Supplementary online appendix

Thomann, E., Ioannidis, G., Zgaga, T. and F. Schwarz. 2025. Explaining Public Sector Corruption: The Hexagon Model. *Governance*, DOI:10.1111/gove.70000.

Table A1. Variables, measurement, data sources and aggregation

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
Context: analytically relevant factors that provide important backgrounds for mechanisms of corruption, but do not directly produce it.	Gender equality in politics (fempoliticians)*	Number of seats occupied by women in the lower and in the upper house of the national parliament.	The mean value of the percentage share of female politicians in both chambers (upper and lower chamber) of the national parliament, January 2020-December 2020. No normalization necessary.	Reverse Coding: Percentages display the share of male politicians.	Inter-Parliamentary Union
	Stability of political system (stability)*	Likelihood of a disorderly transfer of government power, armed conflict, violent demonstrations, social unrest, international tensions, terrorism, as well as ethnic, religious or regional conflicts.	Political stability index, mean value of 2018-2020. Normalized using the value 1.13 (highest value Portugal with 1.13 in 2018). The Political stability index is a composite measure as it is based on several other indexes from multiple sources including the Economist Intelligence Unit, the World Economic Forum, and the Political Risk Services, among others.	Reverse Coding: Lower values display a higher political stability.	TheGlobalEconomy.com
	Economic wealth (gdp)*	The sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products.	Gross Domestic Product (GDP) per capita: gross domestic product divided by midyear population, in current US dollars, mean value of 2018-2020. Normalized using the value 85'420.191 (highest value Ireland with 85'420.191 In 2020).	Reverse Coding: Lower values display a higher GDP.	World Bank
Opportunity: environmental or situational	Strength of civil society (civilsociety)*	Perceived extent to which a country's citizens are able to	**3 dimensions regarding** Civil liberties and **11 variables regarding** Freedom of the net, **time 2018-2020**. Mean value of individual dimensions	Mean value of normalized values of all dimensions and	Freedom House V-Dem

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
weaknesses in the system that make corruption possible if the right person recognizes, exploits, and turns them into reality.		participate in selecting their government, freedom of expression, and freedom of association.	and variables are normalized according to maximum possible scores of individual dimensions and variables before a mean value for all dimensions and variables is calculated. **<u>Civil Liberties</u>:** For each variable: 0-4 points, where a score of 0 represents the smallest degree of freedom and 4 the greatest degree of freedom. - *Freedom of Expression and Belief*; 4 questions, maximum 16 points: **1.** Are there free and independent media? **2.** Are individuals free to practice and express their religious faith or non-belief in public and private? **3.** Is there academic freedom, and is the educational system free from extensive political indoctrination? **4.** Are individuals free to express their personal views on political or other sensitive topics without fear of surveillance or retribution? - *Associational and Organizational Rights*; 3 questions, maximum 12 points: **1.** Is there freedom of assembly? **2.** Is there freedom for nongovernmental organizations, particularly those that are engaged in human rights– and governance-related work? **3.** Is there freedom for trade unions and similar professional or labor organizations? - *Personal Autonomy and Individual Rights*; 4 questions, maximum 16 points: **1.** Do individuals enjoy freedom of movement, including the ability to change their place of residence, employment, or education? **2.** Are individuals able to exercise the right to own property and establish private businesses	variables are reverse coded, so that lower value displays a higher strength of civil society.	

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			without undue interference from state or nonstate actors? **3.** Are individuals able to exercise the right to own property and establish private businesses without undue interference from state or nonstate actors? **4.** Do individuals enjoy equality of opportunity and freedom from economic exploitation? Freedom of the net: the degree to which governments and non-state actors around the world restrict individuals' rights online. Conducted through 11 questions. For each question 0-4 points, whereas a score of 0 represents the smallest degree of freedom and 4 the greatest degree of freedom (Exception: Question 8 → 0-1 points). Ordinal, converted to interval by measurement model: 1. Government Internet filtering in practice (v2smgovfilprc_osp): How frequently does the government censor political information (text, audio, images, or video) on the Internet by filtering (blocking access to certain websites)? 2. Government Internet shut down in practice (v2smgovshut_osp): How often does the government shut down domestic access to the Internet? 3. Government social media shut down in practice (v2smgovsm_osp): How often does the government shut down access to social media platforms? 4. Government social media alternatives (v2smgovsmalt_osp): How prevalent is the usage of social media platforms that are wholly		

