Spring 2022

Measuring Political Corruption from Population Outcomes: An alternative to perception measures

Glenn-Iain Steinbeck
Claremont Graduate University

Measuring Political Corruption from Population Outcomes:

An alternative to perception measures

By

Glenn-Iain Steinback

Claremont Graduate University

2022

Approval of the Dissertation Committee

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Glenn-Iain Steinback as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy in Political Science.

Melissa Rogers, Chair

Claremont Graduate University

Associate Dean, School of Social Science, Policy Evaluation

Associate Professor of International Studies

Jacek Kugler

Claremont Graduate University

Elisabeth Helm Rosecrans Professor of International Relations

Mark Abdollahian

Claremont Graduate University

Full Clinical Professor of International Studies

Abstract

Measuring Political Corruption from Population Outcomes:

An alternative to perception measures

By

Glenn-Iain Steinback

Claremont Graduate University: 2022

The impact of corruption is an increasingly important and visible topic for academics, policy makers, and the public. Yet corruption is exceptionally difficult to directly observe and empirical measurements of corruption remain highly contested. Despite the increasing availability of corruption measures and generally high correlations between them, scholars and practitioners disagree over their applicability, interpretation, and the validity of their methods. With the most frequent complaint being that existing corruption indices are largely based on expert opinion surveys, and therefore potentially open to bias and differences of interpretation. Yet, while corruption itself may be ephemeral its aggregate effects are more concrete. Corruption, especially transactional corruption disproportionately harms the least advantaged, decreases access to services, imposes additional costs, and harms well-being. Consequently, the individual level impacts of corruption impose a measurable collective shadow upon society, and when considered together serve as a latent indicator of its presence.

This work applies machine learning to objectively estimate corruption from this shadow, proposing both a novel theoretical approach to assessing national corruption, as well as a new quantitative corruption index. The theoretical approach outlined here begins by identifying measurable consequences, or outcomes of corruption from previous research. It then develops a Python based extreme gradient boosting machine learning model to validate selected corruption outcome measures as predictors of existing corruption indices and estimate predictive factor importance. Finally, a

weighted Z-score index of comparative national corruption is produced and tested as a replacement for existing perception based measures of political corruption.

The results of this analysis demonstrate that political corruption can be predicted utilizing variation in a combination of readily available socio-economic and public health data. Further, the newly defined Outcome Index of political corruption is correlated with existing perception based corruption measures, and performs comparably as a direct replacement in replication analysis. Additionally, the Outcome Index of corruption is both easier to construct than most existing corruption measures, because it obviates the need to conduct survey analysis, and in contrast to existing approaches it is based on objectively measured and readily available data.

Acknowledgements

I would first like to acknowledge my advisor and committee chair, Dr. Melissa Rogers for her support and guidance, both in this project and across my time at Claremont. Her mentoring, advocacy, and most importantly her open door have been invaluable. Without her this dissertation would not have been possible. I am grateful to Dr. Jacek Kugler for the insights and opportunities he has offered. I would like to thank Dr. Abdollahian for his methodological advice, it has made this project substantially better.

I have had the privilege to work with many talented fellow scholars during my time at Clarmeont. I would especially like to acknowledge Simon Tang, and Nicholas Stowell for their support as colleagues and friends. I would also like to thank my parents and brother, their support and encouragement in this and all my academic endeavours has been invaluable. Finally, I would like to thank my wife, Ellie for her advice when I was stuck, encouragement when I was unsure, and patience at all times. I could not have succeeded without her.

Table of Contents

List of Tables

Introduction

Corruption is simultaneously ubiquitous, and yet exceptionally difficult to directly observe. Consequently, while individual acts of corruption are often clear, corrupt systems are much harder to delineate and especially hard to empirically measure. In part, because participants are often reluctant to discuss them and because corruption is rarely reported in official statistics. As a result, despite the increasing availability of measures of cross-national corruption, and commensurately high correlations between them, scholars disagree over their applicability, interpretation, and methodological validity. With the most frequent complaint being that index measures of corruption are at least partially perception based and therefore arguably open to bias and differences of interpretation. Further, existing measures are both complicated to compile and often expensive to collect because they require conducting population surveys or consulting subject matter experts on individual countries. For these reasons existing corruption indexes are both difficult to extend over time, and difficult to scale at multiple levels of analysis. Consequently resulting in limited time-series, especially for developing nations, as well as relatively little comparable sub-national corruption data. This is especially frustrating in the field of comparative politics because it makes quantitative analysis of differential policy adaptations and institutional effects difficult.

This project proposes a solution by theoretically reframing the question of quantifying corruption and suggests a novel approach to estimating it from the shadow it leaves in society. Comparative estimates of national corruption are difficult to produce precisely because corruption resists easy quantification, in large part because there is no objective standard for measuring it and because even the best approaches rely on subjective judgements. In part, because participants, willing or otherwise, are often reluctant to talk about it while outsiders differ over their perceptions of magnitude. It is precisely because of this property that corruption has been labeled an 'odorless gas' (Chayes, 2015). In fact, corruption is arguably one of the hardest social science concepts to measure and amongst the most contested. However, it is possible to improve on this empirical dilemma by refocusing on the consequences of corruption rather than the impression of its presence. This process starts with a simple question: What if instead of asking how corrupt is this country? We were instead to ask: how much does this country suffer from corruption? Essentially, to find the measurable effects of the shadow of corruption in society and use them to estimate the level of corruption itself.

Building on that idea this project proposes, defines, and tests an alternative measure based on objective symptoms of corruption. The resulting index is an objective, outcome-based indicator of corruption derived from measurable and readily available socio-economic covariates which are predictive of variations in political corruption. This study will achieve this in four steps. First, it will identify population level socio-economic, demographic, or health indicators which have been shown to respond to changes in levels of political corruption in existing corruption literature. With the specific requirement that these measures must be objective direct indicators that are widely available and have a clear directional association with corruption. Second, it will test the predictive importance of these measures against existing corruption indexes in order to identify relevant variables. Third, it will demonstrate the compilation of an index from these variables, representing an empirical proxy for unobserved corruption. Finally, this Outcome Index of Corruption will be tested in cross-national analysis as a replacement for existing perception-based measures.

This study adopts the commonly used political science definition of corruption as ”misuse of public office for private gain” (Treisman, 2000). However, it is largely concerned with what is often called ’petty corruption,’ corruption involving interactions between individuals and the bureaucracy in which bribes are used to facilitate or enable access to services; or the use of bribery to influence legislative decision making (Yoo, 2008). Collectively, this is the form of transactional corruption encountered by average people attempting to obtain services or to interact with government officials in the bureaucracy, and felt by the misappropriation or misdirection of government resources away from public interests. This focus was selected because these effects disproportionately impact the majority, especially the poor, or less connected in society, and represent the widest scale social costs of corruption. The remainder of this paper will begin by discussing existing measures of political corruption. It will then outline the theory and requirements for an outcome measure of political corruption. Review the methodology behind the outcome-based approach. Discuss potential variables drawn from existing literature and explore their empirical applicability to predicting corruption. Finally, it will discuss the construction of an index from these indicators and test the Outcome Index as a replacement for perception based measures in a replication study.

1 Literature Review

The general approach to measuring corruption is to attempt to define how corrupt countries are in relative terms to each other. Before suggesting a new approach to this problem it is worth considering how previous measures have conceptualized this task. Existing measures of political corruption are essentially based on one of three broad approaches, institutional assessments, inferences from audit reports or anti-corruption prosecution data, and corruption perception measures (Kaufmann, Kraay and Mastruzzi, 2007). Each of these techniques attempts to infer the actual, or at least the potential level of corruption in a society, but they arrive at their conclusions in different ways and utilizing different theoretical assumptions and methodological tools. This section will review these processes and explore examples and implications of each methodological technique in order to provide an overview of existing approaches to characterizing corruption.

1.1 Institutional Constraint Measures

The first approach, institutional or legal constraint measures of corruption attempt to estimate the vulnerability to corruption, or the potential a given environment has to produce corruption based on variations of the core assumption that environments with more potential for corruption will experience higher degrees of actual corruption. This can be summarized as the: if it's possible, someone is probably doing it inference. A frequently cited theoretical example of this approach is the corruption equals monopoly plus discretion, minus accountability formula proposed by Robert

Klitgaard (Klitgaard, 1998) [1]. In this framework corruption on the part of individual government agents, or agencies is expected to increase when provision of the service is limited (monopoly), meaning that citizens lack alternatives for accessing a public good or privilege. As well as when individual agents have high degrees of decision making discretion over the allocation of that service. Finally, these effects are conceptually expected to be mitigated by the degree of existing accountability, or oversight framework built into a system (Klitgaard, 1998).

For example, a situation in which an individual official, at a single government agency is in charge of both reviewing applications, and issuing a permit, with no regular review or supervision would entail a high degree of monopoly and discretion and a low degree of accountability. Under these hypothetical circumstances the C = M + D - A framework predicts a high potential, or risk for corruption. Indeed, this approach to conceptualizing corruption risk has been used by the World Bank and cited by practitioners in international development and aid projects as a guideline for targeting areas of potential corruption during economic reforms (Harrison, 2007). However, the majority of these implementations involve applications of the formula as a guideline to individual countries, comparative case studies between corrupt environments, or analysis of the success or failure of individual anti-corruption campaigns (Klitgaard, 2013; Mahoney and Klitgaard, 2019). Further, applications using the C = M + D - A framework disagree on how to measure degrees of monopoly and discretion, and focus on oversight at different levels. Consequently, it remains a theoretical framework, or a potentially helpful heuristic for international development projects, but not an operationalised comparative measure of corruption.

A similar, and in fact related approach utilizing observed institutional characteristics to estimate corruption is the Index of Public Integrity (IPI) (Mungiu-Pippidi and Dadašov, 2016). [2] The IPI quantitatively scores the institutional and legal framework of a country in a range between 1,

---

[1] This is often stated as the (C = M + D - A) framework

[2] Scores are available in biennial waves starting in 2015, and cover countries which reported the six constituent measures to the World Bank, and World Economic Surveys.

for highly susceptible to corruption, and 10 for most resistant to corruption. IPI scores are assigned on the basis of features of formal and informal governing institutions, theoretically motivated by the: "idea that corruption reflects a national-level equilibrium between resources and constraints" (*Index of public integrity*, 2021). Methodologically the IPI innovates by attempting to quantify elements of this equilibrium and seeks to blend both subject matter experts assessments, and statistical measures to assess the potential for corruption in a given environment.

To do this, the IPI utilizes 6 indicators: first, Judicial Independence, based on the assumption that more independent judiciaries are less corruptible, operationalized by the Judicial Independence indicator from the World Economic Survey, Executive opinion Survey. Administrative burden, based on the idea that business environments with fewer government imposed administrative procedures reduce the opportunity for corruption, operationalized as the standardized averages of the number of procedures and days required to start a business, combined with the average number of corporate tax payments and time in days to pay each year, drawn from the World Bank Doing Business Survey. Third, trade openness, motivated by the idea that fewer barrier to trade create less opportunity for transactional corruption during the import and export process, quantitatively based on standardised averages of the time and cost of fulfilling import and export compliance from World Bank Doing Business data.[3]

Fourth, budget transparency, based on the idea that increased public transparency about funding reduces the opportunity for corruption, measured by averaging survey responses from subject matter experts about the transparency of government budgets. Fifth, E-citizenship, based on the idea that more public access to information improves the ability of the public to hold leaders accountable, operationalized as the standardized mean of broadband subscriptions, internet users,

[3] Notably the Index of Public Integrity has removed China, Saudi Arabia, the United Arab Emirates, and Azerbaijan from their 2021 rankings and will drop the measures drawn from the World Bank Doing Business Survey from future editions because of data manipulation (Mungiu-Pippidi, 2021). Further, the Doing Business survey itself has been discontinued by the World bank following the revelation that multiple countries pressured the World Bank to manipulate 2018, and 2020 scores (Machen, 2021).

and Facebook users as a percentage of the population. Finally, media freedom measured by the Freedom House freedom of the press score for each country-year, and based on the assumption that free media act to hold leaders accountable for actions which harm the public good (Mungiu-Pippidi and Dadašov, 2016).

The IPI conceptually builds on the basic idea that corruption is enabled by institutional environments where opportunities for official corruption are more accessible, where less transparency exists, or where there is a lack of oversight. However, like the C = M + D - A framework the IPI remains an indirect indicator. It can at best suggest where we might expect corruption to be more prevalent, but it does not guarantee that a country with a lower score is actually more corrupt. Just, that it is potentially institutionally more susceptible to corruption. Further, some of the indicators are questionable, or based on implicit causal assumptions that are not always true or well demonstrated empirically.

For example, requiring more steps to open a business may indeed encourage the use of speed money to expedite the process, but it does not intrinsically ensure that it will in all cases. Similarly, having more internet access does not guarantee that people will use that information to monitor their officials. Extending from this assumption, the use of Facebook users as a measure of social media, or E-citizen engagement while no doubt the result of concessions to data availability, is suspect because Facebook while large is not the dominant social media platform in every country. [4] Further, research on recent election cycles in the United States and the emergence of social-media bubbles as well as the use of both social-media, and traditional media to generate and share fake news content calls into question the assumed underlying relationship between measures of E-citizenship and political accountability (Spohr, 2017; Nikolov et al., 2015). Finally, measuring freedom of the press as an anti-corruption element generally assumes freedom of the press from government interference, but ignores the increasing reality that press freedom is far more compli-

[4] In fact it is no longer the dominant social media platform for adults under 30 in the United States (Auxier and Anderson, 2021).

cated than simple government intervention. Further, it assumes that journalists see themselves as accountability agents.

A further, although more granular example of attempts to quantify susceptibility to corruption based on relative constraints is the African Integrity Indicators (AII) (Integrity, 2018). The AII uses a blind peer-reviewed subject matter expert survey, utilizing in country researchers and an evidence based framework, which requires documenting scoring decisions and a subsequent peer review of each score. Questions focus on thematic areas covering the rule of law, public official accountability, judicial independence, integrity of the civil service, elections, media freedom, and systems for dealing with conflicts of interest involving public officials. One notable and unique methodological feature of the AII is that it divides this analysis into a de jure component and a de facto component, so that it compares both the legal or legislative framework, or requirements established in a given country, and the de facto implementation of those elements in practice(Indicators, 2021).

Finally, the resulting scores for each indicator are aggregated between 0 and 100, and an overall integrity score is calculated based on the average of indicator responses for each country. The AII stands out amongst institutional corruption propensity measures because it combines a survey approach with a more rigorous technical review. Further, the deliberate inclusion of de facto and de jure scoring for each indicator helps to expose incongruities in anti-corruption controls and practices, based on the assumption that large gaps between theory and implementation are more indicative of environments suffering from corruption. However, like all institutional corruption propensity indicators the AII can at best attempt to detect environments which may be more susceptible to corruption, it does not measure actual corruption, and coverage is limited to Africa.

While potentially useful from a policy reform perspective, because institutional constrain measures can tell us where to look or suspect potential corruption in a bureaucratic system, or what elements might be susceptible to corruption, they are at best indirect indicators of corruption and are fundamentally based on assumptions about how the presence or absence of an institution increases or decreases the potential for corruption. They do not measure actual corruption, or directly assess the perception of the actual manifestation of corruption. Further, the are necessarily based on periodic reviews of institutional characteristics, legal implementation, or subject matter expert interviews. As a result, institutional constraint measures are relatively complicated and expensive to update and maintain.

1.2 Corruption Frequency Measures

The second general approach to measuring corruption is based on using audit reports or data on corruption prosecutions within a country to estimate the prevalence of corruption based on analysis of administrative data indicating the frequency of, or responses to corrupt activities. This technique can best be summed up as the, where there's smoke there's fire concept to corruption detection. One of the best examples of this approach used longitudinal data on random audits of Federal Account Transfers in Brazil to identify the frequency of corrupt activity within the bureaucracy (Mondo, 2016). This analysis was made possible by the implementation of randomized lottery style audits of expenditures of federal transfers at the municipal level. Specifically, each year the Brazilian Federal Comptrollers office selects a random sample of approximately 40 municipalities to undergo a detailed audit. The results of which are then publicly released, along with a summary of the investigative materials, thus allowing Brazilians to see both counts, and types of irregularities uncovered by each audit Mondo (2016).

Using this data researchers were able to build up a picture of state level corruption in Brazil by aggregating the results of successive audits. Further, the randomization approach, because the drawing is transparent and no one knows who will be selected before hand, coupled with the fact that the Comptroller's office is highly politically independent makes this a reasonably credible approach in the Brazilian case. Meaning that from an analytical perspective we can assume that the randomized selection of audit targets represents a reasonable identification strategy.

A further cross-national example is the Bertelsmann Transformation Index (BTI), prosecution of corruption indicator, which assesses the frequency of anti-corruption prosecutions in a given country, coupled with survey information about anti-corruption policy and the presence or absence of anti-corruption regulations (Stiftung, 2018). The BTI uses a 10 point scale based on subject matter expert assessments and institutional reviews, ranging from 1, a condition in which the government makes no effort to contain corruption and no integrity measures exist, to 10, a condition in which robust measures are taken to contain corruption and wide spread integrity measures exist (Teorell et al., 2019) Unfortunately, the BTI only covers 'transitional' countries, states which are considered to be in the process of, or have recently undergone democratic transitions. As such, it covers a wide range of Eastern European, African, Latin-American, and some Asian countries, but excludes most of the OECD and is not fully cross-nationally comparable.

Although sometimes presented as 'concrete' measures of corruption behavior (Mondo, 2016). Audit and prosecution frequency data is potentially highly endogenous and likely to be confounded by political motivations. Two countries with equal levels of real corruption could have markedly different scores depending on the independence of their auditing agencies or courts, and the degree of political salience attached to anti-corruption activities. Further, anti-corruption prosecution or auditing may itself be a political act. For example, used to suppress opponents, or redistribute resources amongst the political elite, especially in autocratic environments, or placate political constituencies or suppress opposition parties in weakly institutionalized democracies. Although it is possible to find examples, such as the randomized audits of federal transfers in Brazil which can be plausibly demonstrated to represent transparent random selection by a strongly empowered and apolitical regulatory entity, such a strong identification rule is the exception, not the norm.

Further, any 'concrete' measure of corruption which is based on prosecution data, or auditing results implies a comparison to a counter-factual base line level of either 'acceptable,' or zero corruption, and raises further concerns even when the actual data is believable. Such as, whether the magnitude, or rate of occurrence, relative to the population, or number of officials should be considered. For example, a clean country may be very good at detecting, and prosecuting minor corruption amongst public officials, thus inflating the event count, while an institutionally weaker country of comparable size may both detect fewer corrupt instances, and suffer from larger corrupt act. Meaning that even if we had perfect information on corruption prosecutions the figures would be far from a real concrete representation of the effects of that corruption on society.

1.3 Corruption Perception Measures

Among these three approaches, corruption perception or experience indicators, using population surveys, surveys of organizations, or averaged subject matter expert assessments to gauge the level or type of corruption in a given country are the most widely available, and most often used indicators in cross-national analysis. This can be summed up as the wisdom of the crowd approach, based on the idea that asking a number of sufficiently informed respondents you can build up a picture of the relative prevalence of observed corruption.

Frequently used examples of this approach include the Corruption Perception Index (CPI) produced by Transparency International, an aggregate measure of public sector corruption primarily drawing on the survey opinions of country experts and the business community. (Transparency International, 2020*a*). CPI currently incorporates thirteen survey sources and aggregates questions from those sources around topical areas. Such as, the prevalence of bribery, diversion of public resources, the use of public office for personal gain, nepotism, and state capture (Transparency International, 2020*b*). This source aggregation methodology allows CPI to utilize multiple similar questions from different sources and then statistically standardize the results into a 0-100 range, with 0 indicating a high level of the response variable (Transparency International, 2020*b*). Finally, the aggregated results for each question are transformed to Z-scores and the results are averaged across sources to produce a final estimate for each element of CPI and then averaged again across elements to derive a final 0-100 corruption score for a country-year (Transparency International, 2020*b*).

The CPI approach contributes two notable methodological innovations. First, although they draw on different questions from multiple sources, they focus on survey questions about the actual level of corruption observed or experienced by people and business executives in a given country. Rather than focusing on the potential for corruption, or institutional characteristics. Second, by aggregating up to thirteen separate sources bias in any one source is less likely to effect an overall country score. This approach also allows CPI to adapt to changes in data availability over time, and to add new surveys as they become available, or to composite multiple similar sources over time to adjust when an individual survey becomes unavailable for a given country-year.

In this sense the CPI methodology is actually super-sampling a global pool of corruption experience related survey work in order to estimate the level of perceived corruption. However, it has limited temporal coverage, from 1995-present and has only recently achieved a global sample of countries. Further, this super-sampling method, which Transparency International argues is a core feature of CPI, allowing it to composite multiple sources over time, has itself been challenged on the grounds that the resulting year to year estimates are incomparable because they are based on different source data, different questions, and in some cases entirely different sources (Moiseeva, 2018). An additional notable criticism of CPI is that some of their sources are proprietary, and do not publish information about their own methodology (De Maria, 2008). For example, the World Economic Forum executive opinion survey, and the Economic Intelligence Unit country risk assessments use proprietary methods to derive their own estimates, meaning that at least some of the underlying elements of the CPI components are not fully transparent. Finally, the strongest, and probably most prevalent criticism of CPI, and in fact all corruption perception based survey estimations is that they are survey based, meaning that they measure people's perception of corruption, which may be politically, or culturally influenced, rather than their actual experience with corruption (Klitgaard, 2017).

A second widely used implementation of a corruption perceptions approach is the Variants of Democracy, Corruption Index (Vdem-CI), which uses the average of subject and country expert responses to a series of questions about corruption in the public sector, the legislature, the executive, and among the courts and then weights these sub-indexes equally in the final corruption index(Lindberg et al., 2014). Vdem-CI values range from 0, corresponding to no corruption in any sector, to 1, indicating the highest levels of corruption in all four sub-sector surveys. Because Vdem uses country expert assessments it has the advantage of being able to collect historical scores, from 1946 forward and therefore covers a larger sample with a more uniform set of coding rules. However, it is based entirely on the perceptions of external experts, not individuals who are actually subject to corruption in a given country, and in the case of historical estimates the opinions of individuals who never personally experienced or observed the environments they are rating on the basis of historical documentation, which raises concerns about the comparability of contemporary and historical estimates.

A third notable approach is the Bertelsmann Transformation Index measure for Anti-corruption Policy (BTI). BTI uses detailed country reports prepared by experts coupled with a 10-point scale used by coders to assess the degree of anti-corruption effort undertaken by a government (Standaert, 2015). The BTI is expressly designed to evaluate so called 'transitional' countries, new or emerging democracies, and assess their progress in corruption control as part of the larger process of democratizing. As a result, it covers 2005-present and is actually a measure of the perception of corruption control, rather than a measure of corruption itself. Although, it does have the advantage of a consistent scale and uniform sources.

Another frequently used measure, especially among non-governmental entities, is the World Bank World Governance Indicators, Corruption Control Estimate (WBGI-CCE). The WBGI-CCE uses all available quantitative measures of corruption for a given country in a given year and standardizes their scores to create an aggregate measure of corruption. This results in similar coverage to the the CPI, 1996-present, and more complete global sampling. However, the number of sources used for any given country-year can be as low as 1, or as high as 17, and the actual corruption questions being measured by those sources range widely. Meaning that WBGI-CCE scores are more open to bias, potentially less comparable over time, and may not actually measure the same elements of corruption from country to country (Kaufmann, Kraay and Mastruzzi, 2010). Meaning that the WBGI-CCE is probably best understood as a high-level aggregate measure of general corruption trends.

Finally, a similar, although more methodologically complex estimate is the Bayesian Corruption Index, (BCI). The BCI uses data from 20 corruption perception surveys of citizens, business professional, NGOs, and public officials from 1984-present (Standaert, 2015). Like CPI this super-sampling approach allows the BCI to extend to a wider range of coverage, and utilize as much data as possible in creating estimations, similarly, it allows the BCI to be less sensitive to the loss of any one source for a given country-year. Values are entered in a raw original format rather than using re-scaling, like the WBGI-CCE or CPI, and bayesian estimation is applied to create a relative scale where 0 would correspond to a situation in which all input data for a country was at a minimum, as little corruption as possible, and 100 would represent all inputs at their maximum, as corrupt as possible. This approach has the advantage of being less biased than the WBGI, however, it is still arguably a blunt aggregate measure of generalized corruption because of differences in the underlying source questions, and is open to many of the same criticism as CPI, despite using a more sophisticated statistical methodology for index creation

Table 1: Descriptive Statistics for Corruption Perception Measures

	Obs	Counties	Coverage	Mean	SD	Var	Min	Max
CPI	3251	182	1995-2018	43.06	21.48	461.25	4.00	100.00
Vdem	9514	173	1946-2017	0.51	0.30	0.09	0.02	0.99
Vdem Leg	8617	174	1946-2017	2.02	0.92	0.84	0.17	3.91
Vdem Public	9561	174	1946-2017	1.89	1.00	1.00	0.13	3.93
BCI	5551	195	1984-2017	52.97	15.79	249.35	25.04	93.55
BTI	862	126	2005-2017	4.30	1.96	3.85	1.00	10.00
WBGI	3542	190	1996-2017	-0.05	1.00	1.00	-1.87	2.47

Table 1 presents a statistical overview of these measures. Perception based measures of corruption have received criticism from many different directions. Perhaps the strongest theoretical concern is that the subjective nature of perception surveys leaves the potential for respondent bias (Razafindrakoto and Roubaud, 2010) (Escresa and Picci, 2017). For example, partisan political affiliations may cause citizens to rate governments with whom they disagree as more corrupt. While, people in repressive environments may fear retaliation and not respond honestly. Similarly, subject matter experts could potentially be biased by stereotypes, and reputations of historical corruption. In addition, the use of survey aggregation, using multiple surveys with different questions or structures to expand coverage has been criticized for creating measures that are theoretically ungrounded (Rothstein, 2012). Resulting in generalized corruption measures potentially drawn from incomparable underlying questions and surveys with different levels of analysis or goals.

Further, scholars have questioned whether concepts like corruption have universal cross-cultural meanings which could cause respondents across different national or ethnic backgrounds to interpret identical questions differently (Klitgaard, 2017) (Treisman, 2007). In addition, survey-based methods are expensive to collect, often resulting in limited individual coverage and a lack of large-scale or comparable sub-national corruption data.

Table 2: Correlations Between Corruption Perception Indexes

Indicator	CPI	Vdem	Vdem Leg	Vdem Public	BCI	WGI
CPI	1					
VDem	.87	1				
VDem Leg	.82	.91	1			
Vdem Public	.86	.95	.82	1		
BCI	.92	.82	.79	.80	1	
WGI	.98	.91	.84	.89	.92	1
Observations:	2636					

Despite these concerns, perception based measures, largely because of their wide data availability are the only real option for quantitative analysis of cross-national corruption. Further, there is strong statistical evidence to suggest that these different measures, each utilizing different underlying assumptions, measurement techniques, and index aggregation approaches, as discussed above, do largely agree with each other. Table 2 above shows correlations between CPI, Vdem-CI, the Vdem sub-component for public sector corruption, the Vdem sub-component for legislative corruption, the BCI, and the WBGI-CCE. [5] The lowest correlation is .79, and the highest is .98. Ignoring the Vdem sub-components all of the major cross-national perception based indicators are correlated above .8. In other words, although perception based measures may be imperfect and biased, in the aggregate they do seem to measure some underlying trends in corruption and arrive at similar conclusions through different methodological approaches. They are likely 'on' to

[5] BTI was omitted because it only covers transitional countries and significantly limits the sample.

something, even if scholars do not universally agree with their interpretation or disagree with their individual methods.

Figure 1 presents a standardized time series graph of CPI, VDEM, BCI, and the WBGI corruption indicators in order to further illustrate the relationships between them in the aggregate over time [6]. CPI generally estimates higher levels of average corruption, suggesting that all else equal CPI assigns higher scores relative to the range of the indicator 0-100, and implies higher general levels of global corruption. CPI also displays the most variance over time. Although, some of that variance is likely attributable to changing sample size and nonrandom initial sampling of included countries. CPI grew from covering 87 countries in 2000, to 128 in 2003, and 154 in 2005, before finally reaching a fully comparable global sample of 171 by 2012.

Conversely, VDEM and WBGI have exhibited a slow convergence toward each other, with general corruption estimates starting at approximately half the range for each scale and remaining largely stable in the aggregate, with a slight downward trend for VDEM. BCI suggests the lowest levels of average global corruption, accounting for slightly less than half of the potential 0-100 point theoretical range with nearly flat average performance over the last decade. This highlights that even though these measures are highly correlated and tend to vary together they do make slightly different judgements about the aggregate level of corruption. Notably, there is a substantial gap, nearly 20 relative points between average corruption levels according to CPI, verses average corruption according to BCI. Put more simply, CPI implies a notably higher level of background corruption in the world than does BCI. Finally, with the exception of CPI, perception based corruption measures tend to have low variance in the aggregate, with relatively flat trend lines over time as shown here.

These differences imply that the world according to CPI is a relatively more corrupt place,

[6] CPI was reverse coded so that higher values correspond to increasing corruption. While WBGI was re-centered to run in a positive range. Values were standardized for comparison as a proportion of the theoretical range for each indicator

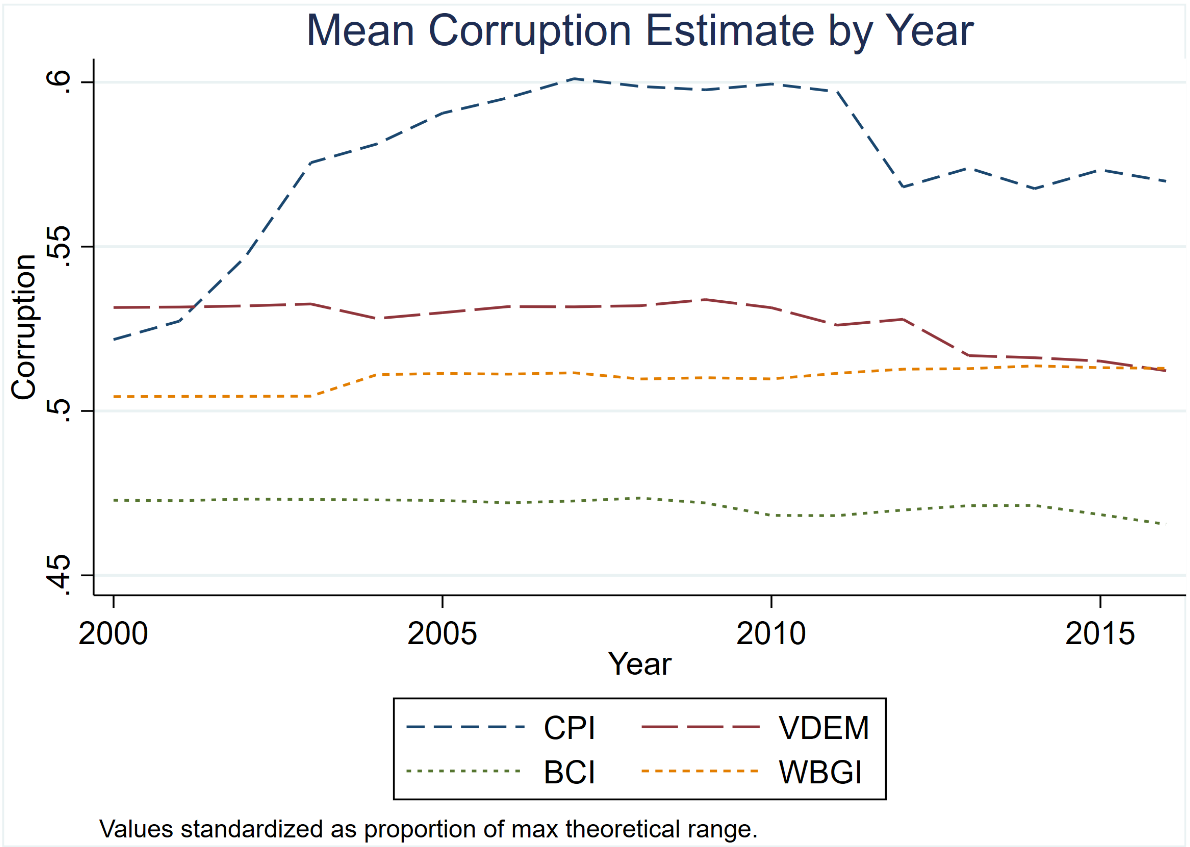


Figure 1: Standardized trend over time for CPI, VDEM, WBGI, and BCI

than the world according to BCI. However, despite differences in the relative levels of corruption implied by these measures they are highly correlated as shown in table 2, and they do tend to vary together. Collectively, this provides strong evidence that existing corruption measures can be used as a bench mark for the alternative outcome based corruption index proposed here. Regardless of imperfections perception based indices are the closest available cross-national estimates of 'real' corruption. The next section will outline the theoretical intuition behind this outcome based indicator of corruption.

2 Theory

2.1 Theoretical Solution

The preceding section discussed existing methods for measuring and modeling corruption including institutional constraint models, corruption frequency measures, and corruption perception based approaches. Among these approaches constraint models effectively measure a counter factual theoretical potential for corruption, rather than corruption itself. In contrast, frequency measures are based on observed corruption events, but they suffer from issues of comparability, data availability, interpretation and trust. Finally, perception based measures evaluate opinions about the prevalence of corruption. Consequently, given their popularity amongst researchers and strong correlations between independent methods perception measures are simultaneously the most reasonable option for quantitative analysis, and the most frequently criticised approach. One solution to this collective problem is to change how corruption itself is measured by attempting to detect the symptoms of corruption, or the latent effect of corruption as an outcome in other indicators. Essentially, to infer the unobserved presence of corruption from the actual quantitative shadow it leaves in other data.

An objective outcome based corruption index based on outcome measures has several potential benefits over existing techniques. First, it bypasses the subjectivity concerns inherent in existing cross-national measures by focusing on socio-economic indicators that have demonstrated associations with transactional corruption. This avoids concerns over respondent bias, cross-cultural interpretations or comparability of survey results, the reliability of subject matter experts judgements, or the larger issue of measuring perceptions of corruption rather than the prevalence of corruption itself. Instead an outcome measure would use changes in population level socio-economic indicators, and the identified causal relationship between corruption and the chosen indicator to directly estimate changes in political corruption. Meaning that an increase in an indicator that is known to correlate with higher levels of corruption would be indicative of an increase in estimated corruption in that country. Simply put, an index of multiple outcome measures could serve as a direct proxy for changes in unobserved corruption.

Moreover, an outcome measure of corruption would offer several additional benefits. First, if it was based on commonly available data it could provide better coverage and overcome the expensive, and time consuming problem of administering corruption perception surveys. Additionally, an outcome based measure would bypass the problem of compositing similar but non-identical survey waves, or the need to develop country specific corruption identification strategies. Second, an outcome based index would allow for extension to the sub-national level. Allowing more effective comparisons of corruption within countries or between regions and more comparability between sub-national corruption research. This would be a significant advantage because sub-national corruption research is currently largely restricted by a lack of data, or the need to create novel, country specific corruption measures. This could potentially be overcome with a corruption index composed of more readily available socio-economic data. The remainder of this section will discuss the theory underlying this approach and evaluate the requirements for potential indicator selection.

2.2 Outcome Measure Theory

The underlying theory behind an outcome measure of corruption is the idea that although corruption itself may be essentially invisible, in the aggregate, the effects of corruption, or the shadow it leaves in other indicators can be detected. Consequently, relative changes in corruption can be predicted based on variation in identified population level 'outcomes,' either individually, or ideally using multiple indicators in an index format. Consequently, corruption outcomes are factors that are indirect indicators of the effect, or presence of corruption. Such that a change in an outcome indicator would be indicative of a change in unobserved corruption. Conceptually this approach is very similar to the use of night time geo-spatial illumination data to estimate the level of economic activity, when official statistics are either unavailable, or unreliable (Henderson, Storeygard and Weil, 2012) (Hu and Yao, 2019). Changes in luminosity are the observable outcome, resulting from more industrial output, of the unobservable level of GDP. Although we can not detect corruption with a satellite, the core idea here is that other indicators that are related to it could be used to estimate it's presence through the shadow it leaves in other data.

In order to be considered these potential outcome indicators would need to meet several requirements. First, they would need to be based on objective, empirical measures with clear criteria. For example, mortality rates, census data, or government expenditures. This requirement automatically excludes any indicator that is based on subjective judgements, expert coding schemes, and unclear or negotiable self-reporting. For example, satisfaction with government performance, survey based assessments of freedom of the press, or categorical subject matter expert classifications of quality of governance. Second, potential indicators must be empirically associated with changes in corruption and must have a clear theoretical causal directional association with corruption. Existing scholarly work must have demonstrated an empirical relationship and proposed a plausible causal mechanism. Third, in order to provide a generalize able index of cross-national corruption potential indicators must be broadly applicable across levels of economic development, as well as government and regime types. For example, political features, such as measures of competition based on the number of effective parties would not be comparable across democracies, because of variations in electoral rules, and would be generally inapplicable to autocracies. Finally, potential indicators must be widely available, both cross nationally and ideally over time. This requirement excludes indicators that are only collected for a certain geographic region, such as OECD statistics, or for certain categories or classifications of countries. As well as indicators which have short time-series availability or are collected in intermittent cross-sectional waves. Building on this, the next section will overview possible indicators, discuss indicator selection, and review the methodology for estimating a corruption outcomes indicator.

3 Methodology

The overall process of building an outcome index of corruption is divided into four major parts. First, identification of potential indicators. Second, testing of these indicators in order to determine their predictive value. Third, the most predictive indicators will be combined into an index. Finally, the resulting index will be applied to longitudinal cross-national analysis as a replacement for existing perception based measures. The following section will address the first three steps.

