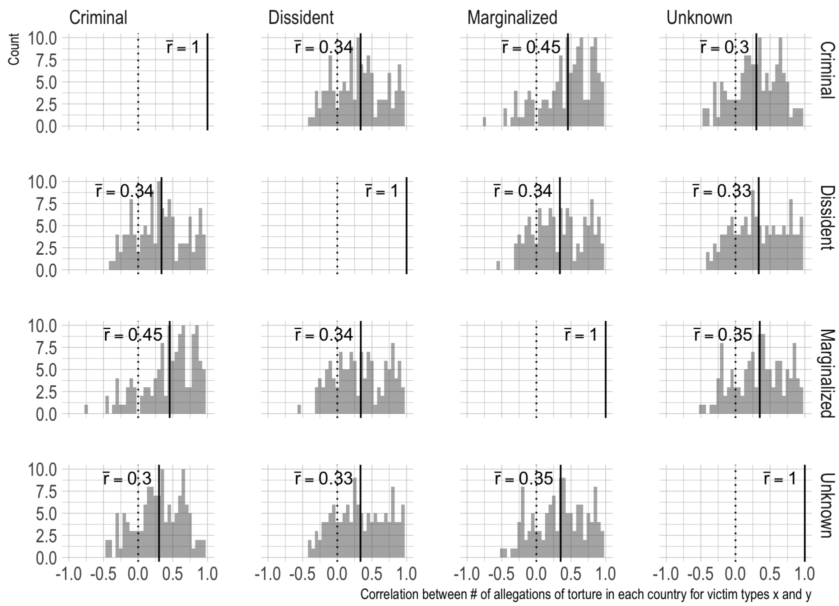


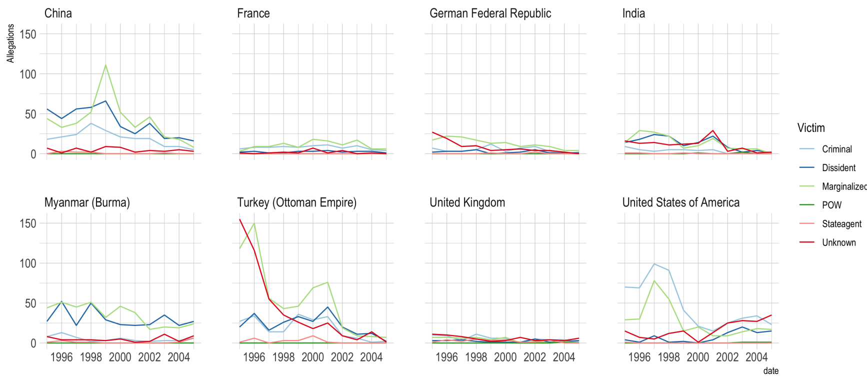
The core models are Poisson mixed effects regression models with state intercepts and three control variables: total population, total GDP, and the ITT restricted access for INGOs indicator. Population and GDP are both logged and then normalized to mean 0 and standard deviation 1. To this basic model we then add each of the variables of interest, one at a time.

$$\hat{Y}_i = \alpha_0 + \alpha_c + \beta_1 \ln \text{Population} + \beta_2 \ln \text{GDP} + \beta_3 \text{ITT\_restricted\_access} + \beta_4 x_j$$

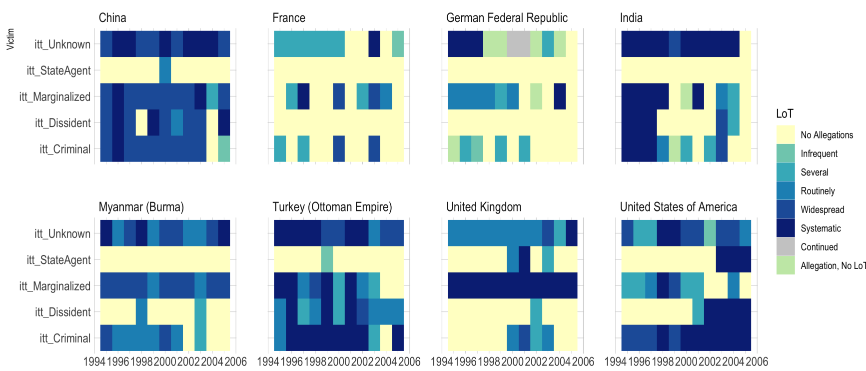
Here i indexes the three outcomes, c states, and j each of the variables of interest. The control model excludes the last term.