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			controlled by either the government or its agents in this country? 5. Government social media monitoring (v2smgovsmmon_osp): How comprehensive is the surveillance of political content in social media by the government or its agents? 6. Government social media censorship in practice (v2smgovsmcenprc_osp): To what degree does the government censor political content (i.e., deleting or filtering specific posts for political reasons) on social media in practice? 7. Internet legal regulation content (v2smregcon_osp): What type of content is covered in the legal framework to regulate Internet? 8. Privacy protection by law exists (v2smprivex_osp): Does a legal framework to protect Internet users' privacy and their data exist? 9. Privacy protection by law content (v2smprivcon_osp): What does the legal framework to protect Internet users' privacy and their data stipulate? 10 Government online content regulation approach (v2smregapp_osp): Does the government use its own resources and institutions to monitor and regulate online content or does it distribute this regulatory burden to private actors such as Internet service providers? 11. Abuse of defamation and copyright law by elites (v2smdefabu_osp): To what extent do elites		

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			abuse the legal system (e.g., defamation and copyright law) to censor political speech online?		
	Independent press (press)*	Degree of freedom that journalists, news organizations, and citizens have in each country, and the efforts made by authorities to respect this freedom.	Yearly Press Freedom Index, created from two sources: 1) a questionnaire completed by media professionals, lawyers and sociologist; and 2) data on abuses and violence against journalists during the period covered, evaluated by a team of team of specialists; score 0-100, mean of 2018- 2020. No normalization necessary. Criteria evaluated in the questionnaire are pluralism, media independence, media environment and self-censorship, legislative framework, transparency, and the quality of the infrastructure that supports the production of news and information.	Reverse Coding: Lower values display a more independent press.	Reporters Without Borders
	Transparency (transparency)*	The available and accessible (cost free) minimal public information, made available by governments to their citizens, required to deter corruption and enable public accountability in a society.	Transparency Index (T-Index): sum of De Jure (what information governments have legally committed to share with citizens, e.g.) and De Facto Transparency (what information governments have actually shared) in 2021, score from 0 to 19. Normalized according to the maximum score of 19. The index is aggregated through a qualitative method similar to building an index of qualitative variation (IQV). The availability of the 14 resources in full is considered the 'anti-corruption transparency' target, and each component adds up equally to fulfill it to 100%, which is the equivalent of the maximum score of 14 points. A country's score represents the percentage to which the target is fulfilled. The same logic is applied to the de jure and to the total T-Index score, bringing the total 100% up to 19 points (value for normalization).	Reverse Coding: Lower values display a higher level of transparency.	CorruptionRisk.org

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			***De Jure* Transparency** Index has five components: 1. UNCAC ratification 2. Membership to Open Government Partnership (OGP) 3. FOI act present in national legislation 4. Ratification of either the OECD Convention against Bribery of International Officials, Extractive Industry Transparency Initiative, World Trade Organization Agreement on Global Procurement (GPA), or the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) 5. Financial Action Task Force against Money Laundering or equivalent ***De Facto* Transparency** Index has 14 components made of questions asked to country experts: **Q1:** Are past public expenditures published online? (1 point) A) Last fiscal year expenditure report is accessible online in its detailed form = 1 B) Last fiscal year expenditure report is accessible online with limited detail = 0.5 C) Not available online or too generic (only aggregated data) = 0 **Q2:** Are current public expenditures published online? (1 point) A) Data is available through an online tracking system with itemized expenditures (e.g. copy machine) = 1 B) Data is available through an online tracking system that is not itemized OR through fairly detailed budget execution reports = 0.5 C) Not available online or too generic (only		