Methodologically, the first step is to identify potential outcome indicators from corruption literature and sort the resulting candidates to find indicators which match the theoretical requirements outlined above. Second, once a pool of appropriate potential indicators has been identified they must be tested to determine which indicators are predictive of changes in corruption. This will be done by testing these variables against the existing cross-national corruption indexes identified in chapter 2. With the objective of determining which factors are generally most predictive of changes in corruption. One of the best ways to do this is to utilize tree based machine learning methods which specialize in determining predictive variable importance, and estimate predictions out of sample, by successively taking each existing measure of corruption as the dependent variable and attempting to predict it with the identified outcome measures. Supervised machine learning methods are especially appropriate for this task because they reduce assumptions regarding the exact functional form, and allow for nonlinear relationships between the variable of interest, in this case corruption, and potential predictors (Steinwart and Christmann, 2008). In other words, machine learning approaches allow this question to be reduced to an optimization problem in which a group of theoretically related indicators are algorithmically applied to predicting the outcome, without needing an exact a priori characterization of their functional relationship.

Therefore, since this is essentially a variable importance and feature selection question, extreme Gradient Boosting (XGBoost) is ideal for this purpose. XGBoost was selected because it provides accurate predictions using iterative rounds of boosted decision tree algorithms, and because it consistently outperforms both linear regression, and alternatives, such as random forest machine learning implementations (Chen and Guestrin, 2016*a*). Additionally, XGboost handles multicollinearity in identified outcome measures (Chen et al., 2018) . This has the advantage of allowing the evaluation of potentially similar, but not fully mutually substitutable outcome measures.

XGBoost is used here in two ways. First, to evaluate the predictive value of the identified candidate indicators against alternative measures of Corruption. Second, relative feature importance weights will be calculated from the XGBoost results in order to determine variable weights for the final index. Once predictors have been selected and tested, the remaining, or most important variables will be combined into an index using a modified version of the simple country average method used by CPI, with the addition of multiplicative feature importance weights. Finally, the resulting index will be tested as a substitute for perception based measures in panel regression. The next section will discuss the search for potential indicators.

3.1 Potential Outcome Measures

As discussed previously a potential outcome indicator for corruption must be an objective measure. It must have a clear theoretical causal relationship with corruption as well as a demonstrated empirical association with changes in corruption. Finally, there must be available data to measure, or proxy the outcome. Based on these requirements corruption literature was searched with special attention to measurable results of corruption, specifically indicators that have been shown to change as a result of changes in the level of corruption. A wide variety of variables were considered. However, the most successful indicators satisfying the above requirements proved to be population level health related measures. Table 3 provides descriptive statistics for the candidate outcome indicators which were found to satisfy these criteria and selected for further investigation.

3.2 Corruption Effects

A relatively large body of public health literature has explored the relationship between corruption and health and repeatedly shown that more corrupt political environments have less effective

Table 3: Descriptive Statistics: Potential Indicators

	obs	mean	sd	Var	min	max
Infant Mortality	8793	50.52	45.79	2096.70	1.60	240.50
Under 5 Mortality	8793	75.26	77.83	6057.52	2.10	441.90
Adult Mortality	3040	191.89	114.14	13028.39	49.00	697.00
Measles Immunization	6748	78.01	22.25	495.22	1.00	99.00
DPT Immunization	6871	79.06	22.79	519.43	1.00	99.00
Health Spending	2834	3.24	2.33	5.42	0.00	24.11
Education Spending	2402	14.63	5.05	25.46	0.00	47.28
GNIPC	7501	7290.75	13798.69	1.90e+08	40.00	186080.00

healthcare systems and less healthy populations overall (Factor and Kang, 2015). Further, health systems are highly vulnerable to corruption as a result of information asymmetry, overall complexity, large numbers of stakeholders, and built in consumer incentives to pay to expedite or ensure treatment when faced with shortages (Savedoff and Hussmann, 2006). Corruption has been argued to affect the health sector in three primary ways, financial corruption, medical supply loss, and human resource issues. (Kirigia and Diarra-Nama, 2008). The effects of medical financial corruption, include diversion of health budgets, administrative embezzlement, and bribery to affect procurement, regulatory, or legislative decisions. Medical supply loss encompasses the theft of medical equipment, treatment supplies, or drugs, often for re-sale, or redistribution. While human resource corruption includes providers charging for services they never rendered, demands for illicit payment for access or services, or speed money payments by patients or their families attempting to jump the queue or secure favorable treatment.

One of the most consistent findings in prior literature is that corruption increases infant mortality rates. Gupta et al. analyzing a decade of data from 89 countries concluded that public corruption increased child and infant mortality rates, resulted in lower birth rates, more frequent births without medical facilities and lower birth weights (Gupta S, 2002). Theorized to be a result of the effect of bribes or other expectations of irregular payments to receive services, treatment, or to gain access to medicine. Gupta further hypothesized that this accentuates existing inequality, creating an environment in which the poorest members of society, who are most likely to need healthcare access cannot afford to regularly access it or are consequently incentivized not to do so.

More recent work using a structural equation modeling approach extended this analysis, arguing that corruption has both indirect and direct effects on public health (Factor and Kang, 2015). Simultaneously diverting resources away from healthcare spending at the legislative level, the indirect effect. Frequently toward more easily captured public services based on the finding that more corrupt societies spend less on healthcare as a percentage of GDP. While also directly resulting in lower life expectancy, increasing infant mortality, and decreasing immunization rates, as a result of transactional corruption in the health sector itself (Factor and Kang, 2015). A relationship between infant mortality and corruption has been further illustrated over the long-term in in-depth country studies. In one of the most rigorous examples Dincer and Teoman utilize a novel corruption indicator based on textual analysis of fifty years of news articles in Turkey (Dincer and Teoman, 2019). They argue that corruption has a synergistic relationship with inequality, in which it both increases inequality because low income citizens pay a larger percentage of their income in bribes, and decreases their healthcare access; because they are less able to pay the expected bribes for healthcare or afford the expensive private options that evolve to serve the elite. Their empirical time-series results for Turkey show a causal relationship between periods of increasing corruption and higher levels of infant mortality (Dincer and Teoman, 2019).

Connected to this corruption has been found to have a negative association with childhood immunization rates (Gupta S, 2002) (Factor and Kang, 2015). This is generally a result of corruption in vaccination programs where payments are based on the number of doses recorded, theft of vaccination supplies for illicit trade or resale, or the misdirection of limited vaccination supplies as a result of bribery. This problem has been so pernicious in Pakistan that it has been identified as a leading cause of the country's failure to fully eradicate Polio (Closser, 2019). Largely as a result of falsified vaccination schemes where health authorities misdirect funds intended to purchase or administer vaccine doses. Collectively, these mechanisms can result in shortages of vaccines, a potential for extortion of patients, or demands for payments to administer vaccines, even when the vaccinations themselves are supposed to be free. Resulting in overall lower population vaccination rates.

Moving to the larger question of the effect of corruption on government social policy, corruption has been shown to reduce spending on education, social services, and healthcare (Rose-Ackerman and Truex, 2012). As well as to correlate with lower levels of human development, specifically life expectancy, educational attainment, and standard of living (Akçay, 2006). Based on the theoretical argument that once corruption becomes engrained in society government officials prioritize opportunities for rent seeking and reduce spending on health, education and social services. Building on this more recent work has asserted a causal relationship between corruption and persistent poverty and increasing inequality (Justesen and Bjørnskov, 2014). Suggesting that corruption creates systematic disadvantages for poorer or less connected members of society and establishes a self-reinforcing vicious cycle. Similarly, countries with larger amounts of corruption have been shown to have larger shadow economies and larger and more sophisticated levels of organized crime (Dell'Anno and Teobaldelli, 2015) (Stinson et al., 2013).

An additional potentially interesting indicator is road deaths and motor vehicle accident fatalities because it speaks to the impact of corruption on infrastructure and regulation, and because enforcement in these areas is logically vulnerable to corruption. Controlling for the level of economic development the relative level of corruption has been shown to impact traffic accident fatality rates (Anbarci, Escaleras and Register, 2006). (Hua, Noland and Evans, 2010) examine this relationship more thoroughly, suggesting that higher levels of corruption are associated with increasing motor vehicle deaths as a result of three major factors, undermining enforcement, decreasing regulation of safety standards, and reduced policing of driving practices. This covers a range of traffic and road safety related issues, from bribing motor vehicle departments to license unsafe vehicles, to paying to obtain driving licenses or shortcut the permit process, resulting in less skilled or unsafe drivers, as well as bribing traffic officers themselves. Collectively, these outcomes could logically directly result in higher rates of traffic fatalities, as well as higher adult mortality rates more generally.

3.3 Indicator Selection

This body of literature suggests several potential indicators that could be used to infer changes in corruption levels. Including, infant mortality rates, the percentage of underweight births, the ratio of unattended births, economic inequality, health spending as a percentage of GDP, immunization rates, education spending and attainment, living standard, the size of the shadow economy, prevalence of organized crime, and road fatalities. This section will briefly explain which indicators were selected or rejected and how they are operationalized.

Infant mortality was selected both because of the strong association and clear causal relationship documented between transactional corruption in the health system and the expectation of higher infant mortality by previous literature. Theoretically speaking this is based on the idea that since infants are often the most vulnerable members of society any deleterious effects of corruption which restrict health access are likely to manifest in higher infant mortality statistics. Further, infant mortality is a widely used measure with excellent data availability and clear reporting criteria. Along these lines, two potential measures were adopted. The infant mortality rate per 1000 live births, defined as deaths under 1 year of age from the World Bank, World Development Indicators (World Bank, 2018). As well as the under 5 mortality rate, the number of deaths under the age of 5 per 1000 live birth cohort (World Bank, 2018).

Low birth weight rates and unattended birth rates were considered but not adopted because of a lack of sufficiently comparable cross-national data, and a greater potential for confounding effects due to the granularity of these measures. Domestic health spending as a percentage of GDP was also selected based on the association between corruption and the misallocation of public health resources, as well as observed lower levels of health spending in more corrupt countries reviewed above. Theoretically, this is based on the idea that corruption diverts resources away from public goods, such as healthcare and the consequent expectation that more corrupt countries will spend less on health. This was operationalized with World Bank data on domestic government health spending as a percentage of GDP (World Bank, 2018).

Further, although numerous articles find an association between increasing inequality and corruption, and although corruption almost certainly increases inequality, inequality itself was not adopted as an outcome measure. This is for two reasons. First, inequality is often measured with GINI coefficients or 20/20 ratios. However, the data for these indicators is frequently poor and there are concerns about their accuracy. Most importantly, there are multiple potential confounding factors that could also increase inequality but are not theoretically related to corruption, defined as misuse of public office for private gain. For example, governments, especially conservative ones, may select fiscal and taxation policies which increase inequality but are not necessarily the result of corruption or indicative of it.

Immunization rates were selected because of the demonstrated empirical association between corruption and lower levels of immunization in previous studies. Theoretically, this is based on the idea that corruption redirects resources away from public health, or encourages their theft, and the consequent expectation that more corrupt environments will have lower basic immunization rates. This was operationalized in two ways. First, with the diphtheria, pertussis, and tetanus (DPT) immunization rate for children under two years of age (World Health Organization, 2018). DPT was selected specifically because it is a critical childhood vaccination and is used around the world and because it was specifically found to be reduced by corruption (Factor and Kang, 2015). Second, the measles, mumps, and rubella (MMR) immunization rate for children younger than two years of age (World Health Organization, 2018). Although these measures are likely to be highly correlated, especially in developed countries, MMR vaccination rates were considered because it is also a widely used vaccine which protects against multiple serious childhood diseases and data is widely available. Theoretically speaking, with the exception of vaccine deniers both parents and public health officials around the world have a strong incentive to maintain high vaccination rates for both DPT and MMR when possible. Consequently, lower rates of vaccination in these categories, controlling for economic development, are plausibly a symptom of corruption.

Education spending as a percentage of GDP and educational attainment are also shown to be related to corruption. Educational spending as a percentage of GDP is available from the World

Bank. Unfortunately, it was excluded because the data is often intermittent and missing for some countries for various intervals of time. Because of this, attempting to include it reduced the overall sample in regression analysis by half. Educational attainment is an additional interesting indicator, and the above literature suggests that educational attainment rates are plausibly related to corruption. Both as a result of misallocation, and potential transactional corruption in education itself. However, educational attainment data is notoriously poor in cross-national data sets and is often characterized by missing data, and limited data for developing countries. Further, the relevant level of educational attainment logically depends on the level of economic development. For example, in a less developed country elementary educational attainment may be the most important, while in a developed country college educational attainment rates are likely to be more relevant. For this reason, attainment is difficult to proxy in the context of cross-national corruption and these measures were dropped from consideration.

Similarly, the standard of living has been measured in past research with the Human Development Index (HDI), as well as with purchasing power parity measures. HDI is limited to 1990 to the present and includes gross national income per capita, which is separately used as a control in the models described below. Further HDI itself is an index, and could both reduce the interpretability of the proposed outcome measure of corruption and complicate the process of producing the final index measure. While purchasing power parity as a measure of the living standard is arguably associated with a range of macro-economic factors that are not corruption related, rendering it a poor theoretical predictor of changes in corruption. As a result, both measures were discarded from consideration. Similarly, although the prevalence of organized crime is quite likely to be related to corruption and it would therefore theoretically increase with increasing corruption, measuring it is fraught with difficulty. Like corruption itself organized crime is an illicit activity, because of this it is difficult to quantify and most cross-national assessments are based on legal or policing data which is likely to be endogenous to corruption. Similar concerns arise with any effort to use the relative size of the shadow economy or gray market. As a result, both measures were excluded from consideration.

Finally, some additional indicators were considered, but were excluded for various reasons. For example, economic growth rates and the presence or absence of growth-oriented policies do appear in corruption literature. However, it was excluded as a potential outcome measure because it is likely to be sensitive to other economic or political forces. Similarly, literacy rates have been used as a proxy for education in some corruption related research. However, it is both a blunt measure which is likely to be less applicable to more developed countries, and for various political reasons subject to falsification, especially in autocratic environments. One final indicator of note is traffic fatalities. As discussed above there is good empirical and theoretical support suggesting that changes in traffic fatality rates, adjusting for economic development, are predictive of corruption. Other research has used the prevalence of road or traffic deaths per 100,000 people. However, this data does not exist in a single easily accessible source and achieving a comparable global sample is difficult.

Following this process, the resulting pool of potential outcome indicators for the creation of a corruption index include: infant mortality, under five mortality, adult mortality, MMR and DPT immunization rates, and government health spending as a percentage of GDP. Gross national income per-capita was included as a control for economic development and because income has been widely shown to correlate with lower levels of corruption in previous studies. GNIPC was selected based on its use in other corruption work and because it can easily be converted to World Bank development classification levels, thus allowing a further division of the sample by level of development if desired. Further, it must be noted that it is impossible to entirely eliminate endogeneity in these measures, as well as the potential for external confounding effects - the goal is to identify a package of indicators that jointly, when indexed together would capture changes in corruption. None of them are likely to be individually perfect. Finally, although this package includes numerous health measures the objective is not to create a measure of healthcare corruption. Instead health measures are being utilized because previous research has identified lower levels of health attainment at the national level as a consequence of corruption. Therefore, it is theoretically reasonable to expect that poorer health outcomes should be correlated with, therefore potentially predictive of higher levels of corruption.

3.4 Outcome Example: Infant Mortality

This section demonstrates an example of the outcome indicator approach in practice, considering an initial evaluation of the feasibility of the approach by centering on the relationship between infant mortality and corruption as a means of demonstrating and explicating the underlying logic. As discussed above increasing infant mortality is associated with higher levels of corruption in existing literature. Put more directly, increasing infant mortality has been demonstrated to be an outcome of higher levels of corruption. The proposed outcome measure approach inverts this relationship and attempts to take variables which are known to respond to changes in the level of corruption and uses them to predict corruption directly with the ultimate objective of producing an index from a group of related variables which could be substituted for a corruption measure.

$$Corruption = \alpha + \beta_1 InfantMort_t + \beta_2 GNIPC_t + \beta_3 Year + \varepsilon \tag{1}$$

For infant mortality, the underlying theory is that since infants are the most vulnerable in society, disturbances in health services, for example because of corruption, are likely to increase infant mortality. Further, increases in infant mortality should be relatively free from positive political effects, everyone should want children to survive - there is not sizable anti-infant vote. Meaning that, politicians may rationally desire to increase inequality or other similar factors under some circumstances, but increasing infant mortality cannot be confounded by rational self-interest policy. They may make policy decisions which have the unintended effect of increasing infant mortality, but increasing it in, and of itself is unlikely to be a direct policy objective. Further, infant mortality is likely to be reduced by better quality of healthcare, better nutrition, or other factors associated with development, this can be controlled for by including a measure of income or development. Equation 1 shows an example of this approach, predicting a measure of corruption as a function of infant mortality, GNIPC and a control for time using robust standard errors clustered by country.

Table 4: Panel Regression Results: DV Corruption

Logged	CPI	VDEM	BCI	WGI
Logged Infant Mortality	-0.120**	-0.343***	-0.0987***	-0.122**
	(0.0378)	(0.0719)	(0.0257)	(0.0413)
Logged GNIPC	0.180***	0.116*	0.0827***	0.132***
	(0.0267)	(0.0499)	(0.0175)	(0.0279)
Constant	2.638***	0.00496	3.685***	0.247
	(0.324)	(0.582)	(0.215)	(0.342)
Randomization P	0.0000	0.0000	0.0000	0.0000
Randomization T	-.1196992	-.3432238	-.0987178	-.1224618
Observations	3314	7113	5070	3706
Adjusted R^2	0.658	0.431	0.574	0.605
Countries	178	170	187	185
RMSE	0.280	0.583	0.187	0.250

Clustered Robust Standard errors in parentheses.
Dummy year variables not shown.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4 shows the results of this analysis for various cross-national perception measures of freedom from political corruption. The full Vdem measure and the BCI measures have been reverse coded so that they match the approach used by CPI, higher values indicate more freedom from corruption, or 'cleaner' countries. This means that a positive regression coefficient actually indicates less corruption, while a negative coefficient indicates more corruption. Further, the specification proposed in equation one has been modified to a log-log model, both to correct GNIPC for normality in a linear model, and to allow the results to be interpreted more clearly as percentage point changes. As shown infant mortality is negative and significant across all corruption measures. Suggesting that increasing infant mortality decreases freedom from corruption, controlling for economic development. Meaning, that infant mortality rates are predictive of corruption in regression analysis, with higher levels of infant mortality driving increasing corruption.

These results were further validated using randomization inference, implemented with the ritest command for Stata (Heß, 2017). Each specification was repeated 1000 times, randomly permuting values of infant mortality each time, and the probability of acquiring the original result was calculated from the distribution of potential outcomes. As shown, the randomization P value for each model is significant. Indicating that this is a rare result within the distribution of permuted results and that they are unlikely to be achieved by chance. Based on these results, if we accept that an indicator like infant mortality predicts controlled changes in corruption then we can infer that changes in infant mortality could be used to predict corruption indirectly. If additional similar indicators could be identified, this process could be used to produce an outcome index of changes in, or the presence of corruption. This will be explored in more detail in the next section by applying machine learning techniques to explore the full list of variables identified in Table 3.

3.5 Evaluating Indicators

Fundamentally, the question of identifying and testing potential indicators for inclusion in an outcome based index of corruption is a problem of prediction. We want to know what factors predict the level of, or changes in corruption in order to evaluate outcome measures and select variables for inclusion in a final index. Conceptually speaking, the intuition here is to identify a group of objective variables which are theoretically associated with changes in corruption, because they have been shown to increase or decrease in response to changing levels of corruption in past literature and then reverse the process by using those corruption outcomes to estimate corruption levels themselves. Essentially, to predict corruption levels from the shadow it leaves in other related indicators. If this can be done successfully, then a plausible case can be made for using an index constructed from those predictors as a measure of correlated, but unobserved corruption.

Building on this idea, machine learning methods which specialize in determining predictive variable importance and estimating predictions out of sample are ideally suited for this problem. One of the best current algorithmic approaches for this task is Extreme Gradient Boosting (XGBoost). XGBoost is a supervised ensemble learning technique which functions by building iterative decision trees, examining the predictive error from each tree, and then fitting a new tree to attempt to explain what the last one could not (Friedman, 2001). This process continues over multiple rounds with the algorithm attempting to minimize the error term and aggregate a better predicting model from a number of individually simple 'weak' learners (Chen and Guestrin, 2016*b*). Like other machine learning algorithms XGBoost functions by dividing the data into a training set, on which the algorithm is trained and it's parameters are incrementally tuned to explain the data, and a testing set, on which final predictions and model evaluation are performed. An additional advantage of XGBoost for this application is that it is robust to multicollinarity amongst predictors(Chen et al., 2018). This is important here because it allows the use of multiple correlated features, such as healthcare spending and vaccination rates, both of which are individually theoretically linked to changes in corruption, but would likely be too highly correlated to function in a traditional statistical model.

XGBoost excels at prediction problems involving structured data and it is especially good at predicting outcomes where theoretical associations are known, but causal pathways are unclear. Academically, it has been used to estimate medical risks from patient data (Noh et al., 2021), to forecast bond default rates (Zhang, Chen et al., 2021), and to predict geological changes over time (Ibrahem Ahmed Osman et al., 2021). In addition to being the top predictive modeling approach for Kaggel predictive data science competitions (Becker, N.d.). This paper implements the XGBoost package for Python in three ways (Chen and Guestrin, 2016*a*). First, it fits a simple gradient boosted model using the predictors in Table 3 to estimate each of the four previously identified corruption metrics and compute feature importance to roughly judge how applicable the identified predictors are to estimating alternative corruption measures. Second, it applies a fully trained gradient boosting model to predict a sample for each corruption measure as a means of estimating the collective applicability of these variables to proxying changes in unobserved corruption. Finally, it derives averaged relative feature importance values from the trained model to use as weights in the final index construction. Note, that the objective here is not to develop the best predictive model for any given corruption measure. For example, this section does not seek to develop a model which could be used to predict future levels of CPI, or any of the other measures directly as an end goal. Instead the intention is to demonstrate that the package of explanatory variables is highly correlated with (can predict) corruption with reasonable accuracy, and therefore can be used as the basis of a new indirect index measure of corruption.

Figure 1 shows estimated feature importance graphs for CPI, WBGI, BCI, and VDEM created with the plot importance function from the Python XGBoost library. The F score indicates how frequently each predictor was used when splitting trees, this approach to feature importance is based on the idea that more frequently used predictors have more explanatory value. The initial results are encouraging because they are quite stable across all four alternative measures of corruption and have similar relative magnitudes compared to other predictors in each model. Although the order of some predictors, such as the immunization rates and under 5 mortality shift between models none of them drop significantly, or fail to predict for a specific corruption measure. Indicating that as a whole, government health spending, infant, under 5 and adult mortality, and immunization rates have generalizable explanatory power across alternative specifications of corruption. In addition, World Bank economic classification levels, and education spending are shown not to be predictive, and can consequently be eliminated from further consideration.

Building on these results the data for each alternative corruption measure was partitioned into training and testing sets using a randomized 80/20 split and separate XGBoost learners were trained on each set. The results are summarized in table 5 below. Because the four corruption measures considered here are continuous this is a regression style predictive problem. A squared error objective function was implemented for the model and hyperparameter tuning was carried out as suggested by the XGBoost documentation (Chen and Guestrin, 2016*b*). Mean absolute error (MAE) was Utilizing as the model training evaluation metric, meaning that for each parameter tuning or boosting round the algorithm sought parameter settings which minimize the MAE, the difference between observed and predicted values of Y.

The XGBoost learners were trained in several steps. First, the optimum number of boosting rounds was determined by allowing the algorithm to run repeated estimates using cross-fold validation until runs converged on a boosting number with the minimum possible MAE on the training data. Second, tree depth and complexity was optimized by tuning maximum node depth and minimum child weight. These parameter control model complexity by limiting tree depth, how many

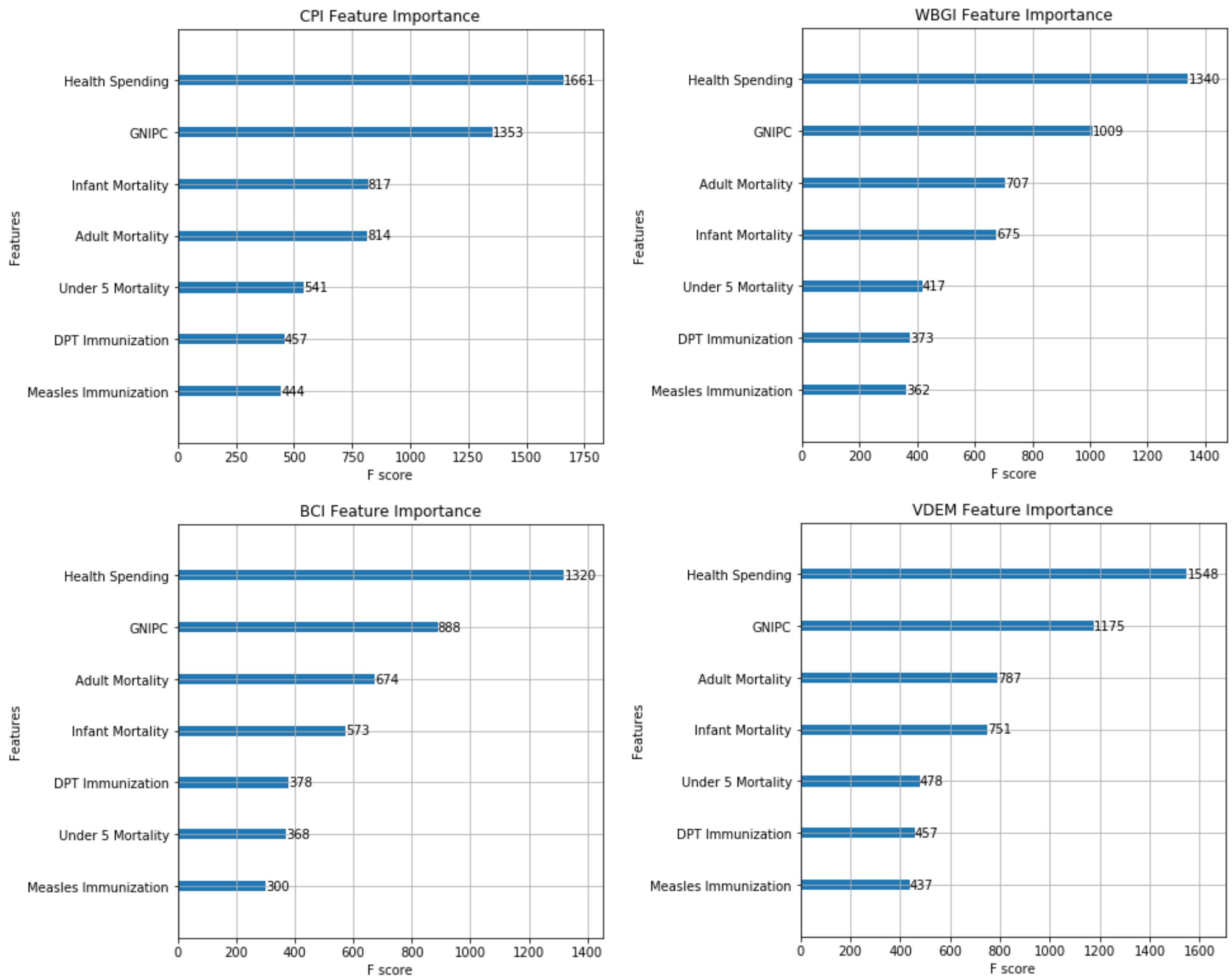


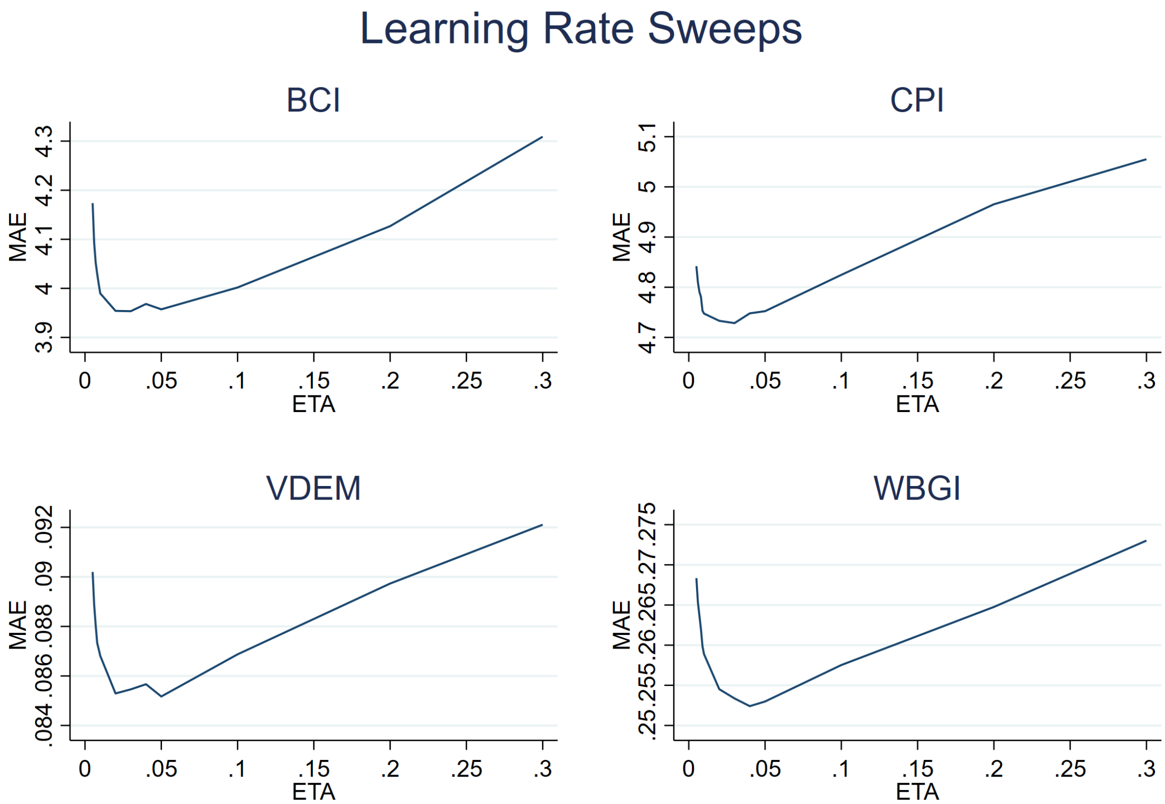
Figure 2: Feature importance for CPI, WBGI, BCI, and VDEM

splits deep the trees are allowed to grow, and the sampling threshold for creating new leaves on each tree. These parameters were determined using broad parameter sweeps and K-fold validation to find the combination with the lowest MAE. Sweep ranges between 2-10 for depth, and 1-10 for child weight were used, and parameter ranges were increased until neither parameter bound to either extreme within the potential range.

Next, the subsampling and column sampling parameters were tuned using cross-fold parameter sweeps between 0, and 1, in .01 increments in order to find the optimum row and feature numbers to sample for each boosting round. Optimizing these parameters helps to reduce over-fitting by managing the proportion of observations, and features used in re-sampling for each successive tree. Following this, The alpha L1, and Lambda L2 regularization feature weight parameters were tuned by incrementally modifying the parameter sweep range until the selected weights were no longer at one of the parameter range extremes. The final parameter sweep ranges selected were 0-.05 for L1 regularization, and 0-.15 for L2 regularization. Finally, the learning weight parameter, which governs how aggressively the algorithm attempts to correct after each boosting round was tuned utilizing a final round of k-fold cross validation parameter sweeps between .005 and .3, once again using MAE as the evaluation criteria. Learning rate sweeps with XGBoost are conducted from largest to smallest, and the final range adopted was .3 - .05 in .01 increments and the learning rate with the lowest MAE was algorithmically selected for each model.

In order to guard against over-fitting, the learning rate (ETA), and MAE was recorded for each sweep during parameter tuning and a graph of the association between ETA and MAE for each corruption measure was produced as shown in figure 2. All four show the expected J-curve relationship, with error terms increasing with larger learning rates, and all four models converge on a learning rate at or below .05. There is some variance in MAE results for Vdem, and BCI, but the actual magnitude is negligible, there is a definitive identifiable learning rate with the lowest MAE, and the overall trend describes the expected relationship between higher learning rates and larger error terms. These results provide strong evidence against model over fitting for the learning rate parameter.

The full set of tuned hyper parameters were then fed into the complete training model and it was run one more time to validate the optimum boosting rounds under the conditions with fully tuned parameters. The trained XGBoost learner was then run on the retained test data, which it had never seen during training, and final out of sample mean absolute error and root mean square error coefficients were calculated against test data for each corruption measure. The results of this process are shown in table 5. In addition, normalized root mean square error coefficients were estimated by standardizing each model RMSE by the standard deviation of Y in order to generate comparable between model estimates. This is necessary because the different models do not share the same dependent variable, meaning that the MAE and RMSE values are not directly comparable. Finally, in order to validate the results and estimate overall robustness against additional simulated



MAE from 10 fold crossvalidation on alternate learning rate values (.3 - .005).

Figure 3: Results of Learning Rate Sweeps

out of sample data a round of 10-fold cross validation was run on each model and cross validation MAE, RMSE, and NRMSE coefficients were calculated from the averaged results.

As an additional step the predicted verses actual values for each indicator in the testing data were computed and graphed in Figure 3. As expected, each predicted measure from the fully trained XGBoost model is linearly correlated with the true value, and all are correlated above 93%. Indicating that the full model does an excellent job of predicting true values of the withheld testing data. The BCI relationship is especially strong, with few outliers and a 95% correlation, while the Vdem predictions are notably nosier, with slightly more outliers and variance. However, this is expected since Vdem has the highest NRMSE in the XGboost results as shown below. Further, the predicted values are still correlated at 93% with the actual observed levels of Vdem for the training data.

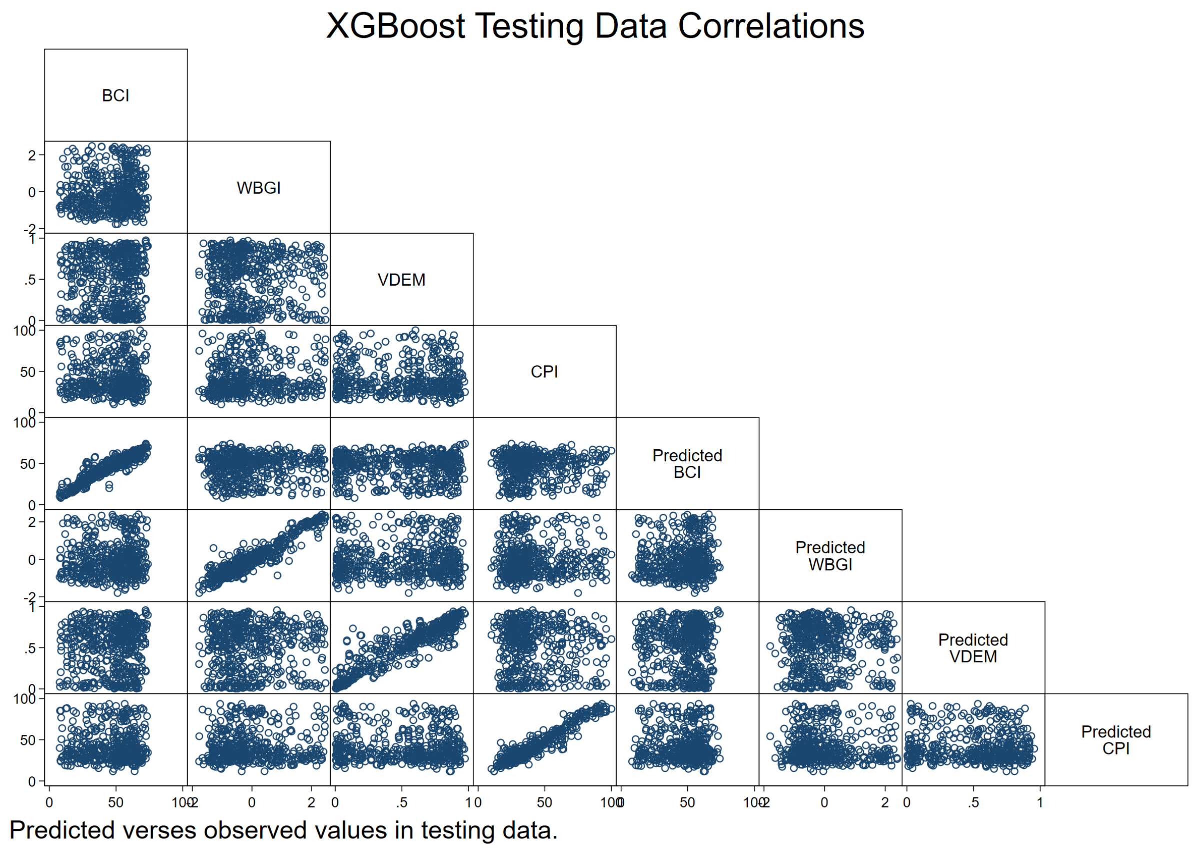


Figure 4: Correlations between predicted and actual training data

An analysis of the results shown in table 5 indicates that overall these predictors are successful at estimating levels of expected corruption out of sample. Based on MAE, this modeling approach estimates expected CPI in the test data to within 4.4 points of the observed value on a 100 point scale, WBGI to within .24 points on a 5 point scale, BCI to within 3.4 points on a 100 point scale, and Vdem to within .08 points on a 0-1 scale. Although, these estimates rise to 8 points for CPI, .3 points for WBGI, 5.7 points for BCI, and .09 points for VDEM once K-fold cross validation is performed to evaluate out of sample accuracy, indicating that the model is slightly over fitted on the training set, despite hyper parameter tuning. However, even when this increased margin is considered the estimates are still accurate, and sufficient to demonstrate that the underlying predictors are highly correlated with corruption levels.