We include country random effects because average levels of allegations are related to factors, including wealth/GDP and levels of democracy, that are also related to our variables of interest. Democracies for example seem to face, on average, higher levels of allegations than similar non-democracies, but to us this appears to be a matter of either higher expectations and/or higher transparency and press freedom, not a higher underlying level of ill-treatment and torture.

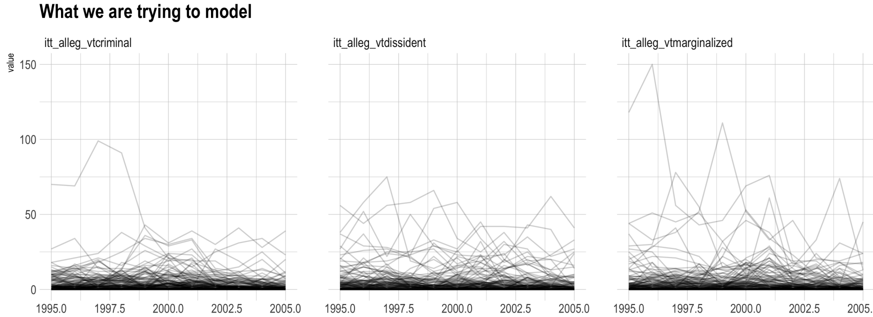
ITT allegations by victim type for select countries

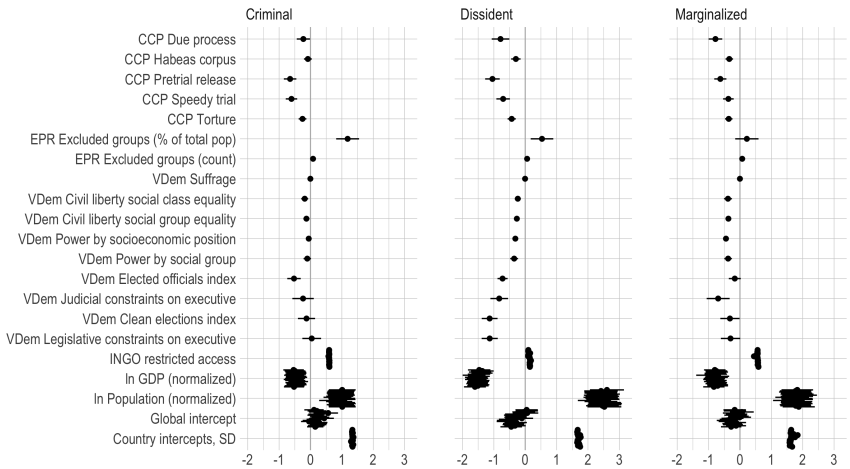


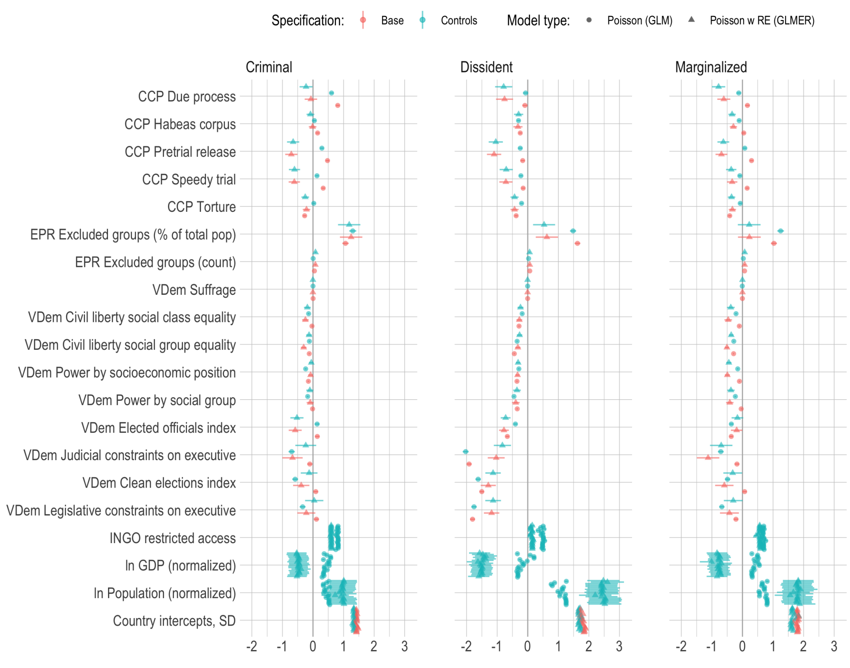
ITT level of torture by victim type for select countries

