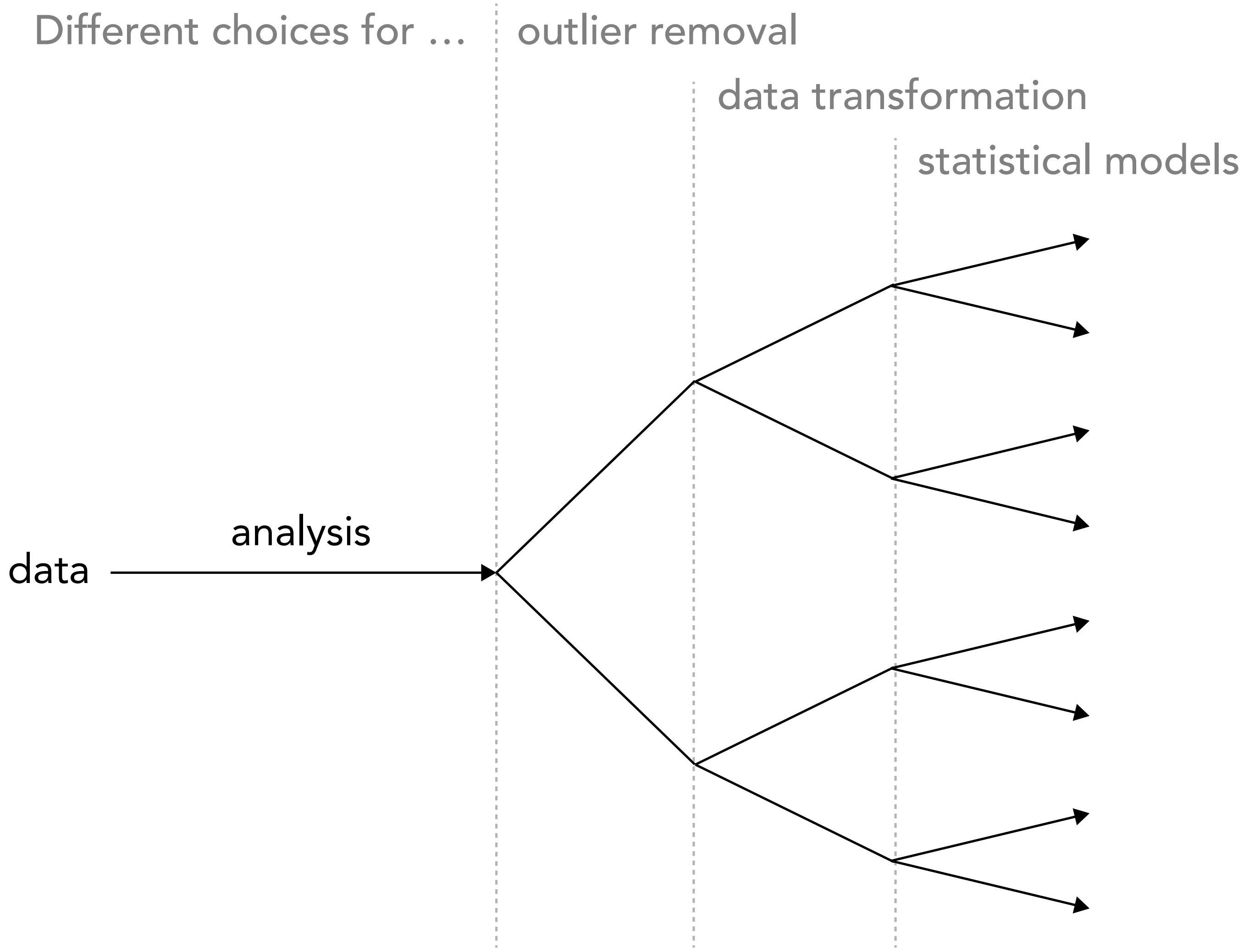
Garden of forking paths [Gelman and Loken 2014]



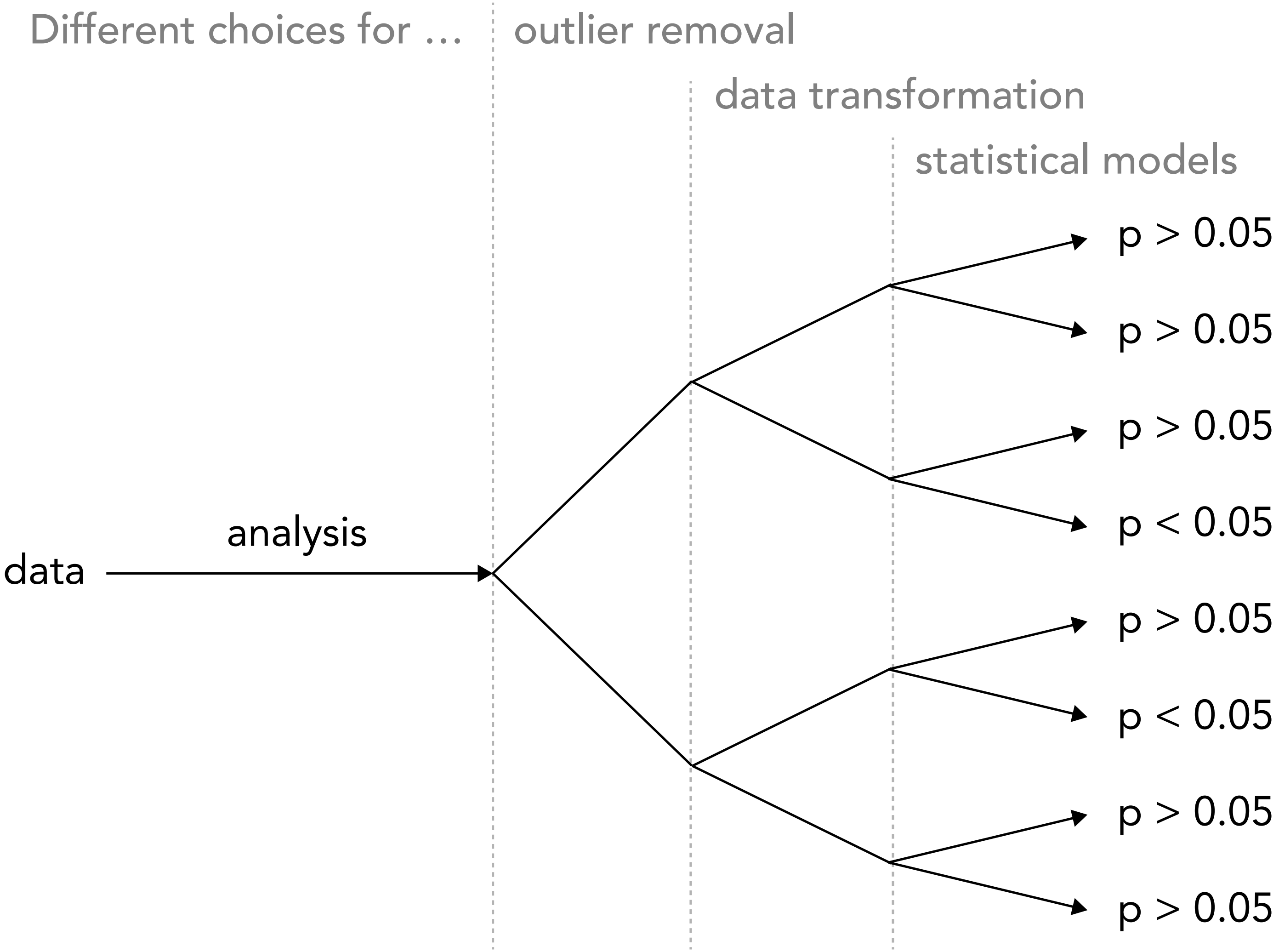
Garden of forking paths [Gelman and Loken 2014]



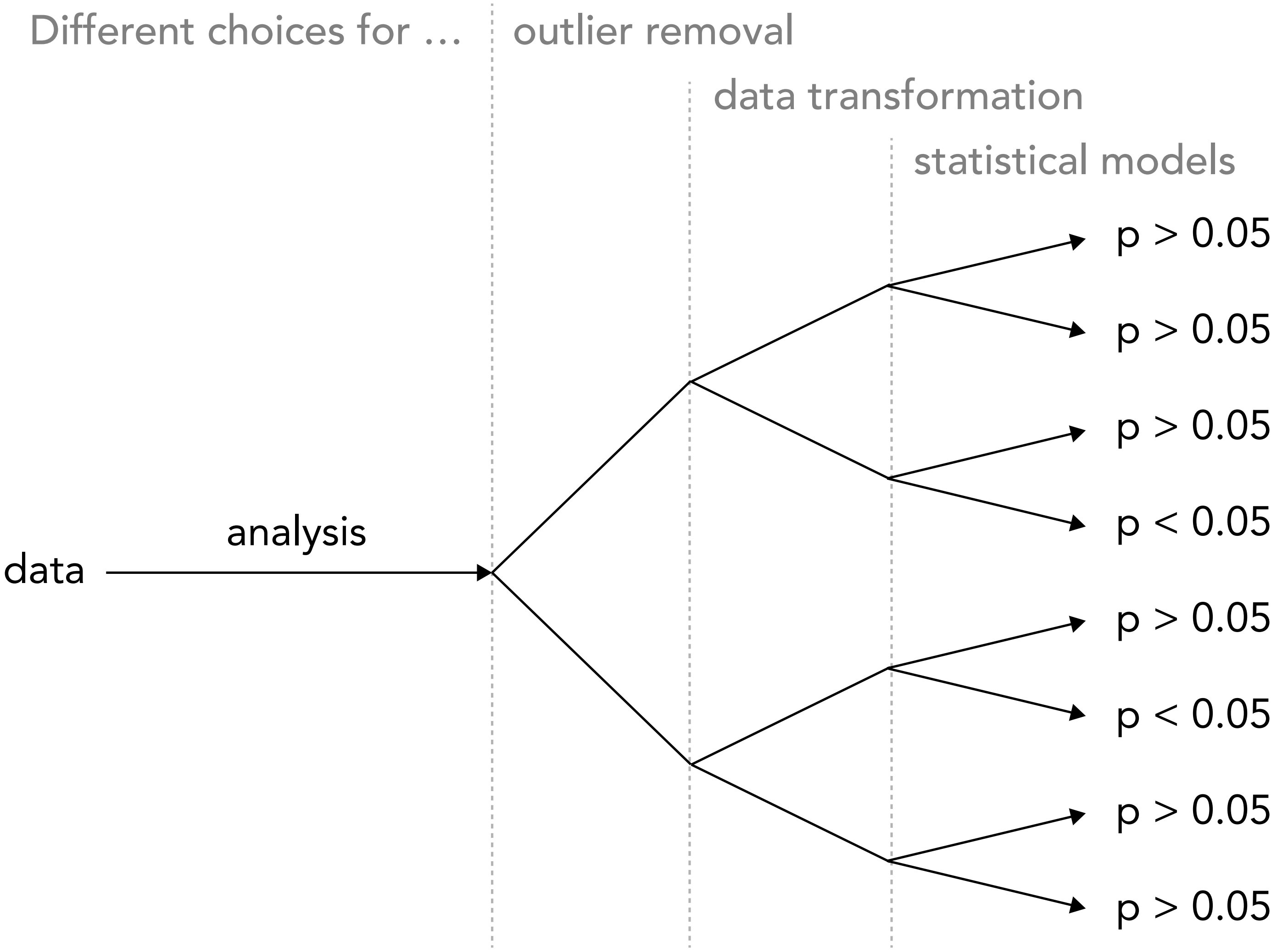
Garden of forking paths [Gelman and Loken 2014]



Garden of forking paths [Gelman and Loken 2014]

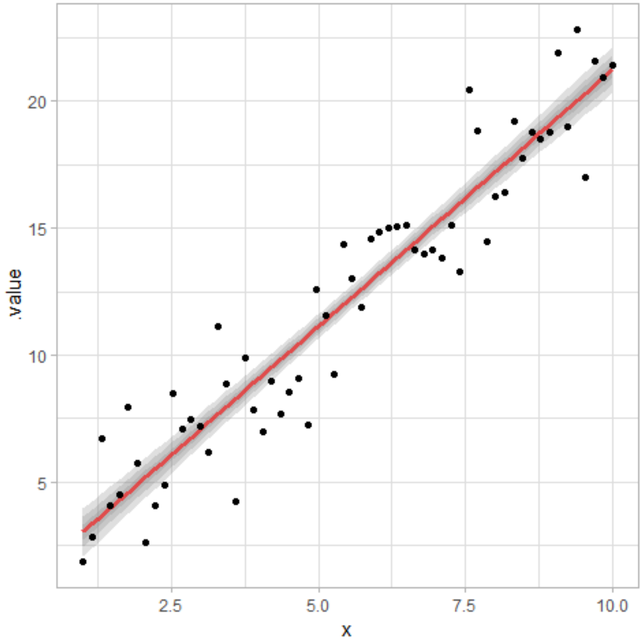


Garden of forking paths [Gelman and Loken 2014]

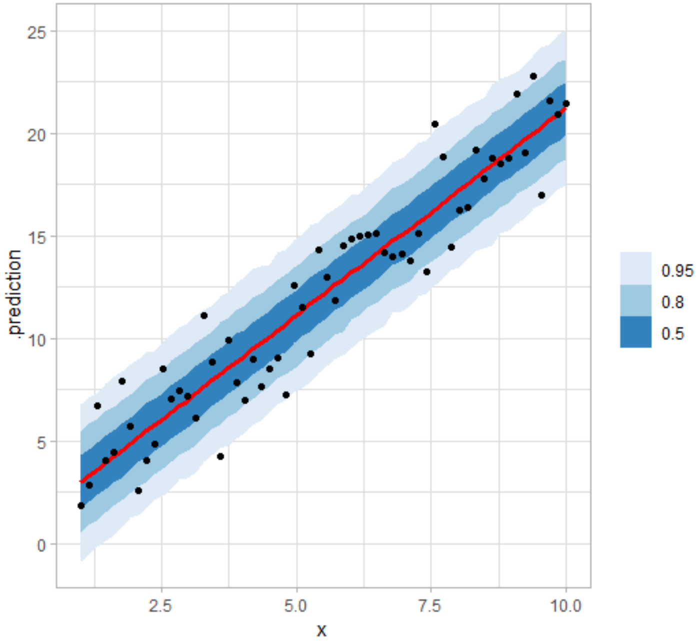


This is **specification uncertainty**

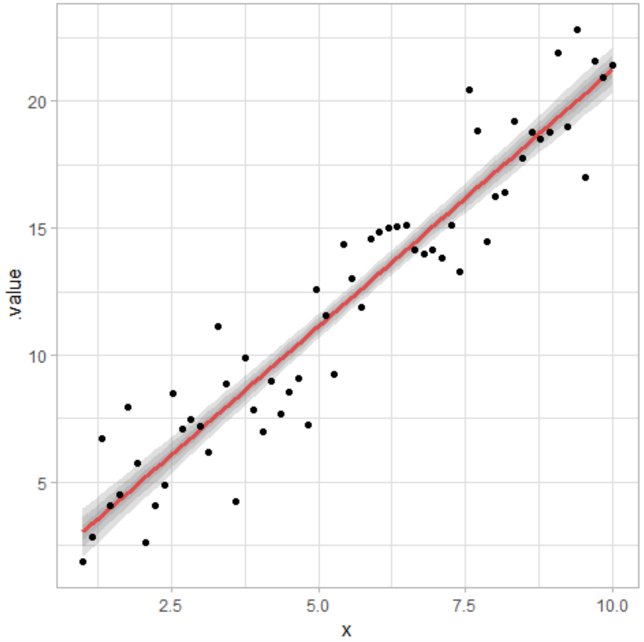
Parameter uncertainty



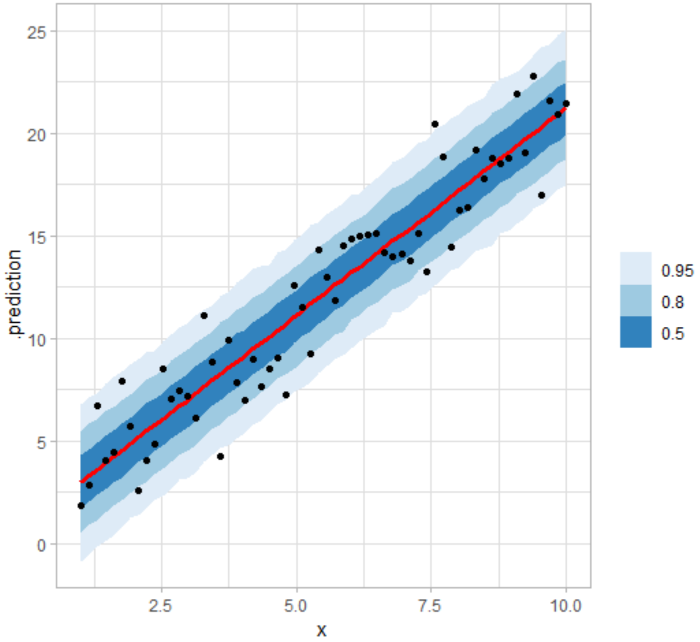
Predictive uncertainty



Parameter uncertainty



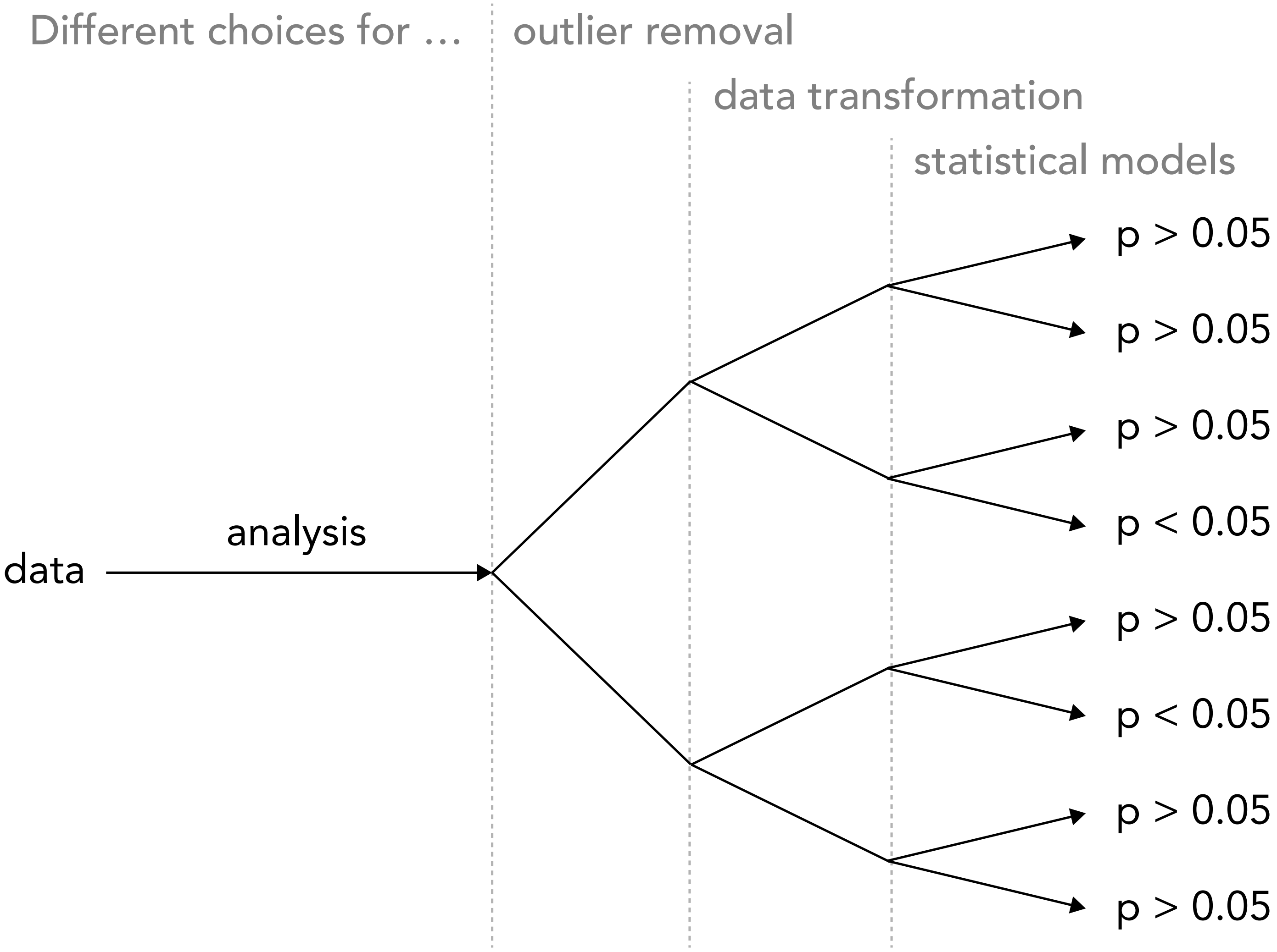
Predictive uncertainty



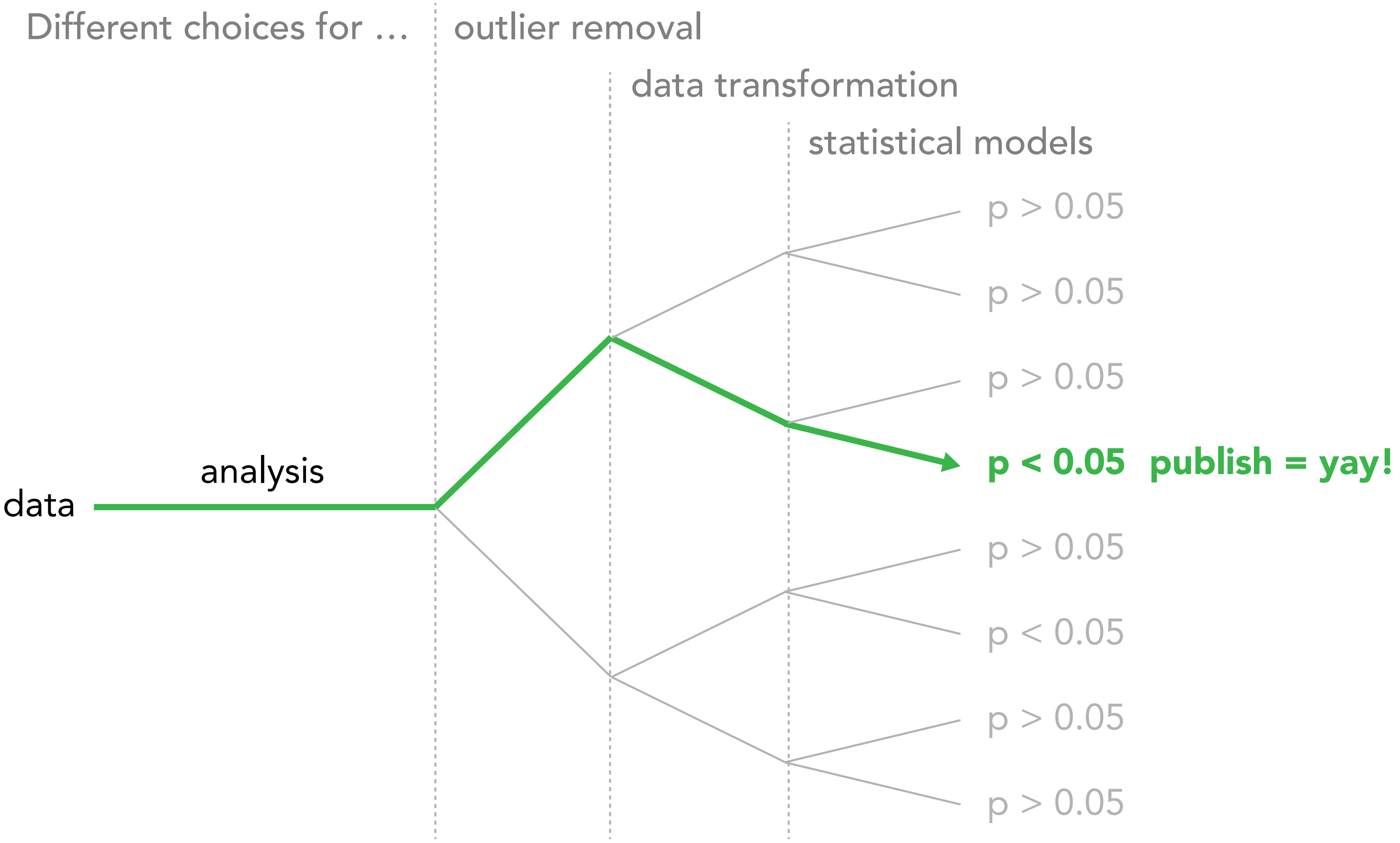
Specification uncertainty

How well does this describe reality?

