

Uniformity Simulations to Evaluate Trial Randomization Quality

Peter Claussen, GDM Solutions

May 12, 2021



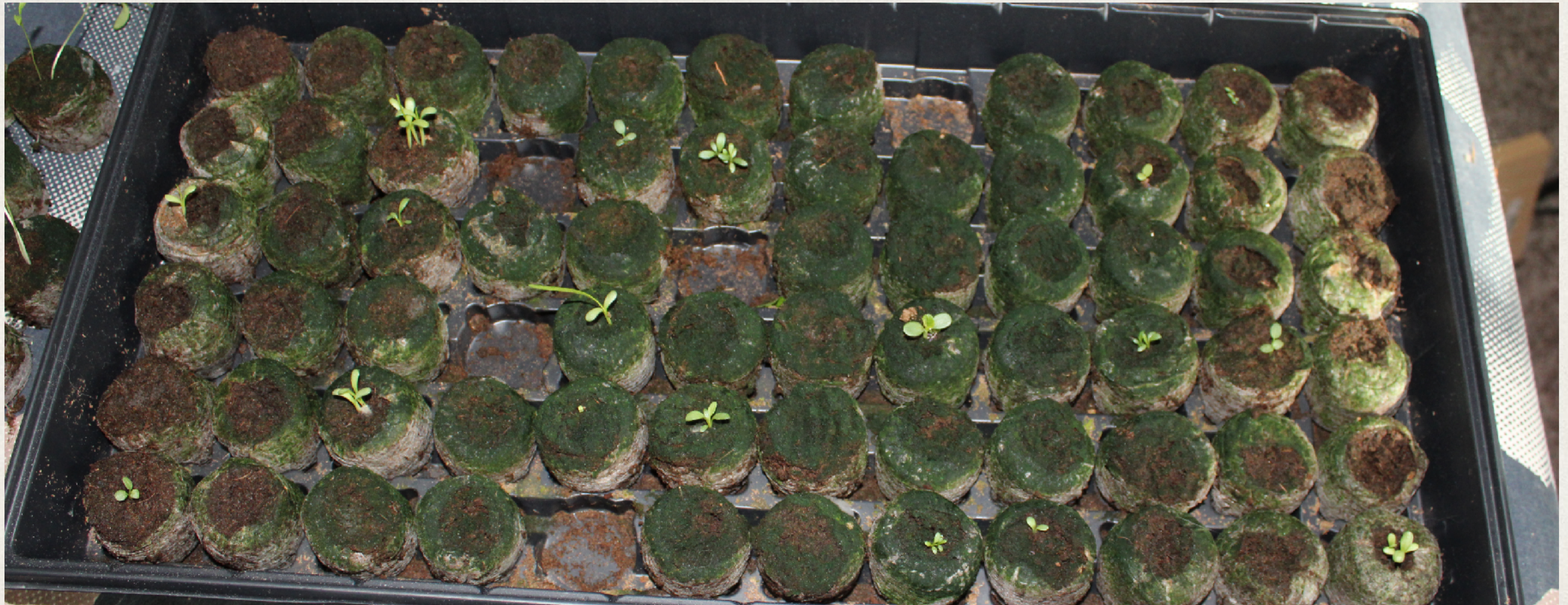
The Project

Restoring Native Prairie

Semi-native Pasture

- ❖ 35 acres in eastern South Dakota
- ❖ Northern 10 acres contains variable terrain and a diverse mix of native grasses and forms.
- ❖ Southern 10 acres has little diversity, large patches of non-native grasses.
- ❖ The goal is to reintroduce native varieties from north to south.





The Experiment

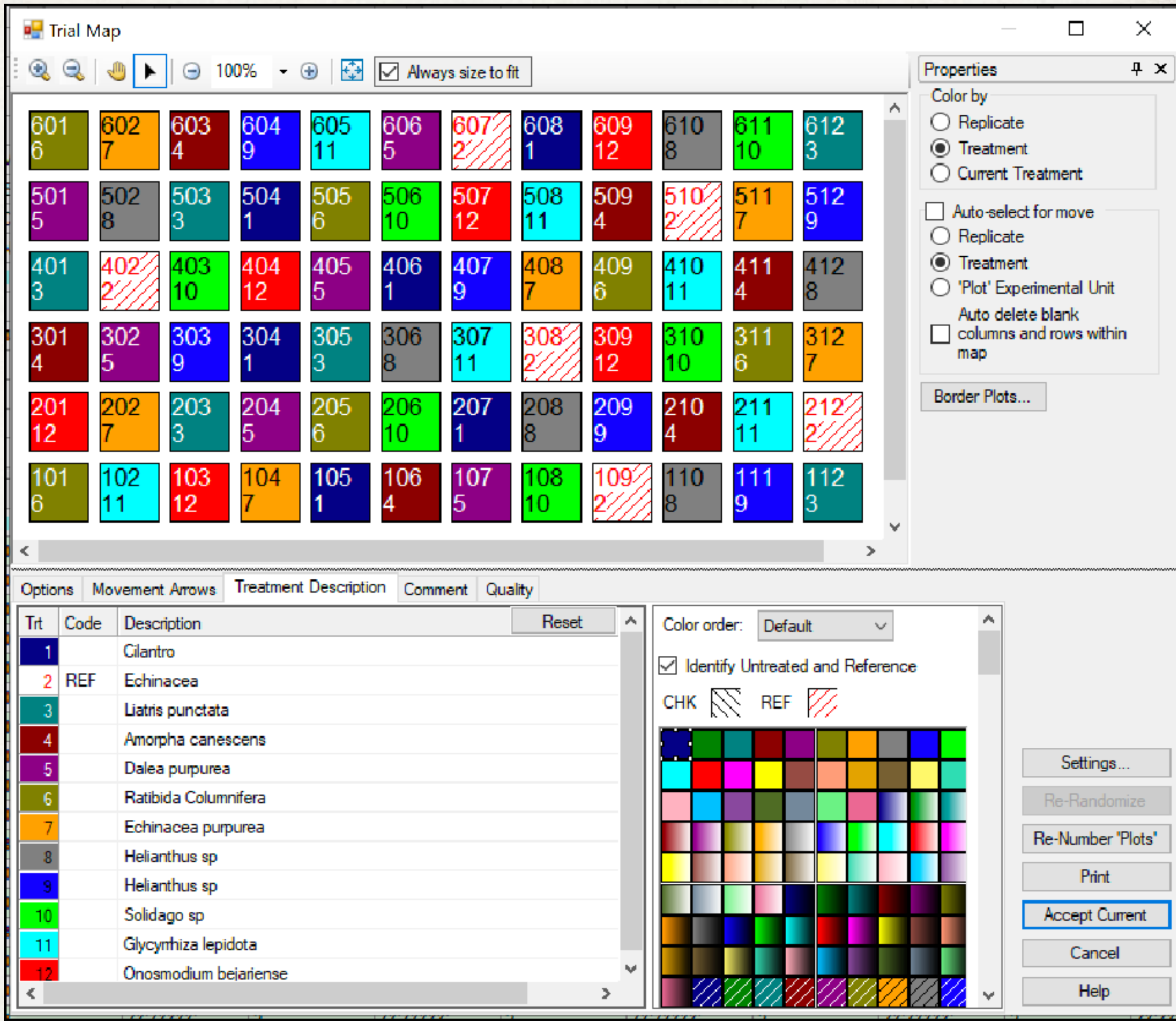
Seed Germination Assay

Determine percent germination of harvests of 10 different native species (plus 2 checks)

Trial Map

Since this is effectively a greenhouse trial (Jiffy Professional Seed Starting Greenhouse), I did not need to randomize as a complete block design.

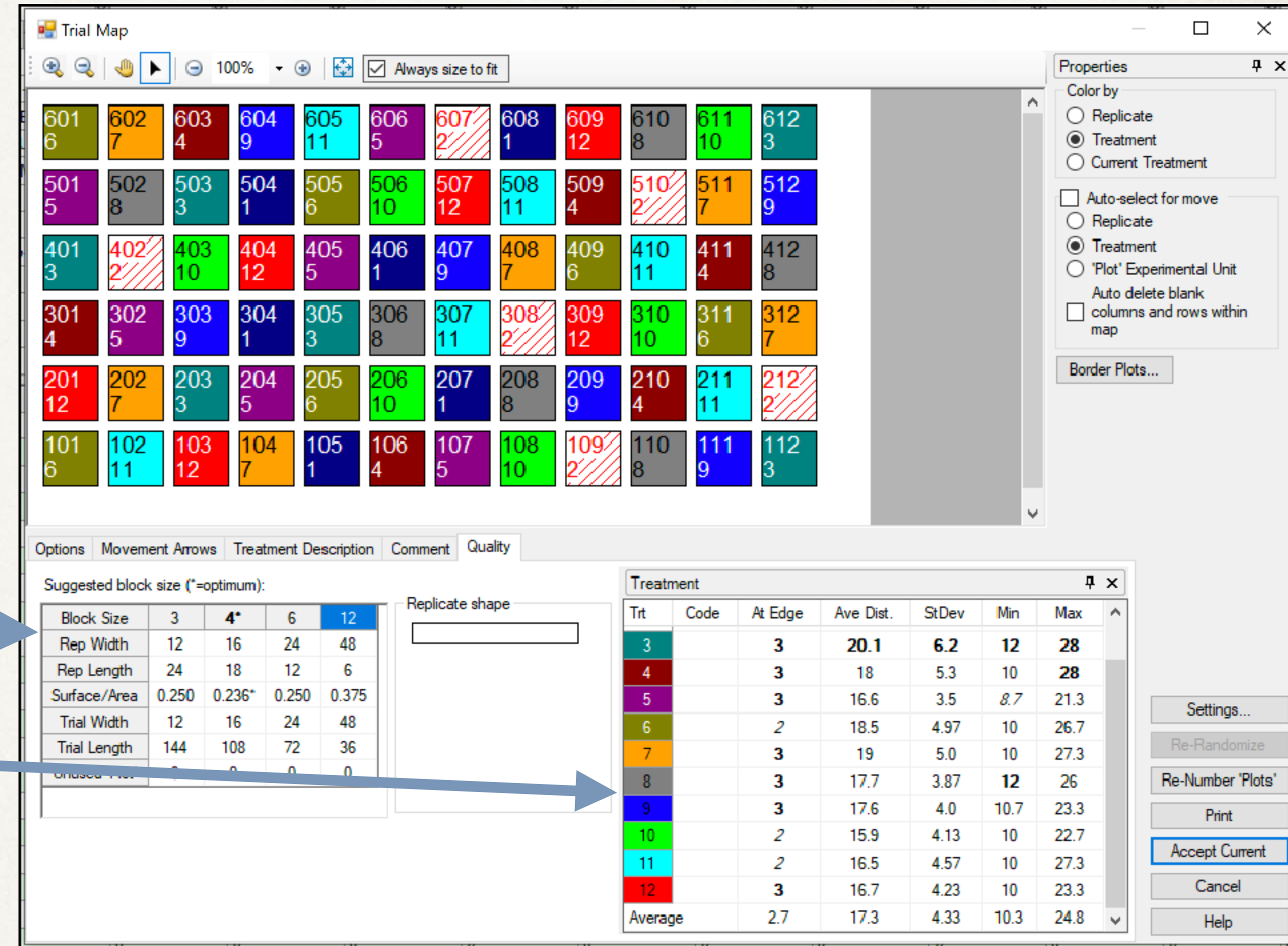
However, I knew the kit would be placed next to a window, creating a natural light gradient. To account for this, I randomized the trial as an RCB.



Trial Quality

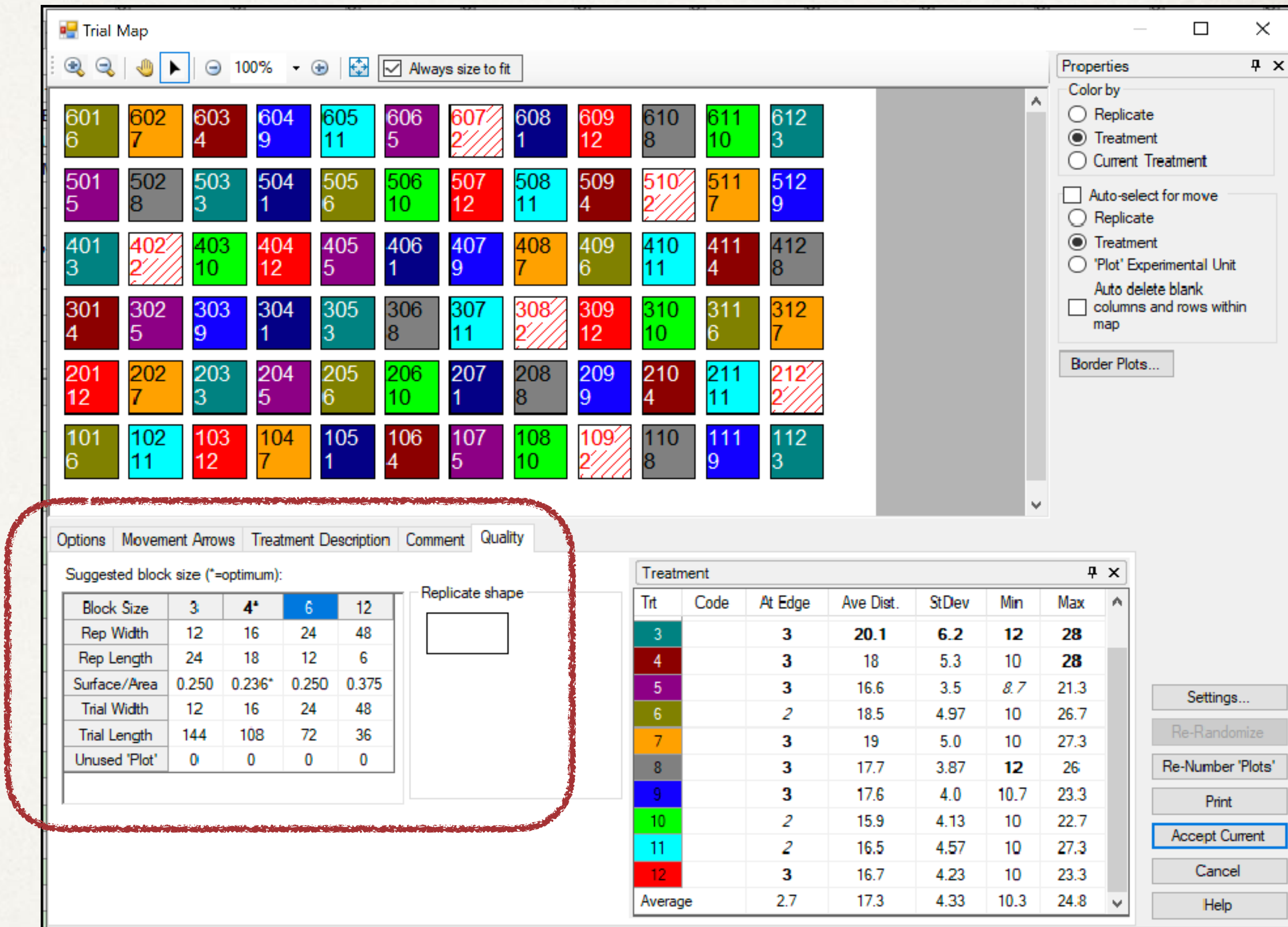
❖ Replicate Shape

❖ Treatment Dispersion



Trial Quality

- ❖ Replicate Shape
- ❖ Suggested Block Size
- ❖ Treatment Dispersion



Trial Quality

- ❖ Replicate Shape
- ❖ Suggested Block Size
- ❖ How will this affect trial map quality?
- ❖ Treatment Dispersion

❖ As Executed

601 6	602 7	603 4	604 9	605 11	606 5	607 2	608 1	609 12	610 8	611 10	612 3
501 5	502 8	503 3	504 1	505 6	506 10	507 12	508 11	509 4	510 2	511 7	512 9
401 3	402 2	403 10	404 12	405 5	406 1	407 9	408 7	409 6	410 11	411 4	412 8
301 4	302 5	303 9	304 1	305 3	306 8	307 11	308 2	309 12	310 10	311 6	312 7
201 12	202 7	203 3	204 5	205 6	206 10	207 1	208 8	209 9	210 4	211 11	212 2
101 6	102 11	103 12	104 7	105 1	106 4	107 5	108 10	109 2	110 8	111 9	112 3

❖ 2x6 (Better?)

1001 12	1002 11	1003 4	1004 2	1005 7	1006 9	1201 2	1202 1	1203 12	1204 8	1205 10	1206 3
901 5	902 8	903 3	904 1	905 6	906 10	1101 6	1102 7	1103 4	1104 9	1105 11	1106 5
801 11	802 2	803 12	804 10	805 6	806 7	701 5	702 12	703 4	704 1	705 8	706 7
601 4	602 5	603 9	604 1	605 3	606 8	501 10	502 3	503 9	504 2	505 11	506 6
201 5	202 10	203 2	204 8	205 9	206 3	401 1	402 8	403 9	404 4	405 11	406 2
101 6	102 11	103 12	104 7	105 1	106 4	301 12	302 7	303 3	304 5	305 8	306 10

❖ 3x4 (Best?)

1201 6	1202 11	1203 4	1204 8	1501 4	1502 2	1503 7	1504 9	1801 12	1802 8	1803 10	1804 3
1101 5	1102 1	1103 9	1104 7	1401 6	1402 10	1403 12	1404 11	1701 11	1702 5	1703 2	1704 1
1001 3	1002 2	1003 10	1004 12	1301 5	1302 8	1303 3	1304 1	1601 6	1602 7	1603 4	1604 9
301 2	302 8	303 9	304 3	601 9	602 4	603 11	604 2	901 12	902 10	903 6	904 7
201 1	202 4	203 5	204 10	501 6	502 10	503 1	504 8	801 3	802 8	803 11	804 2
101 6	102 11	103 12	104 7	401 12	402 7	403 3	404 5	701 4	702 5	703 9	704 1

A bit of review



Two-way Analysis of Variance

Source	Sums of Squares	df	Mean Squares	F
Treatment	$SS_{trt} = J \sum_{i=1}^I (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$	$I - 1$	$MS_{trt} = \frac{SS_{trt}}{df_{trt}}$	$\frac{MS_{trt}}{MS_{error}}$
Block	$SS_{block} = I \sum_{j=1}^J (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2$	$J - 1$	$MS_{block} = \frac{SS_{block}}{df_{block}}$	$\frac{MS_{block}}{MS_{error}}$
Error	$SS_{error} = \sum_{i=1}^I \sum_{j=1}^J (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot})^2$	$(I - 1)(J - 1)$	$MS_{error} = \frac{SS_{error}}{df_{error}}$	

Calculating Replicate Means

$$\bar{y}_{\cdot 6} = \frac{1}{I} \sum_{i=1}^I \bar{y}_{i6}$$

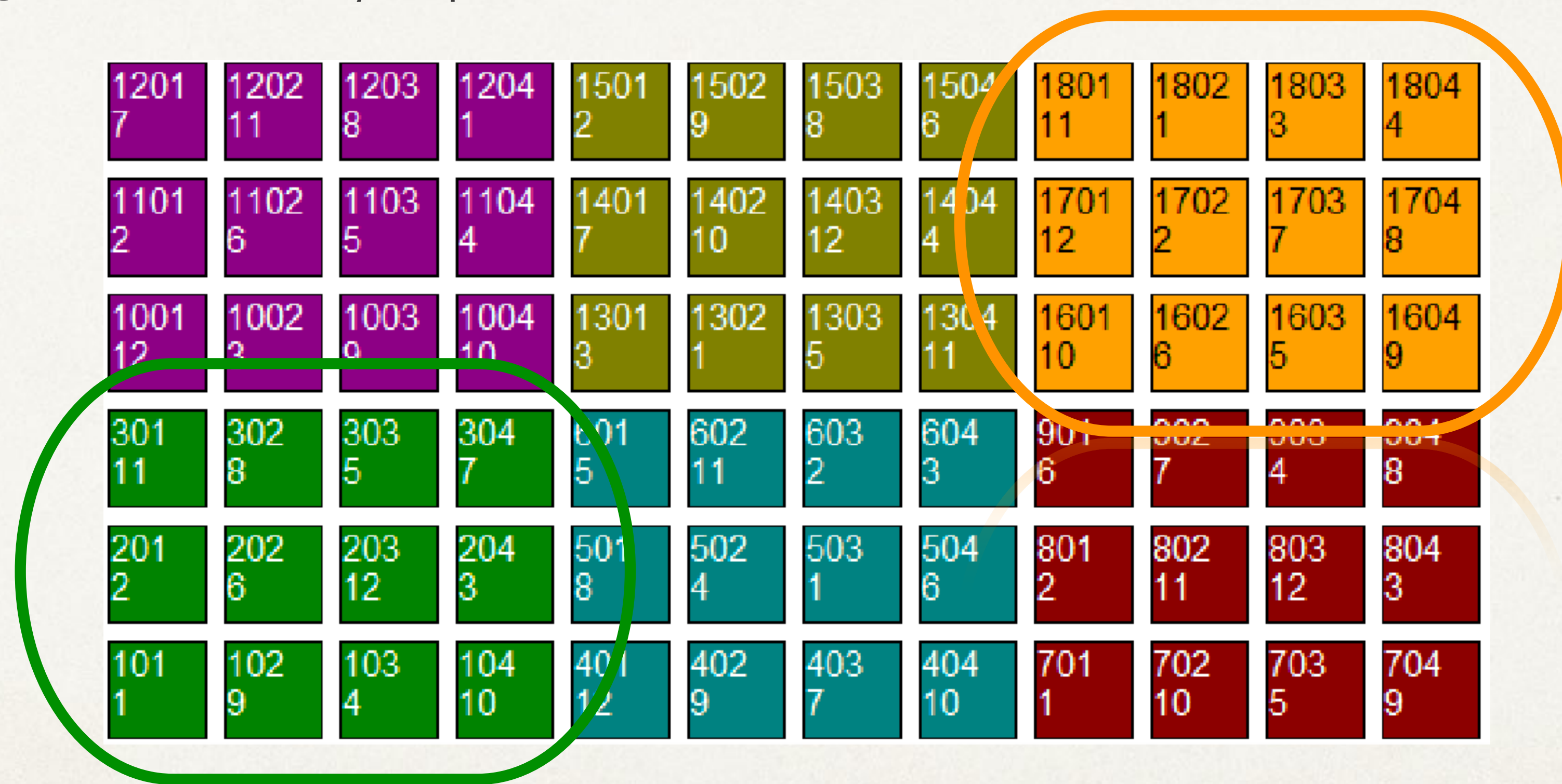
601 6	602 7	603 4	604 9	605 11	606 5	607 2	608 1	609 12	610 8	611 10	612 3
501 5	502 8	503 3	504 1	505 6	506 10	507 12	508 11	509 4	510 2	511 7	512 9
401 3	402 2	403 10	404 12	405 5	406 1	407 9	408 7	409 6	410 11	411 4	412 8
301 4	302 5	303 9	304 1	305 3	306 8	307 11	308 2	309 12	310 10	311 6	312 7
201 12	202 7	203 3	204 5	205 6	206 10	207 1	208 8	209 9	210 4	211 11	212 2
101 6	102 11	103 12	104 7	105 1	106 4	107 5	108 10	109 2	110 8	111 9	112 3

$$\bar{y}_{\cdot 1} = \frac{1}{I} \sum_{i=1}^I \bar{y}_{i1}$$

Calculating Replicate Means

Intuitively, it seems that 3x4 block size may be preferable to 1x12 block size. Can we test this?

$$\bar{y}_{\cdot 1} = \frac{1}{I} \sum_{i=1}^I \bar{y}_{i1}$$



1201 7	1202 11	1203 8	1204 1	1501 2	1502 9	1503 8	1504 6	1801 11	1802 1	1803 3	1804 4
1101 2	1102 6	1103 5	1104 4	1401 7	1402 10	1403 12	1404 4	1701 12	1702 2	1703 7	1704 8
1001 12	1002 3	1003 9	1004 10	1301 3	1302 1	1303 5	1304 11	1601 10	1602 6	1603 5	1604 9
301 11	302 8	303 5	304 7	601 5	602 11	603 2	604 3	901 6	902 7	903 4	904 8
201 2	202 6	203 12	204 3	501 8	502 4	503 1	504 6	801 2	802 11	803 12	804 3
101 1	102 9	103 4	104 10	401 12	402 9	403 7	404 10	701 1	702 10	703 5	704 9

$$\bar{y}_{\cdot 6} = \frac{1}{I} \sum_{i=1}^I \bar{y}_{i6}$$

Simulation



Simulated Uniformity Trials

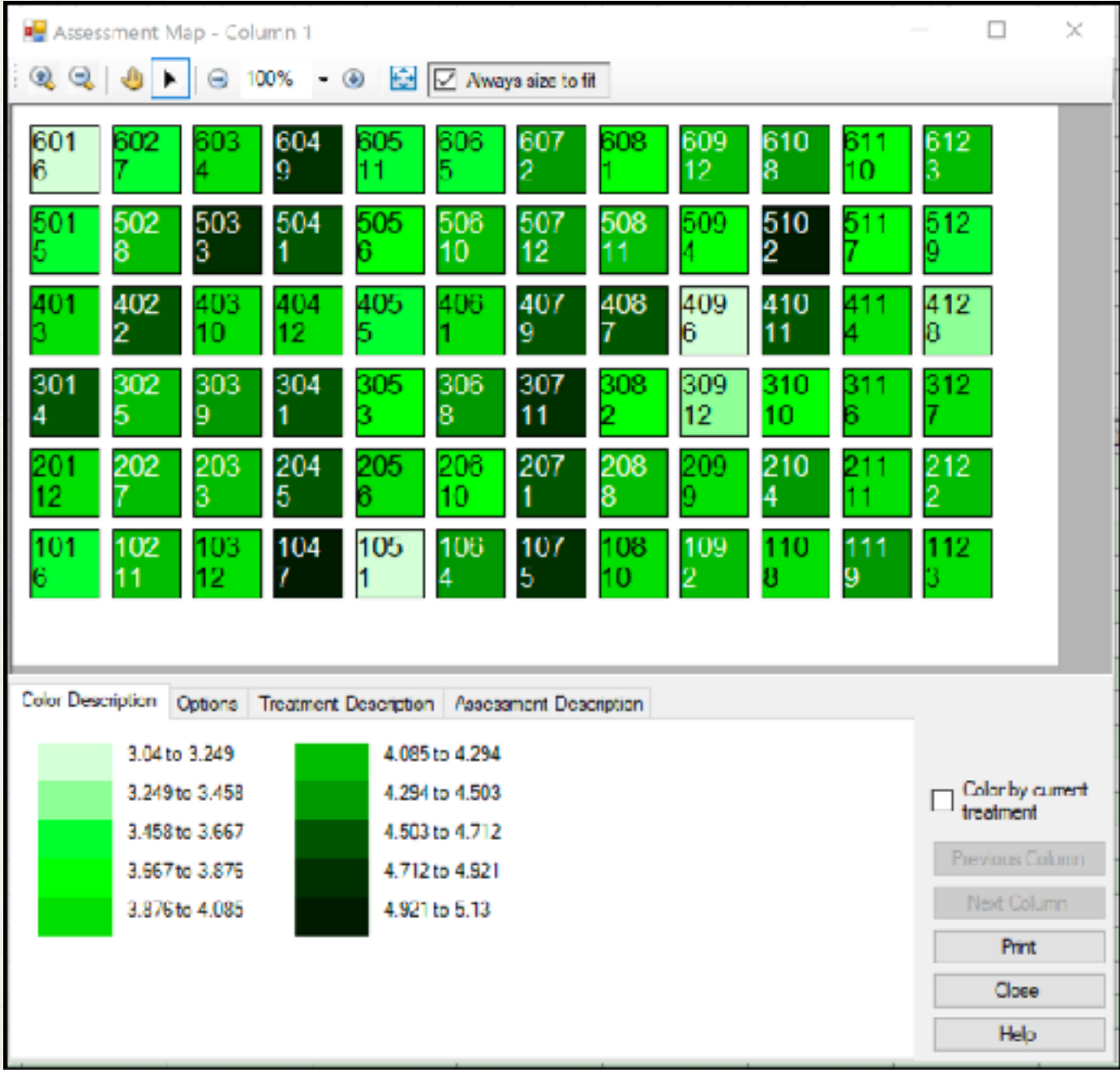
- ❖ Start with a uniformity trial (from many available in the literature)
- ❖ Copy plot assessments to a proposed randomization
- ❖ Analyze as if this were an actual trial
- ❖ Since there is no actual treatments, we expect $P(>F_{\text{trt}})$ to be large.
- ❖ If treatments are confounded with a spatial trend we may get a significant $P(>F_{\text{trt}})$
- ❖ How often do significant $P(>F_{\text{trt}})$ occur over uniform cropland data?

Mercer and Hall

- ❖ Mercer, W. B. and Hall, A. D. 1911. The experimental error of field trials. The Journal of Agricultural Science. 4 (2), pp. 107-132
- ❖ The magnitude of experimental error attaching to one or more field plots is a question of extreme importance in Agricultural Science, because upon its proper recognition depends the degree of confidence which may be attached to the results obtained in field work. A very cursory examination of the results of any set of field trials will serve to show that a pair of plots similarly treated may be expected to yield considerably different results, even when the soil appears to be uniform and the conditions under which the experiment is conducted are carefully designed to reduce errors in weighing and measurement.



Simulated Trial



Randomized Complete Block (RCB) Least square estimation AOV For C 1:1 1 1 Apr-20-2021 (Data Column 1)

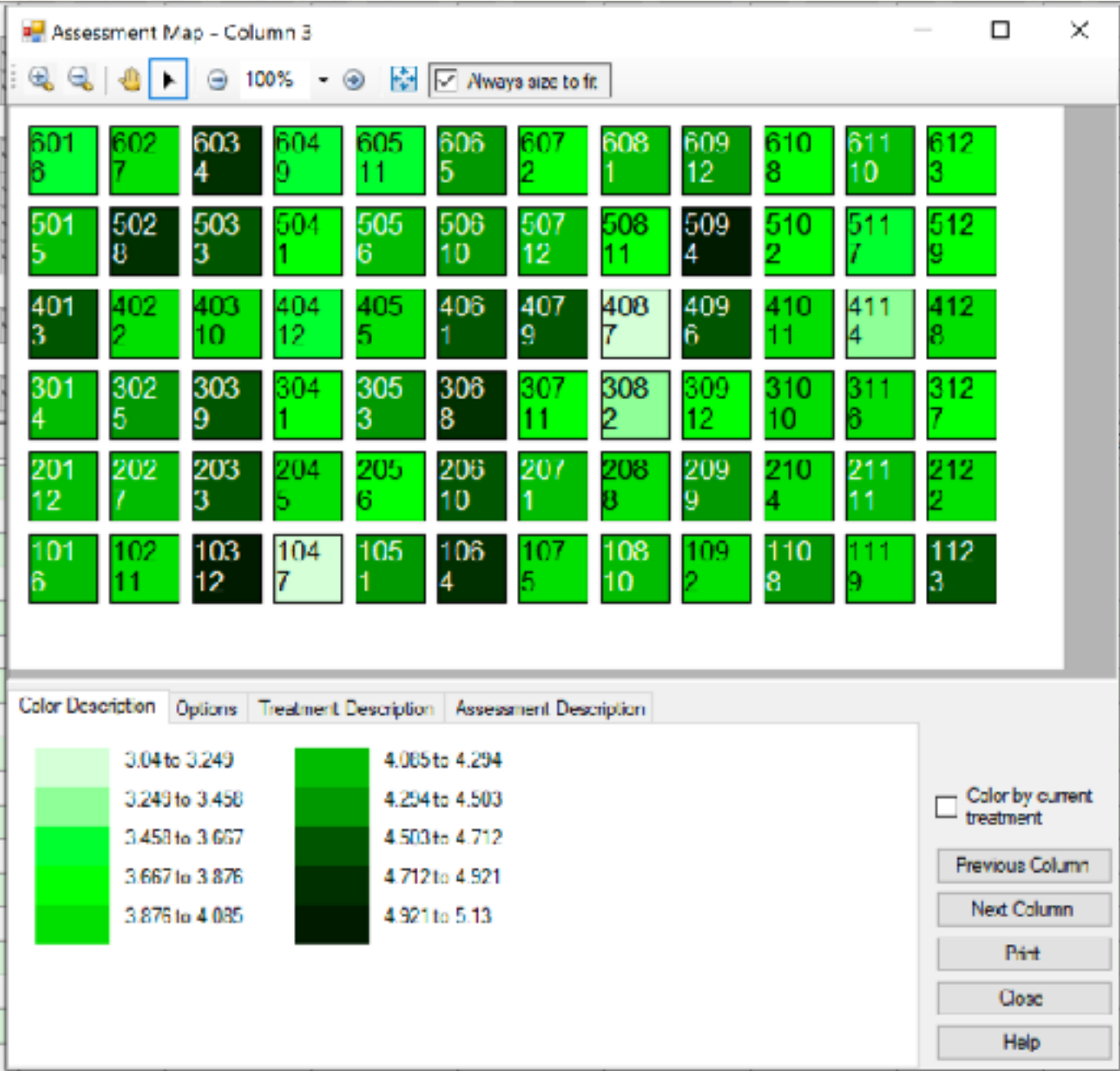
Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total	71	14.454461			
Replicate	5	0.566411	0.113282	0.555	0.7337
Treatment	11	2.667061	0.242460	1.188	0.3166
Error	55	11.220989	0.204018		

See H. M. van Es and C. L. van Es. (1993) Spatial nature of randomization and its effect on the outcome of field experiments. Agron J, 85:420–428 for something similar

Simulated Uniformity Trials

- ❖ Repeat a single randomization over multiple samples from uniformity trials.
- ❖ In the absence of a true treatment effect, we expect to detect significance (at 5% probability) in about 5% of the trials.
- ❖ We can compare the observed distribution of p-values over multiple uniformity trials with the expected distribution.