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			aggregated data) = 0 **Q3:** Is there a centralized public procurement portal where both tenders and contract awards are posted? (1 point) A) Calls for bids AND award notices are published = 1 B) Only call for bids OR award notices are published = 0.5 C) No procurement portal exists OR information published is minimal (selected procedures only) = 0 **Q4:** Is there an online land cadaster where property ownership is disclosed? (1 point) A) Cadaster data is fully accessible online = 1 B) Cadaster data is partial OR limited in geographic coverage OR access requires payment = 0.5 C) Not available online = 0 **Q5:** Is there a register of commerce where shareholders and main data of companies is published? (1 point) A) Business registry is fully and freely available online = 1 1 point also when the register is run by a private company B) Information is partial or access to relevant information is paid = 0.5 C) Not available online = 0 **Q6:** Is the annual report of the Supreme Audit Institution publicly posted? (1 point) A) Annual report is available online with detailed information on individual audit results = 1 Cases where the report is not comprehensive, but		

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			all individual reports are easily accessible are granted a full point as well B) Annual report has information on selected audits (and audit results are not available elsewhere) = 0.5 C) No (current) report is available online = 0 **Q7:** Are supreme court hearing schedules public and accessible online? (1 point) A) All court information available online = 1 B) Not all information public, politically sensitive cases not available = 0.5 C) Not available online = 0 **Q8:** Are supreme court sentences published online? (1 point) A) All court sentences available online = 1 B) Not all information public, politically sensitive cases not available = 0.5 C) Not available online = 0 **Q9:** Are financial disclosures of officials publicly available online? (1 point) A) Available for all officials required to declare = 1 B) Available only for part of the officials required to declare (e.g. top officials) = 0.5 C) Not available online (or only upon request) = 0 **Q10:** Are conflict of interest disclosures of officials publicly available online? (1 point) A) Available for all officials required to declare = 1 B) Available only for part of the officials required to declare (e.g. top officials) = 0.5 C) Not available online (or only upon request) = 0		

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			Q11: Are incoming and outcoming donor funds' allocations published? (1 point) A) Incoming/outcoming donor funds (or both whenever applicable) are available B) Only incoming OR outcoming donor funds are available in a situation when there should be both = 0.5 A half point is also given if information is partial. C) Not available online = 0 **Q12:** Are mining concessions publicly posted? (1 point) A) Information on mining concessions/licenses/titles is fully available = 1 B) Information on mining concessions/licenses/titles is partially available or access is paid = 0.5 C) Not available online = 0 **Q13:** Are construction permits in the country's capital city publicly posted? (1 point) A) Information on issued construction permits is fully available = 1 B) Information on issued construction permits is partially available OR access is paid = 0.5 C) Not available online = 0 **Q14:** Is there an online gazette or a government portal which publishes all official legislation for everybody to access? (1 point) A) Yes = 1 B) Yes, but the access is paid OR the information is only partially provided = 0.5 C) No = 0		
	Rule of law	Extent to which laws	7 variables with responses on a 0-4 scale	Mean value of	V-Dem

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
	(rol)*	are transparently, independently, predictably, impartially, and equally enforced, and to which the actions of government officials comply with the law.	whereas a score of 0 represents the smallest degree of rule of law and 4 the highest degree of rule of law. Ordinal, converted to interval by measurement model Mean value of 2018-2020. Mean value of individual variables is normalized according to the maximum possible score (=4) before a mean value of all 7 variables is calculated **Compliance with high court** (v2juhccomp_osp): How often would you say the government complies with important decisions of the high court with which it disagrees? **Compliance with judiciary** (v2jucomp_osp): How often would you say the government complies with important decisions by other courts with which it disagrees? **High court independence** (v2juhcind_osp): When the high court in the judicial system is ruling in cases that are salient to the government, how often would you say that it makes decisions that merely reflect government wishes regardless of its sincere view of the legal record? **Lower court independence** (v2juncind_osp): When judges not on the high court are ruling in cases that are salient to the government, how often would you say that their decisions merely reflect government wishes regardless of their sincere view of the legal record? **Executive respects constitution** (v2exrescon_osp): Do members of the executive (the head of state, the head of government, and cabinet ministers) respect the constitution? **Access to justice for men** (v2clacjstm_osp): Do men enjoy secure and effective access to justice? **Access to justice for women** (v2clacjstw_osp): Do women enjoy equal, secure, and effective	normalized values of all indicators are reverse coded, so that a lower value displays a higher level of rule of lawThe mean value of normalized values of all indicators is reverse coded so that a lower value displays a higher level of the rule of law.	