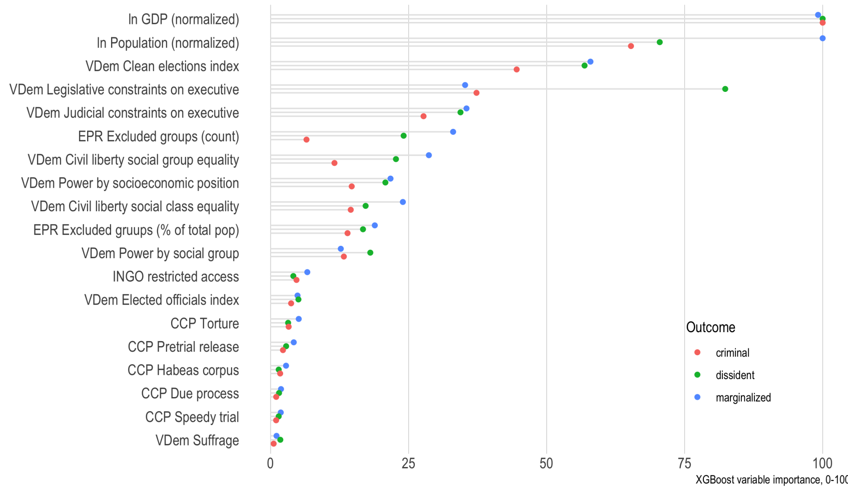


What we are trying to model




Criminal
Dissident
Marginalized
CCP Due process
CCP Habeas corpus
CCP Pretrial release
CCP Speedy trial
CCP Torture
EPR Excluded groups (% of total pop)
EPR Excluded groups (count)
VDem Suffrage
VDem Civil liberty social class equality
VDem Civil liberty social group equality
VDem Power by socioeconomic position
VDem Power by social group
VDem Elected officials index
VDem Judicial constraints on executive
VDem Clean elections index
VDem Legislative constraints on executive
INGO restricted access
ln GDP (normalized)
ln Population (normalized)
Global intercept
Country intercepts, SD
-2 -1 0 1 2 3


Specification: Base Controls Model type: Poisson (GLM) Poisson w RE (GLMER)
Criminal
Dissident
Marginalized
CCP Due process
CCP Habeas corpus
CCP Pretrial release
CCP Speedy trial
CCP Torture
EPR Excluded groups (% of total pop)
EPR Excluded groups (count)
VDem Suffrage
VDem Civil liberty social class equality
VDem Civil liberty social group equality
VDem Power by socioeconomic position
VDem Power by social group
VDem Elected officials index
VDem Judicial constraints on executive
VDem Clean elections index
VDem Legislative constraints on executive
INGO restricted access
ln GDP (normalized)
ln Population (normalized)
Country intercepts, SD
-2 -1 0 1 2 3



Hyperparamaters for the `xgboost` model were picked from a random initial set of hyperparemeters to minimize out-of-sample mean absolute error, based on 11-fold cross-validation.

Fit statistics:

To compare the relative fit of the core models and xgboost, we used cross-validation to obtain out-of-sample predictions from each set of models, and then calculated three fit statistics:

The mean absolute error (MAE) and root mean squared error (RMSE) are based on the absolute and squared deviations of a point prediction from the target value. Both are in the same units as the outcome variables, i.e. allegation counts, and for both lower values indicate better predictive performance. The RMSE penalizes large prediction errors more than the MAE does, i.e. with target value and predicted value pairs of ([10, 11], [100, 110]) where both predicted values are 10% too high, the MAE gives penalties of 1 and 10, while the RMSE penalizes with 1 and 100.

