Judges' Nationality Bias In 2017 World Figure Skating Championships

Alexa Jin
Kexin Ding

Hypothesis:

- Is there judge bias in figure skating?
- Does it vary by country?

Background:

Nationality bias is a documented issue in many judged Olympic sports

How figure skating is judged:

- 5 elements per performance, 9 judges per element

- Elements scored from -3 to 3

ISU World Figure Skating Championships 2017
ICE DANCE SHORT DANCE   JUDGES DETAILS PER SKATER

Rank	Name		Nation	Starting Number	Total Segment Score	Total Element Score		Program Component Score (factored)	Total Deductions
1	Tessa VIRTUE / Scott MOIR		CAN	20	82.43	43.74		38.69	0.00

#	Executed Elements	Info	Base Value	GOE	J1	J2	J3	J4	J5	J6	J7	J8	J9	Ref	Scores of Panel
1	NtMiSt4		8.60	2.99	2	3	3	2	3	2	3	3	3		11.59
2	1MB4+kpYYY		5.00	1.37	3	3	2	1	2	2	3	3	1		6.37
3	PSt4		8.60	2.83	2	2	3	1	2	3	3	3	3		11.43
4	STw4		6.60	1.54	2	3	2	2	3	3	3	2	3		8.14
5	CuLi4		4.50	1.71	3	3	3	3	3	2	3	2	3		6.21
			33.30												43.74

Program Components		Factor												
Skating Skills		0.80	9.25	9.75	9.75	9.00	9.75	9.50	9.50	9.75	9.50		9.57
Transitions		0.80	9.25	9.50	9.75	9.25	9.75	9.50	9.50	9.75	9.50		9.54
Performance		0.80	9.50	9.75	10.00	9.50	9.50	9.75	10.00	10.00	9.75		9.75
Composition		0.80	9.50	9.75	10.00	9.50	9.75	9.50	9.75	9.50	10.00		9.68
Interpretation of the Music / Timing		0.80	9.50	10.00	10.00	9.25	9.75	9.75	10.00	9.75	10.00		9.82
Judges Total Program Component Score (factored)														38.69

| Deductions | | | | | | | | | | | | | | 0.00 |

What Sparked Our Suspicion

ISU World Figure Skating Championships 2017
ICE DANCE SHORT DANCE   JUDGES DETAILS PER SKATER

Rank	Name		Nation	Starting Number	Total Segment Score	Total Element Score		Program Component Score (factored)	Total Deductions
1	Tessa VIRTUE / Scott MOIR		CAN	20	82.43	43.74		38.69	0.00

#	Executed Elements	Info	Base Value	GOE	J1	J2	J3	J4	J5	J6	J7	J8	J9	Ref	Scores of Panel
1	NtMiSt4		8.60	2.99	2	3	3	2	3	2	3	3	3		11.59
2	1MB4+kpYYY		5.00	1.37	3	3	2	1	2	2	3	3	1		6.37
3	PSt4		8.60	2.83	2	2	3	1	2	3	3	3	3		11.43
4	STw4		6.60	1.54	2	3	2	2	3	3	3	2	3		8.14
5	CuLi4		4.50	1.71	3	3	3	3	3	2	3	2	3		6.21
			33.30												43.74

Nat	Official	Role
	Sharon ROGERS	Judge 1
	Maira ABASOVA	Judge 2
	Albert ZAYDMAN	Judge 3
	Christine HURTH	Judge 4
	Chihee RHEE	Judge 5
	Richard KOSINA	Judge 6
	Nicole LEBLANC-RICHARD	Judge 7
	Tanay OZKAN SILAOGLU	Judge 8
	Marta OLOZAGARRE	Judge 9

Our data

- BuzzFeed scoring database from GitHub
  - All major international competitions Oct 2016 – Dec 2017
  - Separate .csv files for judges, events, elements, performances
  - Subsetted for 2017 WC, cleaned and merged data into one data frame
- 32 performances x 5 elements x 9 judges = 1,440 data points

	performance_id	player_nation	aspect_id	judge	judge_nation	score
1	d278a3056b	ARM	179a354b86	J1	USA	−1
2	d278a3056b	ARM	499bc0746f	J1	USA	0
3	d278a3056b	ARM	4ba2f37cb4	J1	USA	−1
4	d278a3056b	ARM	9d2b054be3	J1	USA	−1
5	d278a3056b	ARM	aa96646060	J1	USA	0
6	279d0d9ffc	AZE	32ec1ab732	J1	USA	−1