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			access to justice?		
	Public sector accountability (accountability)*	Assessments of in-law and in-practice efforts to enhance the transparency of public administration and the accountability of public officials.	Mean score (0-100) of 5 indicators for 2020: political financing, financial disclosure, conflict of interests, freedom of information, public procurement. No normalization necessary.	Reverse Coding: Lower values display a higher public sector accountability.	EuroPAM
Supply: actions or actor constellations that actively encourage or pursue corrupt exchanges.	Organized crime (crime)	A continuing criminal enterprise that rationally works to profit from illicit activities that are often in great public demand.	Global Organized Crime Index that assesses the level of criminality and resilience to organized crime in a country assessed by thematic geographic and technical experts. Scale 1-10, year 2020. Normalized according to highest possible score of 10. Based on three key pillars with indicators: 1. **Criminal markets** Human trafficking Human smuggling Extortion and protection racketeering Arms trafficking Trade in counterfeit goods Illicit trade in excisable goods Flora crimes Fauna crimes Non-renewable resource crimes Heroin trade Cocaine trade Cannabis trade Synthetic drug trade Cyber-dependent crimes Financial crimes	Reverse Coding not necessary: Lower values display a lower level of organized crime.	Global Initiative Against Transnational Organized Crime

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
			2. **Criminal actors** Mafia-style groups Criminal networks State-embedded actors Foreign actors Private sector actors 3. **Resilience** Political leadership and governance Government transparency and accountability International cooperation National policies and laws Judicial system and detention Law enforcement Territorial integrity Anti-money laundering Economic regulatory capacity Victim and witness support Prevention Non-state actors		
Capability: public official's personal traits and abilities that are necessary to recognize a particular corruption opportunity and turn it into reality.	Public integrity (publicintegrity)*	The capacity of the public sector to act in an impartial way.	1 variable with responses on a 0-4 scale whereas a score of 0 represents the smallest degree of public integrity and 4 the greatest degree of public integrity. Ordinal, converted to interval by measurement model. Mean value of 2018-2020 Mean value of individual variables is normalized according to the maximum possible score (=4): **Rigorous and impartial public administration** (v2clrspct_osp): Are public officials rigorous and impartial in the performance of their duties?	Normalized mean value is reverse coded so that a low value displays a high level of public integrity.	V-Dem
Pressure/motivation: the public official's individual, subjective	Salary (salary)*	Net remuneration of central bureaucrats in real terms.	Remuneration index of national civil servants in central public administration, mean of 2018-2020. Normalized according to the highest value	Reverse Coding: Lower values display a higher salary-level.	Eurostat

Hexagon corner	Factor	Definition	Measurement	Coding	Data source
or objective need or desire to commit corruption-- occupational, economic, or social.			(Germany with 4282.10 in 2020).		
Rationalization: factors facilitating the public official's efforts to reduce dissonance and normalize the corrupt behavior, by creating a certain mindset and a set of reasons to justify their actions or disengage from guilt.	Interpersonal trust (trust)	The level of trust put in other persons.	Percentage of respondents who answered "totally trust" or "tend to trust" to the question: "Please tell me to what extent do you trust or not the people in (OUR COUNTRY)?". 2020. No normalization necessary	Reverse Coding not necessary: Lower values display a lower level of interpersonal trust.	Eurobarometer 93.1
	Trust in government (govtrust)*	The level of trust put in the national government.	Share of respondents who answered "tend to trust" to the question: "How much trust do you have in certain institutions? For each of the following institutions, do you tend to trust it or tend not to trust it? – The (NATIONALITY) Government" mean of 7 waves from 2018-2020. No normalization necessary	Reverse Coding not necessary: Lower values display a lower level of trust in government. Reverse Coding: low values display a high level of trust in the government.	Eurobarometer 89 - 97
Outcome					
High corruption risks	Corruption risks in public procurement in the healthcare market (COR)	Granted leeway and ambiguity in legal rules, procedures, institutional arrangements, and contracts that are commonly associated with corruption, measured as average of the Corruption Risk Index (CRI) for the healthcare market in 2020.	Average CRI score for healthcare markets between the first quarter of 2020based on the observation of seven red flags over all available procurement contracts in 2020: and the fourth quarter of 2020. 1. Single bidder contracts 2. Non-open procedures 3. Lack of public call for tenders 4. Period for submitting bids 5. Period for selecting the winning bid 6. Spending concentration 7. Share of suppliers registered in jurisdictions offering limited company and banking transparency	Higher normalized mean values display a higher risk of corruption in the healthcare market.	Government Transparency Institute