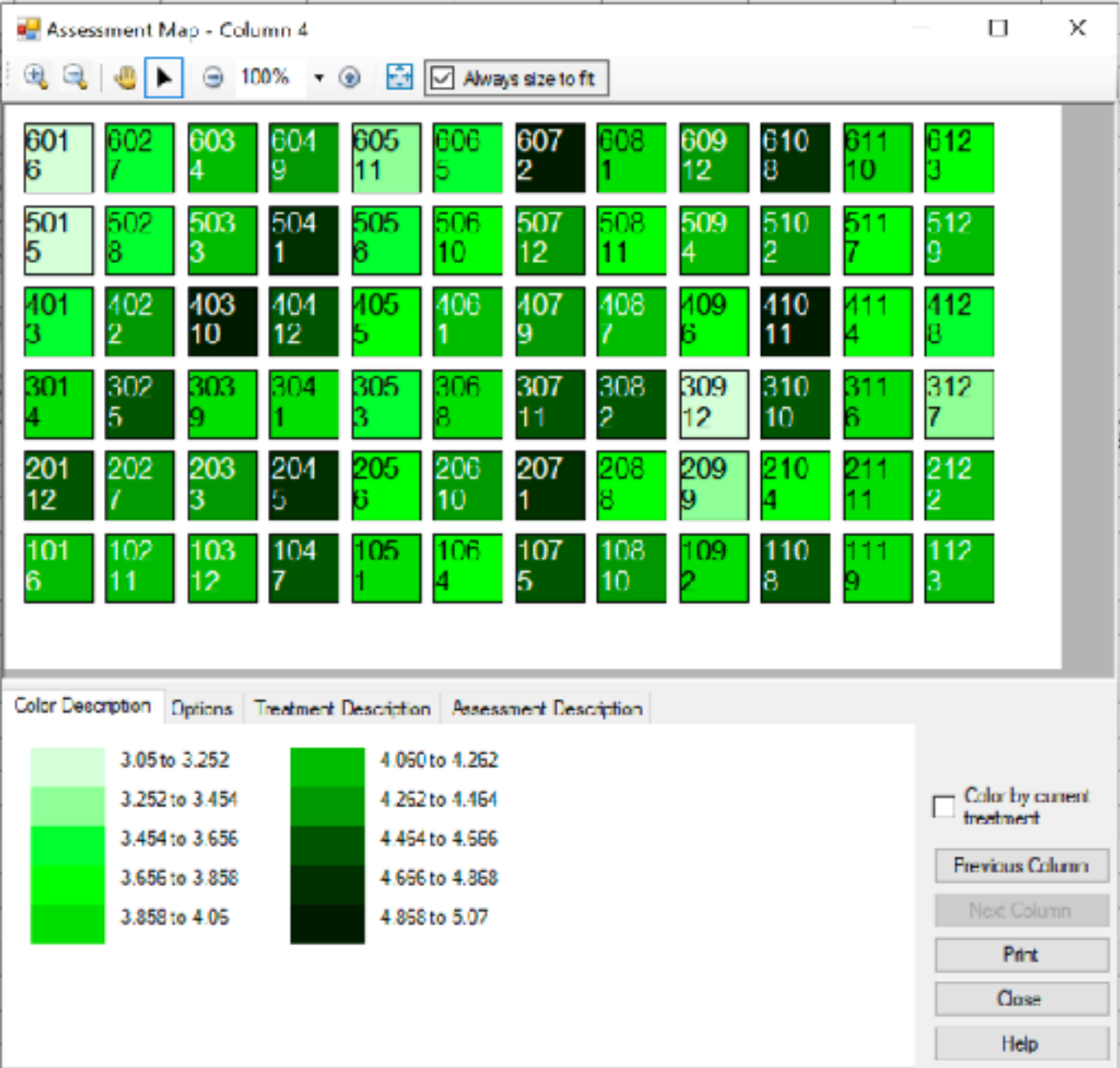
Simulated Trial (1:2)



Randomized Complete Block (RCB) Least square estimation AOV For C 1:2 1 May-5-2021 (Data Column 3)

Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total	71	13.545044			
Replicate	5	0.568444	0.113689	0.673	0.6457
Treatment	11	3.684144	0.334922	1.982	0.0482
Error	55	9.292456	0.168954		

Simulated Trial (2:1)

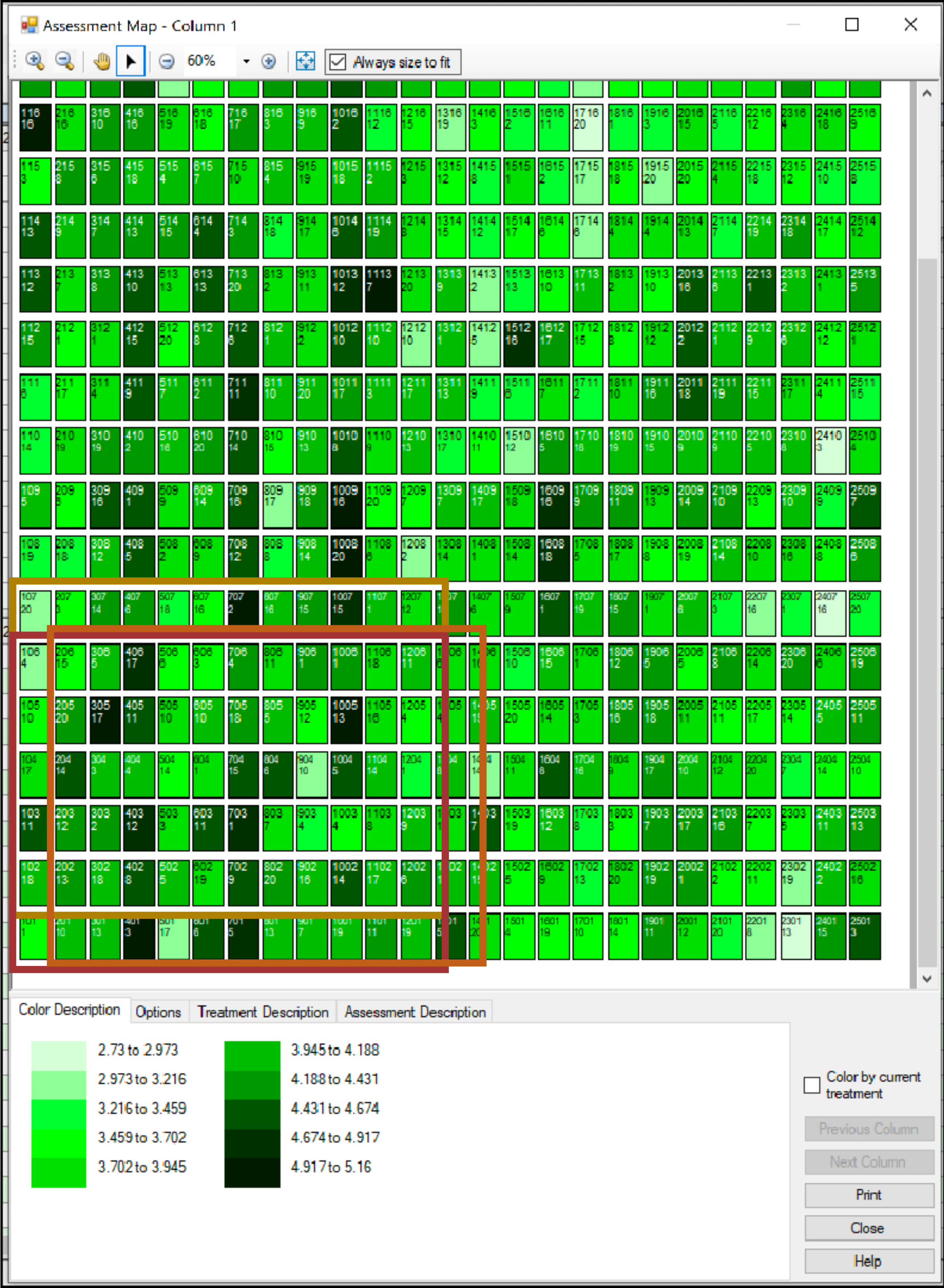


Randomized Complete Block (RCB) Least square estimation AOV For C 2:1 1 May-5-2021 (Data Column 4)

Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total	71	15.105788			
Replicate	5	0.619863	0.123973	0.560	0.7298
Treatment	11	2.317671	0.210697	0.952	0.4994
Error	55	12.168254	0.221241		

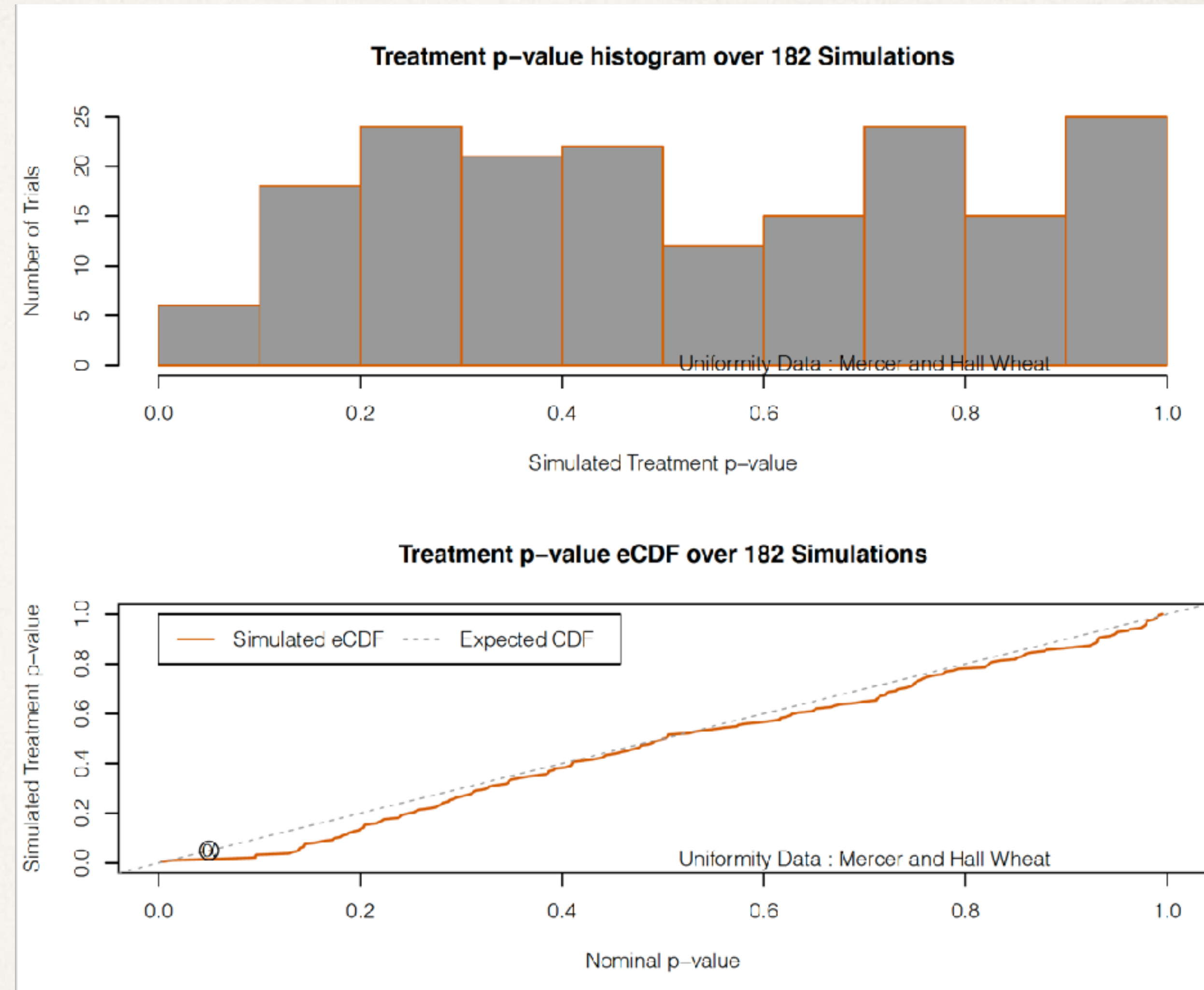
Simulated Trial (2:1)

Crop Type, Code	1:1		C - 1:2		C - 2:1	
Description	1		1		1	
Number of Subsamples	1		1		1	
Data Entry Date	Apr-20-2021		May-5-2021		May-5-2021	
Entry No.	1		2		3	
Entry Name	StDev		StDev		StDev	
1 Cilantro	4.113 -	0.645	4.177 abc	0.338	4.287 -	0.358
2 E chinacea	4.365 -	0.452	3.845 cd	0.285	4.365 -	0.336
3 Liatris punctata	4.142 -	0.405	4.447 a	0.363	3.922 -	0.355
4 Amorpha canescens	4.208 -	0.345	4.358 ab	0.631	3.938 -	0.250
5 Dalea purpurea	4.077 -	0.554	4.173 abc	0.208	4.062 -	0.653
6 Ratibida Columnifera	3.595 -	0.406	4.033 a-d	0.360	3.718 -	0.321
7 E chinacea purpurea	4.218 -	0.555	3.613 d	0.483	3.955 -	0.498
8 Helianthus sp	4.117 -	0.410	4.322 ab	0.445	4.035 -	0.527
9 Helianthus sp	4.280 -	0.392	4.180 abc	0.452	4.033 -	0.402
10 Solidago sp	3.898 -	0.186	4.195 abc	0.236	4.310 -	0.452
11 Glycyrrhiza lepidota	4.208 -	0.417	3.892 bcd	0.224	4.143 -	0.614
12 Onosmodium bejariense	4.022 -	0.381	4.172 abc	0.578	4.177 -	0.577
LSD P=.05	0.5226		0.4756		0.5442	
Standard Deviation	0.4517		0.4110		0.4704	
CV	11.01		9.98		11.53	
Grand Mean	4.1036		4.1172		4.0788	
Levene's F^	0.667		1.382		0.535	
Levene's Prob(F)	0.764		0.205		0.871	
Rank X2	.		.		.	
P(Rank X2)	.		.		.	
Skewness^	-0.2603		-0.0764		-0.1001	
Kurtosis^	-0.1663		-0.3224		-0.4554	
Replicate F	0.555		0.673		0.560	
Replicate Prob(F)	0.7337		0.6457		0.7298	
Treatment F	4.188		1.882		0.652	
Treatment Prob(F)	0.3166		0.0482		0.4994	



Visualization

- ❖ Histogram
 - ❖ Shows the frequency (or proportion) of observed values.
 - ❖ In this case, the observed values are p-values
- ❖ empirical Cumulative Distribution
 - ❖ Cumulative frequency (or proportion) of observed values.
 - ❖ Interpret this as the probability of an observed treatment p-value $<$ nominal p-value



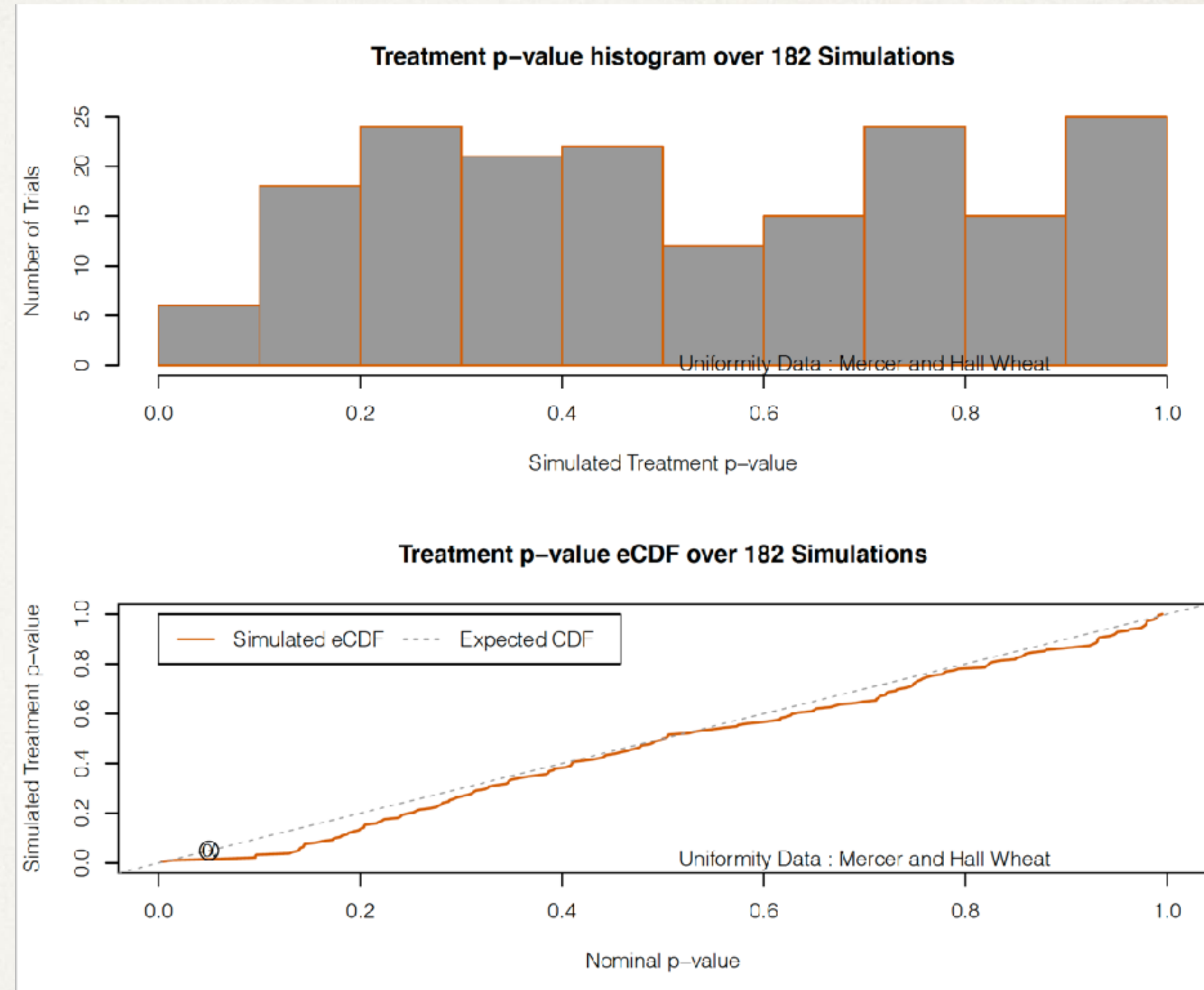
Visualization

- ❖ Histogram

- ❖ We see, over this uniformity data, fewer trials than expected that find significant treatment effects.

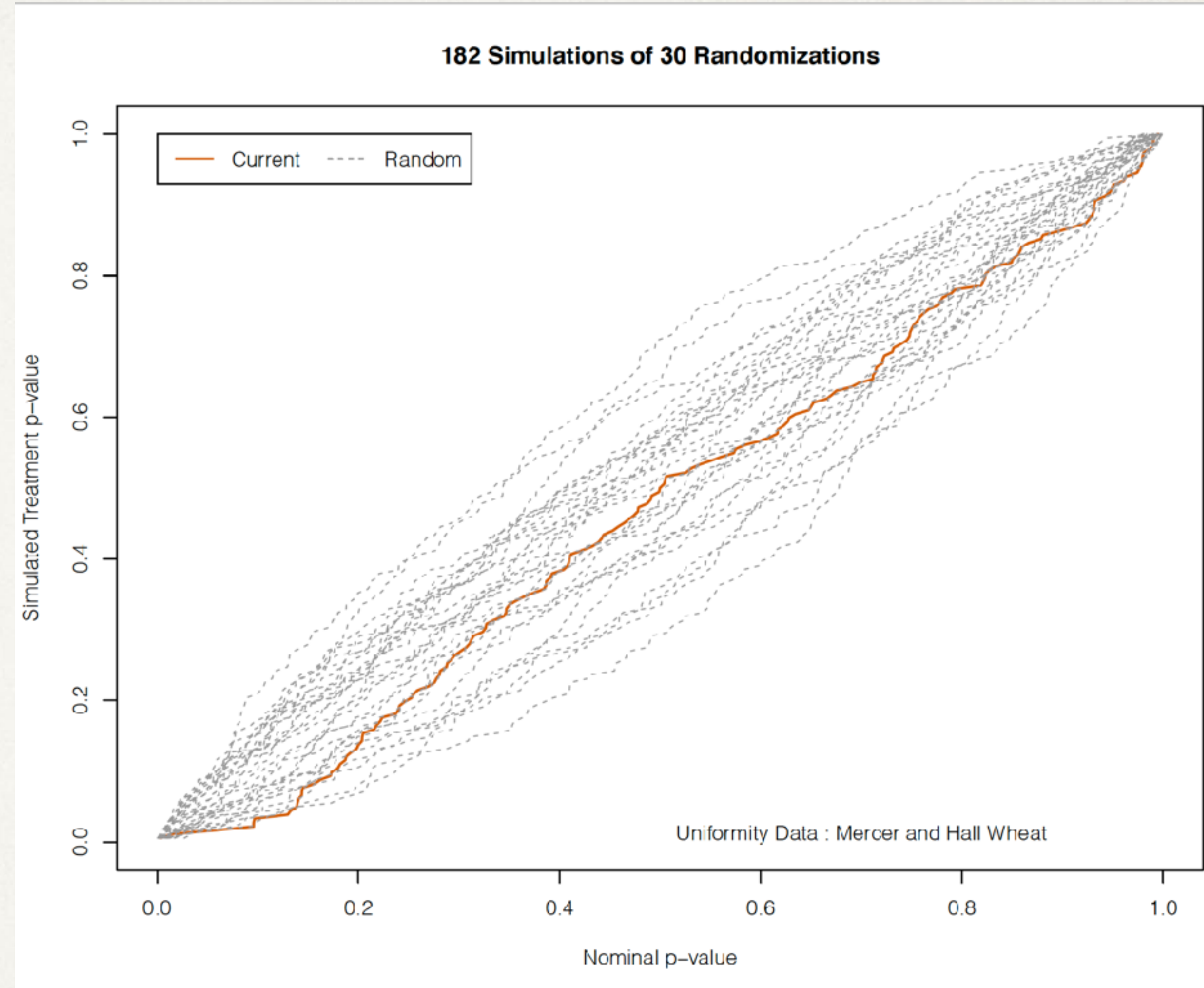
- ❖ empirical Cumulative Distribution

- ❖ Observed treatment p-values tend to be fewer than expected, for any critical value.



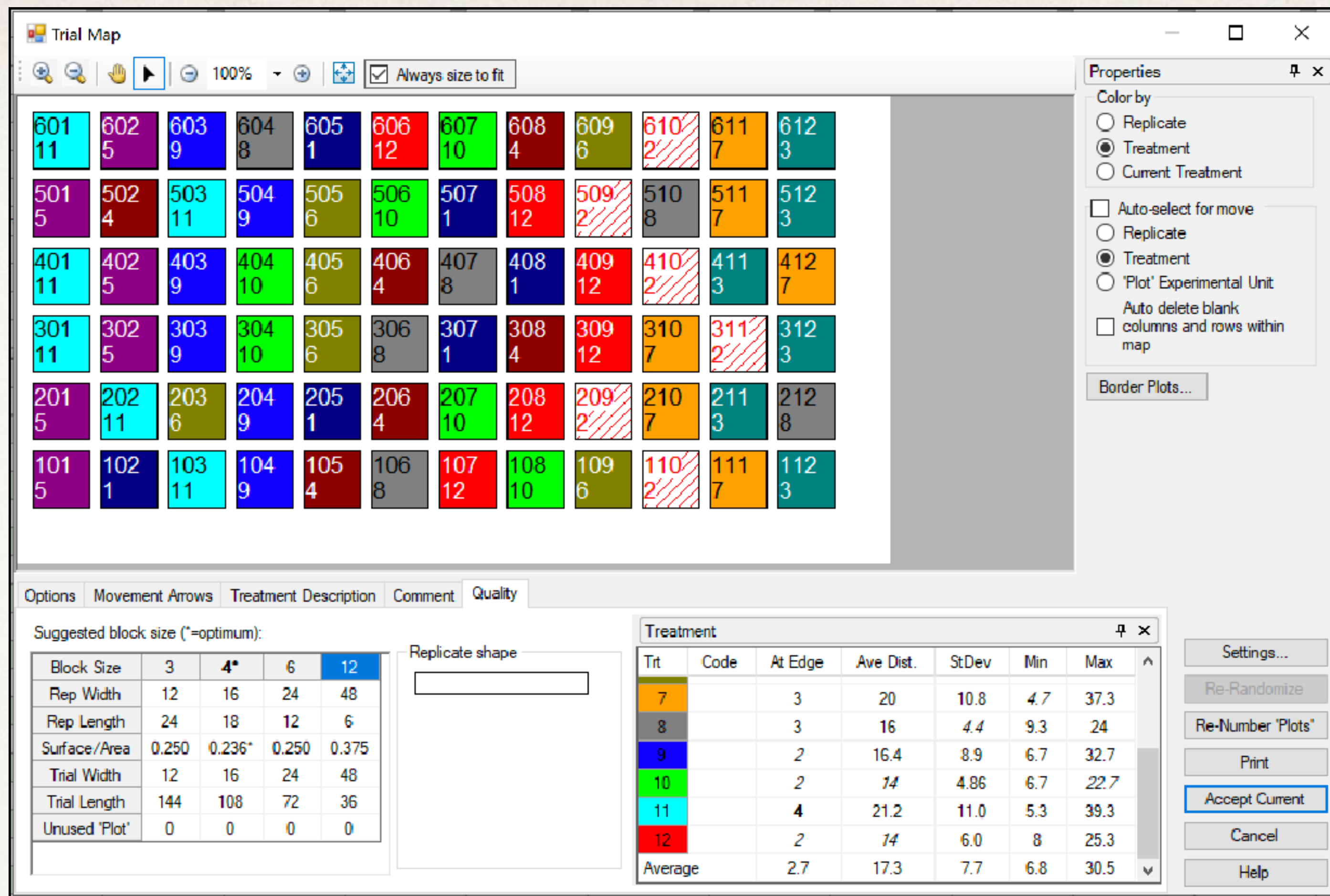
Comparison

- ❖ Repeat simulation with alternate randomizations.
- ❖ Some plans have more convex curves, suggesting better control of Type I error rates.
- ❖ Other trials have concave curves, suggesting higher rate of Type I error.



Type I Error

- ❖ Recall, we are inferring plot yields from uniformity trial data.
- ❖ There are no treatments applied in uniformity trials; these trials are established to study spatial variation under agricultural trial conditions
- ❖ The null hypothesis is, a priori, true for all simulated trials. We expect to fail to reject the null hypothesis at predefined nominal rate (i.e. 5%).
- ❖ For some randomizations, we fail to reject the null hypothesis in more trials than expected; thus, we achieve a greater Type I error rate than our nominal rate.
- ❖ What increases Type I error?

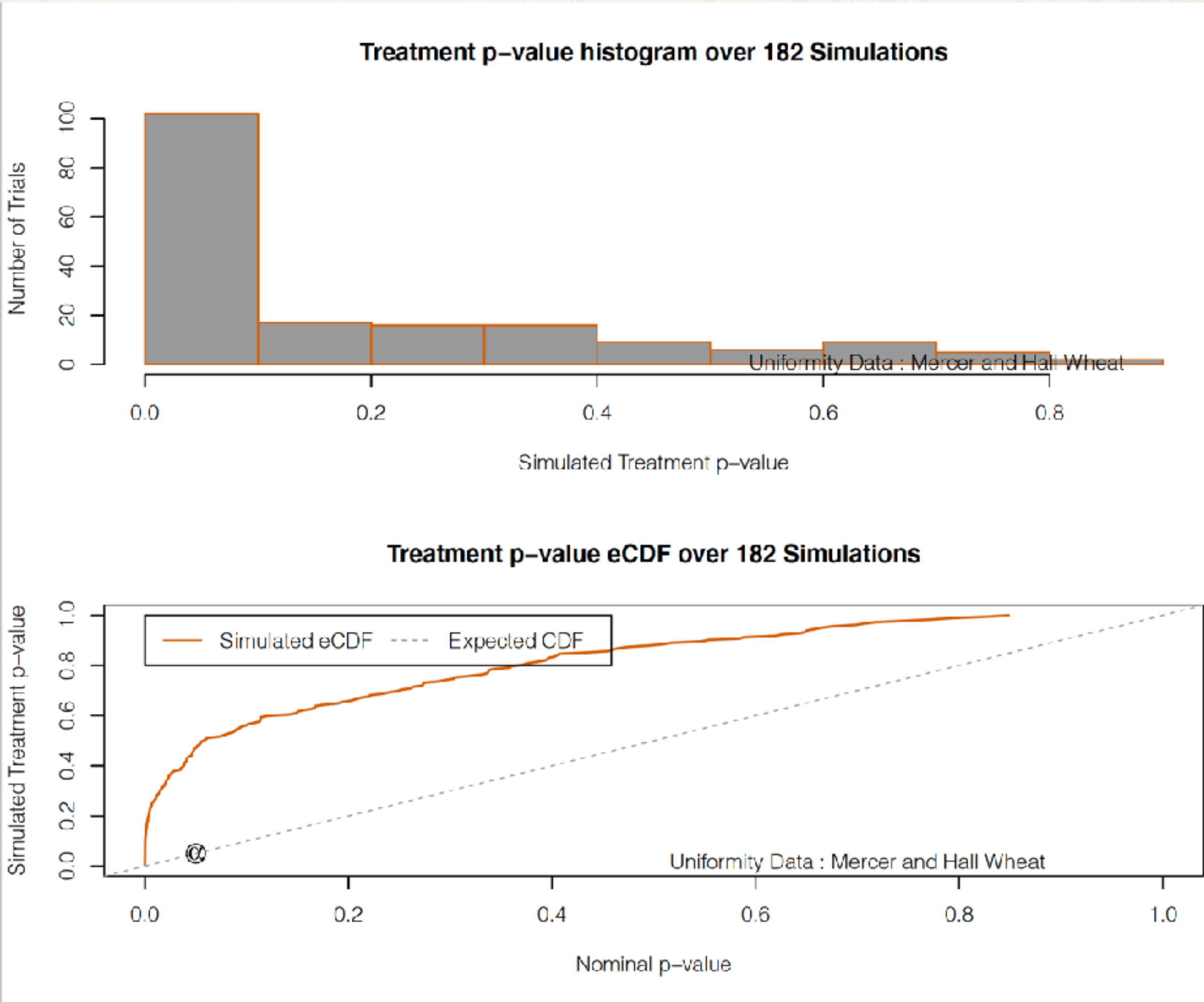
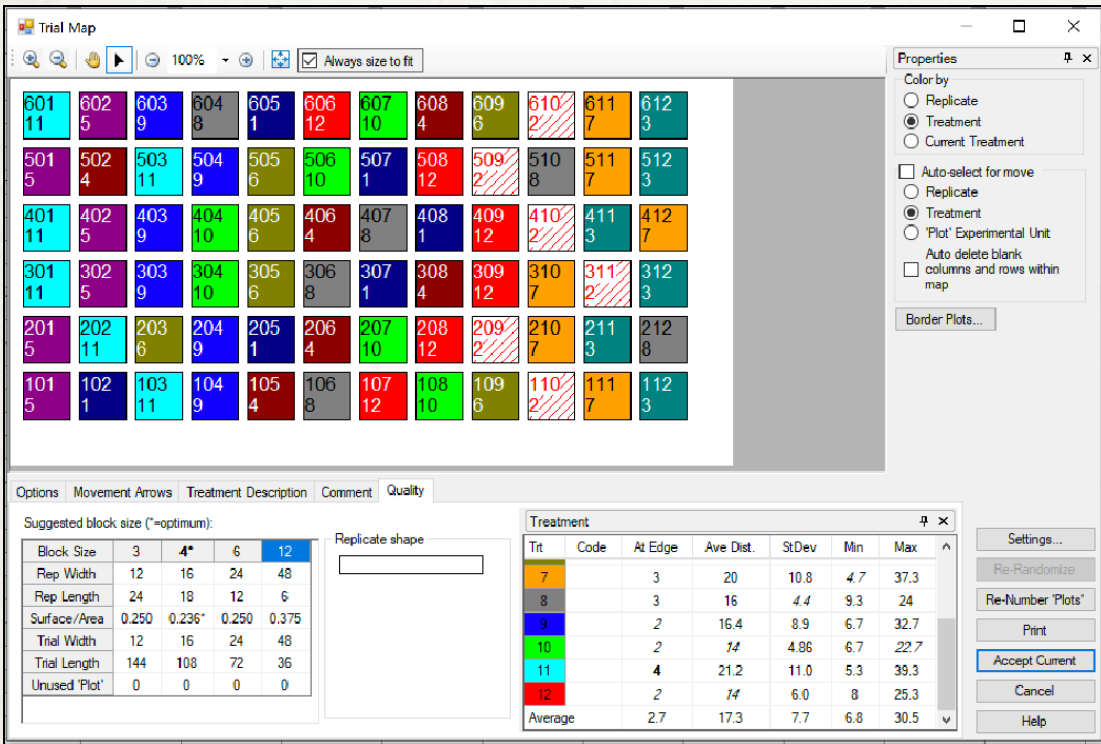


Undesirable Randomization

Trial was manually editing to group treatments in columns.

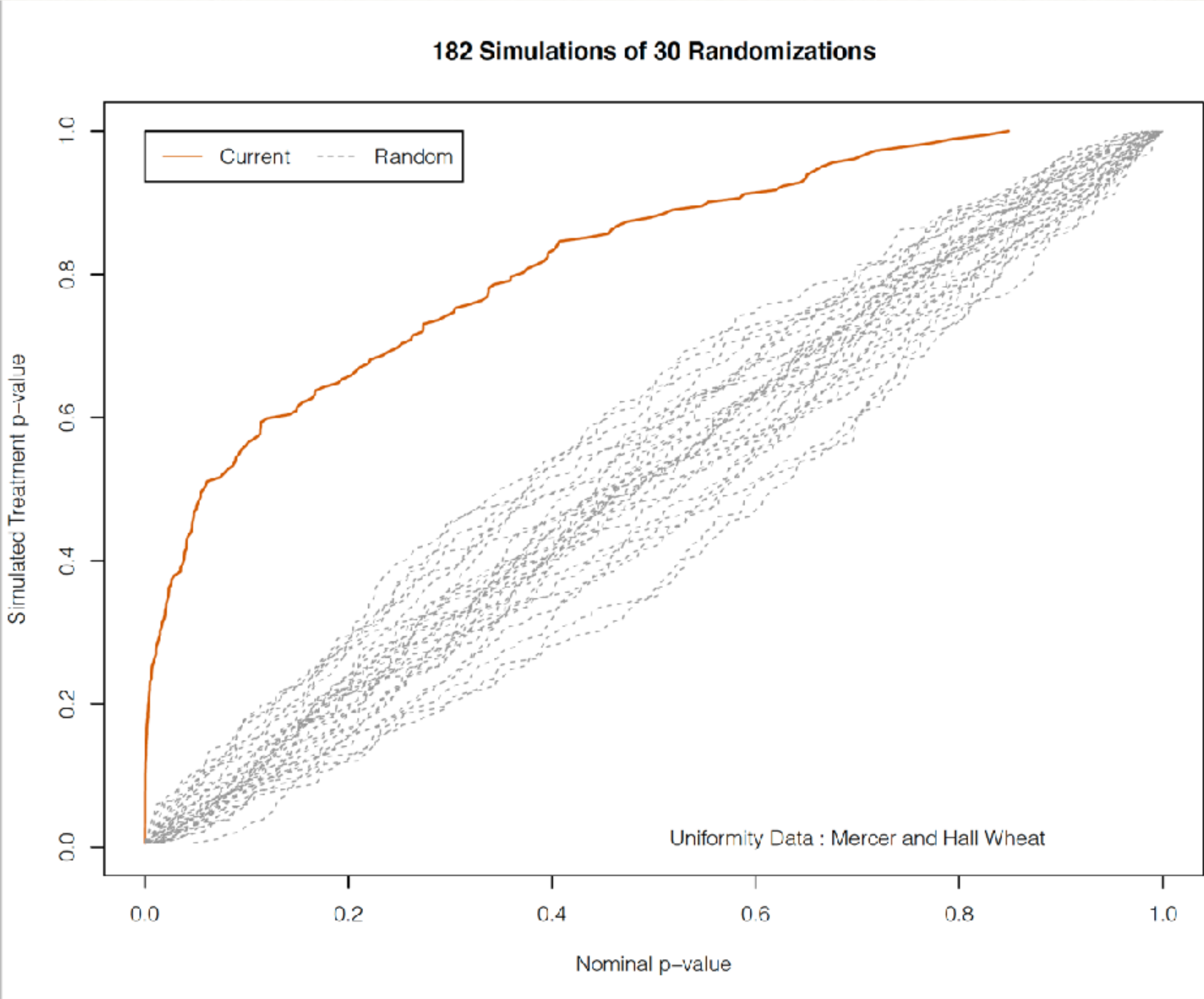
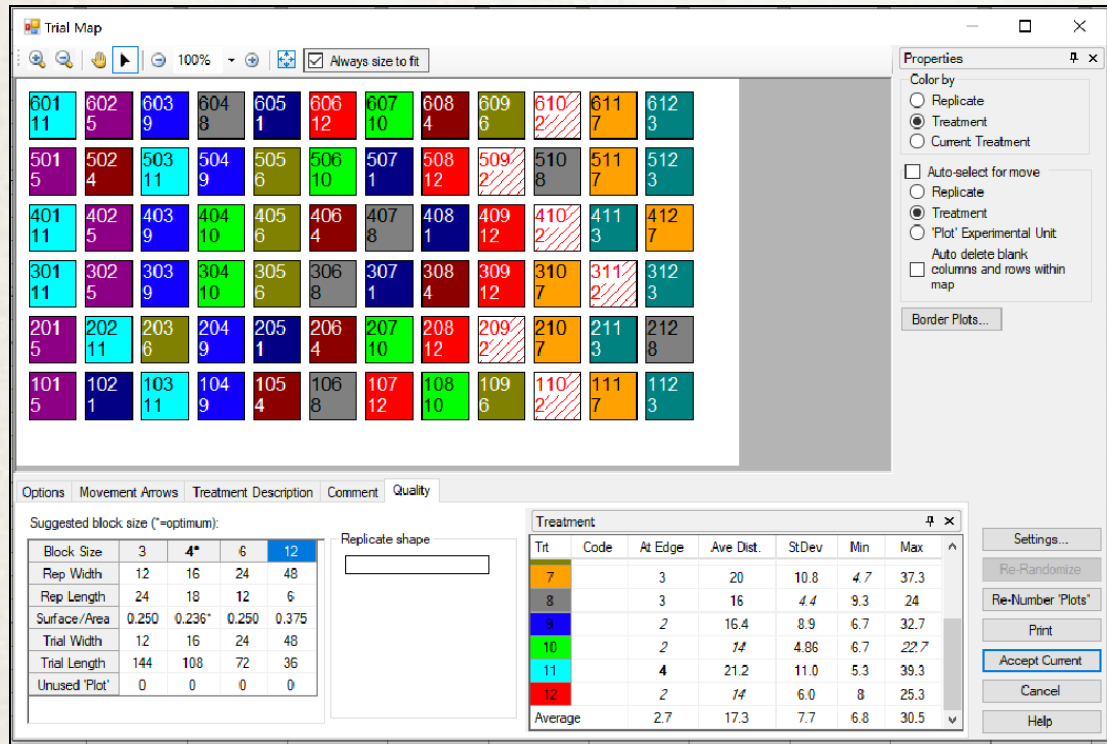
Undesirable Randomization

- ❖ Treatments aligned in columns are more likely to be confounded with spatial trends.
- ❖ If spatial variation is significant across columns, then we expect this to suggest significant treatment effects.