Methodology

- Mean = mean score of the panel
- Standard score = individual score - mean of panel
- AoA: Average of all standard scores of a single judge
- AoM: Average of matched standard scores of a single judge
- DoD=AoM - AoA
- Overall DoD= overall AoM - overall AoA
- Matches between Judge Nation & Athlete Nation

	performance_id	player_nation	aspect_id	judge	judge_nation	score	mean	standard_score
155	c16ce12907	USA	98bef55278	J1	USA	2	1.7777778	0.2222222
156	c16ce12907	USA	a4e260d057	J1	USA	3	2.3333333	0.6666667
157	89ed79b609	USA	a916aa348d	J1	USA	3	2.4444444	0.5555556
158	c16ce12907	USA	a949c4986f	J1	USA	3	2.1111111	0.8888889
159	0729cba3e7	USA	cbd69707ea	J1	USA	3	2.1111111	0.8888889
160	0729cba3e7	USA	d72a54de00	J1	USA	3	2.5555556	0.4444444

Overall Distribution of Standard Scores
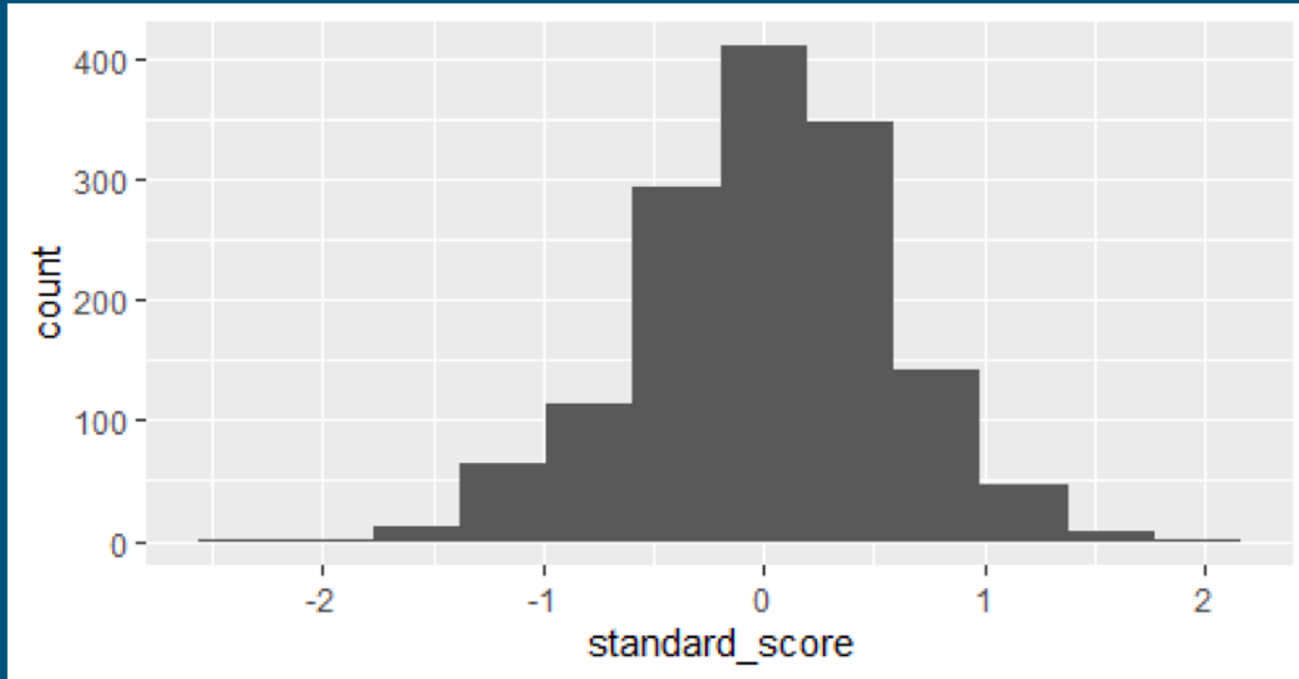
Distribution of Matched VS All Standard Scores
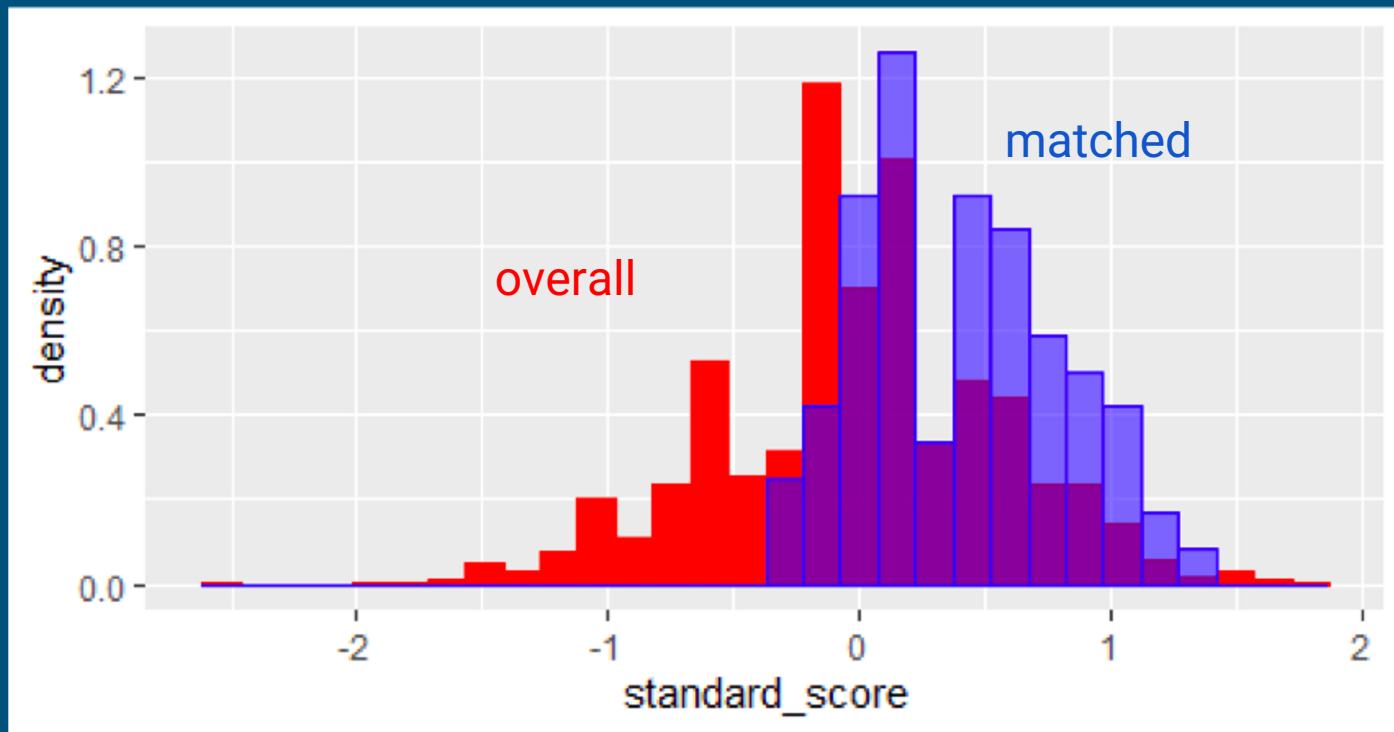
A first look at bias

- Number of match elements: 80
- Average standard score among elements: 0.4152778
- Testing for significance
- 10,000-trial permutation tests