Further, between model results are strongly clustered, indicating that the package of predic-

Table 5: XGBoost Corruption Prediction Models: DV Corruption

	CPI	WBGI	BCI	VDEM
Range	(0 - 100)	(-2.5 - 2.5)	(0 - 100)	(0-1)
Model Fit				
MAE	4.369	0.242	3.396	0.075
RMSE	6.172	0.346	5.967	0.114
Model Cross Validation				
MAE CV	6.553	0.273	5.650	0.093
RMSE CV	8.751	0.373	7.610	0.133
Model Comparison				
NRMSE	0.304	0.342	0.356	0.367
NRMSE CV	0.431	0.368	0.455	0.427
Observations	2464	2745	2908	2718
Countries	176	178	178	164

NRMSE normalized by standard deviation
CV indicates statistic from K-fold cross validation.

tor variables is generalizable to alternative corruption definitions as shown by the normalized root mean squared error (NRMSE) in model comparison. When normalized using the standard deviation of the dependent variable NRMSE represents the ratio of variation which is not explained by the predictors as a ratio of overall variation in the DV. The smaller this value is the better the predictions, a value of zero would indicate that the regression perfectly predicts variation in the dependent variable. While a value of one would indicate that variance in the predictions exceeds variation in the variable of interest. NRMSE values are clustered between .3 and .37, increasing to .37 to .46 once cross validation is performed, suggesting that even in a worse case scenario this vector of economic and public health indicators is accounting for at least 54% of the variation in observed levels of corruption. Put more concretely, based on the MAE from cross validation against the range of each variable. The candidate predictors outlined here could be expected to produce an average estimate within 5-9% of the true value on out of sample data using this model. Collectively, this strongly supports the idea that variations in public health spending, mortality rates, especially amongst the young, and vaccination rates are predictive of changes in unobserved levels of corruption and good candidates for an outcome based measure.

3.6 The Prediction Process

In order to further explore the predictive capability of the XGBoost model described above and better explain the results presented in table 5 this section will examine some examples of predicted verses actual levels of BCI. Although this process was run for each of the four major corruption measures discussed above, and the overall results are shown in table 5. The BCI predictions are a good point of focus because BIC is the most methodologically sophisticated, and robust of the major cross-national measures of corruption and the best single point of comparison. As discussed earlier BCI uses a super sampling approach, meaning that it incorporates the largest number of sources of the available corruption measures. Further, it enters sources directly, without re-scaling, and implements better internal bias correction and outlier suppression than other perception based measures (Standaert, 2015).

Consequently, on theoretical grounds BCI is the single best measure to conduct a direct comparison against, and a good measure for illustrating the underlying methodology and predictive capability of the XGBoost model implemented here. In this data BCI has the most available country-year observations, at 2,908, and in the testing data the mean absolute error for BCI was 3.396, meaning that on average the predicted value of BCI was plus or minus 3.5 points from the true value for a given country-year. Suggesting that in addition to the theoretical grounds for using BCI as a bench mark for comparison, the BCI predictive model was also quite accurate. The following figures demonstrate the true observed value of BCI in red for a number of exemplar countries, while the blue lines indicate the predicted values generated by re-estimating corruption for each country using the fully trained XGBoost model[7].

[7]The objective of these graphs is to illustrate the comparison between observed and predicted values for BCI from the full XGBoost model, in order to illustrate the predictive accuracy of the model. Consequently, because we are interested in the variation within country predictions, Rather than between countries these graphs are not standardized.

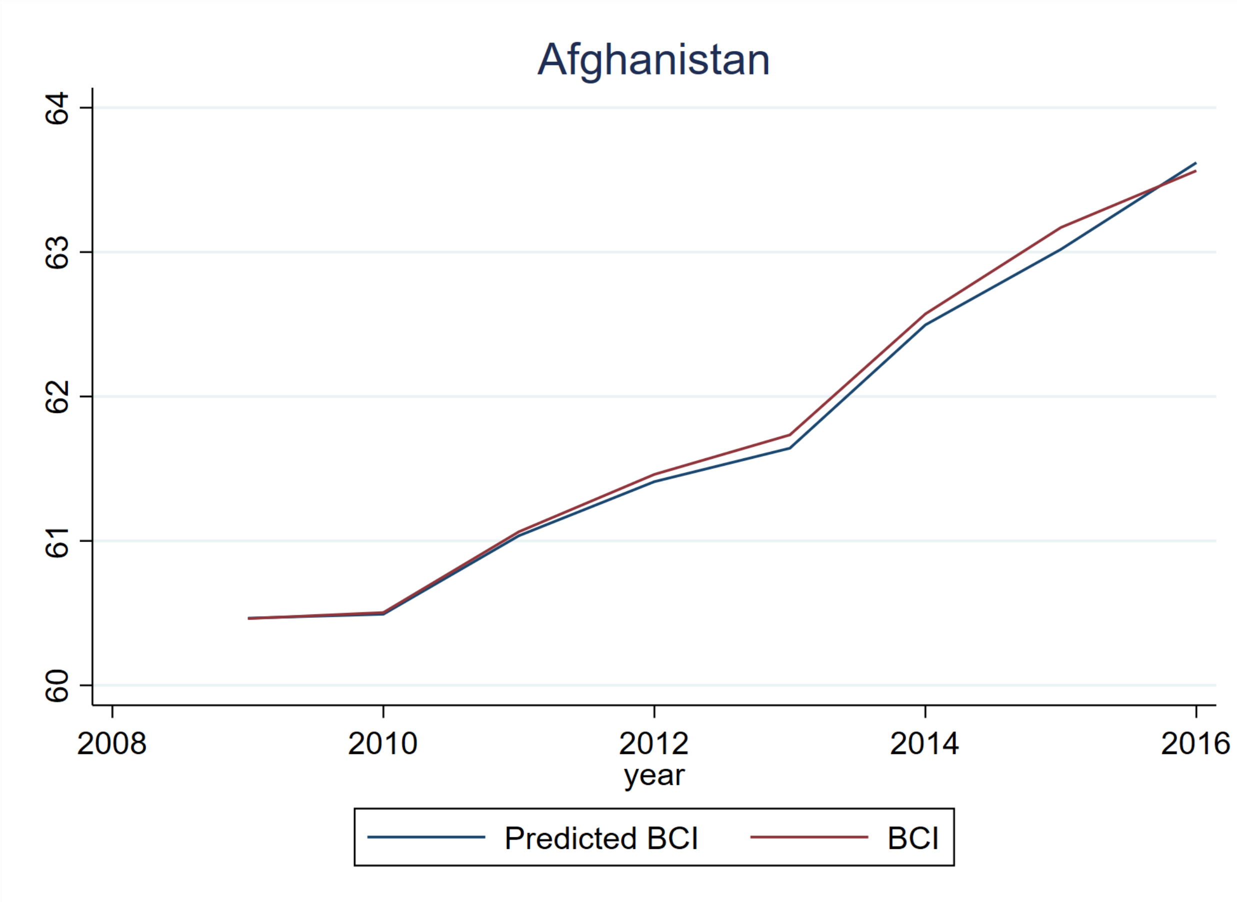


Figure 5: Correlations between predicted and actual training data

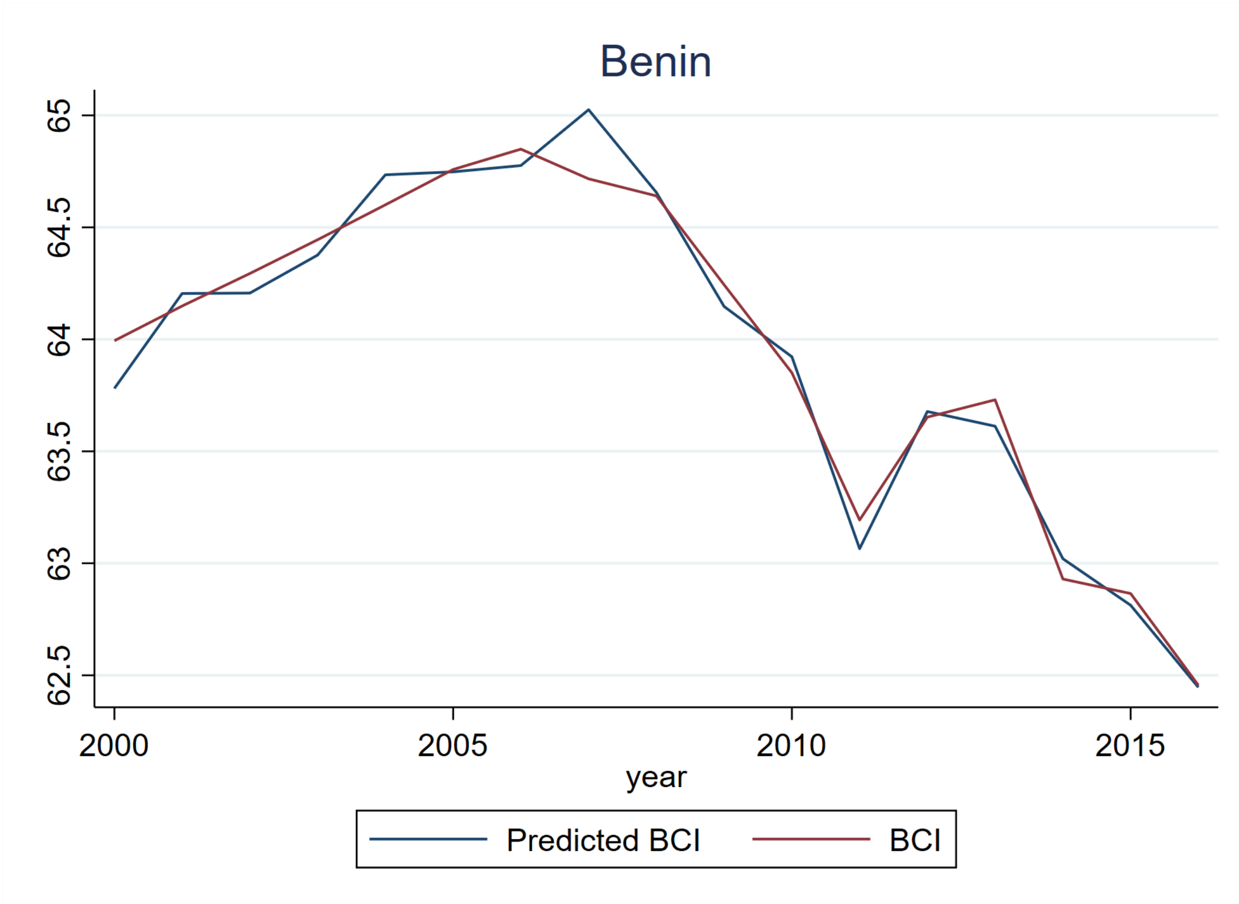


Figure 6: Correlations between predicted and actual training data

For example, according to BCI data, Afghanistan experienced a gradual increase in corruption between 2009 and 2016. The XGBoost estimate does an excellent job of capturing this, as shown in figure 4, with the blue predicted line closely mirroring the red real observed values. Specifically, the model predicted the true value of BCI within 2 points, on a 100 point scale between 2009 and 2016. Although Afghanistan shows a steady linear increase in corruption over time, the model does equally well with a country with more variance. For example, Benin experienced increasing corruption for most of the first decade of the century, before decreasing sharply, with a brief reversal between 2011 and 2013, based on observed values of BCI. As shown in figure 5 the model accurately tracks both the direction, and magnitude of these shifts, and correctly estimates the inflection points in 2007, and 2013 respectively.

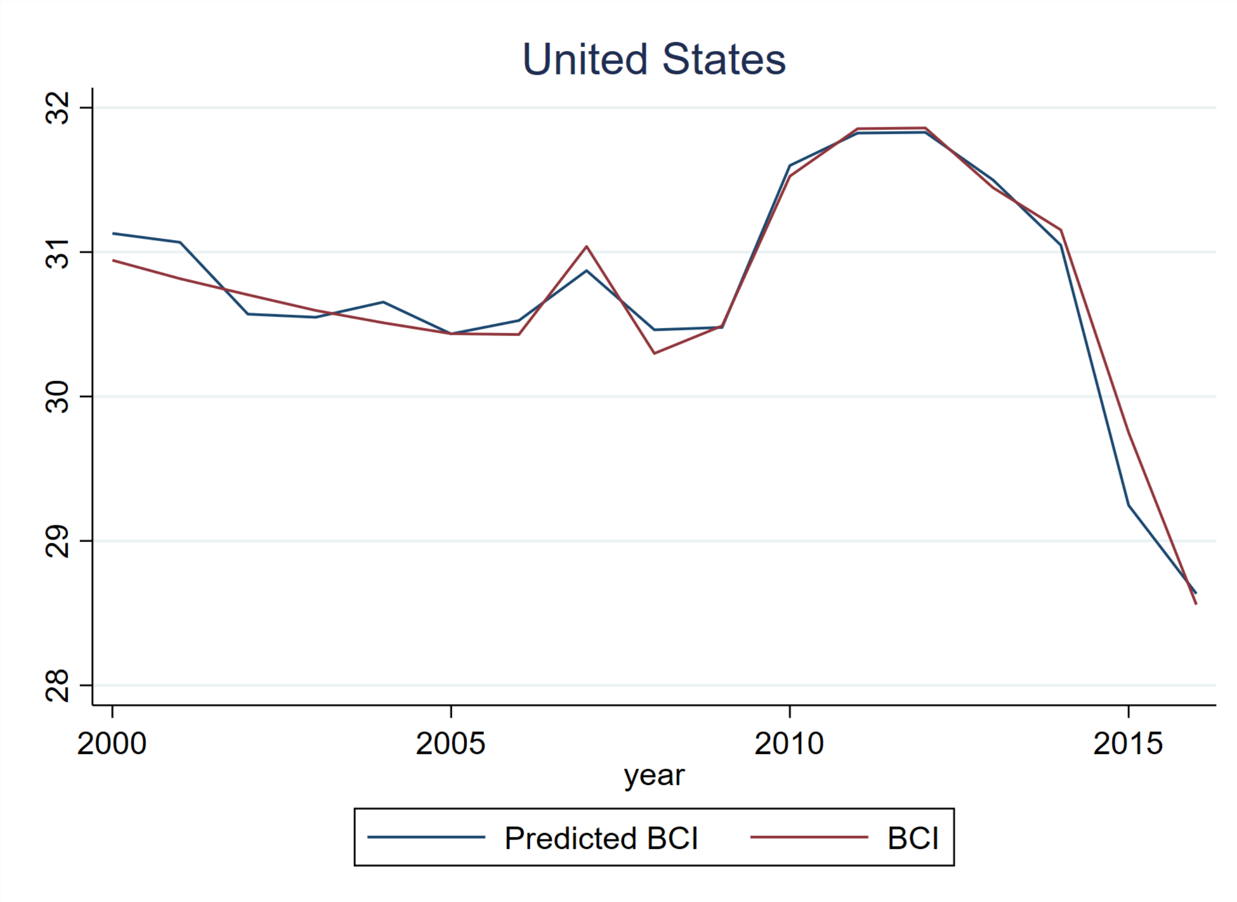


Figure 7: Correlations between predicted and actual training data

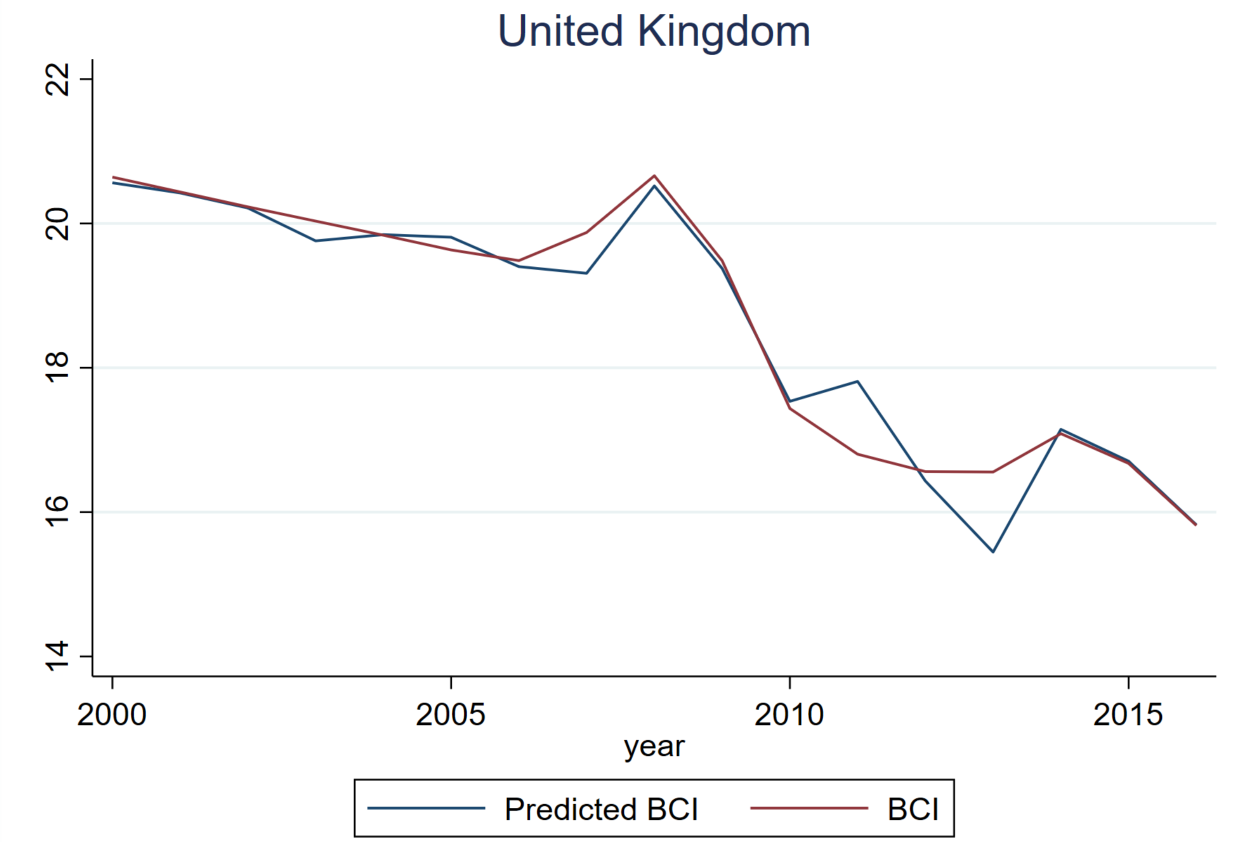


Figure 8: Correlations between predicted and actual training data

Similarly, the model works well for both developing and developed economies. For example, the BCI prediction for the United States between 2000, and 2016 is within a half point of the true estimate and tracks closely with changes over time. While, the estimate for the United Kingdom, lies within 1 point of the true estimate, and tracks changes BCI, although it slightly over predicts in 2011, and under predicts in 2013.

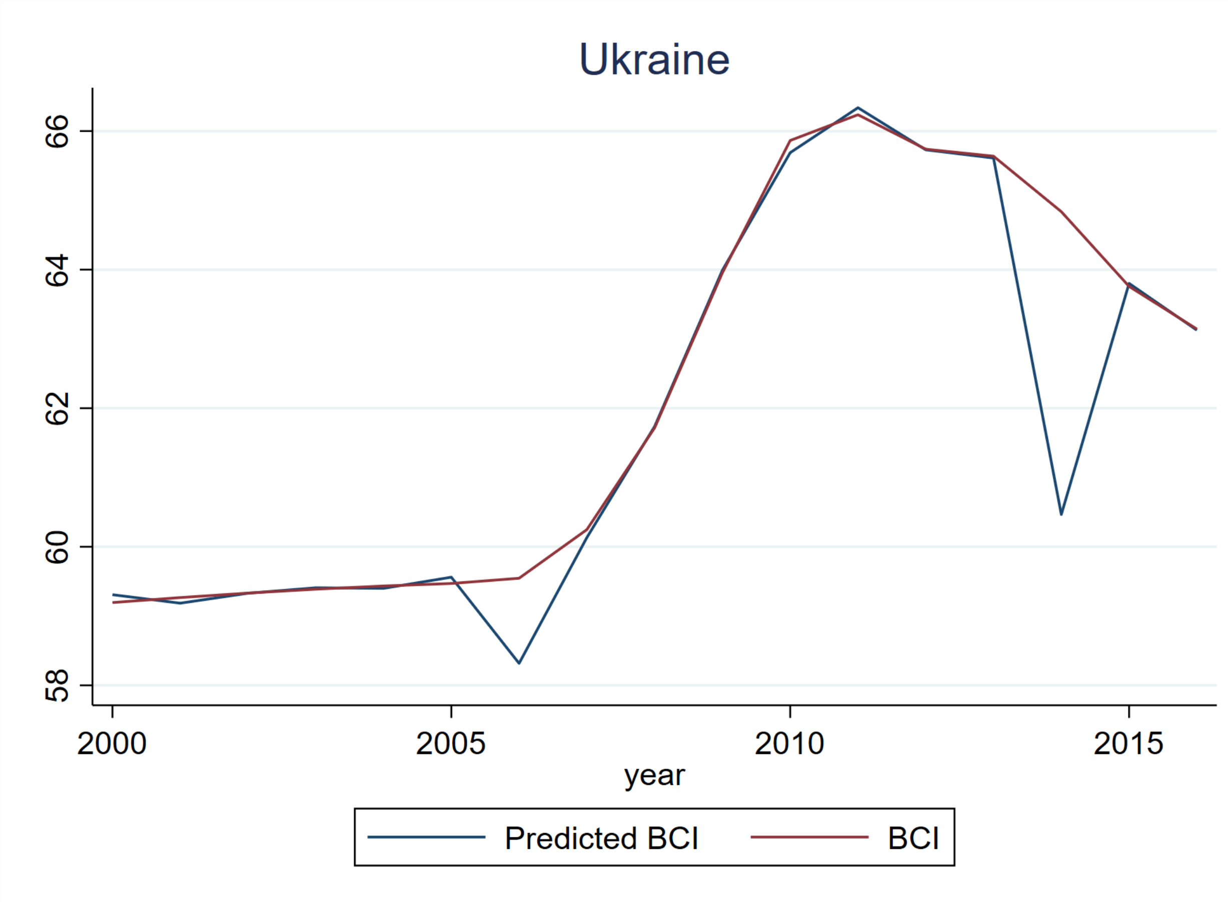


Figure 9: Correlations between predicted and actual training data

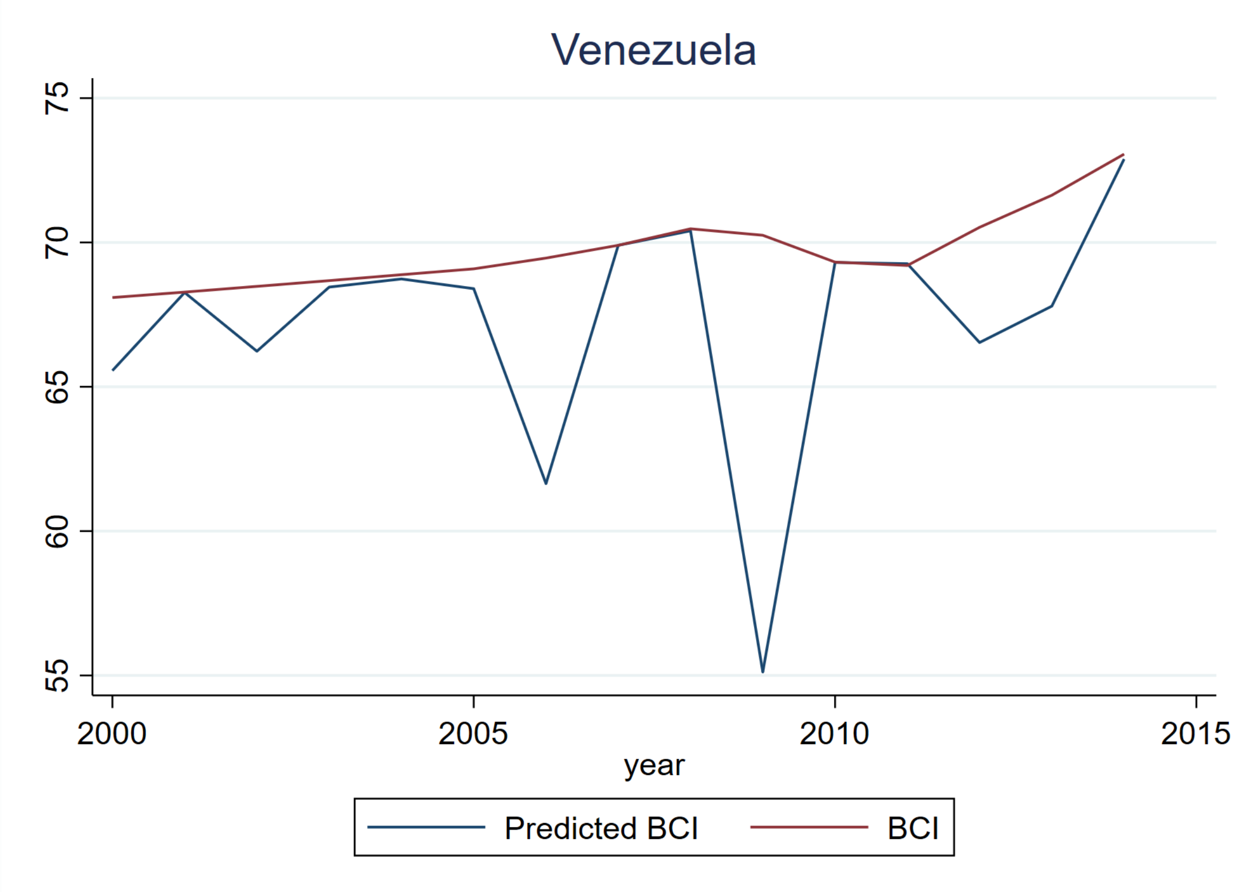


Figure 10: Correlations between predicted and actual training data

However, there are limitations to the predictive capability of this model. Specifically, periods of significant economic upheaval tend to result in under prediction of BCI. For example, as shown in figure 8, the estimated predictions for BCI were lower for Ukraine in 2006, when Ukraine had a significant political upheaval, and 2014, which corresponds to the outbreak of the Donbass war. Similarly, there is a high degree of variance, in the estimations for Venezuela, corresponding to 2006, 2009, and 2012-2013, during which corruption estimates are markedly low. Especially for 2009, which likely reflects the results of the 2008-2009 global economic crisis. These examples suggest that while accurate in the aggregate this estimation approach is likely to under predict during a substantial shock. Despite this, the overall results are very close to observed values. In most cases where deviations between actual and predicted BCI occur they are substantively small, even when crisis shocks impact the model the estimation tends to return to accuracy within one year. Although the objective of this section is not to estimate BCI directly, but rather to validate that the selected vector of predictors is able to reasonably estimate existing levels of corruption as a justification for extending this approach to estimating a new index directly in the next section.

4 Outcome Index of Corruption

The objective of this project is to estimate an outcome based measure of corruption based on the shadow of corruption from related objective socio-economic data. The proceeding sections have laid out a theoretical basis for this endeavor, considered and tested potential indicators, and validated those indicators as predictors of existing cross-national measures of political corruption. This section builds on that work to produce an outcome index of corruption.

4.1 Index Construction

Methodologically, the Outcome Index of corruption largely adopts the index construction technique implemented by CPI involving base year selection, Z-score predictor standardization, and a simple average across the final predictors to establish a country-year score (Transparency International, 2020*b*). With the additional inclusion of feature importance weights derived from machine learning to establish the relative importance of each outcome predictor in the final estimate. The first step in the process is to select a base year, this makes the estimates comparable and allows Z-score standardization of the data. Mirroring CPI, 2000 was selected as the base year and a sample mean and standard deviation were calculated for each predictor using the 2000 cross-section. Following the data management approach used by CPI if data was missing for 2000, it was estimated by interpolation. Note that this was only done in order to calculate the base year means and standard deviations for predictors, no interpolated data was used in the final calculations.

Second, feature importance weights were calculated using the frequency weight function in XGBoost on the final fully tuned model with the training dataset. This was done by collecting the feature importance estimates, representing how often each variable was used in splitting trees during boosting, for each variable of interest under each alternative specification of corruption. These weights were then converted to relative feature importance weights by first calculating the ratio of frequency for each feature over the total number of splits for that model. For example, the share of splits using infant mortality divided by the total number of splits on the CPI model. Finally, in order to estimate generalizable feature weights the relative feature scores for each variable were averaged across all models in order to establish a multiplicative weight for the final index element. For example, taking the average of relative feature importance for infant mortal across the CPI, BCI, WBGI, and VDEM models to establish a final average weight of 0.133 as shown in table 6 below. These weights were then used as multiplicative weighting factors to adjust the relative importance to assign to each variable in a given country-year during the final index calculation, as shown in equation 2.

Table 6: Average Relative Feature Weights from Trained XGBoost Model

Indicator	Feature Importance
Health Spending	0.278844
GNIPC	0.208698
Adult Mortality	0.141993
Infant Mortality	0.133270
DPT Immunization	0.079144
Measles Immunization	0.072827
Under 5 Mortality	0.085223

Next Z-scores were calculated for each predictor as shown in equation 2 and then re-scaled to fit the desired 0-100 range. To do this the base year mean was subtracted from the country-year observation for each predictor and then multiplied by the average feature importance weight in order to adjust the indicator by it's relative predictive importance. The weight adjusted country year indicator was then divided by the standard deviation of the base year in order to create Z-scores centered around zero, with a standard deviation of approximately one. next, following the example set by CPI, the resulting Z-scores for each indicator were re-scaled to have a mean of approximately 45, with a standard deviation of 20. This process was repeated with each of the predictors in the model to produce weighted re-scaled country-year indicators for adult, infant, and under five mortality, both measles and DPT immunization, and GNI per capita.

$$Z_{it} = \frac{(Indicator_{it} - \mu Indicator_{t2000}) * \mu FeatureWeight}{\sigma Indicator_{t2000}} * 20 + 45$$

(2)

Finally, adjusted country year indicators were aggregated using a simple average formula to incorporate the results from all seven predictors for each country-year to construct the final country-year outcome indicator of corruption as shown in equation 3. Final scores are only calculated for country-years for which all seven indicators are available. Lastly, the result was inversed, so that higher scores correspond to higher levels of corruption like the BCI, and VDEM corruption scales.

$$Corruption_{it} = (\frac{1}{n}\sum_{i=1}^{n} Indicator_{it})$$

(3)

4.2 Index Comparisons

The resulting indicator runs in a theoretical range between 0, no corruption, and 100 the highest level of corruption possible with a mean value of 54, and a range of 23-80 as shown in table 7. The resulting indicators is correlated with existing corruption measures at levels between 0.49, and 0.58. Additionally, turning to an analysis of the comparative mean and range the outcome based indicator is largely comparable in the aggregate to exiting indicators of corruption. With a similar observed range and mean to other major indicators and a comparable level of observations and coverage. Further, it maintains the expected positive correlation with other corruption measures.

Table 7: Outcome Index Comparison to Other Corruption Indicators 2000-2016.

Indicator	Obs	N	Mean	Std. Dev.	Min	Max	Index Correlation
Outcome Index	2908	177	54.224	6.360	23.009	79.524	1
CPI	2464	175	57.485	21.146	0	96	0.580
Vdem	2718	166	.516	0.309	0.005	0.974	0.491
BCI	2908	177	47.115	16.075	6.450	74.889	0.543
WBGI	2745	177	-0.034	0.994	-1.805	2.469	0.578

CPI reverse coded to match direction of other indicators.
WBGI runs between -2.5 and 2.5.

Continuing this comparative analysis with existing corruption measures Figure 11 shows a standardized aggregate time series of the annual mean value between the Outcome Index and other comparable corruption indicators[8]. One immediate observation is that CPI has a notably higher variance and range than other measures. As discussed in the literature review CPI generally estimates higher relative levels of corruption compared to VDEM, BCI, or WBGI, and has the highest degree of variance of the four major perception indices. In comparison, as shown in figure 11 the Outcome Index ranks second, occupying a position slightly higher than VDEM and WBGI, and roughly half way between CPI and VDEM/WBGI since 2012. However, in contrast to CPI it shows less extreme variance and a flatter trend line, with stable performance in the aggregate over time. From this perspective the Outcome Index indicates higher relative corruption compared to VDEM, BCI, or WBGI. However, like these alternative measures it exhibits less volatility over time, especially when compared to CPI. In fact, the Outcome Index has the lowest standard de-

[8] CPI was reverse coded so that higher values correspond to increasing corruption. While WBGI was re-centered to run in a positive range. Values were standardized for comparison as a proportion of the theoretical range for each indicator

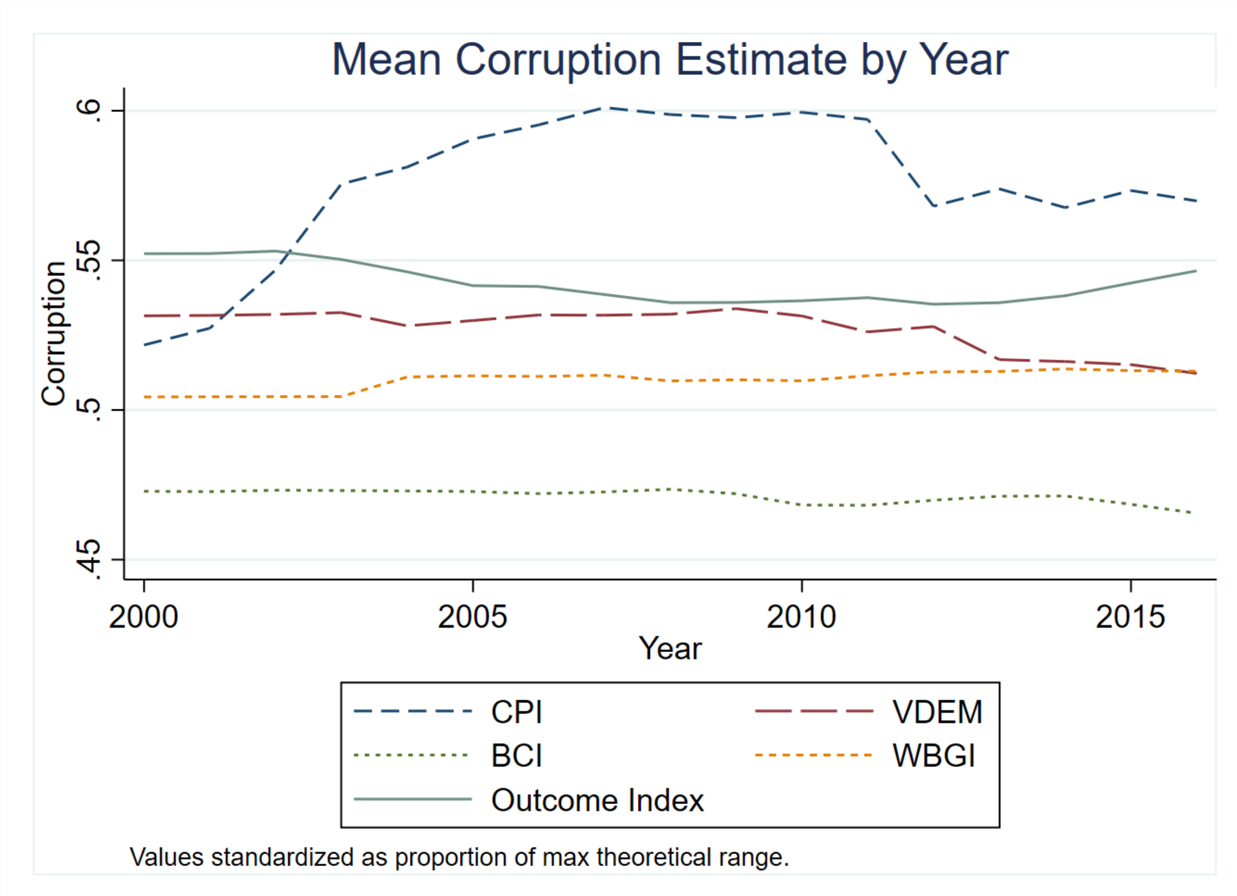


Figure 11: Comparative Standardized trend over time.

viation relative to either the theoretical range, or the maximum observed range, of any indicator in Table 7. Indicating that on average Outcome Index values are closer to the mean than other indicators, and notably substantially lower than CPI.

This relative stability compared to CPI is a strength because corruption measures should theoretically be expected to be relatively stable in the aggregate over time. The global population of countries may gradually become more or less corrupt, however, sudden global population level shifts are theoretically unlikely. Consequently, the relatively stable aggregate performance of the outcome measure over time is both expected, and strong evidence that it is robust to external global shocks. Taken collectively, this evaluation suggests that the Outcome Index has a similar aggregate profile to existing corruption measures. Demonstrating a comparable range and average level existing measures. Further, it is stable over time and has a low standard deviation. These results suggest that using an index composed of commonly available socio-economic predictors, such as age based mortality rates, vaccination rates, health spending, and gross national income per-capita has similar properties to existing corruption measures and could be a plausible candidate to proxy levels of corruption.