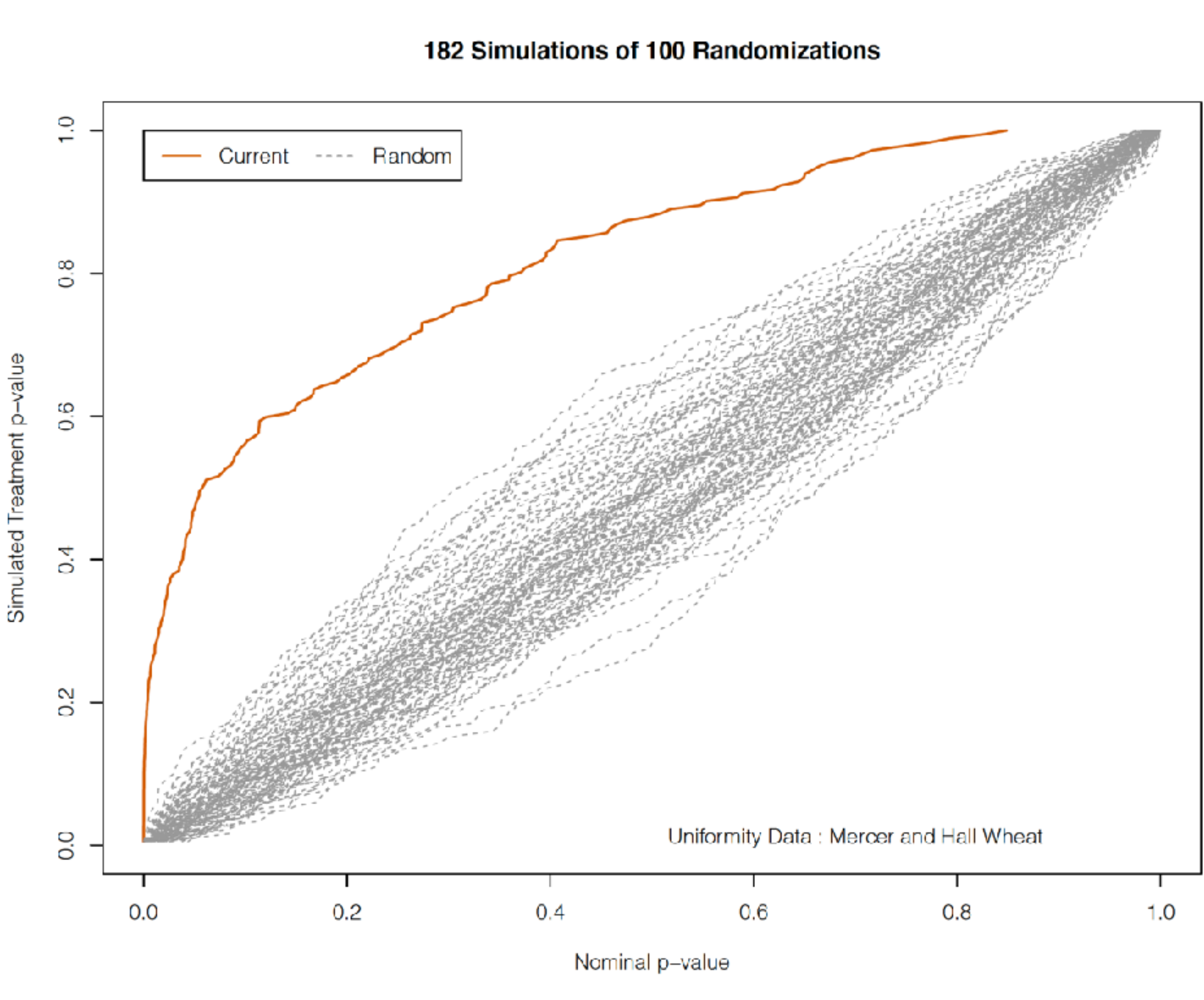
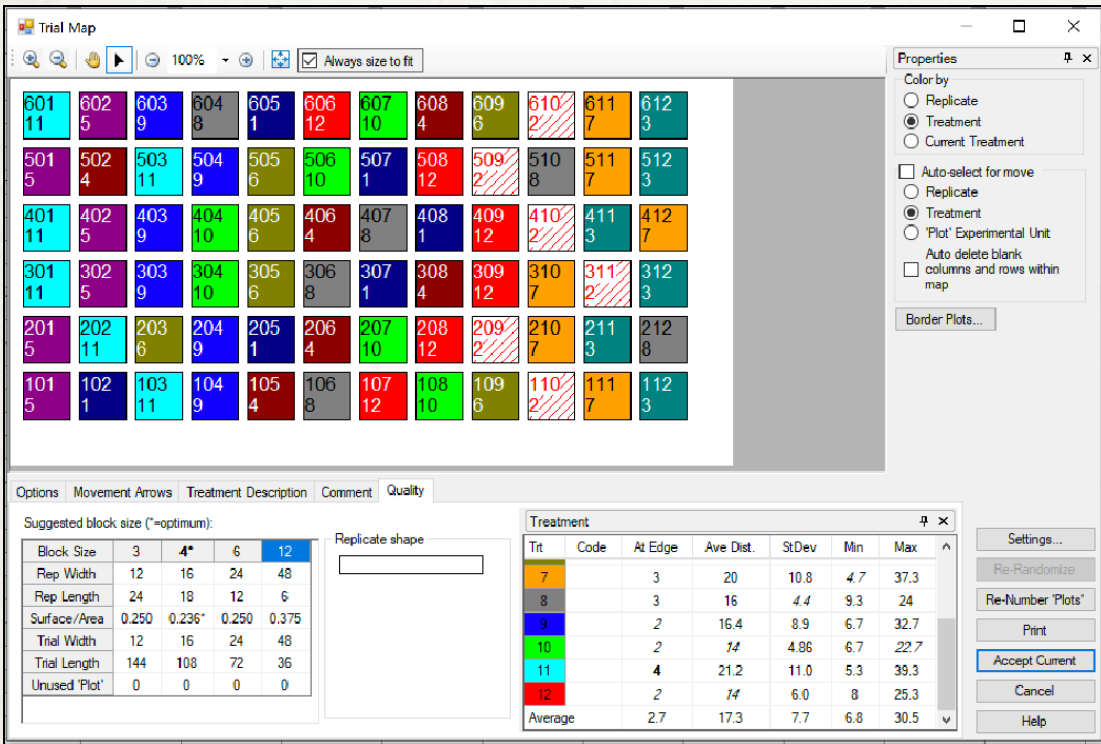
Undesirable Randomization

- ❖ While this is a valid randomization, it is an unlikely randomization



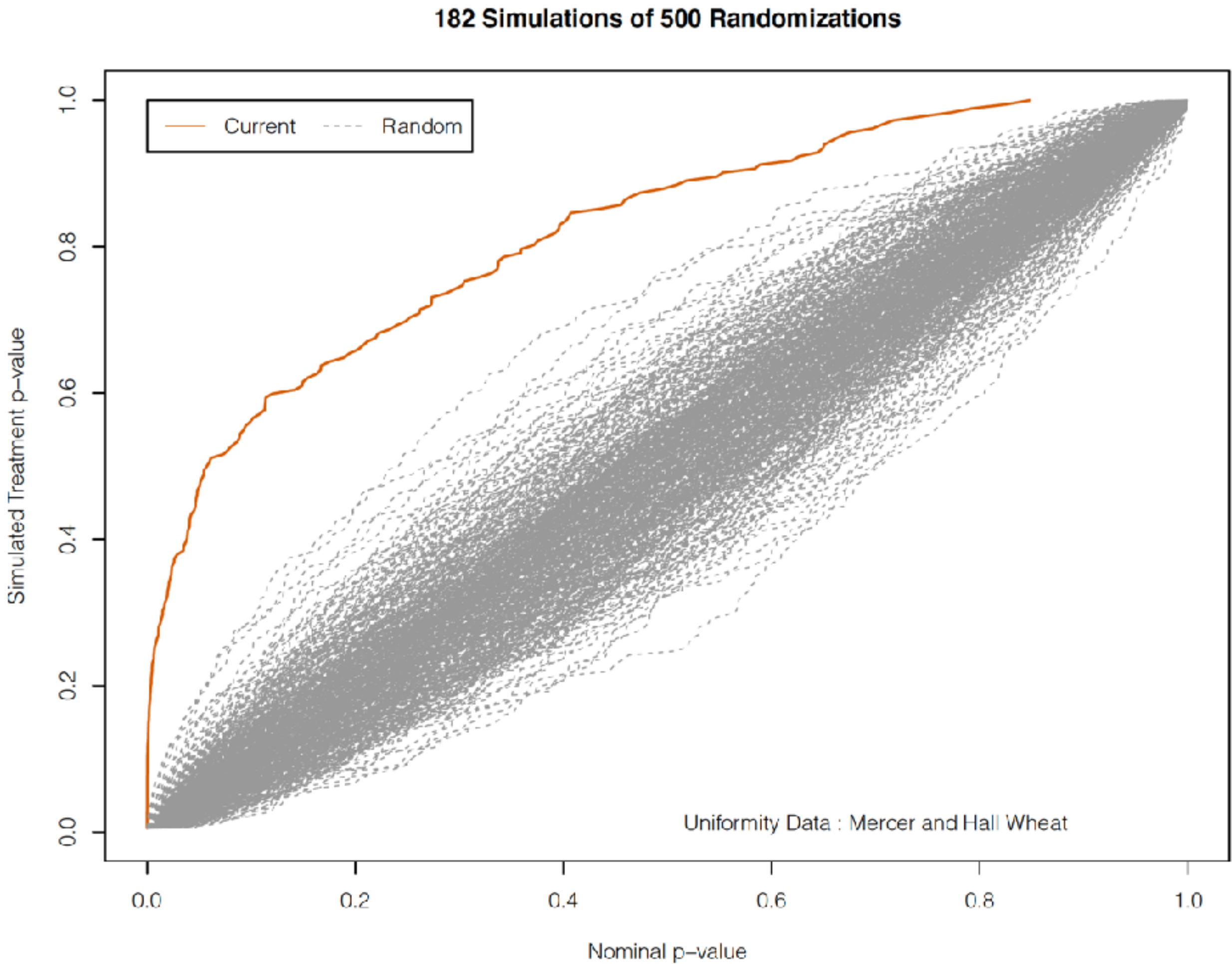
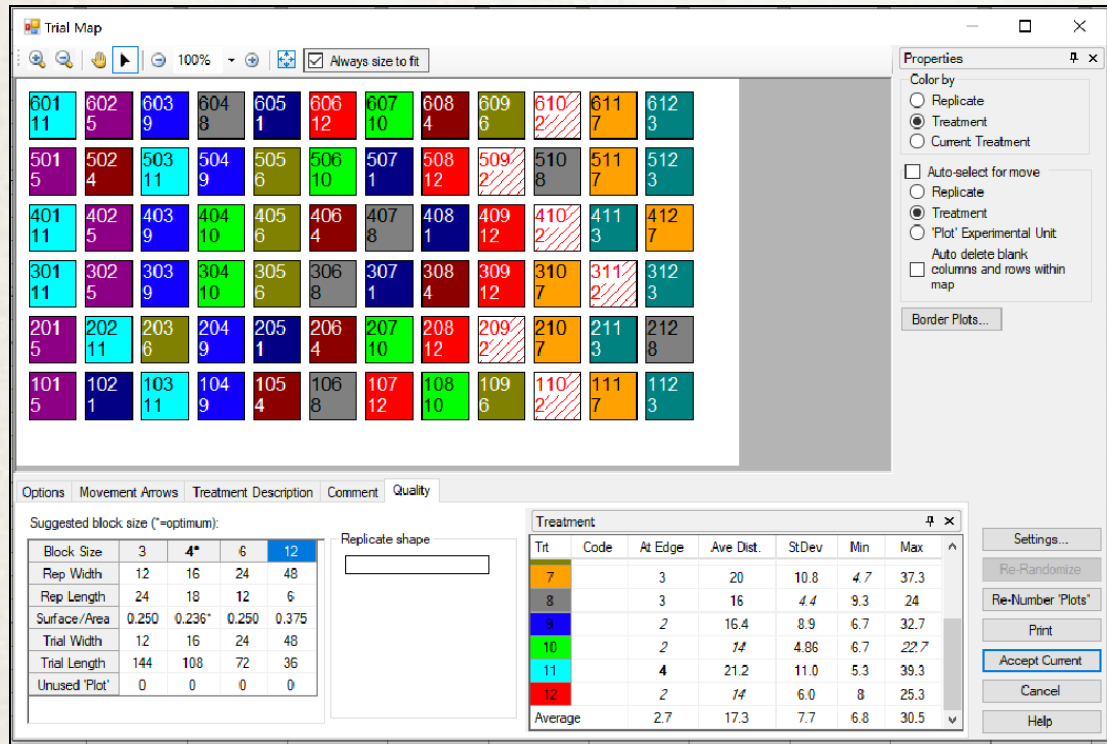
Undesirable Randomization

- ❖ While this is a valid randomization, it is a very unlikely randomization

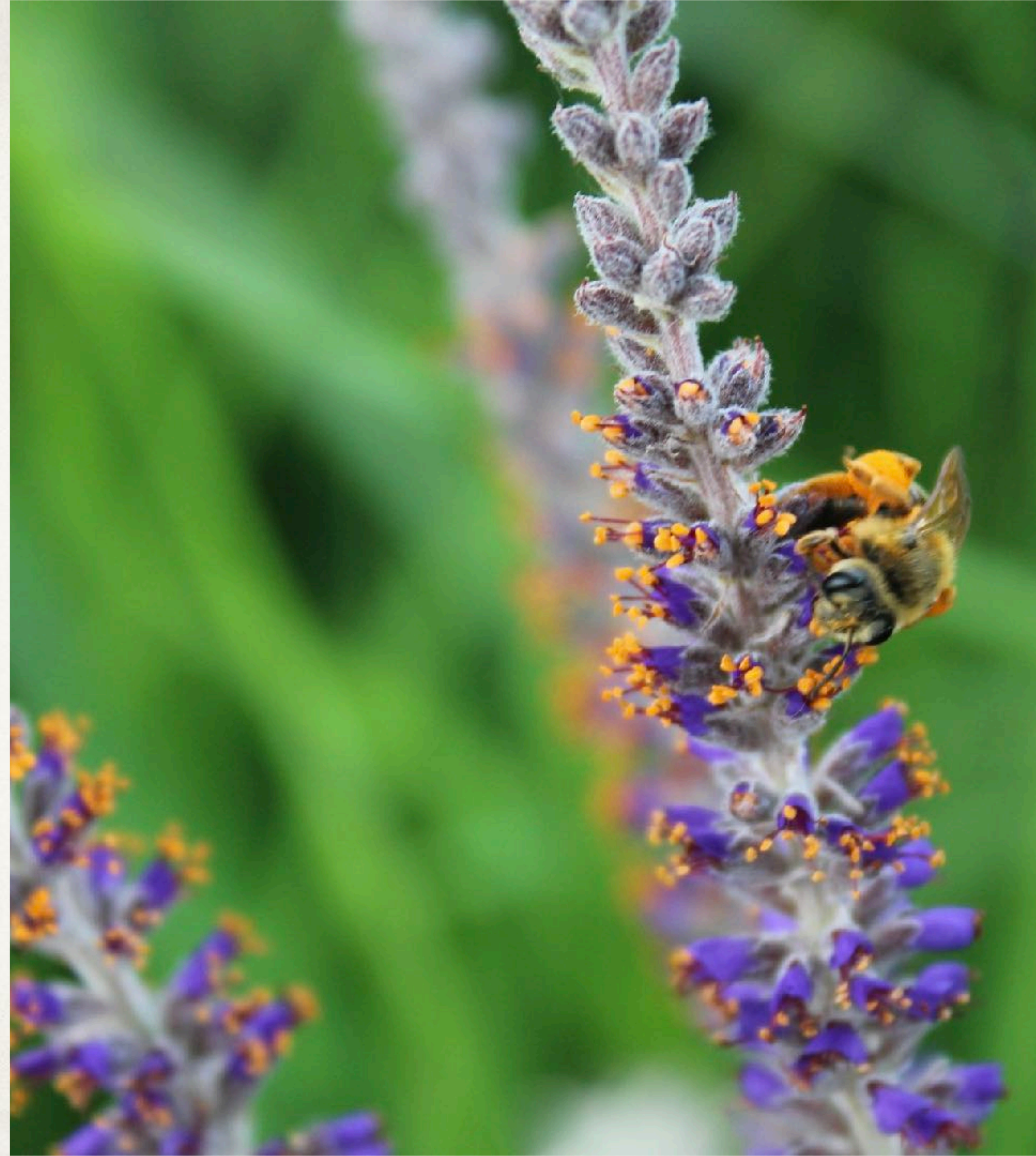


Undesirable Randomization

- ❖ While this is a valid randomization, it is an extremely unlikely randomization



Back to Block Dimensions



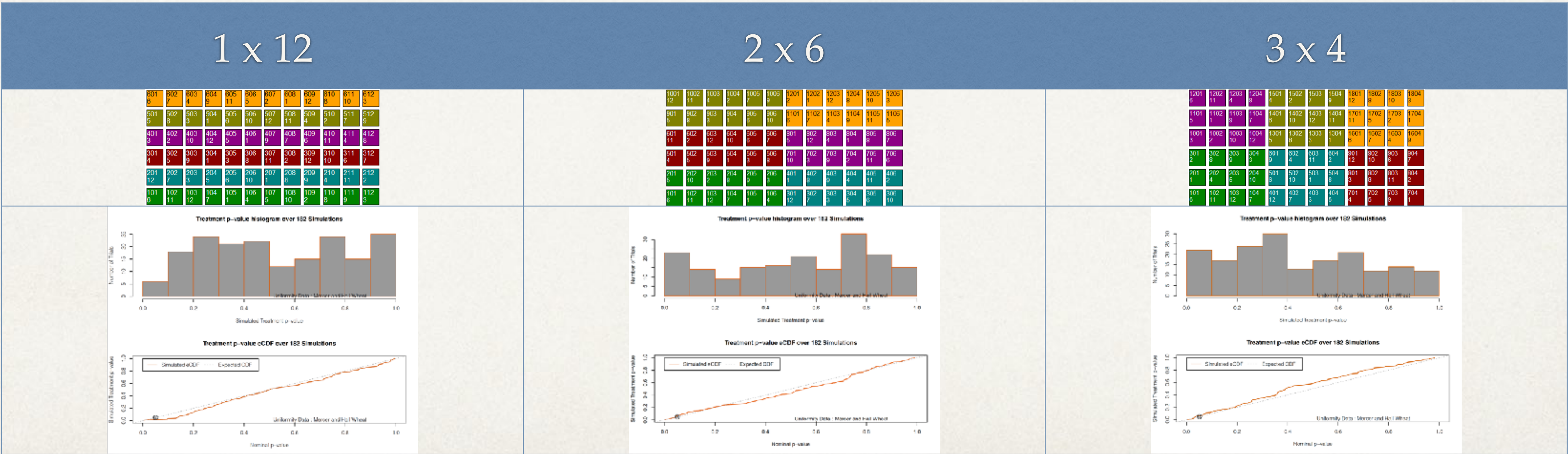
Back to Block Dimensions

- ❖ We can use uniformity simulations to understand how replicate shape may influence Type I error rates
- ❖ We compare three block shapes:

1 x 12												2 x 6												3 x 4															
601 6	602 7	603 4	604 9	605 11	606 5	607 2	608 1	609 12	610 8	611 10	612 3	1001 12	1002 11	1003 4	1004 2	1005 7	1006 9	1201 2	1202 1	1203 12	1204 8	1205 10	1206 3	1201 6	1202 11	1203 4	1204 8	1501 4	1502 2	1503 7	1504 9	1801 12	1802 8	1803 10	1804 3				
501 5	502 8	503 3	504 1	505 6	506 10	507 12	508 11	509 4	510 2	511 7	512 9	901 5	902 8	903 3	904 1	905 6	906 10	1101 6	1102 7	1103 4	1104 9	1105 11	1106 5	1101 5	1102 1	1103 9	1104 7	1401 6	1402 10	1403 12	1404 11	1701 11	1702 5	1703 2	1704 1				
401 3	402 2	403 10	404 12	405 5	406 1	407 9	408 7	409 6	410 11	411 4	412 8	601 11	602 2	603 12	604 10	605 6	606 7	801 5	802 12	803 4	804 1	805 8	806 7	1001 3	1002 2	1003 10	1004 12	1301 5	1302 8	1303 3	1304 1	1601 6	1602 7	1603 4	1604 9				
301 4	302 5	303 9	304 1	305 3	306 8	307 11	308 2	309 12	310 10	311 6	312 7	501 4	502 5	503 9	504 1	505 3	506 8	701 10	702 3	703 9	704 2	705 11	706 6	301 2	302 8	303 9	304 3	601 9	602 4	603 11	604 2	901 12	902 10	903 6	904 7				
201 12	202 7	203 3	204 5	205 6	206 10	207 1	208 8	209 9	210 4	211 11	212 2	201 5	202 10	203 2	204 8	205 9	206 3	401 1	402 8	403 9	404 4	405 11	406 2	201 1	202 4	203 5	204 10	501 6	502 10	503 1	504 8	801 3	802 8	803 11	804 2				
101 6	102 11	103 12	104 7	105 1	106 4	107 5	108 10	109 2	110 8	111 9	112 3	101 6	102 11	103 12	104 7	105 1	106 4	301 12	302 7	303 3	304 5	305 6	306 10	101 6	102 11	103 12	104 7	401 12	402 7	403 3	404 5	701 4	702 5	703 9	704 1				

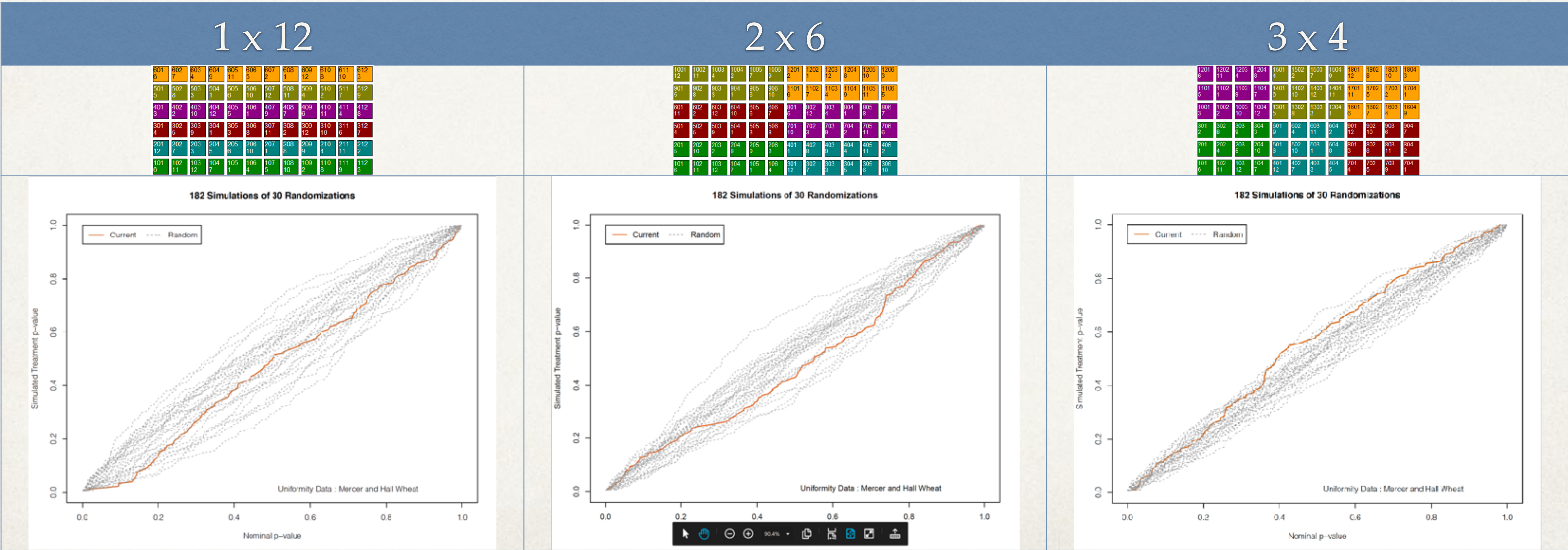
Block Dimensions

- ❖ It will not be particularly informative to consider the distribution of simulated trials for any single randomization



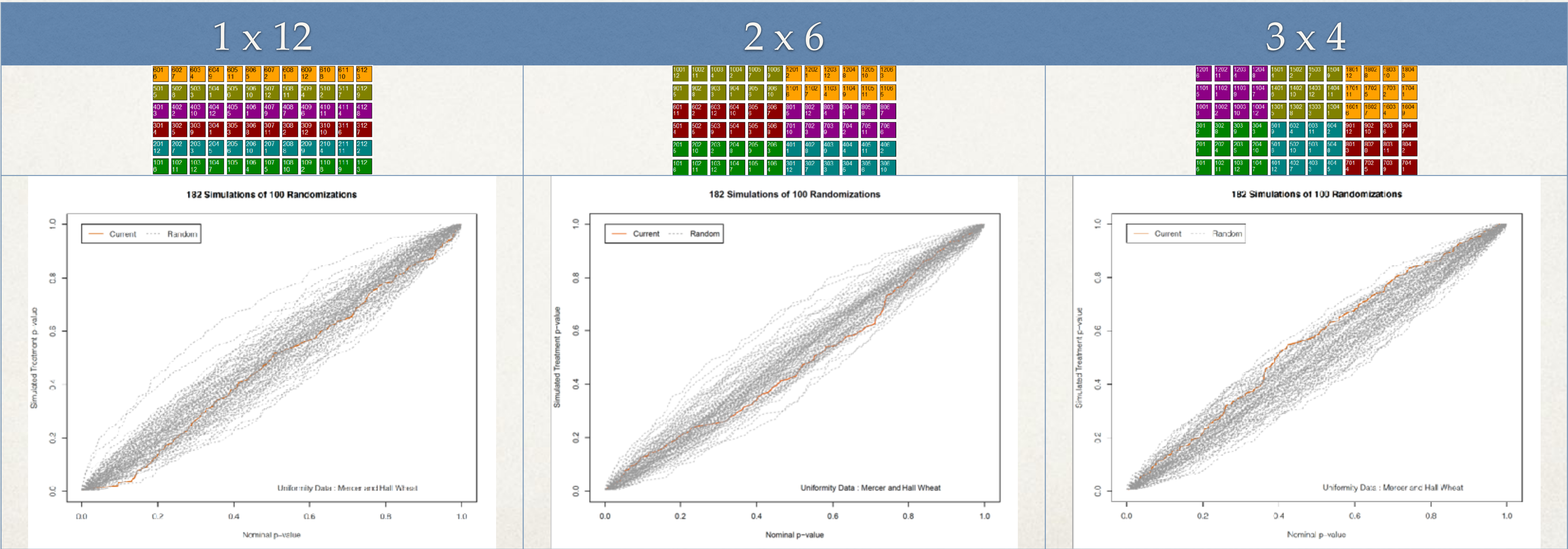
Block Dimensions

❖ But we can consider the quality of multiple (30) randomizations



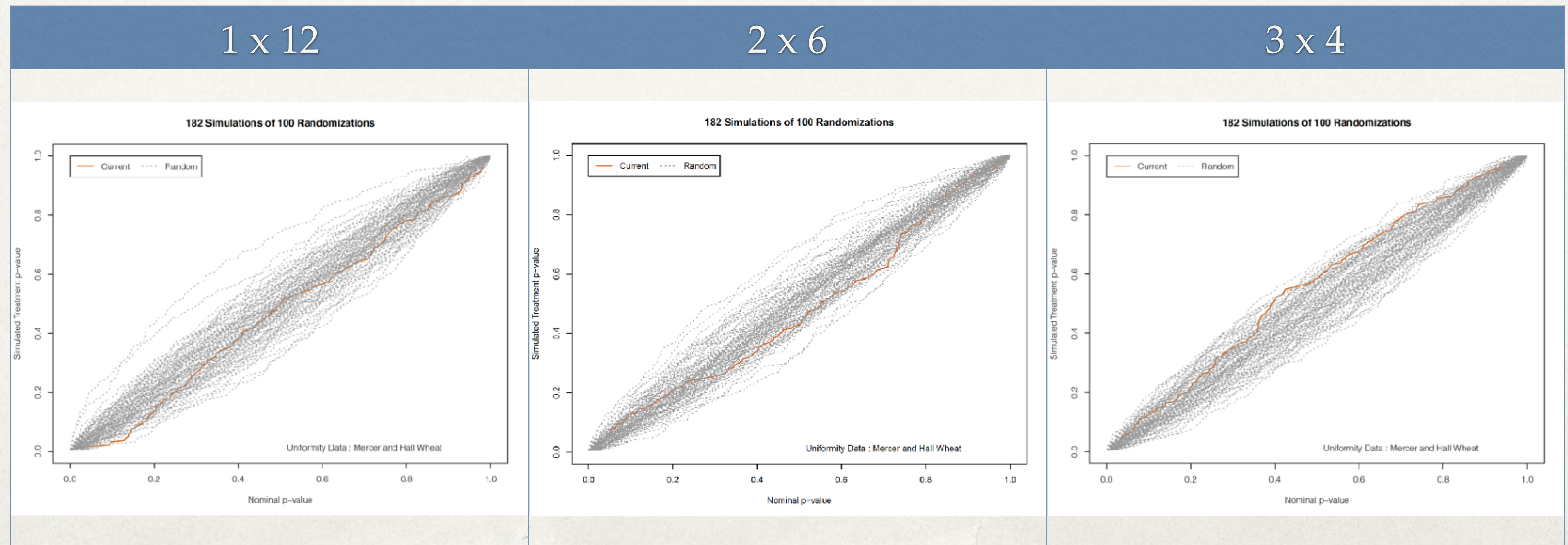
Block Dimensions

❖ But we can consider the quality of multiple (100) randomizations



Block Dimensions

- ❖ More uniform block dimensions lead to more uniform Type I error rates across possible randomizations.



What about power?



What about power?

- ❖ Recall that we can have two types of error of inference:
- ❖ Type I error (false positive) occurs when we reject a true null hypothesis
- ❖ But we also want to control for Type II error (false negative), when we reject fail to reject then null hypothesis when there is a true treatment effect.
- ❖ We tend to focus on the former, but we should not ignore that latter when designing an experiment.

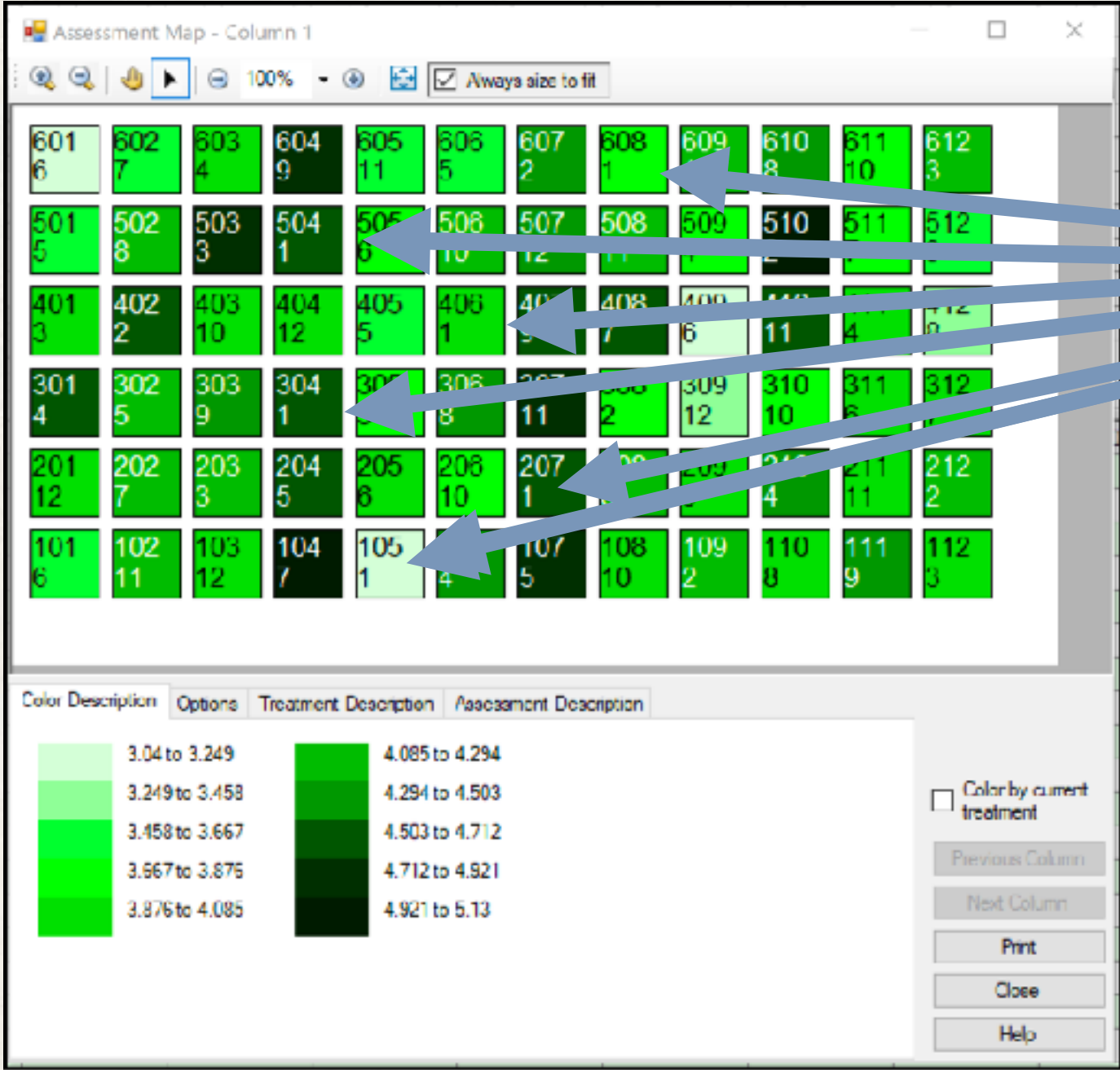
What about power?