P-value: 0.0459; Overall DoD: 0.415

Permutation test for overall bias
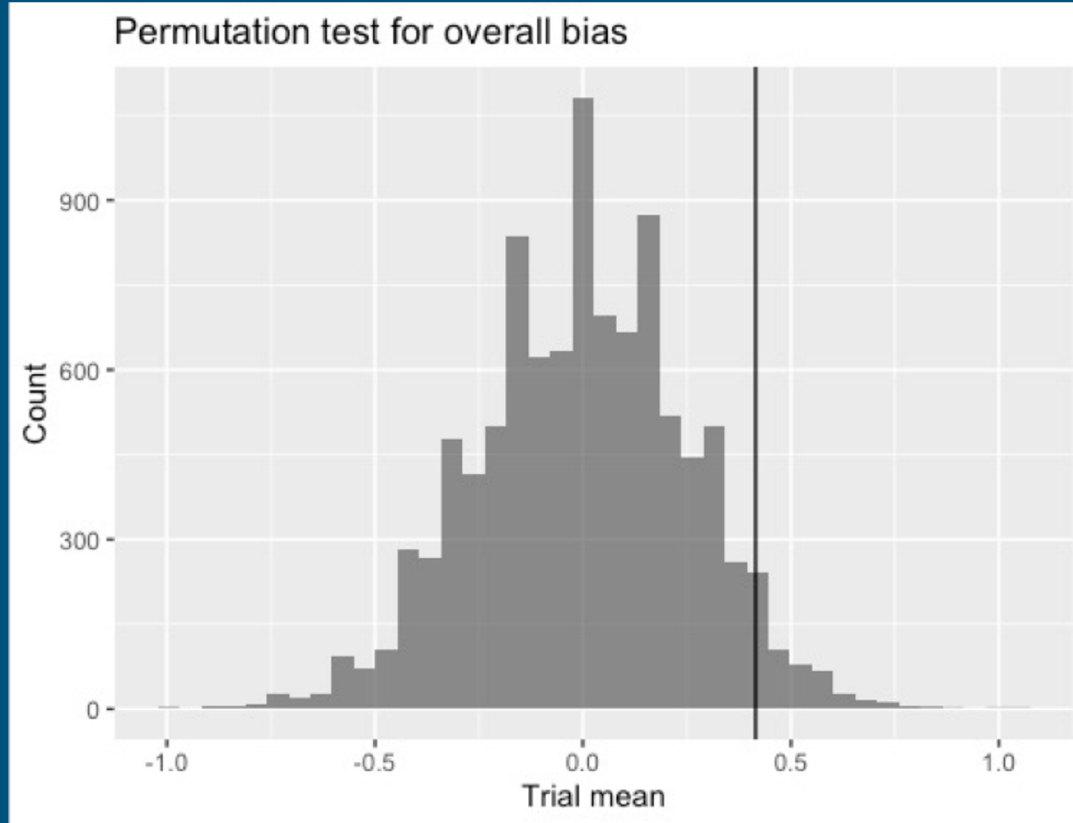
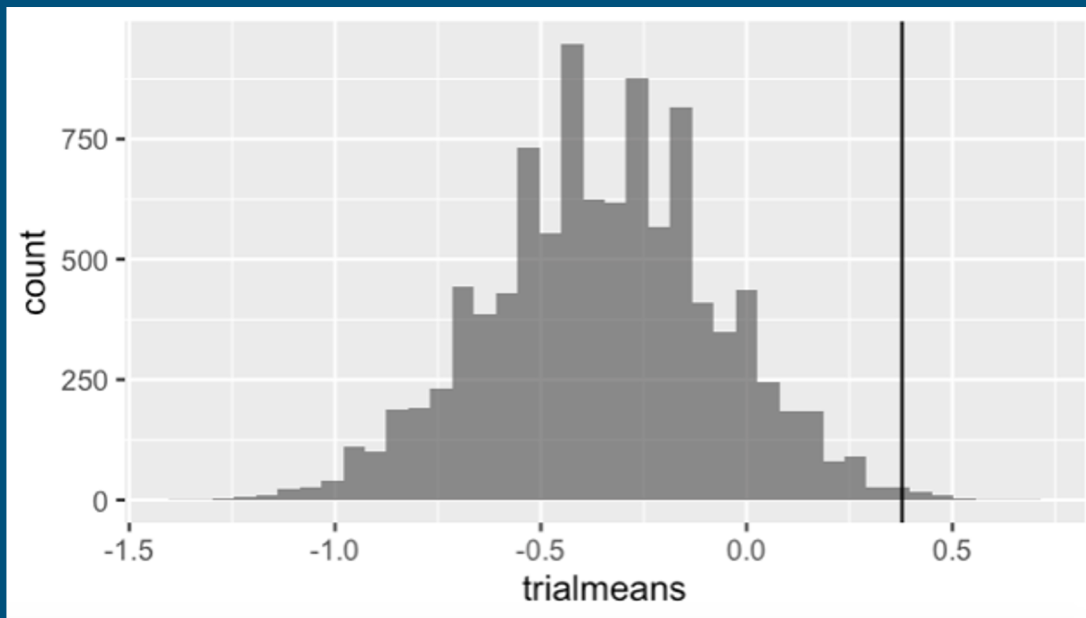# France