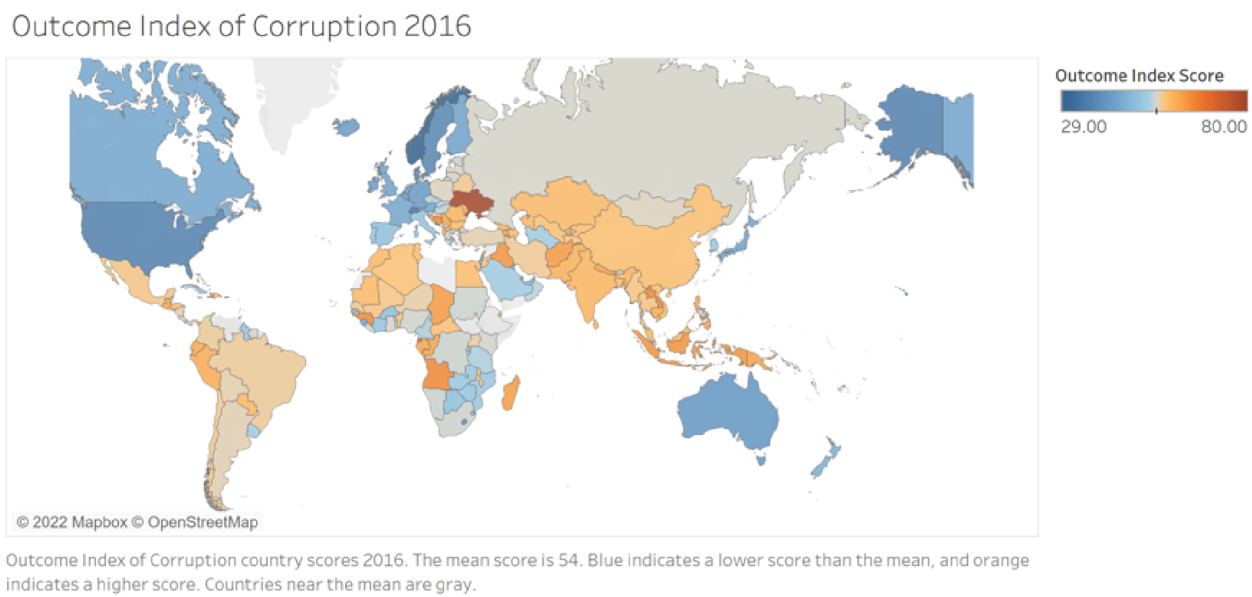


Figure 12: Outcome Index Scores 2016

Figure 12 provides a graphical overview of the 2016 Outcome Index showing comparative levels of estimated corruption [9] Blue indicates lower corruption, and orange indicates higher corruption relative to the mean, countries close to the mean are gray. Most of Europe, the United States and Canada, parts of Asia and a notable cluster in south eastern Africa have lower corruption. While eastern Europe, central Asia and parts of Africa have relatively higher corruption levels. In 2016 the most corrupt country is assessed to be Ukraine, followed by countries like Laos, Angola, Iraq and Indonesia. Figure 13 presents a similar map showing the average Outcome Index score over time for each country between 2000 and 2016. The overall trend is largely similar, with

[9] A principle issue with comparative maps of corruption is picking a reference point. Since there is no empirical justification for a 'good' level of comparative corruption this color gradient is centered on the mean and color and intensity measure divergence from the mean.

Syria, Iraq, Ukraine and portions of central and south-east Asia standing out as more corrupt than the data set average over time.

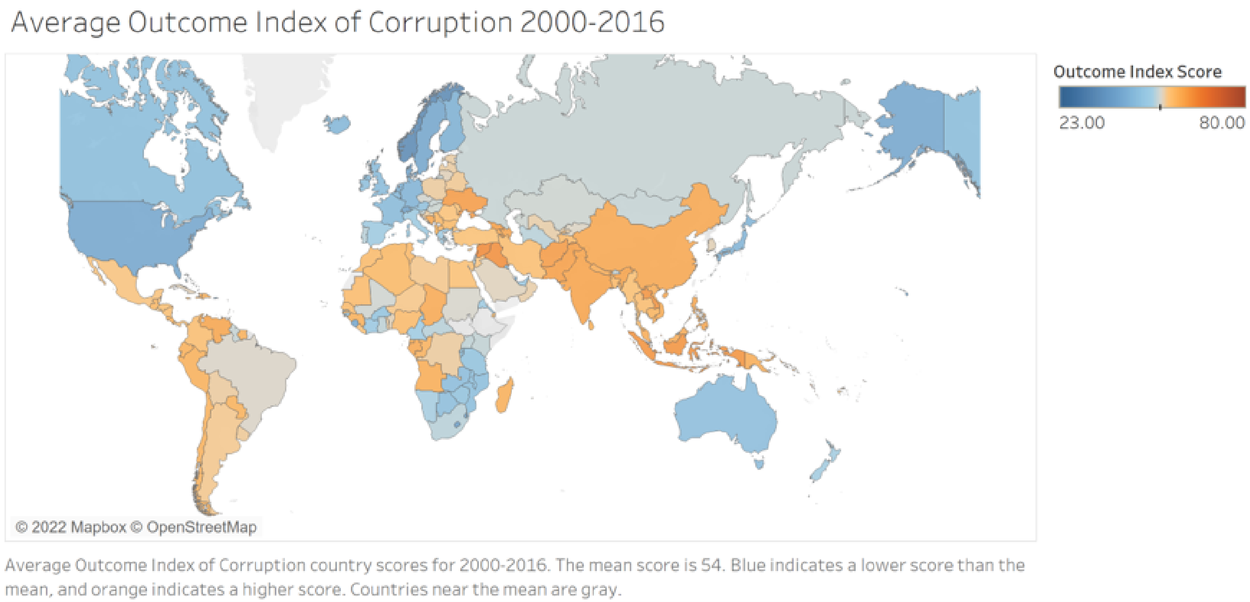


Figure 13: Average Outcome Index Scores 2000-2016

Table 8 extends this exploratory results analysis with a freedom from corruption style ranking comparing the Outcome Index ranking for 2016, to alternative BCI, CPI, VDEM, and WBGI ranks for 157 countries. In addition, it includes a calculated difference between the outcome rank for a given country year and the average alternative rank[10]. The point of this comparison is to directly compare a representative year and highlight similarities and differences between the Outcome Indicator and other corruption measures. The alternative ranks are presented in addition to a field average rather than comparing to a single 'best' indicator largely because there is no good theoretical reason for arguing that one of them is inherently more correct, especially given their high aggregate correlations. Overall, The mean difference between the Outcome Index and the mean of alternative measures is -0.347, indicating that on average the index scores countries slightly lower than the collective mean across alternative measures.

[10]Calculated as the difference between the observation index rank and the average rank for that country across alternative measures of corruption.

Table 8: Comparison of Outcome Index and Alternative Corruption Rankings 2016.

Country	Outcome Index	Average Difference	BCI	CPI	VDEM	WBGI
Norway	1	3.5	7	6	1	4
Switzerland	2	4.5	6	5	6	9
United States	3	18.75	26	17	24	20
Luxembourg	4	6.5	5	10	21	6
Denmark	5	-2.75	4	1	1	3
Sweden	6	0	10	4	5	5
Qatar	7	26.75	13	29	63	30
Lesotho	8	62.5	76	74	73	59
Australia	9	4.75	15	13	13	14
Germany	10	3.25	20	10	10	13
Iceland	11	.25	14	14	7	10
Belgium	12	2.75	18	15	11	15
Ireland	13	1.75	9	18	16	16
Netherlands	14	-4.25	11	8	9	11
Japan	15	3.25	16	19	20	18
Finland	16	-11.25	2	3	12	2
United Kingdom	17	-4	12	10	18	12
Austria	18	.5	24	16	17	17
Sierra Leone	19	107.25	139	112	130	124
Canada	20	-9.5	17	9	8	8
France	21	-.5	22	22	19	19
Singapore	22	-17.5	3	7	1	7
New Zealand	23	-21.25	1	1	4	1
United Arab Emirates	24	-3.25	8	23	27	25
Kuwait	25	44.25	59	66	80	72

Israel	26	1.5	29	26	31	24
Burkina Faso	27	49.5	121	63	58	64
Spain	28	11.25	58	35	23	41
Botswana	29	5	37	32	39	28
Italy	30	24	64	52	45	55
Cuba	31	19.5	42	52	57	51
Cote d'Ivoire	32	58.5	86	98	81	97
Zimbabwe	33	98.75	113	143	125	146
Guyana	34	55.75	118	98	63	80
Korea, South	35	10.5	49	44	47	42
Mozambique	36	80.75	130	131	77	129
Burundi	37	111	145	148	154	145
Zambia	38	38.75	88	78	55	86
Malta	39	1.5	39	39	48	36
Portugal	40	-9	36	27	32	29
Saudi Arabia	41	6.75	25	54	65	47
Turkmenistan	42	108.75	154	143	152	154
Uruguay	43	-22	27	20	15	22
Gambia	44	55.25	52	134	90	121
Guinea-Bissau	45	108.75	157	155	146	157
Cameroon	46	97.75	143	134	156	142
Trinidad and Tobago	47	27.5	110	91	28	69
Tanzania	48	49.25	116	105	75	93
Slovakia	49	16.25	114	46	53	48
Czech Republic	50	-2	70	39	43	40
Slovenia	51	-13	44	29	46	33
Togo	52	64.75	134	105	121	107

Oman	53	-10	31	56	41	44
Comoros	54	64.25	124	142	102	105
Hungary	55	7.25	82	49	65	53
Bahrain	56	5.75	23	61	106	57
Bhutan	57	-29.75	32	25	25	27
Greece	58	1.25	74	60	42	61
Namibia	59	-12.5	62	45	34	45
Maldives	60	33	83	86	97	106
Estonia	61	-42	21	21	13	21
Congo, Democratic Rep.	62	86	144	145	153	150
Sudan	63	75.75	108	157	135	155
South Africa	64	-2	78	56	62	52
Lithuania	65	-27	47	33	35	37
Liberia	66	33	75	81	127	113
Nigeria	67	69.25	149	125	133	138
Kenya	68	46.25	109	134	83	131
Russia	69	46.75	104	120	114	125
Ghana	70	10.75	102	61	95	65
Latvia	71	-29	55	37	33	43
Suriname	72	-2.5	95	56	51	76
Rwanda	73	-34.5	19	42	54	39
Barbados	74	-45.25	34	29	29	23
Mongolia	75	11.5	93	78	85	90
Poland	76	-45.5	38	27	22	35
Costa Rica	77	-35.5	57	35	40	34
Nicaragua	78	48	117	134	125	128
Panama	79	7.75	92	78	85	92

Argentina	80	12.75	129	86	82	74
Cyprus	81	-44	40	39	37	32
Turkey	82	-10	45	66	109	68
Bolivia	83	29.25	140	102	93	114
Niger	84	0	56	91	89	100
Uganda	85	50	142	140	119	139
Sao Tome and Principe	86	-31.75	53	54	50	60
Brazil	87	5.75	127	70	91	83
Colombia	88	1.75	120	81	79	79
Senegal	89	-28.75	80	56	49	56
Iran	90	20.75	89	120	123	111
Chile	91	-65	30	23	25	26
Belarus	92	-31.25	51	70	51	71
Mauritius	93	-41.75	41	42	73	49
Croatia	94	-35.25	79	47	59	50
Mexico	95	18.5	123	112	107	112
El Salvador	96	-2	107	86	87	96
Mali	97	15.25	125	105	116	103
Thailand	98	-3.75	81	91	121	84
Myanmar	99	11.5	147	125	71	99
Malawi	100	15.25	122	109	111	119
Jamaica	101	-40.5	72	74	30	66
Tajikistan	102	18.5	50	140	151	141
Benin	103	-9.75	132	86	60	95
Algeria	104	-2.75	99	98	100	108
Central African Rep.	105	37	152	148	120	148
Uzbekistan	106	22.5	91	145	135	143

Jordan	107	-57	33	49	72	46
Kyrgyzstan	108	27.25	128	125	148	140
Tunisia	109	-50.5	61	66	44	63
China	110	-44.75	54	70	67	70
Dominican Republic	111	13	141	109	128	118
Vietnam	112	-23	90	102	76	88
Azerbaijan	113	2.5	69	112	155	126
Egypt	114	-15	63	98	134	101
Bulgaria	115	-36.5	94	66	87	67
Honduras	116	1	111	112	141	104
Bangladesh	117	15.25	131	134	137	127
Kazakhstan	118	-3.25	77	120	139	123
Morocco	119	-48.75	60	81	78	62
Mauritania	120	16.25	150	131	149	115
Cape Verde	121	-84	46	33	38	31
Paraguay	122	2.25	151	112	124	110
Djibouti	123	-23.75	73	112	110	102
Malaysia	124	-69	35	47	84	54
India	125	-60.25	48	70	68	73
Romania	126	-63	84	49	61	58
Georgia	127	-92.25	28	37	36	38
Moldova	128	-.25	133	112	131	135
Sri Lanka	129	-53.75	68	86	70	77
Albania	130	-39	87	74	116	87
Pakistan	131	-19.25	100	105	112	130
Solomon Islands	132	-56.75	65	63	92	81
Peru	133	-46.25	105	91	69	82

Cambodia	134	3	112	145	144	147
Armenia	135	-32.75	71	102	138	98
Ecuador	136	-32.25	103	109	94	109
North Macedonia	137	-59	43	81	113	75
Gabon	138	-32.75	98	91	115	117
Serbia	139	-56.25	85	63	105	78
Philippines	140	-43.5	106	91	98	91
Guatemala	141	-25	115	125	108	116
Timor-Leste	142	-56.5	101	91	56	94
Haiti	143	4.25	146	148	146	149
Congo	144	-1.75	148	148	129	144
Papua New Guinea	145	-22.25	136	125	96	134
Nepal	146	-36.5	97	120	101	120
Madagascar	147	-10.25	137	134	144	132
Lebanon	148	-25	126	125	104	137
Afghanistan	149	1	138	156	150	156
Chad	150	3.5	156	148	157	153
Indonesia	151	-68	67	81	99	85
Iraq	152	-13.25	119	154	131	151
Guinea	153	-14	153	131	139	133
Bosnia and Herzegovina	154	-63.75	96	74	102	89
Angola	155	-4.25	155	153	143	152
Laos	156	-42	66	112	142	136
Ukraine	157	-33.25	135	120	118	122

An example of this is Denmark, which is assigned a ranking of 5th best in the world, approximately 2.75 ranks lower than the average 2016 ranking for the country between other indicators. Similarly, countries like Norway, Switzerland, Luxembourg, Denmark, Sweden, and Germany to name a few are often found near the tops of these indices, and are unsurprisingly also within the top 10 for the Outcome Index. Notably, the United States scores substantially better, about 18 ranks higher than the average, which is likely a result of the impact of healthcare spending. On the opposite side of the scale the Outcome Index ranks Afghanistan and Haiti within 1 rank and 4 ranks respectively of the average. Considering that both countries are benchmark cases for high corruption the close coherence between the Outcome Index rank and alternative measures strongly suggests that the approach works in high corruption cases. Overall, in 2016 half the country scores fall bellow the average prediction, indicating that they predict lower corruption levels than the average of other indicators, while the other half predict higher corruption for a given country [11]. 49 countries are within plus or minus 10 ranks of their average position, and 74 countries are within 20 ranks of the overall average.

However, this comparative ranking also highlights several other notable cases. For example, Lesotho comes in at number 8, a full 62 ranks higher than the average, and somewhat improbably high for a country which experienced a failed coup in 2014, and has numerous senior officials embroiled in political scandals tied to their conduct in office (House, 2021). Qatar, is also surprisingly high, at number 7, although there is wide disagreement on corruption in Qatar, with BCI placing it as high as rank 13, and Vdem as low as 63rd in the world. Similarly, Sierra Leone is ranked at number 19, one higher than Canada, which is highly improbable and indicates that there are some countries that this index approach does not perfectly capture.

However, it is also important to note that the inverse does not occur frequently. There are few cases of countries which other indicators agree are very clean which the Outcome Index scores as highly corrupt. For example, there are no top 30 countries according to alternative rankings for

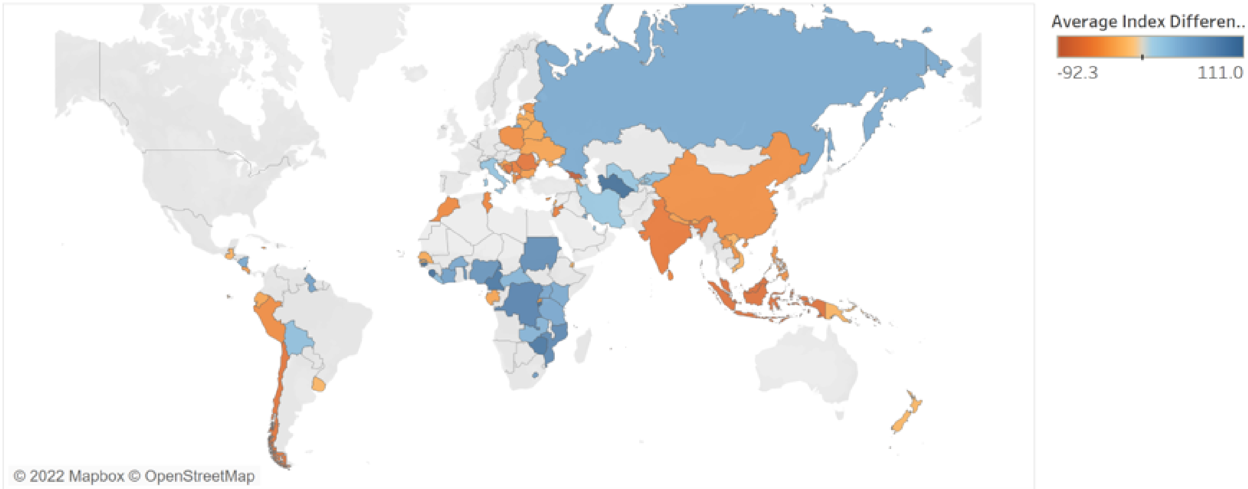
[11] 78 countries have an average difference less than 0, 79 have an average difference greater than or equal to 0.

2016 among the bottom ranks for the Outcome Index. The closest is Georgia which trends in the middle 30s, and is ranked 127 by the Outcome Index, followed by Chile at rank 91. In order to further visualize this Figure 14 maps countries which fall more than 20 ranks higher or lower than their average across alternative measures. [12] Blue countries are ranked better by the Outcome index than their average, while orange indicates countries that are ranked lower. Interestingly in 2016 eastern European countries are geographically associated with lower than average ranks, as well as China, India and Indonesia, while Russia, some Central Asian states and a group of central African states tend to be ranked higher than average. Conversely, North American, western European, north African and many south American countries tend to fall within 20 ranks of the average across alternative measures. While this is interesting it does not inherently say anything about the quality of the Outcome measure, and it is difficult to judge whether it reflects a feature of the outcome measure itself, or a structural result of the alternative comparative measures. Consequently, despite some notable outlier cases the average difference of -0.347 indicates that in the aggregate Outcome Index ranks are close to alternative measures and suggests that they are plausible.

This cross-section also highlights a separate issue in comparative cross-national corruption measures, specifically that there is often a wide range in estimates between alternative measures. It is generally easy to gain consensus on the least corrupt countries, but there is often more disagreement on the middle and lower ranking countries. Laos for example is assigned the second to last ranking 156, Outcome Index measure 42 points below the cross-indicator average. However, there is a 76 point spread in assessed positions for Laos, BCI places it as high rank position 66, while VDEM places it at 142, representing a 76 point spread between the two measures. Similarly Indonesia, a major developing economy has a 32 point spread in assessed corruption values. While Tajikistan is assessed to be in a relatively good 50th position by BCI, but placed a full 100 ranks lower by VDEM, and 90 ranks lower at 141 by WBGI, with the Outcome Index position squarely

[12]Plus or minus 20 ranks was selected as a diagnostic cutoff because the standard deviation of the average difference is 40.

Rank Difference -20 to 20 2016

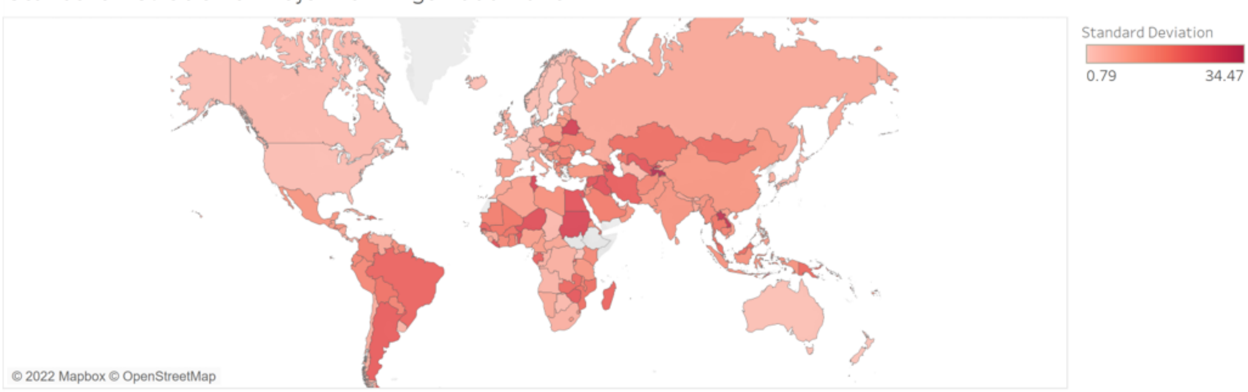


Comparison of difference between Outcome Index rank and average of alternative indicator ranks by country for 2016. Orange countries indicate lower than average ranks, blue indicates higher than average ranks. Countries within 20 ranks of their average position excluded.

Figure 14

in the middle of this range, at 102nd. Similarly, Costa Rica, often considered one of the 'cleanest' states in Central America ranks between 34-40 across CPI, VDEM, and WBGI, but is substantially lower at 57 according to CPI. This highlights that while in some cases it is clear that the Outcome Index has missed, for example by scoring Lesotho as the 8th best country for corruption in the world, in other cases it is more difficult to assess accuracy when alternative sources have wide variance in estimates.

Standard Deviation of Major Rankings 2000-2016



Comparison of average variance in alternative corruption rankings 2000-2016. Standard deviation of cumulative rankings from CPI, VDEM, WBGI, and BCI. Darker shaded countries have higher variation between alternative measures, and indicate less agreement.

Figure 15

Figure 15 shows a map of the most divergent country scores between CPI, VDEM, WBI, and BCI, with darker countries indicating higher mean standard deviations [13]. This calculation serves to estimate the variance between annual corruption rankings by country and year on the theoretical basis that larger deviations between measures indicate less agreement about real levels of corruption. This analysis immediately highlights two points. First, Tajikistan and Laos are the least consistent countries between corruption indices, with mean rankings as high as 148 according to VDEM, or as low as 73 according to BCI. Equating to the difference between being one of the most corrupt countries in the world, or a decidedly middle of the pack average country. In 2016 specifically this range was even wider with BCI suggesting that Tajikistan was the 50th best country. Followed by Belarus, which despite autocratic one-party rule since 1994 is ranked between the 50th and 90th positions on average, indicating substantial disagreement over estimated levels of corruption there between alternate measures.

Second, this map reveals a curious pattern in south America, where aside from Venezuela, Uruguay and Chile, every country has wide divergence in average rankings between measures. In fact, South America is visually unique in that it is the only region in which the majority of countries have wide divergence in their rankings. Despite the fact that there is no theoretical reason this should be the case. However, this does not imply that existing corruption measures are inherently wrong, but it does suggest that it is difficult to select a single ideal existing index. Further, it highlights the potential value of an outcome based index by reducing subjective codding and opaque judgements from the process of producing corruption measures.

[13] Calculated as the average standard deviation across alternative corruption rankings for 2000 - 2016. Calculated using the mean standard deviation of corruption ranks for a given year, using the rowsd, or row standard deviation command in Stata.

5 Applying the Outcome Index to Replication

The proceeding sections have outlined the theoretical and empirical basis for an outcome measure of corruption, calculated an index, and presented a set of comparative corruption rankings. This section will further demonstrate the outcome based approach by applying it to a re-analysis and replication of published cross-national corruption research. The objective of this section is to demonstrate that the proposed outcome index can function as an effective substitute for traditional perception based measures in cross-national research, and to evaluate the results of applying it to a well studied relationship in the literature on corruption.

For this purpose ”Corruption and Economic Growth. New Empirical Evidence” written by Klaus Gründler and Niklas Potrafke, and published by the *European Journal of Political Economy* was selected (Gründler and Potrafke, 2019). This paper was selected for three reasons. First, it is a cross-national analysis of a ’traditional’ topic in corruption literature utilizing CPI as a corruption measure. Consequently, replicating this research provides a comparison against a representative research topic in cross-national corruption literature, and allows a direct comparison against one of the most common corruption metrics used in the field. Second, corruption is used as an independent variable in this work with an argued causal relationship to economic growth as a national level socio-economic outcome. This allows for a direct comparison of the impact of substituting the Outcome Index. Third, the authors made both their data and replication files publicly available, ensuring that the resulting re-analysis is an exact replication using the originally published data and methods, but substituting the Outcome Index for CPI.

Gründler and Potrafke utilize the Corruption Perception Index from Transparency International, coupled with a range of World Bank indicators to evaluate the relationship between economic growth and corruption for a panel of 175 countries (Gründler and Potrafke, 2019). Their work seeks to test two long-standing political-economic theories of corruption. First, the grease the wheels’ hypothesis, which holds that corruption should actually increase economic growth by reducing inefficiency, and allowing business to get around bureaucracy. Second, the ’sand the wheels' hypothesis, which holds that corruption should be expected to reduce growth, by increasing inefficiency, decreasing innovation, and representing an additional marginal cost on the economy. Further, building on recent work (Acemoglu et al., 2019) they operationalize economic growth with the annual logged real GDP per capita, rather than absolute growth in GDP, or annual percentage change. This has the advantage of both adjusting for inflation, and provides a population adjusted growth measure. Finally, following the general trend in CPI based corruption research they reverse code CPI, so that higher values indicate more corruption.

Gründler and Potrafke test the relationship between growth and corruption in several different ways. In fact, they present no fewer than 7 empirical tables in the course of establishing their results. The core of their argument is a dynamic panel model using logged per capita GDP as the dependent variable and incorporating multiple auto-regressive lags of GDPPC, CPI as a measure of corruption and multiple control variables, including democracy, government effectiveness, and rule of law (Gründler and Potrafke, 2019). Separately, they also test for confounding effects with additional controls, such as globalization, trade openness, migration, and infrastructure, as well as endogeneity between corruption and economic growth using a system general method of moment model, an instrumental variable model, and a jack-knifed regional corruption approach. This replication study will focus on their main model evaluating the conditional effects of corruption on economic growth in the context of government effectiveness, rule of law, and democracy, presented in table 7 of the original article, and reproduced from their replication files below (Gründler and Potrafke, 2019).

5.1 Original Replication

As shown in table 9, Gründler and Potrafke model the interaction between corruption and growth via a dynamic panel model using logged GDPPC as the dependent variable, four auto-regressive lags of GDPPC to account for what they term economic 'GDP dynamics,' reverse coded CPI, as a measure of corruption, and three theoretical measures of government quality or effectiveness, drawn from the World Bank index of government effectiveness, Rule of Law, drawn from

Table 9: Replication of Gründler and Potrafke (2019) Table 7: DV log real per capita GDP.

	Effectiveness		Rule of Law		Democracy	
	(1)	(2)	(3)	(4)	(5)	(6)
Corruption	-0.0149***	-0.00720*	-0.0163***	-0.00799*	-0.0209**	-0.00994**
	(0.00520)	(0.00366)	(0.00575)	(0.00413)	(0.0104)	(0.00432)
Log (GDPPC)(t-1)		0.570***		0.561***		0.448***
		(0.177)		(0.178)		(0.170)
Log (GDPPC)(t-2)		0.0829*		0.0704		0.116
		(0.0492)		(0.0427)		(0.104)
Log (GDPPC)(t-3)		0.0818		0.0752		0.235
		(0.0871)		(0.0849)		(0.155)
Log (GDPPC)(t-4)		-0.00708		-0.0163		0.0664
		(0.0457)		(0.0463)		(0.0648)
Corruption x Gov. Effect.	0.00472**	0.00220*				
	(0.00183)	(0.00118)				
Corruption x Rule of Law			0.00542**	0.00257*		
			(0.00211)	(0.00137)		
Corruption x Democracy					0.0205*	0.00952*
					(0.0123)	(0.00495)
Government Effectiveness	0.310***	0.125*				
	(0.108)	(0.0639)				
Rule of Law			0.425***	0.184**		
			(0.149)	(0.0880)		
Democracy					0.779	0.407*
					(0.491)	(0.209)
Observations	1001	999	1001	999	839	837
Countries	172	172	172	172	173	173
R-Squared	0.228	0.587	0.283	0.596	0.241	0.590
RMSE	0.0613	0.0448	0.0591	0.0443	0.0578	0.0425
F Stat	14.42	94.25	20.44	85.72	11.54	79.09

Cluster Robust Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the World Bank Rule of Law index (Gründler and Potrafke, 2019). Additionally, they include a support vector machine based democracy estimate drawn from earlier work by Gründler, which is conceptually similar to the Polity index of democracy but arguably less biased (Gründler and Krieger, 2018). Finally, they test multiplicative interactive effects between each of these political institutional quality of governance indicators and corruption on economic growth. The theoretical idea here is that the impact of corruption should show up both directly, as a positive or negative effect of corruption on the current economy, as well as indirectly through an interactive effect with institutions. While, the lagged GDPPC factors control for economic cycle effects.

Their headline finding is that corruption is always negatively, and significantly correlated with economic growth. As reported in their paper, and shown in table 9 above when controlling for cyclical economic effects, and institutional quality a one standard deviation (19.56 point) increase in corruption is associated with a 14 to 19 percentage decrease in economic growth, between models 2, 4, and 6. All else equal, the authors show that more corrupt countries experience lower levels of growth in the aggregate. Further, on the basis of these results they assert that corruption has a significant negative effect on economic growth in countries with poor institutions, and that the effect of corruption on growth is mitigated by more effective governance, better rule of law, and more institutionalized democracy (Gründler and Potrafke, 2019). They further extend this argument with an analysis of interactive conditional marginal effects between corruption and government effectiveness, rule of law, and democracy, as shown in figure 10. The results shows that the effect of corruption on growth is highly negative in countries with lower levels of democracy, rule of law, or government effectiveness, but becomes insignificant at high levels of democracy, and moderate levels of government effectiveness or rule of law. Collectively, their results indicate that corruption does decrease economic growth in the aggregate, all else equal an identical counter factual country with lower corruption would be expected to experience more rapid growth. Further, this tendency is somewhat mitigated by institutional characteristics, the quality of democracy, established rule of law, and government effectiveness. Higher levels of democracy reduce the negative impact of corruption on growth, while countries with below average effectiveness and rule of law suffer the most serious negative effects of corruption on growth rates.

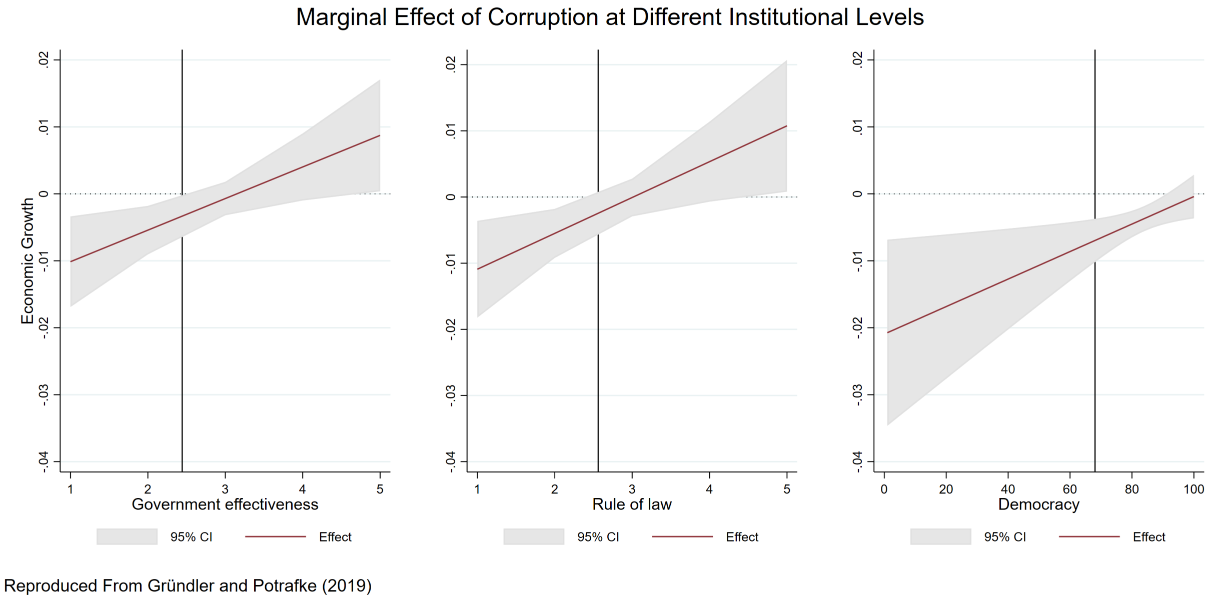


Figure 16: Interactive Marginal Effects of Corruption and Institutional Effects

The above model was re-estimated using the Outcome Index of corruption in place of CPI and the results are presented in table 10 below.[14] This was done by merging the Outcome Index with the data provided by Gründler and Potrafke and utilizing the replication code from the original article. Consequently the data and methods are identical, except for the notable loss of ten countries, because the underlying data was not available to calculate the Outcome Index. This could bias the result, specifically because it excludes several potentially low capacity, and high corruption cases. Such as Syria, Libya, North Korea, and Venezuela, on the other hand, it also excludes Taiwan and Hong Kong. As a robustness check the original CPI model was re-estimated in Appendix A using only the overlapping cases from Table 10. The results are theoretically consistent with the original work and are both statistically less significant, and have poorer explanatory power than the Outcome Index based results shown here.

[14]The Outcome Index and CPI are correlated at .70 in this data.

Table 10: Re-estimation Using Outcome Index of Corruption: DV log real per capita GDP.

	Effectiveness		Rule of Law		Democracy	
	(1)	(2)	(3)	(4)	(5)	(6)
Corruption	-0.00910**	-0.0129**	-0.00853*	-0.0126*	-0.0123***	-0.0120**
	(0.00456)	(0.00589)	(0.00451)	(0.00651)	(0.00411)	(0.00495)
Log (GDPPC)(t-1)		0.751***		0.752***		0.738***
		(0.0973)		(0.0976)		(0.0821)
Log (GDPPC)(t-2)		-0.145		-0.157		-0.170*
		(0.101)		(0.103)		(0.0861)
Log (GDPPC)(t-3)		0.0184		0.0197		0.0266
		(0.0458)		(0.0479)		(0.0385)
Log (GDPPC)(t-4)		-0.0201		-0.0236		0.0251
		(0.0394)		(0.0405)		(0.0468)
Corruption x Gov. Effect.	0.00156	0.00375*				
	(0.00160)	(0.00192)				
Corruption x Rule of Law			0.00140	0.00357		
			(0.00170)	(0.00221)		
Corruption x Democracy					0.00949*	0.0106**
					(0.00488)	(0.00494)
Government Effectiveness	0.144	0.192**				
	(0.0875)	(0.0918)				
Rule of Law			0.167*	0.207*		
			(0.0953)	(0.116)		
Democracy					0.399*	0.472**
					(0.211)	(0.217)
Observations	781	781	781	781	796	796
Countries	162	162	162	162	165	165
R-Squared	0.352	0.682	0.359	0.683	0.298	0.652
RMSE	0.0362	0.0255	0.0360	0.0254	0.0393	0.0277
F Stat	21.51	80.31	28.33	82.22	16.32	69.65

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.2 Comparison to Published Results

As shown in Table 10 the overall results from re-estimation match the original findings. Corruption is negative and significant across all six model specifications, confirming the core argument from the original paper. Although the effect size measured by the expected change from a one standard deviation (6.37) increase in corruption ranges between 7.6 and 8.2 percent. Accounting for approximately half the effect size when using CPI in the original analysis. Further, as noted earlier higher levels of effective governance, rule of law, and democracy are positively and significantly associated with growth in the full models, and the hypothesised interactive effects are noted for government effectiveness and democracy. Consequently, when the Outcome Index is used as a substitute corruption measure it is once again demonstrated that countries with less effective governance, or lower levels of institutionalized democracy are expected to experience lower growth rates.

An evaluation of the interactive marginal effects shown in Figure 11 highlights the similarity of the findings. Democracy is positively correlated with higher growth rates. In fact, the association is actually stronger when the Outcome Index is used, and persists even to high levels of democracy, with the 95 percent confidence interval remaining below 0 even at maximum levels of democracy. Similarly, government effectiveness is shown to have both a positive and significant direct effect, and an interactive effect with corruption, sustaining the original findings by Gründler and Potrafke using CPI. Although, the slope is shallower, and therefore the effect less substantively significant. The one notable difference is the institutional effect of the rule of law. Rule of law remains positive, both directly, and in interaction with corruption, and the marginal effects graph indicates the expected impact for lower levels of rule of law. However, it is insignificant in the interactive term and no conclusion can be drawn that better rule of law increases growth rates in the context of corruption.

Finally, since the objective of this replication is to test how well the Outcome Index can be used as a substitute measure for corruption the root mean squared error (RMSE) can be used as a means to directly compare the results. Since both estimations share logged real GDP per capita as the dependent variable RMSE functions as an absolute measure of fit expressed in terms of the standard deviation of the unexplained variance in logged real GDP per capita. Lower values indicate less unexplained variance, and therefore more accurate model predictions. The RMSE values from table 10 are about half the original published values from table 9, and slightly lower than the RMSE shown in the overlapped model from Appendix A. The comparability of these results is notable, and can be interpreted as strong support for the outcome approach especially considering that the index is derived from substantially different data than CPI.

As discussed in the literature review CPI is based on survey responses from subject matter experts and members of the business community about how corrupt they perceive governance to be in a given country, it is fundamentally a perception based approach. In contrast, the Outcome Index is based on measuring variation in population level impacts of corruption and is fundamentally designed to estimate unobserved levels of corruption in a given country. These are markedly different approaches, yet even the most conservative comparison of these estimations indicates that on the basis of RMSE the Outcome Index performs at least as well, and certainly no worse than CPI at predicting the impact of corruption on economic growth.