- ❖ Planning experiments to achieve a desired power (say, 80%) is supported in ARM, but would be a topic for another day.
- ❖ Instead, we consider the implications of a randomization with greater or lesser simulated Type I on the corresponding Type II error.
- ❖ That is, we can also simulate the proportion of trials that detect true treatment differences added to uniformity data.

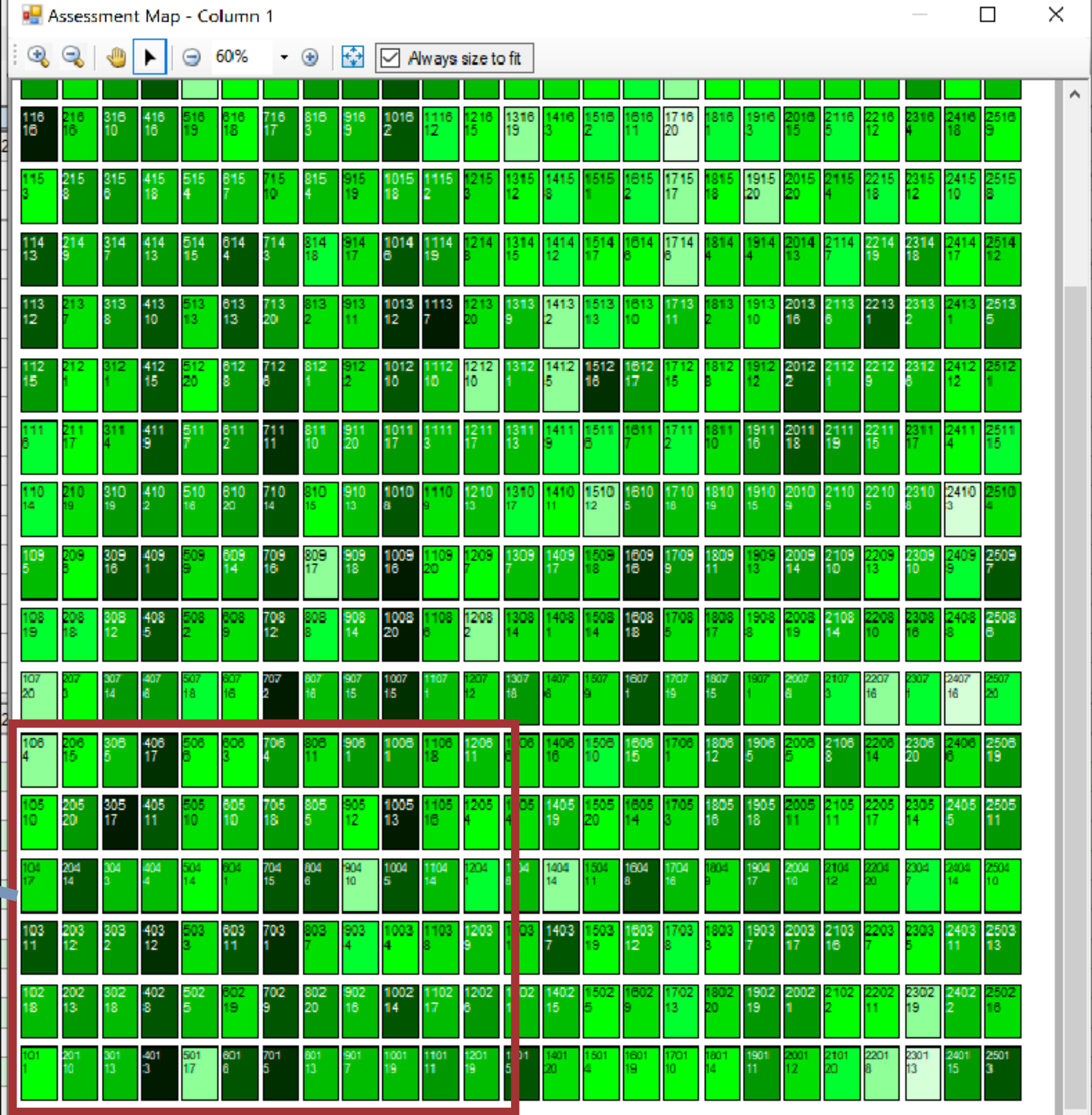
Power Simulations

- ❖ Start with a uniformity trial (from many available in the literature)
- ❖ Copy plot assessments to a proposed randomization
- ❖ Select treatment and add a percentage (Percent Mean Difference) to the plot assessments, from the prior test, for that treatment only
- ❖ Analyze as if this were an actual trial; continue over entire uniformity data as before.
- ❖ Increase Percent Mean Difference and repeat
- ❖ Do this for each treatment in turn

Simulated Trial



+ 3% (Trial Mean)



Randomized Complete Block (RCB) Least square estimation AOV For 1:1 1 Apr-20-2021 (Data Column 1)					
Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total	71	14.454461			
Replicate	5	0.566411	0.113282	0.555	0.7337
Treatment	11	2.667061	0.242460	1.188	0.3166
Error	55	11.220989	0.204018		

Randomized Complete Block (RCB) Least square estimation AOV For C 1:1+3% 1 May-4-2021 (Data Column 2)					
Source	DF	Sum of Squares	Mean Square	F	Prob(F)
Total	71	14.552179			
Replicate	5	0.566411	0.113282	0.555	0.7337
Treatment	11	2.764779	0.251344	1.232	0.2888
Error	55	11.220989	0.204018		

Add a percent ...

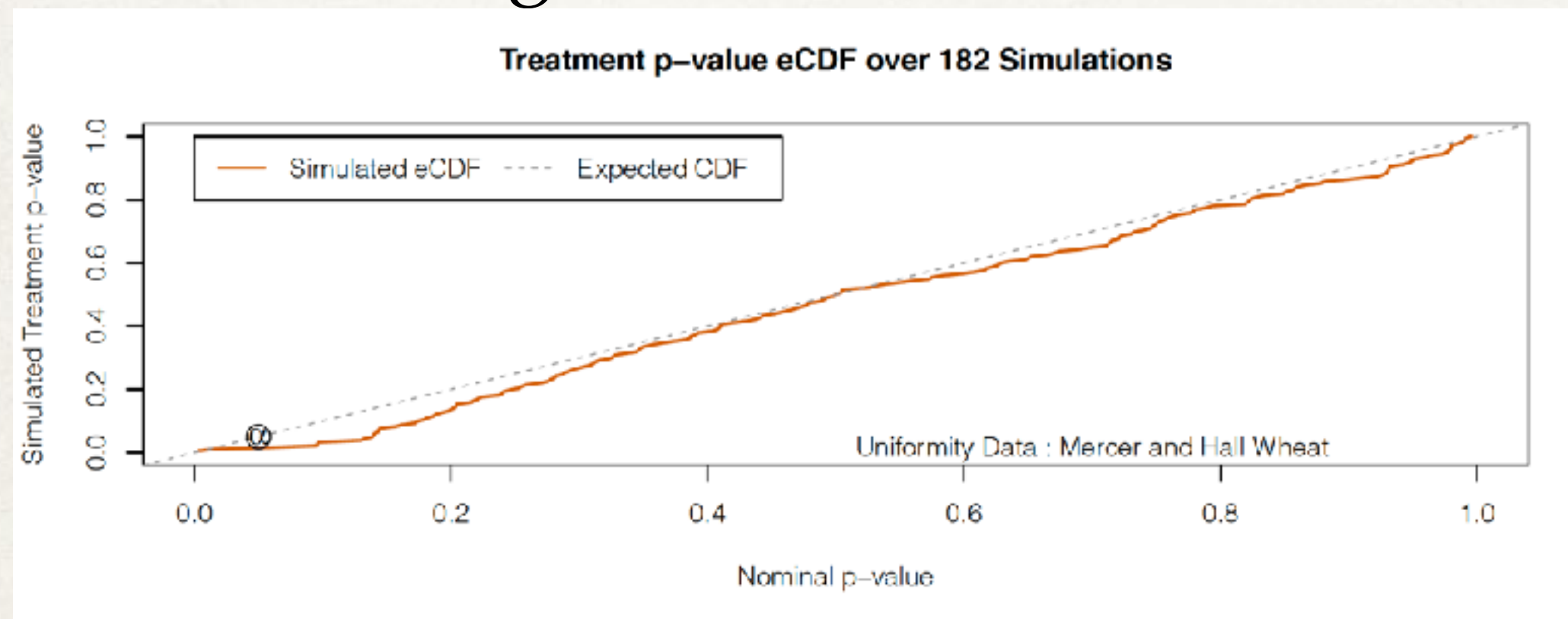
[illegible]

... and analyze

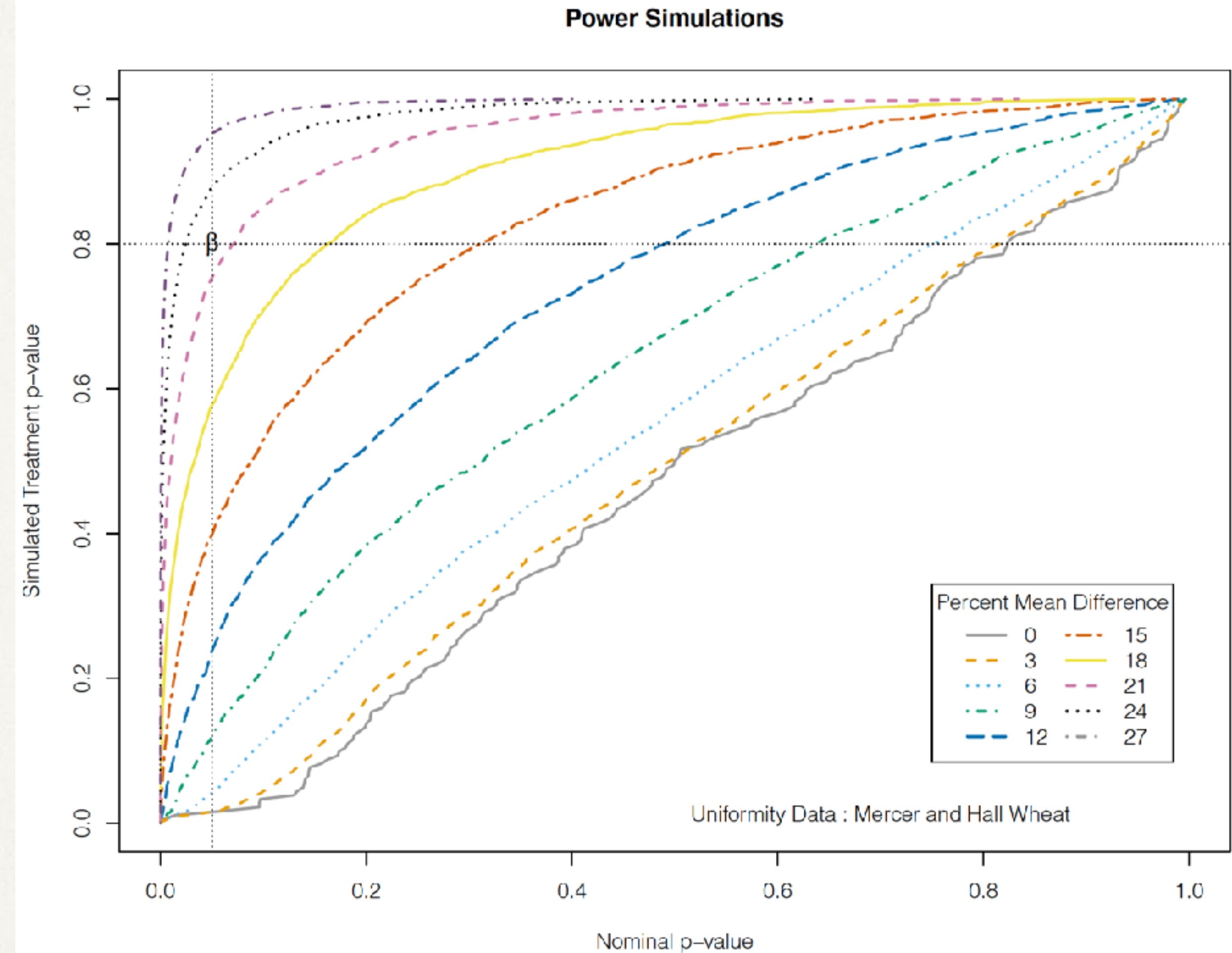
Crop Type, Code Description	1:1	C - 1:1+3%	C - 1:1+6%	C - 1:1+9%	C - 1:1+12%	C - 1:1+15%	C - 1:1+18%	C - 1:1+21%
Number of Subsamples	1	1	1	1	1	1	1	1
Data Entry Date	Apr-20-2021	May-5-2021	May-5-2021	May-5-2021	May-5-2021	May-5-2021	May-5-2021	May-5-2021
Entry No.	1	4	5	6	7	8	9	10
Entry Name								
1 Cilantro	4.113 -	4.236 -	4.360 -	4.483 -	4.606 -	4.729 a	4.852 a	4.975 a
2 Echinacea	4.365 -	4.365 -	4.365 -	4.365 -	4.365 -	4.365 ab	4.365 ab	4.365 b
3 Liatris punctata	4.142 -	4.142 -	4.142 -	4.142 -	4.142 -	4.142 b	4.142 b	4.142 b
4 Amorpha canescens	4.208 -	4.208 -	4.208 -	4.208 -	4.208 -	4.208 ab	4.208 b	4.208 b
5 Dalea purpurea	4.077 -	4.077 -	4.077 -	4.077 -	4.077 -	4.077 bc	4.077 bc	4.077 bc
6 Ratibida Columnifera	3.595 -	3.595 -	3.595 -	3.595 -	3.595 -	3.595 c	3.595 c	3.595 c
7 Echinacea purpurea	4.218 -	4.218 -	4.218 -	4.218 -	4.218 -	4.218 ab	4.218 b	4.218 b
8 Helianthus sp	4.117 -	4.117 -	4.117 -	4.117 -	4.117 -	4.117 bc	4.117 bc	4.117 bc
9 Helianthus sp	4.280 -	4.280 -	4.280 -	4.280 -	4.280 -	4.280 ab	4.280 b	4.280 b
10 Solidago sp	3.898 -	3.898 -	3.898 -	3.898 -	3.898 -	3.898 bc	3.898 bc	3.898 bc
11 Glycyrrhiza lepidota	4.208 -	4.208 -	4.208 -	4.208 -	4.208 -	4.208 ab	4.208 b	4.208 b
12 Onosmodium bejariense	4.022 -	4.022 -	4.022 -	4.022 -	4.022 -	4.022 bc	4.022 bc	4.022 bc
LSD P=.05	0.5226	0.5226	0.5226	0.5226	0.5226	0.5226	0.5226	0.5226
Standard Deviation	0.4517	0.4517	0.4517	0.4517	0.4517	0.4517	0.4517	0.4517
CV	11.01	10.98	10.95	10.93	10.9	10.87	10.84	10.82
Grand Mean	4.1038	4.1139	4.1241	4.1344	4.1446	4.1549	4.1652	4.1754
Levene's F^	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667
Levene's Prob(F)	0.764	0.764	0.764	0.764	0.764	0.764	0.764	0.764
Rank X2
P(Rank X2)
Skewness^	-0.2603	-0.2603	-0.2603	-0.2603	-0.2603	-0.2603	-0.2603	-0.2603
Kurtosis^	-0.1663	-0.1663	-0.1663	-0.1663	-0.1663	-0.1663	-0.1663	-0.1663
Replicate F	0.555	0.555	0.555	0.555	0.555	0.555	0.555	0.555
Replicate Prob(F)	0.7337	0.7337	0.7337	0.7337	0.7337	0.7337	0.7337	0.7337
Treatment F	1.188	1.232	1.350	1.542	1.808	2.149	2.564	3.053
Treatment Prob(F)	0.3166	0.2888	0.2233	0.1432	0.0747	0.0314	0.0107	0.0030

A picture is worth a thousand words.

- ❖ Recall the original simulation distribution:

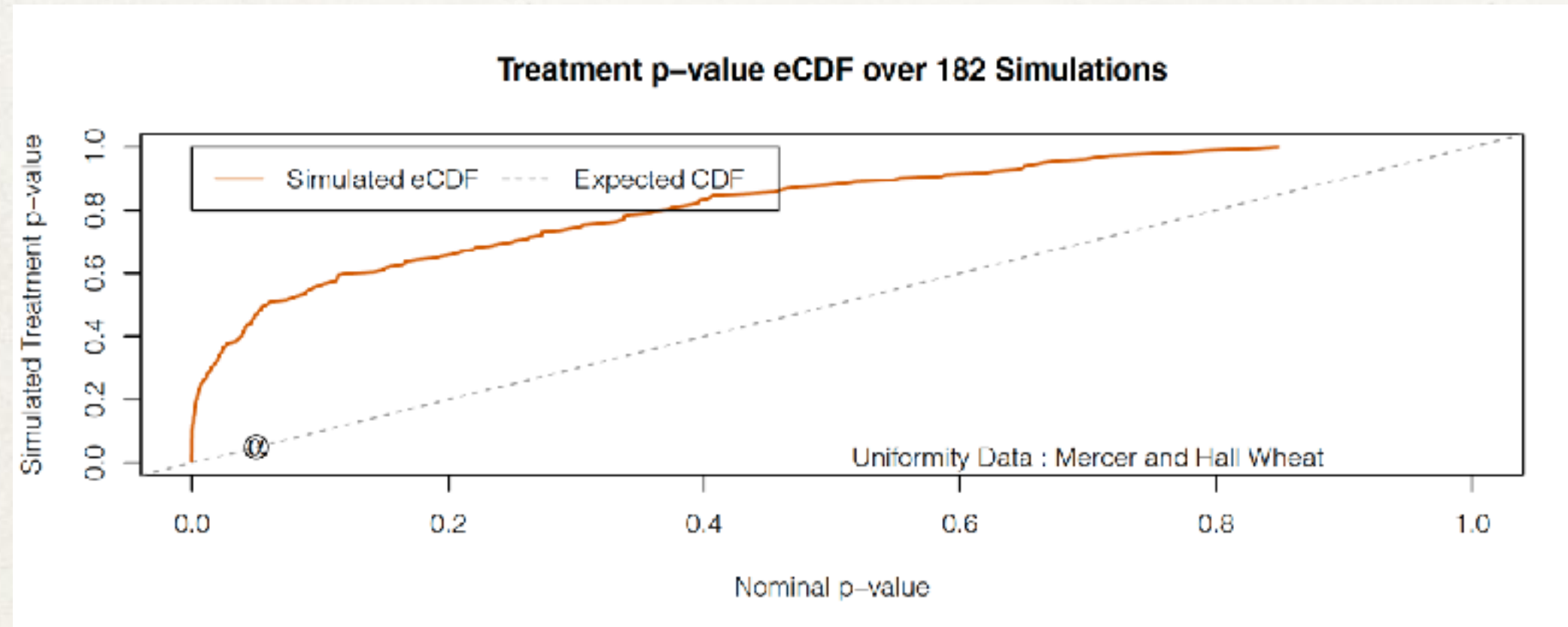


- ❖ and compare to a series of trials with an added treatment effect ranging from 3-27 Percent Mean Difference

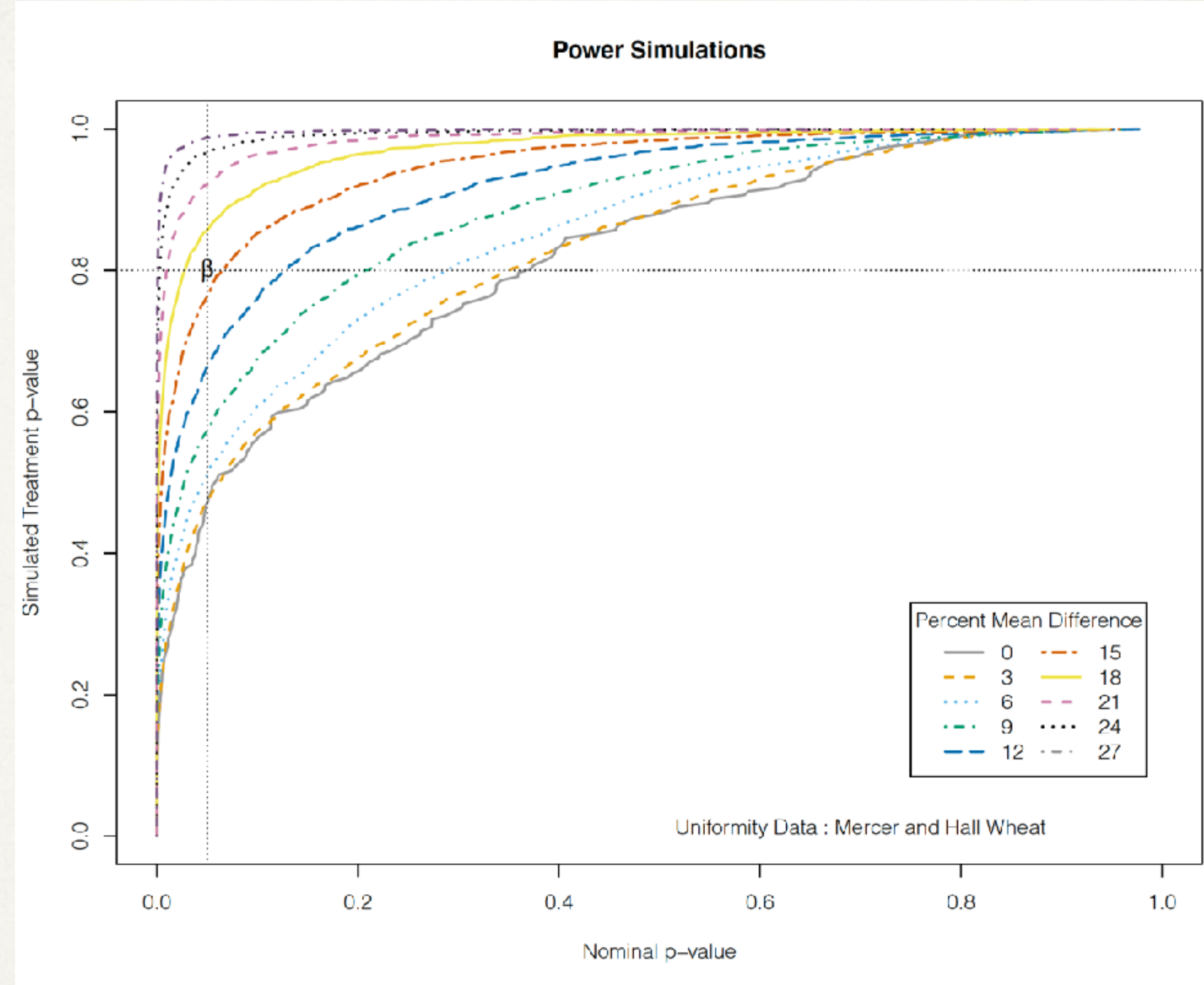


A picture is worth a thousand words.

- ❖ Recall the undesirable randomization:

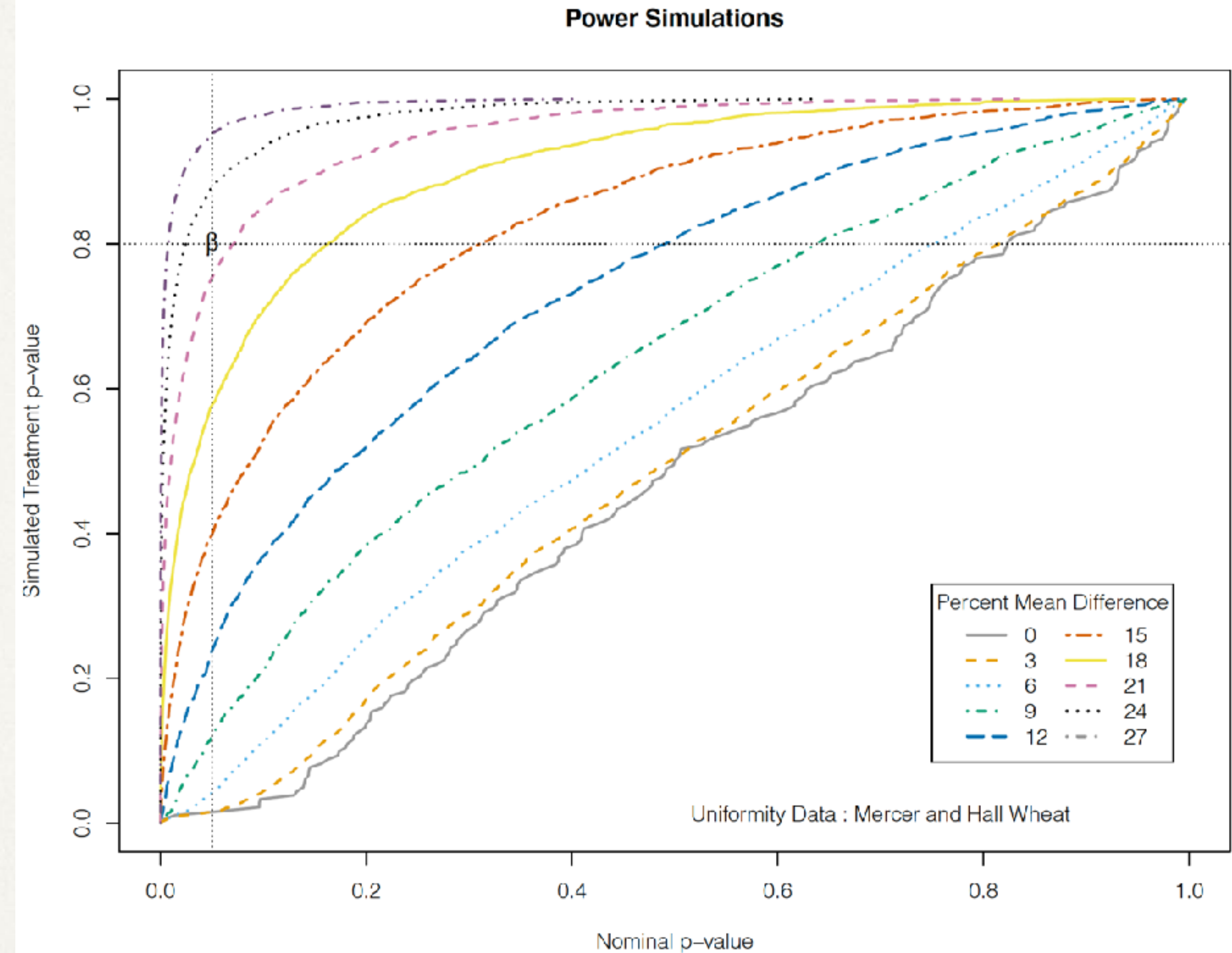


- ❖ and remember that when we have a true treatment effect, we expect a greater number of trials to have significant treatment p-values.



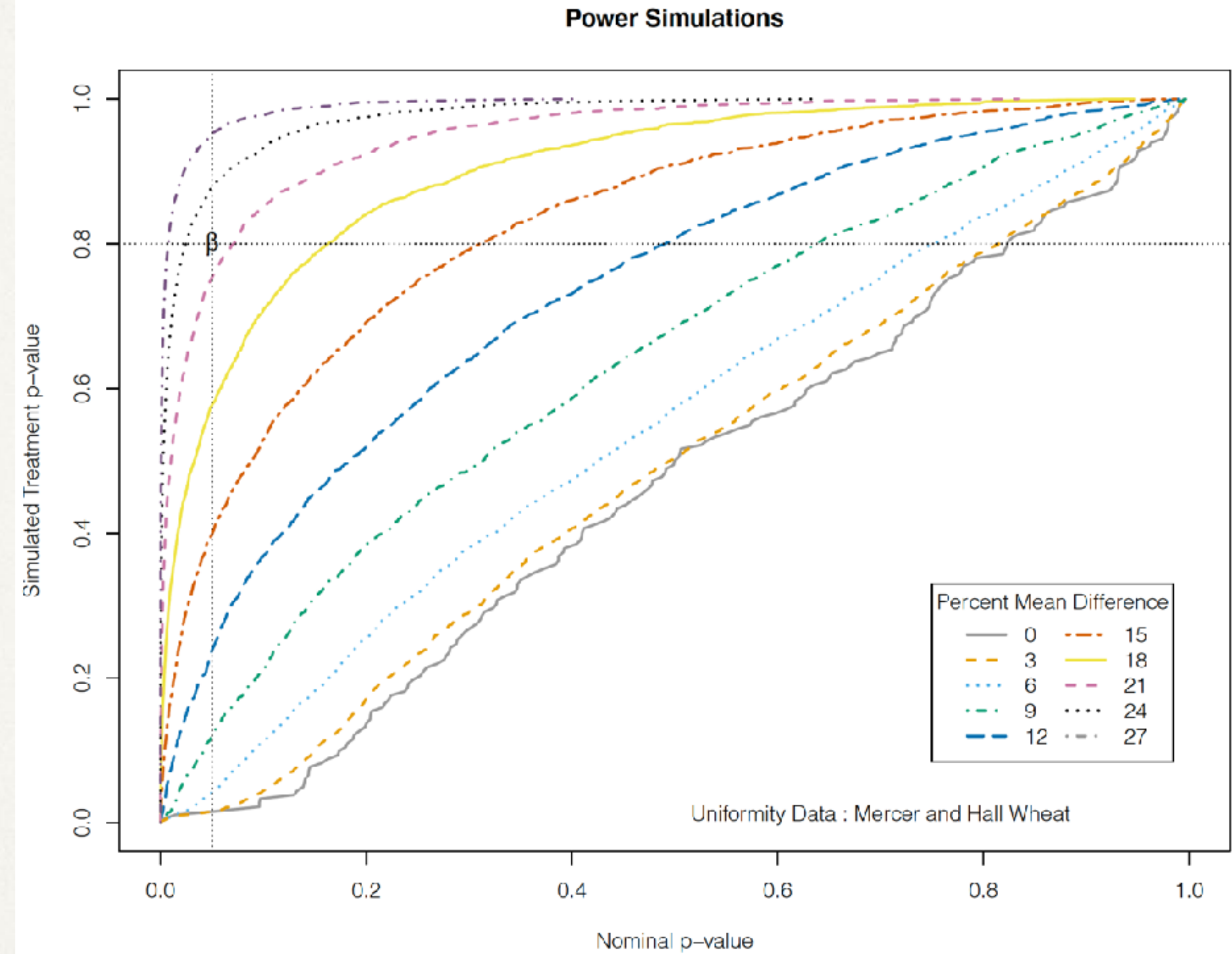
A picture is worth a thousand words.

- ❖ Our goal is to approximate a Percent Mean Difference that gives us significant treatment values at our desired statistical power
- ❖ In this case, that occurs at about a 22% Mean Difference.



A caveat

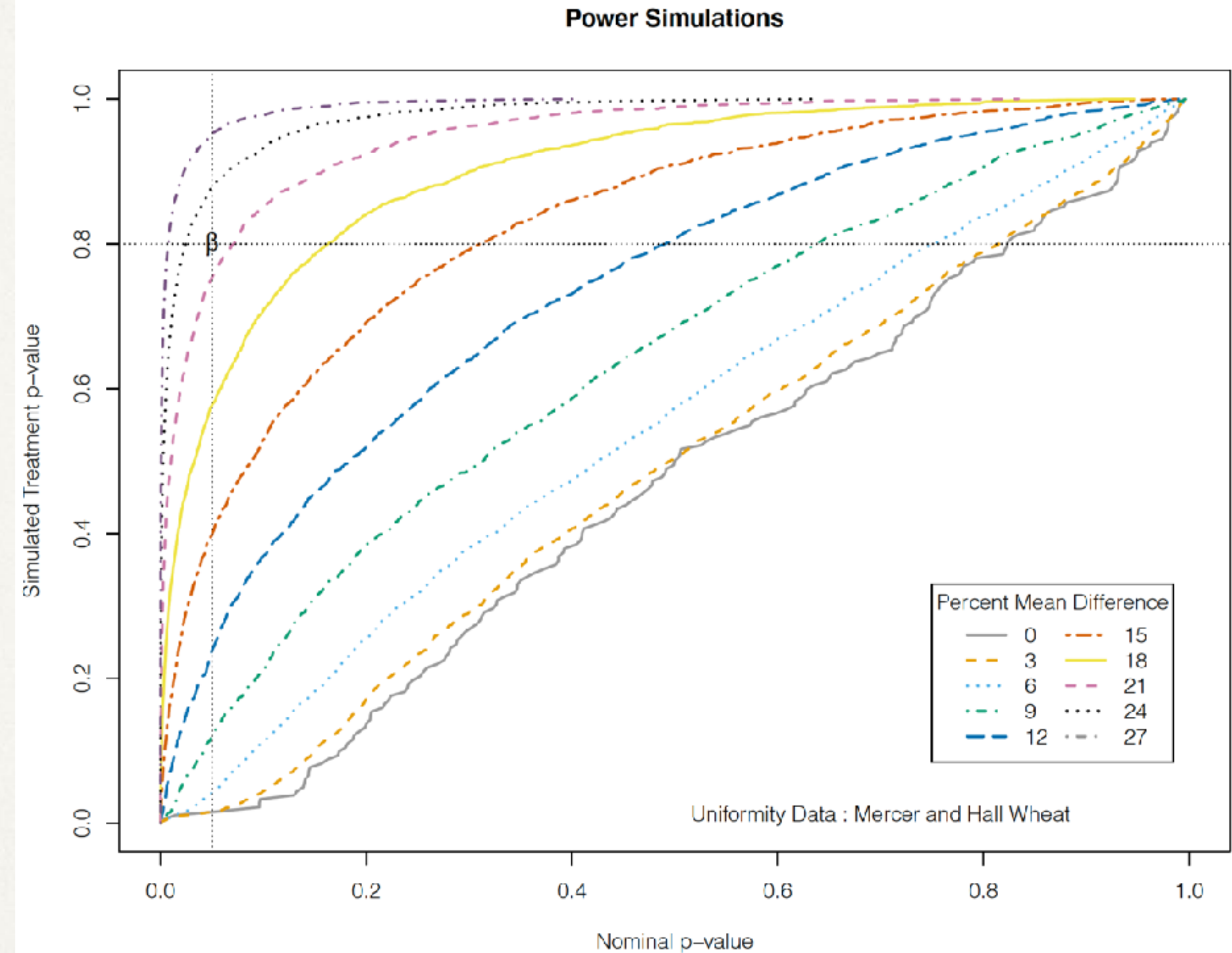
- ❖ Statistical power as suggested by these simulations should not be interpreted as exactly the power used in planning experiments.