- DoD=0.658
- p-value=0.0577
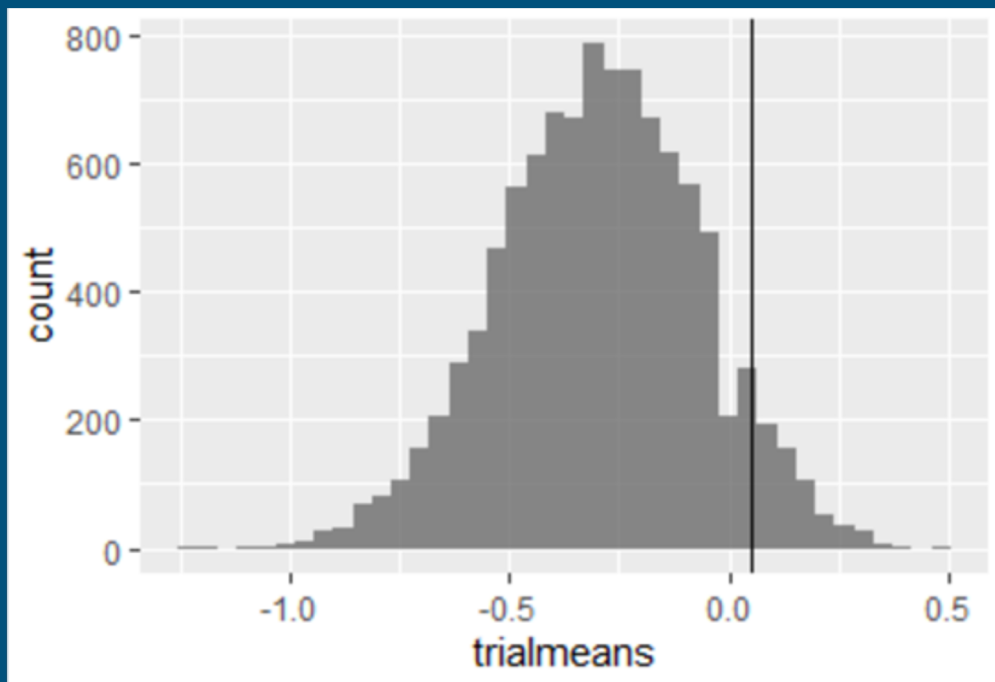
Canada

- DoD=0.545
- p-value=0.0575
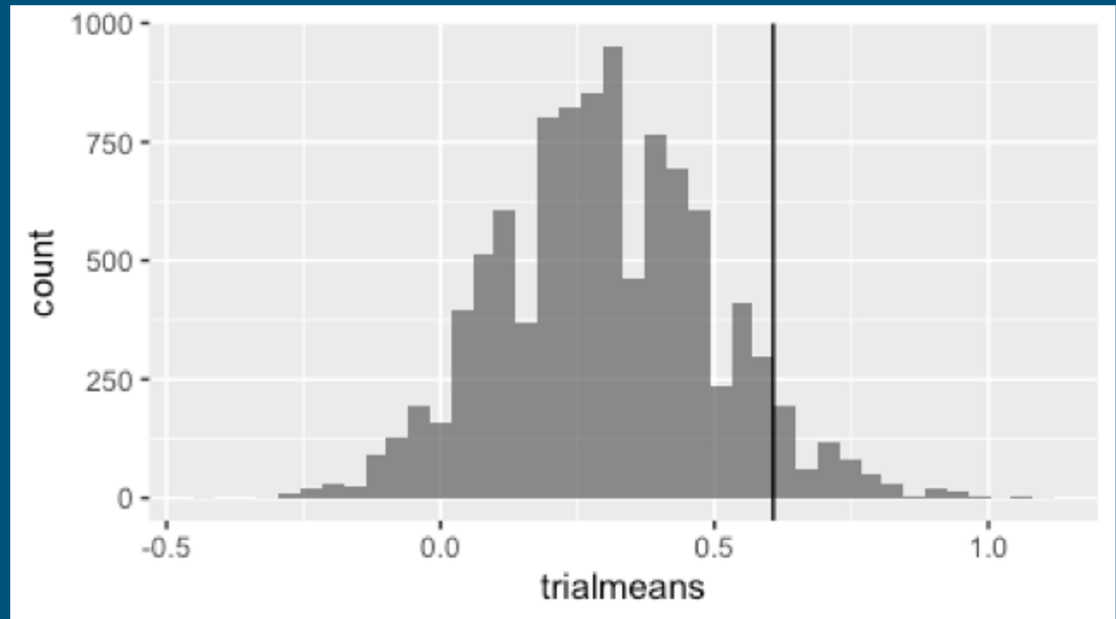
Korea

- DoD=0.439
- p-value= 0.0128
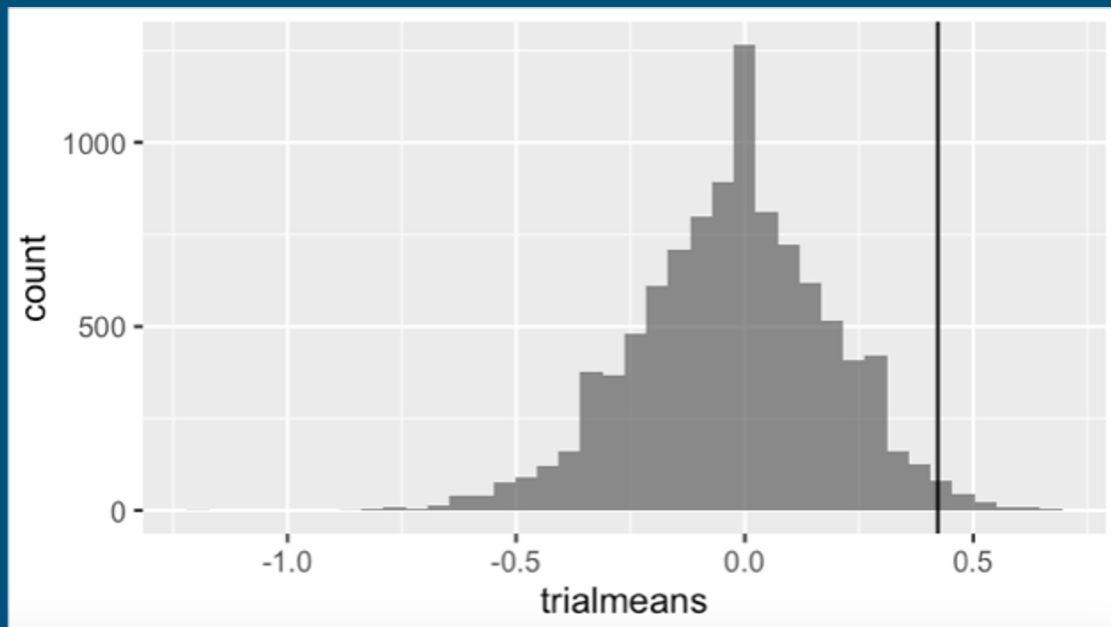
Russia

- DoD=0.234
- p-value=0.007
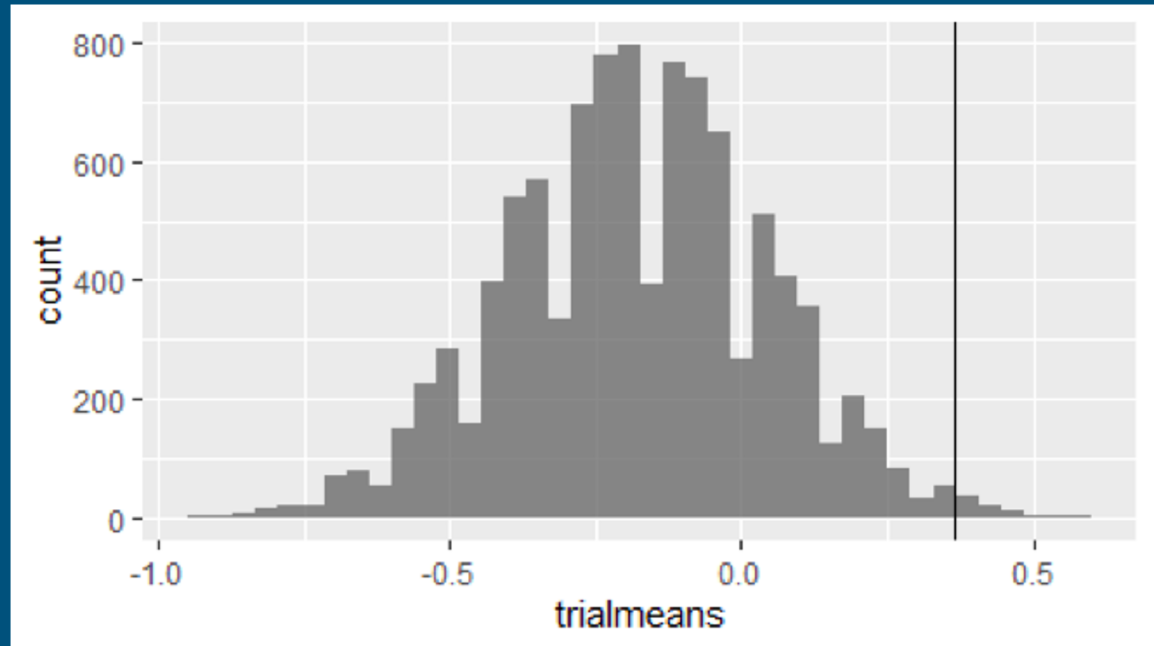
United States

- DoD=0.729
- p-value=0.0035

Conclusion

Judge's nationality	DoD	P-value
🇺🇸	0.7285	0.0035
🇫🇷	0.6579	0.0577
🇨🇦	0.5449	0.0575
🇮🇱	0.4646	0.0175