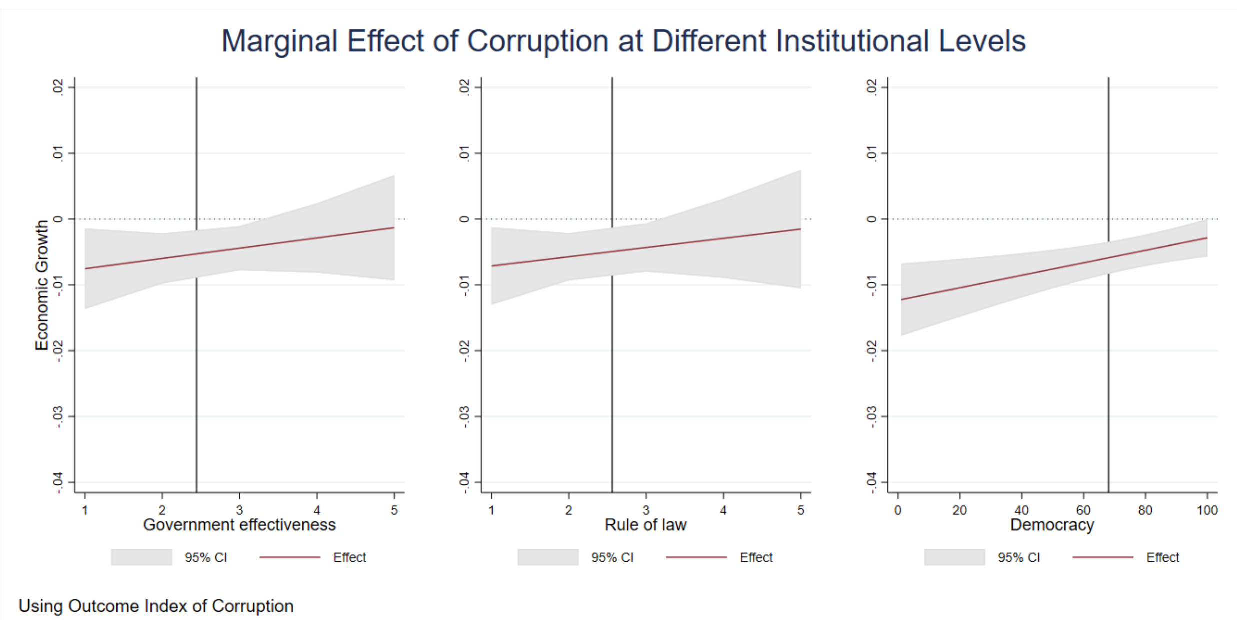


Figure 17: Interactive Marginal Effects of Corruption and Institutional Effects

Overall, these results provide strong support for the use of the outcome measure as a substitute for CPI in cross-national political corruption research, and demonstrate that it is a functional measure of corruption which can produce expected results. As shown above the outcome measure demonstrates similar results when substituted for an established measure of political corruption, both when compared to the original model, and to a robustness check using only the matching cases. Further, the relationship between corruption and economic growth is a long running topic, although in recent scholarship quantitative analysis has largely settled on a consensus that corruption reduces growth (d'Agostino, Dunne and Pieroni, 2016; Cieślik and Goczek, 2018).

Consequently, this topic is a good replication target for testing the robustness of the outcome measure proposed here, because we can confidently predict the expected result. If the outcome measure has construct validity, that is to say, if it measures corruption accurately, then we should expect a negative and significant relationship between it and economic growth. As shown above this expected relationship is observed in all six of the replicated specifications. Further, this main result, a negative sign and significance, is robust to the inclusion of alternative indexes of government capability, and democracy, as well as GDPPC. This is important because it means that the index is not simply measuring a correlation between good governance and low corruption, or between increasing economic development and low corruption. Previous sections have demonstrated that the components of the Outcome Index can predict existing measures of corruption, in keeping with the theoretical concept of the outcome index approach. This section validates the theory that corruption can be effectively measured from population level outcomes and demonstrates that the resulting index estimates can be used as an effective proxy for commonly used corruption measures.

6 Conclusion

This project seeks to contribute toward solving a long-running problem in the quantitative study of political corruption. Specifically, the issue of how to identify and measure it by defining a new cross-national index based on variation in population level outcomes from, or symptoms of corruption. The primary issue in quantitative corruption research is actually finding corruption, measuring it, and scoring it to facilitate analysis. This is necessarily difficult because corruption is by it's very nature at least partially hidden, rarely tabulated in official statistics, and almost never directly observed.

As a result, this dissertation innovates in two primary ways by both suggesting a new way to look for corruption, and by applying a novel machine learning technique to predict it. First, it inverts the usual logic of corruption research by looking for symptoms, or outcomes of corruption in order to identify novel predictors. Rather than asking how corrupt a country is, or how corrupt it could plausibly be? It instead seeks to identify population level markers with known associations to corruption and to then track variation in these indicators as a means to measure unobserved changes in corruption itself. Second, it innovates by applying an extreme gradient boosting predictive machine learning model to estimate corruption for the first time, and to validate the selected indicators. Finally, it has proposed, defined and constructed a new outcome based index of corruption and demonstrated it as a substitutable measure of corruption.

There are currently three existing approaches to measuring political corruption, each based on a different conceptual technique and set of underlying assumptions. First, institutional constrain measures which attempt to infer levels of corruption based on the presence or absence of specific legal, governmental, and social institutions. This school operates from the assumption that specific institutional and structural elements mediate corruption, and that although they cannot measure corruption directly, countries with weaker institutions will be more vulnerable to, or likely to be corrupt. Constraint measures are often produced in waves and rely on detailed information about the political, legal, and social institutions of a country. Further, they are dependent on assumptions about causal associations between specific institutional characteristics and corruption. Second, corruption frequency measures, which attempt to infer corruption from data about the presence, absence, or level of response to corrupt activities within a country. This school operates from the assumption that they can estimate corruption based on how frequently corrupt activities are reported, prosecuted, or exposed in a country. Corruption frequency measures are very difficult to construct cross-nationally, largely because of data variability between countries, because of differences in what is considered 'corruption.'

The third, and most notably approach are perception based measures which attempt to infer the level of corruption through survey analysis of individual opinions. The process varies from one organization to the next, but it generally involves surveying a group of experts on a given country. Often including various combinations of academics, business people, non-governmental officials, diplomats, and sometimes the actual population itself. These responses are then standardized, sometimes peer reviewed, and then aggregated in a way that attempts to correct for observable bias. Notable, examples include the Corruption Perception Index produced by Transparency International, and the World Bank Control of Corruption Index, as well as several less frequently used measures. Perception based measures are the most viable and frequently used option for cross-national research, largely because of availability and implicit comparability. Further, perception based measures are highly correlated with each other, even when using different aggregation and indexing methodologies, suggesting that they do posses some degree of accuracy. However, they are fundamentally based on perceptions, opinions, or qualitative judgements by subject matter experts. Simultaneously, they are also complicated and expensive to produce, time consuming, and their actual sources are rarely transparent. Further, they are unfortunately always open to criticism of the validity of the underlying survey opinions.

This project has contributed a theoretical forth approach, an outcome based measure derived from novel population level factors which vary in associated with political corruption, and demonstrated how those measures can be used to directly estimate expected levels of corruption from objective data. Specifically, this project implements a weighted index utilizing infant, under five, and adult mortality, two measures of vaccination rates, health spending, and gross national income per capita to estimate national levels of corruption. Put more simply, it allows the detection of the shadow of corruption from readily available aggregate national level data and produces corruption estimates in a more transparent, less time consuming, and less costly way. This approach is potentially superior to existing measures because it is quantitatively based on indicators which have a proven relationship as outcomes of changes in relative corruption. Further, it does not rely on assumptions about how institutional settings, or environments 'should' impact corruption. It can be calculated using a small number of variables all of which are commonly available, even in developing countries, and it does not require a deep knowledge of the circumstances, governance, or history of a specific country.

Substantial consideration was given to selecting and testing potential corruption outcome indicators. Identified indicators were required to be objective quantitative evaluations with a clear scale and criteria. Further, potential indicators were required to have a proven published association with corruption and inferred causal relationship. This project has reviewed available literature on the effects of corruption to identify seven of these measures. Including infant, and under five mortality, adult mortality rates, immunization rates, healthcare spending, and gross national income per capita, as an economic control factor. Health outcomes were particularly valuable because they are discreetly measured, widely collected internationally, and most importantly highly responsive to corruption in part because corruption disproportionately impacts the well being of the most vulnerable in society and manifests as a second order effect in poorer health, increased mortality and less healthcare expenditure.

Further, this project has innovated by exploring the use of supervised machine learning techniques as a means of predicting the relevance of socio-economic outcomes to estimating out of sample corruption for the first time. Specifically utilizing extreme gradient boosting to evaluate indicators and prove their predictive importance by comparing them as predictors of existing corruption measures and to generate predictive importance weights to allow for differential contributions by predictors in the final index. Building on this it has constructed an aggregated index of cross-national relative political corruption directly from the source variables using a weighted Z-score standardization approach. Additionally, it has presented a comparative ranking of countries to allow comparison to the ranking approaches used by most published corruption indexes. Finally, it has tested the resulting index in a cross-national replication study to prove that they can be used interchangeably with CPI as a measure of political corruption and achieve expected quantitative results.

Undoubtedly, this index could be further refined. Additional predictors could potentially be added, and more data sources could be aggregated to extend the coverage. However, as presented here it represents a functional first step in producing a direct estimate of corruption without requiring subject mater expert analysis. As such it has the potential to be used both academically, as a substitute for, or a robustness check against other perception based corruption measures. Specifically, it resolves the long-running concern about unobserved bias in perception measures, or the idea that respondents could skew the results. It further addresses the concern that for many countries few of the respondents actually come from the culture, or national community they are assessing. As such it also escapes the need to bias correct those estimates or triangulate by aggregation. Further, although it was developed with reference to traditional corruption measures, the final aggregation and predictive estimation is separate from them, meaning that although source predictor selection was based on correlations with perception measures, the final estimate was drawn directly from objective predictors. Further, beyond academic research an outcome measure of corruption could be used in both business and government applications as a corruption sentinel. In this capacity the primary advantage is that the required predictors are widely available population level data and are often available even for less open societies. Further, although this project has demonstrated index creation at the national level, it could in fact be estimated at a sub-national, or even potentially regional level as a granular estimate of corruption.

A Replication Robustness Check

There is a difference in the number of observations between the original replication of Gründler and Potrafke (2019) presented in Table 9, and the re-estimation of their work with the Outcome Index presented in Table 10. This re-estimation was conducted a second time using only the overlapping cases [15] and the original CPI data as a means of testing the robustness of the results.

As shown in Table 11 below the overlapping re-estimation looks like a weaker version of the original Table 9. Corruption retains the expected negative sign, although it loses significance in models 2 and 4. Similarly, the interactive terms, and the institutional effects retain expected positive signs, with some losing significance. This performance is expected, since the reduction in cases is likely to weaken the results when compared to the original from Table 9. Notably, none of the variables reverse signs and none of the results become inconsistent or are theoretically contradictory to the original. Further, the RMSE from this modified replication is comparable to the results shown in Table 11, indicating that for this data the outcome measure is a substitutable corruption estimator, and that the results are not driven by the reduction in sample size.

[15] Those cases for which both the original article, and the Outcome Index have data.

Table 11: Re-estimation of Gründler and Potrafke (2019) Using Overlapping Observations

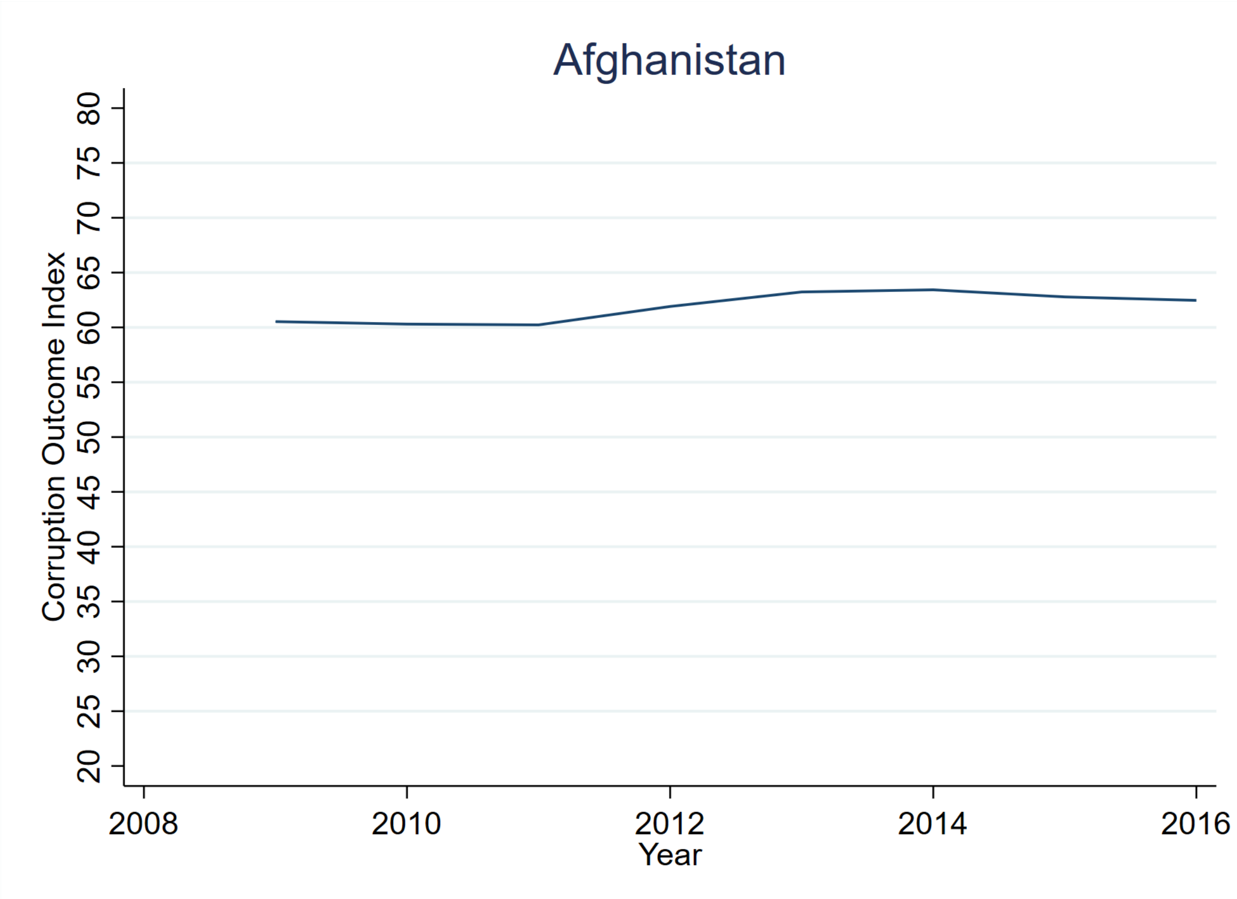
	Effectiveness		Rule of Law		Democracy	
	(1)	(2)	(3)	(4)	(5)	(6)
Corruption	-0.00582**	-0.00171	-0.00609**	-0.00150	-0.00556***	-0.00221*
	(0.00256)	(0.00131)	(0.00289)	(0.00165)	(0.00201)	(0.00117)
Log (GDPPC)(t-1)		0.748***		0.747***		0.733***
		(0.108)		(0.110)		(0.0884)
Log (GDPPC)(t-2)		-0.172		-0.175		-0.170*
		(0.107)		(0.109)		(0.0883)
Log (GDPPC)(t-3)		0.0259		0.0230		0.0281
		(0.0476)		(0.0493)		(0.0432)
Log (GDPPC)(t-4)		-0.00690		-0.0137		0.0188
		(0.0424)		(0.0423)		(0.0482)
Corruption x Gov. Effect.	0.00153*	0.000342				
	(0.000821)	(0.000395)				
Corruption x Rule of Law			0.00157*	0.000306		
			(0.000922)	(0.000563)		
Corruption x Democracy					0.00303	0.00115
					(0.00206)	(0.00124)
Government Effectiveness	0.125**	0.0304				
	(0.0511)	(0.0233)				
Rule of Law			0.156**	0.0526		
			(0.0622)	(0.0410)		
Democracy					0.0930	0.0524
					(0.0681)	(0.0422)
Observations	781	781	781	781	796	796
Countries	162	162	162	162	165	165
R-Squared	0.343	0.656	0.351	0.660	0.288	0.629
RMSE	0.0365	0.0265	0.0363	0.0263	0.0395	0.0286
F Stat	19.38	60.45	28.64	68.60	13.28	47.34

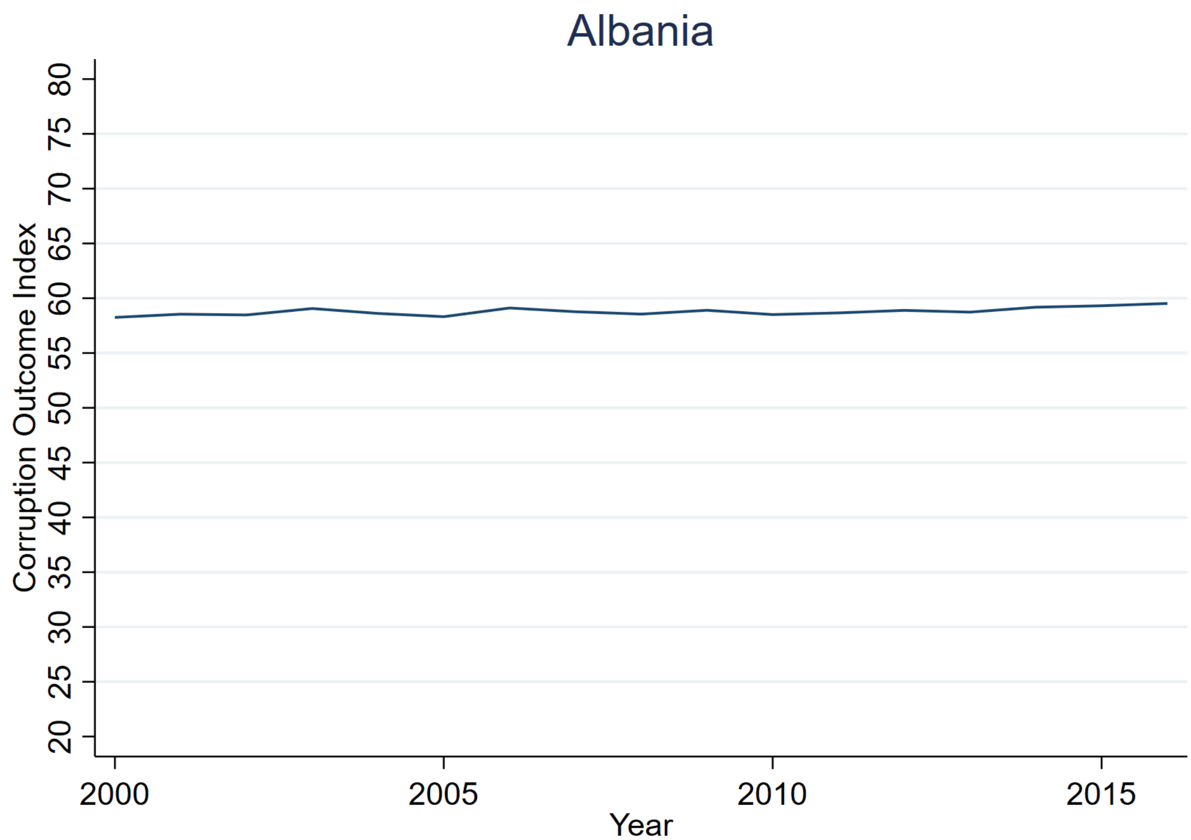
Standard errors in parentheses

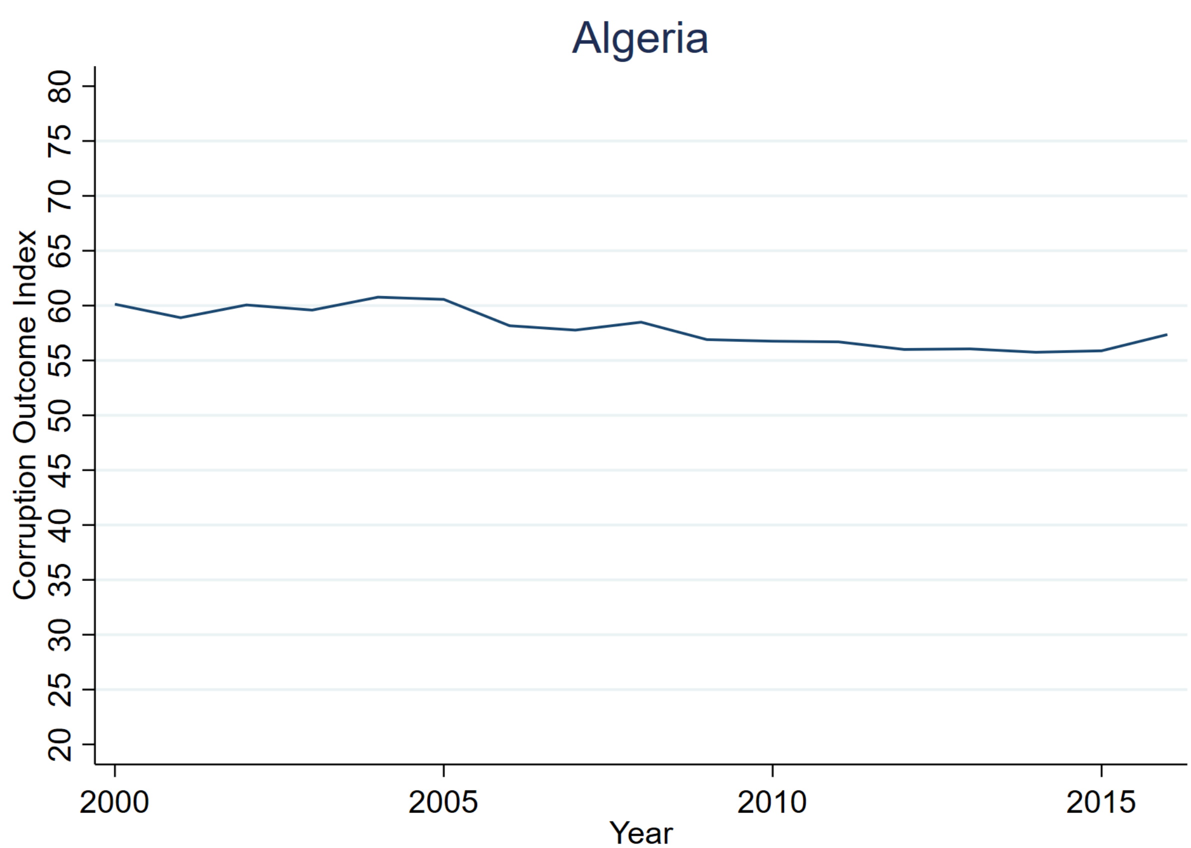
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

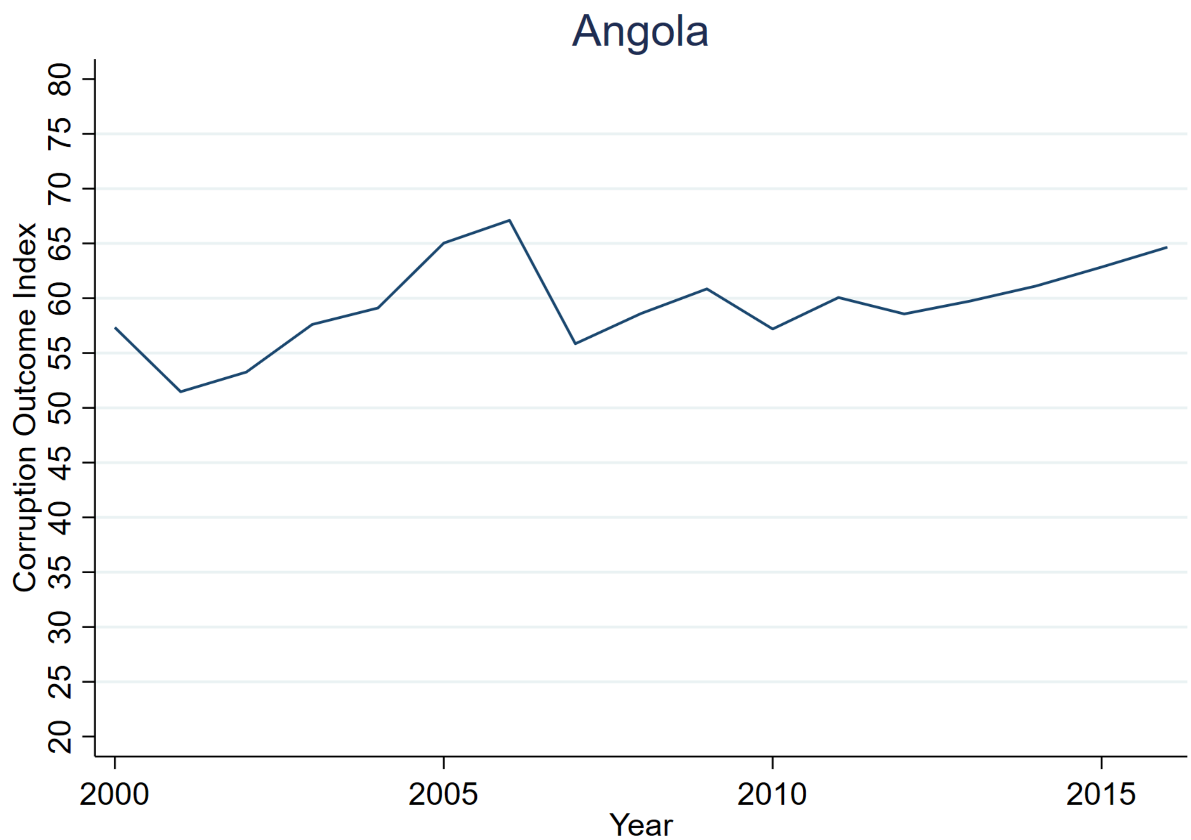
B Outcome Index of Corruption Graphs

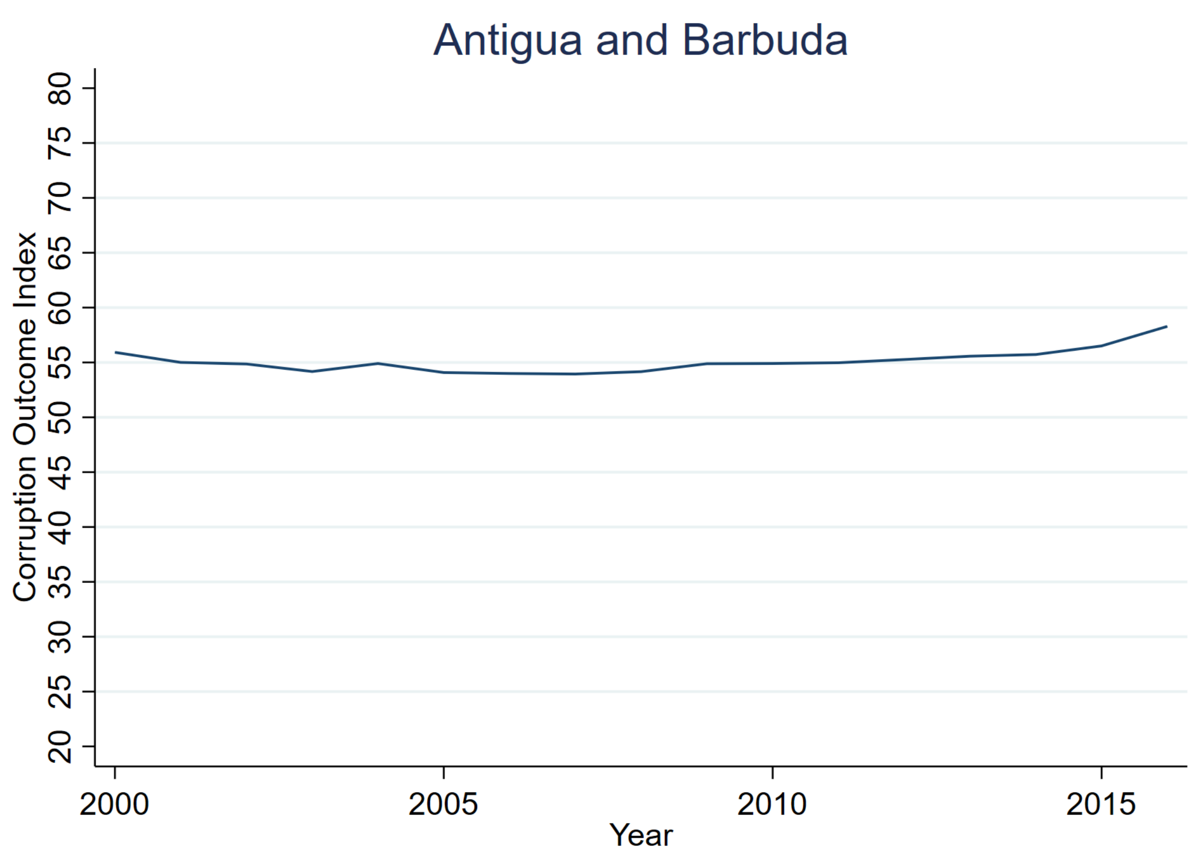
This appendix contains graphs for all 177 countries for which the Outcome Index of corruption has been calculated. Higher values indicate increasing levels of corruption. These graphs have been standardized to run between the highest, and lowest observed levels of the outcome indicator and are presented in alphabetical order.

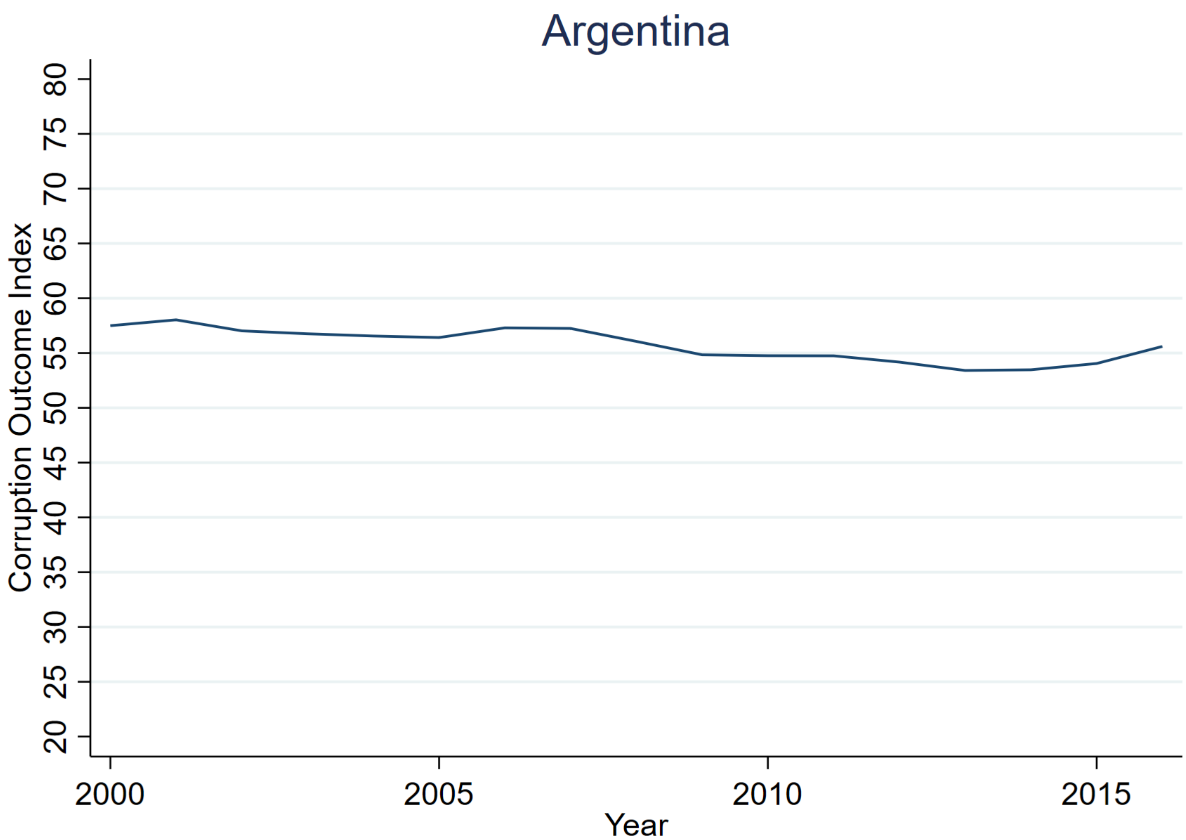


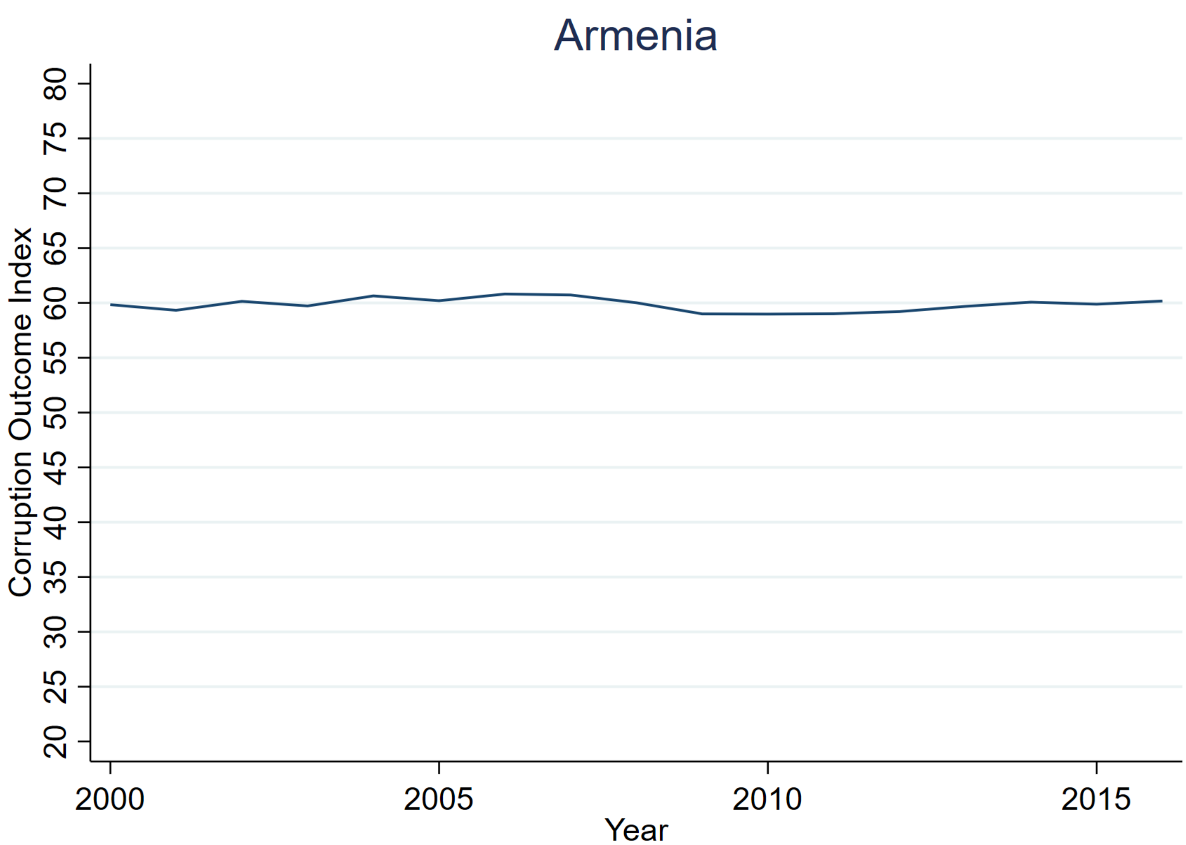

Albania
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

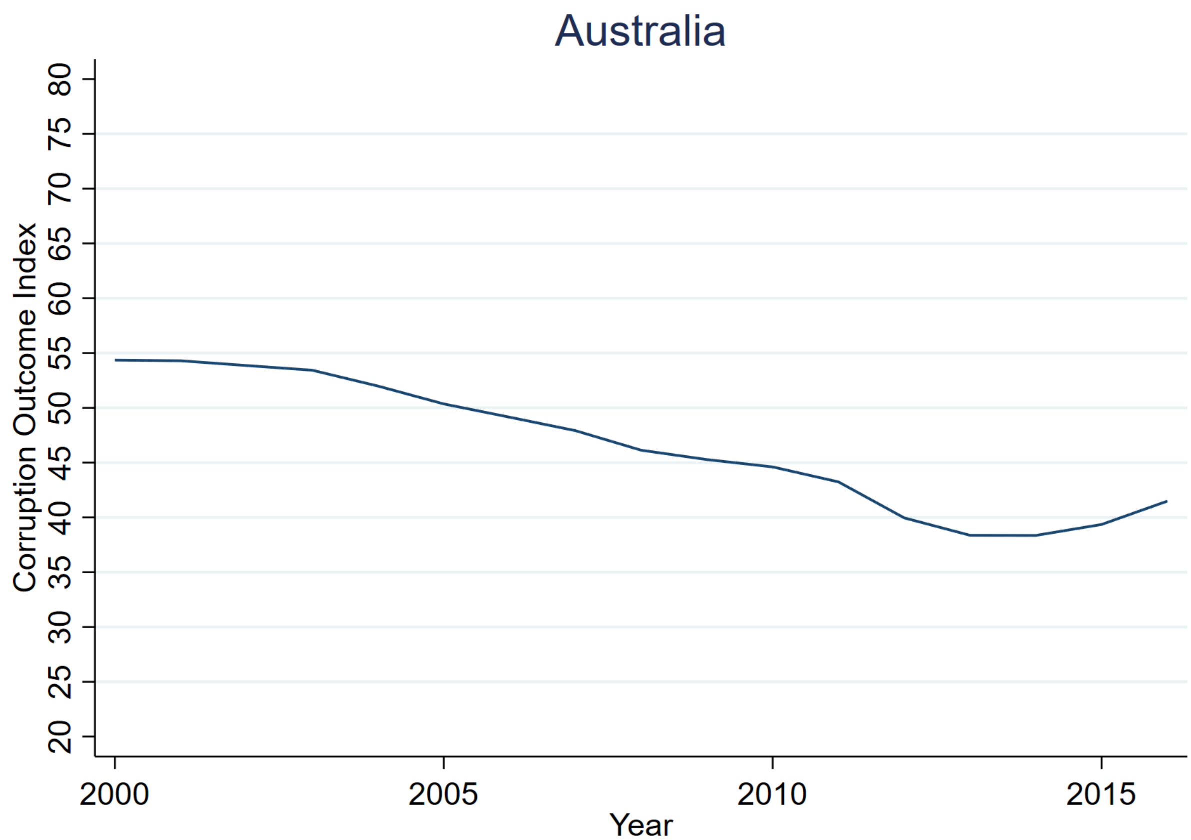

Algeria
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

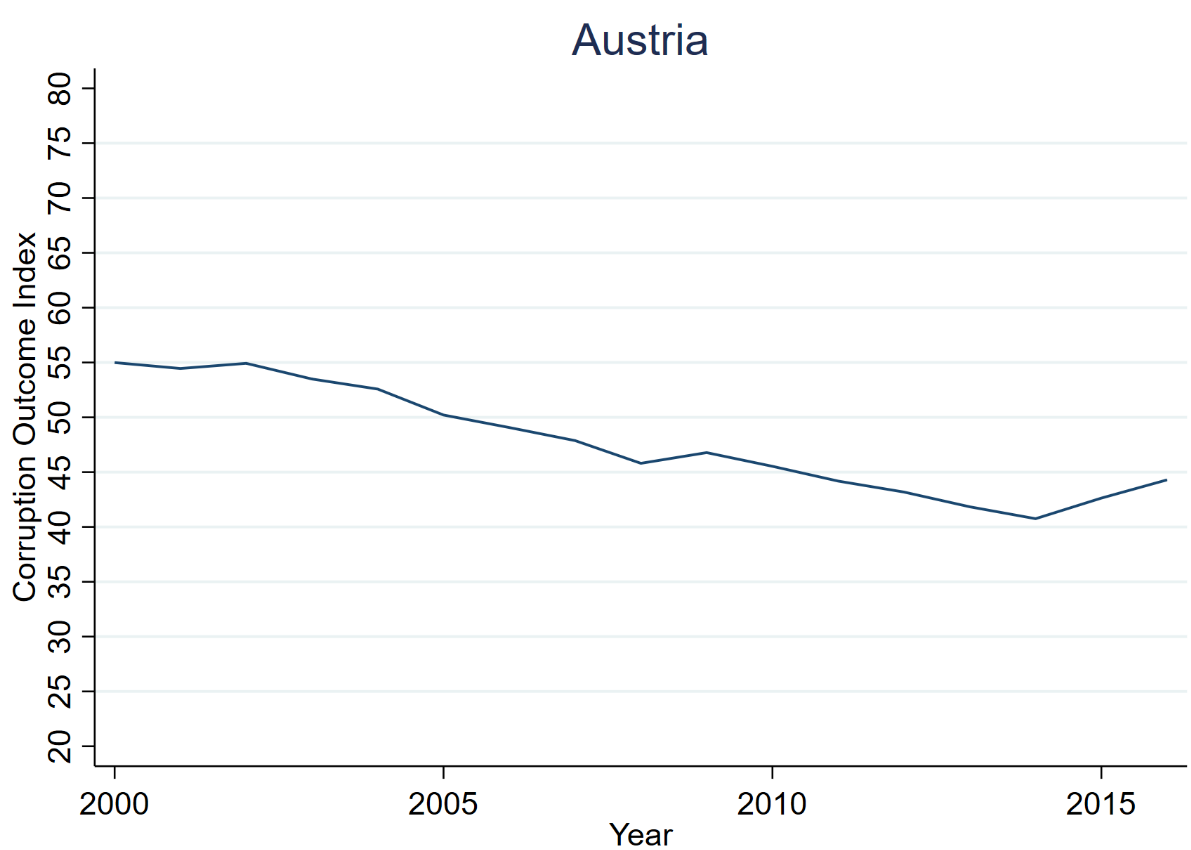

Angola
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Antigua and Barbuda
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

