

Supplementary Materials

Barbalat G, Reddy G, Franck N. **Association of globalization with the burden of opioid use disorders 2019. A country-level analysis using targeted maximum likelihood estimation.**

Supplementary Table 1. Performance of machine learning algorithms for the exposure mechanism and the outcome regression: example.

Supplementary Table 2. Outlier observations for the outcome variables.

Supplementary Figure 1. Globalization and burden of mental and substance use disorders 2019 after removing outlier observations.

Supplementary Figure 2. Sub-indices of globalization and burden of opioid use disorders 2019 after removing outlier observations.

Supplementary Table 1. Performance of machine learning algorithms for the exposure mechanism and the outcome regression: example^a.

Algorithm	Outcome Regression ^b		Exposure Mechanism ^c	
	Coefficient	Performance (mean CV-R ²)	Coefficient	Performance (mean CV-AUC)
Linear Model	0.31	0.68	0.01	0.73
Stepwise	0.04	0.65	0.23	0.75
LASSO	0.05	0.67	0.03	0.74
MARS	0.14	0.64	0.21	0.76
Random Forest	0.46	0.69	0.52	0.84
SuperLearner ^d	-	0.71	-	0.82

^aIn the analysis where globalization is the exposure and the burden of OUD 2019 is the outcome, trimming probabilities of exposure to the 99.9th percentile and not removing outliers

^bFits the covariates on 2019 OUD DALYs across the four levels of globalization with a SuperLearner ensemble (defined below). Performance is measured by the Cross-Validated R² (20 folds for cross-fitting)

^cFits the covariates on the exposure across the four levels of globalization with a SuperLearner ensemble (defined below). Performance is measured by the multiclass Cross-Validated Area Under the Receiving Operating Curve (20 folds for cross-fitting)

^dDefined as a weighted linear combination (that would sum up to 1) of the following basic learners: Linear model with main effects only (Linear Model), Stepwise regression with a step forward procedure (Stepwise), Linear regression with L1-regularization (LASSO), Multivariate adaptive regression splines (MARS), Random Forest. For each algorithm, we used the default hyperparameters of the *SuperLearner* R package

