Appendix:
Soft governance against Superbugs:
How effective is the international regime against antimicrobial resistance?

Table of Contents

1. Descriptive statistics

Table A1: descriptive statistics (time-varying models)

Variables	N	Mean	SD	Min	Max
NAP approved (participation)	3364	0.035	0.183	0	1
Antibiotic consumption	3629	12.574	7.044	2.800	45.900
Implementing NAP	3629	0.072	0.258	0	1
Adopted NAP	3629	0.034	0.180	0	1
Population (log)	3043	15.996	1.688	11.304	21.050
Government health expenditure (share of total)	3428	0.500	0.216	0.023	0.947
Bureaucratic capacity	3154	0.416	1.202	-2.610	3.600
GDPpc (log)	3420	8.426	1.534	4.488	12.163
Economic growth	3229	1.039	0.066	0.387	2.247
Democracy	3219	0.536	0.262	0.014	0.948
Regional peers	3629	0.105	0.175	0	1

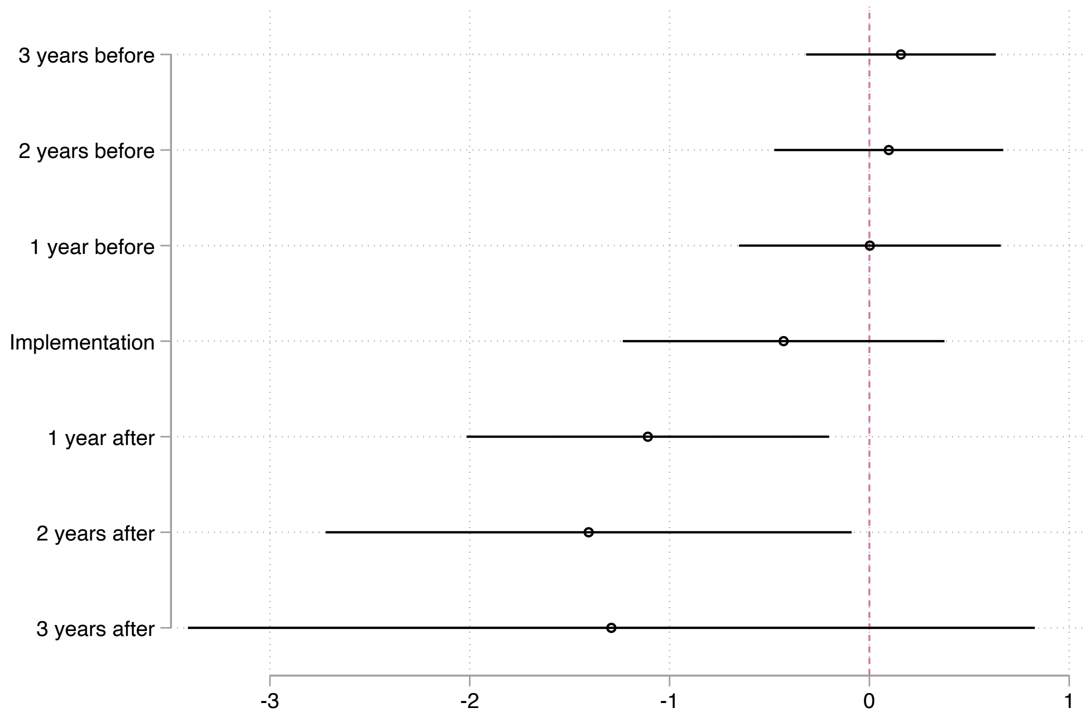
Table A2: descriptive statistics (time-invariant models)

Variables	N	Mean	SD	Min	Max
Implementing NAP	191	0.581	0.495	0	1
Antibiotic consumption	191	14.692	7.452	4.600	44.200
Bureaucratic capacity	170	0.397	1.229	-2.369	3.600
Regional peers	191	0.581	0.164	0.353	1
Population (log)	169	16.101	1.684	11.445	21.039
Government health expenditure (share of total)	191	0.515	0.217	0.047	0.941
GDPpc (log)	191	8.547	1.496	4.504	12.136
Economic growth	191	1.027	0.061	0.545	1.258
Democracy	170	0.540	0.252	0.022	0.932

2. Consumption models

Common trends are the primary identifying assumption of our Two-Way Fixed Effects (TWFE) regressions. The models assume that the treated observations would have been similar to the control observations without the treatment. Figure A1 presents a plausibility probe of this assumption. It shows no statistically significant differences between countries that implement a NAP and countries that do not in the three years preceding the first year of NAP implementation. Differences start to appear in the first year of implementation and are statistically significant in the two years following the start of NAP implementation. Hence, the results do not imply that the common trends assumption is violated.

Figure A1: Common trends assumption



The results presented so far strongly imply that the TWFE models are valid. Nevertheless, we conducted an additional test to understand whether our estimations could be affected by the negative weight problem. Recent advances in the literature on TWFE models have shown that estimations are sometimes based on problematic comparisons within groups over time. Since no observations are always treated, our TWFE models are a weighted average of two types of difference-in-difference estimators: (1) comparison of treated units over time and (2) comparisons of treated and control units (Goodman-Bacon, 2021). Results can be problematic when treated groups are compared to other treated groups, and it is essential to understand which groups drive our results (Goodman-Bacon, 2021; Baker et al., 2022). To this end, we present results from a Bacon decomposition in Table A5 for the fully specified Model 3 from the main body of the article. Overall, the TWFE estimators appear to make appropriate comparisons. The decomposition of Model 3 implies that the estimation is based on valid comparisons. The coefficients for the never treated vs. timing group (weighted at 83%) is similar to the timing group coefficient. However, the always treated vs. timing group comparison is positive and weighted very low. To ensure that the weighting of this group does not bias results, we present additional estimates that can overcome the negative weights problem below.