A caveat

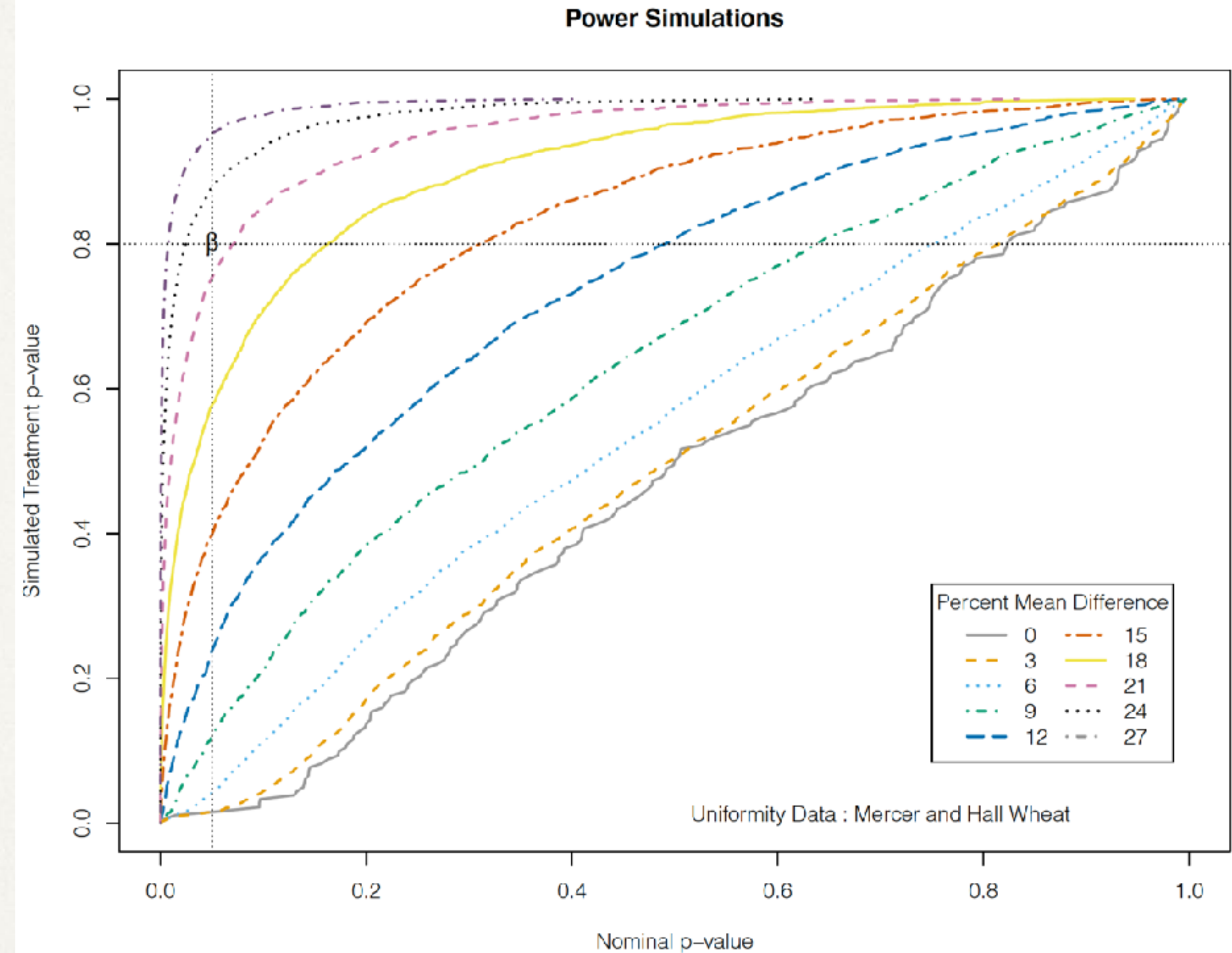
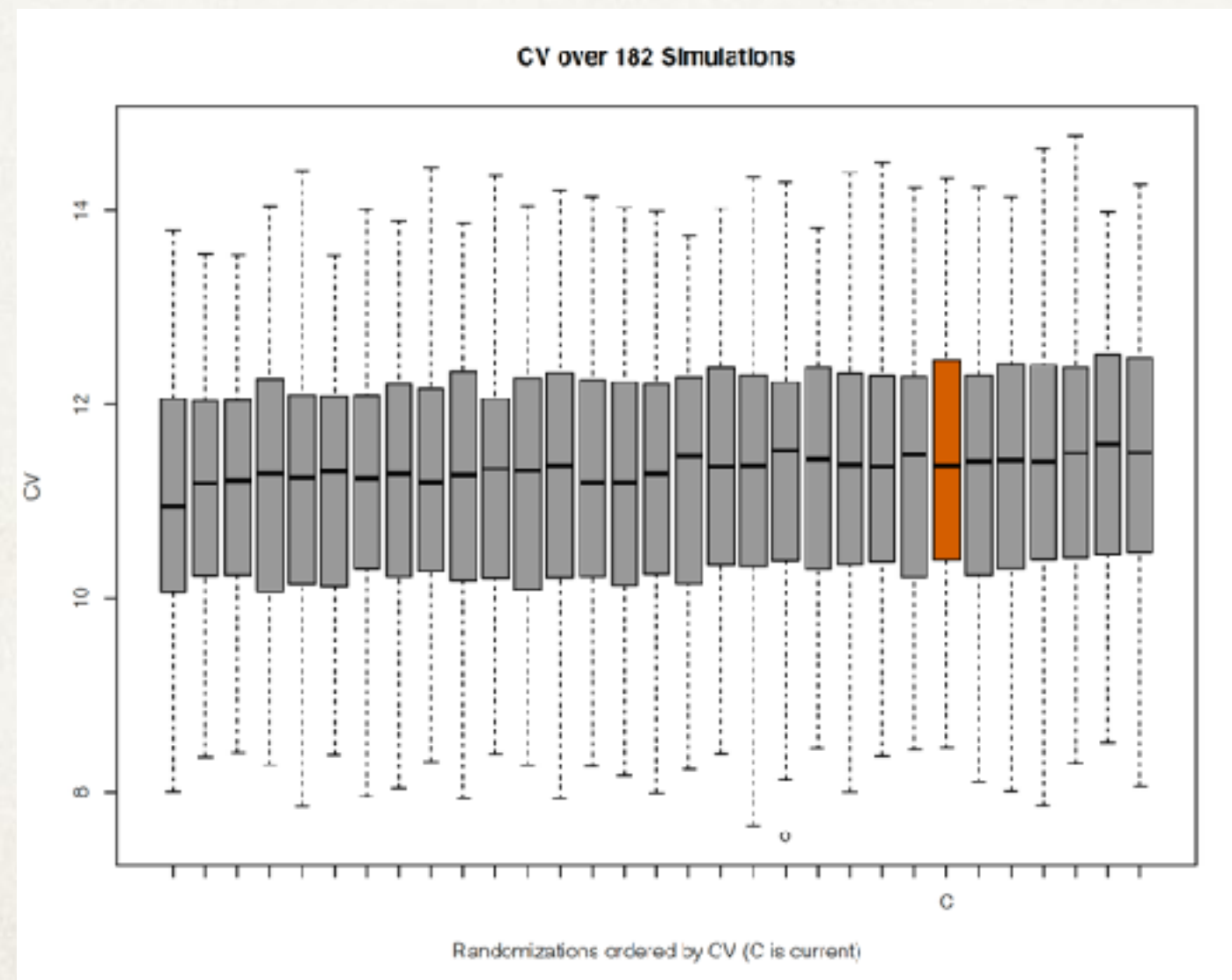
- ❖ The Power and Efficiency table for this design would suggest a Percent Mean Difference of ~20 is detectable with 80% power.

Power and Efficiency						
CV	12.0	Reps	6	Power	80	α SL 5%
Lock at <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>						
CV	Reps	Power	α SL	% Mean Diff	Error DF	'Plot' EUs
12	6	53.3	5%	14.4	55	72
		78		19.2		
		93		24		
		98		28.8		
		>99		33.6		
		80		19.8		
		85		21.1		
		90		22.9		
		95		25.4		



A caveat

- ❖ But we find that by holding the Coefficient of Variance constant at 12%. In the simulated trials, CV can vary over different parts of the field.



A caveat

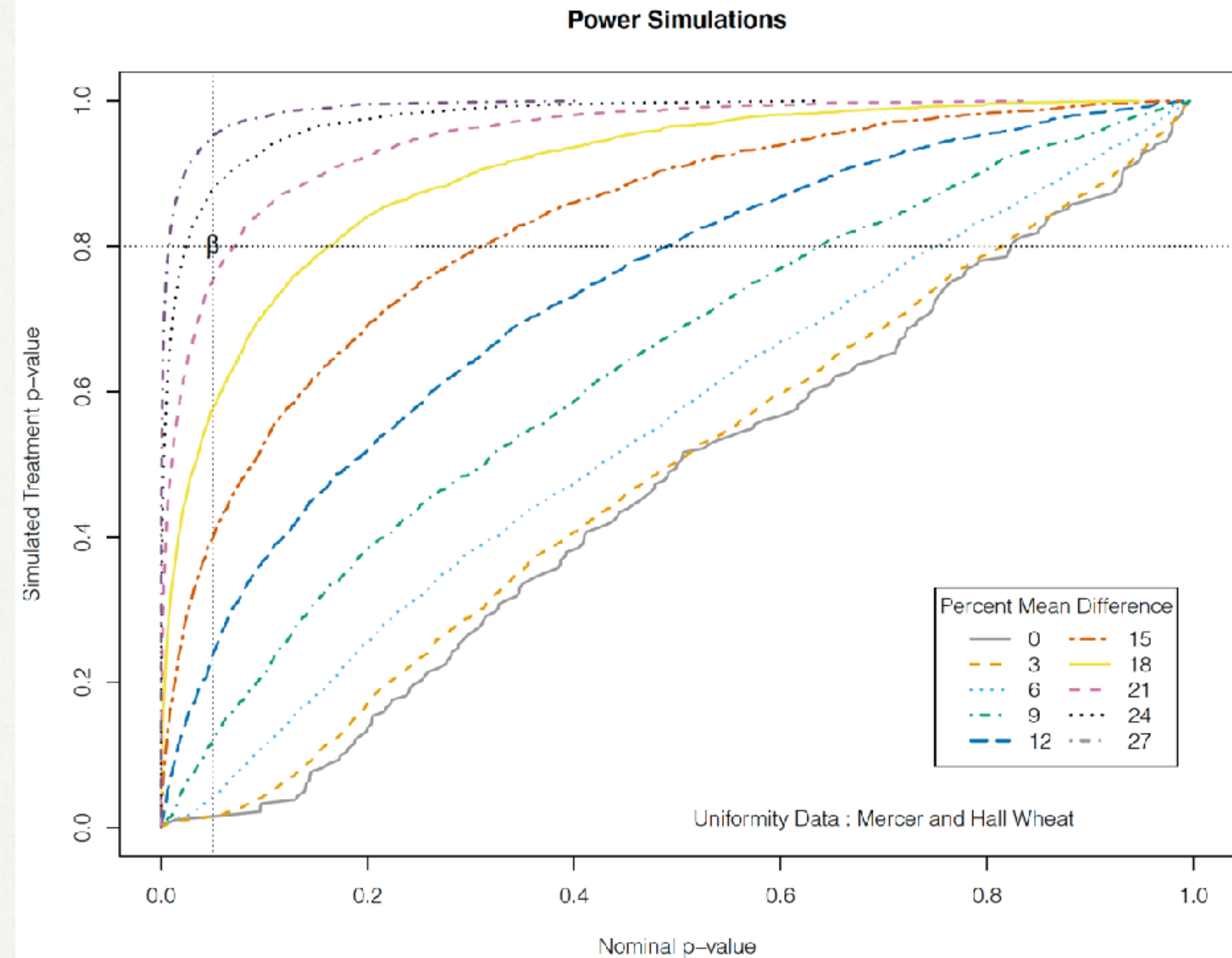
- ❖ Further, we base power calculations on the two sample t-test, while simulations derive p-values from the treatment F-test.

Power and Efficiency

CV 12.0 Reps 6 Power 80 α SL 5% % Mean Diff 24.0

Lock at ☐ ☒ ☒ ☒ ☐

CV	Reps	Power	α SL	% Mean Diff	Error DF	'Plot' EUs
8.75	6	80	5%	14.4	55	72
11.7				19.2		
14.6				24		
17.5				28.8		
20.4				33.6		
7.2				11.9		
9.6				15.8		
12				19.8		
14.4				23.7		
16.8				27.7		



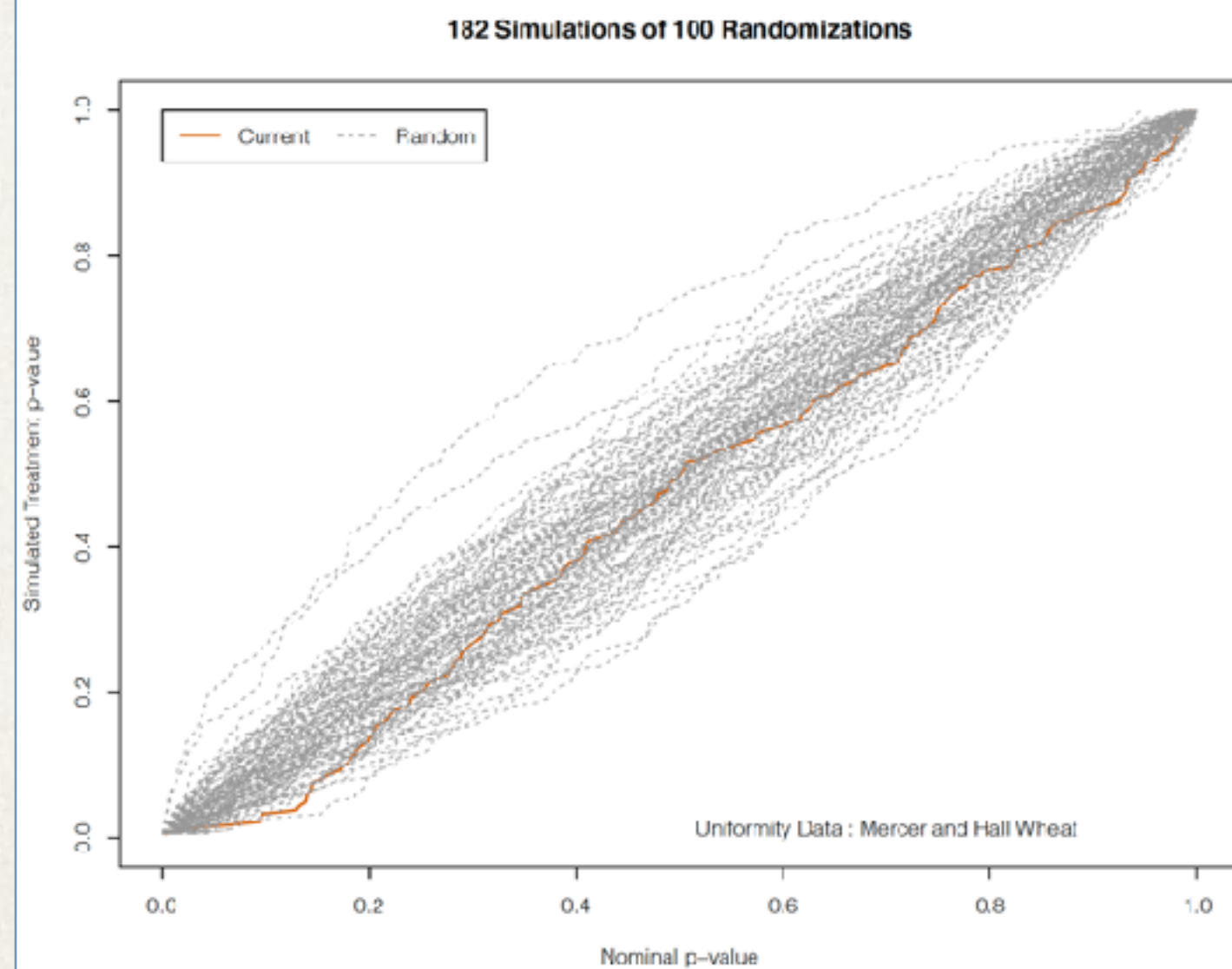
Block Dimensions



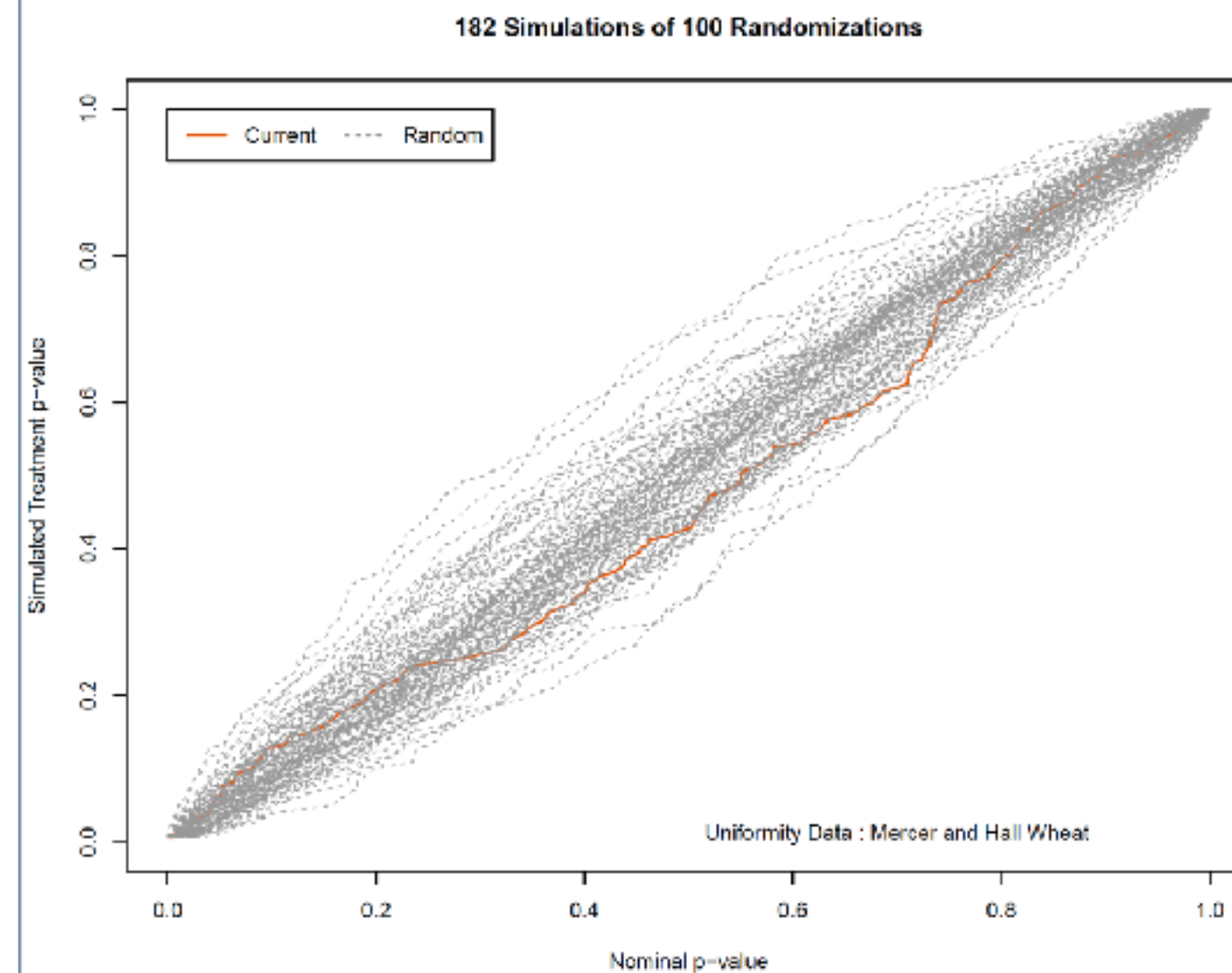
Block Dimensions

- ❖ Simulations, then, may give us insight into the impact of block dimensions on trial outcomes

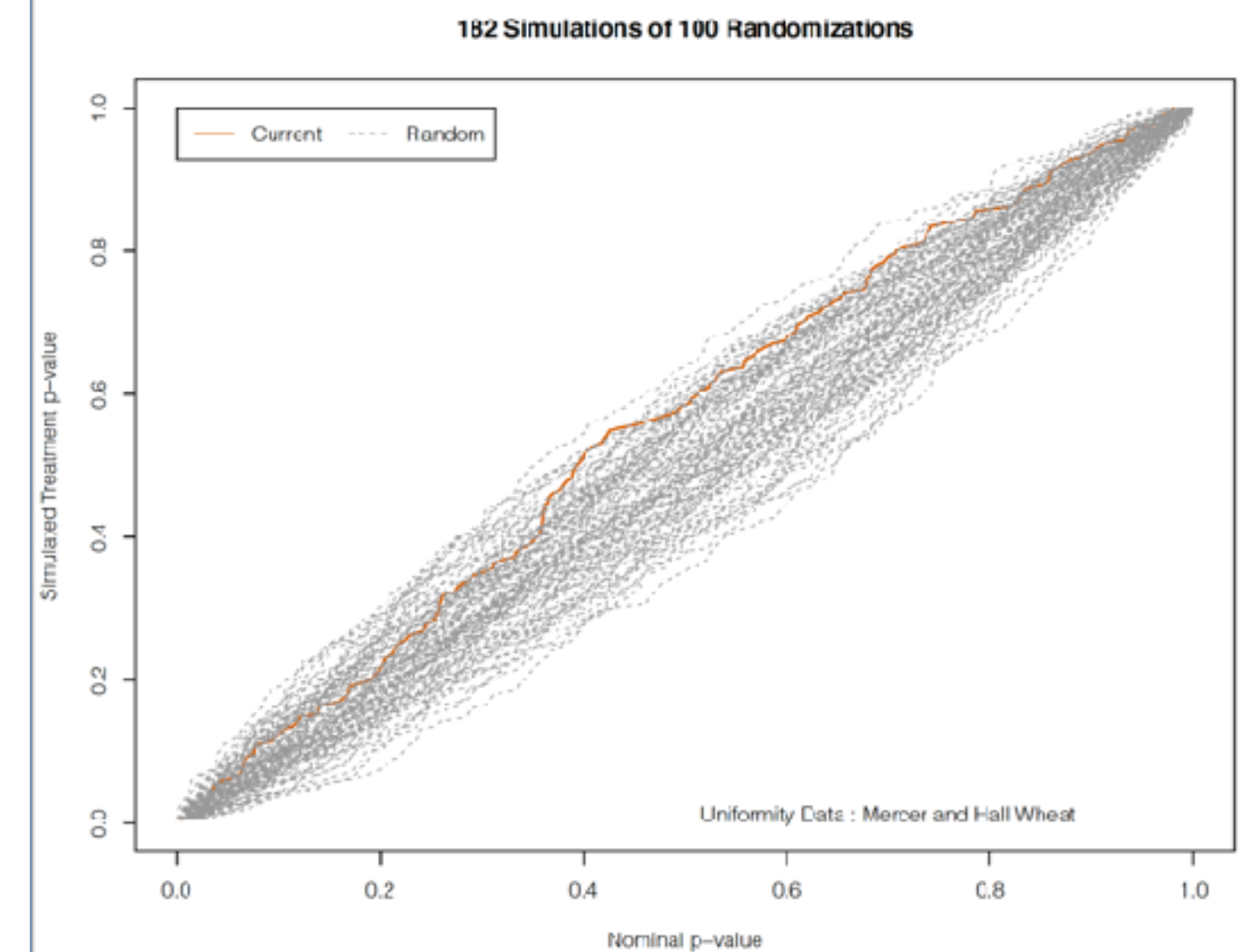
1 x 12



2 x 6



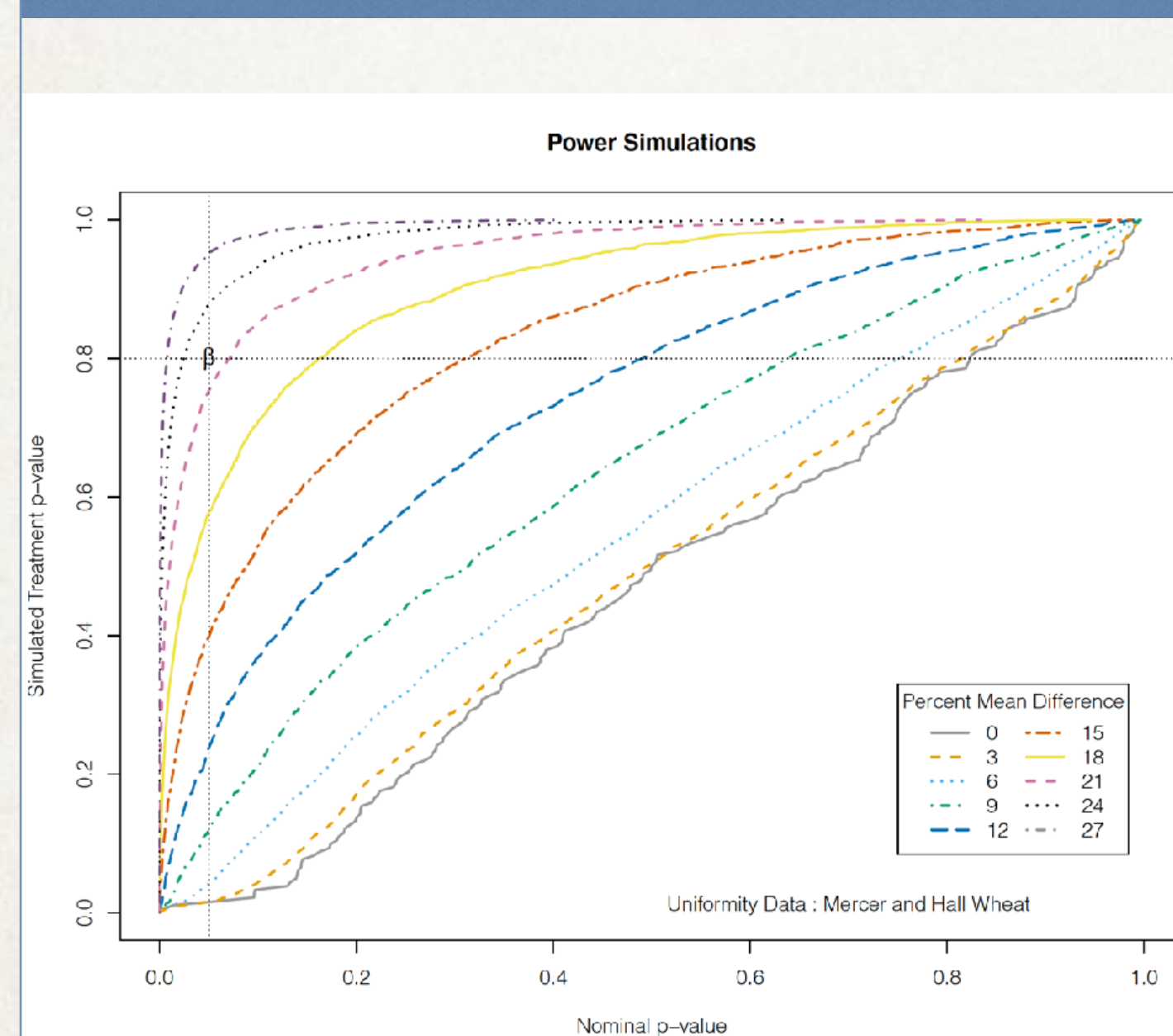
3 x 4



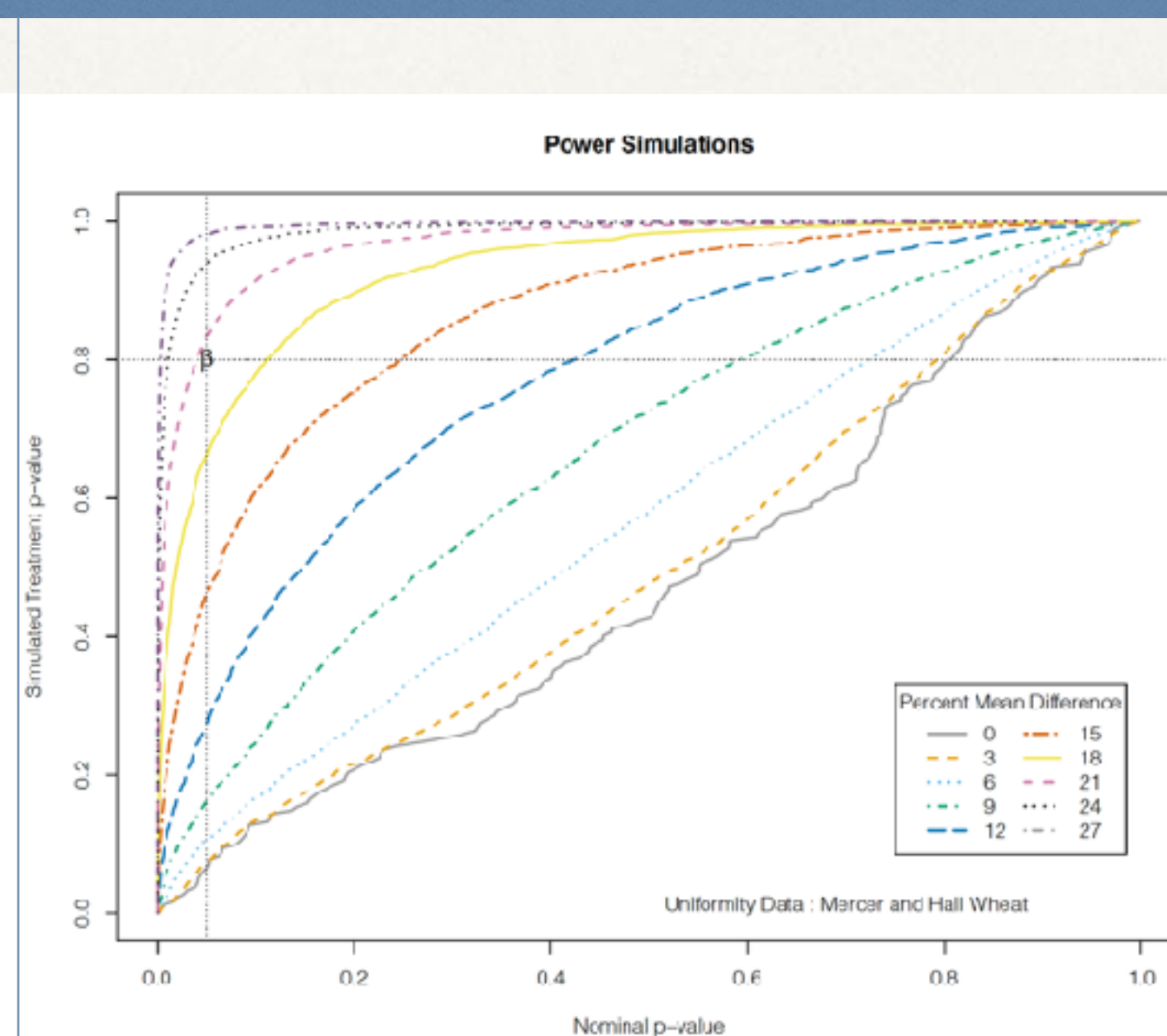
Block Dimensions

- ❖ Simulations, then, may give us insight into the impact of block dimensions on trial outcomes

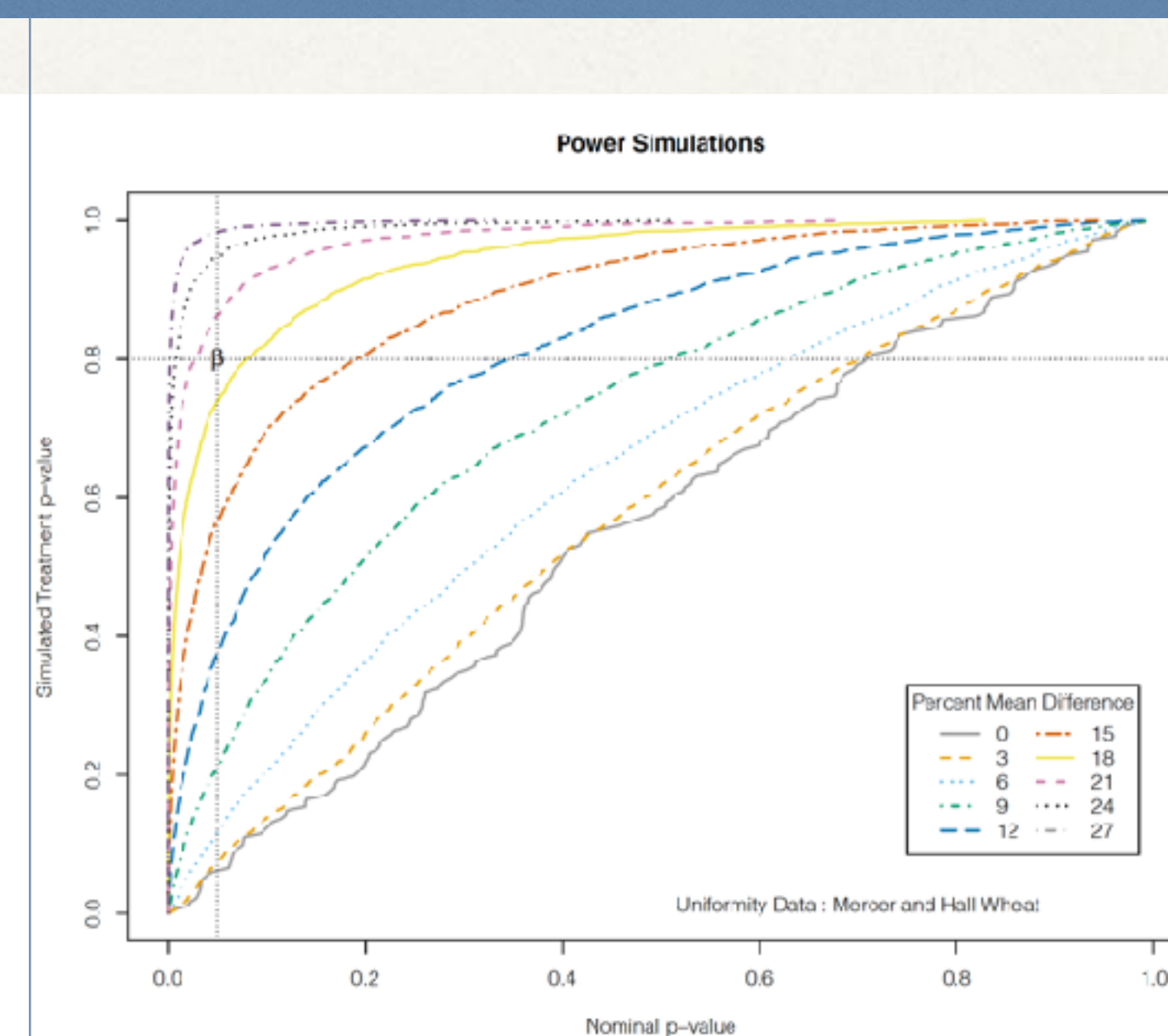
1 x 12 (%MD ~ 22)



2 x 6 (%MD ~ 21)



3 x 4 (%MD ~ 20)