*Variable was reversecoded for radar charts

Table A2: Descriptive statistics of normalized reversecoded raw data—used for calibration

Statistic	N	Mean	St. Dev.	Min	Max
con	23	57.666	26.552	0.000	100.000
opp	23	45.557	29.063	0.000	100.000
sup	23	59.810	24.560	0.000	100.000
cap	23	41.987	29.472	0.000	100.000
mot	23	56.057	27.050	0.000	100.000
gov	23	47.030	31.273	0.000	100.000
int	23	43.243	27.064	0.000	100.000
cor	23	50.004	28.002	0.000	100.000

Table A3: Calibrated fuzzy dataset

Case	*COR*	*CON*	*SEC*	*OPP*	*SUP*	*CAP*	*MOT*	*INT*	*GOV*
AUT	0,84	0,15	0,22	0,70	0,30	0,22	0,25	0,41	0,89
BEL	0,10	0,37	0,67	0,70	0,41	0,13	0,24	0,64	0,52
BGR	0,71	0,87	0,05	0,57	0,88	0,95	0,90	0,12	0,13
HRV	0,95	0,70	0,66	0,31	0,76	0,90	0,86	0,25	0,05
CYP	0,69	0,79	0,21	0,93	0,35	0,65	0,40	0,32	0,51
CZE	0,18	0,43	0,39	0,65	0,53	0,64	0,62	0,41	0,40
DNK	0,34	0,08	0,95	0,53	0,25	0,09	0,09	0,95	0,95
EST	0,82	0,61	0,54	0,05	0,18	0,08	0,41	0,73	0,79
FIN	0,10	0,11	0,83	0,19	0,05	0,17	0,13	0,89	0,91
FRA	0,51	0,71	0,55	0,25	0,93	0,16	0,38	0,08	0,18
DEU	0,60	0,34	0,11	0,59	0,68	0,05	0,05	0,64	0,88
GRC	0,46	0,95	0,60	0,87	0,70	0,71	0,86	0,19	0,09
HUN	0,93	0,73	0,06	0,92	0,47	0,87	0,88	0,41	0,79
IRL	0,60	0,05	0,31	0,23	0,68	0,18	0,23	0,89	0,72
ITA	0,57	0,66	0,93	0,57	0,95	0,71	0,63	0,05	0,15
LVA	0,05	0,82	0,21	0,05	0,16	0,33	0,82	0,50	0,26
POL	0,14	0,81	0,05	0,95	0,29	0,84	0,87	0,41	0,36
PRT	0,49	0,22	0,82	0,26	0,49	0,67	0,69	0,73	0,79
ROU	0,63	0,93	0,05	0,37	0,51	0,88	0,95	0,14	0,15
SVK	0,14	0,70	0,10	0,06	0,57	0,62	0,73	0,12	0,19
SVN	0,41	0,53	0,62	0,17	0,39	0,42	0,67	0,25	0,17
ESP	0,95	0,68	0,88	0,21	0,95	0,21	0,26	0,67	0,09
SWE	0,24	0,31	0,94	0,82	0,80	0,07	0,26	0,89	0,95