Table A3: Decomposed TWFE estimators

	Table 1, Model 1	
	Beta	Weight
Timing groups	-1.484	0.190
Always treated vs. timing group	4.111	0.014
Never treated vs. timing group	-1.128	0.795

The recent difference-in-difference literature has pointed to three problems when interpreting TWFE models as causal effects: negative weights, the assumption of no feedback, and constant treatment effects (Blackwell & Glynn, 2018; Goodman-Bacon, 2021; Imai & Kim, 2021). First, as discussed, the weights of the individual two-time-period difference-in-difference estimators are sometimes negative due to variation in the timing of treatments, leading to bias in the estimates (De Chaisemartin & d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Liu et al., 2022). Second, TWFE models must assume that past outcomes do not affect treatment assignment. Third, the effects of treatments are constant and do not carry over into the following year. There is reason to believe that these assumptions are violated in our case. Therefore, we test the robustness of a recently developed estimation procedure that relaxes these assumptions: the fixed effects counterfactual estimator (Liu et al., 2022). Specifically, the estimators take "observations under the treatment condition as missing, use data under the control condition to build models and impute counterfactuals of treated observations based on the estimated models" (Liu et al., 2022, p. 2). The approach to estimate counterfactuals is similar to the synthetic control method (Abadie et al., 2010) but can be applied to time-series-cross-sectional data (Liu et al., 2022).

Figure A2 displays the impact of NAP adoption—coded as one for each year that a NAP was in place—and Figure A3 shows the effect of NAP implementation—coded as one for each year that a NAP was being implemented. NAP adoption does not appear to impact antibiotic consumption. The coefficient fails to attain statistical significance ($p < 0.05$). However, the models clearly show that NAP implementation is associated with a statistically significant decrease in antibiotic consumption. The average treatment effect starts decreasing at the time of the treatment and becomes statistically significant compared to the counterfactual group in the third year after implementation started. In substantive terms, the models estimate an average yearly decrease in consumption of 0.811 due to the implementation of NAPs. The average consumption in the control group was approximately 16 DDD per 1000 per day. Hence, our models estimate that NAP implementation is associated with a decrease in consumption of approximately 5% every year.

Figure A2: Fixed effects counterfactual estimator: Adopting NAPs

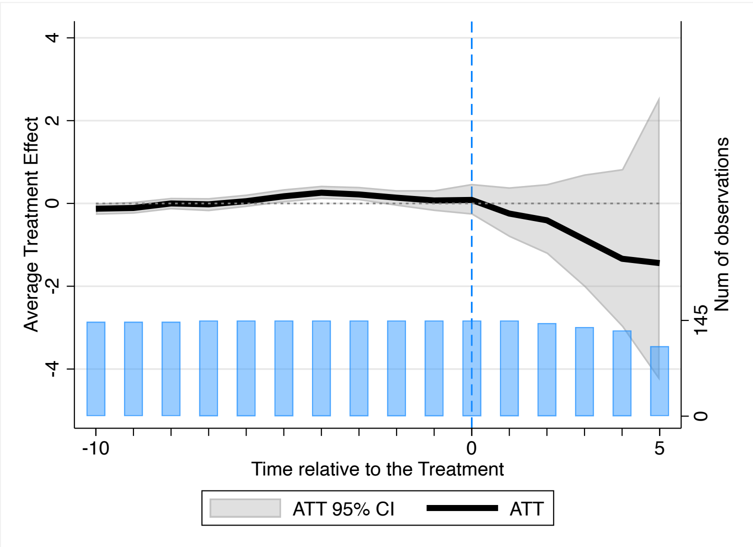


Figure A3: Fixed effects counterfactual estimator: Implementing NAPs

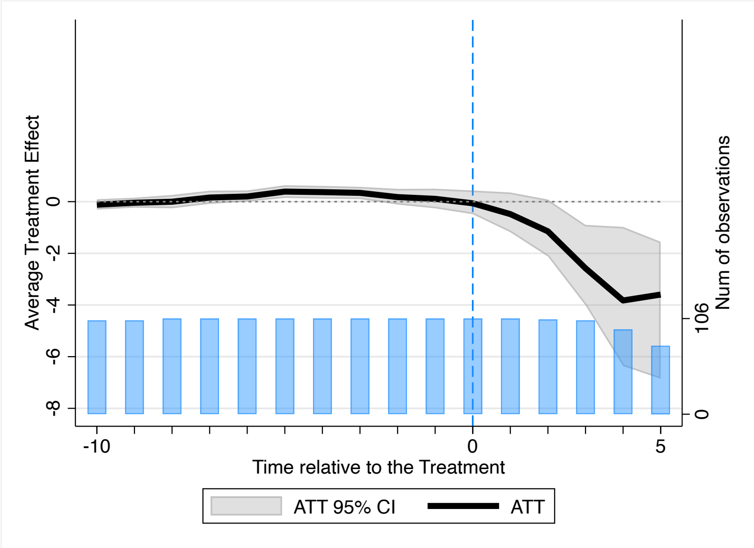


Figure A4 displays an equivalence test to verify the common trends assumption further, and Figure A5 is a placebo test. The equivalence test indicates that the common trend assumption is not violated. All ten pre-treatment years are within the equivalence bounds, and the estimated average treatment effect is nearly precisely zero in year 0. The placebo check presents further evidence of the validity of our estimations. We follow Liu et al. (2022) and define t-2 to 0 as our placebo periods. The placebo check assumes that the treatment starts three years earlier than it does and applies the same counterfactual estimator used to produce the average treatment effects for the main model. The p-value for the placebo check is far from any conventionally accepted level of statistical significance, and the plot shows that the placebo check estimates an average treatment effect close to zero. Therefore, the diagnostics confirm the validity of the model.