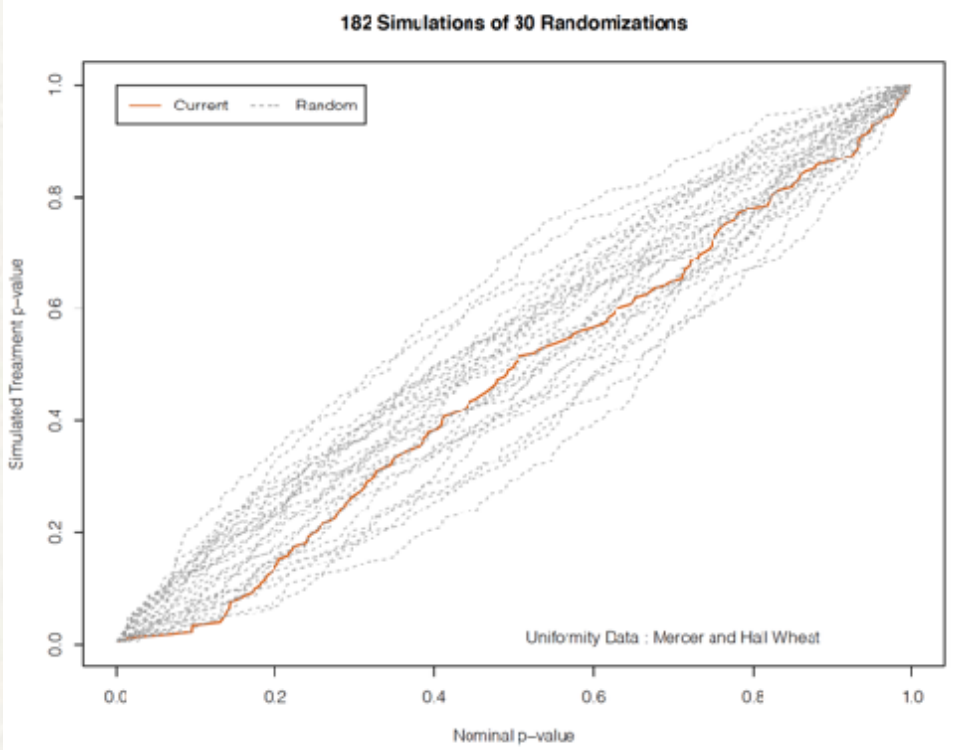
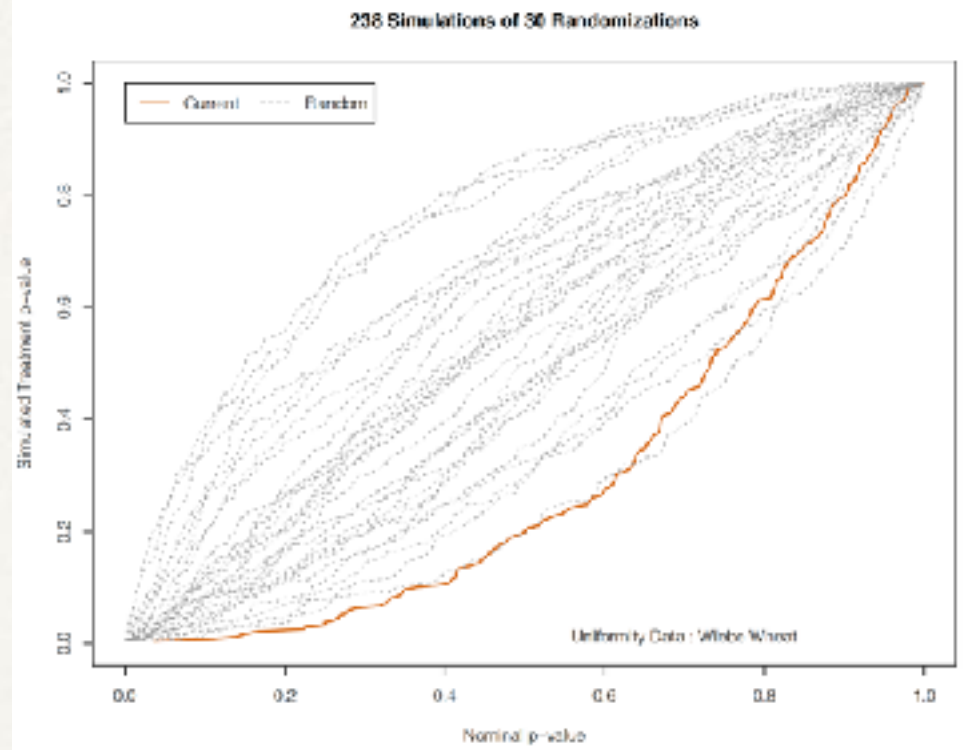
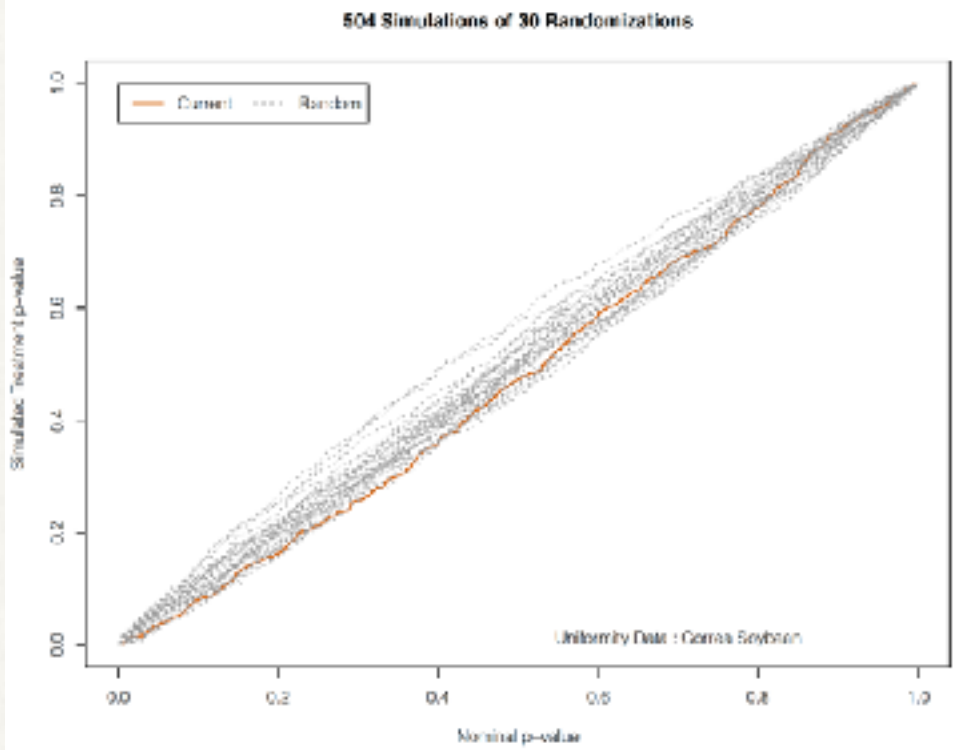
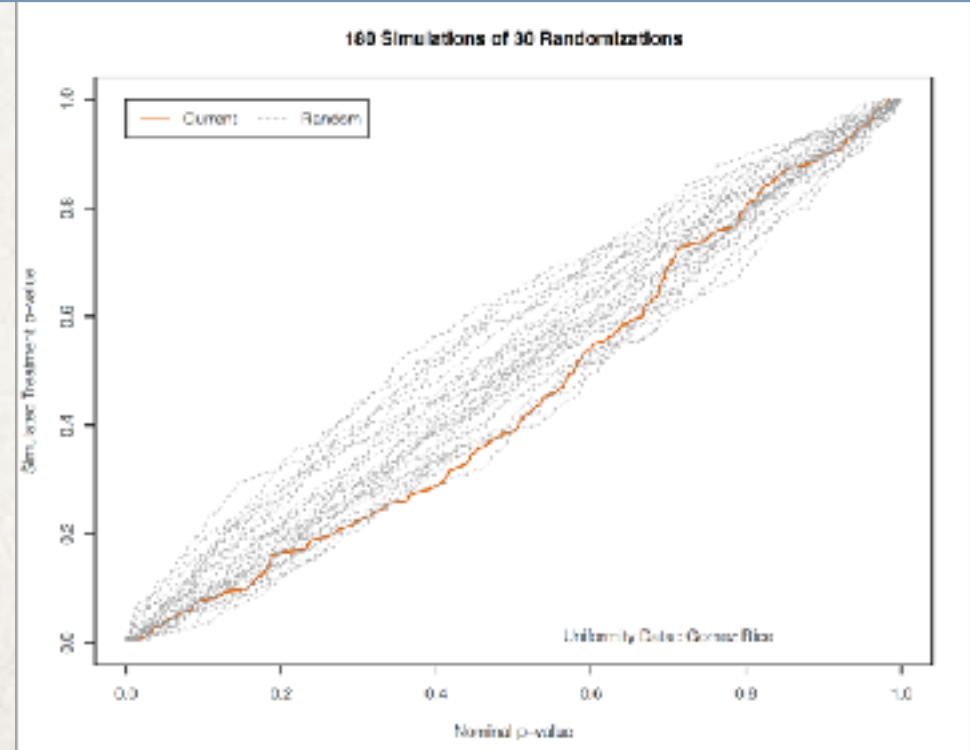
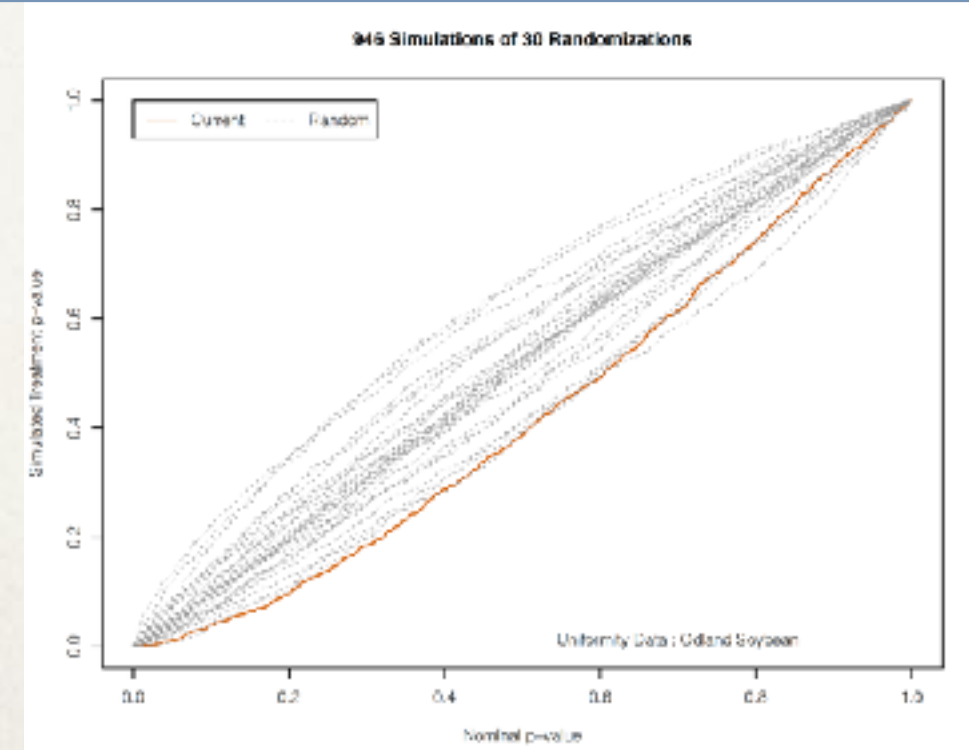
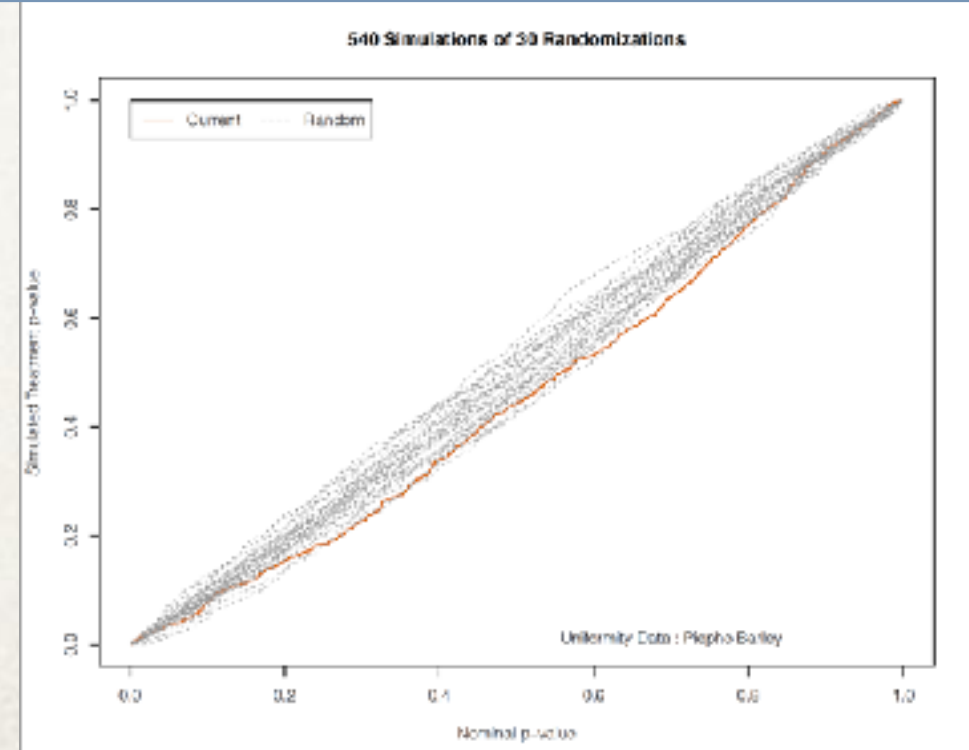
Data Sources

- ❖ For these simulations, I've used data from the R library 'agridat'
- ❖ Kevin Wright
<https://cran.r-project.org/web/packages/agridat/index.html>



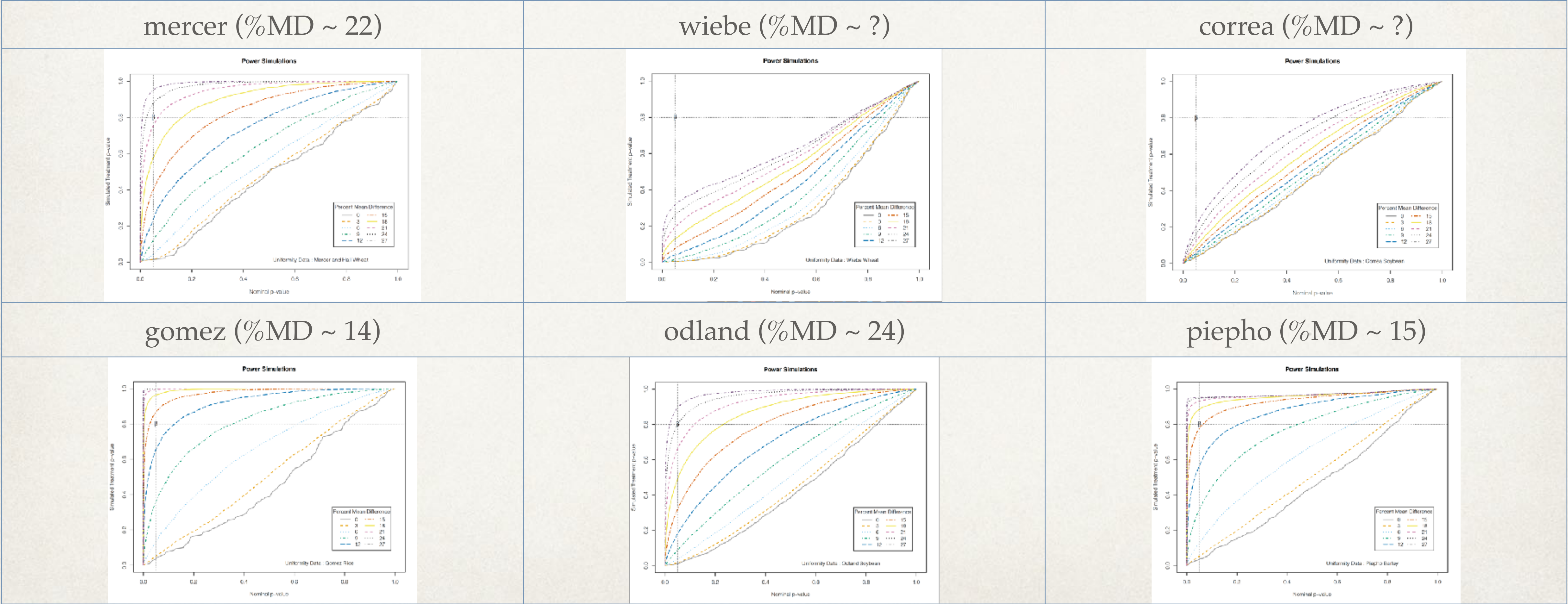
Uniformity Data Sources (agridat)

Repeated Randomizations and Number of Trials

<p>mercer.wheat.uniformity (182)</p>  <p>182 Simulations of 30 Randomizations</p> <p>Simulated Treatment p-value</p> <p>Nominal p-value</p> <p>Uniformity Data : Mercer and Hall Wheat</p>	<p>wiebe.wheat.uniformity (238)</p>  <p>238 Simulations of 30 Randomizations</p> <p>Simulated Treatment p-value</p> <p>Nominal p-value</p> <p>Uniformity Data : Wiebe Wheat</p>	<p>correa.soybean.uniformity (504)</p>  <p>504 Simulations of 30 Randomizations</p> <p>Simulated Treatment p-value</p> <p>Nominal p-value</p> <p>Uniformity Data : Correa Soybean</p>
<p>gomez.rice.uniformity (180)</p>  <p>180 Simulations of 30 Randomizations</p> <p>Simulated Treatment p-value</p> <p>Nominal p-value</p> <p>Uniformity Data : Gomez Rice</p>	<p>odland.soybean.uniformity (946)</p>  <p>946 Simulations of 30 Randomizations</p> <p>Simulated Treatment p-value</p> <p>Nominal p-value</p> <p>Uniformity Data : Odland Soybean</p>	<p>piepho.barley.uniformity (540)</p>  <p>540 Simulations of 30 Randomizations</p> <p>Simulated Treatment p-value</p> <p>Nominal p-value</p> <p>Uniformity Data : Piepho Barley</p>

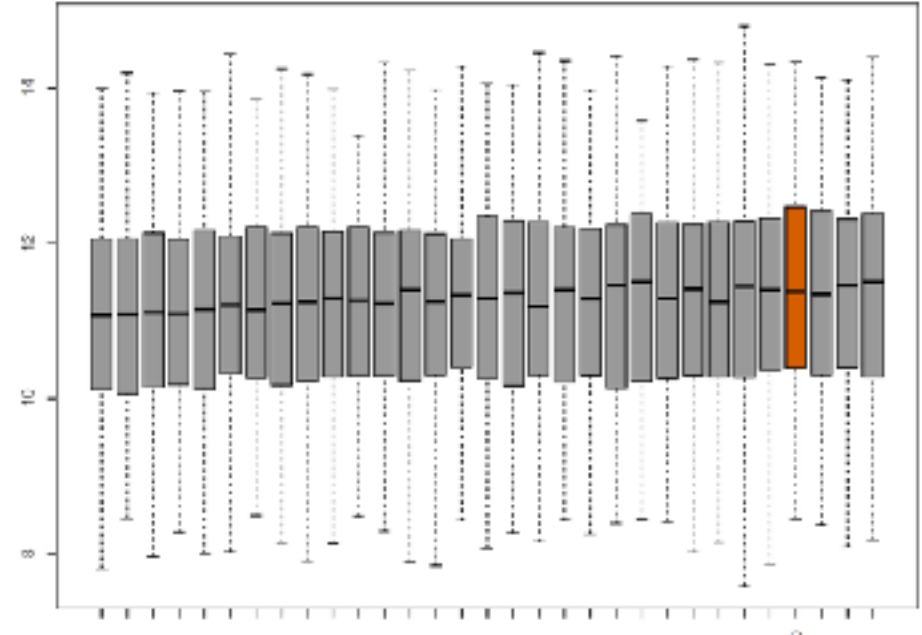
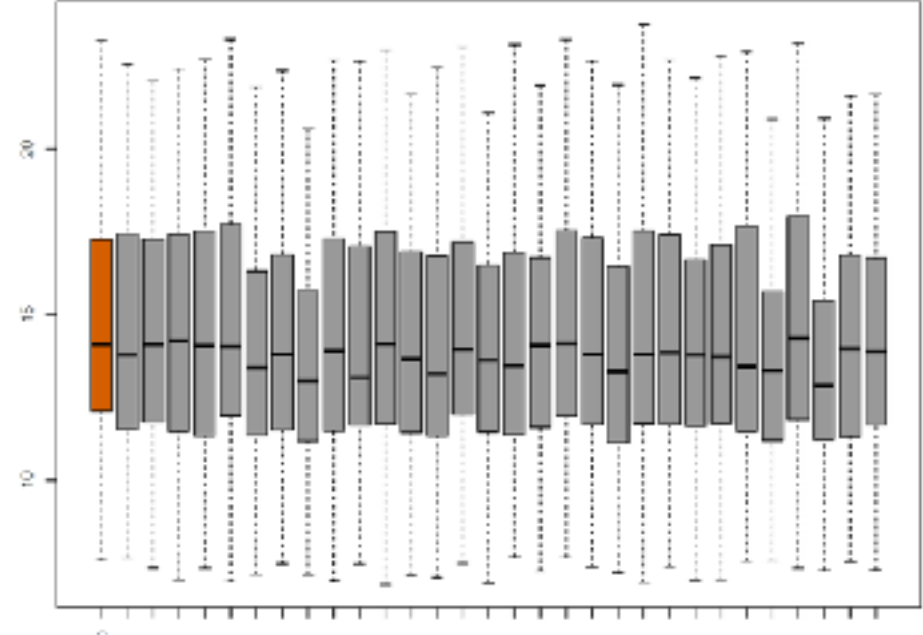
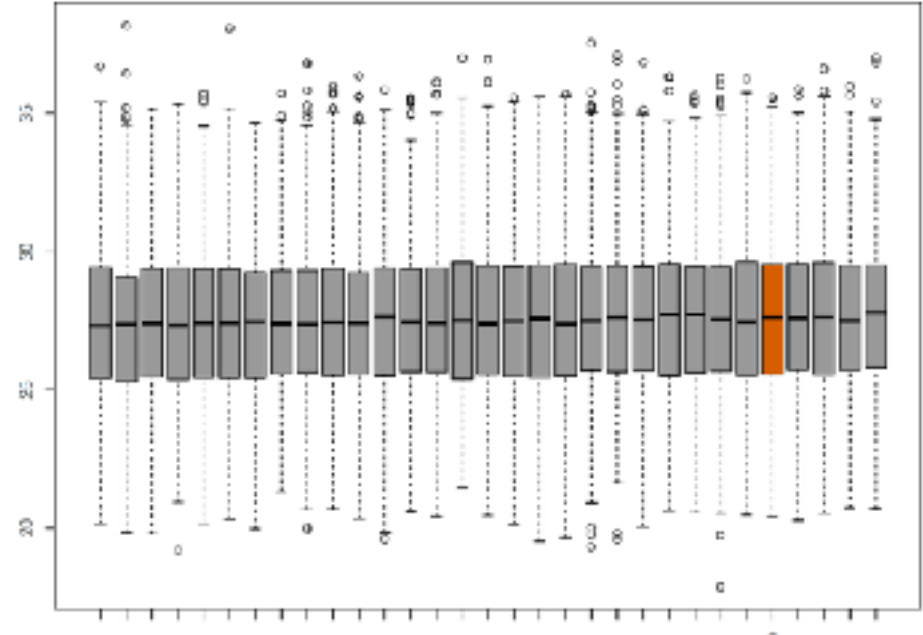
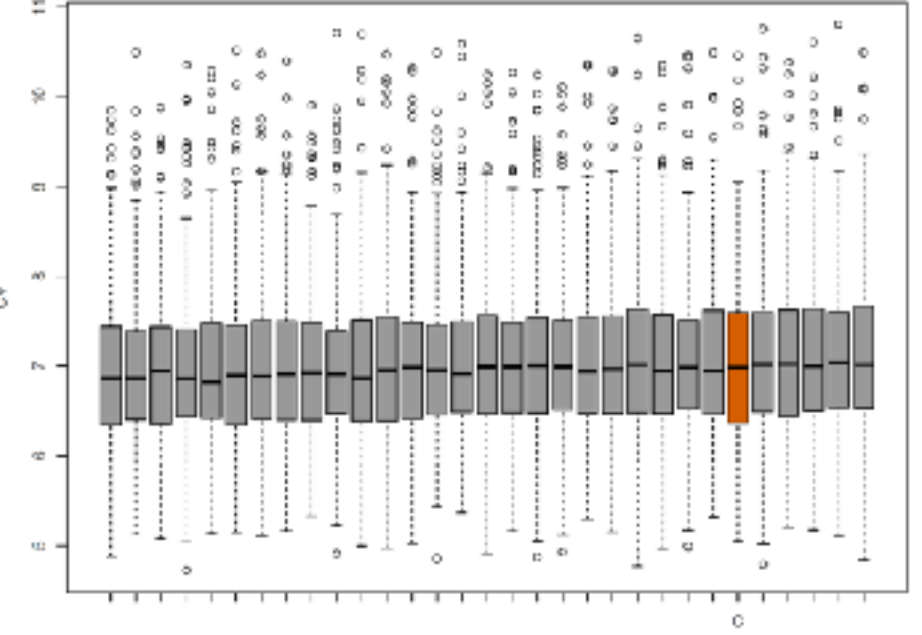
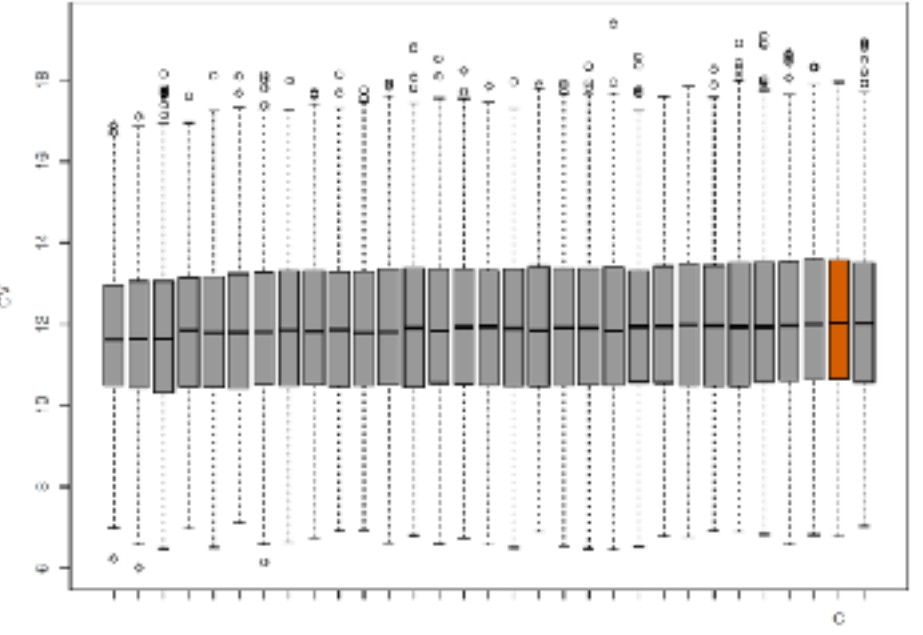
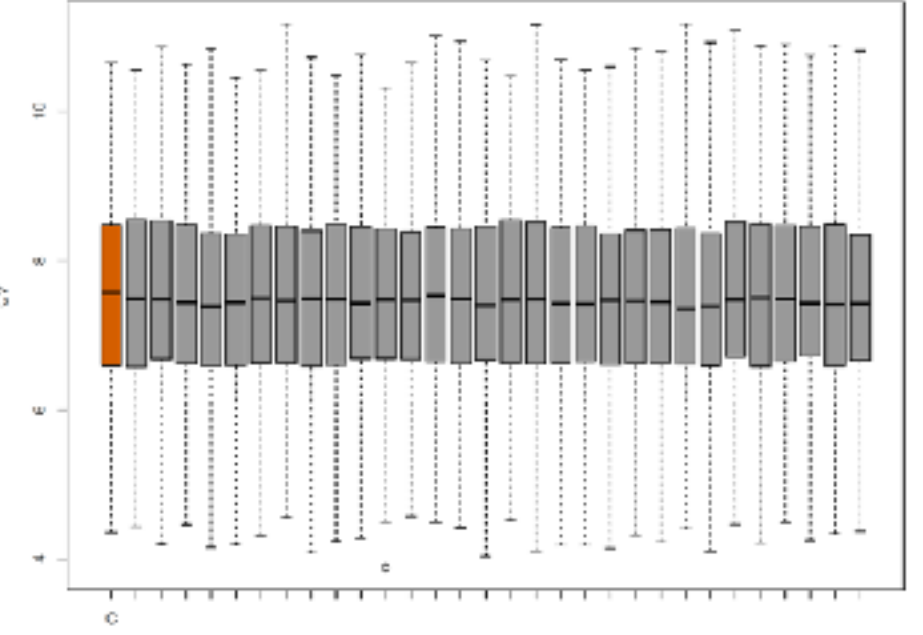
Uniformity Data Sources (agridat)

Power Analysis and Approximate Mean Difference at 80%



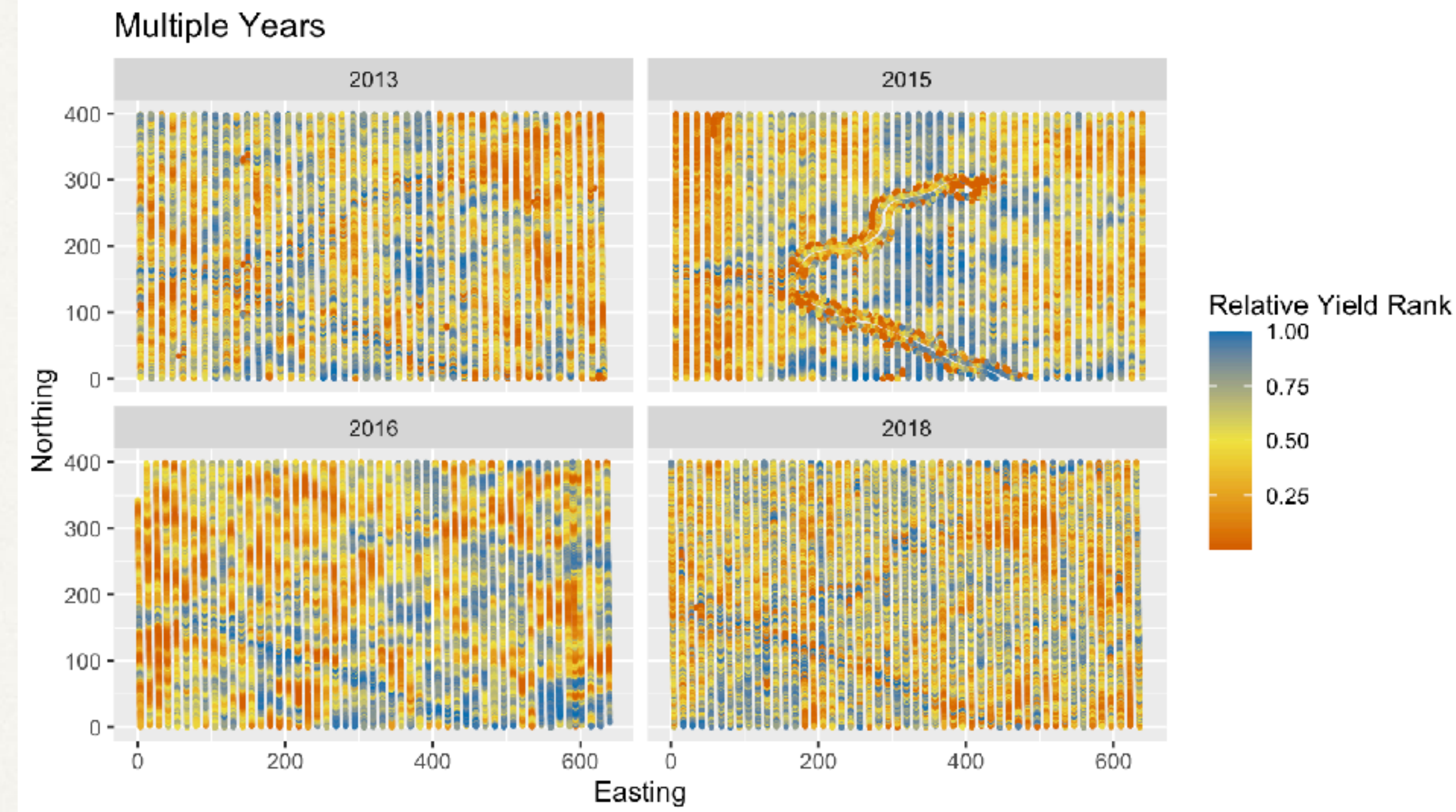
Uniformity Data Sources (agridat)

Coefficient of Variance for Plans and Simulations

<p>mercer : CV ~ 11 (8-14)</p> <div data-bbox="296 671 929 1178"><p>CV over 162 Simulations</p><p>Randomizations ordered by CV (C is current)</p></div>	<p>wiebe : CV ~ 14 (7-24)</p> <div data-bbox="1346 671 1979 1178"><p>CV over 238 Simulations</p><p>Randomizations ordered by CV (C is current)</p></div>	<p>correa : CV ~ 27 (20-35)</p> <div data-bbox="2395 671 3028 1178"><p>CV over 604 Simulations</p><p>Randomizations ordered by CV (C is current)</p></div>
<p>gomez : CV ~ 7 (5-9)</p> <div data-bbox="296 1290 929 1787"><p>CV over 100 Simulations</p><p>Randomizations ordered by CV (C is current)</p></div>	<p>odland : CV ~ 12 (7-18)</p> <div data-bbox="1346 1290 1979 1787"><p>CV over 945 Simulations</p><p>Randomizations ordered by CV (C is current)</p></div>	<p>piepho : CV ~ 7 (4-11)</p> <div data-bbox="2395 1290 3028 1787"><p>CV over 540 Simulations</p><p>Randomizations ordered by CV (C is current)</p></div>

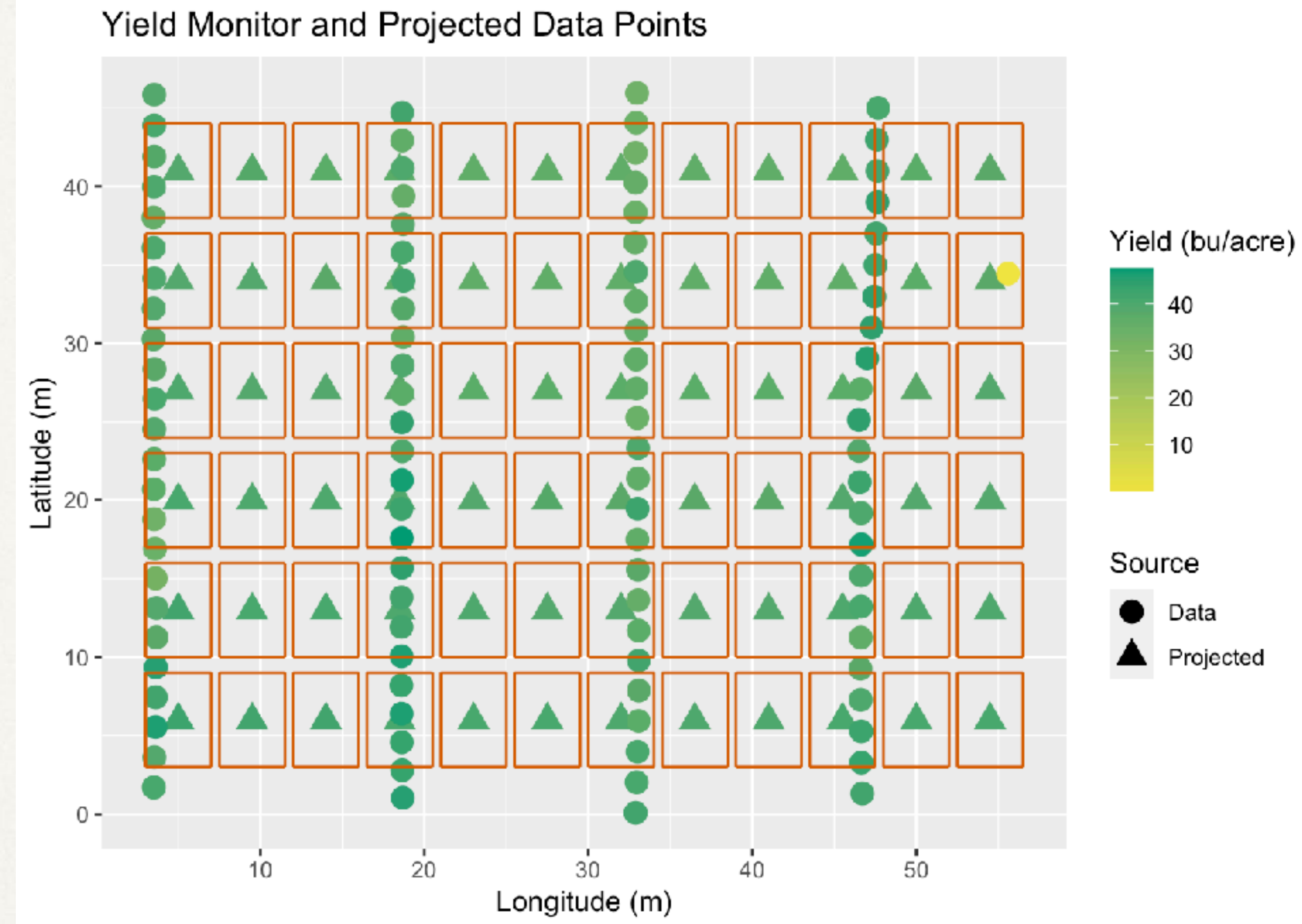
Other Sources

- ❖ It's relatively easy to use published uniformity trial data, but there are some conditions
 - ❖ There may not be data sources similar to your crops or your field conditions
 - ❖ Uniformity trials may not have been conducted using your plot and buffer dimensions
- ❖ Yield monitor data may be adapted for use as uniformity data