Table A4: Individual analyses of necessity

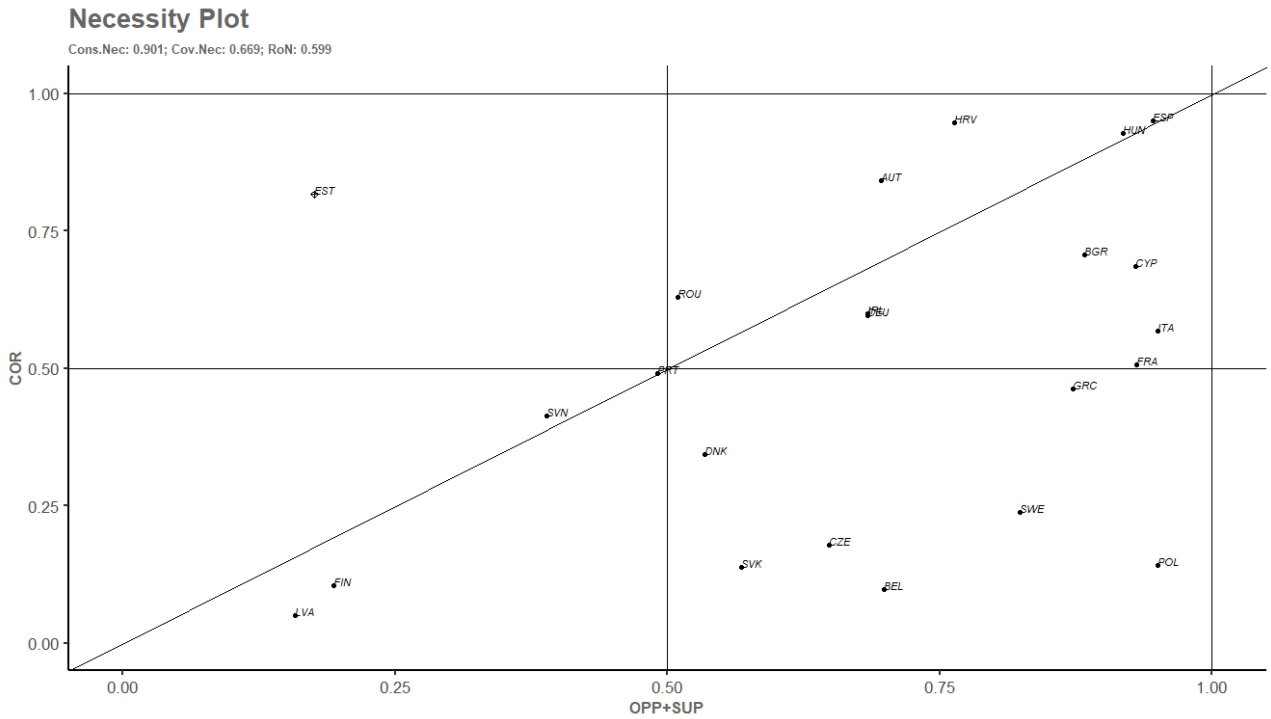
	Outcome:	**COR**		**Outcome:**	**~COR**	
Condition	***Cons.Nec***	*Cov.Nec*	*RoN*	*Cons.Nec*	*Cov.Nec*	*RoN*
CON	0,742	0,675	0,719	0,668	0,615	0,684
OPP	0,66	0,689	0,78	0,644	0,68	0,775
SUP	0,785	0,731	0,764	0,661	0,623	0,699
CAP	0,653	0,708	0,801	0,566	0,621	0,757
MOT	0,703	0,658	0,722	0,689	0,654	0,719
RAT	0,888	0,582	0,433	0,916	0,608	0,449
INT	0,574	0,615	0,75	0,635	0,689	0,788
GOV	0,606	0,634	0,751	0,61	0,646	0,757
~CON	0,577	0,632	0,766	0,647	0,717	0,81
~OPP	0,693	0,658	0,727	0,706	0,678	0,738
~SUP	0,595	0,635	0,758	0,714	0,771	0,833
~CAP	0,65	0,597	0,678	0,733	0,681	0,727
~MOT	0,63	0,667	0,772	0,64	0,685	0,782
~RAT	0,402	0,825	0,947	0,371	0,77	0,932
~INT	0,709	0,658	0,717	0,645	0,605	0,687
~GOV	0,662	0,626	0,708	0,654	0,627	0,708

Table A5: Complex necessary conditions for COR

Condition	*inclN*	*RoN*	*covN*
OPP + SUP	0,9011384	0,59899691	0,66920383
SUP + CAP + ~INT[1]	0,91184108	0,55230414	0,64929228

Thresholds: Consistency 0.9, Coverage 0.55, Relevance of Necessity 0.5.
Read: + = "OR", ~="NOT"
[1]Condition is not conceptually meaningful.