Figure A4: Equivalence test

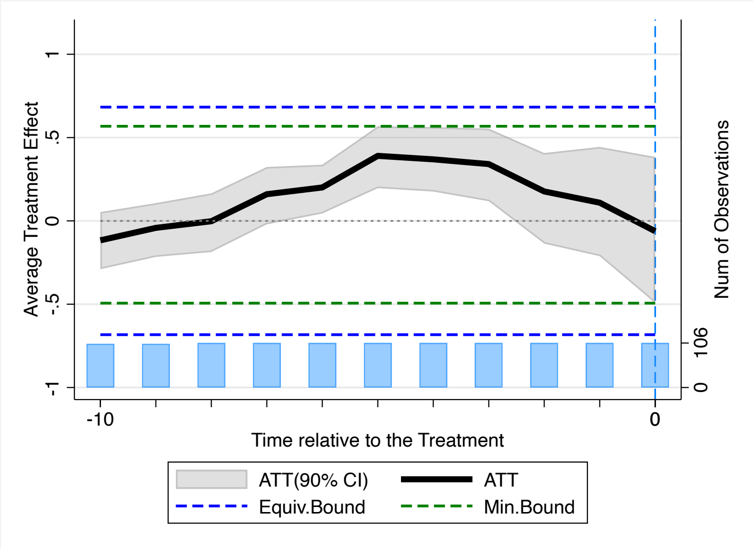
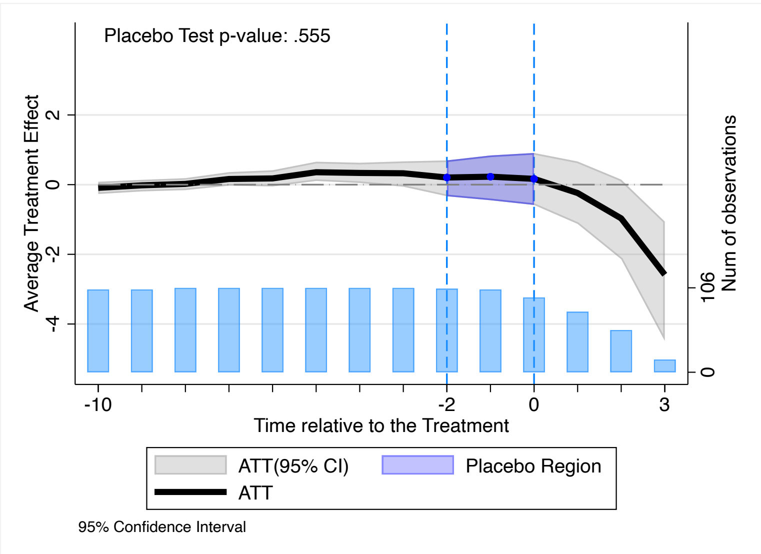


Figure A5: Placebo test



These robustness checks rule out the possibility that differential trends between countries with NAPs and without NAPs drive the results. However, it is still possible that a third variable affects treatment assignment and antibiotic consumption—which would create omitted variable bias. A potentially relevant confounder is political will. Countries may simply have pushed for the GAP and NAPs because they were already implementing domestic policies to address AMR and wanted to ensure other countries would do the same. Data on domestic legislation before 2016 is, unfortunately, not widely available. However, the 2016/2017 TRACSS survey allows us to probe an observable implication of this argument: that a substantial number of countries had already implemented national legislation to steward antibiotic consumption. Specifically, the survey asked whether countries implemented national stewardship legislation that they also enforced and monitored (Question 9.1). According to this survey, only 16 countries had such legislation in 2016/17 (11%)—of which eight countries did not have a NAP yet. We re-estimate the models excluding all 16 countries that could have potentially had robust national stewardship legislation before introducing a NAP in Table A6. The results are virtually similar when excluding these countries. The data from the survey also gives some descriptive evidence of the relevance of NAPs—Five years after the GAP, the number of countries that implemented comprehensive national stewardship legislation had almost tripled (to 53 countries).

Table A4: Excluding 16 countries that could have implemented comprehensive national stewardship legislation before implementing a NAP

	(14)	(15)
NAP implementation	-1.3180**	-1.3704***
	(0.5066)	(0.4549)
NAP adoption without implementation		-0.2033
		(0.3697)
Population (log)		2.9410
		(1.8725)
Government health expenditure (share of total)		1.4676
		(1.6760)
Bureaucratic capacity		-0.6836
		(0.6172)
GDPpc (log)		1.1631
		(0.7667)
Economic growth		-0.1998
		(0.6275)
Democracy		3.5136
		(2.7032)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	3325	2561
R^2	0.919	0.936

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Another way to probe political will is to estimate whether NAP implementation made a bigger difference in countries where AMR is already a bigger problem. The implication would be that countries with more AMR prevalence should, all else equal, be more willing to address the overconsumption of antibiotics. Unfortunately, data on AMR prevalence is very scarce. The only global estimate of the disability-adjusted life years (DALYs) lost to AMR is provided only for 2015 by Murray et al. (2022). We interact this variable with our NAP implementation variable to understand whether NAP implementation led to greater reductions of antibiotic consumption in regions where AMR already leads to more death and disability. The estimates are presented in Table A5.

Table A5: Interaction with the severity of AMR in 20 sub-regions

	(16)	(17)
NAP implementation	-2.1182***	-1.8326***
	(0.6777)	(0.5701)
NAP implementation * Regional AMR DALYs lost	0.0021**	0.0013**
	(0.0008)	(0.0006)
NAP adopted without implementation		-0.0500
		(0.3374)
Population (log)		2.8723
		(1.8217)
Government health expenditure (share of total)		1.1382
		(1.6557)
Bureaucratic capacity		-0.6981
		(0.5800)
GDPpc (log)		0.8544
		(0.7073)
Economic growth		-0.0409
		(0.6218)
Democracy		3.5106
		(2.6777)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	3629	2816
R^2	0.925	0.941

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

The main effect for the Regional AMR DALYs lost variable is absorbed by the country fixed effects. The main coefficient for NAP implementation is negative and statistically significant ($p < 0.001$). The interaction is positive, which implies that NAP implementation is less effective at reducing antibiotic consumption in regions with higher AMR prevalence. The coefficient is relatively small, and our models would estimate that NAP implementation still reduces antibiotic consumption in 17 of the 20 sub-regions included in the data. These findings imply that political will due to a greater AMR prevalence is unlikely to explain the reduction in antibiotics associated with NAP implementation.

Table A6: 2SLS models

	(18)	(19)
NAP implementation	-5.1513*	-3.0426^{+}
	(2.0502)	(1.8236)
Population (log)	1.9864	-0.1952
	(1.9440)	(2.1039)
Government health expenditure (share of total)	0.8551 (1.7280)	1.2489 (1.6859)
Bureaucratic capacity	-0.4862	-0.5903
	(0.6023)	(0.6247)
GDPpc (log)	1.0633	1.3872*
	(0.6946)	(0.6364)
Economic growth	-0.1678	0.0579
	(0.6528)	(0.6366)
Democracy	3.1222	2.6770
	(2.5255)	(2.3349)
Regional peers (consumption)		0.6755*** (0.1942)
Observations	2816	2816
First-stage F-statistic	22.7	22.6

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Given the presented results, we deem it highly unlikely that our results are driven by endogeneity. We present an instrumental variable analysis in Table A6 to minimize this possibility further. An instrumental variable must explain NAP implementation (relevance) but be unrelated to antibiotic consumption in a given country (exclusion). As an instrument, we utilize the share of regional peers implementing their NAPs in the previous year. We demonstrated the instrument's relevance in the main article. We also believe that the instrument is plausibly exogenous as NAP implementation in other countries should not be directly related to the political will to reduce antibiotic consumption in a given country of interest, except through the greater likelihood that a country starts implementing its NAP as well. Model 18 displays Two-Stage-Least-Squares regressions that use this instrument. Model 19 further controls for antibiotic consumption among regional peers to ensure that potential regional clusters of political will do not threaten the validity of the instrument. The 2SLS estimates are negative and statistically significant (Model 18) or marginally significant (Model 19).