The continuous rank probability score (CRPS) evaluates a probabilistic forecast density by comparing it's cumulative distribution function to that of the target value. For discrete forecasts it reduces to the mean absolute error, and it thus has a similar interpretation. Its units are the same as the outcome variable, i.e. allegation counts, and lower values indicate better fit. For the probabilistic predictions that we have here, the CRPS, unlike MAE and RMSE, scores the complete forecast density, and not just the quality of

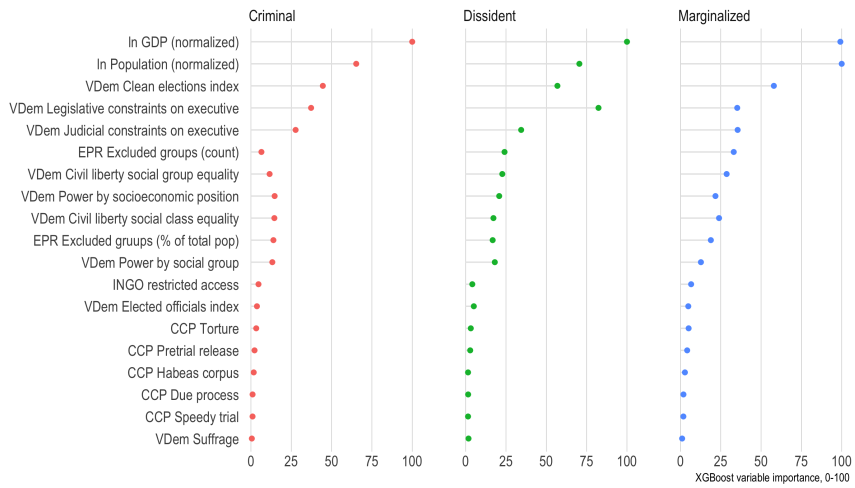
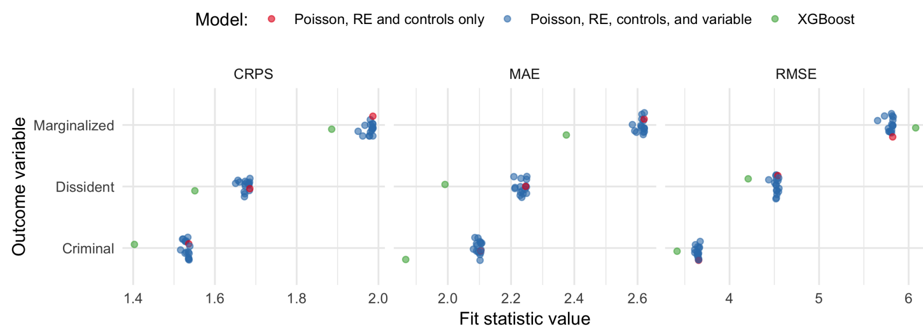


Figure 1: Model fit comparison.



the point forecast.

Table 1:

	Dependent variable:								
	itt_alleg_vtcriminal								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ccp_torture		-0.254^{***}							
		(0.063)							
ccp_prerel			-0.650^{***}						
			(0.102)						
ccp_habcorp				-0.083					
				(0.064)					
ccp_dueproc					-0.227^{**}				
					(0.108)				
ccp_speedtri						-0.606^{***}			
						(0.093)			
v2x_elecoff							-0.526^{***}		
							(0.109)		
v2xel_frefair								-0.127	
								(0.139)	
v2asuffrage									-0.003^{***}
									(0.001)
norm_ln_NY.GDP.MKTP.KD	-0.531^{***}	-0.487^{***}	-0.444^{***}	-0.506^{***}	-0.525^{***}	-0.451^{***}	-0.451^{***}	-0.494^{***}	-0.491^{***}
	(0.164)	(0.159)	(0.163)	(0.163)	(0.164)	(0.161)	(0.165)	(0.169)	(0.164)
norm_ln_pop	1.007^{***}	0.997^{***}	0.953^{***}	1.000^{***}	1.013^{***}	1.014^{***}	0.951^{***}	0.981^{***}	0.995^{***}
	(0.218)	(0.215)	(0.222)	(0.217)	(0.219)	(0.218)	(0.221)	(0.219)	(0.218)
itt_RstrctAccess	0.595^{***}	0.605^{***}	0.592^{***}	0.601^{***}	0.604^{***}	0.599^{***}	0.594^{***}	0.593^{***}	0.592^{***}
	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Constant	0.135	0.276^{**}	0.240^{*}	0.183	0.152	0.264^{**}	0.564^{***}	0.203	0.437^{***}
	(0.132)	(0.135)	(0.136)	(0.137)	(0.133)	(0.134)	(0.160)	(0.151)	(0.157)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
Log Likelihood	−3,834.972	−3,826.947	−3,814.044	−3,834.140	−3,832.781	−3,812.073	−3,823.631	−3,834.563	−3,828.828
Akaike Inf. Crit.	7,679.944	7,665.893	7,640.088	7,680.280	7,677.562	7,636.145	7,659.262	7,681.127	7,669.655
Bayesian Inf. Crit.	7,706.999	7,698.359	7,672.554	7,712.746	7,710.028	7,668.611	7,691.728	7,713.592	7,702.121