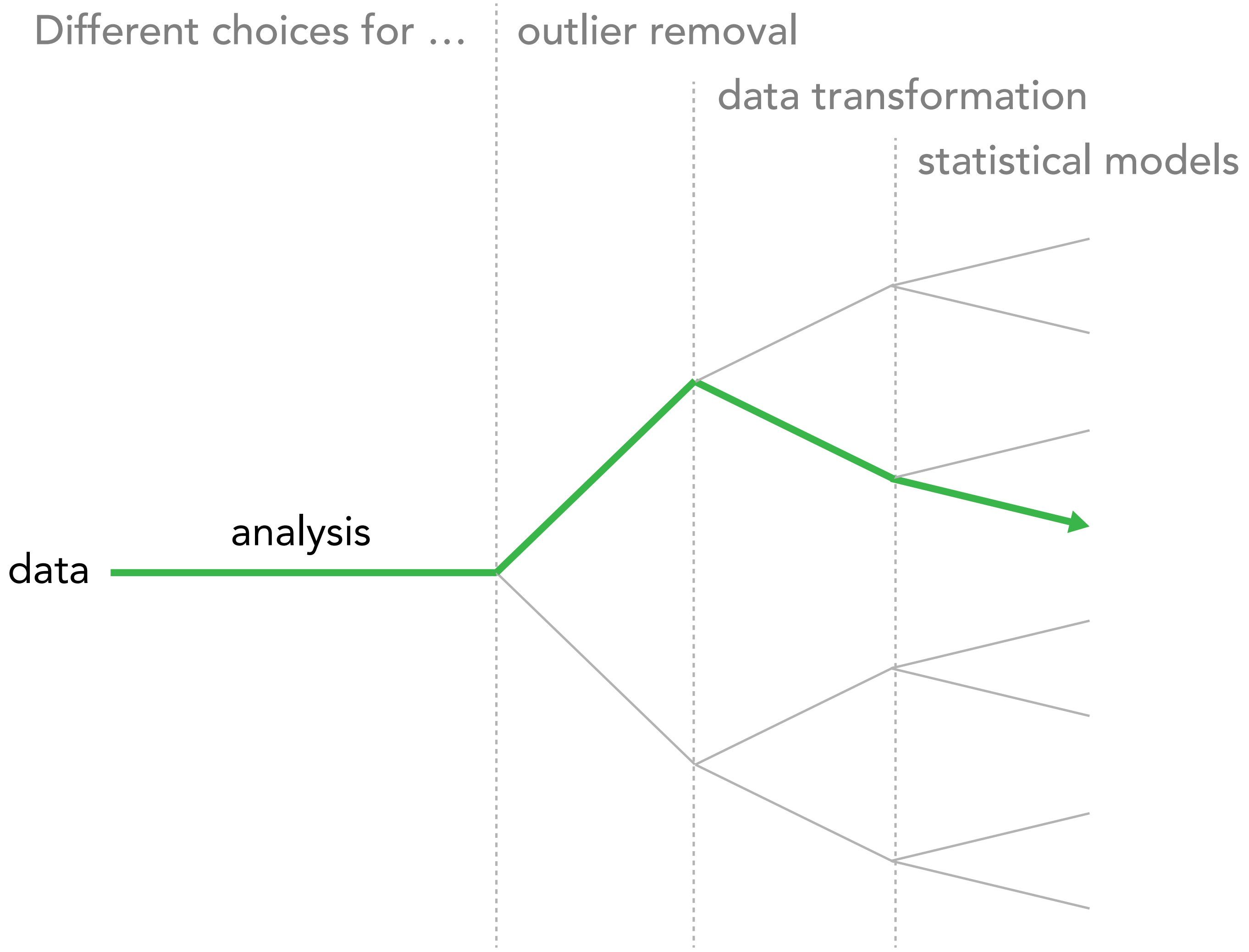
Garden of forking paths [Gelman and Loken 2014]



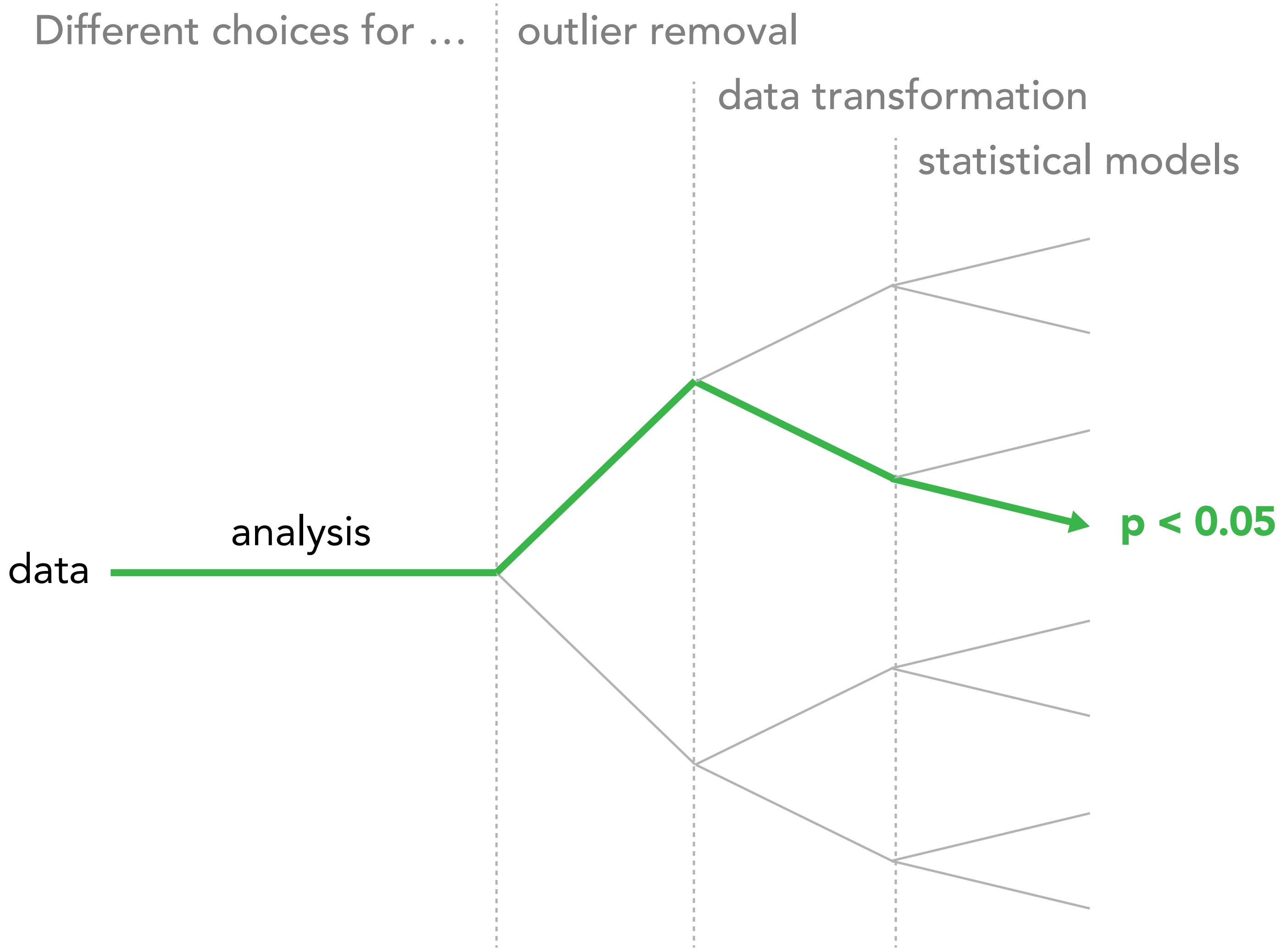
Garden of forking paths [Gelman and Loken 2014]



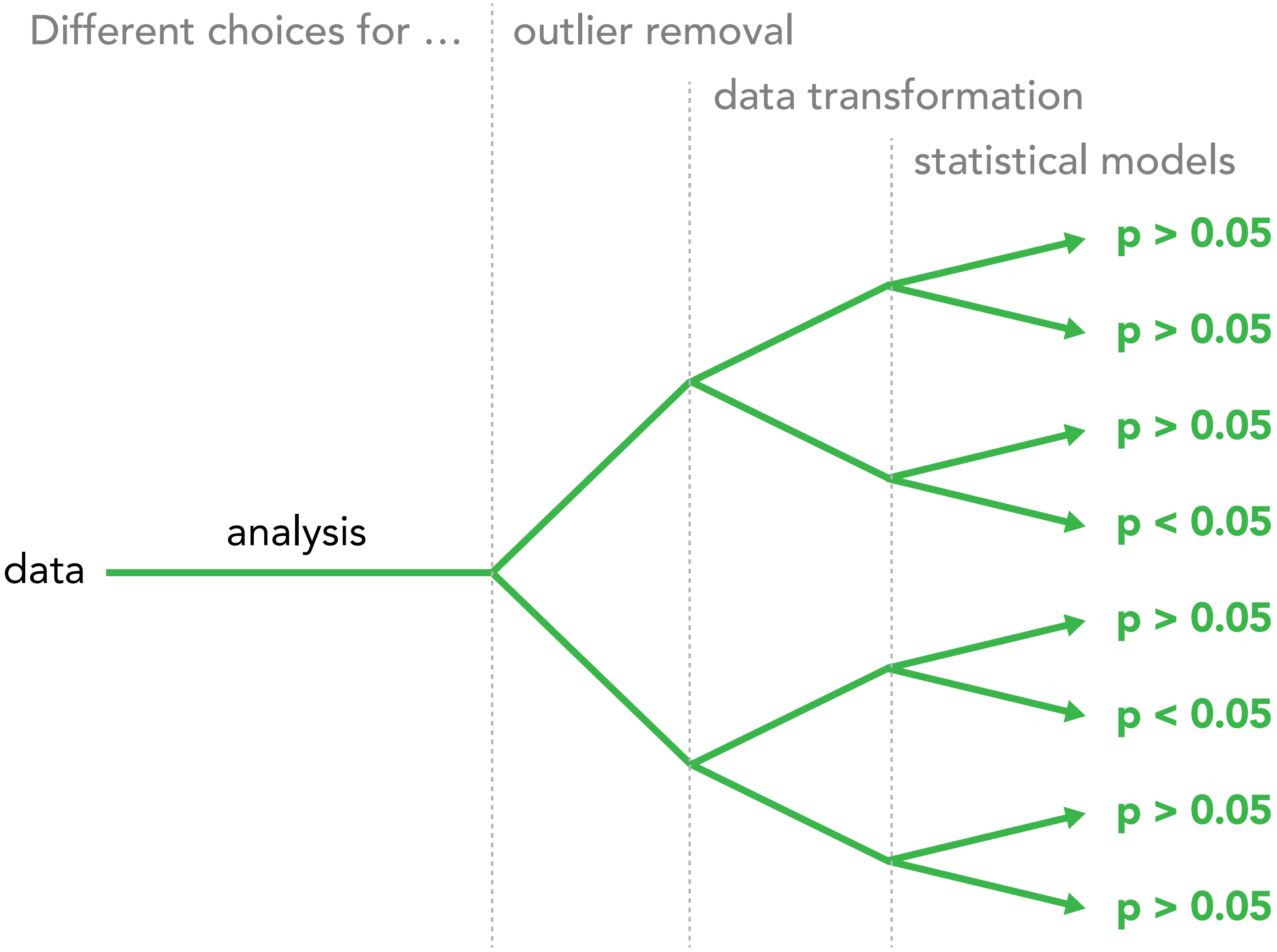
(pre-registration / hold-out)



(pre-registration / hold-out)



(multiverse analysis) [Steegen, Tuerlinckz, Gelman, Vanpaemel 2014]



Religiosity (Study 2)

				R1					R2					R3		
			F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
NMO1	ECL1	EC1	0	0	0.03	0.02	0.08	0.02	0.02	0.03	0.05	0.06	0	0	0.02	0.01	0.04
		EC2	0.01	0	0.08	0.03	0.16	0.07	0.05	0.18	0.24	0.41	0.01	0	0.09	0.06	0.23
	ECL2	EC1	0	0	0.06	0.04	0.37	0.02	0.03	0.07	0.08	0.21	0	0	0.05	0.03	0.23
		EC2	0.01	0	0.13	0.08	0.44	0.06	0.03	0.22	0.24	0.52	0.01	0	0.14	0.09	0.43
NMO2	ECL1	EC1	0	0	0.03	0.01	0.08	0.15	0.07	0.17	0.07	0.14	0.02	0.01	0.06	0.02	0.07
		EC2	0	0	0.02	0.01	0.06	0.2	0.05	0.42	0.23	0.44	0.03	0	0.14	0.05	0.19
	ECL3	EC 1	0.01	0.01	0.05	0.01	0.1	0.39	0.2	0.45	0.11	0.26	0.08	0.04	0.17	0.03	0.13
		EC 2	0.01	0	0.05	0.02	0.11	0.33	0.09	0.59	0.26	0.55	0.09	0.02	0.26	0.08	0.27
NMO3	ECL1	EC1	0.01	0.01	0.02	0.1	0.28	0.11	0.09	0.43	0.26	0.85	0.02	0.02	0.12	0.12	0.51
		EC2	0.01	0.01	0	0.07	0.06	0.07	0.1	0.11	0.14	0.23	0.01	0.02	0.02	0.06	0.08
	ECL2	EC1	0.02	0.01	0.06	0.11	0.36	0.06	0.04	0.3	0.13	0.66	0.02	0.01	0.13	0.07	0.46
		EC2	0.02	0.01	0.02	0.15	0.13	0.04	0.05	0.07	0.07	0.16	0.01	0.02	0.03	0.05	0.09
	ECL3	EC1	0.07	0.04	0.12	0.09	0.16	0.16	0.11	0.54	0.32	0.77	0.07	0.04	0.25	0.13	0.39
		EC2	0.02	0.02	0.01	0.06	0.02	0.06	0.1	0.07	0.16	0.17	0.02	0.03	0.02	0.07	0.05

Social political attitudes

				R1					R2					R3		
			F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
NMO1	ECL1	EC 1	0	0	0	0	0	0.18	0.19	0.3	0.07	0.21	0.01	0.01	0.04	0	0.03
		EC 2	0.01	0.01	0.02	0	0.02	0.93	0.76	0.99	0.52	0.78	0.24	0.15	0.37	0.06	0.23
	ECL2	EC 1	0	0	0	0	0	0.17	0.26	0.3	0.13	0.27	0.02	0.02	0.06	0.01	0.05
		EC 2	0.02	0.01	0.04	0	0.01	0.89	0.86	0.98	0.47	0.66	0.24	0.2	0.42	0.05	0.19
NMO2	ECL1	EC 1	0.01	0.01	0.01	0	0	0.47	0.45	0.21	0.06	0.07	0.09	0.11	0.04	0.01	0.01
		EC 2	0.06	0.05	0.02	0.01	0	0.94	0.93	0.4	0.24	0.13	0.4	0.41	0.13	0.06	0.02
	ECL3	EC 1	0.01	0.01	0	0	0	0.49	0.45	0.22	0.08	0.1	0.1	0.11	0.04	0.01	0.02
		EC 2	0.1	0.11	0.02	0.02	0	0.87	0.91	0.32	0.27	0.18	0.42	0.48	0.12	0.09	0.04
NMO3	ECL1	EC 1	0	0	0.01	0	0.02	0.19	0.22	0.1	0.07	0.29	0.02	0.03	0.02	0.01	0.07
		EC 2	0.01	0	0.01	0	0.01	0.46	0.62	0.08	0.08	0.1	0.1	0.14	0.02	0.01	0.02
	ECL2	EC 1	0.01	0	0.01	0	0.01	0.37	0.27	0.18	0.09	0.23	0.07	0.05	0.03	0.01	0.04
		EC 2	0.03	0.02	0.01	0.02	0.01	0.76	0.84	0.13	0.16	0.1	0.24	0.29	0.04	0.04	0.02
	ECL3	EC 1	0.01	0.01	0.02	0	0	0.37	0.35	0.18	0.09	0.24	0.08	0.08	0.05	0.01	0.04
		EC 2	0.05	0.04	0.03	0.01	0	0.9	0.97	0.22	0.18	0.16	0.36	0.45	0.09	0.04	0.02

Voting preferences

				R1					R2					R3		
			F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
NMO1	ECL1	EC 1	0	0	0	0.01	0	0.04	0.04	0.02	0.07	0.02	0.01	0.01	0	0.03	0.01
		EC 2	0.11	0.14	0.01	0.08	0	0.38	0.6	0.19	0.38	0.16	0.22	0.37	0.07	0.2	0.05
	ECL2	EC 1	0.01	0.02	0	0.03	0	0.03	0.05	0.01	0.08	0.03	0.01	0.02	0	0.04	0.01
		EC 2	0.13	0.15	0.01	0.07	0	0.27	0.36	0.14	0.27	0.14	0.16	0.22	0.05	0.13	0.04
NMO2	ECL1	EC 1	0.01	0.01	0	0	0.01	0.04	0.06	0.03	0.04	0.06	0.01	0.02	0.01	0.02	0.02
		EC 2	0.05	0.03	0.01	0	0	0.19	0.22	0.08	0.09	0.12	0.08	0.09	0.03	0.03	0.03
	ECL3	EC 1	0.01	0.01	0	0	0.01	0.05	0.07	0.02	0.05	0.08	0.01	0.02	0.01	0.02	0.03
		EC 2	0.08	0.04	0.01	0	0	0.22	0.25	0.06	0.14	0.15	0.11	0.11	0.02	0.04	0.04
NMO3	ECL1	EC 1	0.11	0.13	0.03	0.08	0.02	0.05	0.09	0.05	0.07	0.08	0.04	0.06	0.02	0.05	0.03
		EC 2	0.42	0.32	0.04	0.18	0	0.59	0.68	0.23	0.4	0.23	0.45	0.5	0.09	0.28	0.06
	ECL2	EC 1	0.07	0.09	0.01	0.07	0.01	0.08	0.12	0.08	0.08	0.11	0.04	0.07	0.02	0.05	0.03
		EC 2	0.28	0.28	0.02	0.18	0	0.47	0.54	0.16	0.37	0.19	0.31	0.38	0.05	0.25	0.04
	ECL3	EC 1	0.08	0.1	0.02	0.04	0.01	0.11	0.14	0.08	0.14	0.19	0.06	0.09	0.03	0.07	0.06
		EC 2	0.28	0.27	0.04	0.09	0	0.54	0.66	0.22	0.44	0.31	0.37	0.47	0.09	0.25	0.07

Donation preferences

				R1					R2					R3		
			F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
NMO1	ECL1	EC 1	0	0	0	0	0	0.03	0.04	0.01	0.04	0.01	0.01	0.01	0	0.01	0
		EC 2	0.07	0.1	0.01	0.06	0	0.19	0.33	0.09	0.35	0.14	0.1	0.18	0.03	0.17	0.04
	ECL2	EC 1	0.01	0.01	0	0.01	0	0.03	0.04	0.01	0.05	0.01	0.01	0.01	0	0.02	0
		EC 2	0.08	0.11	0.01	0.06	0	0.12	0.16	0.06	0.25	0.11	0.07	0.09	0.02	0.11	0.03
NMO2	ECL1	EC 1	0.01	0.01	0	0	0.01	0.03	0.05	0.02	0.03	0.05	0.01	0.02	0	0.01	0.02
		EC 2	0.03	0.02	0	0	0	0.07	0.09	0.03	0.05	0.06	0.03	0.04	0.01	0.02	0.01
	ECL3	EC 1	0.01	0.01	0	0	0.01	0.06	0.09	0.02	0.06	0.09	0.02	0.03	0.01	0.02	0.03
		EC 2	0.08	0.05	0.02	0	0	0.16	0.19	0.04	0.1	0.1	0.08	0.08	0.02	0.03	0.03
NMO3	ECL1	EC 1	0.08	0.17	0.02	0.06	0.01	0.03	0.08	0.02	0.04	0.04	0.02	0.07	0.01	0.03	0.01
		EC 2	0.42	0.4	0.04	0.24	0.01	0.37	0.41	0.11	0.32	0.16	0.31	0.35	0.05	0.26	0.05
	ECL2	EC 1	0.05	0.12	0.01	0.05	0.01	0.04	0.09	0.03	0.05	0.05	0.02	0.06	0.01	0.03	0.01
		EC 2	0.28	0.37	0.02	0.24	0.01	0.27	0.3	0.07	0.3	0.12	0.2	0.25	0.02	0.22	0.03
	ECL3	EC 1	0.08	0.18	0.02	0.03	0.01	0.08	0.18	0.06	0.09	0.12	0.04	0.13	0.02	0.04	0.04
		EC 2	0.37	0.44	0.07	0.14	0.01	0.48	0.56	0.19	0.41	0.27	0.37	0.47	0.09	0.26	0.08

[Steegen, Tuerlinckz, Gelman, Vanpaemel. Increasing Transparency Through a Multiverse Analysis. Perspectives on Psychological Science, 2016]

Explorable Multiverse Analysis Reports

[Dragicevic, Jansen, Sarma, **Kay**, and Chevalier. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. CHI 2019: https://explorablemultiverse.github.io/. **Best Paper**]