We presented extensive tests to ensure that our results are not driven by endogeneity at the national level. In a final step, we now quantify the degree to which our main estimates would be robust to potential confounding variables based on the procedure developed by Oster (2019). We present estimates quantifying the potential confounders necessary to render our primary variable of interest insignificant in Table A7. The estimates imply that a potential confounder would have to explain 21% of the variation of the existing control variables in Model 1, 27% in Model 2, 151% in Model 3, and 182% in Model 4. Due to the high R-squared of Model 4, 0.99 (within R-squared 0.85), and given the results of the various tests presented in the main body of the article and the appendix, we believe such a confounder is very unlikely.

Table A7: Test for possible confounders (Table 2)

Models	**Variable**	**Bound estimate**
Table 3, Model 1	NAP implemented	-0.215
Table 3, Model 2	NAP implemented	-0.276
Table 3, Model 3	NAP implemented	-1.517
Table 3, Model 4	NAP implemented	-1.829

Nevertheless, we provide further sensitivity analyses below to ensure the robustness of our estimations to alternative specification choices. Table A8 includes three additional control variables: health IGO and INGO membership (Heinzel & Koenig-Archibugi, 2022), development assistance for AMR (Micah et al., 2023), as well as government ideology and policy (Lindberg et al., 2022). In Table A9, we control for additional country-specific trends. Specifically, we control for an interaction between the average global antibiotic consumption and country dummies in Model 23. In Model 24, we control for an interaction between the average antibiotic consumption in a country's WHO region and country dummies. Finally, Model 25 includes an interaction between the years since the GAP and country dummies. Table A10 uses an error correction model. Table A12 re-estimates models using an alternative dependent variable: overall antibiotic consumption (rather than per 100.000 people) and Poisson models to account for overdispersion in this variant of the measure. In Table A13, we restrict the sample to only those countries that ever adopted a NAP (Models 36 and 37) and exclude all countries that did not fill out all five rounds of the TRACSS surveys (Models 38 and 39). Our headline results are consistent throughout these alternative specification choices.

In Table A11, we interact our main variable of interest with the share of government and out-of-pocket health spending to understand how the political economy of healthcare provision relates to our main findings. The interaction is statistically significant and negative, showing that the effect of NAP implementation on antibiotic consumption is conditional on the share of health spending that is provided by governments. The results indicate that the effect reaches conventional levels of statistical significance in roughly half of the countries—those with larger than median values in government share of total health expenditure.

Table A8: Additional control variables

	(20)	(21)	(22)
NAP implementation	-1.1700**	-1.2068**	-1.1000*
	(0.4220)	(0.4166)	(0.4393)
NAP adoption without implementation	-0.1168	-0.0838	-0.0864
	(0.3472)	(0.3411)	(0.3483)
Population (log)	3.0113	3.0859+	4.3872*
	(1.9416)	(1.8313)	(2.0194)
Government health expenditure (share of total)	1.2318	1.1747	1.2245
	(1.7087)	(1.6608)	(1.8262)
Bureaucratic capacity	-0.7573	-0.7290	-0.9500
	(0.5795)	(0.5786)	(0.5859)
GDPpc (log)	1.1220	0.8961	1.6562+
	(0.7551)	(0.7222)	(0.8719)
Economic growth	-0.2199	-0.1098	-0.1838
	(0.7010)	(0.6188)	(0.7942)
Democracy	3.7686	3.6150	3.8320
	(2.6835)	(2.6586)	(3.3351)
Health IO membership	-0.9757		
	(1.2153)		
Health INGO membership	1.1963		
	(0.8016)		
Development Assistance for AMR (log)		0.1331	
		(0.0920)	
Government Ideology			0.0783
			(0.1479)
Government Anti-elitism			0.0510
			(0.1434)
Government Welfare			-0.0132
			(0.1649)
Country fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	2775	2816	2492
R^2	0.939	0.940	0.944

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A9: Alternative time trends

	(23)	(24)	(25)
NAP implementation	-0.6750**	-0.8985***	-0.9178+
	(0.2432)	(0.2560)	(0.5207)
NAP adoption without implementation	0.0462	-0.1560	-0.3676
	(0.2131)	(0.2519)	(0.3180)
Population (log)	-3.0637	-3.3902	2.3681
	(2.2007)	(3.9000)	(1.8619)
Government health expenditure (share of total)	0.1216	1.1156	1.0709
	(0.8863)	(1.0086)	(1.6669)
Bureaucratic capacity	-0.1917	-0.0625	-0.6950
	(0.2142)	(0.2203)	(0.6738)
GDPpc (log)	1.2745**	0.9772*	1.2502
	(0.4037)	(0.4511)	(0.7763)
Economic growth	0.3416	0.4499	-0.1478
	(0.2884)	(0.3551)	(0.5982)
Democracy	0.2718	0.5316	1.8399
	(0.5678)	(0.7901)	(2.2272)
Country fixed effects	Yes	Yes	Yes
Country-specific global consumption trends	Yes	No	No
Country-specific regional consumption trends	No	Yes	No
Country-specific GAP trend	No	No	Yes
Year fixed effects	Yes	Yes	Yes
Observations	2816	2816	2816
R^2	0.986	0.984	0.961