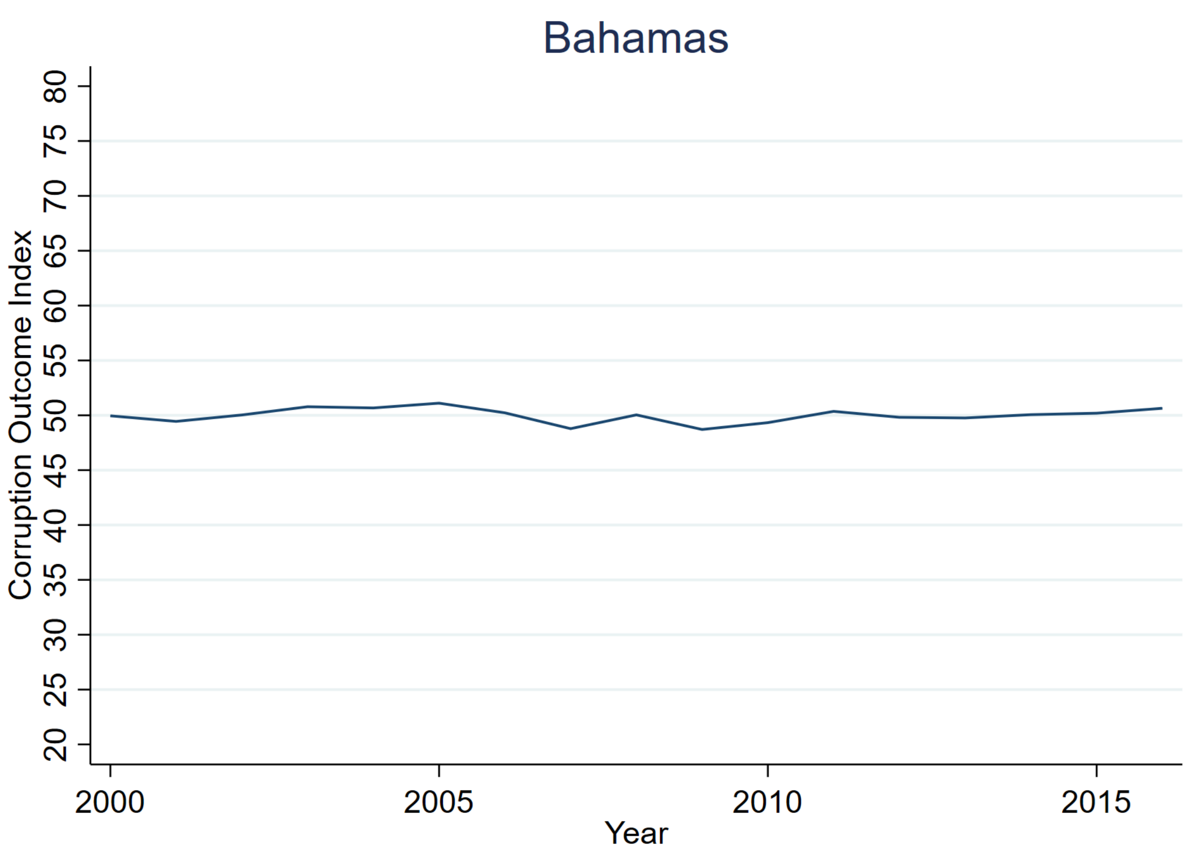

Argentina
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

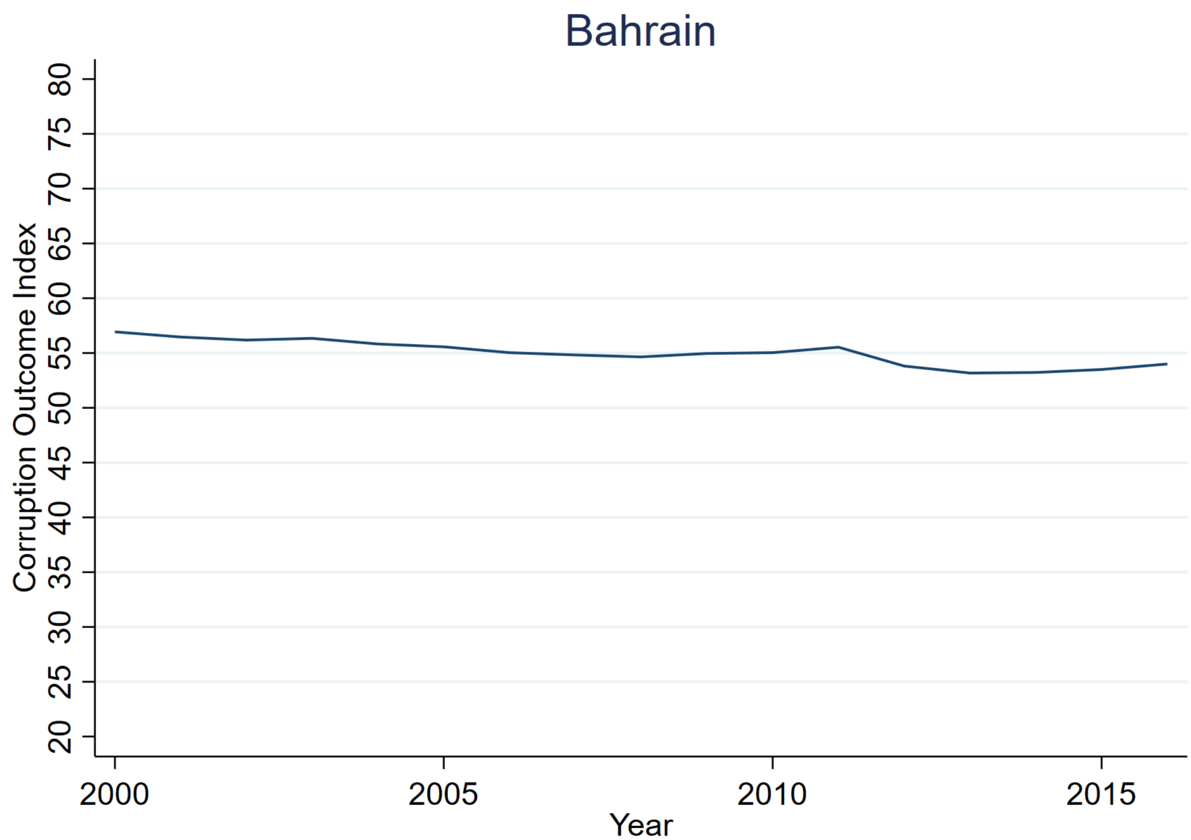

Armenia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

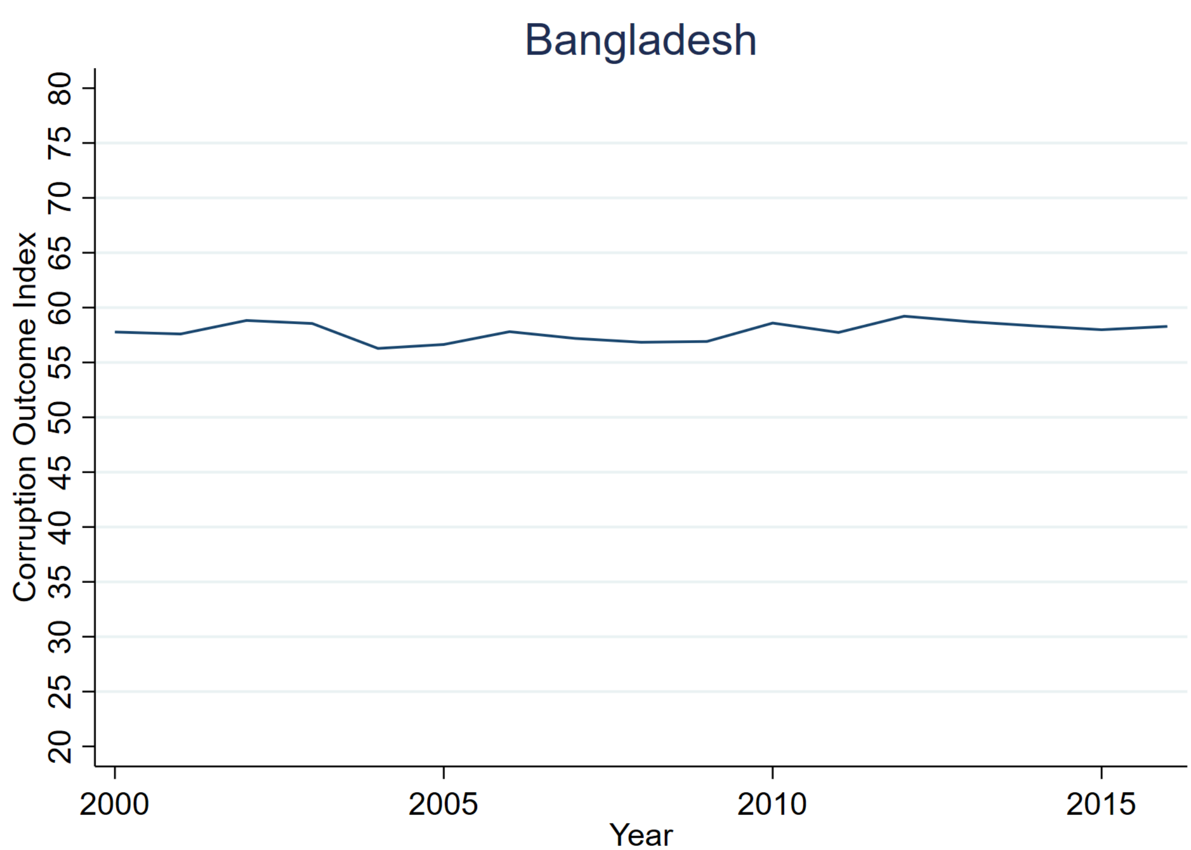

Australia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

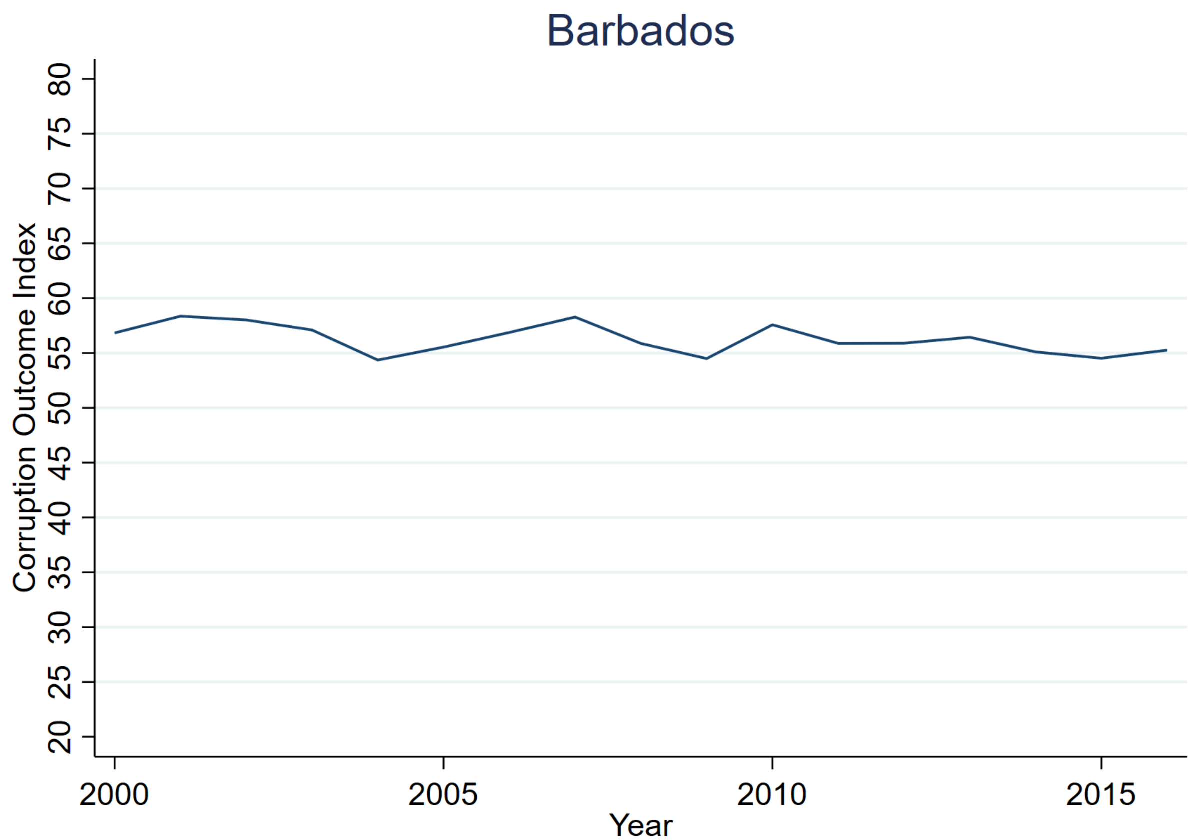

Austria
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

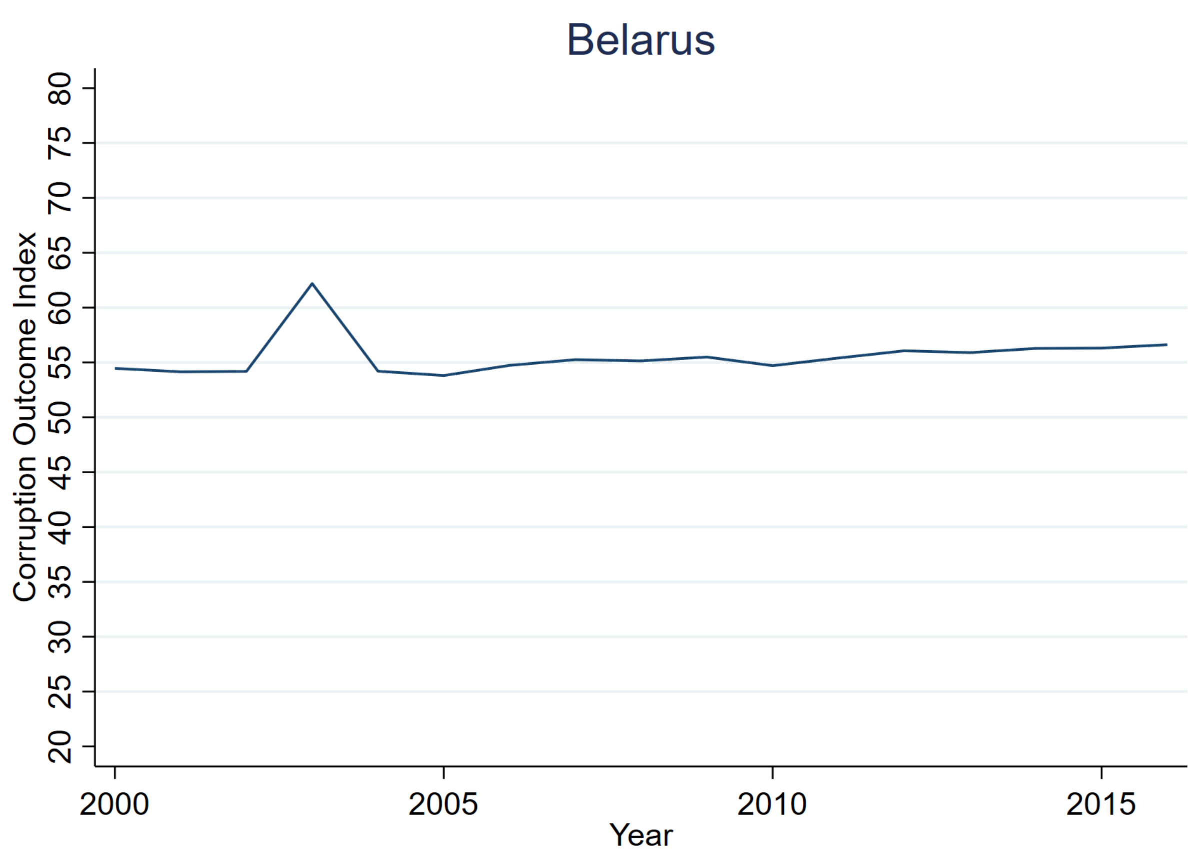

Azerbaijan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

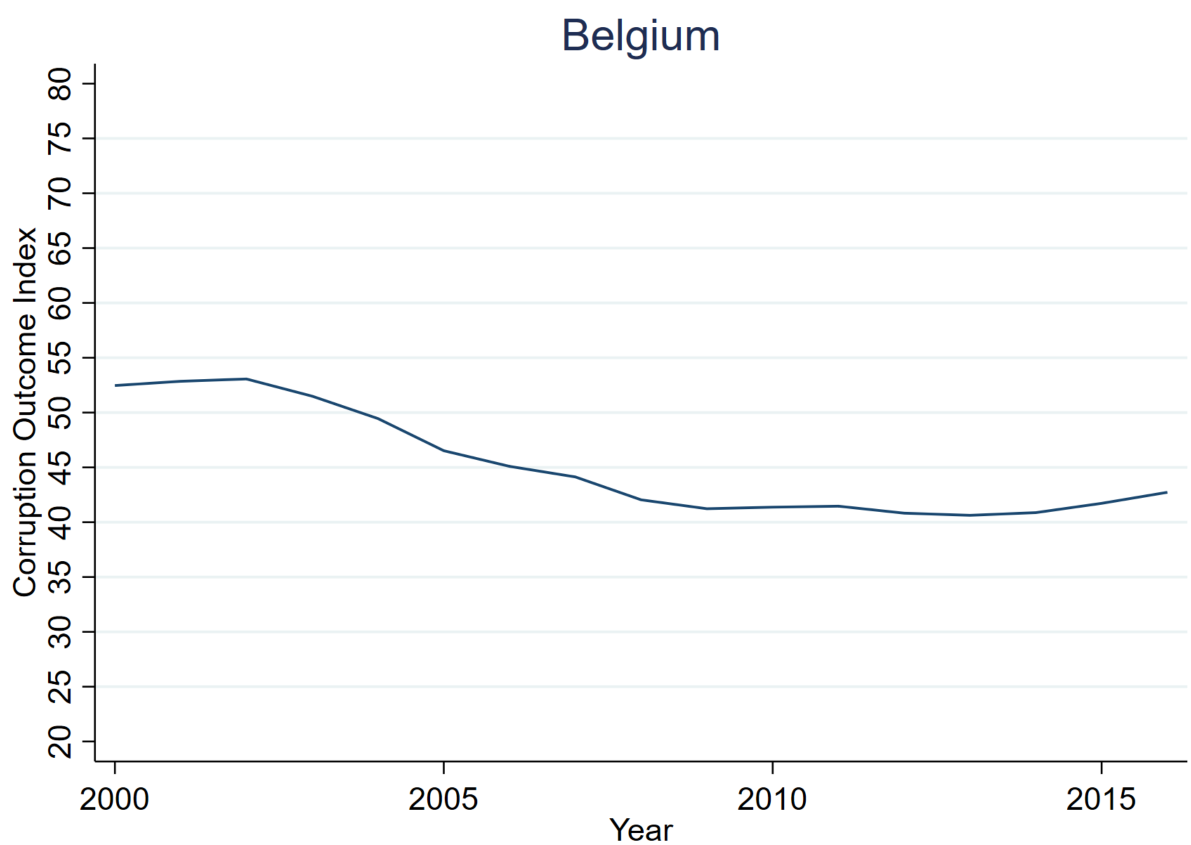

Bahamas
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

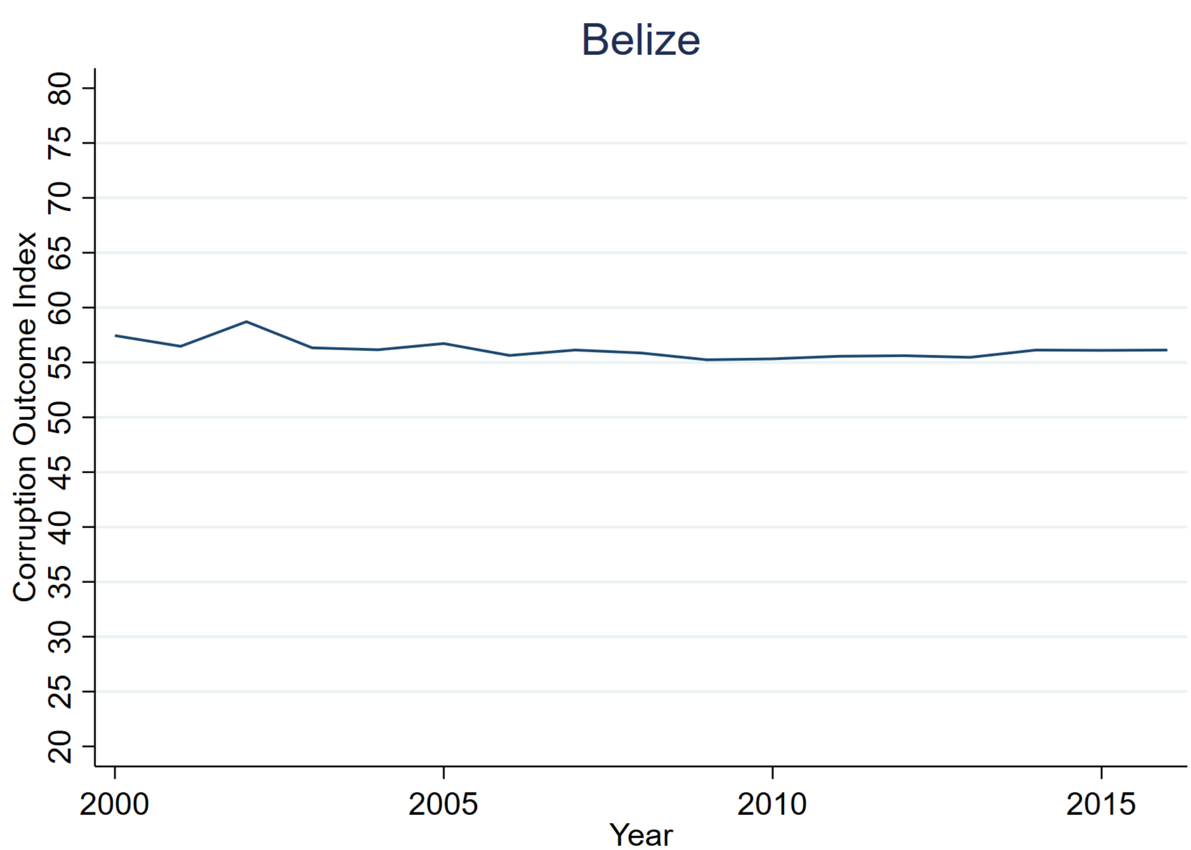

Bahrain
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

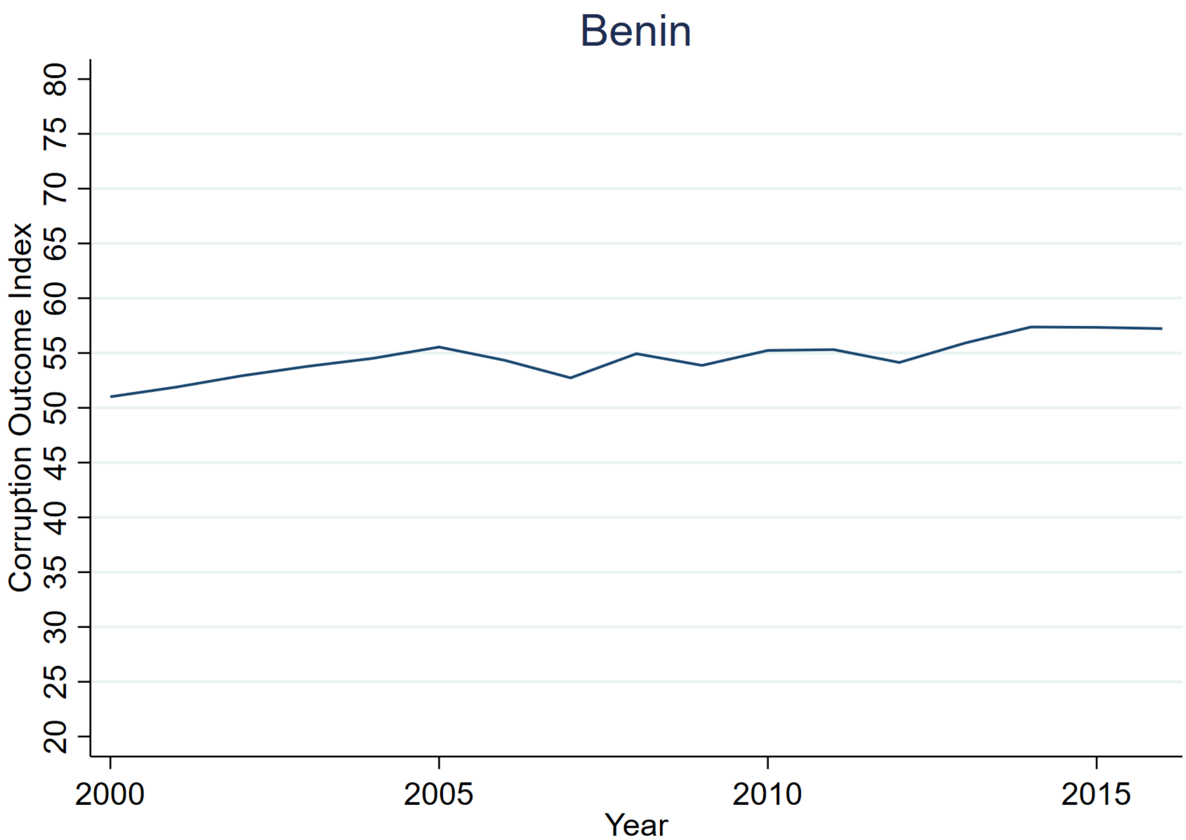

Bangladesh
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

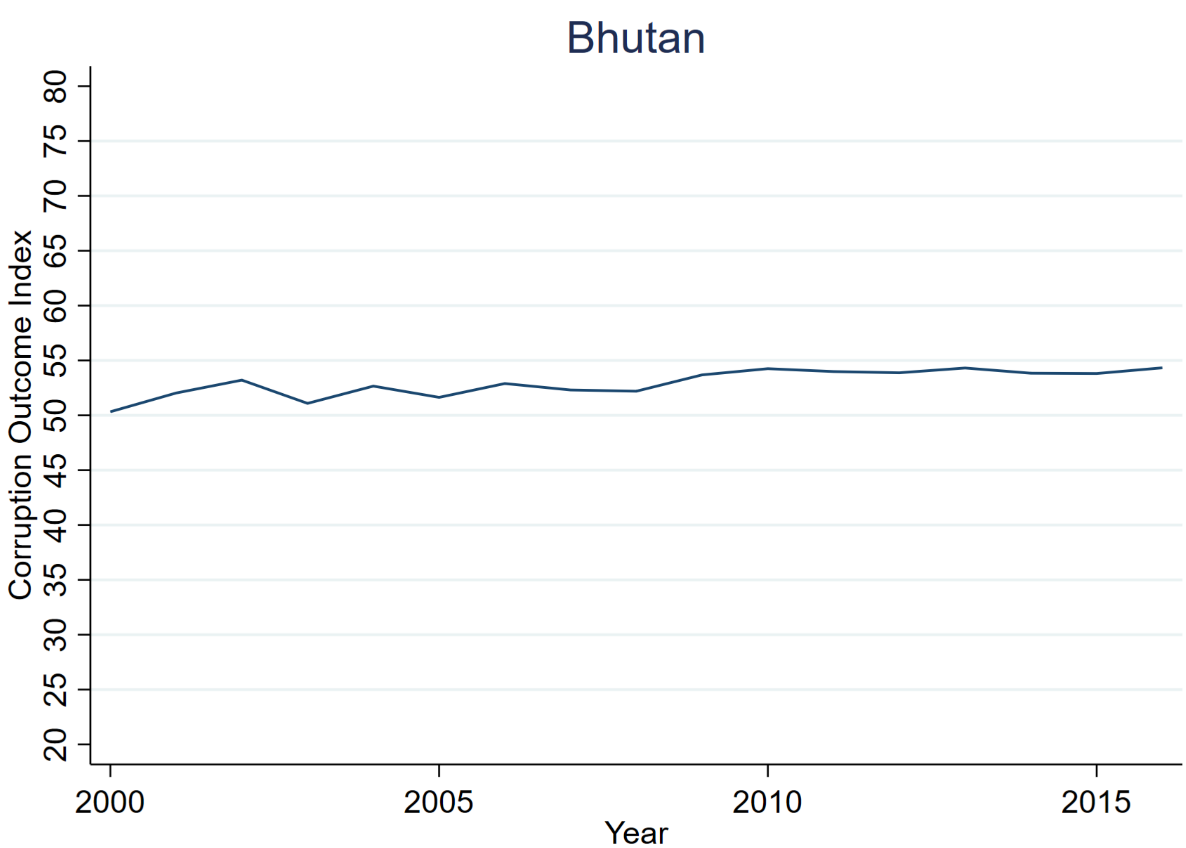

Barbados
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

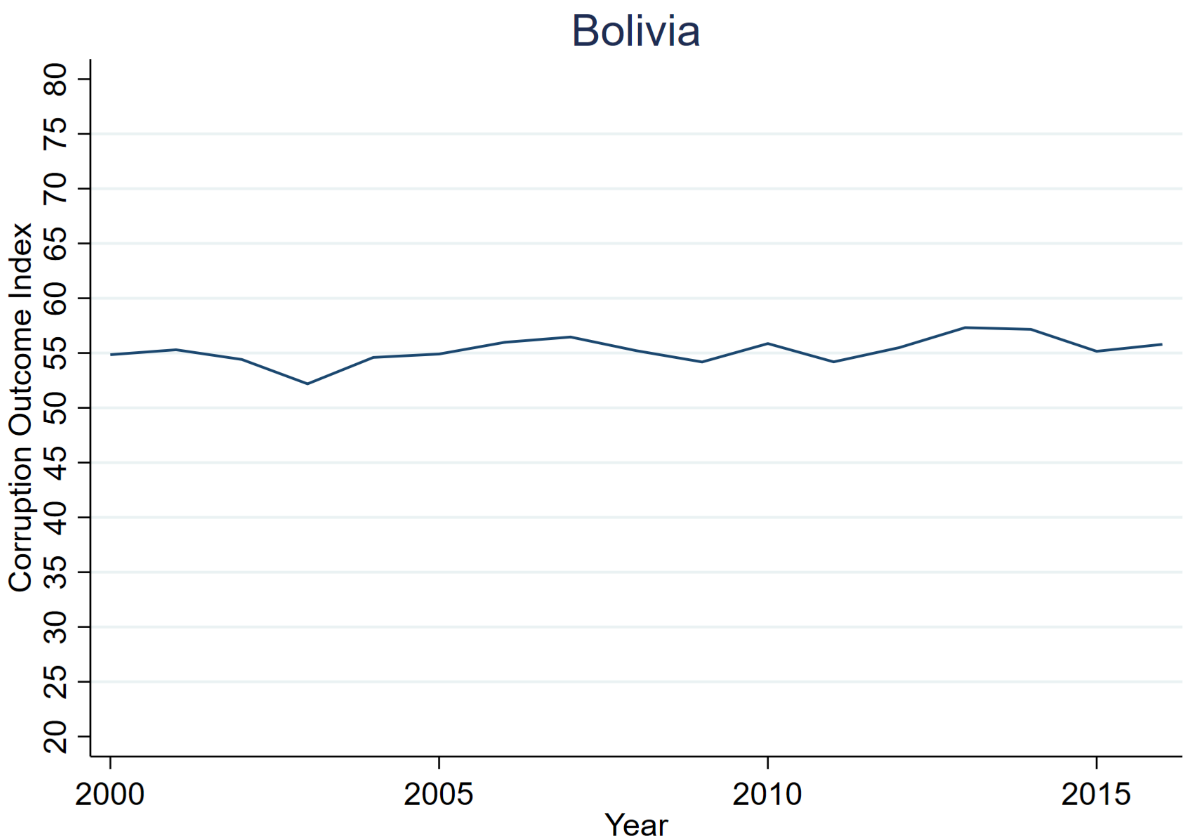

Belarus
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

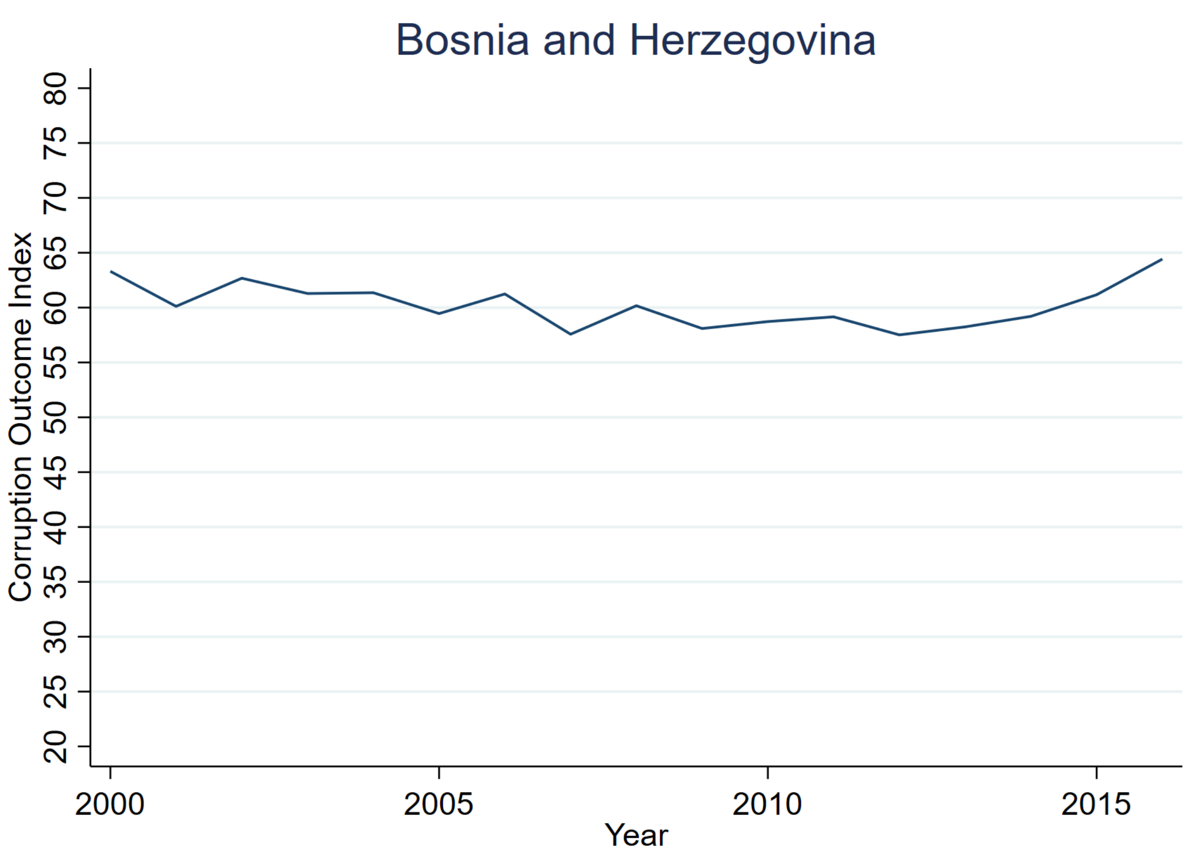

Belgium
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

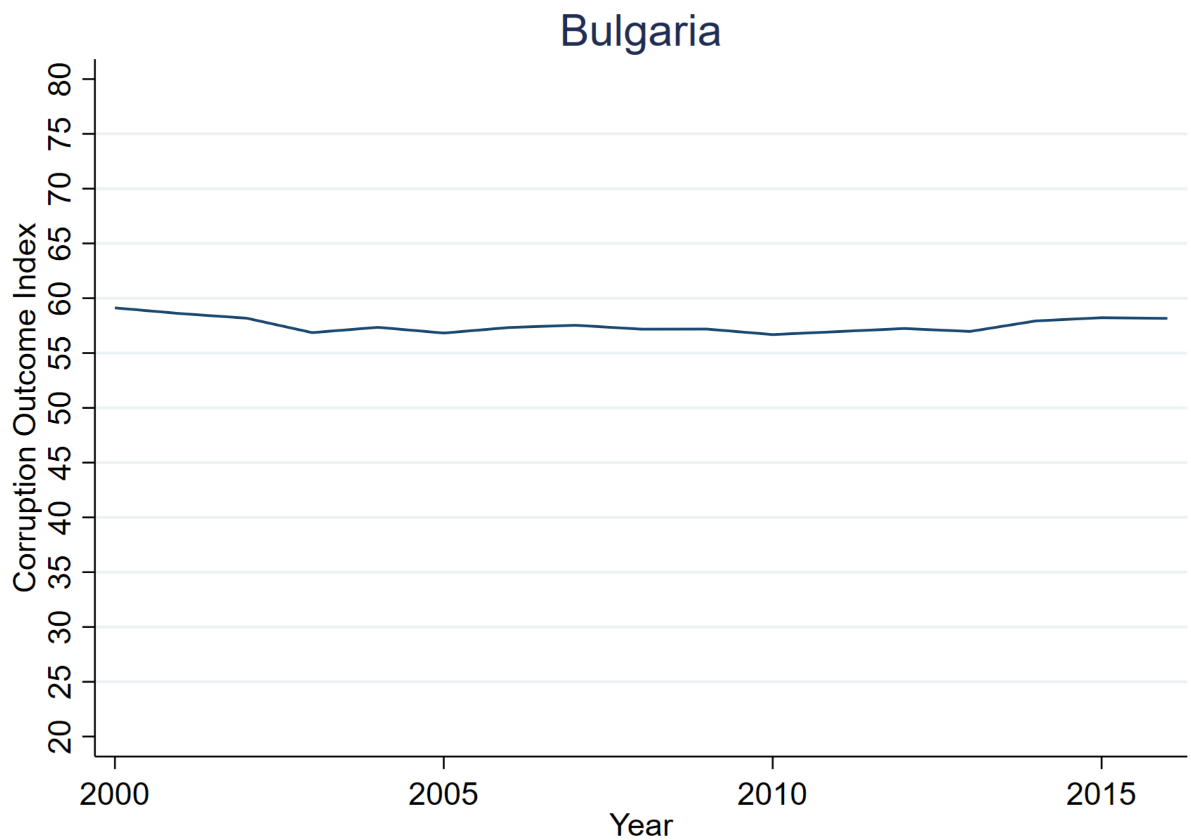

Belize
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

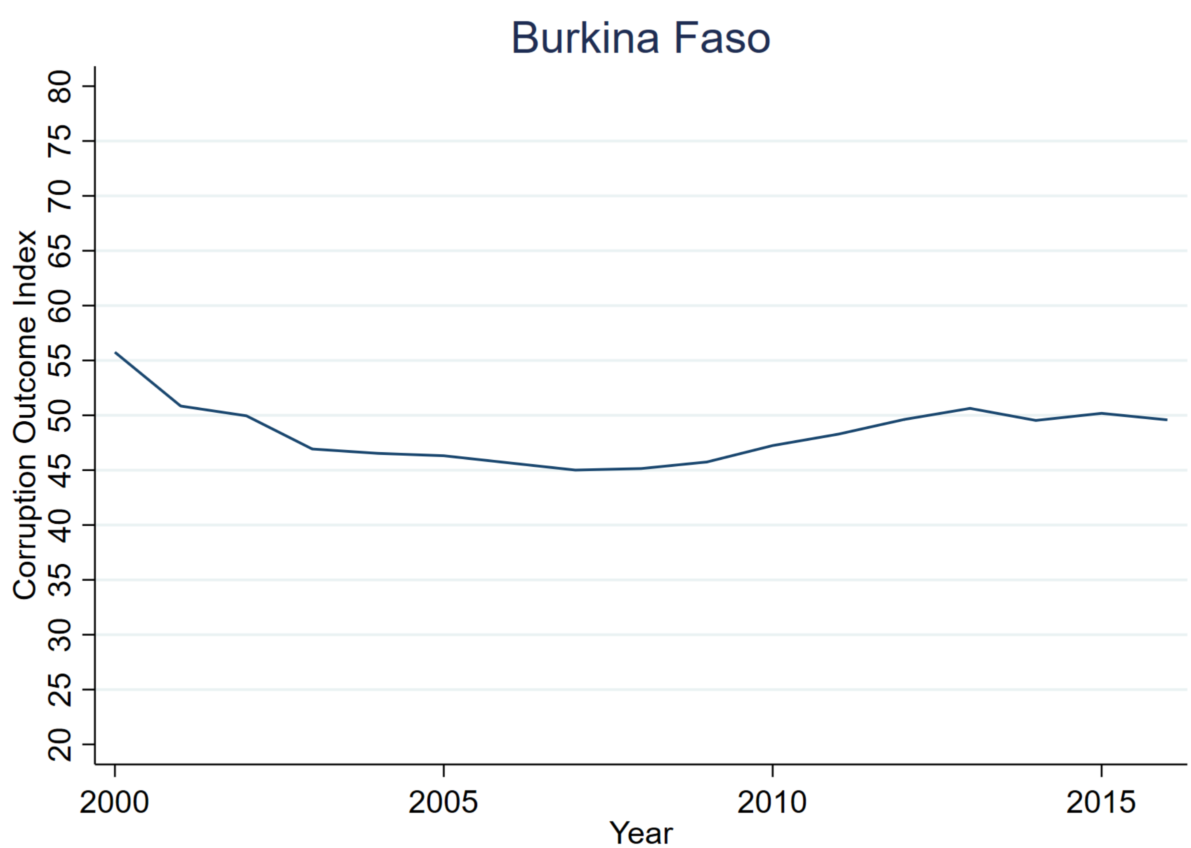

Benin
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

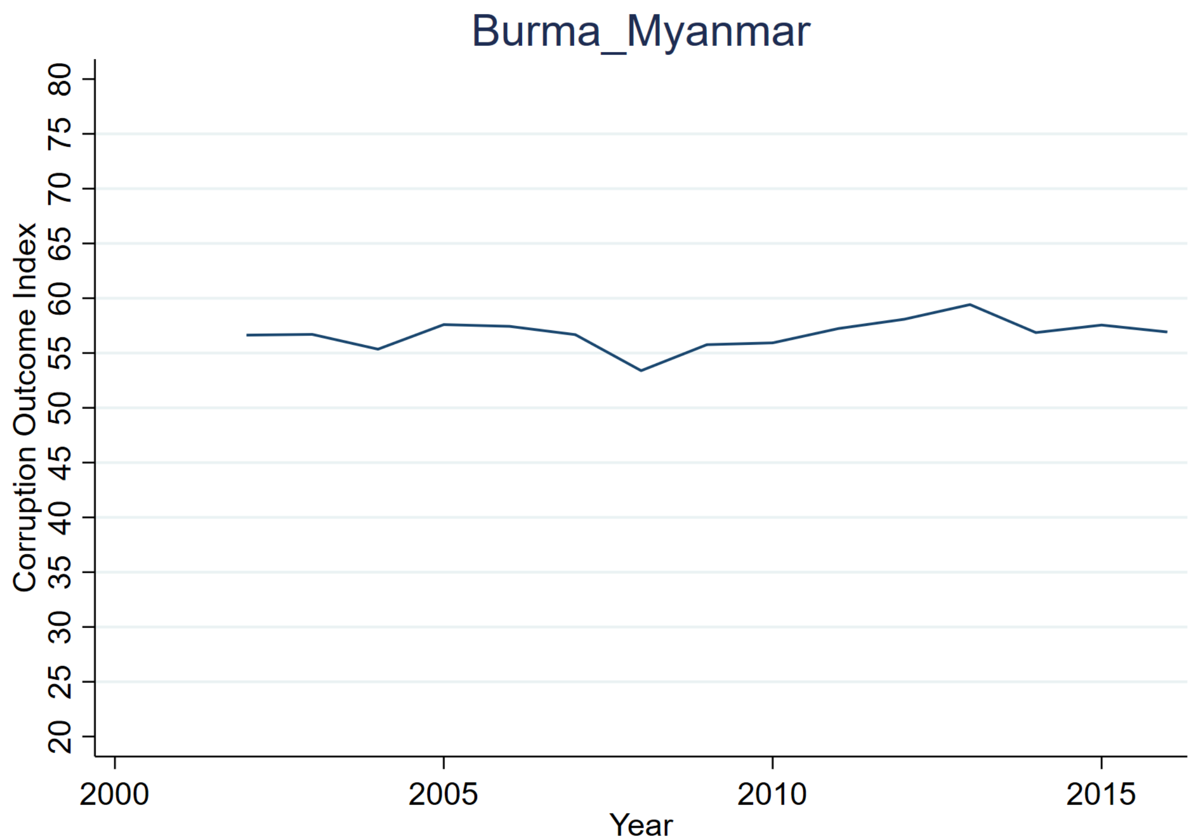

Bhutan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

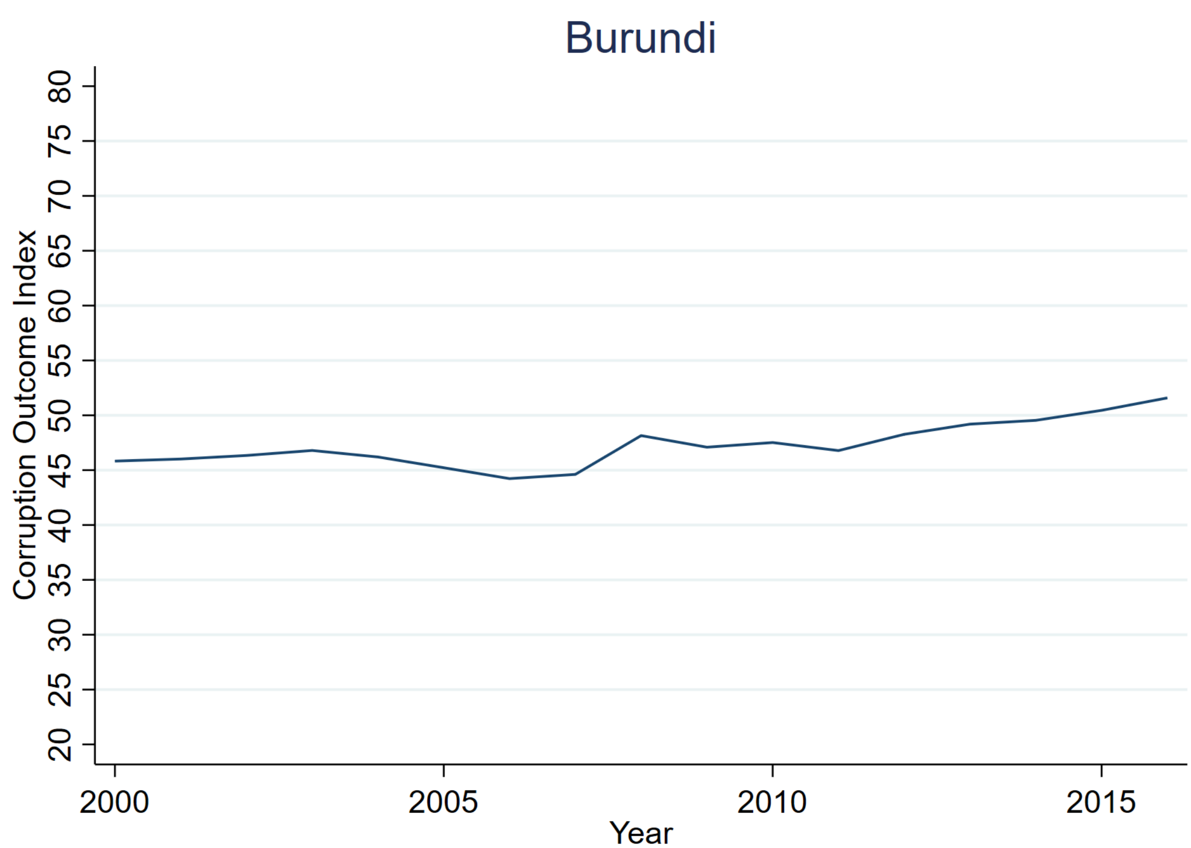

Bolivia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

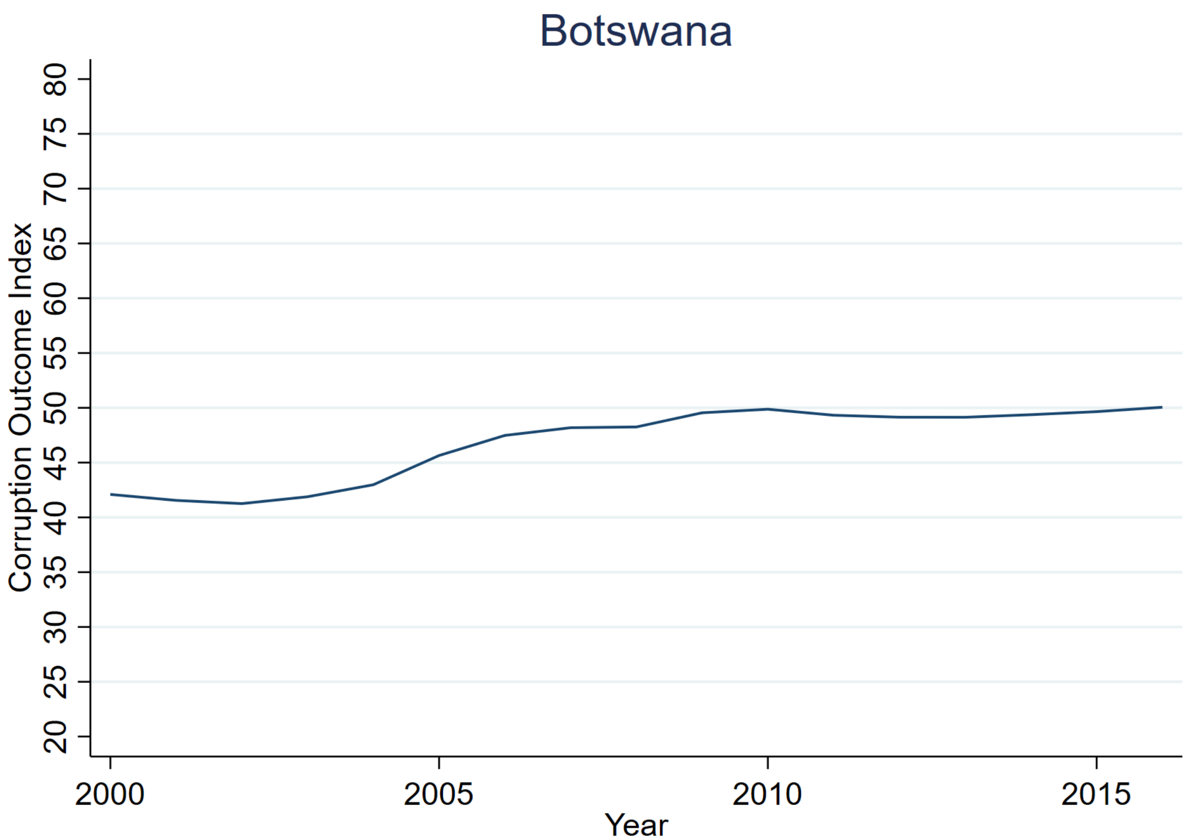

Bosnia and Herzegovina
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Bulgaria
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