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2:

	Dependent variable:							
	itt_alleg_vtcriminal							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
v2x_jucon	−0.237							
	(0.174)							
v2xlg_legcon		0.040						
		(0.153)						
v2clacjust			−0.184***					
			(0.057)					
v2clsocgrp				−0.129***				
				(0.048)				
v2pepwrses					−0.056			
					(0.040)			
v2pepwrsoc						−0.102*		
						(0.060)		
epr_excluded_groups_count							0.081***	
							(0.018)	
epr_excluded_group_pop								1.184***
								(0.186)
norm_ln_NY.GDP.MKTP.KD	−0.470***	−0.539***	−0.444**	−0.461***	−0.516***	−0.473***	−0.486***	−0.391**
	(0.169)	(0.167)	(0.173)	(0.166)	(0.165)	(0.168)	(0.164)	(0.161)
norm_ln_pop	0.960***	1.012***	0.861***	0.936***	0.984***	0.967***	0.724***	0.915***
	(0.219)	(0.219)	(0.227)	(0.220)	(0.218)	(0.218)	(0.228)	(0.213)
itt_RstrctAccess	0.590***	0.596***	0.595***	0.571***	0.595***	0.601***	0.605***	0.594***
	(0.041)	(0.041)	(0.041)	(0.042)	(0.041)	(0.041)	(0.041)	(0.041)
Constant	0.271	0.112	0.313**	0.257*	0.166	0.204	0.017	−0.045
	(0.165)	(0.158)	(0.146)	(0.140)	(0.134)	(0.138)	(0.136)	(0.131)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
Log Likelihood	−3,834.055	−3,834.939	−3,829.660	−3,831.367	−3,833.977	−3,833.546	−3,824.012	−3,812.880
Akaike Inf. Crit.	7,680.110	7,681.878	7,671.319	7,674.734	7,679.955	7,679.092	7,660.025	7,637.761
Bayesian Inf. Crit.	7,712.576	7,714.343	7,703.785	7,707.199	7,712.420	7,711.558	7,692.491	7,670.227

Note: *p<0.1; **p<0.05; ***p<0.01

Table 3:

	Dependent variable:								
	itt_alleg_vtdissident								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ccp_torture		-0.431^{***}							
		(0.067)							
ccp_prerel			-1.045^{***}						
			(0.118)						
ccp_habcorp				-0.298^{***}					
				(0.076)					
ccp_dueproc					-0.788^{***}				
					(0.143)				
ccp_speedtri						-0.704^{***}			
						(0.110)			
v2x_elecoff							-0.723^{***}		
							(0.082)		
v2xel_frefair								-1.134^{***}	
								(0.129)	
v2asuffrage									-0.004^{***}
									(0.001)
norm_ln_NY.GDP.MKTP.KD	-1.574^{***}	-1.523^{***}	-1.486^{***}	-1.509^{***}	-1.603^{***}	-1.533^{***}	-1.379^{***}	-1.368^{***}	-1.524^{***}
	(0.176)	(0.171)	(0.175)	(0.174)	(0.178)	(0.176)	(0.173)	(0.184)	(0.175)
norm_ln_pop	2.475^{***}	2.534^{***}	2.526^{***}	2.462^{***}	2.523^{***}	2.548^{***}	2.410^{***}	2.416^{***}	2.447^{***}
	(0.270)	(0.266)	(0.273)	(0.266)	(0.275)	(0.272)	(0.266)	(0.273)	(0.269)
itt_RstrctAccess	0.147^{***}	0.140^{***}	0.140^{***}	0.156^{***}	0.149^{***}	0.151^{***}	0.157^{***}	0.093^{**}	0.149^{***}
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)	(0.043)	(0.042)
Constant	-0.504^{***}	-0.283^{*}	-0.378^{**}	-0.326^{*}	-0.446^{**}	-0.368^{**}	0.064	0.052	−0.108
	(0.172)	(0.171)	(0.174)	(0.175)	(0.176)	(0.174)	(0.180)	(0.181)	(0.183)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
Log Likelihood	−3,855.179	−3,834.206	−3,813.040	−3,847.519	−3,839.089	−3,833.308	−3,818.482	−3,816.788	−3,838.051
Akaike Inf. Crit.	7,720.357	7,680.413	7,638.080	7,707.039	7,690.179	7,678.617	7,648.964	7,645.575	7,688.103
Bayesian Inf. Crit.	7,747.412	7,712.878	7,670.546	7,739.504	7,722.645	7,711.082	7,681.429	7,678.041	7,720.569

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4:

	Dependent variable:							
	itt_alleg_vtdissident							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
v2x_jucon	−0.829***							
	(0.143)							
v2xlg_legcon		−1.132***						
		(0.131)						
v2clacjust			−0.237***					
			(0.050)					
v2clsocgrp				−0.267***				
				(0.046)				
v2pepwrses					−0.315***			
					(0.036)			
v2pepwrsoc						−0.353***		
						(0.064)		
epr_excluded_groups_count							0.061***	
							(0.014)	
epr_excluded_group_pop								0.535***
								(0.184)
norm_ln_NY.GDP.MKTP.KD	−1.441***	−1.467***	−1.560***	−1.480***	−1.629***	−1.473***	−1.480***	−1.502***
	(0.181)	(0.182)	(0.184)	(0.179)	(0.186)	(0.183)	(0.174)	(0.176)
norm_ln_pop	2.411***	2.605***	2.405***	2.456***	2.411***	2.486***	2.207***	2.414***
	(0.268)	(0.277)	(0.279)	(0.272)	(0.276)	(0.271)	(0.271)	(0.267)
itt_RstrctAccess	0.104**	0.100**	0.142***	0.091**	0.093**	0.139***	0.184***	0.159***
	(0.043)	(0.043)	(0.042)	(0.043)	(0.043)	(0.042)	(0.042)	(0.042)
Constant	−0.048	0.046	−0.309*	−0.297*	−0.332*	−0.306*	−0.574***	−0.575***
	(0.185)	(0.182)	(0.182)	(0.175)	(0.177)	(0.172)	(0.169)	(0.171)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
Log Likelihood	−3,838.513	−3,817.959	−3,843.958	−3,837.750	−3,816.799	−3,839.851	−3,845.627	−3,850.900
Akaike Inf. Crit.	7,689.026	7,647.919	7,699.916	7,687.500	7,645.597	7,691.702	7,703.254	7,713.801
Bayesian Inf. Crit.	7,721.491	7,680.385	7,732.382	7,719.966	7,678.063	7,724.168	7,735.720	7,746.266

Note: *p<0.1; **p<0.05; ***p<0.01

Table 5:

	Dependent variable:								
	itt_alleg_vtmarginalized								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ccp_torture		-0.357^{***}							
		(0.061)							
ccp_prerel			-0.625^{***}						
			(0.099)						
ccp_habcorp				-0.340^{***}					
				(0.061)					
ccp_dueproc					-0.782^{***}				
					(0.111)				
ccp_speedtri						-0.369^{***}			
						(0.086)			
v2x_elecoff							-0.165^{*}		
							(0.097)		
v2xel_frefair								-0.320^{**}	
								(0.156)	
v2asuffrage									-0.001
									(0.001)
norm_ln_NY.GDP.MKTP.KD	-0.833^{***}	-0.738^{***}	-0.714^{***}	-0.710^{***}	-0.831^{***}	-0.763^{***}	-0.783^{***}	-0.754^{***}	-0.813^{***}
	(0.173)	(0.166)	(0.170)	(0.168)	(0.174)	(0.171)	(0.174)	(0.178)	(0.173)
norm_ln_pop	1.824^{***}	1.790^{***}	1.763^{***}	1.777^{***}	1.874^{***}	1.818^{***}	1.803^{***}	1.774^{***}	1.817^{***}
	(0.256)	(0.250)	(0.253)	(0.250)	(0.261)	(0.253)	(0.254)	(0.257)	(0.255)
itt_RstrctAccess	0.564^{***}	0.577^{***}	0.563^{***}	0.583^{***}	0.590^{***}	0.573^{***}	0.566^{***}	0.562^{***}	0.565^{***}
	(0.038)	(0.038)	(0.038)	(0.039)	(0.039)	(0.038)	(0.038)	(0.038)	(0.038)
Constant	-0.327^{**}	-0.123	-0.225	-0.125	-0.280^{*}	-0.244	-0.194	-0.160	-0.238
	(0.163)	(0.162)	(0.162)	(0.163)	(0.166)	(0.162)	(0.179)	(0.182)	(0.179)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
Log Likelihood	−4,203.589	−4,186.348	−4,182.659	−4,188.072	−4,178.377	−4,194.042	−4,202.182	−4,201.535	−4,202.937
Akaike Inf. Crit.	8,417.178	8,384.697	8,377.317	8,388.144	8,368.754	8,400.084	8,416.364	8,415.071	8,417.874
Bayesian Inf. Crit.	8,444.233	8,417.162	8,409.783	8,420.609	8,401.220	8,432.550	8,448.829	8,447.536	8,450.340
Note:									*p<0.1; **p<0.05; ***p<0.01

Table 6:

	Dependent variable:							
	itt_alleg_vtmarginalized							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
v2x_jucon	-0.696^{***}							
	(0.187)							
v2xlg_legcon		-0.301^{*}						
		(0.157)						
v2clacjust			-0.382^{***}					
			(0.065)					
v2clsocgrp				-0.369^{***}				
				(0.039)				
v2pepwrses					-0.446^{***}			
					(0.044)			
v2pepwrsoc						-0.378^{***}		
						(0.062)		
epr_excluded_groups_count							0.072^{***}	
							(0.017)	
epr_excluded_group_pop								0.220
								(0.190)
norm_ln_NY.GDP.MKTP.KD	-0.732^{***}	-0.802^{***}	-0.751^{***}	-0.692^{***}	-1.012^{***}	-0.772^{***}	-0.757^{***}	-0.805^{***}
	(0.179)	(0.175)	(0.187)	(0.178)	(0.195)	(0.184)	(0.171)	(0.174)
norm_ln_pop	1.753^{***}	1.828^{***}	1.664^{***}	1.814^{***}	1.899^{***}	1.820^{***}	1.560^{***}	1.807^{***}
	(0.258)	(0.257)	(0.274)	(0.263)	(0.283)	(0.269)	(0.258)	(0.255)
itt_RstrctAccess	0.553^{***}	0.562^{***}	0.561^{***}	0.445^{***}	0.516^{***}	0.563^{***}	0.576^{***}	0.565^{***}
	(0.039)	(0.038)	(0.039)	(0.041)	(0.039)	(0.039)	(0.038)	(0.038)
Constant	0.060	−0.168	−0.005	−0.050	−0.139	−0.108	-0.425^{***}	-0.359^{**}
	(0.193)	(0.183)	(0.182)	(0.167)	(0.183)	(0.174)	(0.161)	(0.164)
Observations	1,654	1,654	1,654	1,654	1,654	1,654	1,654	1,654
Log Likelihood	−4,196.657	−4,201.765	−4,185.625	−4,156.688	−4,148.376	−4,184.121	−4,194.799	−4,202.917
Akaike Inf. Crit.	8,405.314	8,415.529	8,383.249	8,325.377	8,308.752	8,380.243	8,401.599	8,417.834
Bayesian Inf. Crit.	8,437.779	8,447.995	8,415.715	8,357.842	8,341.218	8,412.708	8,434.064	8,450.300
Note:								*p<0.1; **p<0.05; ***p<0.01