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A10: Error-correction model

	(26)	(27)	(28)
NAP implementation	-0.2610*	-0.3130**	-0.3558*
	(0.1107)	(0.1112)	(0.1638)
NAP adoption without implementation	-0.0470	-0.1763	-0.1501
	(0.1359)	(0.1245)	(0.1576)
Antibiotic consumption	-0.0947***	-0.1120***	-0.3396***
	(0.0204)	(0.0211)	(0.0461)
Population (log)		0.4802	-1.6474
		(0.3046)	(1.4221)
Government health expenditure (share of total)		0.1124	0.3634
		(0.2915)	(0.5113)
Bureaucratic capacity		0.1289	0.0540
		(0.1060)	(0.1197)
GDPpc (log)		0.0244	0.0552
		(0.1032)	(0.2091)
Economic growth		0.2838	0.4429+
		(0.1760)	(0.2296)
Democracy		0.2660	-0.2310
		(0.3351)	(0.4009)
Country fixed effects	Yes	Yes	Yes
Country-specific time trends	No	No	Yes
Year fixed effects	Yes	Yes	Yes
Observations	3438	2816	2816
R^2	0.185	0.217	0.352

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A11: Interaction with health spending

	(29)	(30)	(31)	(32)
NAP implementation	1.1556	0.1288	-2.3744***	-1.8354***
	(0.7996)	(0.6770)	(0.5880)	(0.5399)
Government health expenditure (share of total)	1.5685	1.2336		
	(1.2908)	(1.6571)		
Out-of-pocket health expenditure (share of total)			-0.4508	1.4944
			(1.8596)	(1.7752)
Interaction	-4.0833***	-2.4193*	3.9981**	2.0022
	(1.2131)	(1.0829)	(1.4522)	(1.2460)
Population (log)		2.9166		3.3288+
		(1.8185)		(1.8250)
Bureaucratic capacity		-0.7112		-0.7092
		(0.5757)		(0.5783)
GDPpc (log)		0.8428		0.9497
		(0.7155)		(0.7302)
Economic growth		-0.0224		-0.1037
		(0.6151)		(0.6071)
Democracy		3.5163		3.6745
		(2.6881)		(2.7203)
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	3428	2816	3428	2816
R^2	0.930	0.940	0.930	0.940

Country-clustered standard errors in parentheses; + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12: alternative dependent variable: total consumption

	(33)	(34)
NAP implementation	-0.1700***	-0.1147*
	(0.0384)	(0.0567)
NAP adoption without implementation		-0.0637
		(0.0432)
Government health expenditure (share of total)		-1.0215**
		(0.3867)
Bureaucratic capacity		-0.0844
		(0.0581)
GDPpc (log)		0.3370*
		(0.1609)
Economic growth		-0.0645
		(0.2166)
Democracy		0.1248
		(0.2491)
Country fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Observations	3428	2652
Pseudo R^2	0.990	0.992

Country-clustered standard errors in parentheses; + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13: Restricted samples

	(35)	(36)	(37)	(38)
NAP implementation	-1.5305**	-1.2635*	-1.5901*	-1.2937*
	(0.5713)	(0.5351)	(0.6105)	(0.5131)
NAP adoption without implementation		-0.1515		-0.0295
		(0.4098)		(0.4693)
Population (log)		3.3231		3.4683
		(2.3963)		(2.5193)
Government health expenditure (share of total)		1.7465		2.0629
		(2.5334)		(2.6899)
Bureaucratic capacity		-0.8103		-1.1142
		(0.7118)		(0.8275)
GDPpc (log)		0.9863		0.7067
		(0.8746)		(0.7876)
Economic growth		0.2239		1.0332
		(0.7015)		(0.9404)
Democracy		4.8235		6.3507+
		(3.5280)		(3.7914)
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	2109	1803	1938	1677
R^2	0.910	0.933	0.893	0.923

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

3. Participation models

In addition to the robustness checks presented above, we estimate several additional analyses for the regressions explaining participation in the global regime to combat AMR through adopting a NAP. First, we test for robustness to the same additional control variables introduced in the previous section for the analysis of consumption (Table A14). Second, we control for linear time trends (Table A15). Third, we estimate an interaction between the GAP and the NAP adoption of regional peers (Table A16). Fourth, we re-estimate models using logistic regressions since the dependent variable is binary. We do not include country fixed effects in these models to retain the entire sample.

Table A14: Additional control variables

	(39)	(40)	(41)
GAP	0.0777***	0.0816***	0.0769***
	(0.0178)	(0.0169)	(0.0190)
Adoption by regional peers	0.6674***	0.6570***	0.7166***
	(0.1277)	(0.1255)	(0.1296)
Population (log)	0.0818	0.0475	0.0730
	(0.0550)	(0.0486)	(0.0790)
Government health expenditure (share of total)	0.0519	0.0778	0.1026
	(0.0834)	(0.0802)	(0.0919)
Bureaucratic capacity	0.0292	0.0288	0.0306
	(0.0181)	(0.0186)	(0.0200)
GDPpc (log)	0.0767**	0.0406	0.0313
	(0.0272)	(0.0247)	(0.0369)
Economic growth	-0.0427	-0.0260	0.0041
	(0.0383)	(0.0337)	(0.0394)
Democracy	-0.0733	-0.0703	-0.0771
	(0.0696)	(0.0669)	(0.0761)
Antibiotic consumption	-0.0009	-0.0018	-0.0019
	(0.0024)	(0.0023)	(0.0026)
Health IO membership	-0.0375		
	(0.0432)		
Health INGO membership	-0.0239		
	(0.0255)		
Development Assistance for AMR (log)		0.0248**	
		(0.0074)	
Government Ideology			-0.0045
			(0.0079)
Government Anti-elitism			-0.0036
			(0.0080)
Government Welfare			-0.0171
			(0.0107)
Country fixed effects	Yes	Yes	Yes
Observations	2529	2567	2260
R^2	0.218	0.223	0.225

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A15: Controlling for time trends

	(42)	(43)	(44)	(45)
GAP	0.1843***	0.1061***	0.1706***	0.0747***
	(0.0131)	(0.0151)	(0.0159)	(0.0178)
Adoption by regional peers		0.5513***		1.1049***
		(0.0718)		(0.1558)
Population (log)			-0.0012	0.1645
			(0.2508)	(0.3028)
Government health expenditure (share of total)			0.1282	0.1258
			(0.1291)	(0.1746)
Bureaucratic capacity			0.0437	0.0397
			(0.0311)	(0.0302)
GDPpc (log)			-0.0643	0.0148
			(0.0595)	(0.0529)
Economic growth			0.0068	-0.0620
			(0.0330)	(0.0416)
Democracy			-0.0841	-0.0312
			(0.0859)	(0.0898)
Antibiotic consumption			-0.0118*	-0.0071
			(0.0047)	(0.0048)
Country fixed effects	Yes	Yes	Yes	Yes
Country-specific time trends	Yes	Yes	Yes	Yes
Observations	3593	3399	2567	2567
R^2	0.280	0.331	0.287	0.370