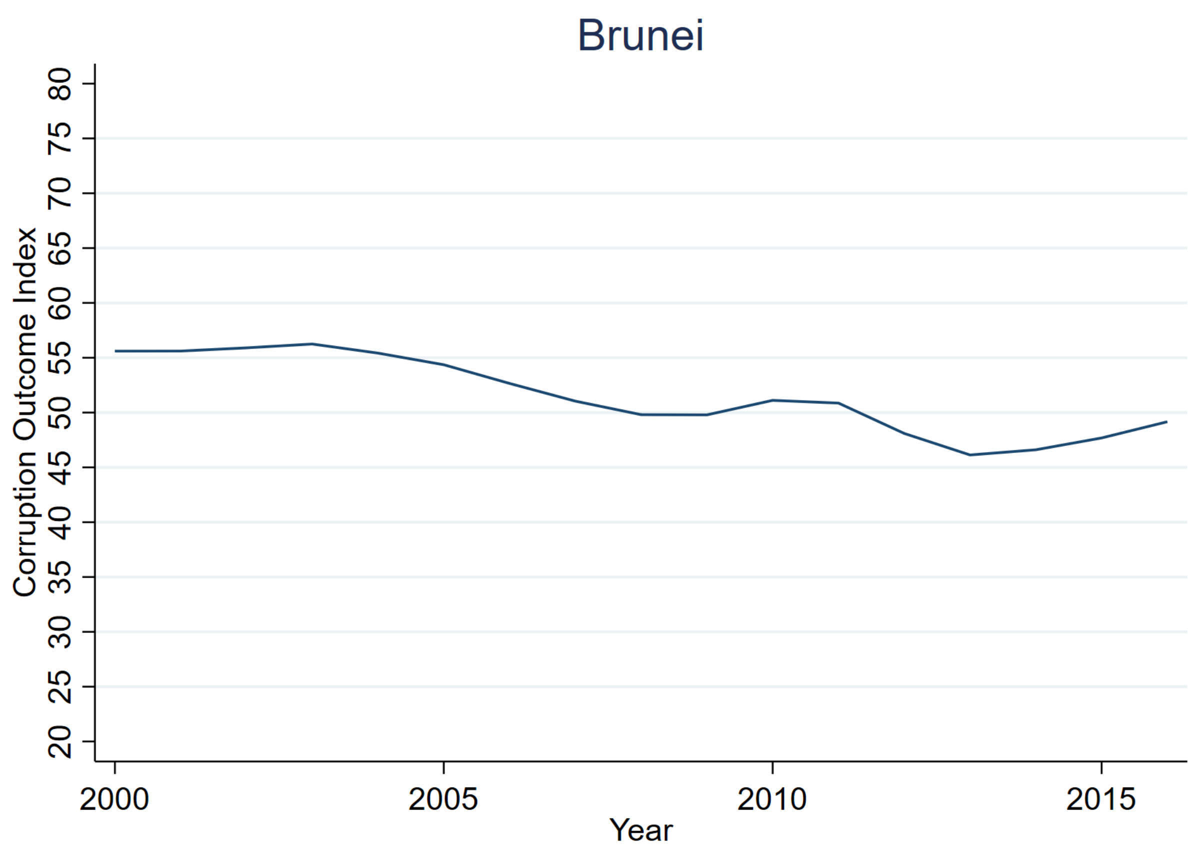

Burkina Faso
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

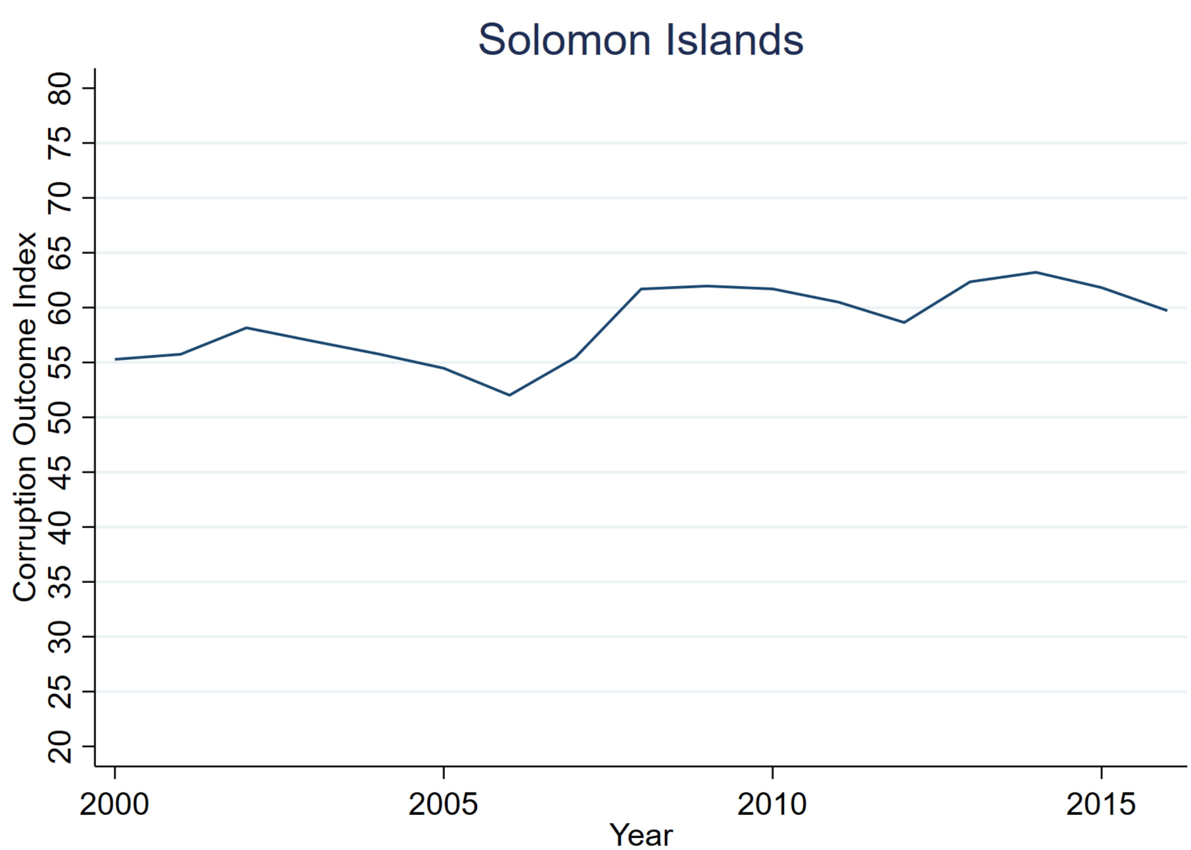

Burma_Myanmar
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

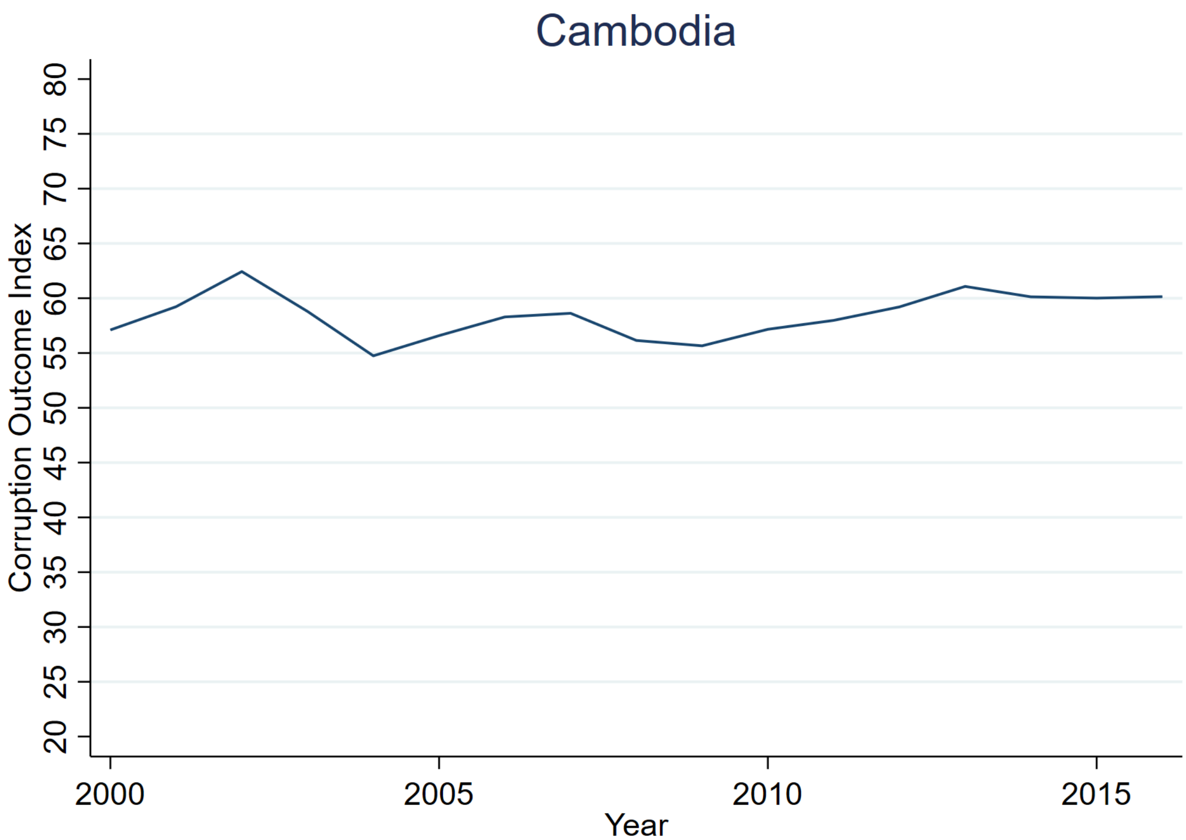

Burundi
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

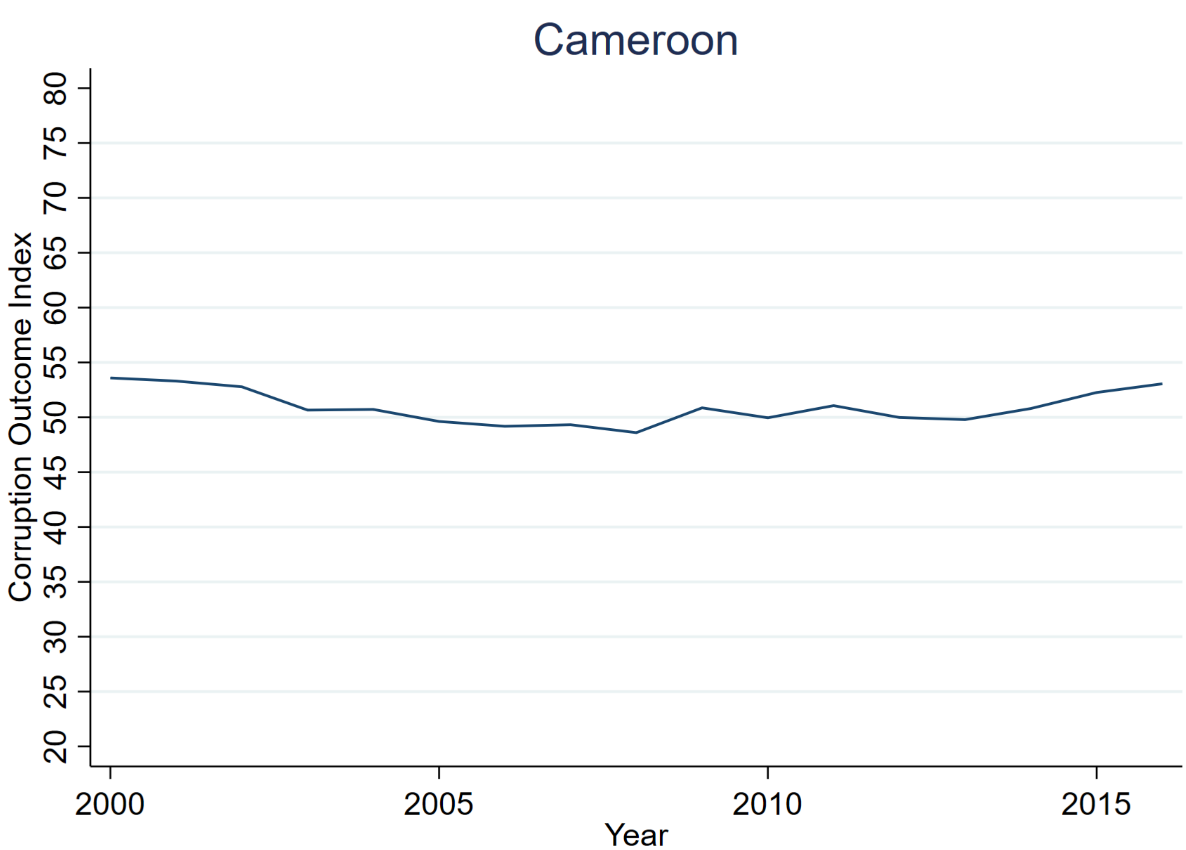

Botswana
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

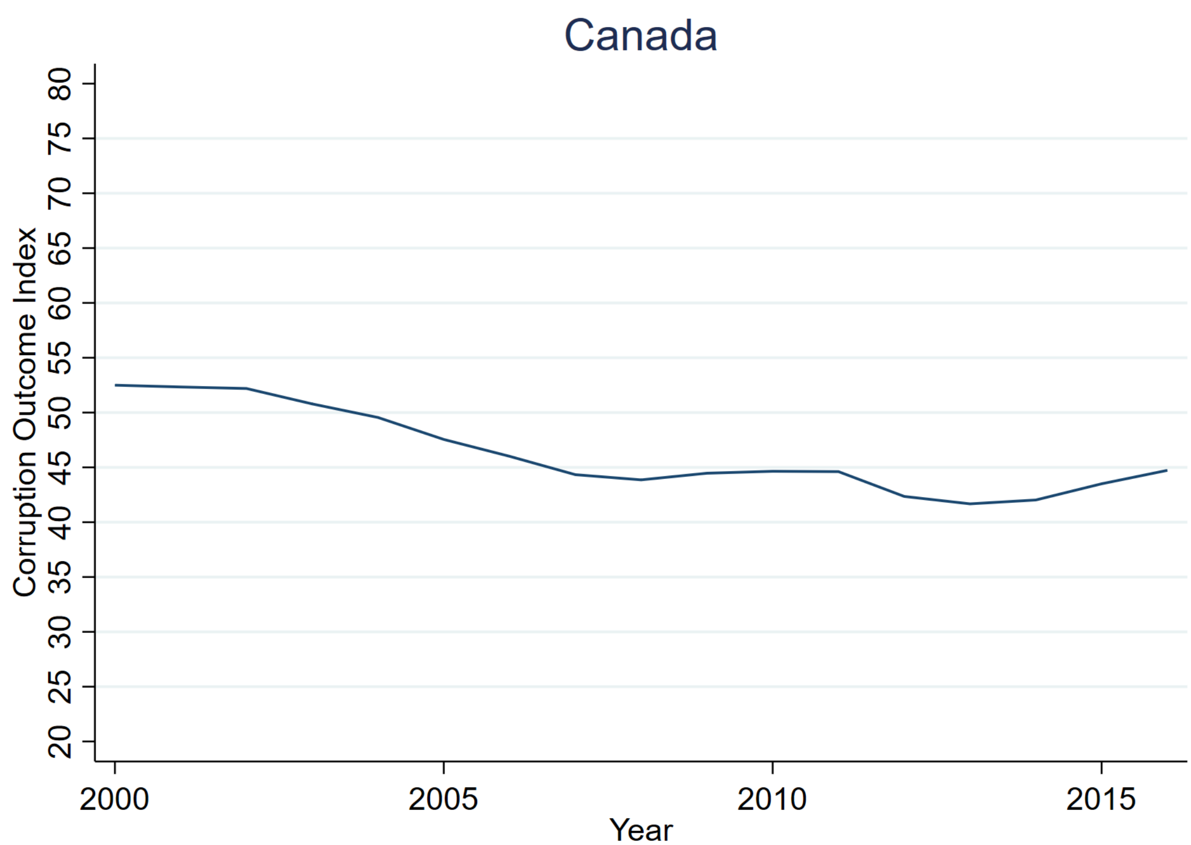

Brazil
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

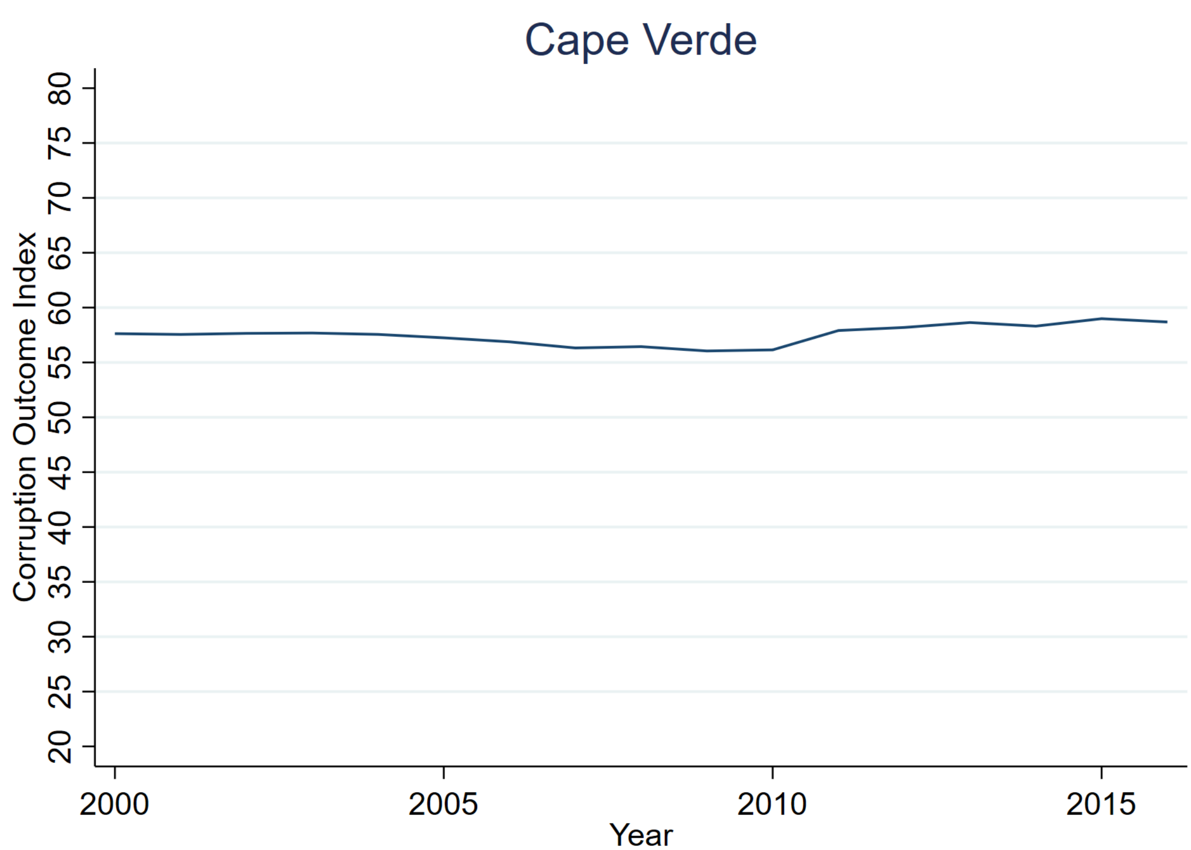

Brunei
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

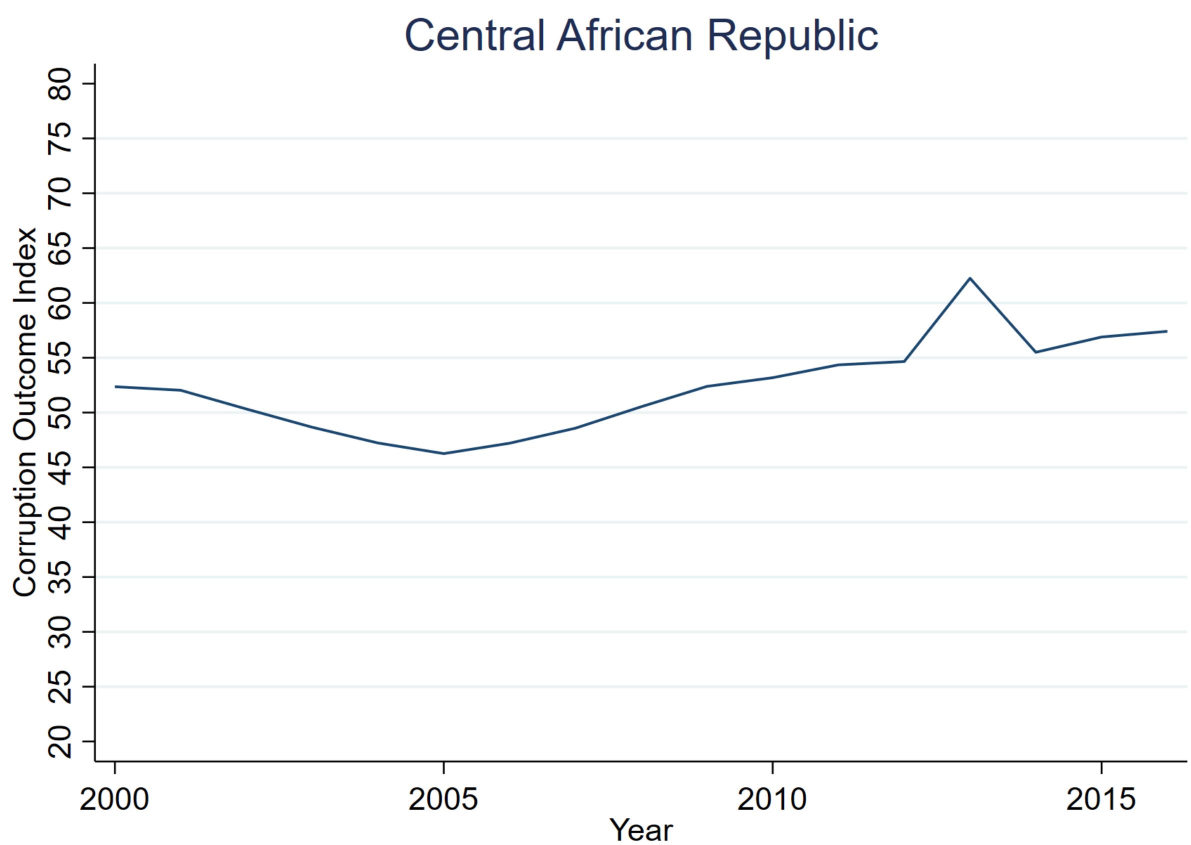

Solomon Islands
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

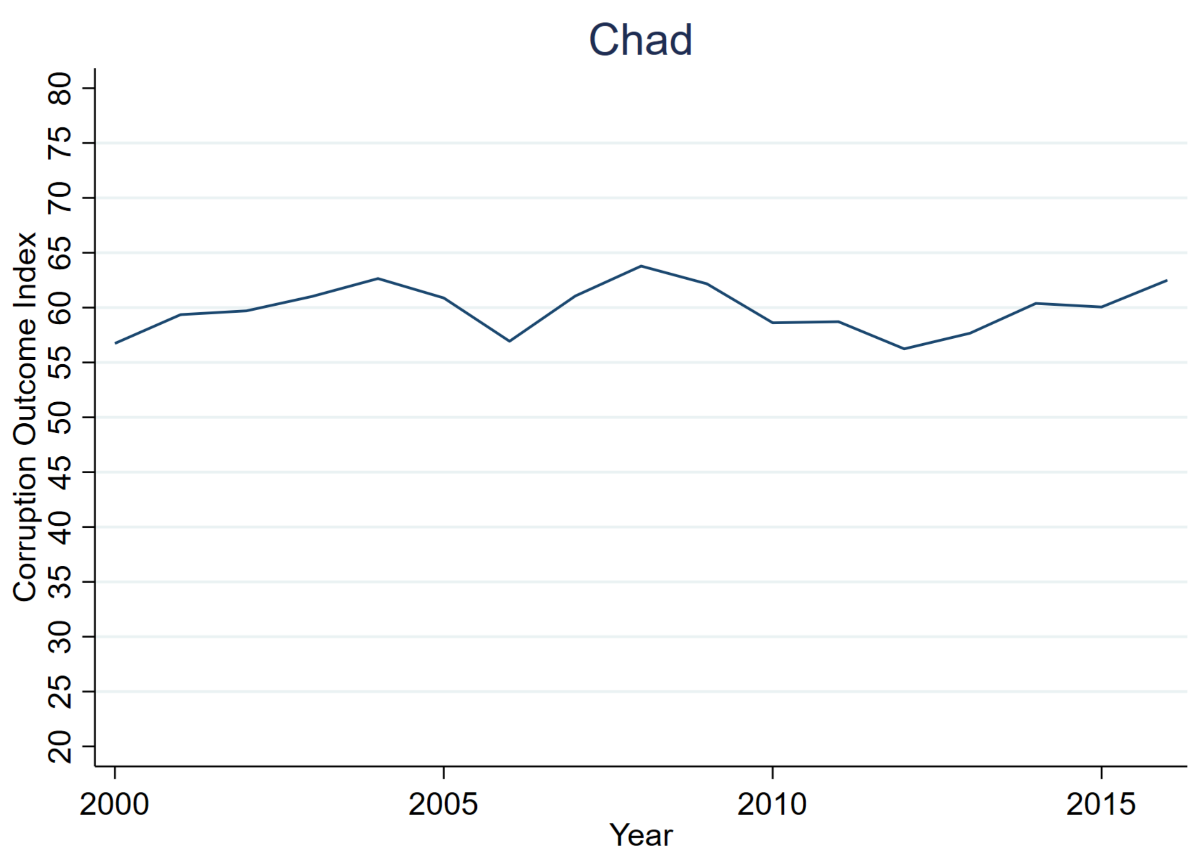

Cambodia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

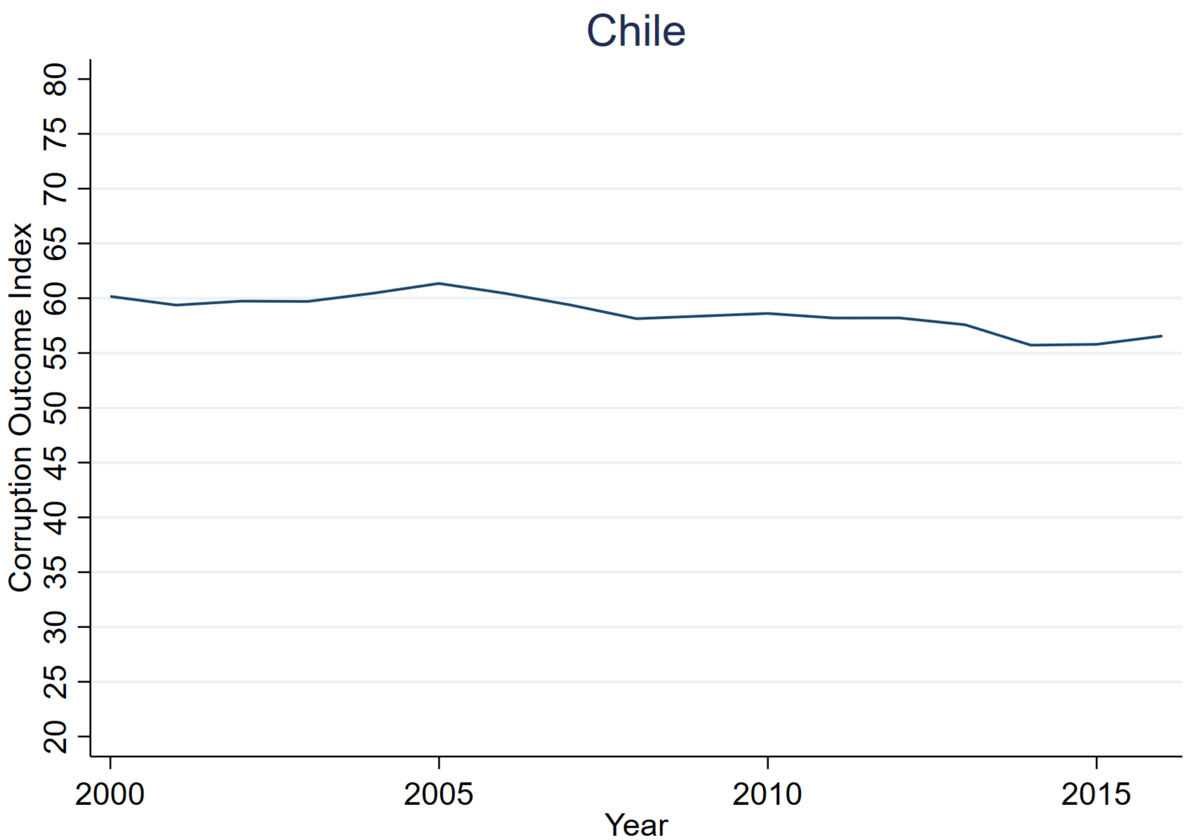

Cameroon
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