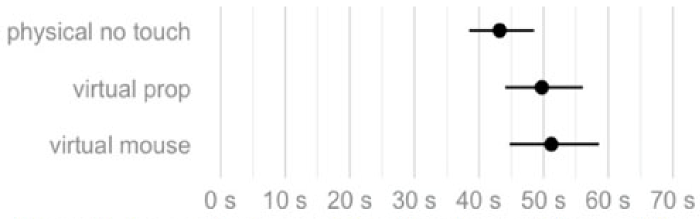
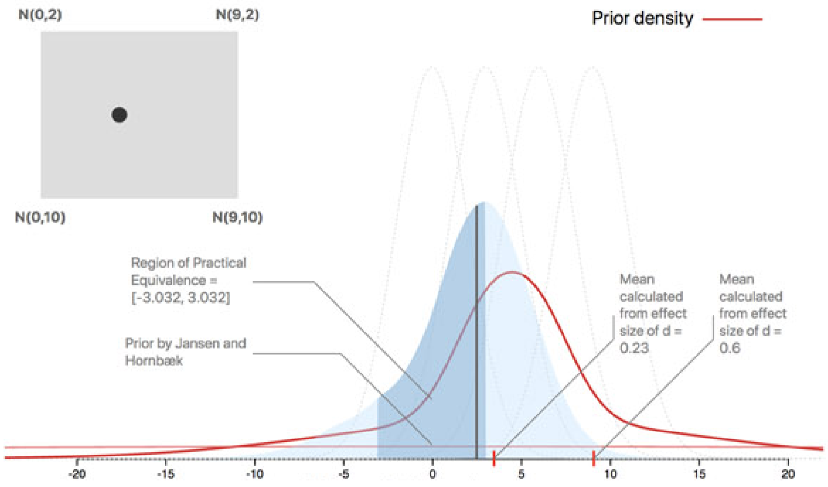
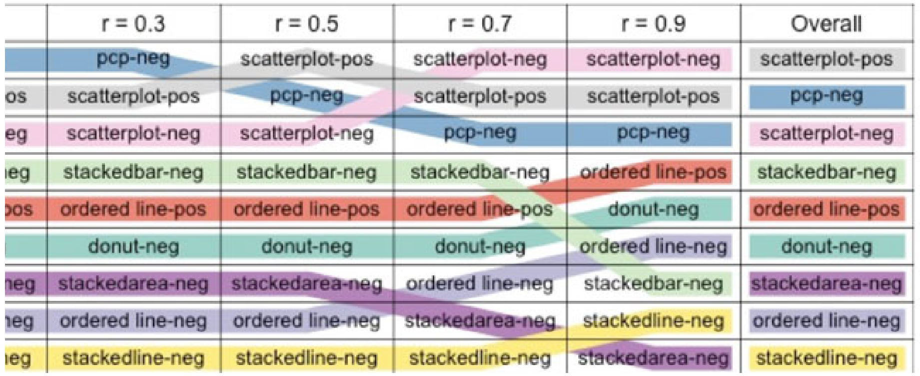
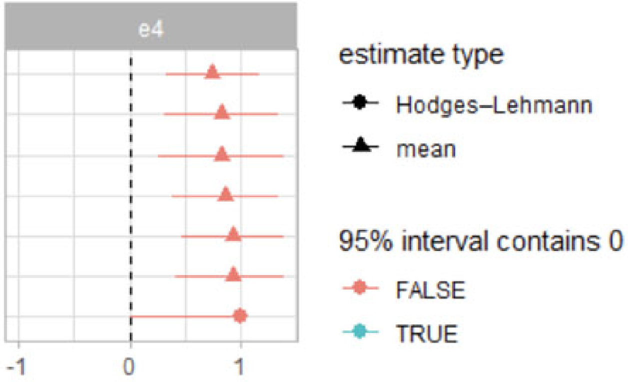


Figure 3. Average task completion time (geometric mean) for each condition. Error bars are 95% t-based CIs.

We focus our analysis on task completion times, reported in Figures 3 and 4. Dots indicate sample means, while error bars are 95% confidence intervals computed on log-transformed data [6] using the t-distribution method. Strictly speaking, all we can assert about each interval is







Explorable Multiverse Analysis Reports

[Dragicevic, Jansen, Sarma, **Kay**, and Chevalier. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. CHI 2019: https://explorablemultiverse.github.io/. **Best Paper**]

We need better ways to acknowledge specification uncertainty and have a conversation about it through the literature

Explorable Multiverse Analysis Reports

[Dragicevic, Jansen, Sarma, **Kay**, and Chevalier. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. CHI 2019: https://explorablemultiverse.github.io/. **Best Paper**]

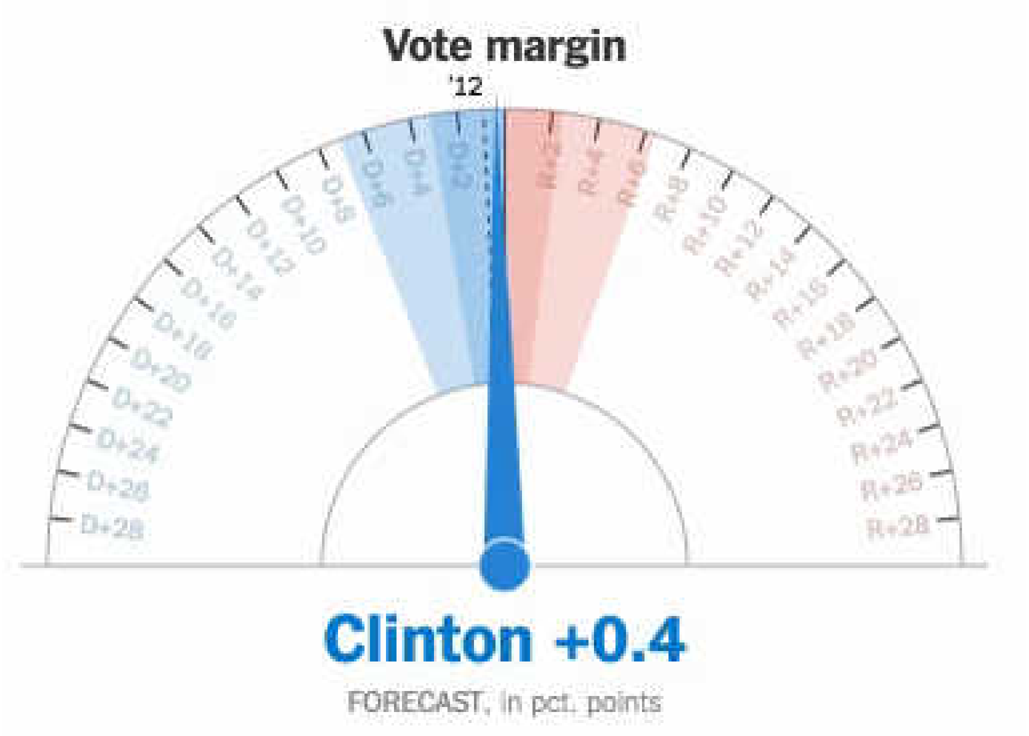
We need better ways to acknowledge specification uncertainty and have a conversation about it through the literature

Currently building an R package [Abhraneel Sarma] and a visualization design space [Brian Hall]

Going back to election data...

New York Times Election Needle

[https://www.nytimes.com/interactive/2016/11/08/us/elections/trump-clinton-election-night-live.html]



The Fake Twitchy Hell Dials of the New York *Times'* Forecast Only Made Last Night Worse

By **Jake Swearingen**

Photo: rhyselsmore/Twitter

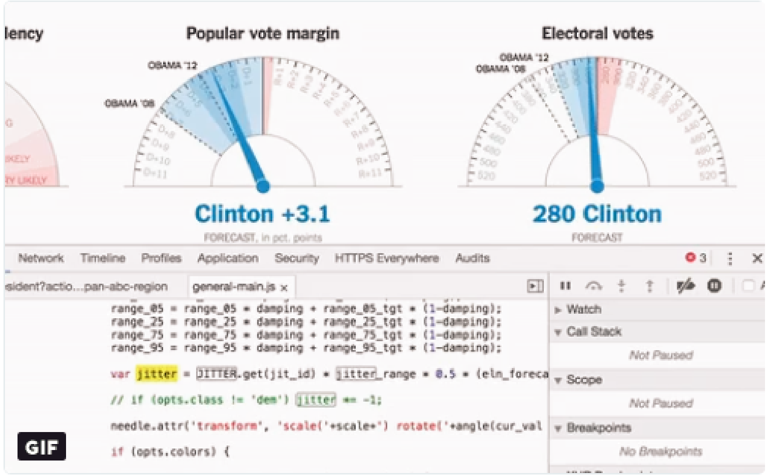
Around 9:30 last night, this tweet popped up on my timeline:

> stop tweeting the fucking hell dial
>
> — **erictoral vote (@ericlimer)** November 9, 2016

Alp Toker @atoker

Follow

Looking for trends in @nytimes's presidential forecast needle? Don't look too hard - the bounce is random jitter from your PC, not live data



Richard Porczak @tsiro

Follow

straight up: the NYT needle jitter is irresponsible design at best and unethical design at worst and you should stop looking at it

9:58 PM - 8 Nov 2016

509 Retweets **882** Likes

17 509 882

But shouldn't anxiety be proportional to uncertainty?

Uncertainty visualization as a moral imperative

We should…

present well-calibrated uncertainty
that cannot be ignored
in ways people can actually understand

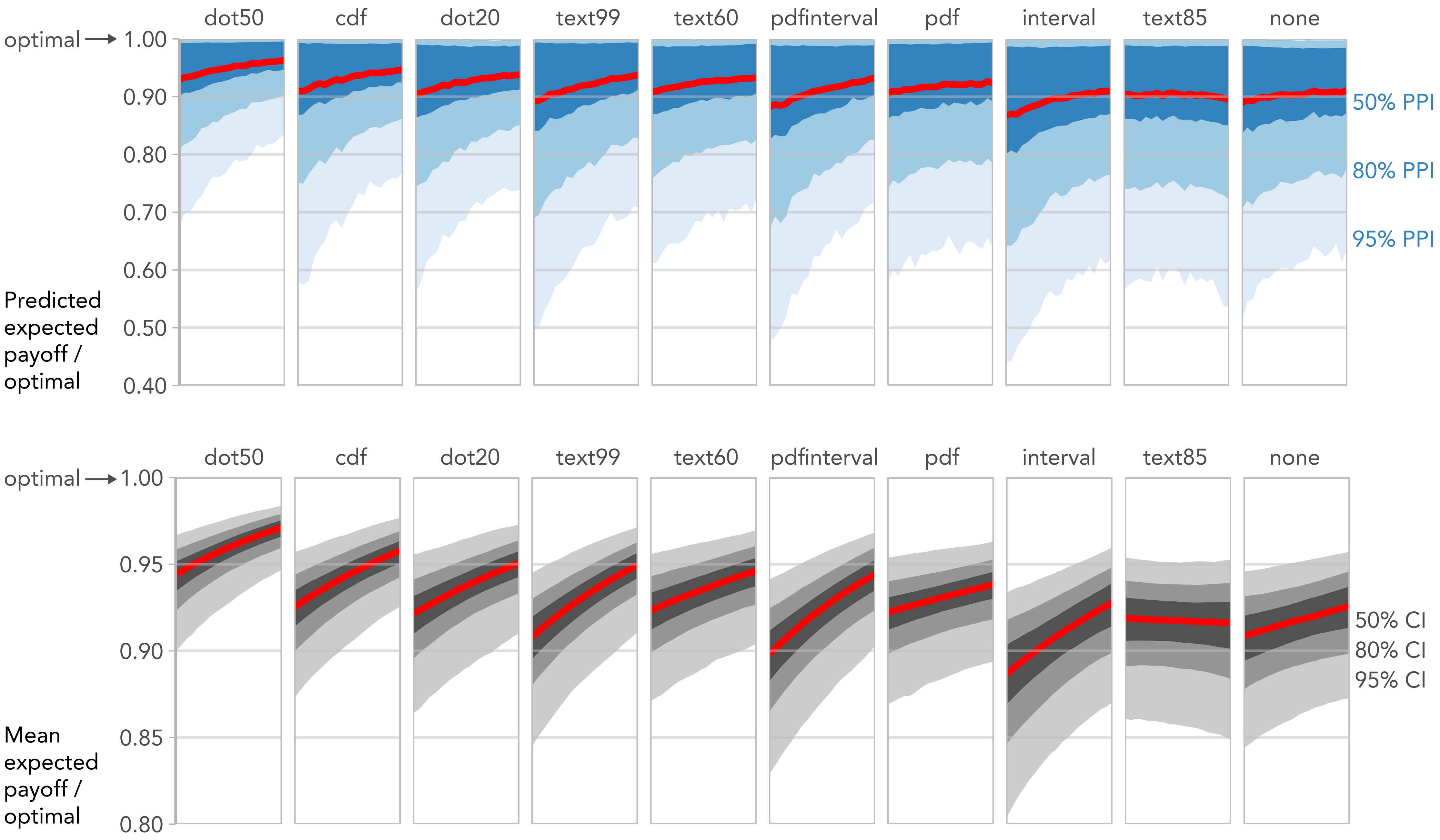
Thanks!

Students: Xiaoying Pu, Brian Hall, Abhraneel Sarma, Puhe Liange, Tara Kola, Michael Fernandes, Logan Walls

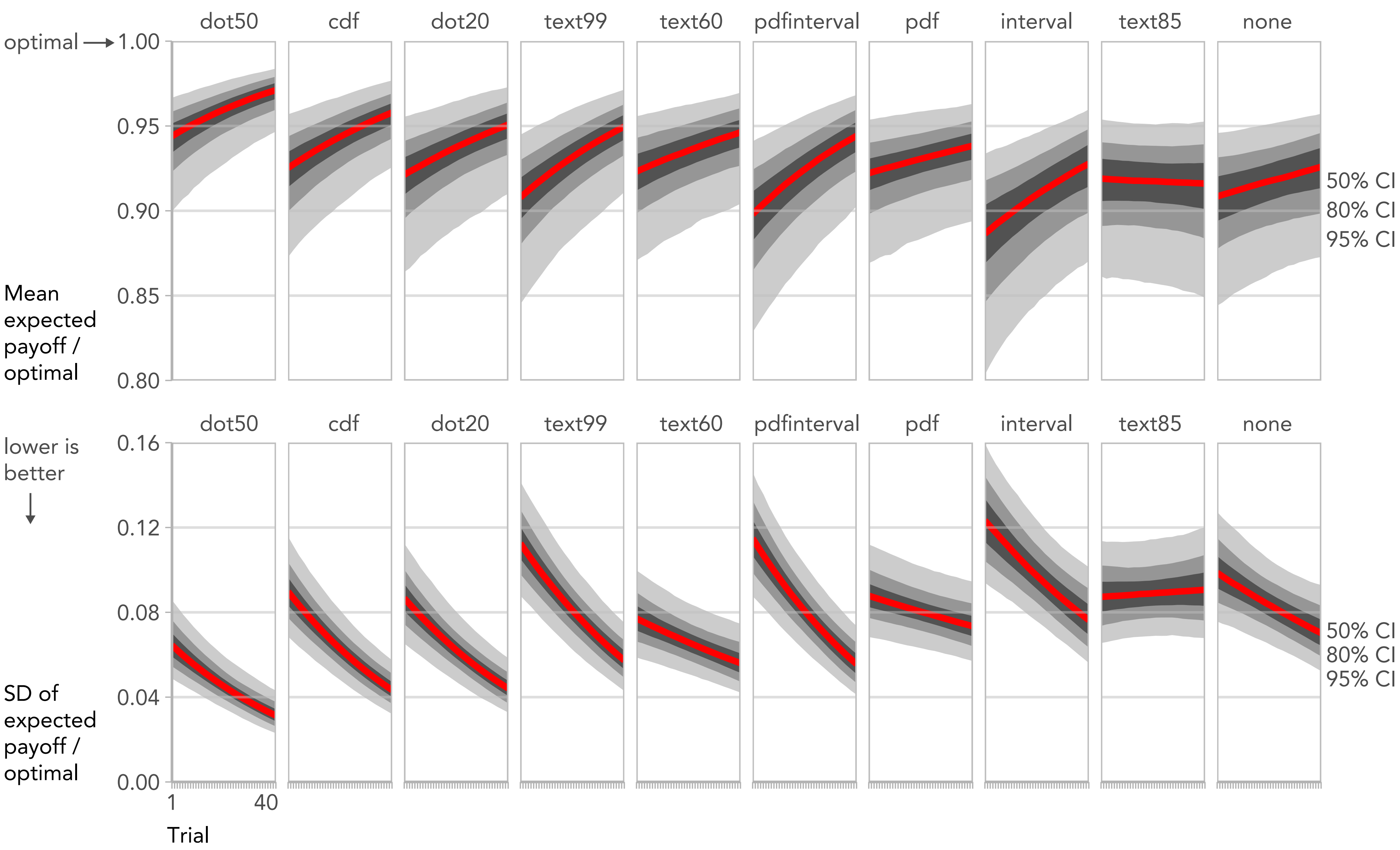
Collaborators: Jessica Hullman, Sean Munson, Julie Kientz, Shwetak Patel, Alex Kale, Gregory Nelson, Eric Hekler, Jeff Heer, Steve Haroz, Pierre Dragicevic, Yvonne Jansen, Fanny Chevalier

Matthew Kay
mjskay@umich.edu
University of Michigan School of Information

http://mjskay.com/
http://mucollective.co/



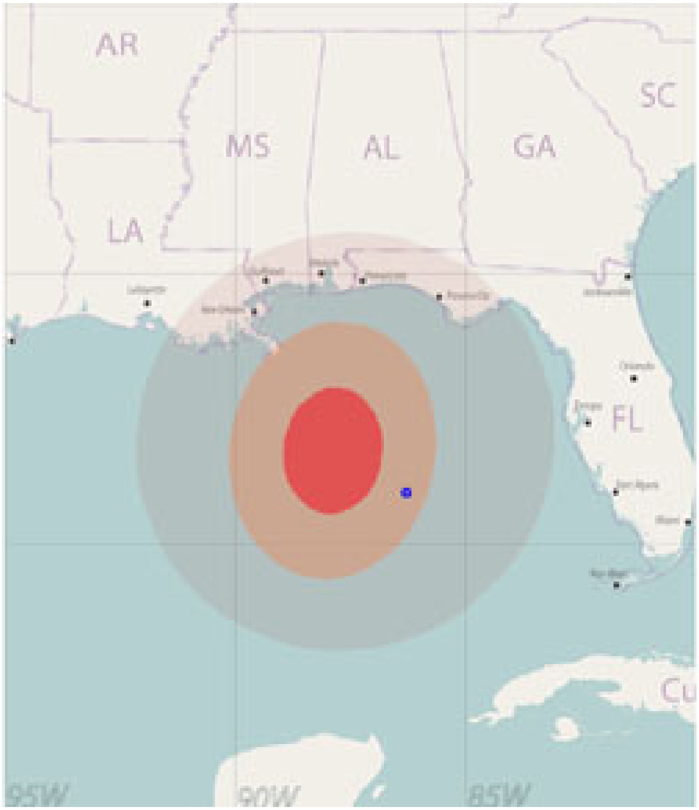
dot50
cdf
dot20
text99
text60
pdfinterval
pdf
interval
text85
none
optimal
1.00
0.90
0.80
0.70
0.60
0.50
0.40
Predicted expected payoff / optimal
50% PPI
80% PPI
95% PPI
dot50
cdf
dot20
text99
text60
pdfinterval
pdf
interval
text85
none
optimal
1.00
0.95
0.90
0.85
0.80
Mean expected payoff / optimal
50% CI
80% CI
95% CI



optimal → 1.00
0.95
0.90
0.85
0.80
Mean expected payoff / optimal
dot50
cdf
dot20
text99
text60
pdfinterval
pdf
interval
text85
none
50% CI
80% CI
95% CI
lower is better
0.16
0.12
0.08
0.04
0.00
SD of expected payoff / optimal
1
40
Trial

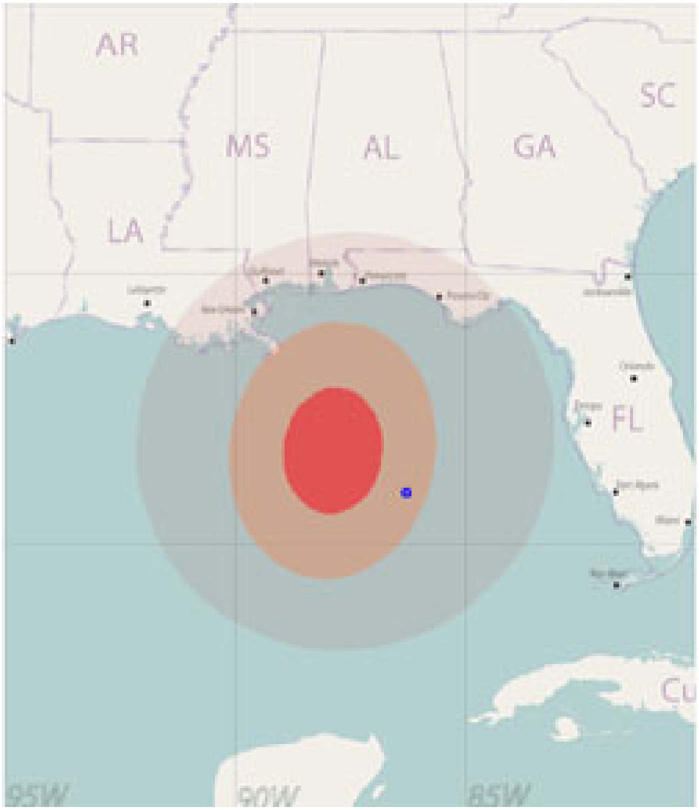
Hurricane location at a time slice...