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A16: Interaction between GAP and regional peers

	(46)	(47)
GAP	0.1275***	0.0980***
	(0.0153)	(0.0183)
Adoption by regional peers	0.4746**	0.9790***
	(0.1816)	(0.2370)
Interaction	-0.2496	-0.3142
	(0.1749)	(0.1981)
Population (log)		0.0332
		(0.0459)
Government health expenditure (share of total)		0.0684
		(0.0820)
Bureaucratic capacity		0.0299
		(0.0186)
GDPpc (log)		0.0487+
		(0.0262)
Economic growth		-0.0192
		(0.0368)
Democracy		-0.0737
		(0.0681)
Antibiotic consumption		-0.0022
		(0.0024)
Constant	-0.0157*	-0.9181
	(0.0073)	(0.6738)
Country fixed effects	Yes	Yes
Observations	3399	2567
R^2	0.195	0.220

Country-clustered standard errors in parentheses; + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A17: Logistic regression

	(48)	(49)	(50)
GAP	3.6120***		3.1549***
	(0.2841)		(0.3137)
Adoption by regional peers		7.7705***	3.1687***
		(0.8272)	(0.7543)
Population (log)	0.2248***	0.2919***	0.2756***
	(0.0577)	(0.0679)	(0.0696)
Government health expenditure (share of total)	-1.0987	-2.2528**	-1.5646*
	(0.7017)	(0.8139)	(0.7759)
Bureaucratic capacity	0.1769	0.2160	0.2094
	(0.1107)	(0.1320)	(0.1290)
GDPpc (log)	0.2475*	0.0899	0.1944
	(0.1187)	(0.1283)	(0.1268)
Economic growth	-0.2856	-2.7610*	-0.8001
	(1.4187)	(1.2133)	(1.6048)
Democracy	0.6186	0.2788	0.5822
	(0.5693)	(0.6680)	(0.6518)
Antibiotic consumption	0.0308+	0.0393*	0.0236
	(0.0185)	(0.0173)	(0.0191)
Observations	2572	2572	2572
Pseudo R^2	0.316	0.202	0.339

Country-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

4. Implementation models

In a final step, we estimate several robustness checks for our implementation models. First, we use logistic models since the dependent variable is binary (Table A18). Second, we account for the additional control variables used in the previous sections, in models where our main variables are introduced one at a time (Table A19, A20, and A21). Third, we employ alternative measures for bureaucratic capacity: the “bureaucracy quality” from the International Country Risk Guide (PRS Group, 2021) and the “government effectiveness” variable from the Worldwide Governance Indicators (World Bank, 2022). Finally, participation and implementation are not independent decisions. Therefore, we re-estimate participation and implementation jointly using the conditional mixed process estimator—a variant of seemingly unrelated regression (Table A22).

The findings on the role of antibiotic consumption and regional peer influences are confirmed by these checks. By contrast, the link between bureaucratic capacity and NAP implementation appears to be less robust. The variable fails to attain statistical significance at conventional thresholds when we control for development aid for AMR and government ideology/policy and when we employ the ICRG measure of bureaucratic quality. Nevertheless, the government effectiveness variable attains statistical significance. As the ICRG measure is methodologically less transparent than the V-Dem measure we used to capture bureaucratic capacity in our main analyses, we have higher confidence in the results based on the latter, but we note the discrepancy.

Table A18: Logistic regression

	(51)	(52)	(53)	(54)
Bureaucratic capacity	0.4946*			0.3797**
	(0.2071)			(0.1211)
Antibiotic consumption		0.0648**		0.0370+
		(0.0198)		(0.0190)
Implementation by regional peers			3.3182**	2.3207+
			(1.0929)	(1.2055)
GDPpc (log)	0.4011	0.3494	0.4287+	0.2875
	(0.2966)	(0.2761)	(0.2589)	(0.2487)
Democracy	-0.2628	1.0555	1.6751	0.5784
	(1.2639)	(1.0769)	(1.0258)	(1.1548)
Economic growth	-4.7789*	-2.9909	-2.0995	-3.2657+
	(2.0714)	(2.3752)	(1.9355)	(1.8770)
Government health expenditure (share of total)	-2.0349	-1.9293	-2.2386	-2.5582+
	(1.4388)	(1.7206)	(1.3809)	(1.3703)
Population (log)	0.4611*	0.4675*	0.4873*	0.4503*
	(0.2247)	(0.2089)	(0.2334)	(0.2164)
Observations	169	169	169	169
Pseudo R^2	0.157	0.157	0.172	0.186

Region-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A19: additional control variables (antibiotic consumption and regional peers omitted)

	(55)	(56)	(57)
Bureaucratic capacity	0.0783^{+}	0.0836	0.0893
	(0.0357)	(0.0452)	(0.0530)
GDPpc (log)	0.0883	0.0712	0.0752
	(0.0617)	(0.0640)	(0.0701)
Democracy	-0.0027	-0.0468	-0.0350
	(0.1716)	(0.2410)	(0.2723)
Economic growth	-0.5160	-0.9412^{+}	-0.9032
	(0.4276)	(0.3921)	(1.0857)
Government health expenditure (share of total)	-0.3860	-0.3511	-0.4377
	(0.2812)	(0.2784)	(0.3127)
Population (log)	0.1018*	0.0812	0.0873^{+}
	(0.0288)	(0.0415)	(0.0370)
Health IO membership	0.1296		
	(0.0960)		
Health INGO membership	-0.0625		
	(0.0635)		
Development Assistance for AMR (log)		0.0049	
		(0.0360)	
Government Ideology			0.0050
			(0.0399)
Government Anti-elitism			0.0191
			(0.0211)
Government Welfare			0.0543
			(0.0384)
Constant	-1.1015^{+}	-0.1321	-0.2771
	(0.4894)	(0.7211)	(1.5052)
Observations	166	169	150
R^2	0.190	0.179	0.204

Region-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A20: Additional control variables (bureaucratic capacity and regional peers omitted)