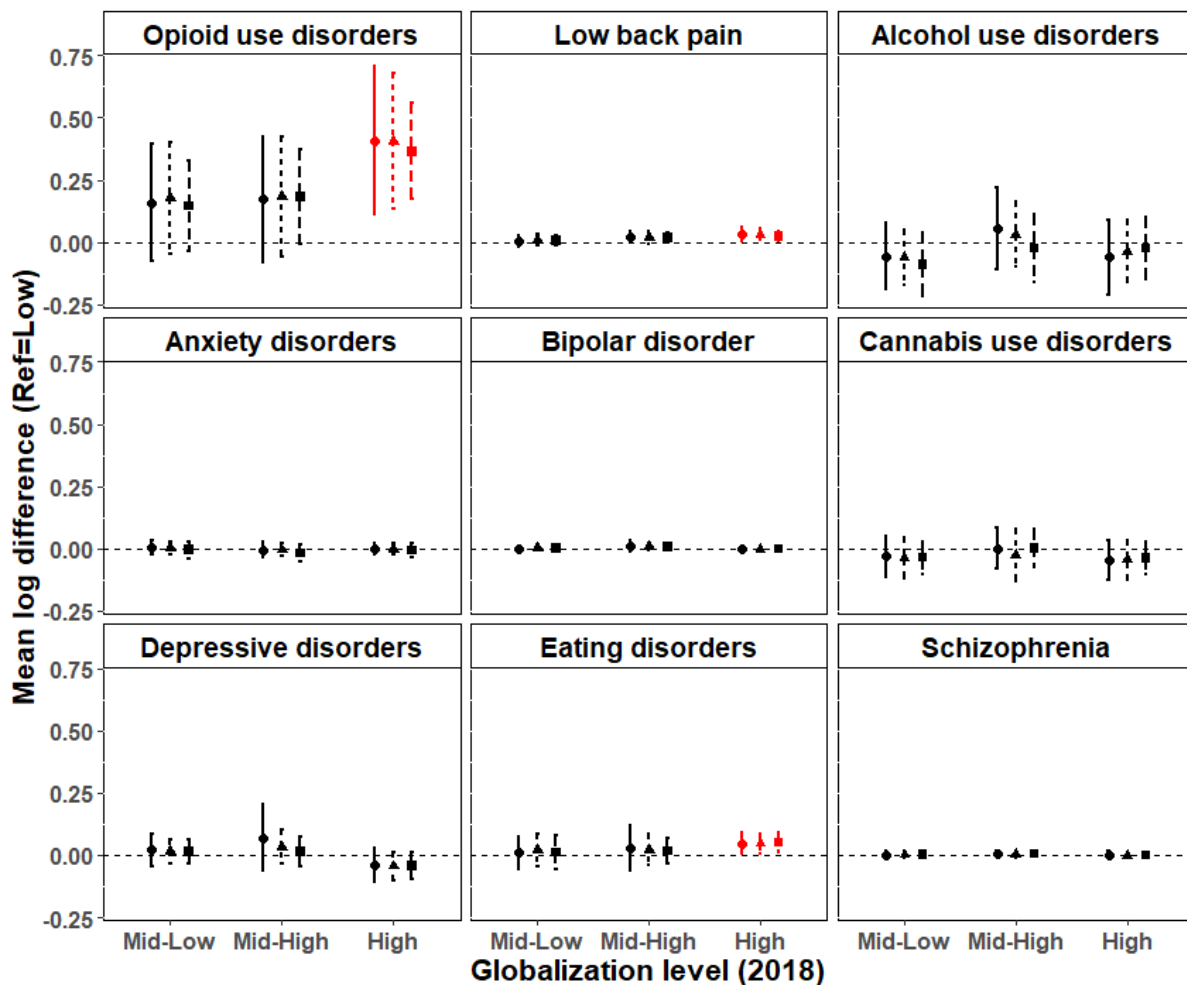
LEGEND. CV: Cross-Validated; AUC: Area Under the receiving operating Curve.

Supplementary Table 2. Outlier observations^a for the outcome variables.

Disorder	Outliers	
	< 1 st pct	> 99 th pct
Opioid use disorders	Singapore (33.8)	United States of America (1481.4)
Low back pain	China (579.1)	United States of America (1402.9)
Alcohol use disorders	Kuwait (46.6)	Belarus (1007.6)
Anxiety disorders	Uzbekistan (194.1)	Portugal (767.7)
Bipolar disorder	China (39.6)	New Zealand (324.1)
Cannabis use disorders	Turkey (2.5)	Canada (35.1)
Depressive disorders	Brunei Darussalam (260.3)	Greece (927.8)
Eating disorders	Tajikistan (19.2)	Australia (218.7)
Schizophrenia	South Africa (141.1)	United States of America (270.7)

^a Expressed as age-standardized Disability Adjusted Life Years rates (per 100000 inhabitants)

Supplementary Figure 1. Globalization and burden of mental and substance use disorders 2019 after removing outlier observations.