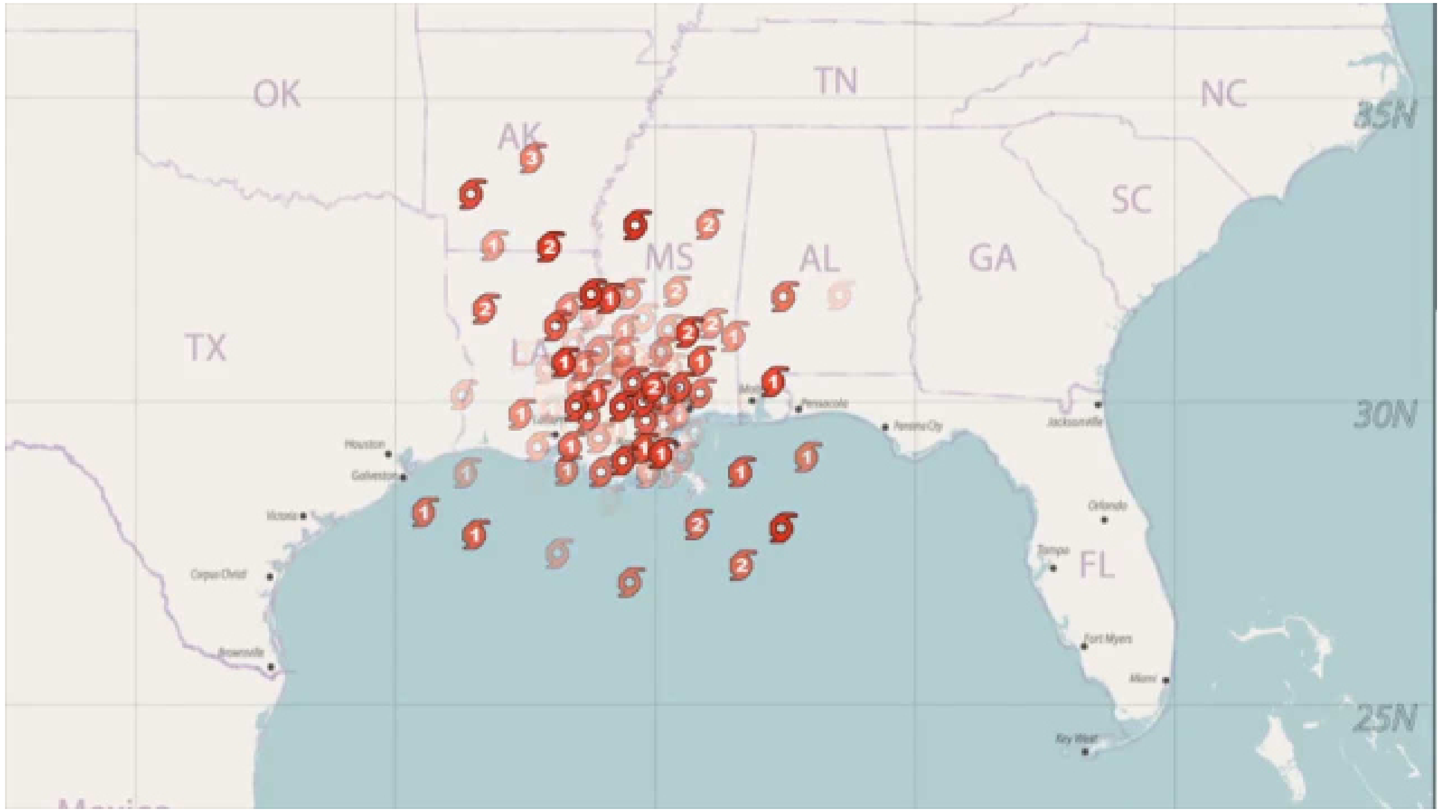
[Liu, Boone, Ruginski, Padilla, Hegarty, Creem-Regehr, … House. Uncertainty Visualization by Representative Sampling from Prediction Ensembles. IEEE Transactions on Visualization and Computer Graphics, PP(99), 2016]



Hurricane location at a time slice...

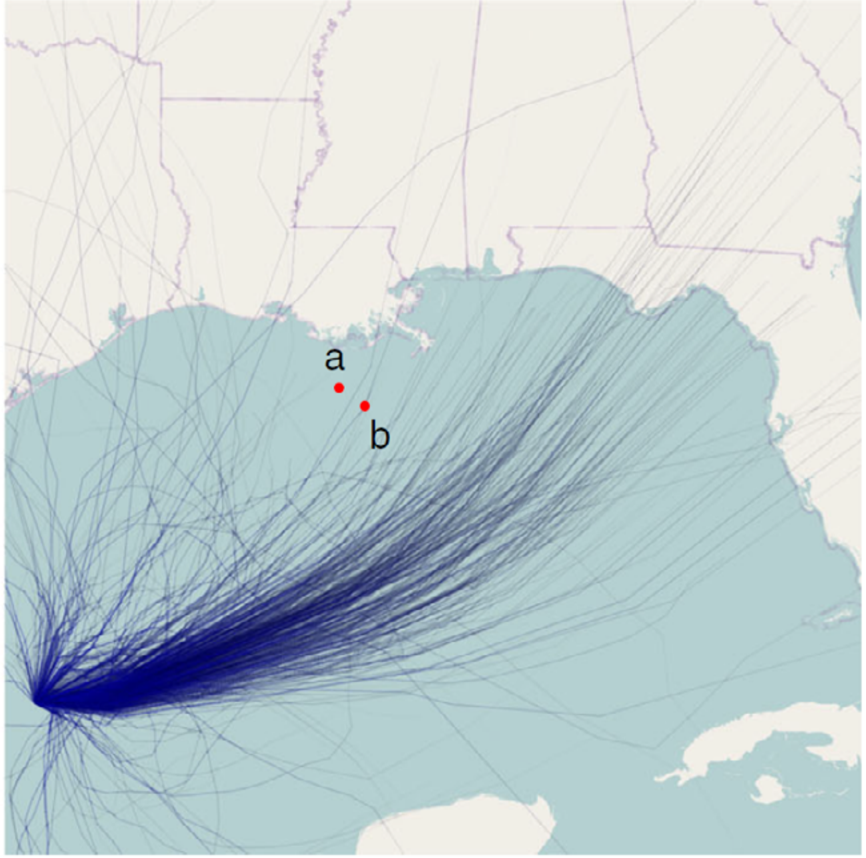
[Liu, Boone, Ruginski, Padilla, Hegarty, Creem-Regehr, … House. Uncertainty Visualization by Representative Sampling from Prediction Ensembles. IEEE Transactions on Visualization and Computer Graphics, PP(99), 2016]





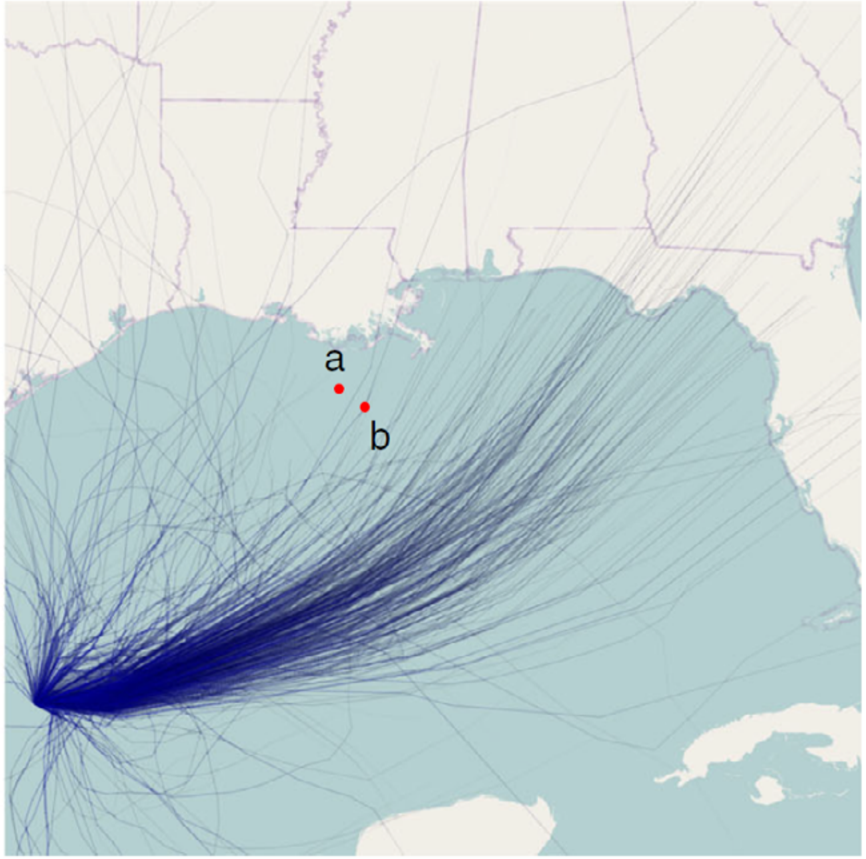
(but problems with ensembles...)

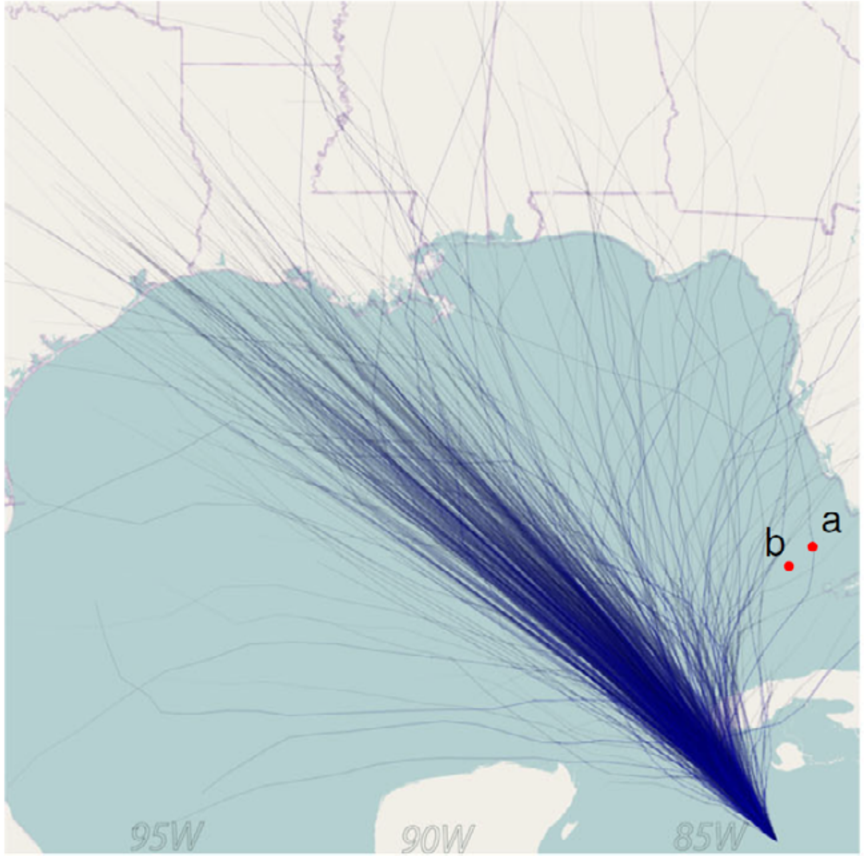
[Padilla, Ruginski, Creem-Regehr. Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. Cognitive Research: Principles and Implications, 2(1), 40, 2017]



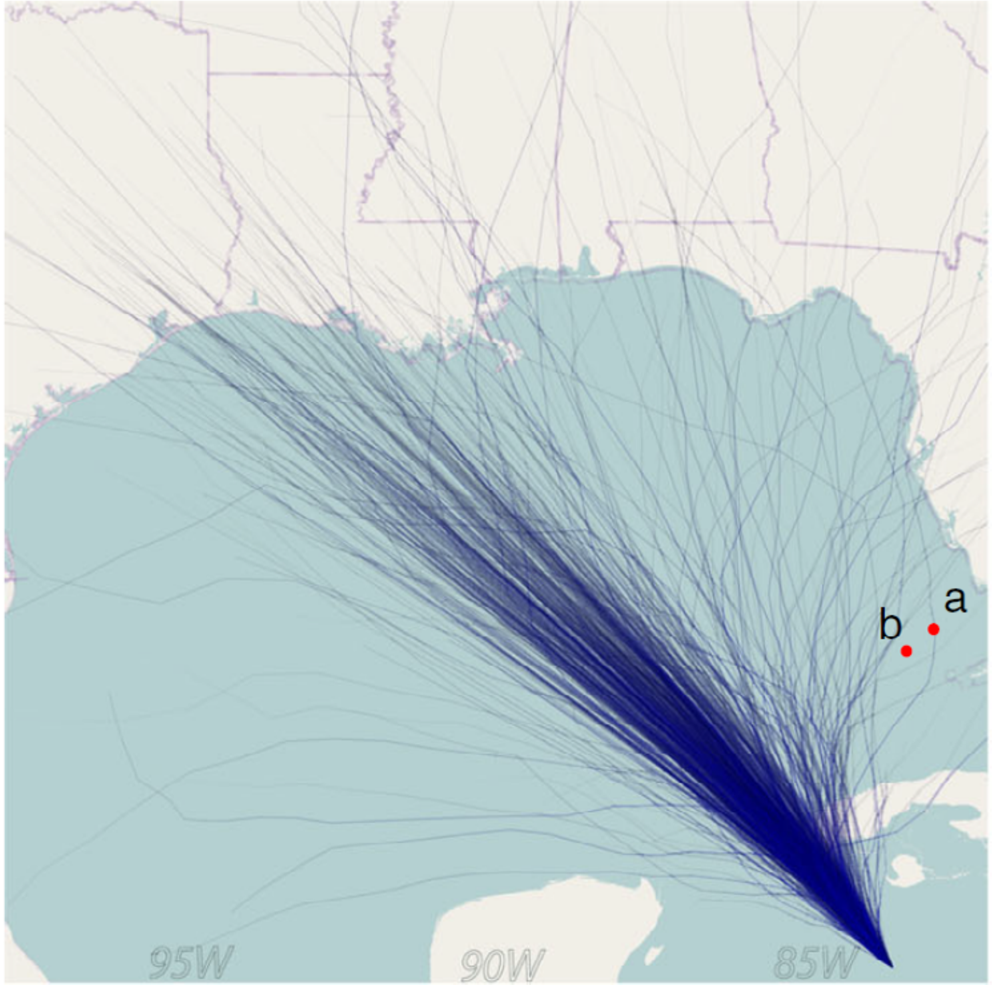
(but problems with ensembles…)

[Padilla, Ruginski, Creem-Regehr. Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. Cognitive Research: Principles and Implications, 2(1), 40, 2017]

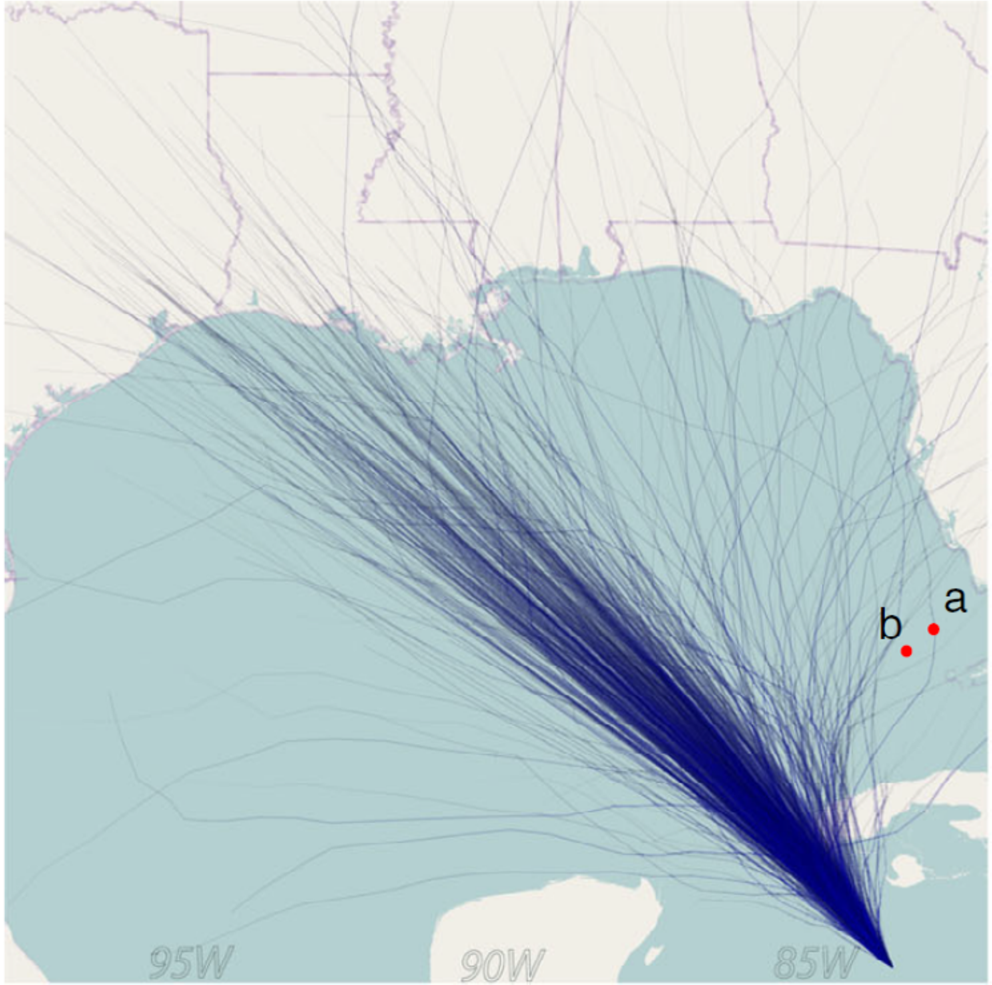


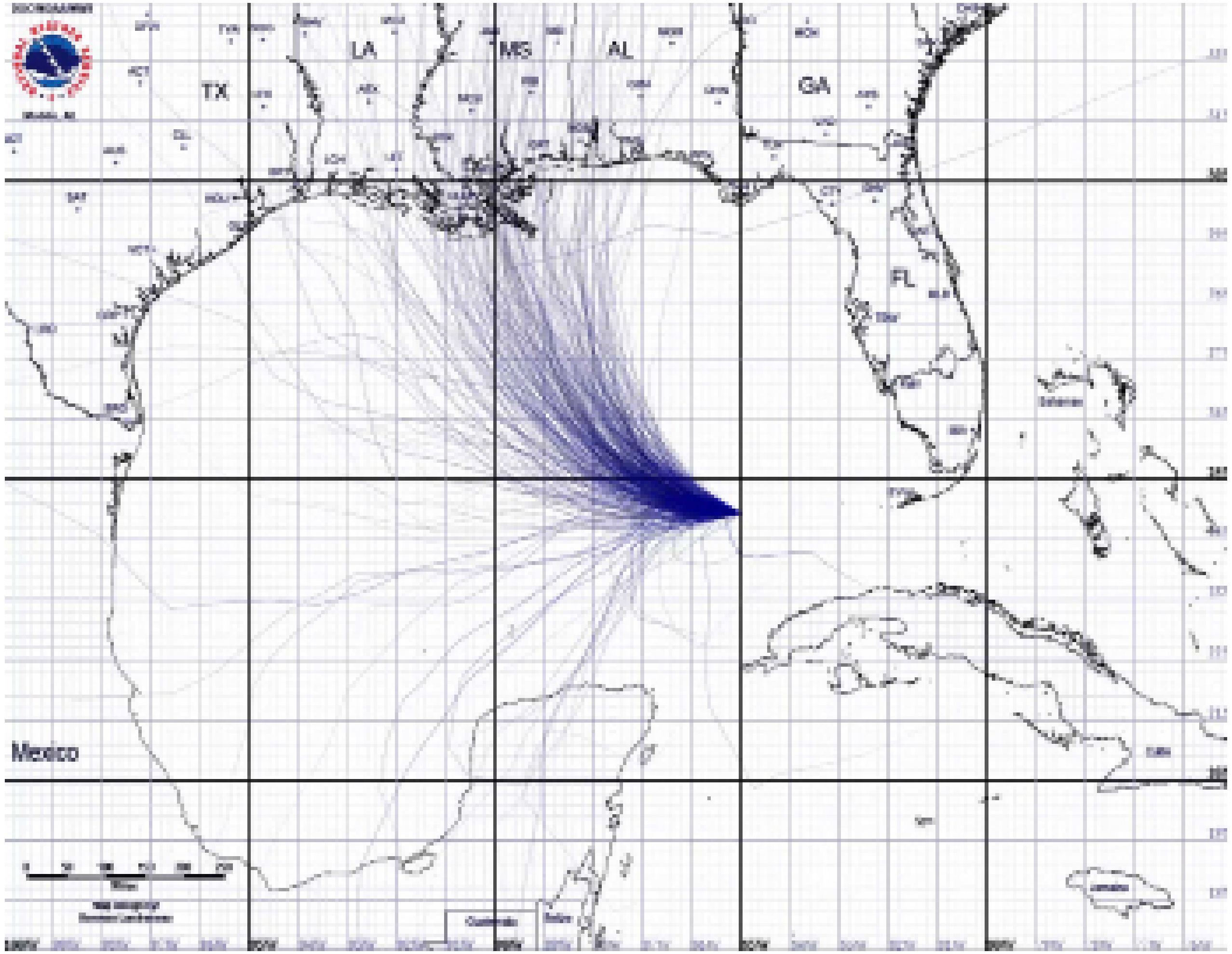


HOPs might aid deterministic construal errors



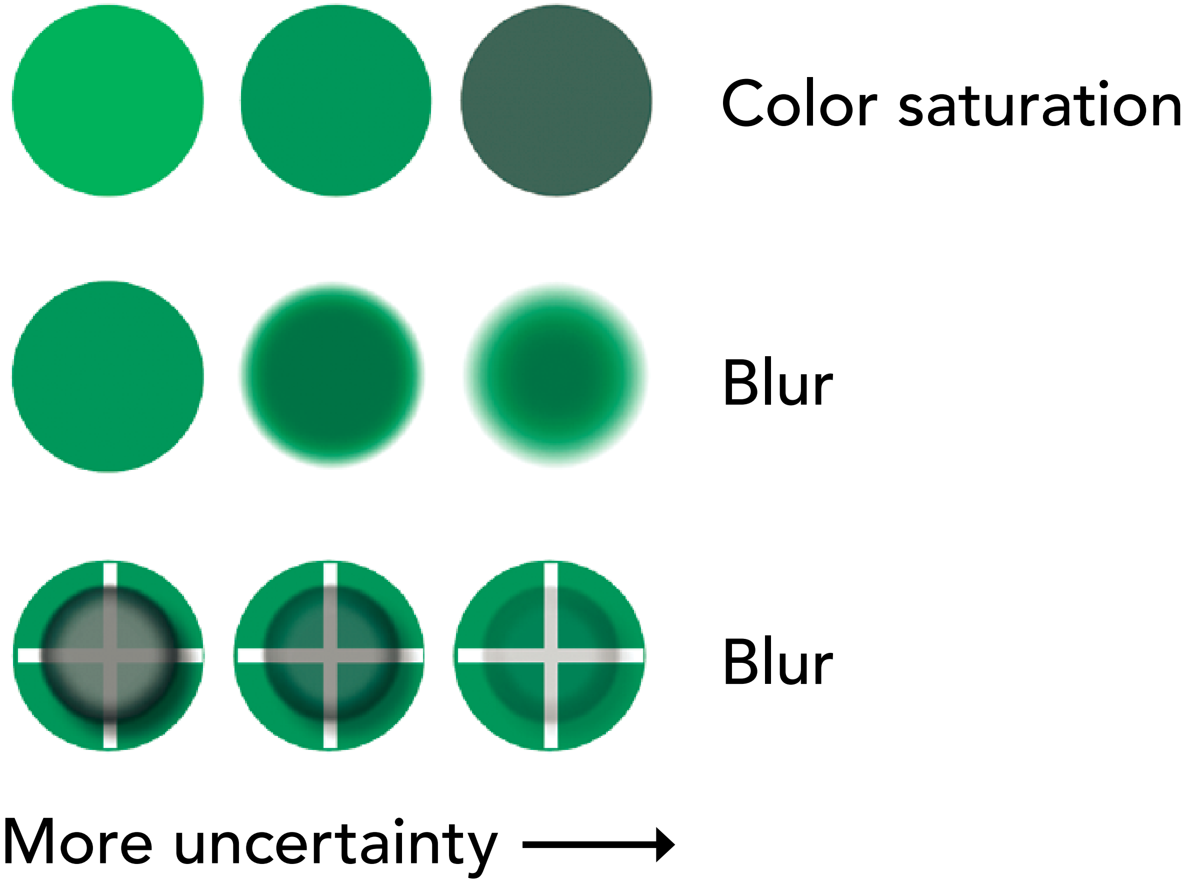
HOPs might aid deterministic construal errors





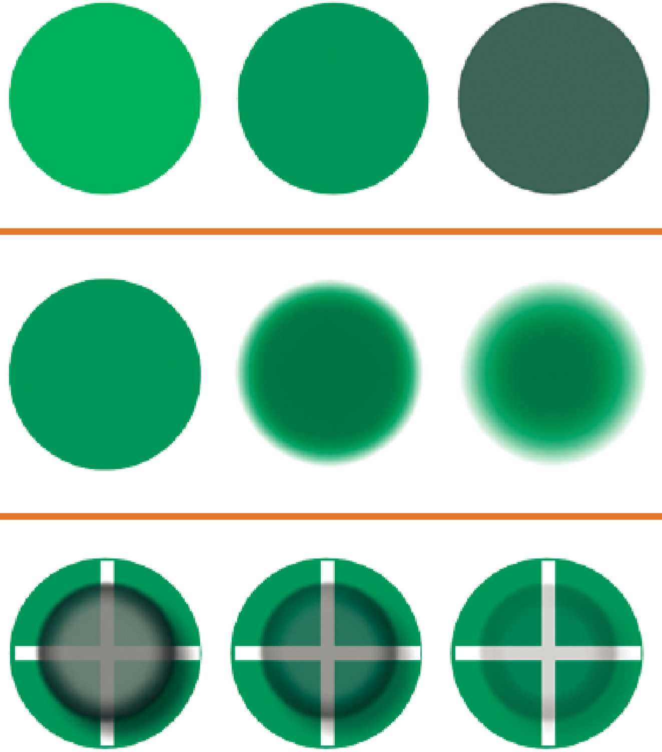
Glyph-based uncertainty

[MacEachren, Robinson, Hopper, Gardner, Murray, Gahegan, Hetzler. Visualizing geospatial information uncertainty: What we know and what we need to know. Cartography and Geographic Information Science, 32(3), 139-160, 2005]



Glyph-based uncertainty

[MacEachren, Robinson, Hopper, Gardner, Murray, Gahegan, Hetzler. Visualizing geospatial information uncertainty: What we know and what we need to know. Cartography and Geographic Information Science, 32(3), 139-160, 2005]



Color saturation

Blur

More intuitive?