	(58)	(59)	(60)
Antibiotic consumption	0.0101*	0.0110*	0.0097*
	(0.0031)	(0.0037)	(0.0034)
GDPpc (log)	0.0873	0.0692	0.0840
	(0.0578)	(0.0573)	(0.0637)
Democracy	0.1902	0.1676	0.1567
	(0.1588)	(0.2213)	(0.2851)
Economic growth	-0.1483	-0.6274	-0.5673
	(0.5378)	(0.4315)	(1.1437)
Government health expenditure (share of total)	-0.4074	-0.3641	-0.4745
	(0.3128)	(0.3216)	(0.3595)
Population (log)	0.1035*	0.0817	0.0922+
	(0.0319)	(0.0414)	(0.0361)
Health IO membership	0.1349		
	(0.0783)		
Health INGO membership	-0.0636		
	(0.0644)		
Development Assistance for AMR (log)		0.0178	
		(0.0342)	
Government Ideology			0.0087
			(0.0436)
Government Anti-elitism			0.0081
			(0.0207)
Government Welfare			0.0625
			(0.0364)
Constant	-1.7112+	-0.6823	-0.9692
	(0.6867)	(0.6242)	(1.4386)
Observations	166	169	150
R^2	0.191	0.182	0.201

Region-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A21: Additional control variables (antibiotic consumption and bureaucratic capacity omitted)

	(61)	(62)	(63)
Implementation by regional peers	0.5116*	0.5989*	0.5992*
	(0.1361)	(0.1718)	(0.1758)
GDPpc (log)	0.0946	0.0806	0.0961
	(0.0603)	(0.0539)	(0.0514)
Democracy	0.2668	0.2636	0.2616
	(0.1613)	(0.2070)	(0.2884)
Economic growth	-0.1593	-0.4774	-0.1551
	(0.5271)	(0.3961)	(1.1025)
Government health expenditure (share of total)	-0.4140	-0.4016	-0.5104
	(0.2780)	(0.2626)	(0.3095)
Population (log)	0.0997*	0.0862+	0.0903+
	(0.0248)	(0.0400)	(0.0364)
Health IO membership	0.0804		
	(0.0519)		
Health INGO membership	-0.0428		
	(0.0568)		
Development Assistance for AMR (log)		-0.0081	
		(0.0454)	
Government Ideology			-0.0036
			(0.0399)
Government Anti-elitism			0.0198
			(0.0150)
Government Welfare			0.0494
			(0.0392)
Constant	-1.8545*	-1.2279	-1.7123
	(0.6279)	(0.6310)	(1.4618)
Observations	166	169	150
R^2	0.202	0.198	0.224

Region-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A22: Additional control variables (fully specified models)

	(64)	(65)	(66)
Bureaucratic capacity	0.0628*	0.0656*	0.0644+
	(0.0209)	(0.0250)	(0.0263)
Antibiotic consumption	0.0073**	0.0074*	0.0054*
	(0.0016)	(0.0021)	(0.0019)
Implementation by regional peers	0.3298+	0.4201*	0.4399*
	(0.1449)	(0.1563)	(0.1492)
GDPpc (log)	0.0675	0.0499	0.0670
	(0.0598)	(0.0573)	(0.0571)
Democracy	0.0949	0.0777	0.0909
	(0.1825)	(0.2110)	(0.2879)
Economic growth	-0.3198	-0.6576	-0.4852
	(0.4262)	(0.3814)	(1.1813)
Government health expenditure (share of total)	-0.4533	-0.4498	-0.5370
	(0.2850)	(0.2657)	(0.3314)
Population (log)	0.0932*	0.0793+	0.0841+
	(0.0260)	(0.0380)	(0.0351)
Health IO membership	0.0631		
	(0.0467)		
Health INGO membership	-0.0434		
	(0.0523)		
Development Assistance for AMR (log)		-0.0155	
		(0.0465)	
Government Ideology			-0.0025
			(0.0407)
Government Anti-elitism			0.0185
			(0.0185)
Government Welfare			0.0479
			(0.0385)
Constant	-1.2294+	-0.5741	-0.9334
	(0.5454)	(0.6330)	(1.5174)
Observations	166	169	150
R^2	0.215	0.213	0.236

Region-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A23: Alternative measures of bureaucratic capacity

	(67)	(68)
ICRG Bureaucracy quality	0.0643	
	(0.0476)	
WGI Government effectiveness		0.2826***
		(0.0278)
GDPpc (log)	0.0576	-0.0373
	(0.0453)	(0.0354)
Democracy	0.1540	-0.1292
	(0.2647)	(0.1232)
Economic growth	-0.4723	-1.3559*
	(0.3937)	(0.3886)
Government health expenditure (share of total)	-0.3862	-0.4298
	(0.2375)	(0.2923)
Population (log)	0.0963*	0.0802+
	(0.0341)	(0.0367)
Constant	-0.9354	1.3766+
	(0.5945)	(0.6289)
Observations	135	169
R^2	0.162	0.232

Region-clustered standard errors in parentheses; + p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Table A24: Conditional Mixed Process models

	(69)
Implementation	
Bureaucratic capacity	0.0625**
	(0.0205)
Regional peers	0.4050**
	(0.1319)
Antibiotic consumption	0.0081***
	(0.0019)
GDPpc (log)	0.0460
	(0.0525)
Democracy	0.0801
	(0.2040)
Economic growth	-0.6444+
	(0.3815)
Government health expenditure (share of total)	-0.4129
	(0.2766)
Population (log)	0.0758*
	(0.0370)
Participation	
GAP	0.0840***
	(0.0174)
Regional peers	0.6528***
	(0.1243)
Population (log)	0.0488
	(0.0507)
Government health expenditure (share of total)	0.0712
	(0.0678)
Bureaucratic capacity	0.0305
	(0.0199)
GDPpc (log)	0.0517
	(0.0508)
Economic growth	-0.0304
	(0.0421)
Democracy	-0.0719
	(0.0653)
Antibiotic consumption	-0.0016
	(0.0026)
Country fixed effects	Yes
Observations	2590

Region-clustered standard errors in parentheses; + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5. Countries and NAP adoption years