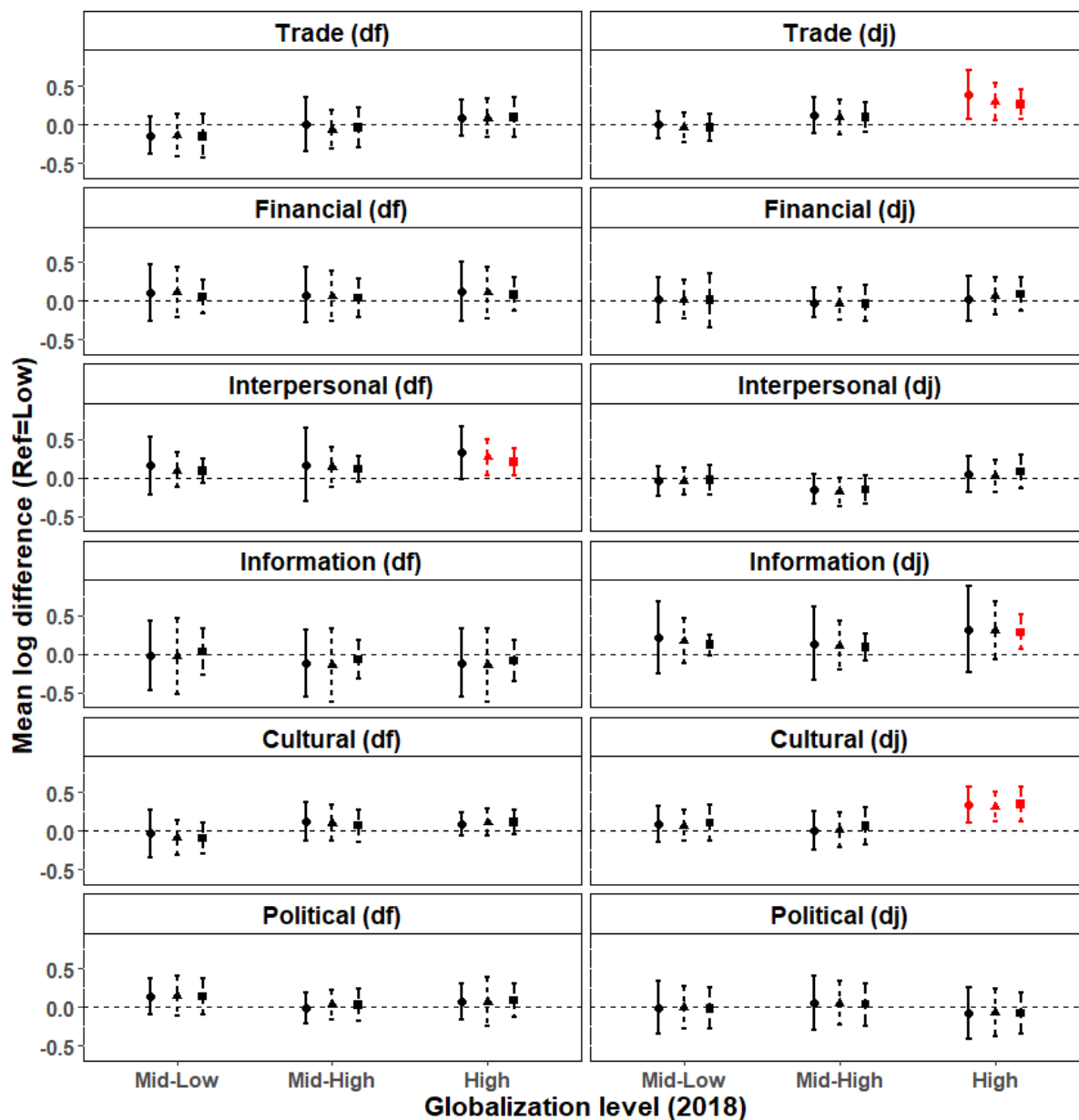


Mean log differences in 2019 DALYs between each globalization level (Mid-Low, Mid-High and High) vs. the reference (Low).

LEGEND. Circles and solid bars: trimming probabilities of exposure to globalization to the 99.9th percentile; triangles and dashed bars: trimming to the 99th percentile; squares and longer dashed bars: trimming to the 97.5th percentile. Error bars denote 95% confidence intervals (CI). Significant differences are shown in red.

df: *de facto*; dj: *de jure*

Supplementary Figure 2. Sub-indices of globalization and burden of opioid use disorders 2019 after removing outlier observations.



Mean log differences in 2019 DALYs between each globalization level (Mid-Low, Mid-High and High) vs. the reference (Low).

LEGEND. Circles and solid bars: trimming probabilities of exposure to globalization to the 99.9th percentile; triangles and dashed bars: trimming to the 99th percentile; squares and longer dashed bars: trimming to the 97.5th percentile. Error bars denote 95% confidence intervals (CI). Significant differences are shown in red.

df: *de facto*; dj: *de jure*