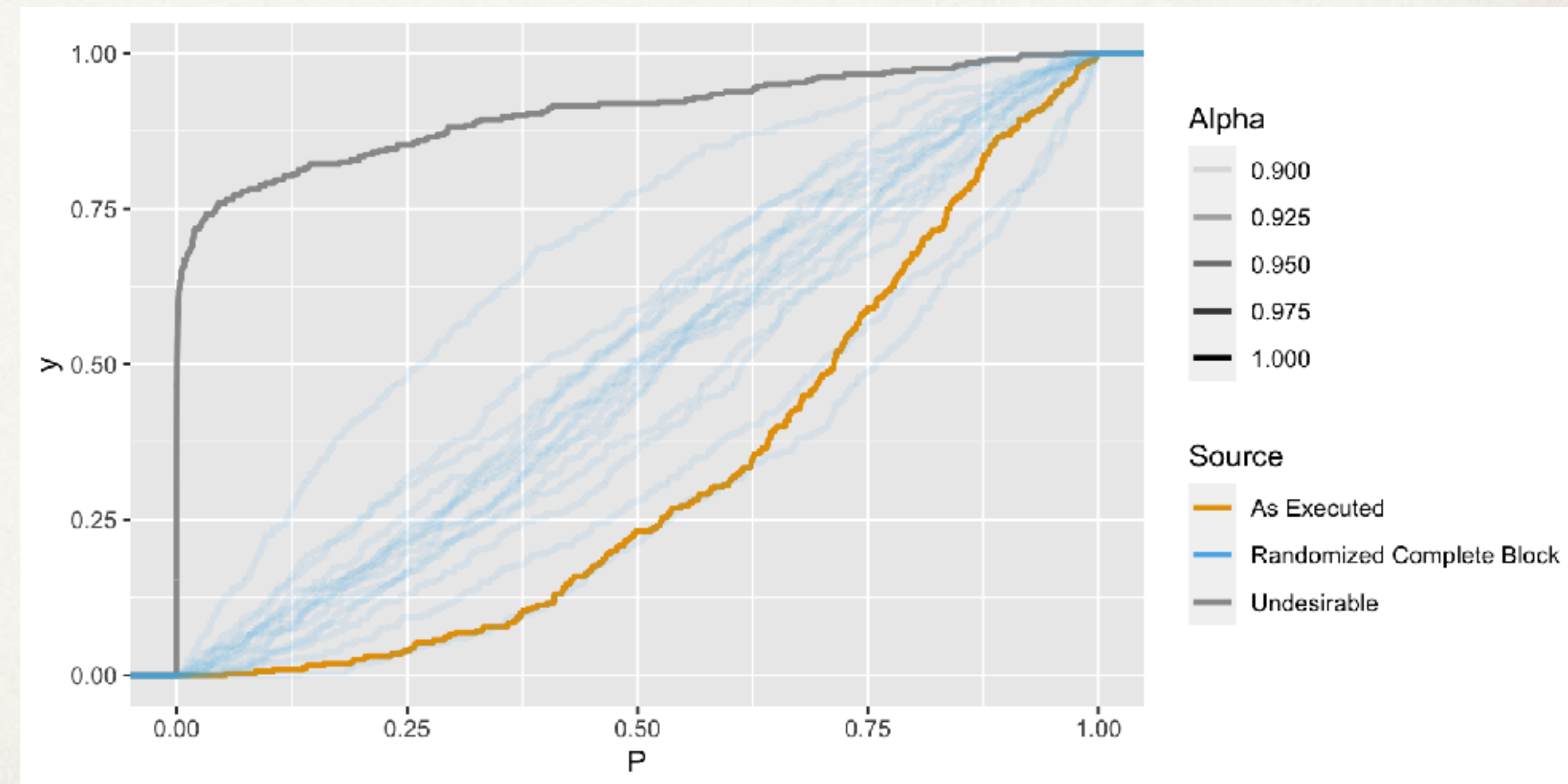
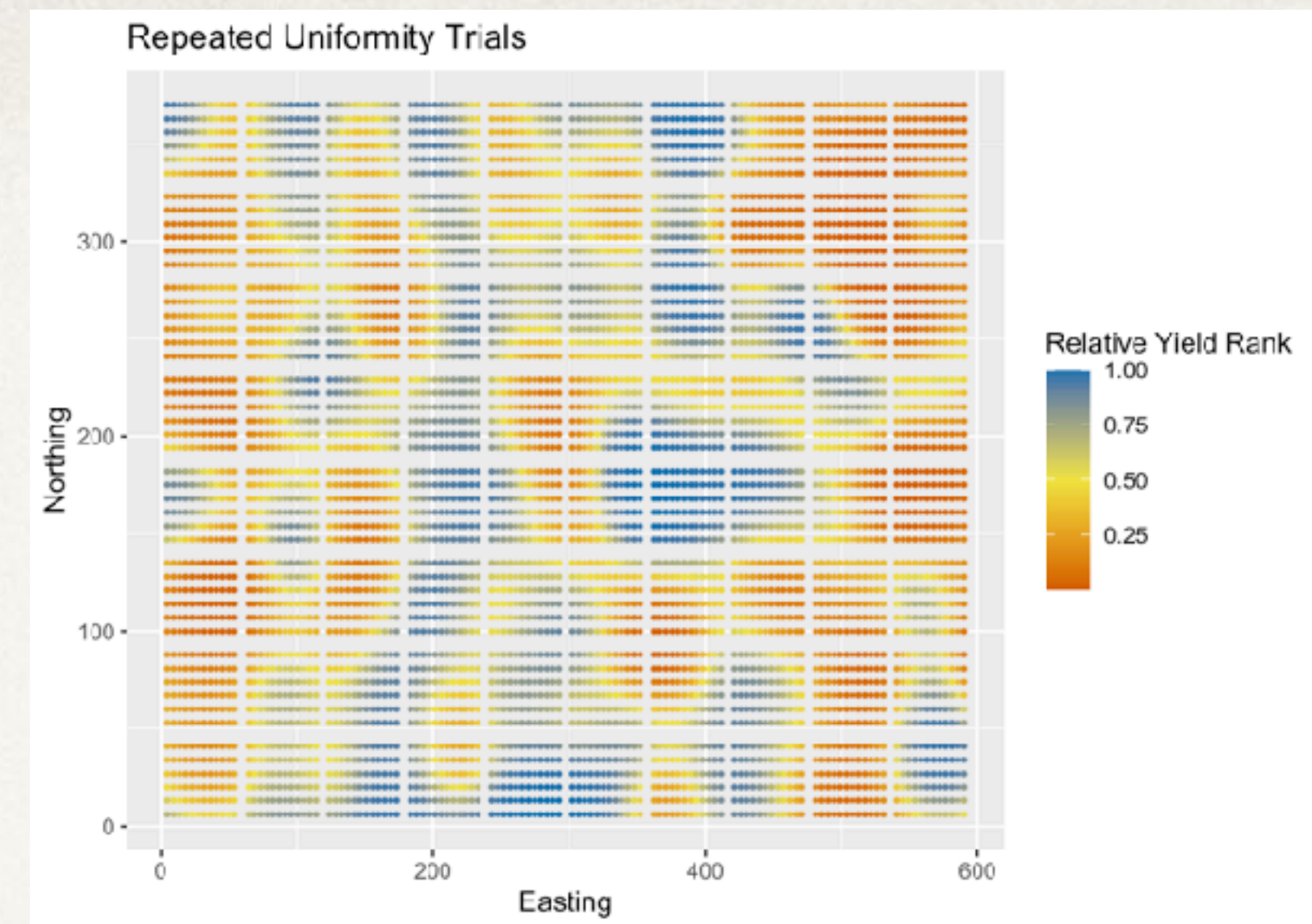
Yield Monitor Data

- ❖ Fit a statistical model to yield monitor data (gam)
- ❖ Project predicted yields onto a trial map, using the plot centers
- ❖ Repeat this over the entire map, and over individual maps from multiple years



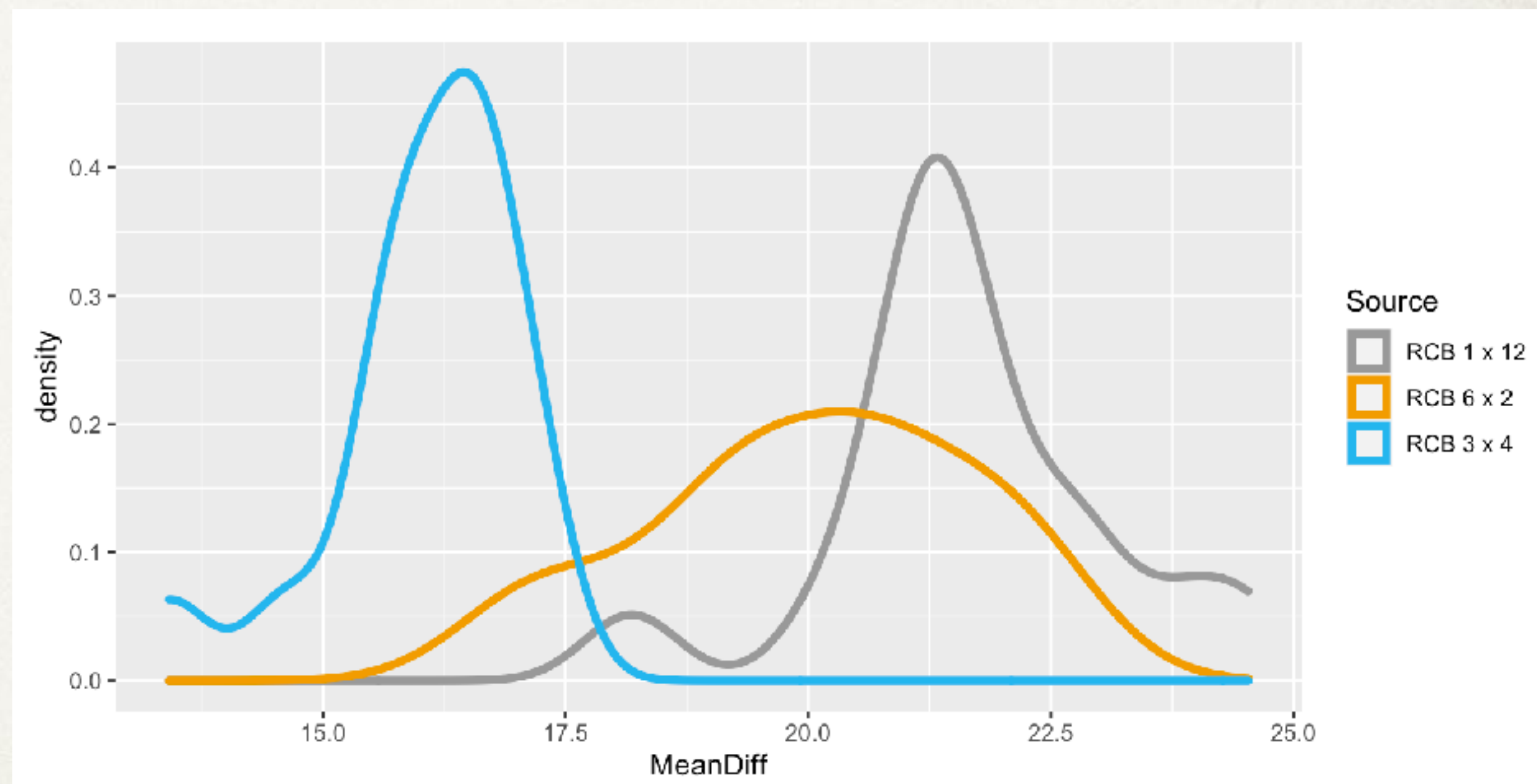
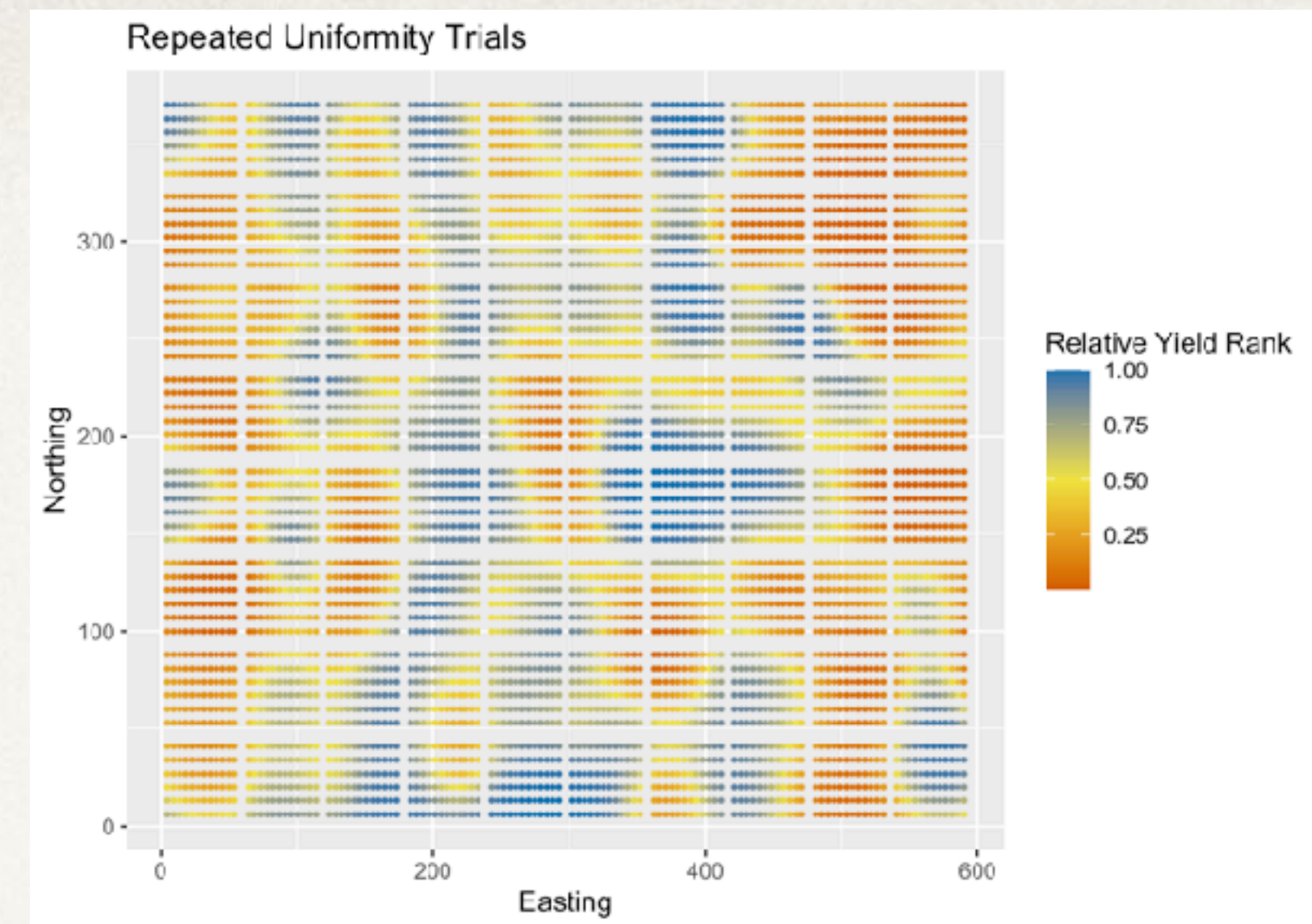
Yield Monitor Data

- ❖ Given a field of $\sim 600\text{m} \times 360\text{m}$, and a trial of 72 plots, 6 rows by 12 columns, 4x6 m plots with 0.5 and 1 m buffers, we can fit 8x10 trials per year over 4 years for 320 (non-overlapping) simulated trials
- ❖ I've used this in the past to explore ARM randomization options, but this will be more difficult to include as an ARM feature.



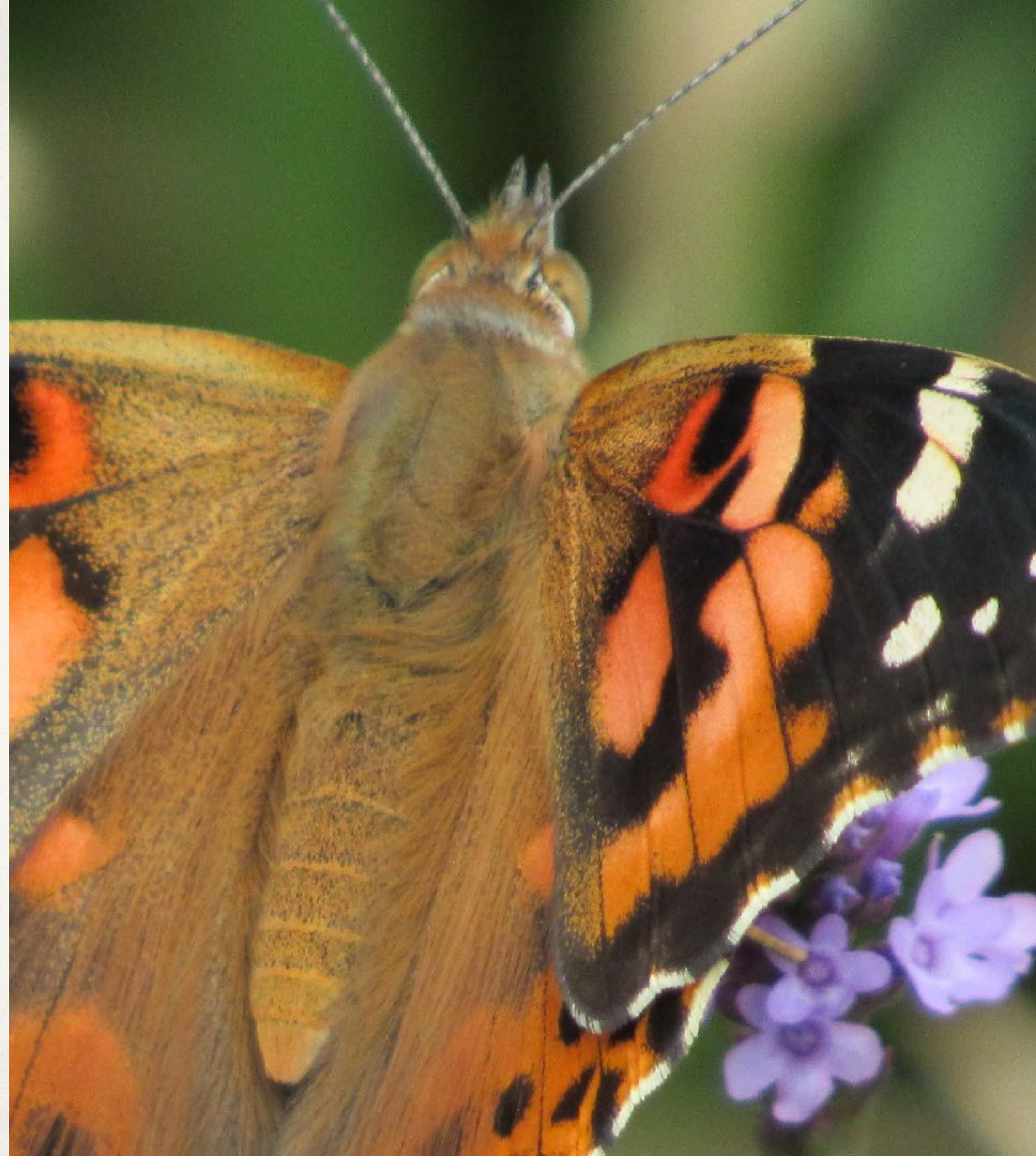
Yield Monitor Data

- ❖ Simulating uniformity trials and power analysis with yield monitor data in R allows me to compare many different randomizations at once, with better separation of statistical power.
- ❖ I can create as many trials, without overlap
- ❖ But, these take a very long time (hours), so may not yet provide a general solution.



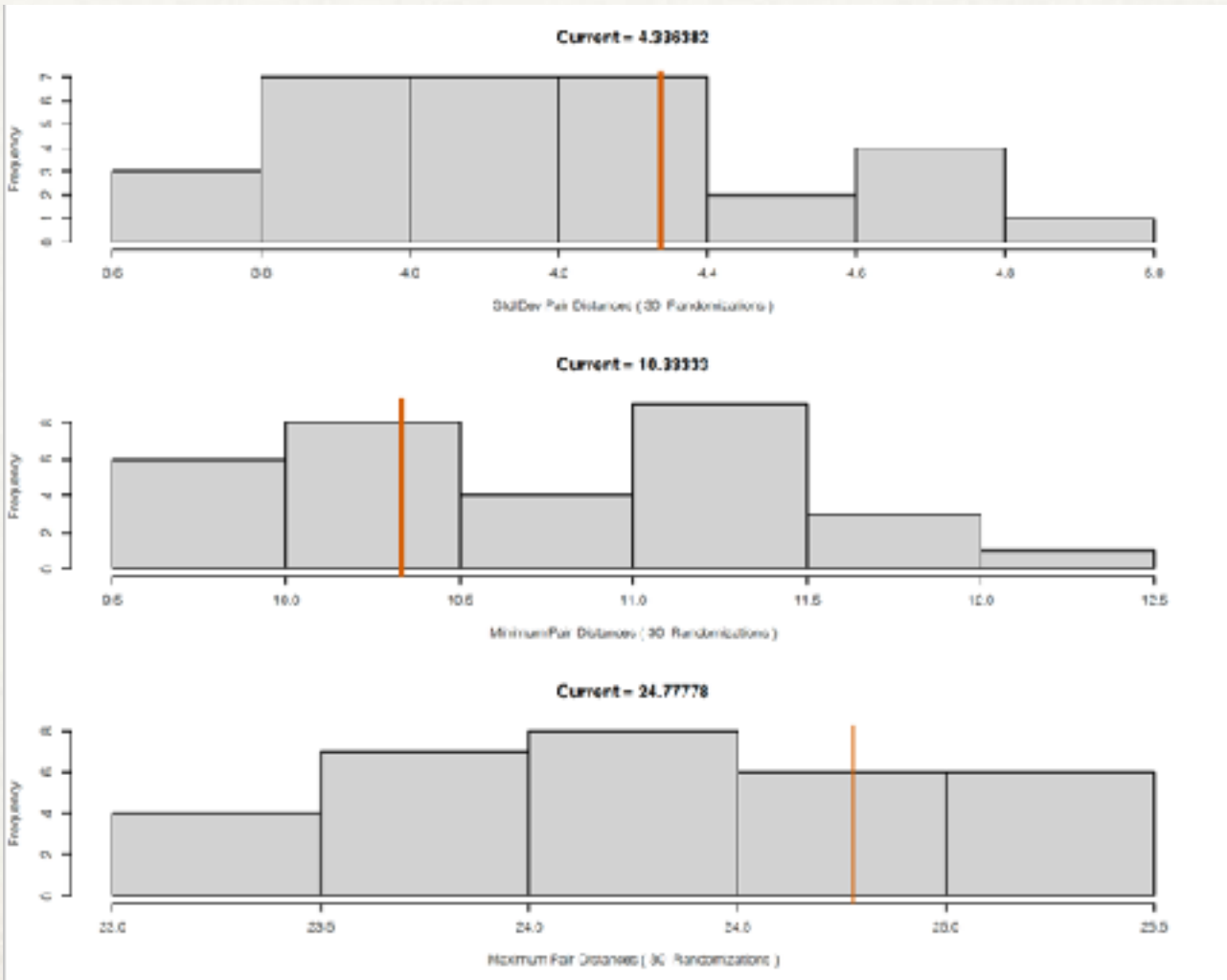
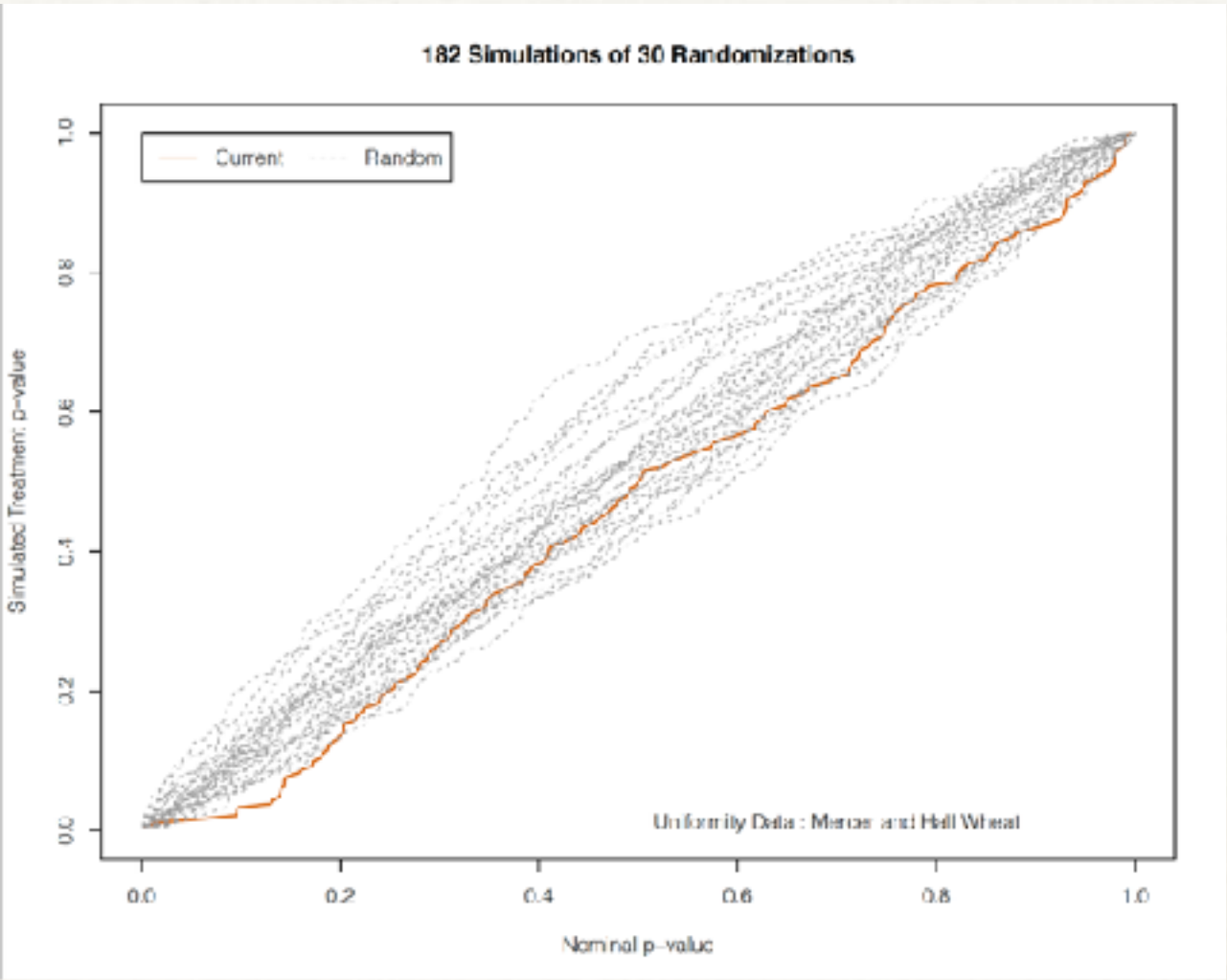
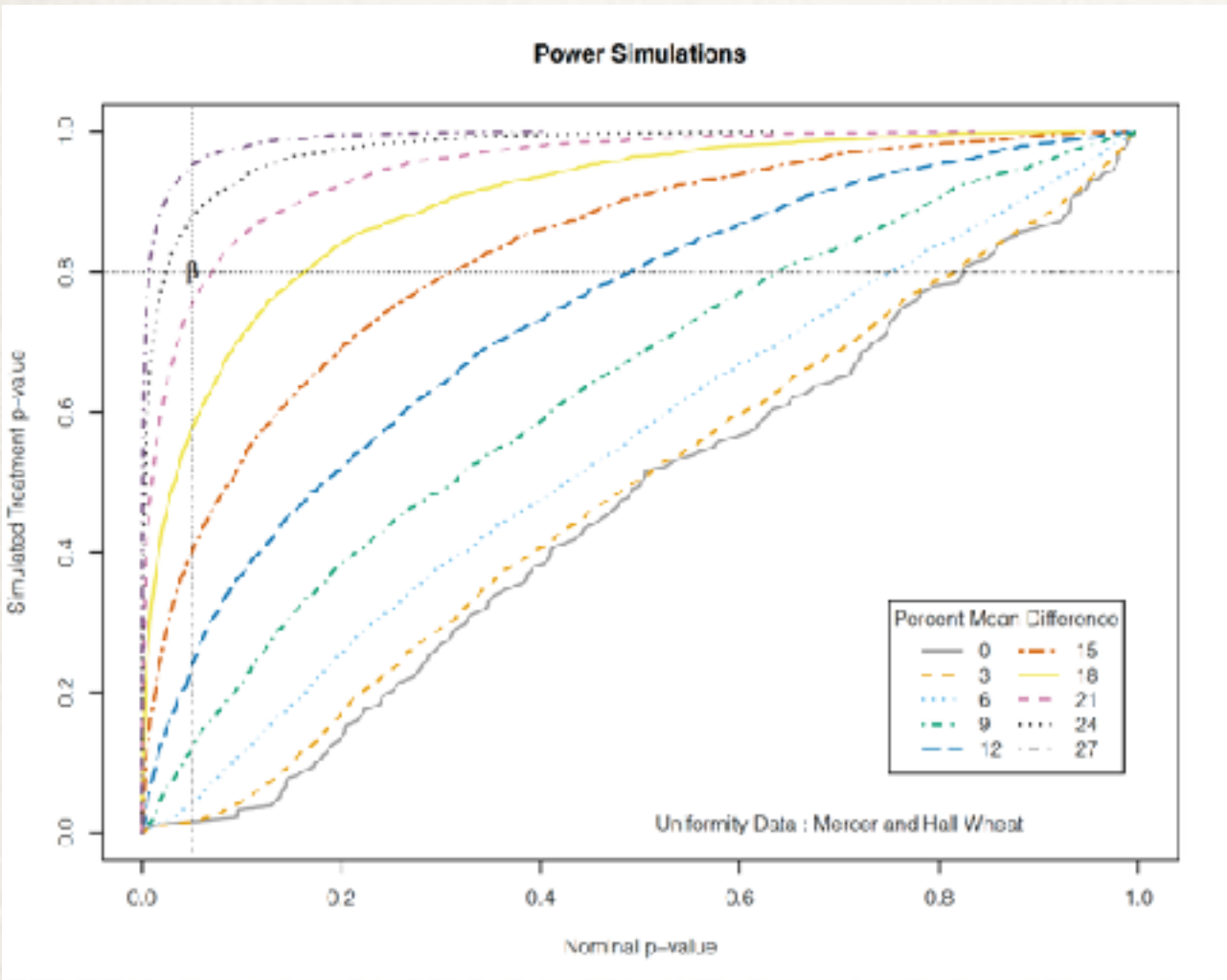
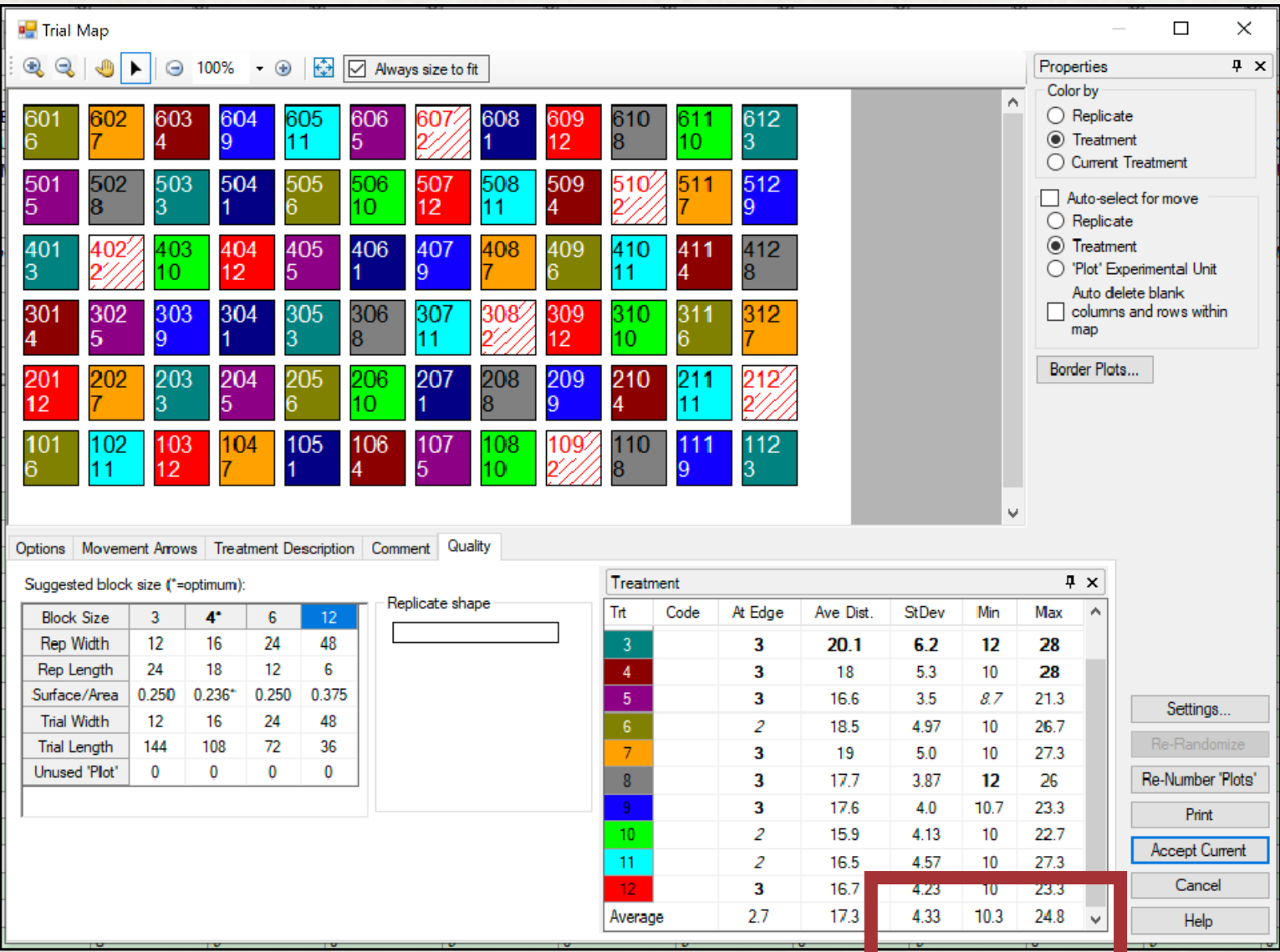
Future Discussions