But how accurate?

Blur

More uncertainty →

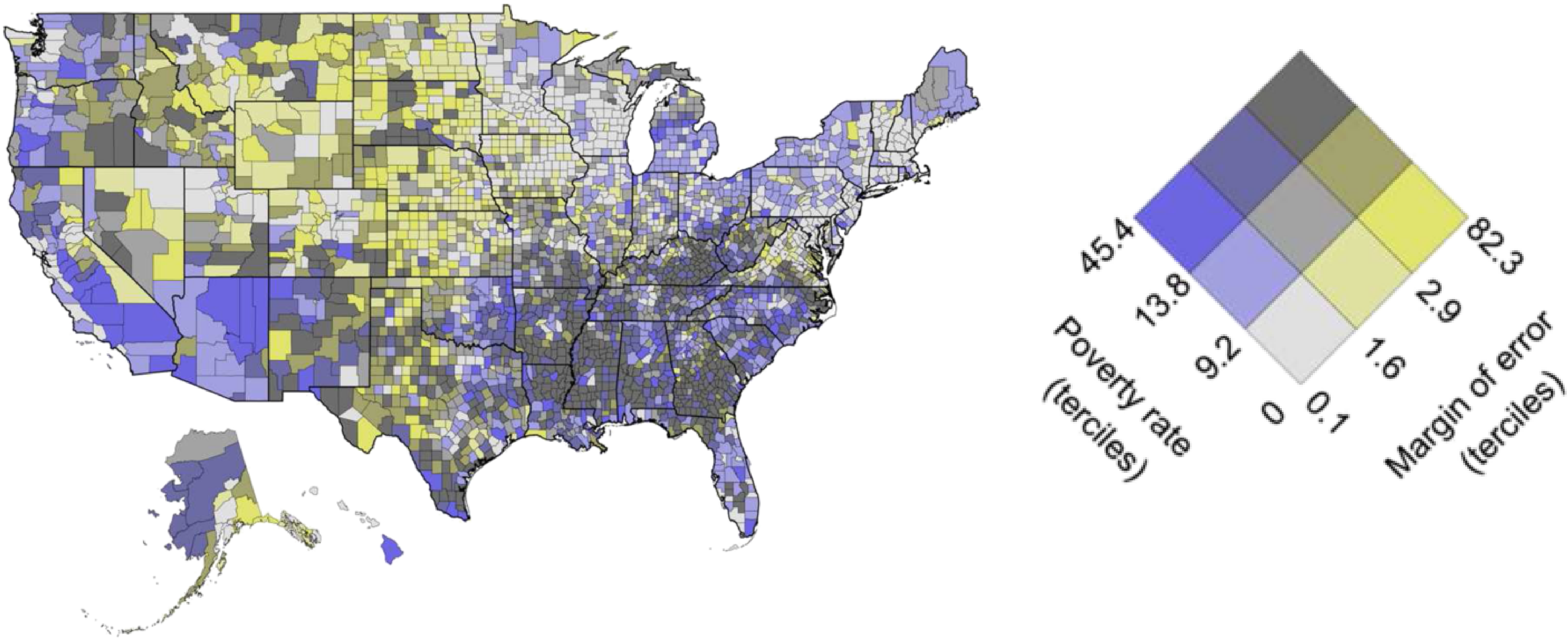
I'm not a GIS person, so let's take a little detour

One example of prototyping
(because it's a fun one)

Cartographic uncertainty

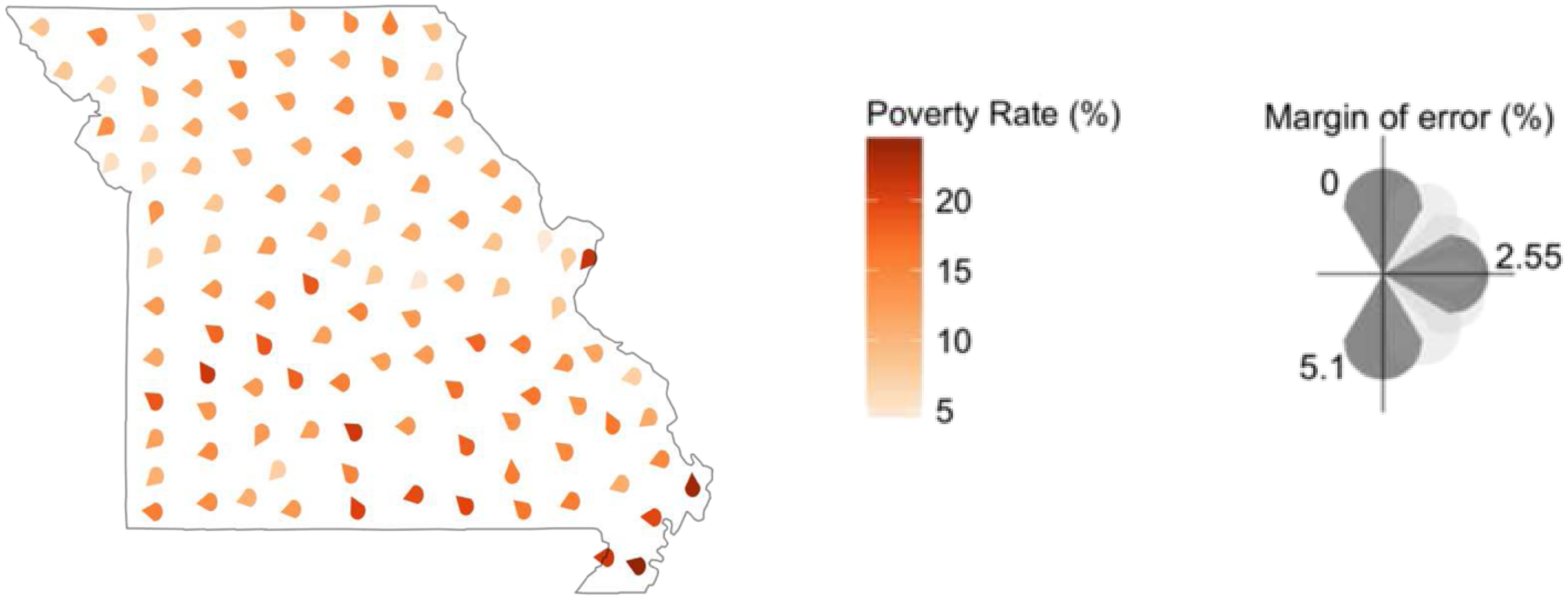
Just map to another visual channel, right?

[Lucchesi & Wikle. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. Stat, 292–302, 2017]



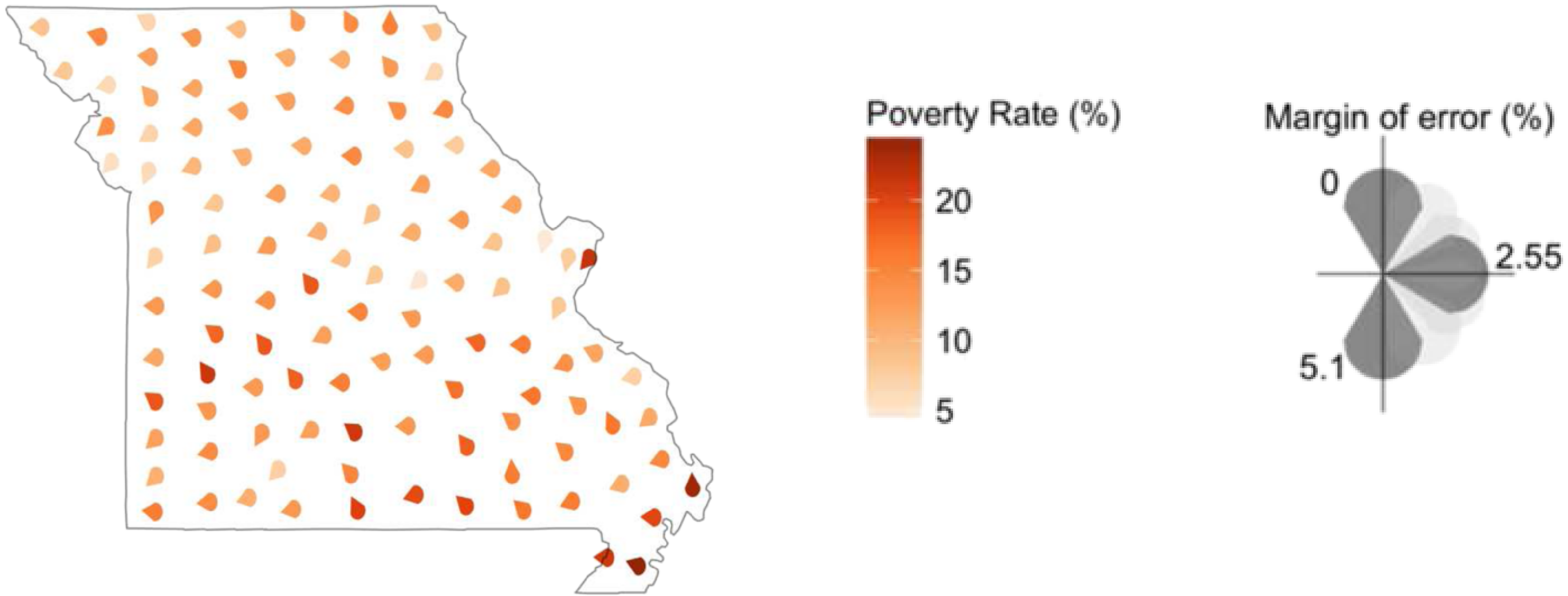
Just map to another visual channel, right?

[Lucchesi & Wikle. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. Stat, 292–302, 2017]



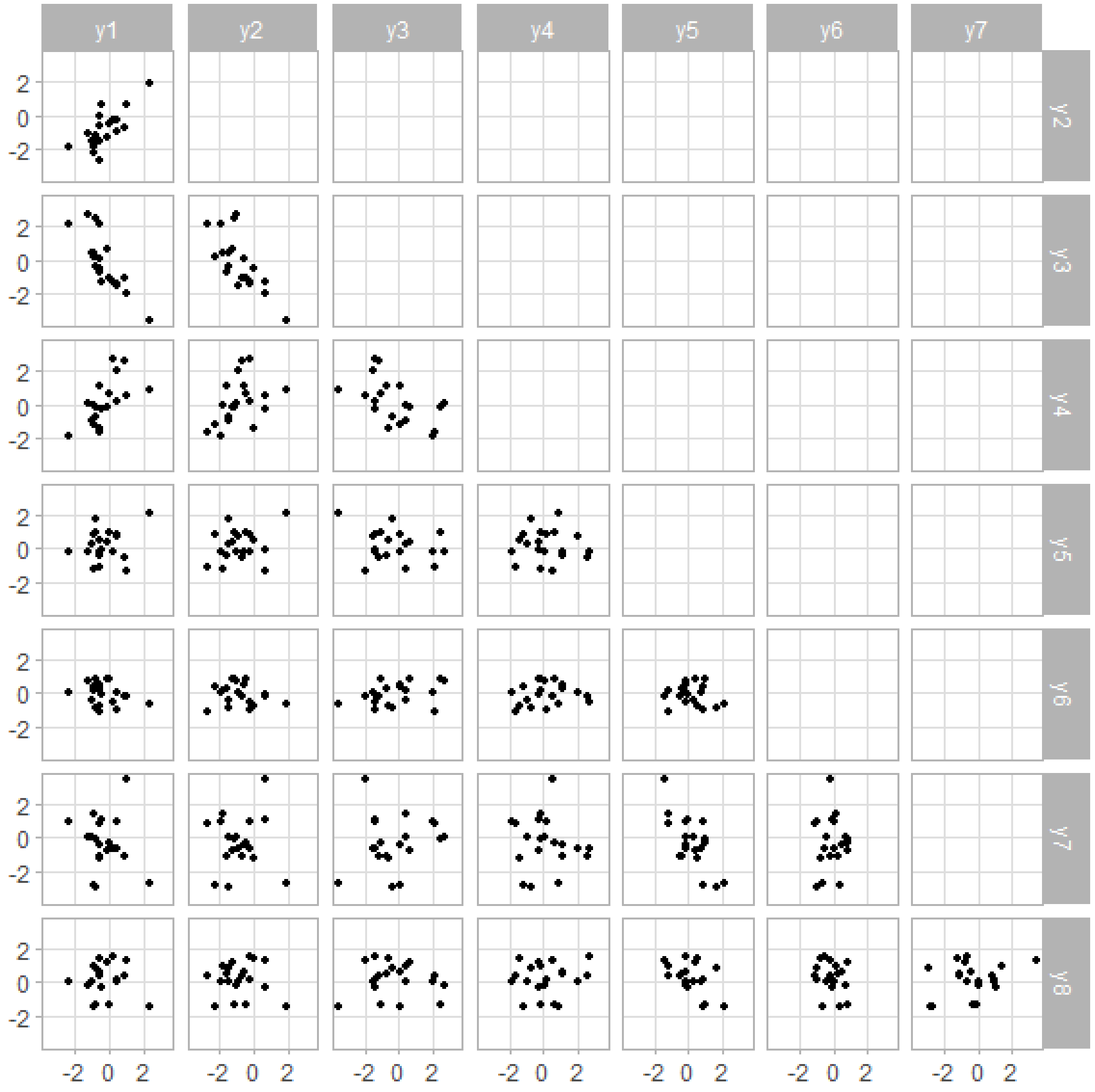
Just map to another visual channel, right?

[Lucchesi & Wikle. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. Stat, 292–302, 2017]

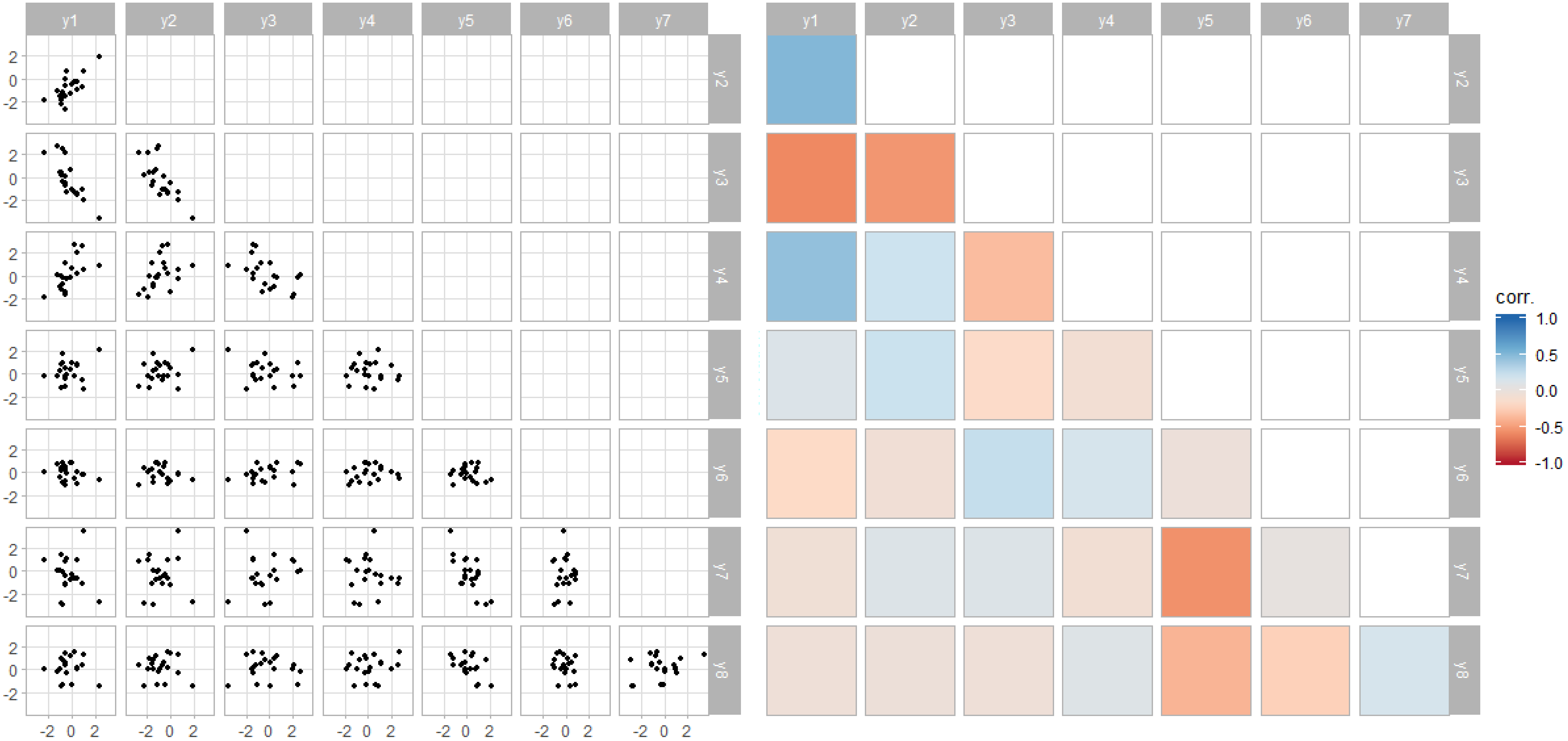


Very abstract…

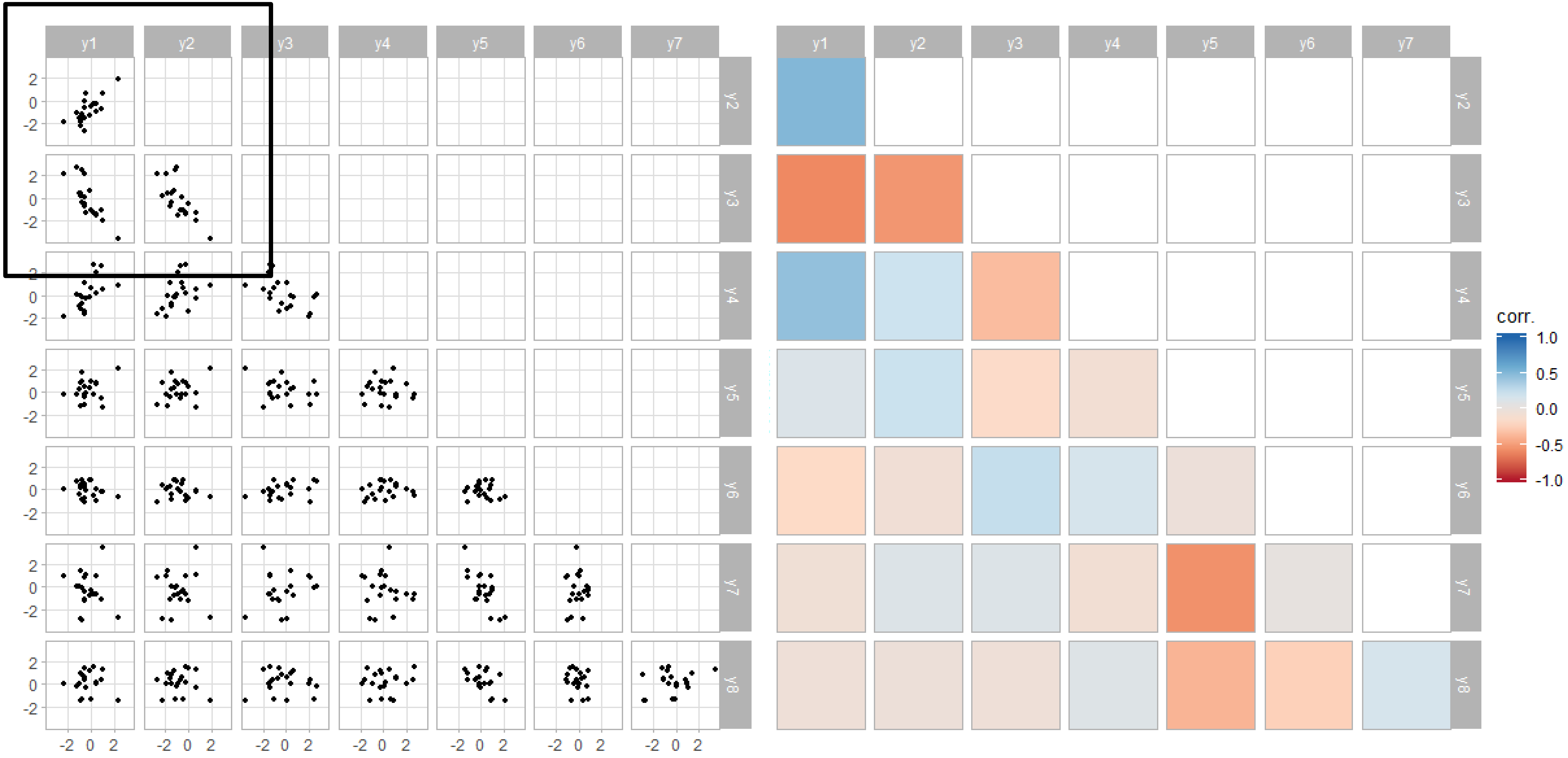
I'm not a map vis person...