Figure A1: XY-plot, OPP + SUP ← COR



Read: + = "OR", ~ ="NOT" ,← = "is necessary for"

Table A6: Truth table for COR

Row Nr.	*CON*	*OPP*	*SUP*	*CAP*	*MOT*	*RAT*	*OUT*	*n*	*incl*	*PRI*	*cases*
53	1	1	0	1	0	0	1	1	0.913	0.518	CYP
32	0	1	1	1	1	1	0	1	0.886	0.605	**CZE**
64	1	1	1	1	1	1	1	3	0.876	0.639	BGR,**GRC,**ITA
17	0	1	0	0	0	0	1	1	0.868	0.520	AUT
10	0	0	1	0	0	1	1	1	0.862	0.489	IRL
55	1	1	0	1	1	0	1	1	0.859	0.533	HUN
48	1	0	1	1	1	1	1	3	0.854	0.593	HRV,ROU,**SVK**
8	0	0	0	1	1	1	0	1	0.848	0.327	**PRT**
42	1	0	1	0	0	1	1	2	0.846	0.490	FRA,ESP
34	1	0	0	0	0	1	0	1	0.830	0.420	EST
56	1	1	0	1	1	1	0	1	0.813	0.436	**POL**
26	0	1	1	0	0	1	0	2	0.808	0.384	DEU,**SWE**
18	0	1	0	0	0	1	0	2	0.778	0.271	**BEL,DNK**
2	0	0	0	0	0	1	0	1	0.715	0.220	**FIN**
36	1	0	0	0	1	1	0	2	0.687	0.178	**LVA,SVN**

Bold: Case has membership in COR < 0.5.

Rules for setting the raw consistency threshold and including or excluding truth table rows: Raw consistency threshold was set at 0.815, below which no more sufficiency claims were supported by individual cases. Second, empirical rows with deviant cases consistency in kind (DCCK) were included into logical minimization only if only 1 DCCK exists and it either has a value on COR very close to the crossover point (ambiguous membership in outcome set), or the DCCK constitutes 1/3 or less of the cases in that row. Third, empirical or remainder rows contradicting the necessity claim (displaying ~OPP*~SUP) were excluded. Fourth, only logical remainder rows that represent easy simplifying counterfactual assumptions (see Table 2) were included. This led to the exclusion of empirical rows 32 (Czech republic as clear DCCK), 8 (Portugal, DCCK and contradicting necessity claim) and 34 (Estonia, contradicting necessity claim), and entailed the counterfactual assumptions listed in Table A8.

Conservative solution (Consistency 0.819, PRI 0.631, coverage 0.773): CON*OPP*~SUP*CAP*~RAT + CON*SUP*CAP*MOT*RAT + ~OPP*SUP*~CAP*~MOT*RAT + ~CON*OPP*~SUP*~CAP*~MOT*~RAT → COR

Parsimonious solution (Consistency 0.808, PRI 0.611, Coverage 0.804): CON*SUP + ~OPP*SUP + OPP*~RAT → COR

2 intermediate solution models, differing on 1 condition in 1 path. M1 (reported in paper) is chosen for interpretation since it displays the highest consistency and fewer ambiguous (multiple covered) cases than M1.

Table A7: Intermediate solution model 2, above-average corruption risk (COR)

	Consistency	*PRI*	*Coverage*	*Unique coverage*	*cases*
~ CON* OPP* ~ RAT	0.866	0.551	0.309	0.025	AUT
~ OPP* SUP* RAT	0.849	0.619	0.611	0.175	IRL; FRA,ESP; HRV,ROU,SVK
OPP* CAP* ~ RAT	0.880	0.629	0.309	0.025	CYP; HUN
CON* SUP* CAP* MOT* RAT	0.801	0.565	0.509	0.012	HRV,ROU,SVK; BGR,GRC,ITA
Solution	0.815	0.623	0.799		