🇨🇿	0.4615	0.0114
🇰🇷	0.4394	0.0128
🇪🇸	0.4137	0.0338
🇹🇷	0.2581	0.0063
🇷🇺	0.2339	0.007

DoD
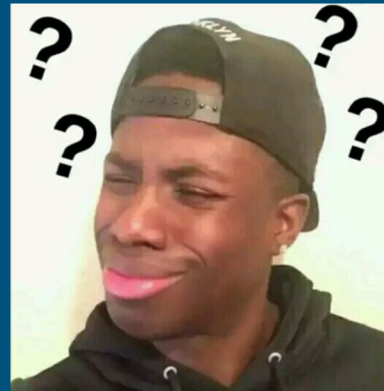
Extension-Practical Significance

Will the overall standing change without nationality bias?

Pl.	Qual.	Name	Nation	TSS =
1	Q	Tessa VIRTUE / Scott MOIR	CAN	82.43
2	Q	Gabriella PAPADAKIS / Guillaume CIZERON	FRA	76.89
3	Q	Madison HUBBELL / Zachary DONOHUE	USA	76.53
4	Q	Madison CHOCK / Evan BATES	USA	76.25
5	Q	Maia SHIBUTANI / Alex SHIBUTANI	USA	74.88
6	Q	Kaitlyn WEAVER / Andrew POJE	CAN	74.84
7	Q	Anna CAPPELLINI / Luca LANOTTE	ITA	73.70
8	Q	Ekaterina BOBROVA / Dmitri SOLOVIEV	RUS	73.54
9	Q	Piper GILLES / Paul POIRIER	CAN	72.83
10	Q	Alexandra STEPANOVA / Ivan BUKIN	RUS	69.07
11	Q	Charlene GUIGNARD / Marco FABBRI	ITA	67.56
12	Q	Isabella TOBIAS / Ilia TKACHENKO	ISR	66.27

Average Influence of nationality bias on the score of a single matched performance:
0.23071

Thank you