- ❖ This feature may become a tool for
 - ❖ validating randomization not generated by ARM
 - ❖ visualizing the uniformity of randomizations across multiple trials
 - ❖ understanding restricted randomizations (adjacency, spatially balanced designs)



Trial Quality

- ❖ Replicate Shape
- ❖ Treatment Dispersion

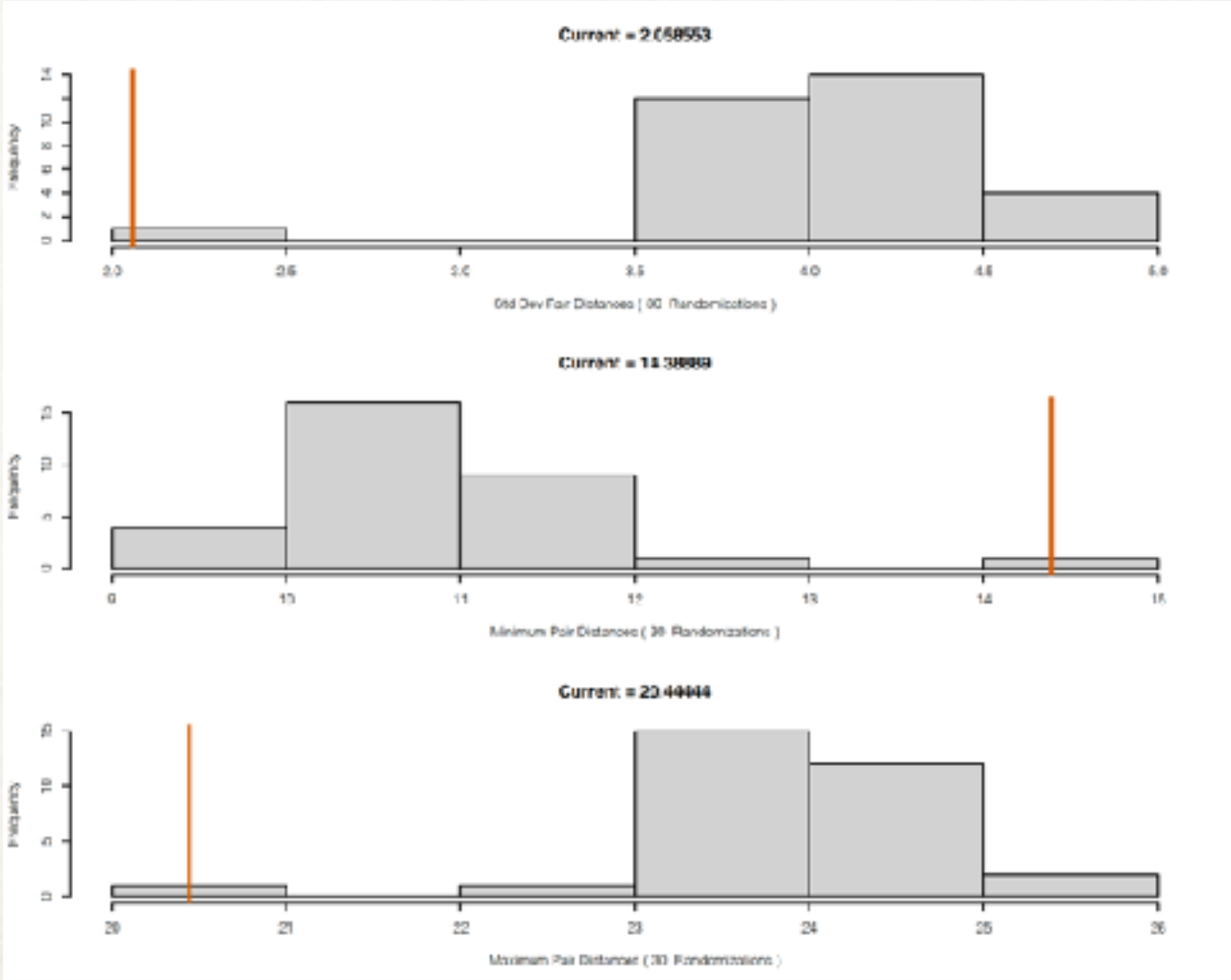
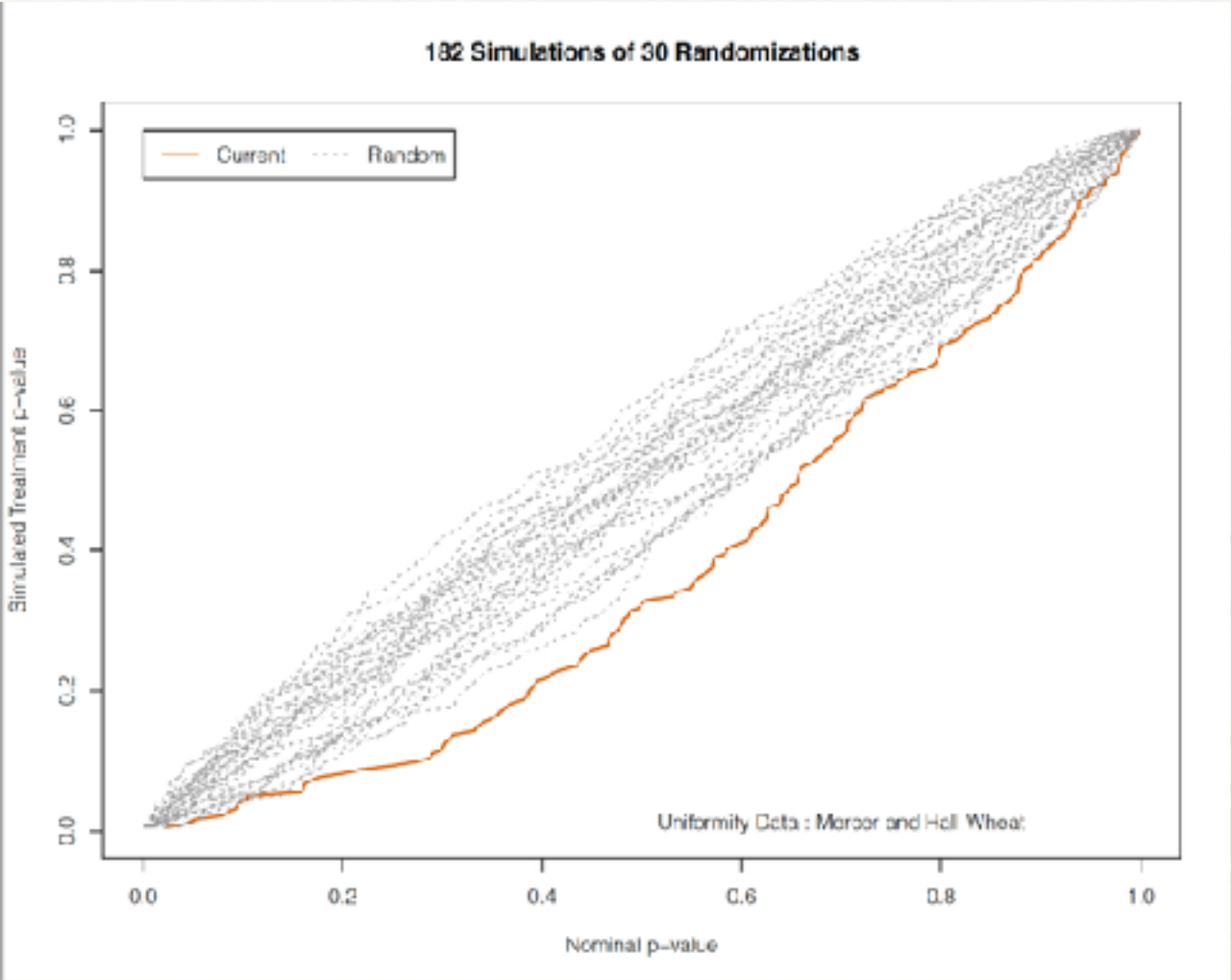
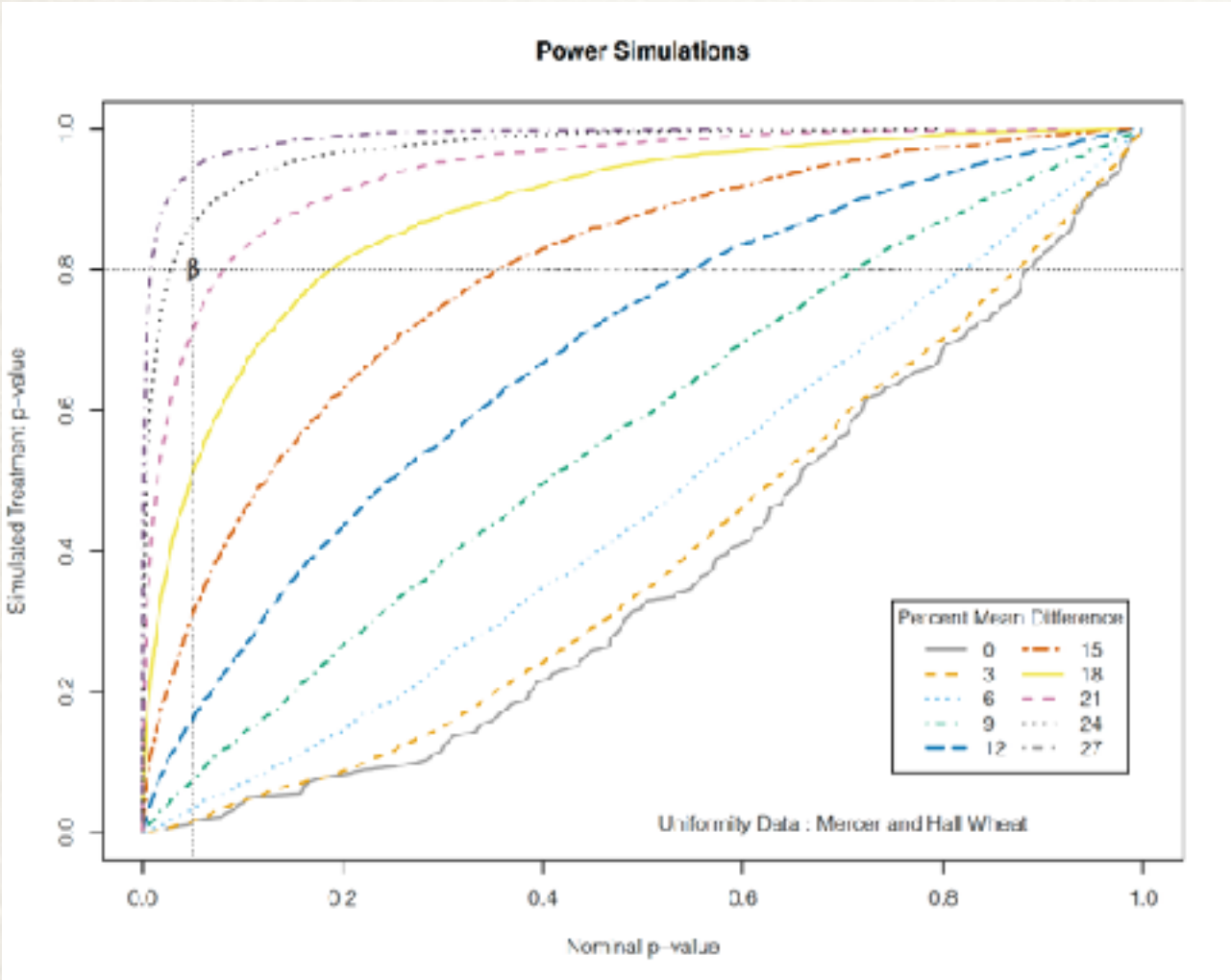
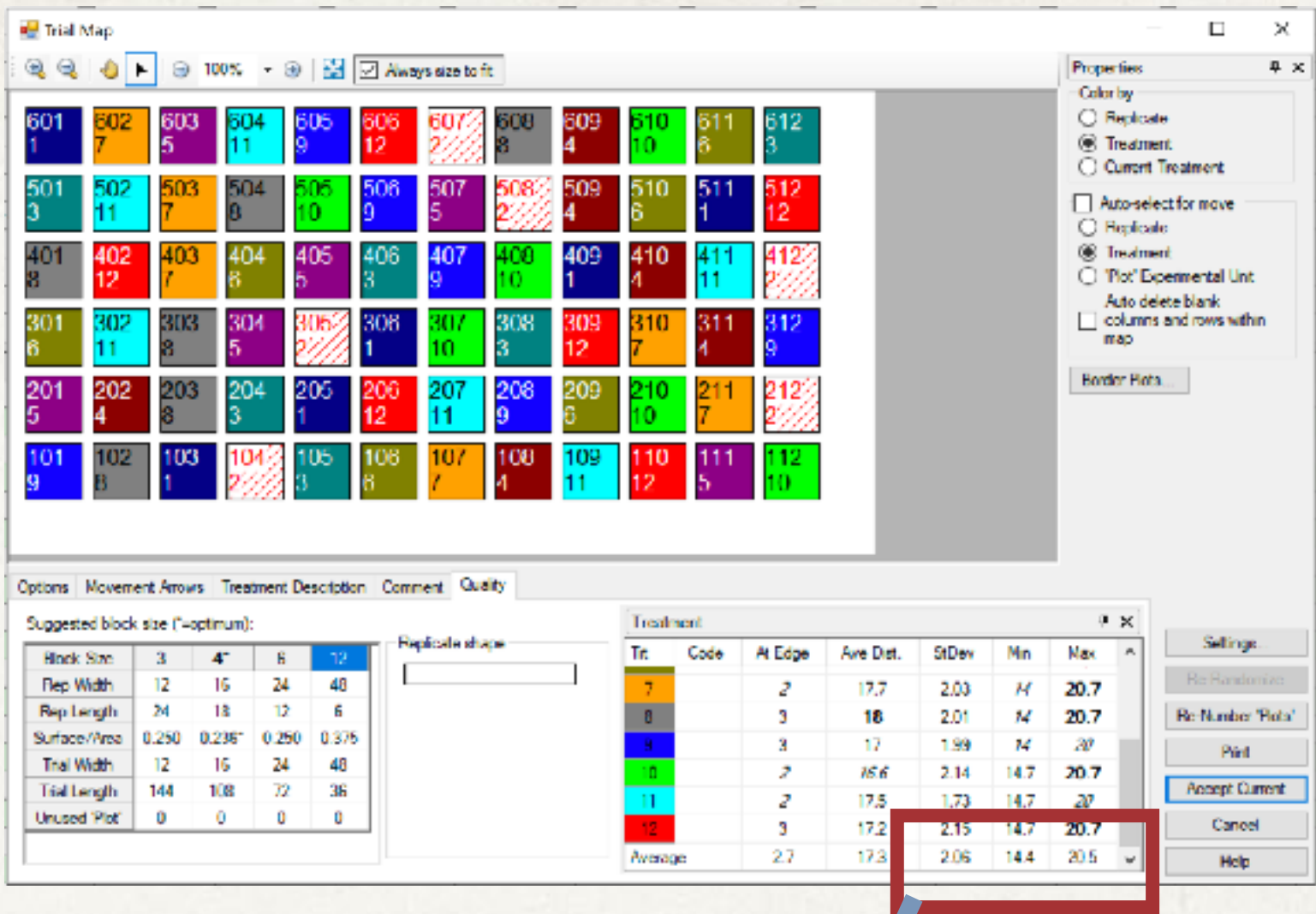


Trial Quality

- ❖ Replicate Shape
- ❖ Treatment Dispersion

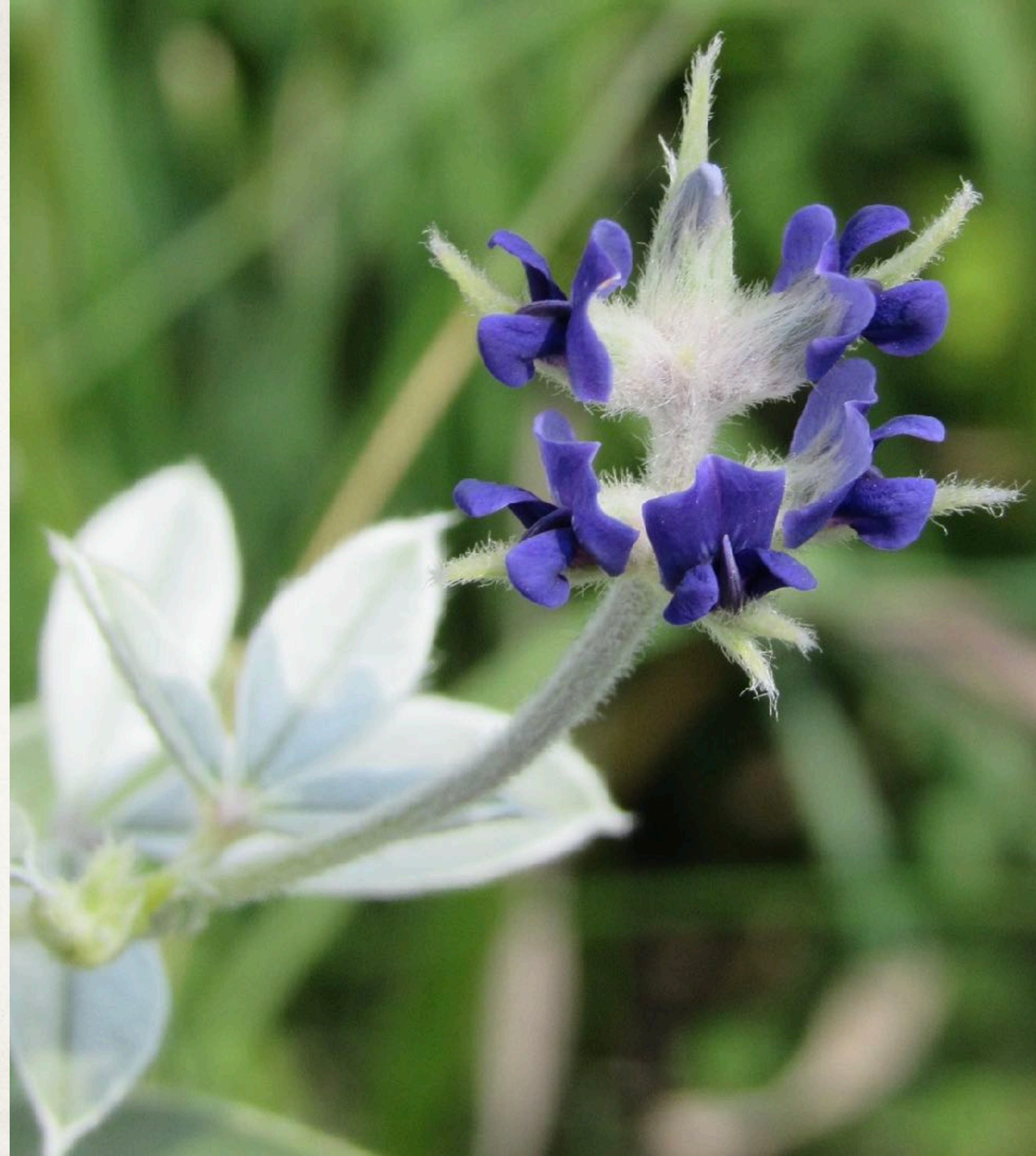
Extended design settings

☒ Spatially balance treatments (for maximum 100 treatments and 20 replicates)



Thanks and Good-day

For now, saving the discussion of ADTC for our next visit.



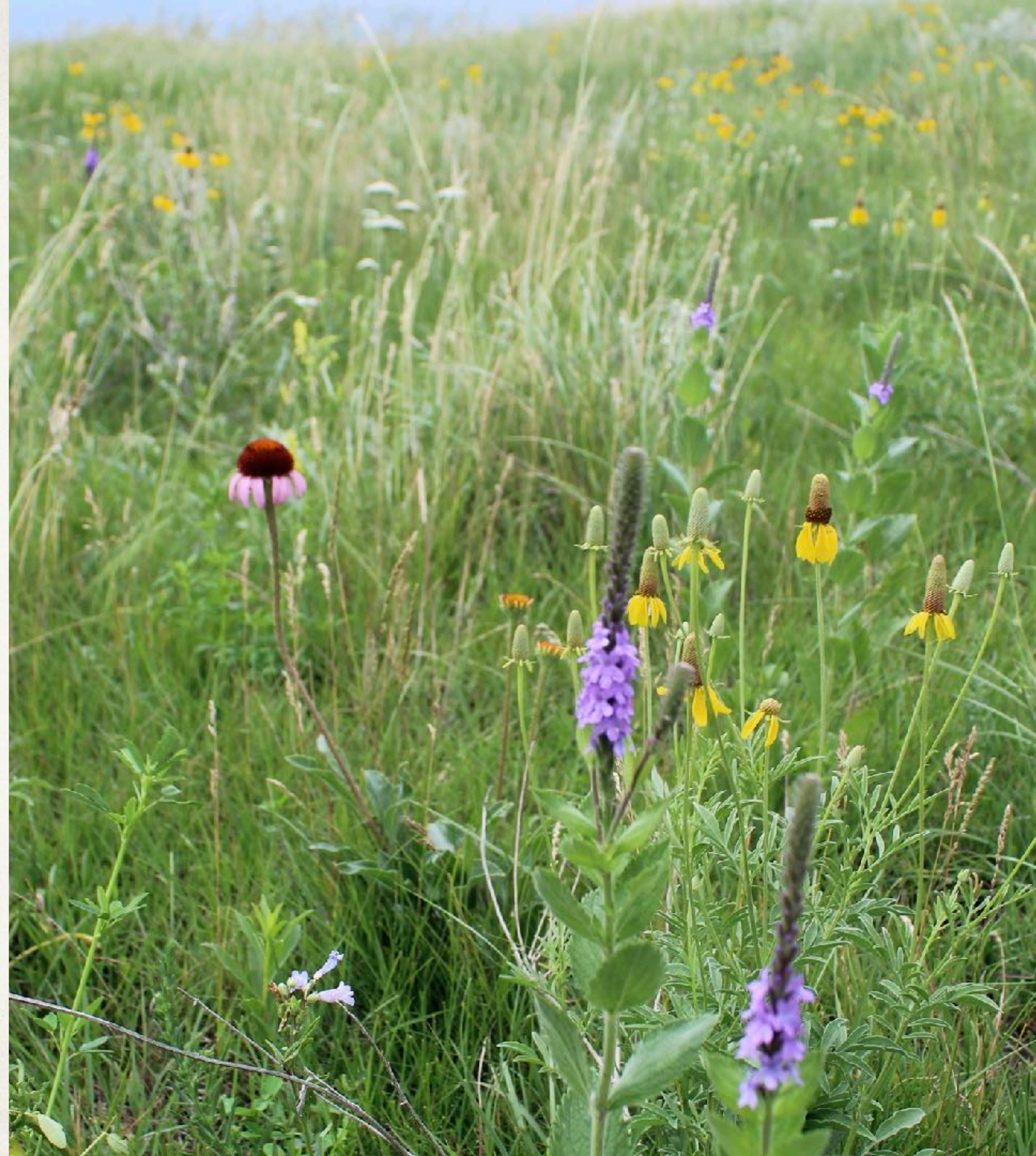
PS

- ❖ All the trial maps shown so far were randomized by ARM.
- ❖ What if we get a trial map that was not created using ARM? Can we check the quality of the randomization?



An ad hoc Experiment

- ❖ Part of this project involves producing *seed bombs* - individual seeds encased in a marble-size portion of dirt/clay mix.
- ❖ Seeds bombs may also include *seed treatments* - fertilizer, biologicals, growth regulators.
- ❖ I want to test the effectiveness of different seed bomb formulations in establishing seedlings without cultivation.



An ad hoc Experiment

- ❖ In one trial, I compared the base seed bomb mix (control) and the seed bomb mix with added commercial product
 - ❖ Dr Earth Flower Girl Organic & Natural Hand Crafted Blend Bud & Bloom Booster
- ❖ I coated 5 commercial varieties of sunflower with both mixes, for a 2x5 factorial design.
- ❖ I didn't carefully count the total number of seeds per treatment and wasn't of the size of my available land, so I didn't plan the trial map until I arrived in the field.

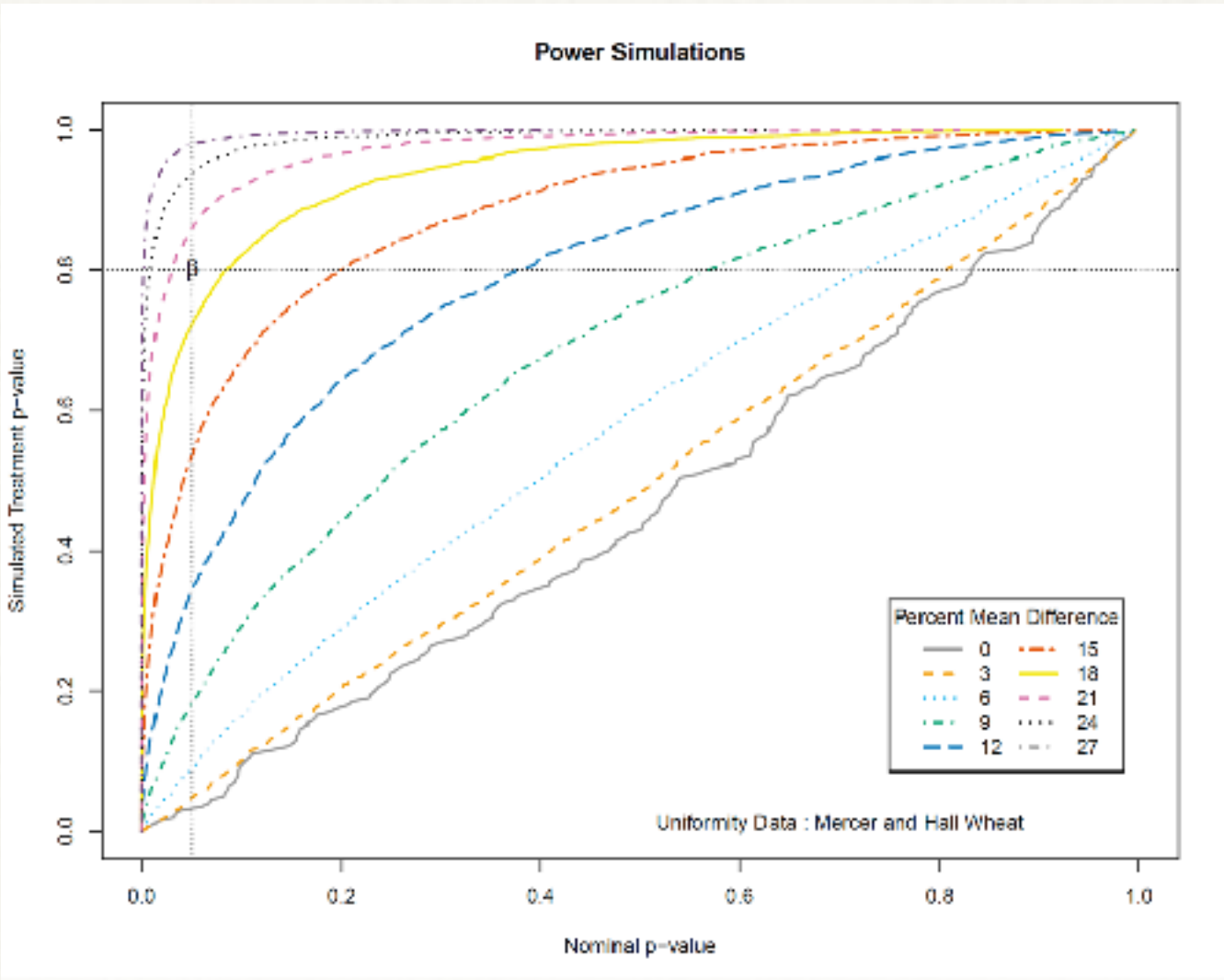
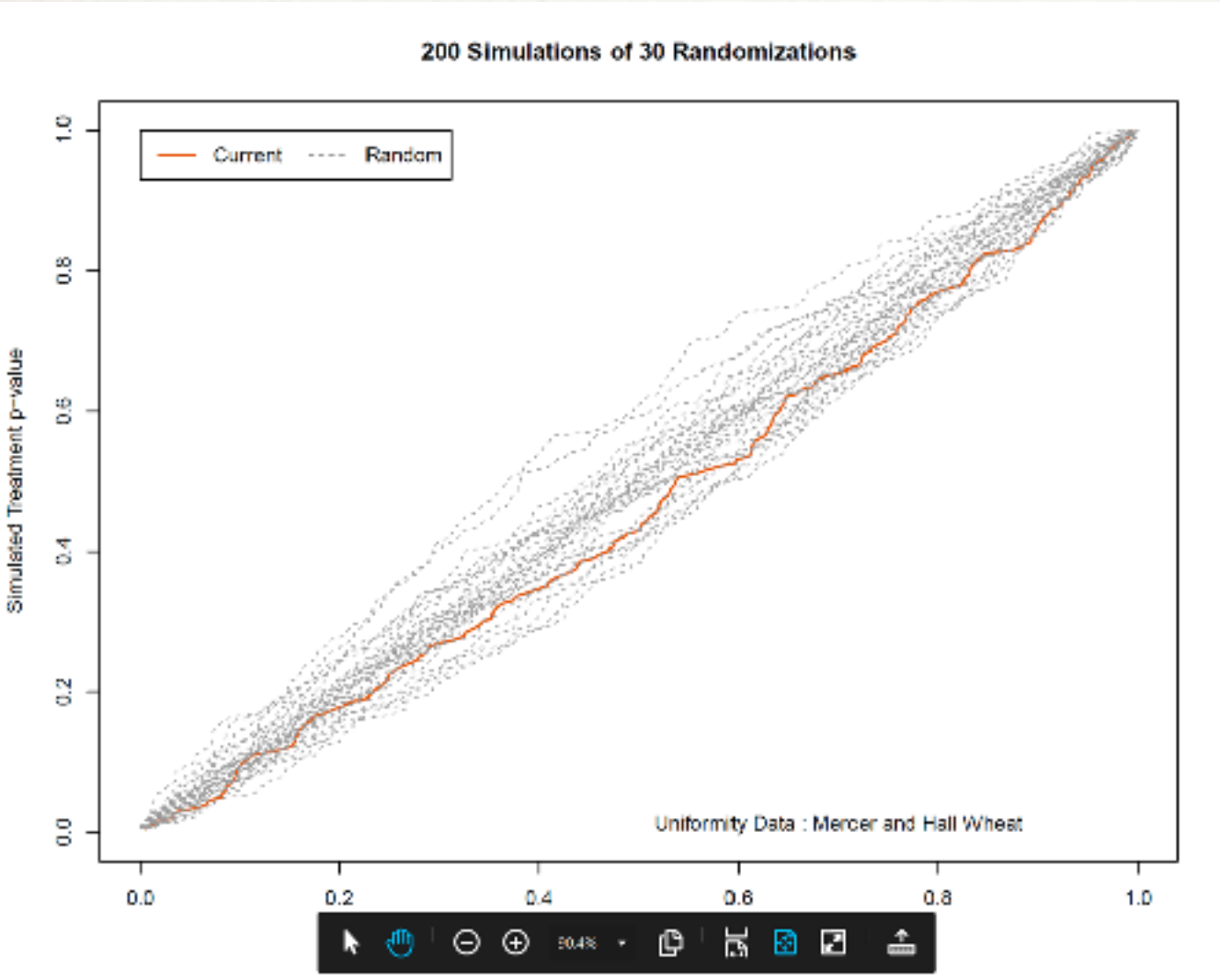


An ad hoc Experiment

- ❖ I finally decided on 5 replicates of 10 plots each, and seeding plots by moving through the trial at random.
- ❖ I then entered the trial map into ARM.
- ❖ How good was my randomization?



Looks good?



110 10	210 3	310 4	410 7	510 1
109 1	209 2	309 5	409 6	509 10
108 8	208 4	308 6	408 9	508 3
107 3	207 1	307 9	407 10	507 5
106 4	206 6	306 2	406 4	506 7
105 5	205 10	305 7	405 3	505 6
104 2	204 9	304 10	404 1	504 8
103 7	203 8	303 3	403 8	503 9
102 9	202 5	302 1	402 2	502 4
101 6	201 7	301 8	401 5	501 2

Could be better?

- ❖ I did not consider dispersion of the most important factor (A - seed bomb treatment)
- ❖ These treatments were not particularly well dispersed.
- ❖ Could I have improved this with planning?

110 10	210 3	310 4	410 7	510 1
109 1	209 2	309 5	409 6	509 10
108 8	208 4	308 6	408 9	508 3
107 3	207 1	307 9	407 10	507 5
106 4	206 6	306 2	406 4	506 7
105 5	205 10	305 7	405 3	505 6
104 2	204 9	304 10	404 1	504 8
103 7	203 8	303 3	403 8	503 9
102 9	202 5	302 1	402 2	502 4
101 6	201 7	301 8	401 5	501 2

Variations on Factorial

- ❖ Consider, for example, a split-plot randomization, with seed bomb treatment as the subplot (B) factor.
- ❖ We lose some accuracy in comparing varieties (now factor A, at the whole plot level), but these comparisons are inherently less interesting.
- ❖ I included varieties mainly to test for interaction with seed treatment, and not for variety main effect.

as Split-Plot

110 6 A3 B2	210 9 A5 B1	310 7 A4 B1	410 6 A3 B2	510 6 A3 B2
109 5 A3 B1	209 10 A5 B2	309 8 A4 B2	409 5 A3 B1	509 5 A3 B1
108 8 A4 B2	208 1 A1 B1	308 5 A3 B1	408 10 A5 B2	508 1 A1 B1
107 7 A4 B1	207 2 A1 B2	307 6 A3 B2	407 9 A5 B1	507 2 A1 B2
106 2 A1 B2	206 3 A2 B1	306 3 A2 B1	406 8 A4 B2	506 9 A5 B1
105 1 A1 B1	205 4 A2 B2	305 4 A2 B2	405 7 A4 B1	505 10 A5 B2
104 4 A2 B2	204 5 A3 B1	304 1 A1 B1	404 4 A2 B2	504 4 A2 B2
103 3 A2 B1	203 6 A3 B2	303 2 A1 B2	403 3 A2 B1	503 3 A2 B1
102 10 A5 B2	202 7 A4 B1	302 9 A5 B1	402 2 A1 B2	502 7 A4 B1
101 9 A5 B1	201 8 A4 B2	301 10 A5 B2	401 1 A1 B1	501 8 A4 B2

as Factorial

110 10 A2 B5	210 3 A1 B3	310 4 A1 B4	410 7 A2 B2	510 1 A1 B1
109 1 A1 B1	209 2 A1 B2	309 5 A1 B5	409 6 A2 B1	509 10 A2 B5
108 8 A2 B3	208 4 A1 B4	308 6 A2 B1	408 9 A2 B4	508 3 A1 B3
107 3 A1 B3	207 1 A1 B1	307 9 A2 B4	407 10 A2 B5	507 5 A1 B5
106 4 A1 B4	206 6 A2 B1	306 2 A1 B2	406 4 A1 B4	506 7 A2 B2
105 5 A1 B5	205 10 A2 B5	305 7 A2 B2	405 10 A1 B3	505 6 A2 B1
104 2 A1 B2	204 9 A2 B4	304 10 A2 B5	404 1 A1 B1	504 8 A2 B3
103 7 A2 B2	203 8 A2 B3	303 3 A1 B3	403 8 A2 B3	503 9 A2 B4
102 9 A2 B4	202 5 A1 B5	302 1 A1 B1	402 2 A1 B2	502 4 A1 B4
101 6 A2 B1	201 7 A2 B2	301 8 A2 B3	401 5 A1 B5	501 2 A1 B2

Future developments

- ❖ Can we apply the treatment dispersion metrics and randomization criteria (adjacency or spatially balanced) to factor levels within factorial designs?



One last picture