Table A25: list of countries and their NAP adoption dates

Country	NAP	Country	NAP	Country	NAP	Country	NAP
Afghanistan	2017	El Salvador	2022	Maldives	2017	Slovakia	2019
Albania	2017	Equ. Guinea	N/A	Mali	2019	Slovenia	2019
Algeria	N/A	Eritrea	2021	Malta	2018	Solomon Islands	2021
Angola	N/A	Estonia	2021	Marshall Islands	2019	Somalia	2020
Antigua & Barbuda	N/A	Ethiopia	2015	Mauritania	N/A	South Africa	2018
Argentina	2015	Micronesia	2019	Mauritius	2017	South Korea	2016
Armenia	2015	Fiji	2015	Mexico	2016	South Sudan	N/A
Australia	2015	Finland	2017	Moldova	N/A	Spain	2014
Austria	2014	France	2001	Monaco	N/A	Sri Lanka	2017
Azerbaijan	N/A	Gabon	2019	Mongolia	2017	Sudan	2017
Bahrain	2019	Georgia	2017	Montenegro	2016	Suriname	2018
Bangladesh	2017	Germany	2008	Morocco	2018	Swaziland	2017
Barbados	2017	Ghana	2017	Mozambique	2019	Sweden	2000
Belarus	2016	Greece	2008	Myanmar	2017	Switzerland	2015
Belgium	2014	Grenada	N/A	Namibia	2017	Syria	2022
Belize	N/A	Guatemala	N/A	Nauru	2021	Tajikistan	2018
Benin	2019	Guinea	2020	Nepal	2016	Tanzania	2017
Bhutan	2017	Guinea-Bissau	N/A	Netherlands	2015	Thailand	2017
Bolivia	N/A	Guyana	N/A	New Zealand	2017	The Bahamas	N/A
Botswana	2017	Haiti	2017	Nicaragua	2014	The Gambia	N/A
Brazil	2018	Honduras	N/A	Niger	2019	Timor-Leste	2017
Brunei	2019	Hungary	N/A	Nigeria	2017	Togo	2019
Bulgaria	2019	Iceland	2017	North Korea	2018	Tonga	2017
Burkina Faso	2018	India	2017	Norway	2000	Trinidad & Tobago	2018
Burundi	2020	Indonesia	2017	Oman	2017	Tunisia	2019
Cambodia	2014	Iran	2016	Pakistan	2017	Turkey	2016
Cameroon	2018	Iraq	2018	Palau	N/A	Turkmenistan	2017
Canada	1997	Ireland	2017	Panama	N/A	Tuvalu	2021
Cape Verde	2018	Israel	N/A	Papua New Guinea	2019	Uganda	2018
CAR.	N/A	Italy	2017	Paraguay	2017	Ukraine	2019
Chad	2018	Jamaica	N/A	Peru	2016	UAE	2019
Chile	2017	Japan	2016	Philippines	2015	United Kingdom	2000
China	2014	Jordan	2018	Poland	2004	United States	2001
Colombia	2018	Kazakhstan	N/A	Portugal	2013	Uruguay	2018
Comoros	N/A	Kenya	2017	Qatar	2016	Uzbekistan	2018
Congo	N/A	Kiribati	N/A	Romania	N/A	Vanuatu	N/A
Cook Islands	2016	Kuwait	N/A	Russia	2017	Venezuela	N/A
Costa Rica	2017	Kyrgyzstan	2022	Rwanda	2020	Vietnam	2013
Cote d'Ivoire	2021	Laos	2019	St. Kitts & Nevis	2018	Yemen	2022
Croatia	2017	Latvia	2019	St. Lucia	N/A	Zambia	2017
Cuba	2019	Lebanon	2019	St. Vincent & Gren.	N/A	Zimbabwe	2017
Cyprus	2012	Lesotho	2020	Samoa	2016		
Czech Rep.	2011	Liberia	2018	San Marino	N/A		
DR Congo	2018	Libya	2019	Sao Tome & Principe	N/A		
Denmark	2017	Lithuania	2017	Saudi Arabia	2017		
Djibouti	N/A	Luxembourg	2018	Senegal	2017		
Dominica	N/A	Macedonia	2012	Serbia	2019		
Dominican Rep.	N/A	Madagascar	2018	Seychelles	2018		
Ecuador	2019	Malawi	2017	Sierra Leone	2018		
Egypt	2018	Malaysia	2017	Singapore	2017		

6. References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493-505.

Blackwell, M., & Glynn, A. N. (2018). How to make causal inferences with time-series cross-sectional data under selection on observables. *American Political Science Review*, *112*(4), 1067-1082.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964-2996.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254-277.

Heinzel, M., & Koenig-Archibugi, M. (2022). Harmful side effects: How government restrictions against transnational civil society affect global health. *British Journal of Political Science*, *Published online 10 November*.

Imai, K., & Kim, I. S. (2021). On the use of two-way fixed effects regression models for causal inference with panel data. *Political Analysis*, *29*(3), 405-415.

Lindberg, S. I., Düpont, N., Higashijima, M., Kavasoglu, Y. B., Marquardt, K. L., Bernhard, M., Döring, H., Hicken, A., Laebens, M., Medzihorsky, J., Neundorf, A., Reuter, O. J., Ruth–Lovell, S., Weghorst, K. R., Wiesehomeier, N., Wright, J., Alizada, N., Bederke, P., Gastaldi, L., Grahn, S., Hindle, G., Ilchenko, N., Römer, J. v., Wilson, S., Pemstein, D., & Seim, B. (2022). *Varieties of Party Identity and Organization (V–Party) Dataset V2*. Varieties of Democracy (V-Dem) Project. https://doi.org/10.23696/vpartydsv2.

Liu, L., Wang, Y., & Xu, Y. (2022). A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data. *American Journal of Political Science*, *Published online August 2*.

Micah, A. E., Bhangdia, K., Cogswell, I. E., Lasher, D., Lidral-Porter, B., Maddison, E. R., Nguyen, T. N. N., Patel, N., Pedroza, P., & Solorio, J. (2023). Global investments in pandemic preparedness and COVID-19: development assistance and domestic spending on health between 1990 and 2026. *The Lancet Global Health*, *11*(3), e385-e413.

Murray, C. J., Ikuta, K. S., Sharara, F., Swetschinski, L., Aguilar, G. R., Gray, A., Han, C., Bisignano, C., Rao, P., & Wool, E. (2022). Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, *399*(10325), 629-655.

PRS Group. (2021). *International Country Risk Guide*. PRS Group.

World Bank. (2022). *Worldwide Governance Indicators (WGI)* World Bank. https://info.worldbank.org/governance/wgi/.