Figure A2: XY plot, intermediate solution for COR

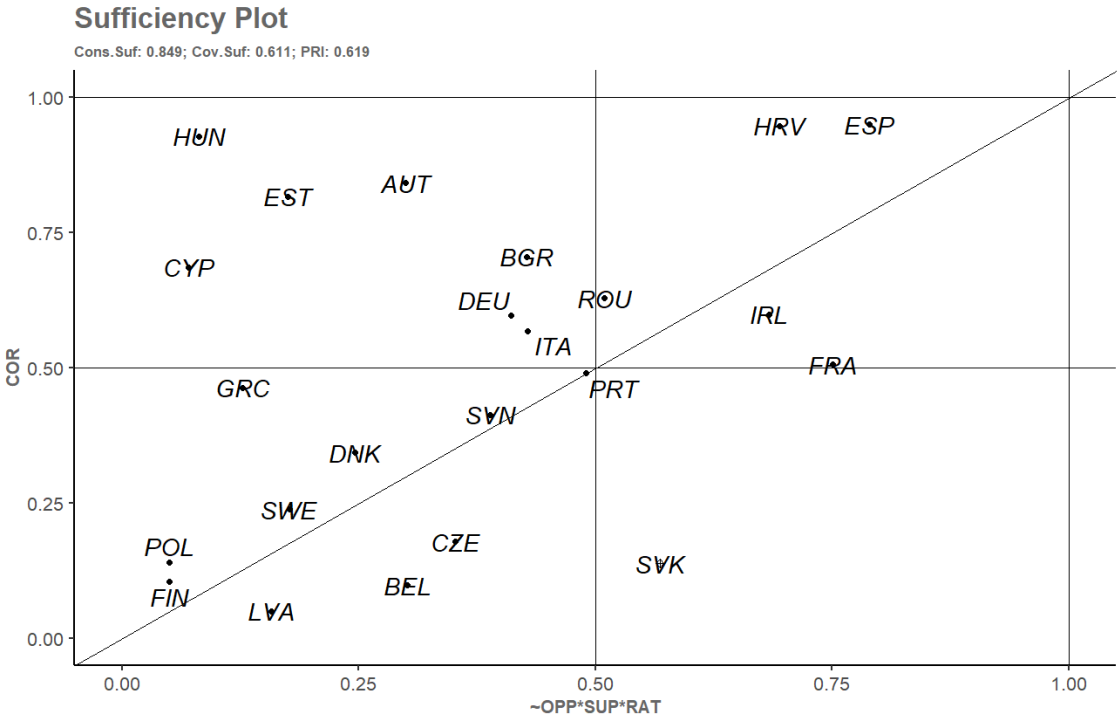


Table A8: Easy simplifying counterfactual assumptions for intermediate solution M1 (COR)

Row Nr.	*CON*	*OPP*	*SUP*	*CAP*	*MOT*	*RAT*
12	0	0	1	0	1	1
14	0	0	1	1	0	1
16	0	0	1	1	1	1
19	0	1	0	0	1	0
21	0	1	0	1	0	0
23	0	1	0	1	1	0
25	0	1	1	0	0	0
27	0	1	1	0	1	0
29	0	1	1	1	0	0
31	0	1	1	1	1	0
44	1	0	1	0	1	1
46	1	0	1	1	0	1
61	1	1	1	1	0	0
63	1	1	1	1	1	0

Robustness test

Given the sample-driven calibration procedure, we assessed sensitivity ranges for the calibration of five conditions (except RAT, which is a composed set) and the outcome set, using the protocol by Oana et al. (2021, 152ff), by increasing/decreasing each threshold by steps of 5, up to 40 times in total. The original calibration thresholds were set by using the sample minimum (0), sample mean (Table A2), and sample maximum (100). The results reported here general suggest that for most conditions, the crossover point (column C)—which determines differences in kind—would have to be altered dramatically in order to obtain a different parsimonious solution. For example, for CON, the parsimonious solution does not substantively change within a range of crossover point from 12,7 to 97,7. These tests support a high robustness of our results, despite the sample-driven calibration procedure. E = exclusion (full set non-membership), C = crossover point, I = inclusion (full set membership). NA means that using 40 iterations, no upper/lower bound could be identified.

Context CON

Original crossover point: 57.666

	E	C	I
Lower bound	-160	12,666	60
Upper bound	55	97,666	NA

Opportunity OPP

Original crossover point: 45.557

	E	C	I
Lower bound	-175	5,557	50
Upper bound	45	95,557	NA

Supply SUP

Original crossover point: 59.810

	E	C	I
Lower bound	-115	29,81	60
Upper bound	55	99,81	NA

Capability CAP

Original crossover point: 41.987

	E	C	I
Lower bound	NA	6,987	45
Upper bound	40	96,987	NA

Motivation MOT

Original crossover point: 56.057

	E	C	I
Lower bound	-175	16,057	60
Upper bound	55	96,057	NA

Corruption COR

Original crossover point: 50.004

	E	C	I
Lower bound	-150	15,004	55
Upper bound	50	95,004	NA