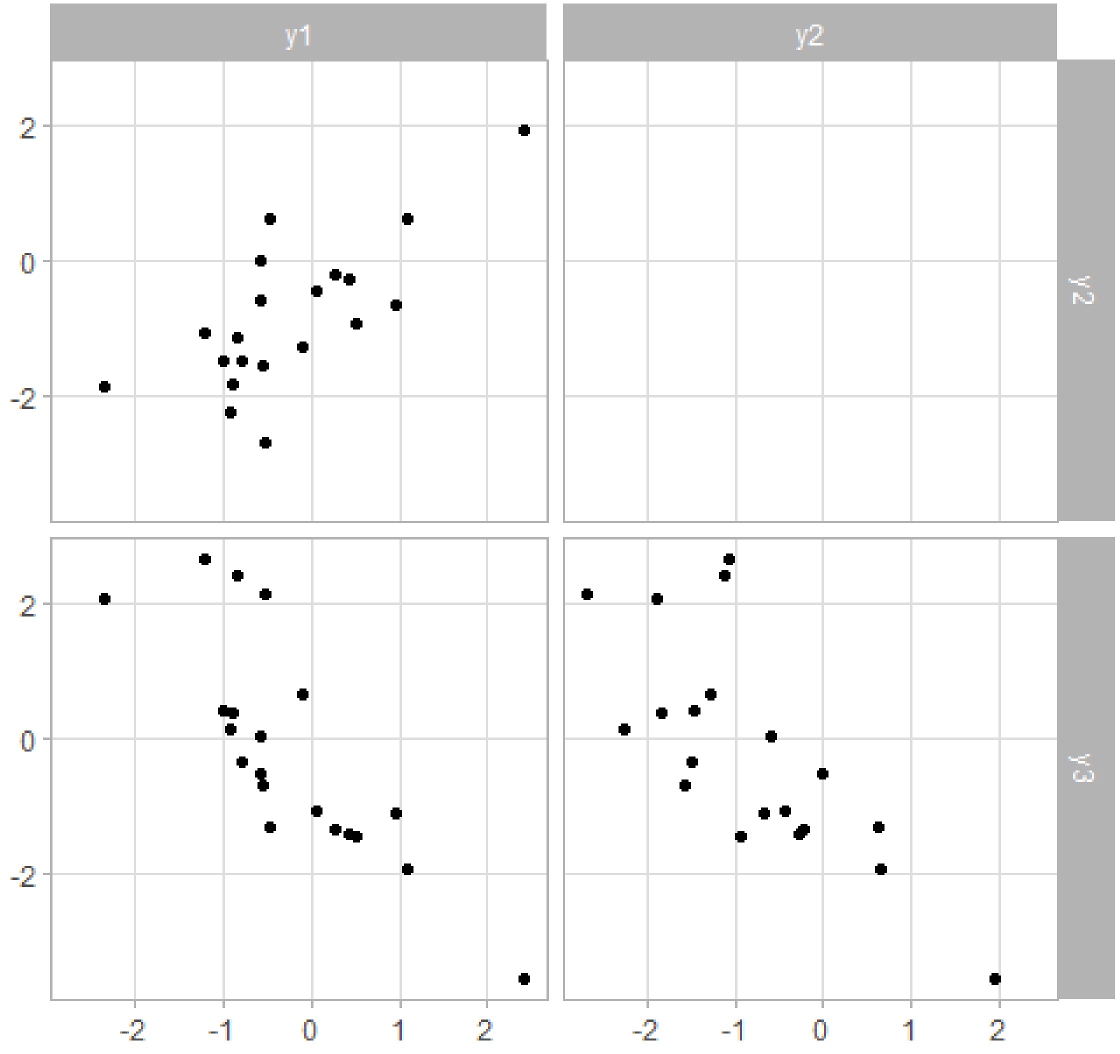
y1
y2
y3
y4
y5
y6
y7
y2
y3
y4
y5
y6
y7
y8
2
0
-2
-2 0 2

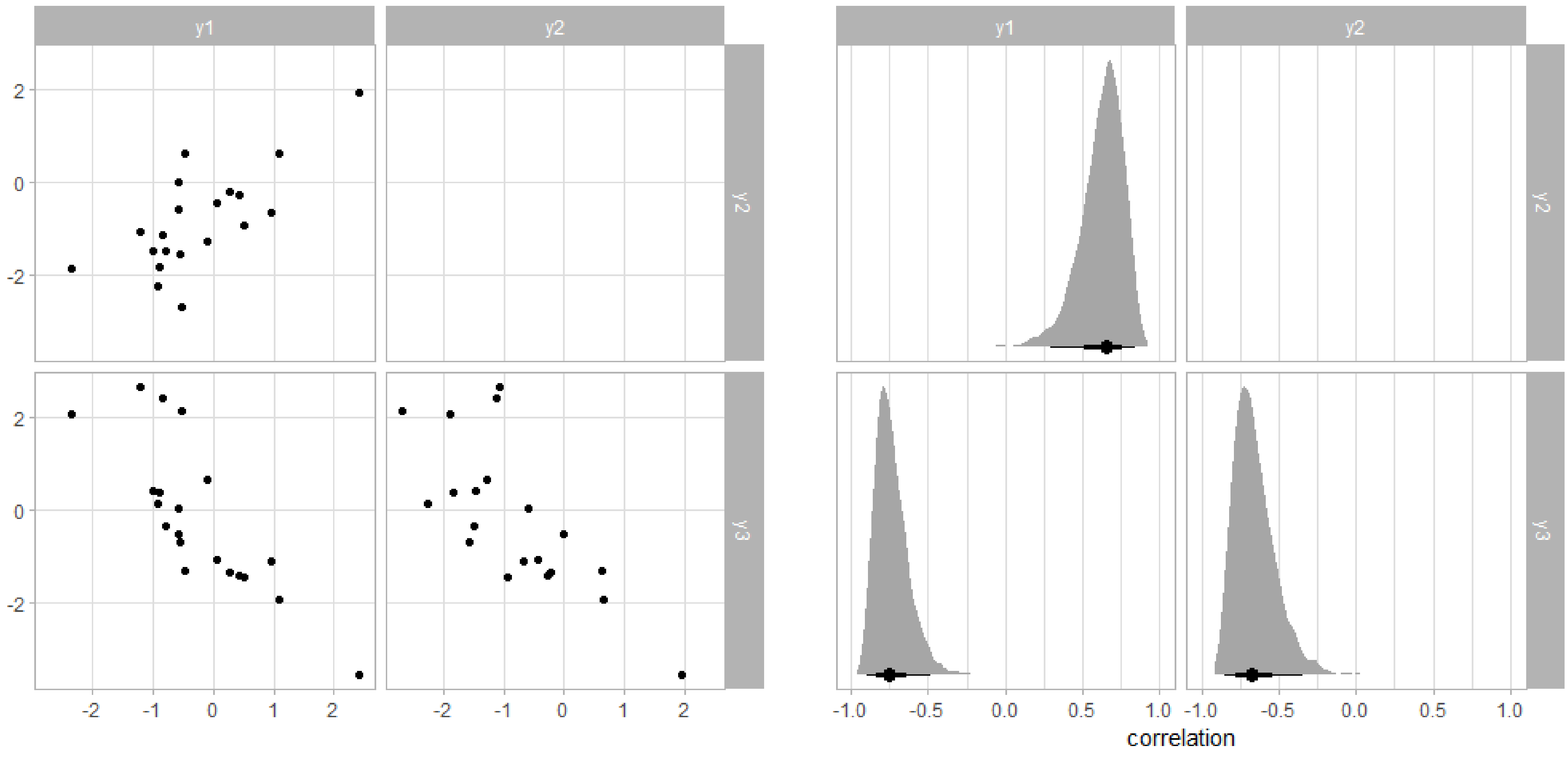


y1
y2
y3
y4
y5
y6
y7
y8
2
0
-2
-2 0 2
corr.
1.0
0.5
0.0
-0.5
-1.0

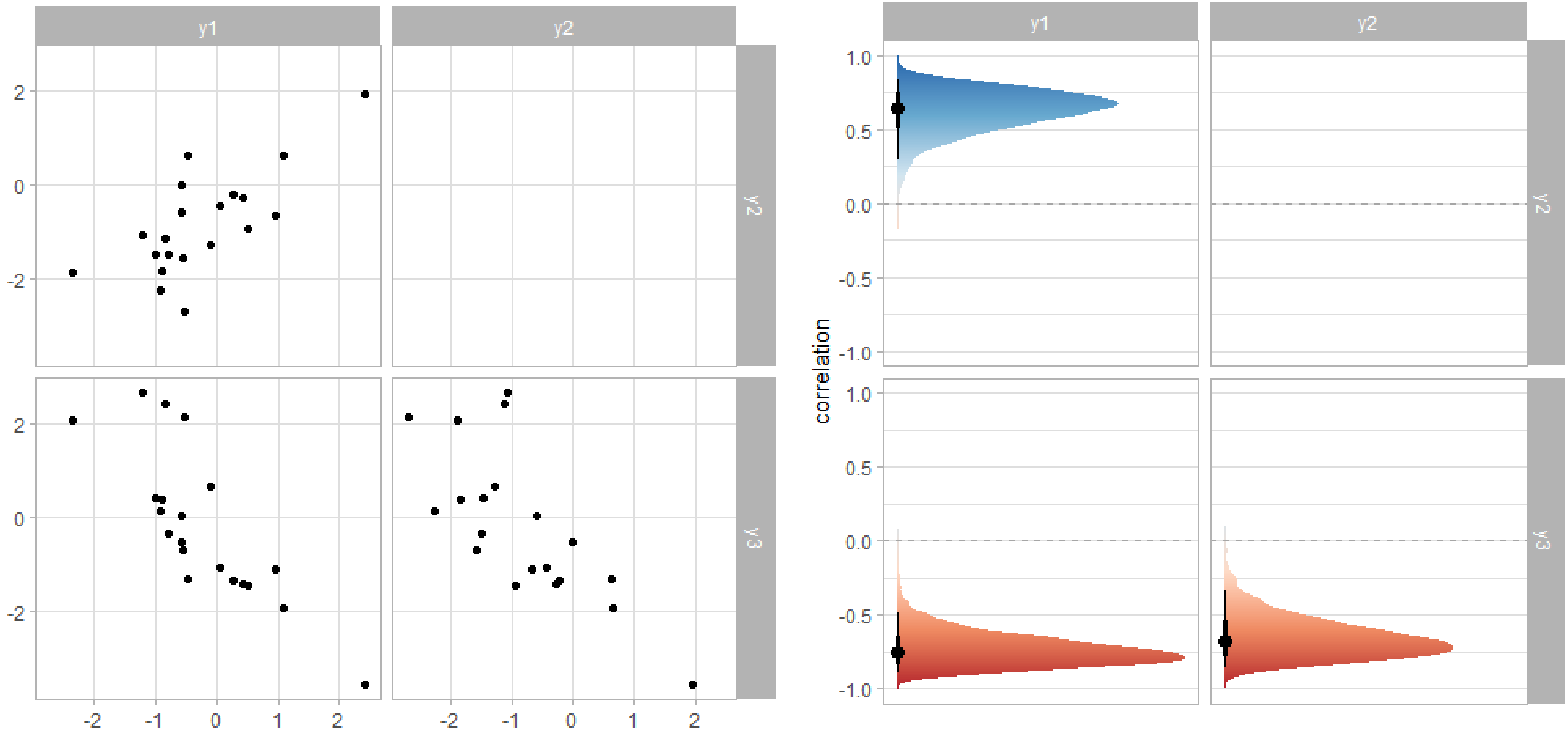


y1
y2
y3
y4
y5
y6
y7
y8
2
0
-2
corr.
1.0
0.5
0.0
-0.5
-1.0

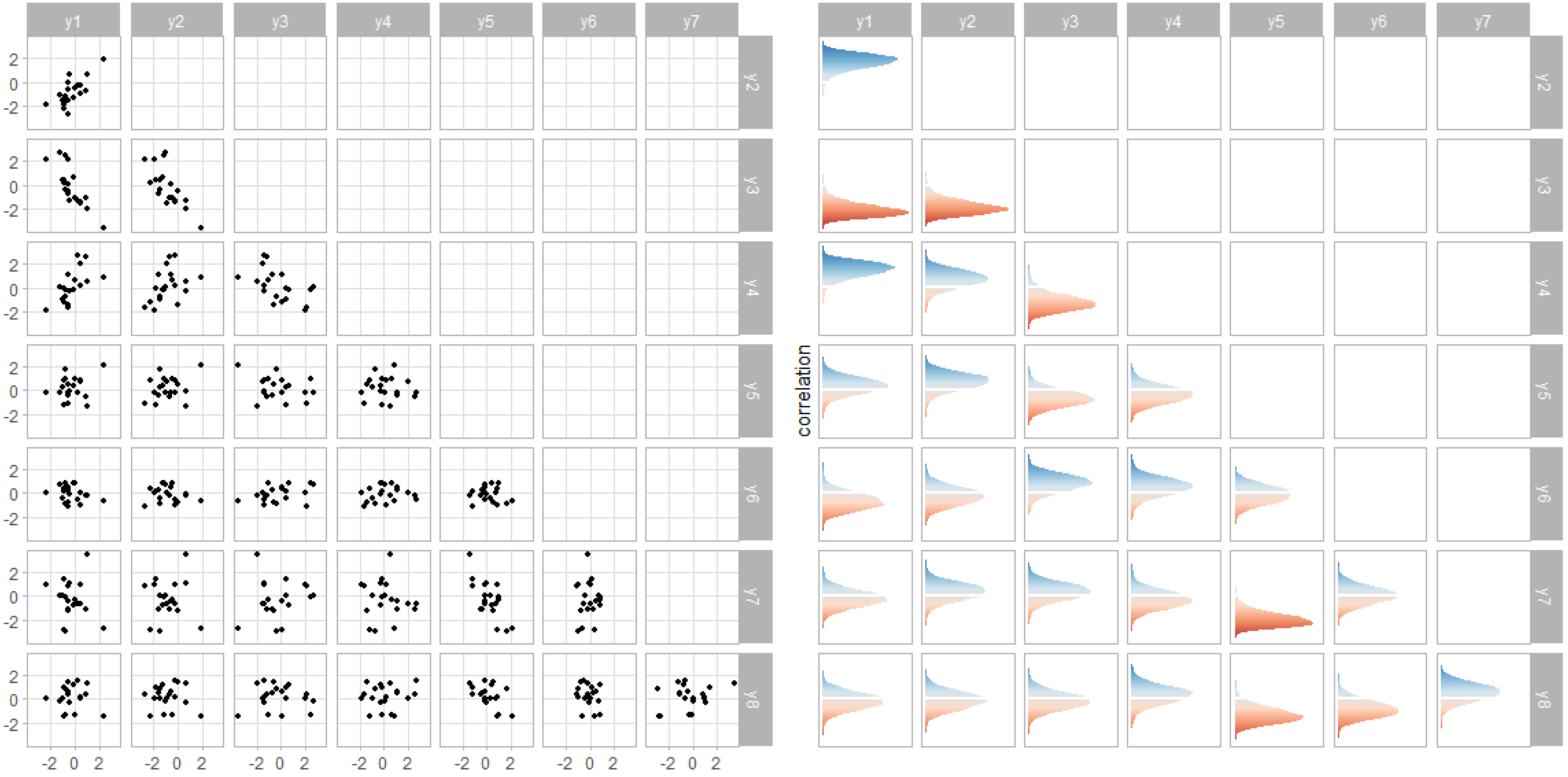

y1
y2
y2
y3
2
0
-2
2
0
-2
-2
-1
0
1
2
-2
-1
0
1
2



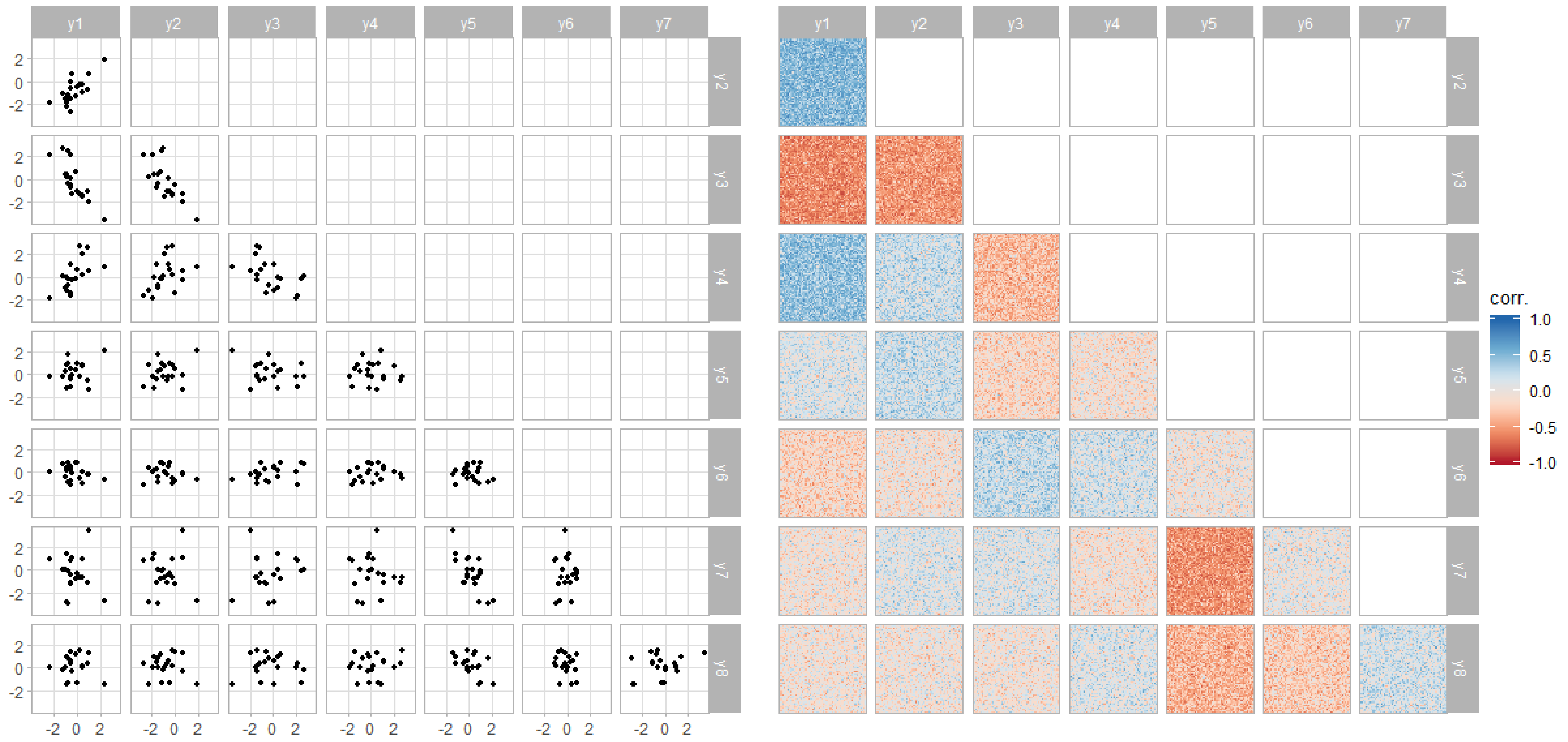
y1
y2
y2
y3
y1
y2
y2
y3
correlation



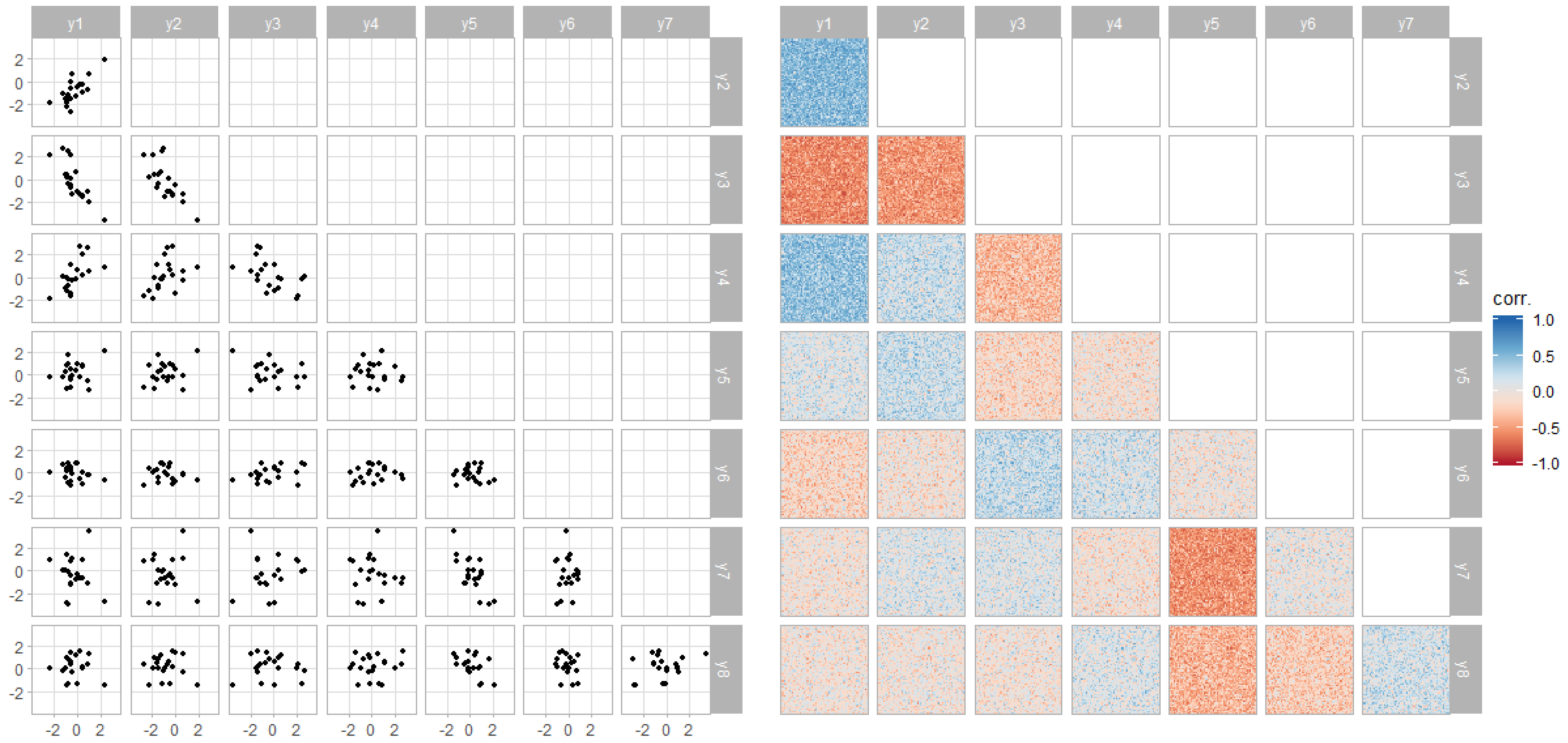
y1
y2
y2
y3
-2
-1
0
1
2
correlation
1.0
0.5
0.0
-0.5
-1.0



y1
y2
y3
y4
y5
y6
y7
y8
2
0
-2
correlation



y1
y2
y3
y4
y5
y6
y7
y8
corr.
1.0
0.5
0.0
-0.5
-1.0

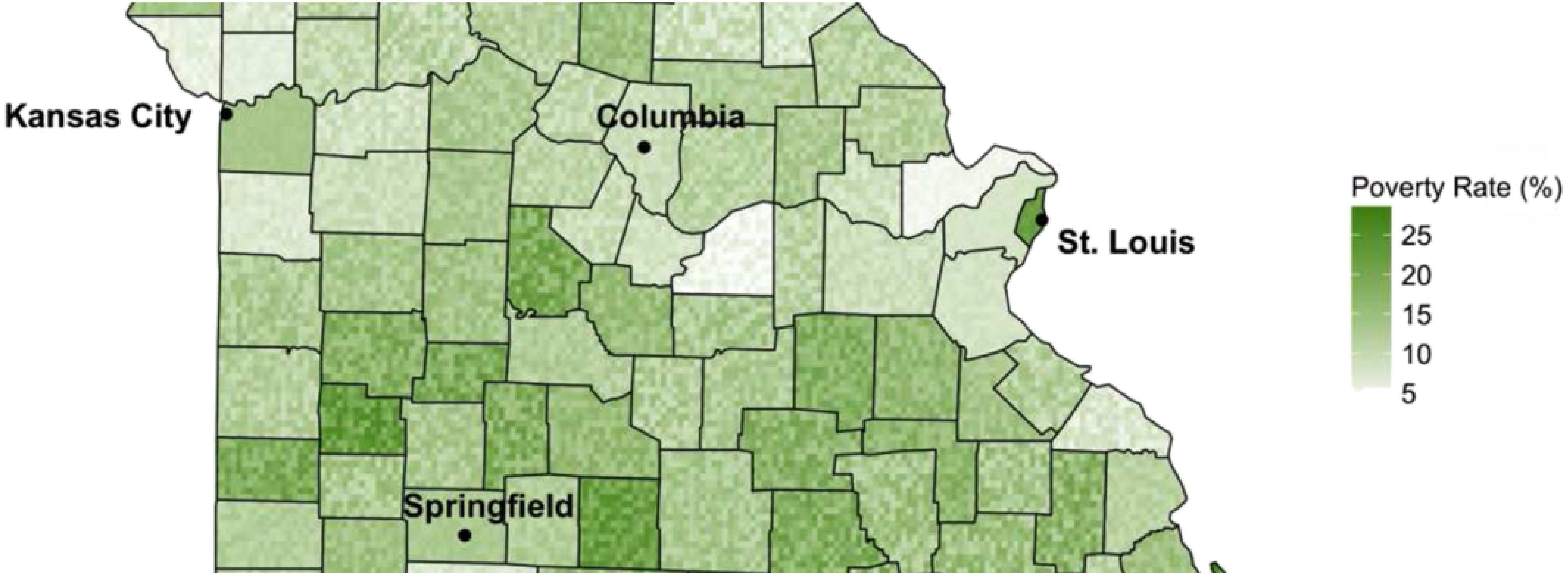



y1
y2
y3
y4
y5
y6
y7
y8
corr.
1.0
0.5
0.0
-0.5
-1.0

and back to map-land...

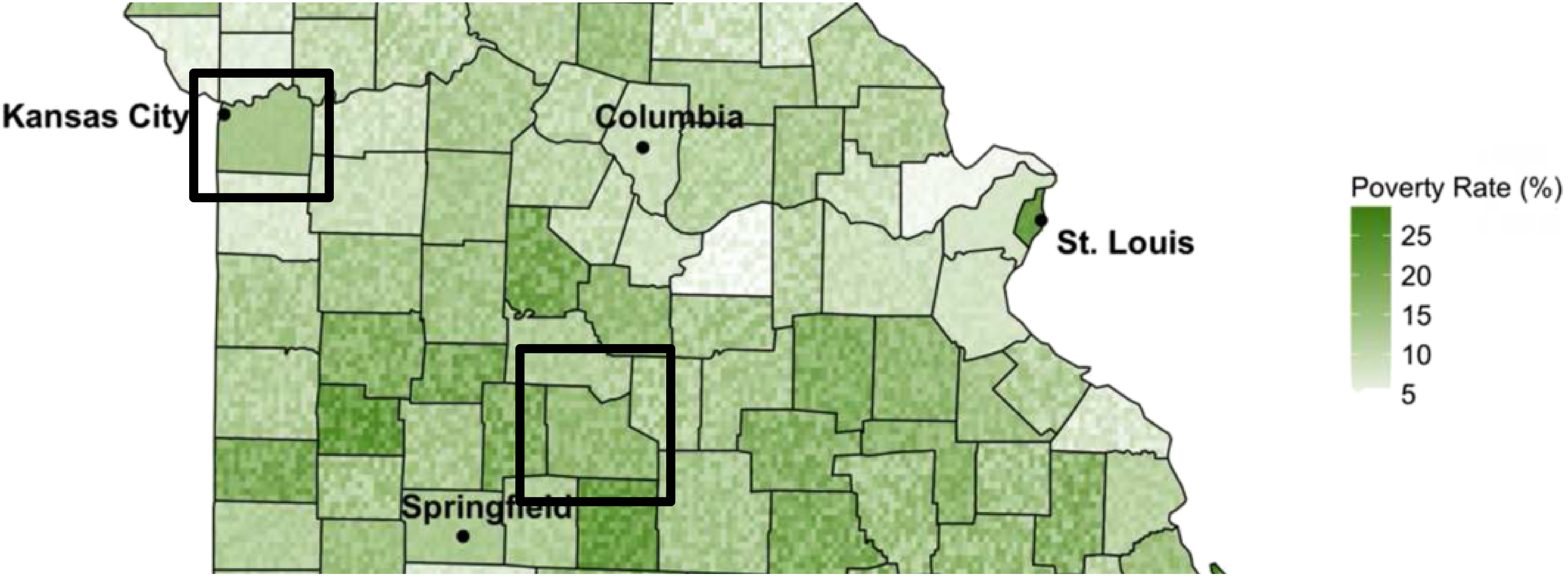
Uncertainty -> ~dither (samples from dist)

[Lucchesi & Wikle. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. Stat, 292–302, 2017]



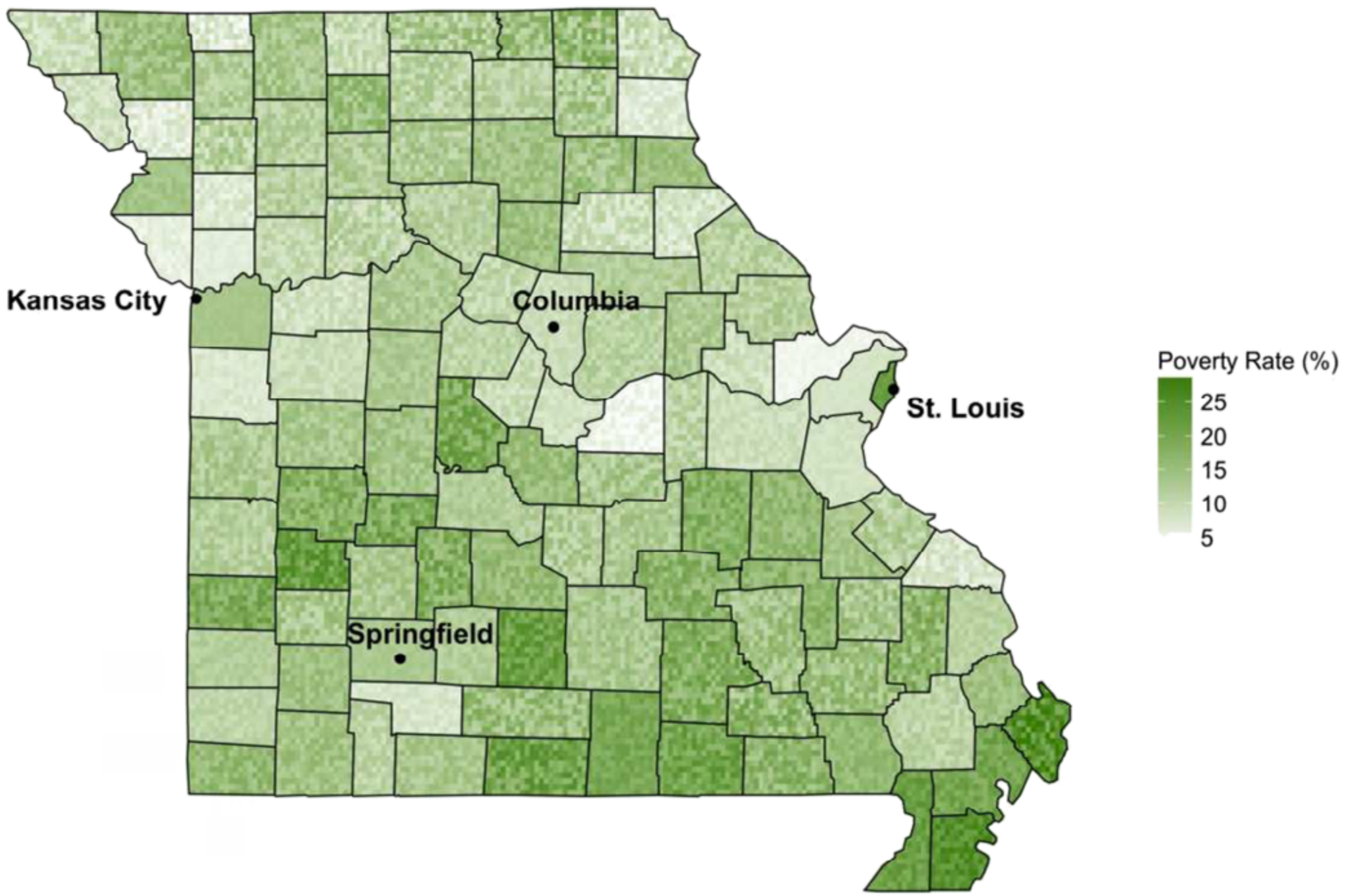
Uncertainty -> ~dither (samples from dist)

[Lucchesi & Wikle. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. Stat, 292–302, 2017]



Uncertainty -> ~dither (samples from dist)

[Lucchesi & Wikle. Visualizing uncertainty in areal data with bivariate choropleth maps, map pixelation and glyph rotation. Stat, 292–302, 2017]



Discrete outcomes

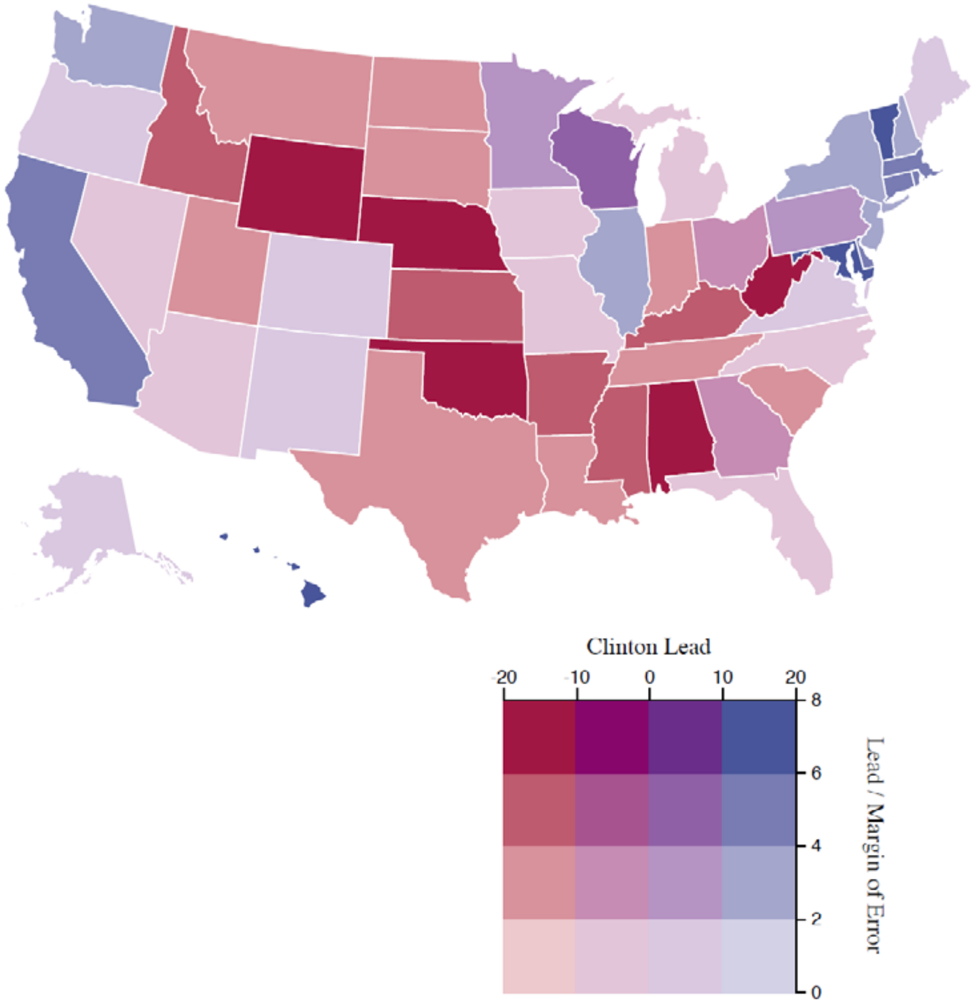
Maybe more intuitive, maybe less?

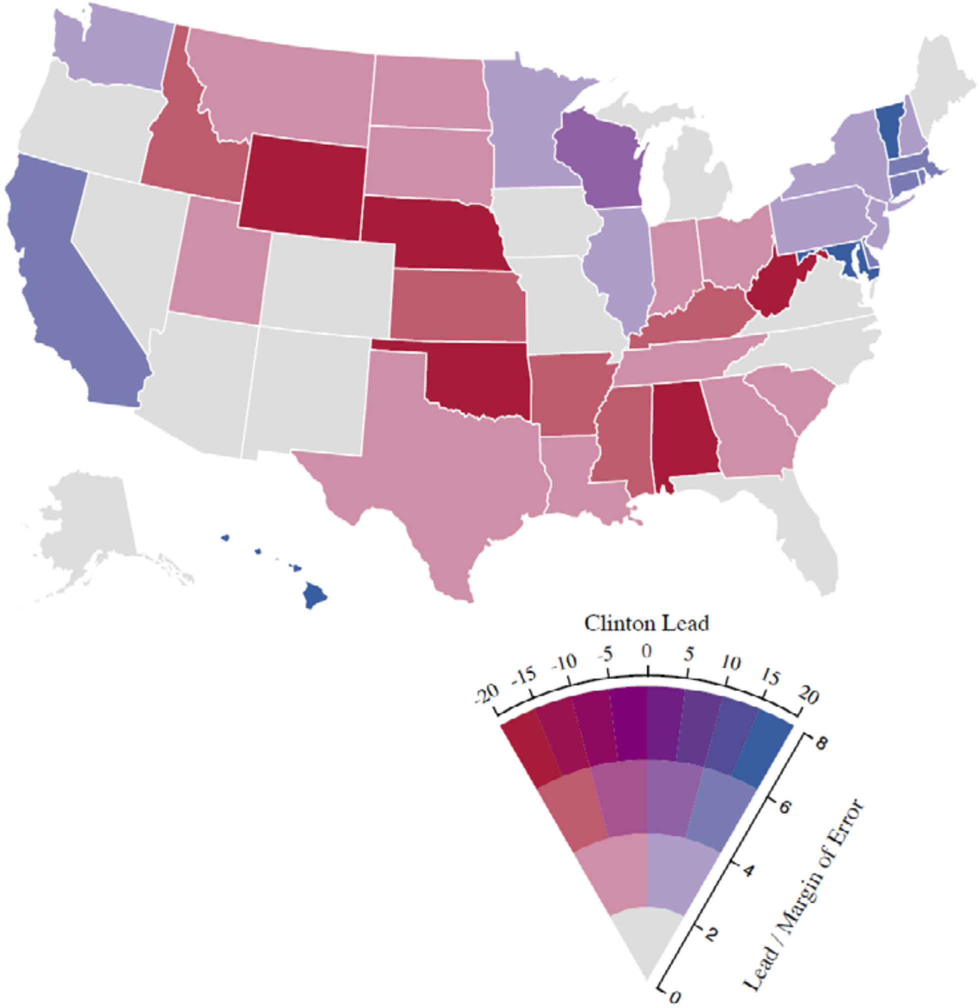
Possible deterministic construal errors

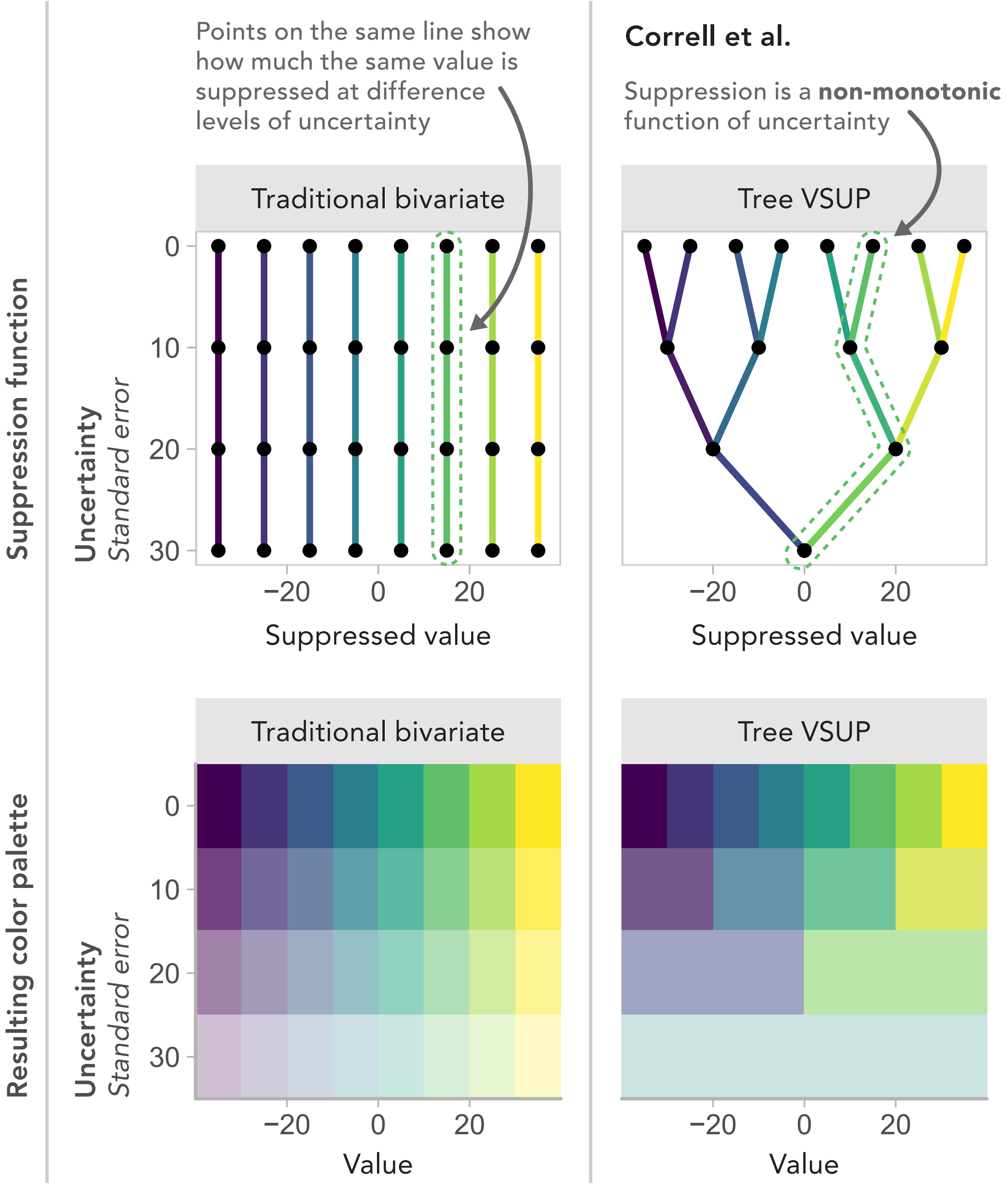
Addressing bias in perception of probability...

Value-suppressing uncertainty palettes

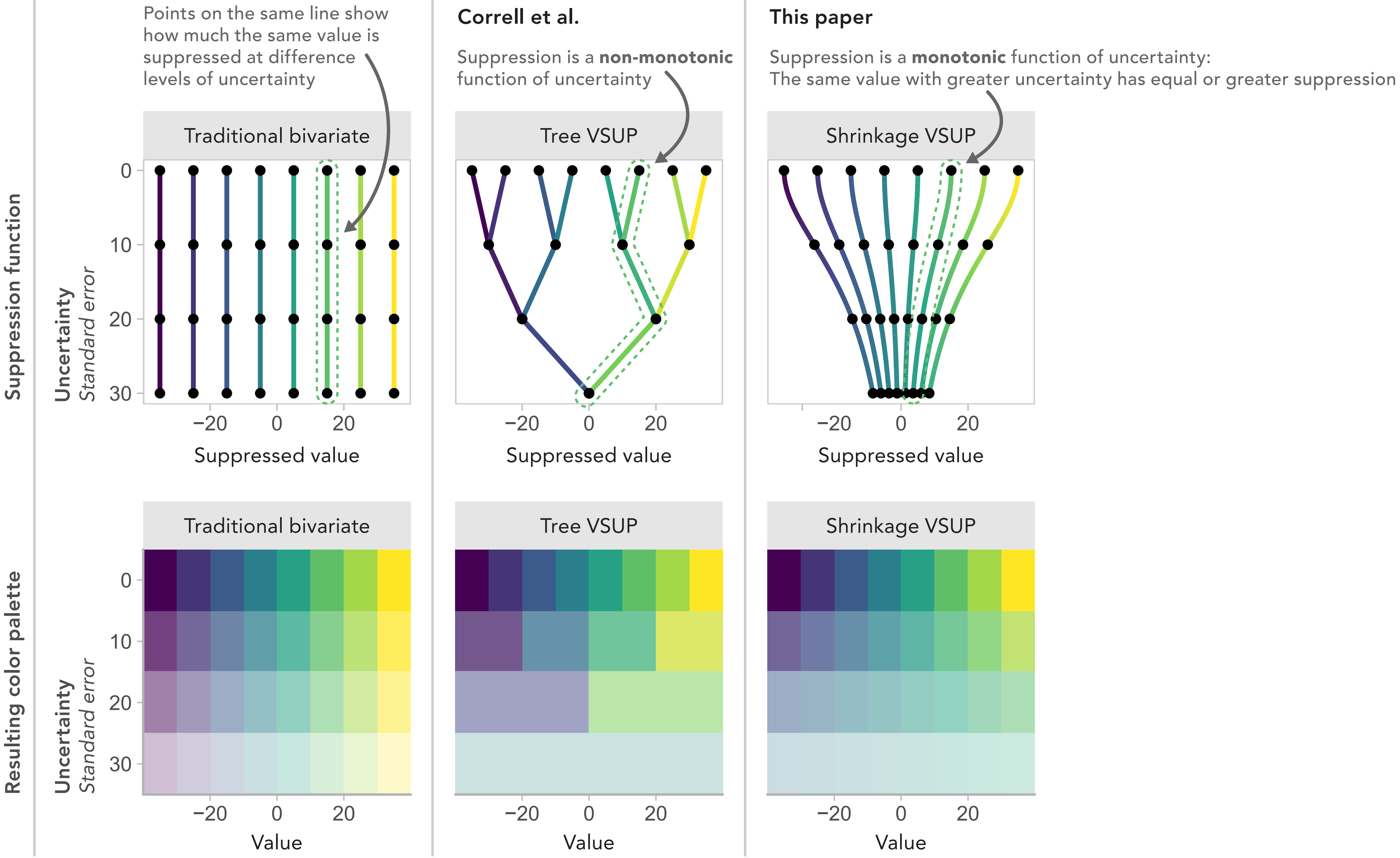
[Correll, Moritz, Heer. Value-Suppressing Uncertainty Palettes. CHI 2018]







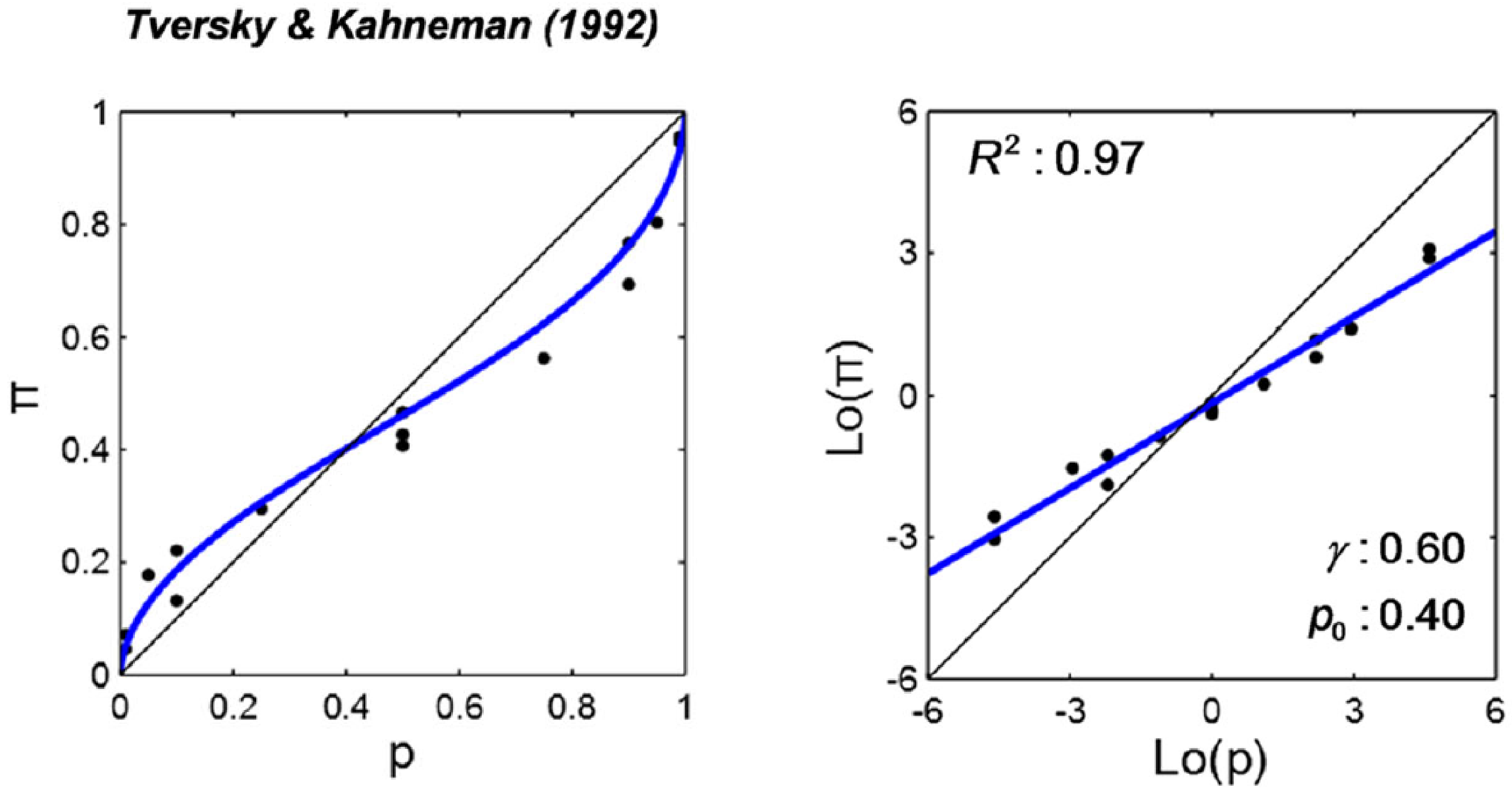
Points on the same line show how much the same value is suppressed at difference levels of uncertainty
Correll et al.
Suppression is a **non-monotonic** function of uncertainty
Suppression function
Resulting color palette
Traditional bivariate
Tree VSUP
Uncertainty
Standard error
0
10
20
30
−20
0
20
Suppressed value
Value



Suppression function
Resulting color palette
Points on the same line show how much the same value is suppressed at difference levels of uncertainty
Correll et al.
Suppression is a **non-monotonic** function of uncertainty
This paper
Suppression is a **monotonic** function of uncertainty: The same value with greater uncertainty has equal or greater suppression
Traditional bivariate
Tree VSUP
Shrinkage VSUP
Uncertainty
Standard error
0
10
20
30
−20
0
20
Suppressed value
Value

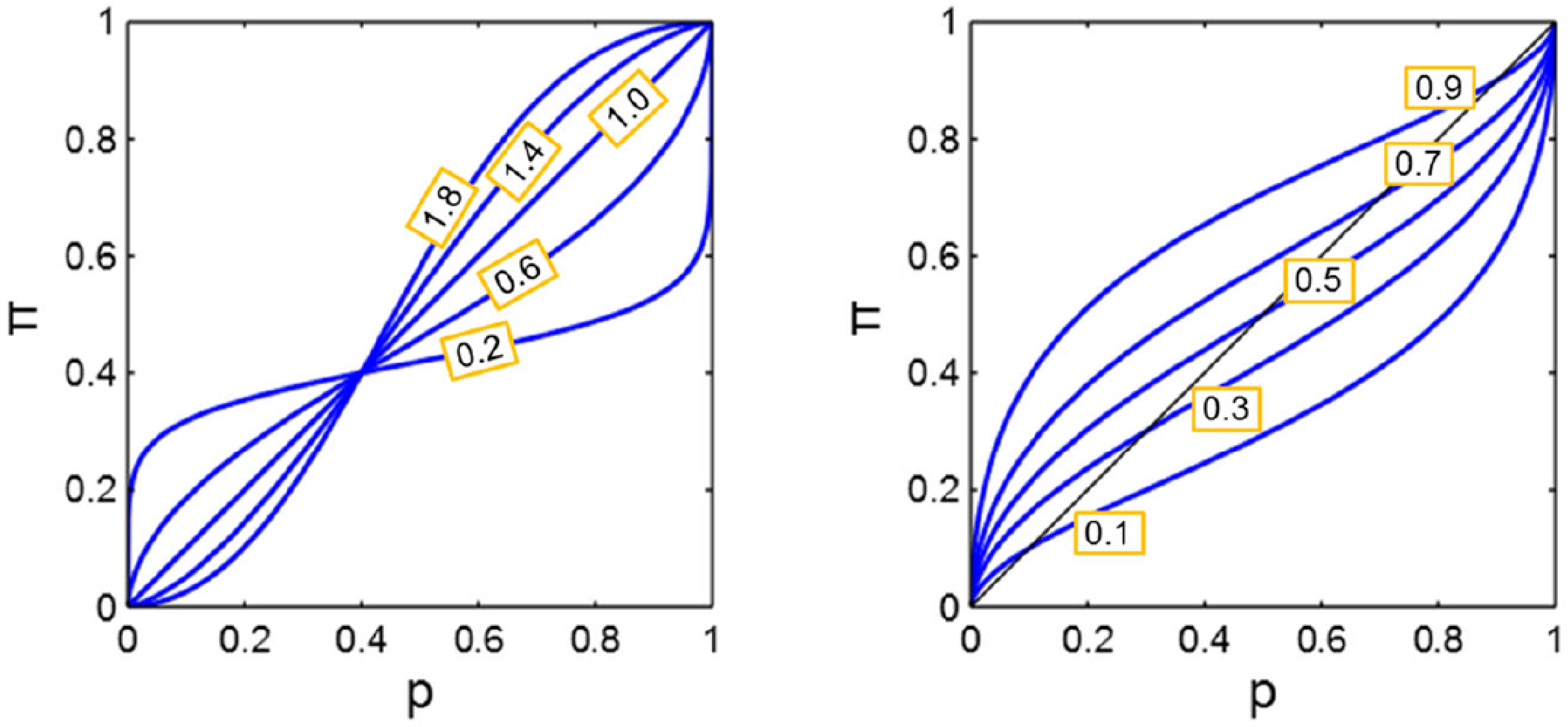
Linear-in-log-odds perception of proportions

[Zhang & Maloney. Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. Frontiers in Neuroscience, 6(JAN), 1–14, 2012]



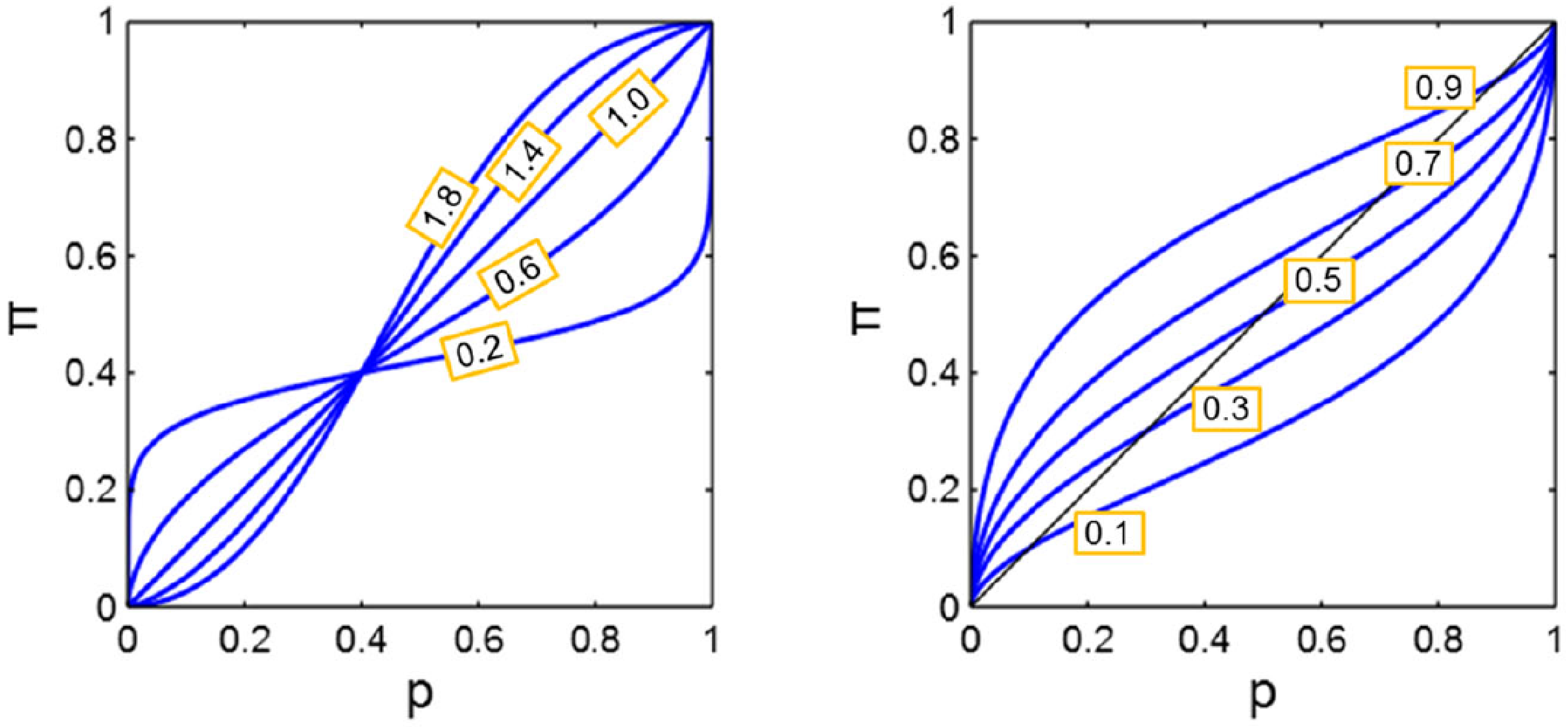
Linear-in-log-odds perception of proportions

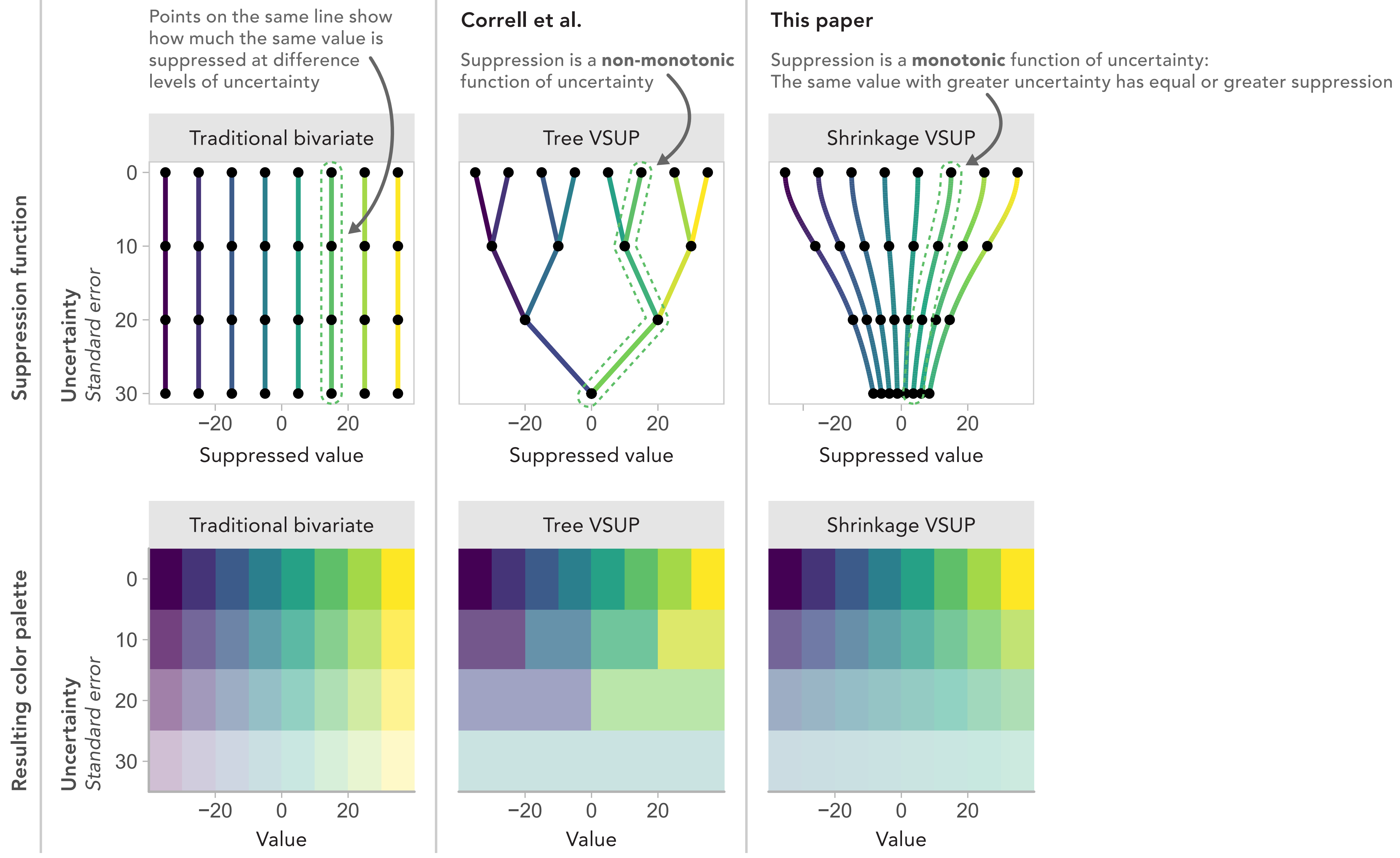
[Zhang & Maloney. Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. Frontiers in Neuroscience, 6(JAN), 1–14, 2012]



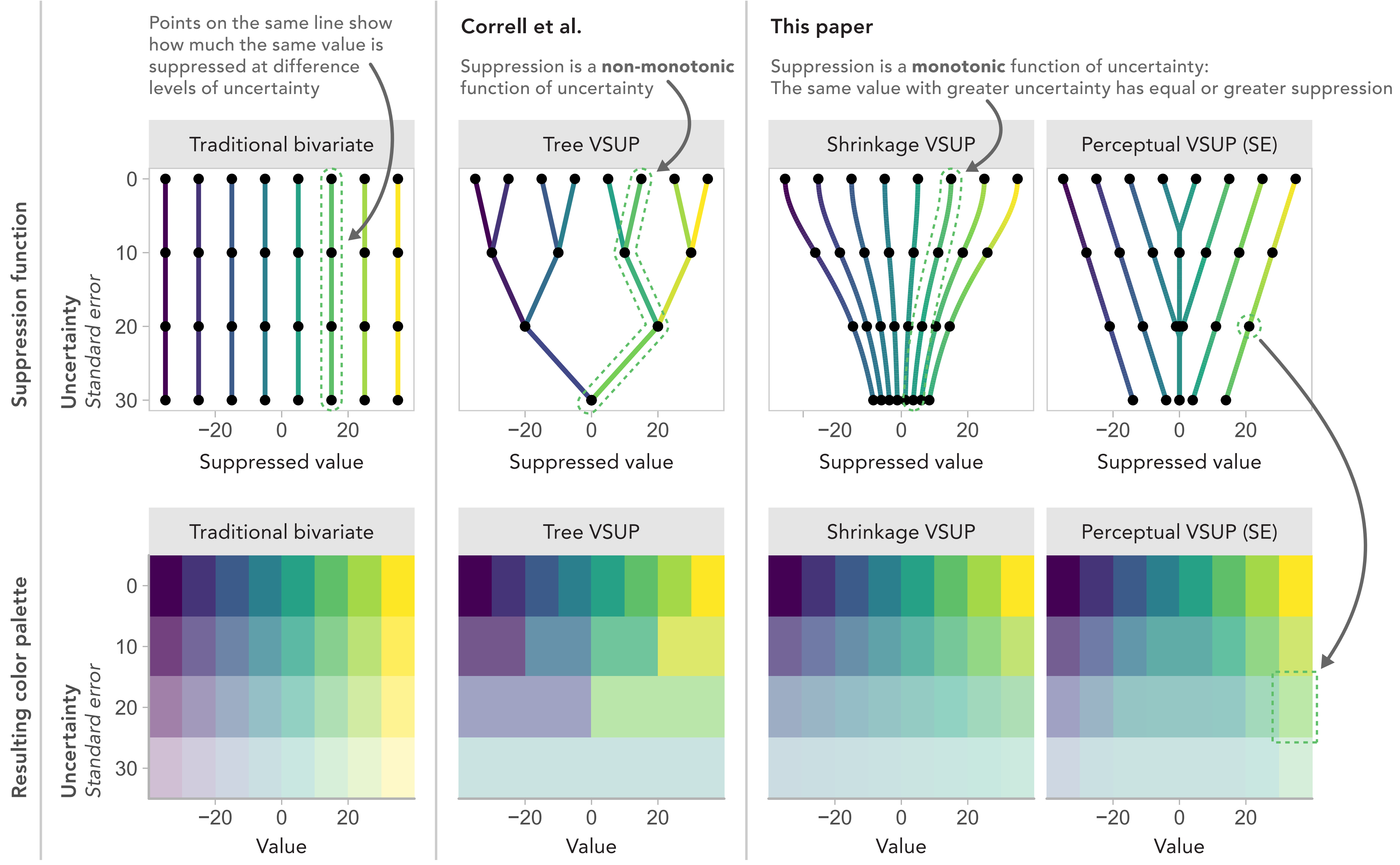
Linear-in-probit perception of proportions

[Zhang & Maloney. Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. Frontiers in Neuroscience, 6(JAN), 1–14, 2012]





Points on the same line show how much the same value is suppressed at difference levels of uncertainty
Correll et al.
Suppression is a **non-monotonic** function of uncertainty
This paper
Suppression is a **monotonic** function of uncertainty:
The same value with greater uncertainty has equal or greater suppression
Suppression function
Resulting color palette
Traditional bivariate
Tree VSUP
Shrinkage VSUP
Uncertainty
Standard error
0
10
20
30
−20
0
20
Suppressed value
Value



Points on the same line show how much the same value is suppressed at difference levels of uncertainty
Correll et al.
Suppression is a **non-monotonic** function of uncertainty
This paper
Suppression is a **monotonic** function of uncertainty:
The same value with greater uncertainty has equal or greater suppression
Suppression function
Resulting color palette
Uncertainty
Standard error
Traditional bivariate
Tree VSUP
Shrinkage VSUP
Perceptual VSUP (SE)
0
10
20
30
−20
0
20
Suppressed value
Value

Deterministic construal errors

[Joslyn & LeClerc. Decisions With Uncertainty: The Glass Half Full. Current Directions in Psych. Science, 22(4), 2013]