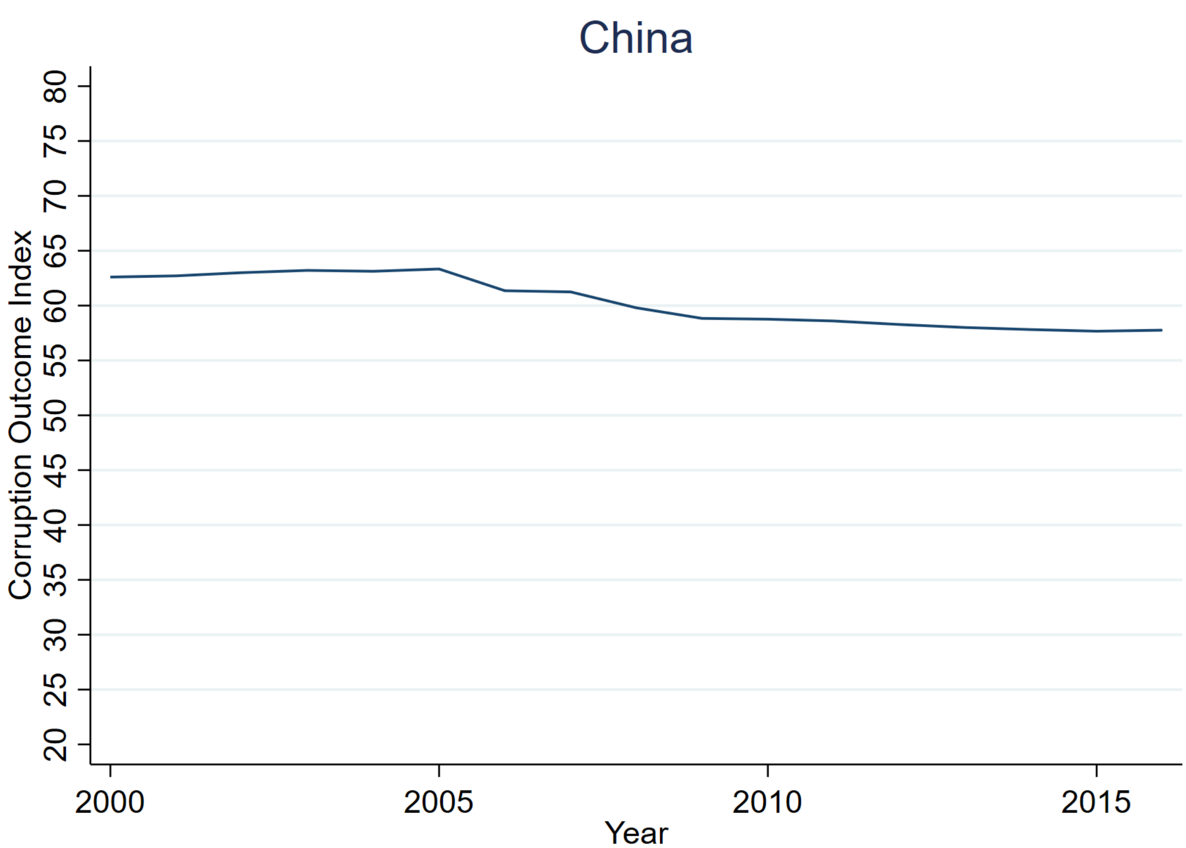

Canada
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

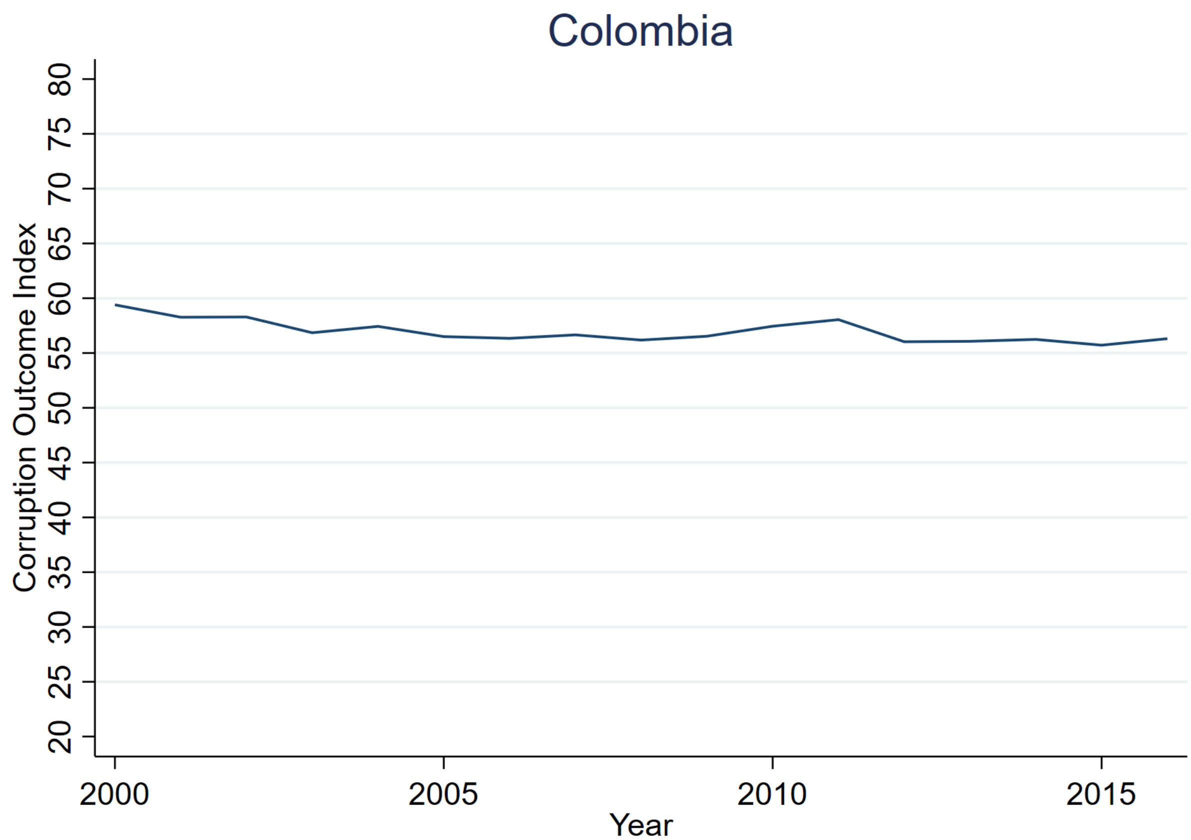

Cape Verde
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

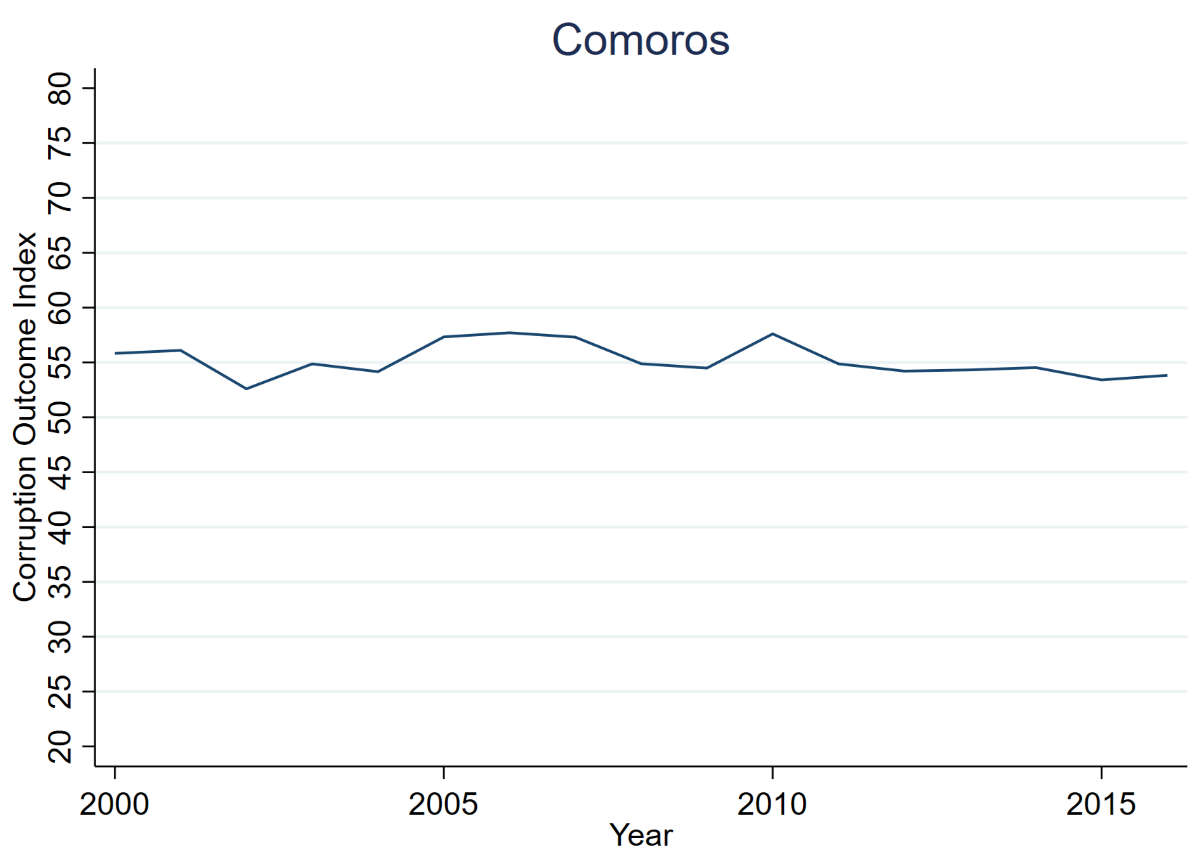

Central African Republic
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

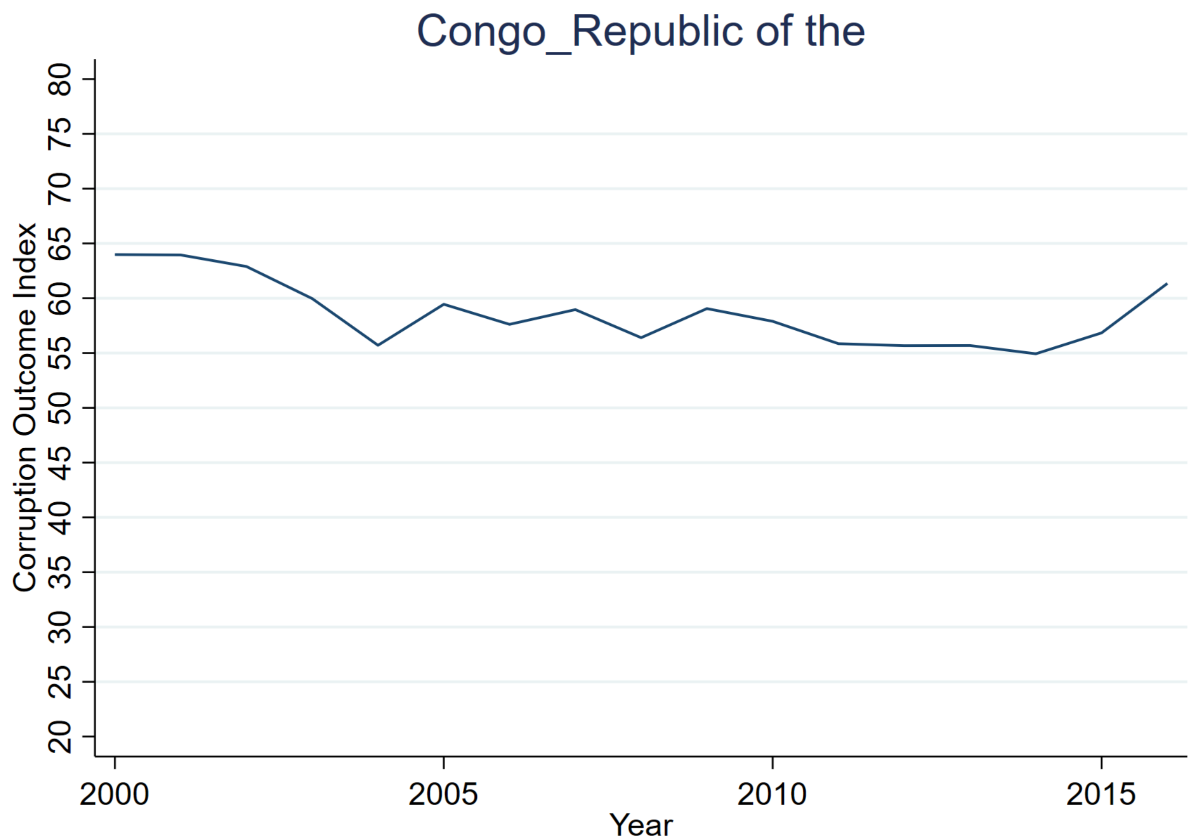

Chad
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

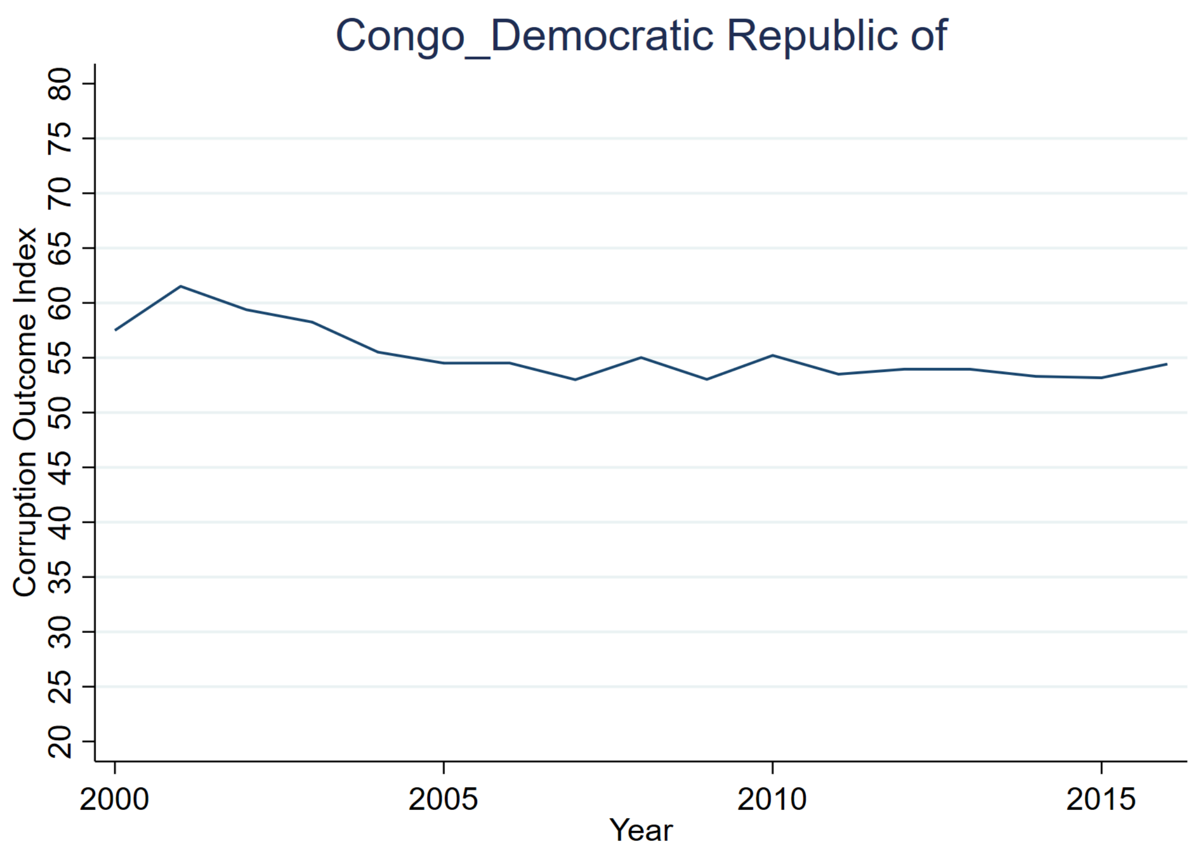

Chile
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

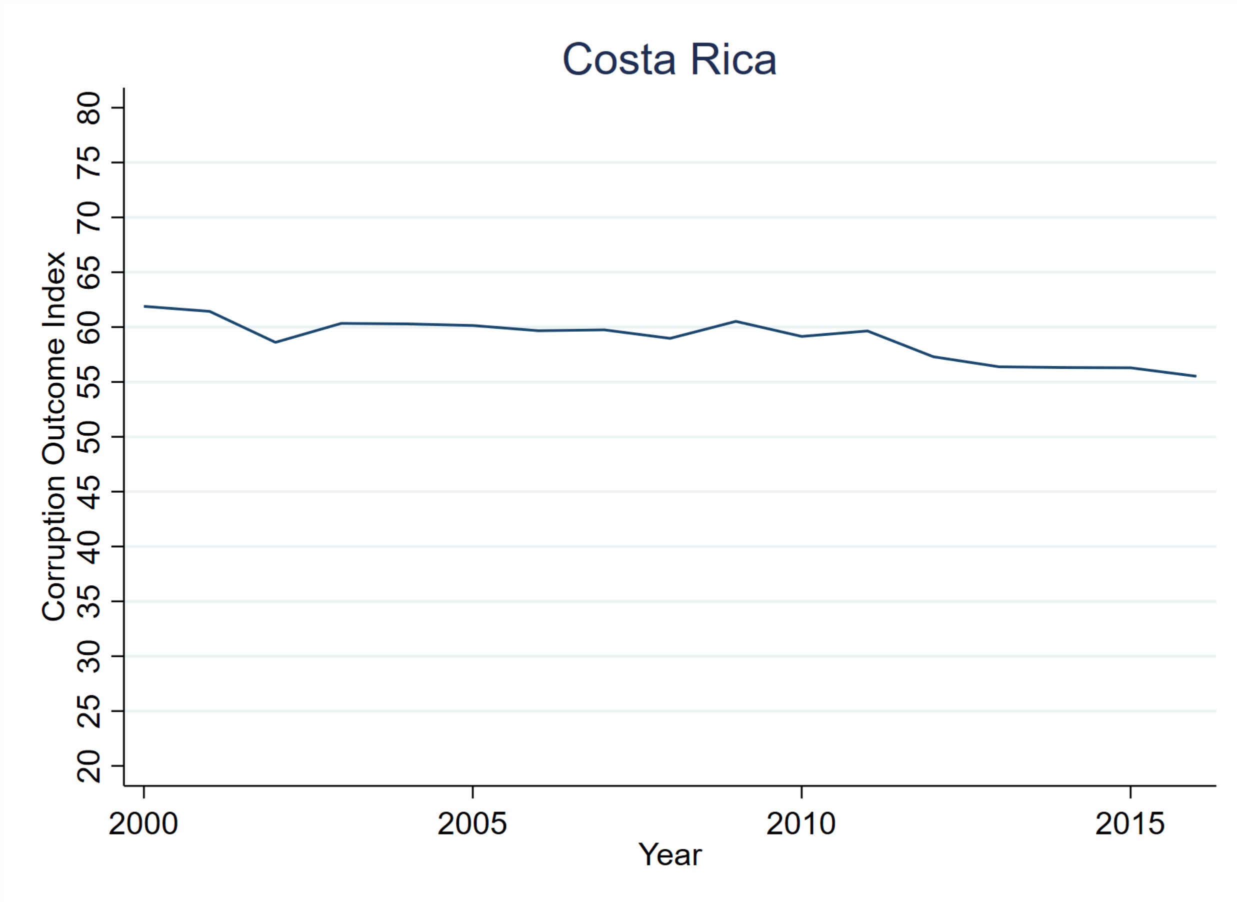

China
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

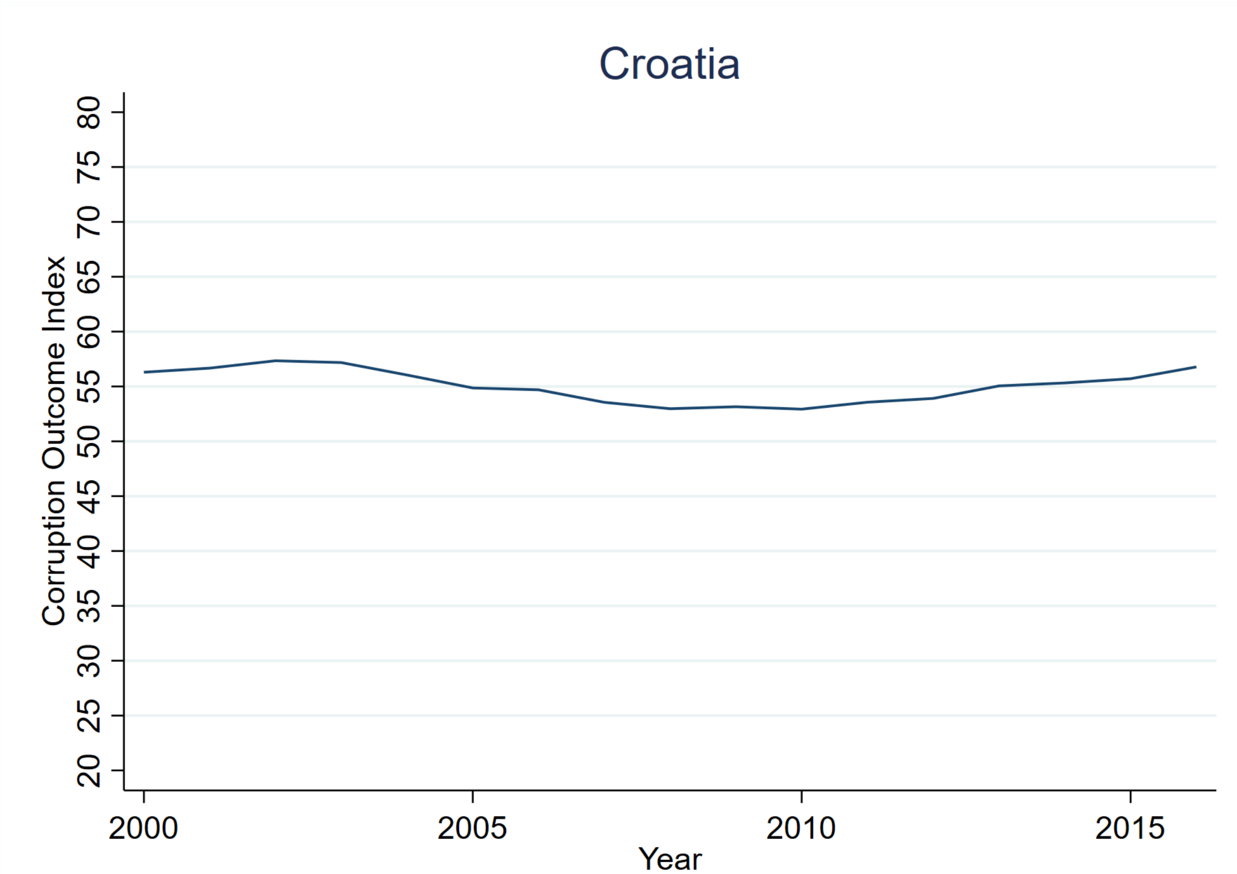

Colombia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

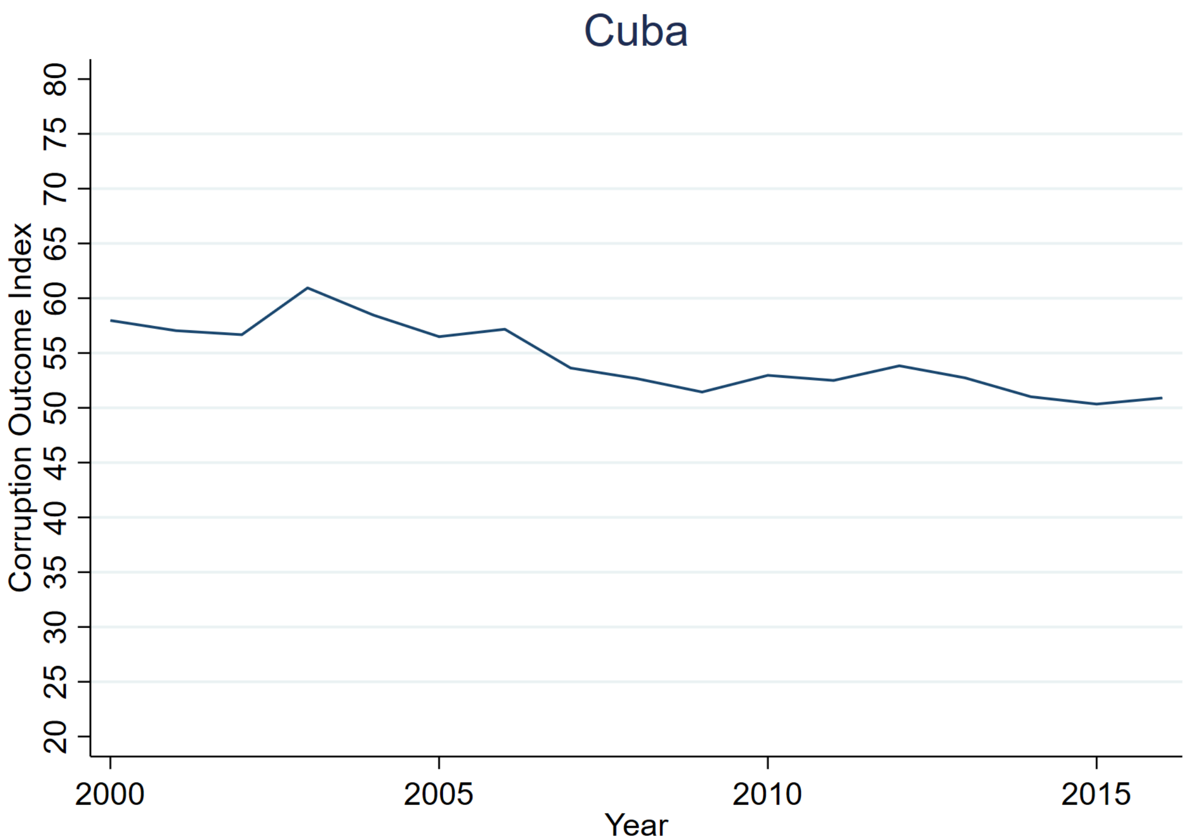

Comoros
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

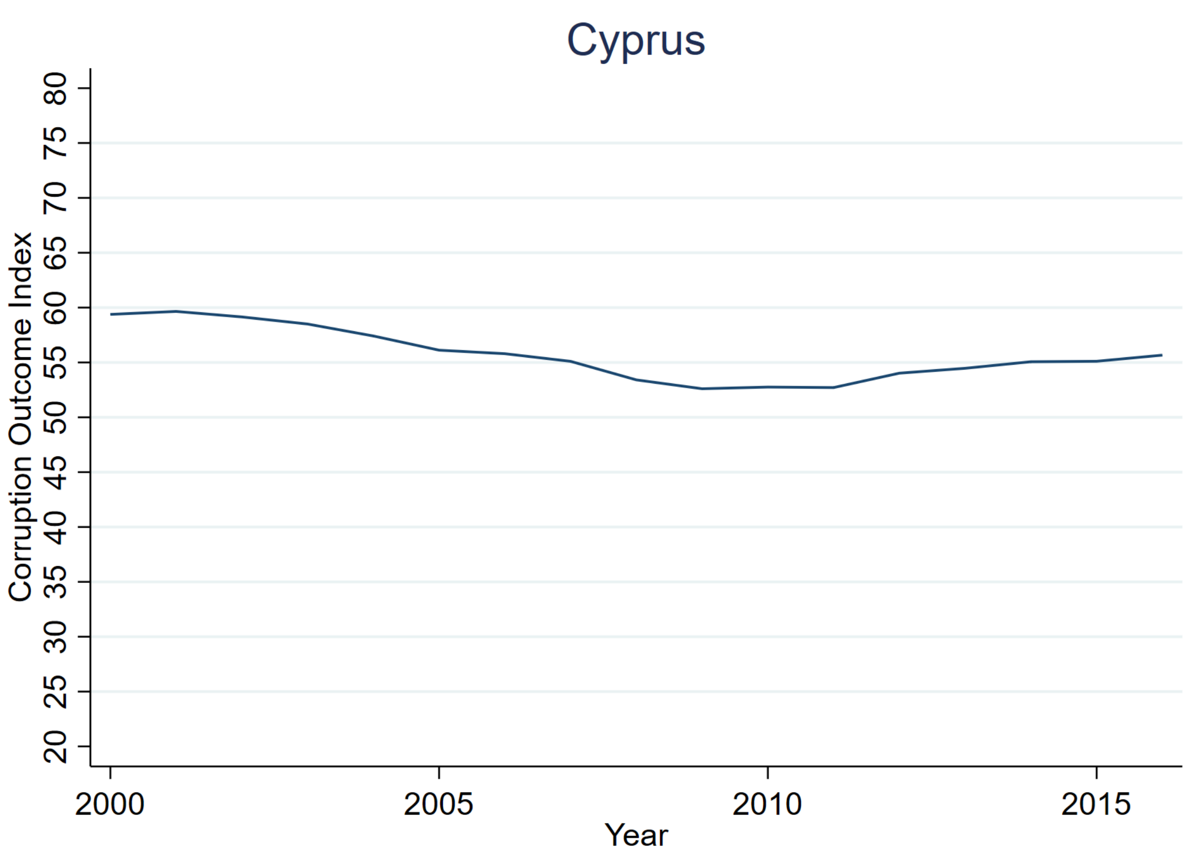

Congo_Republic of the
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

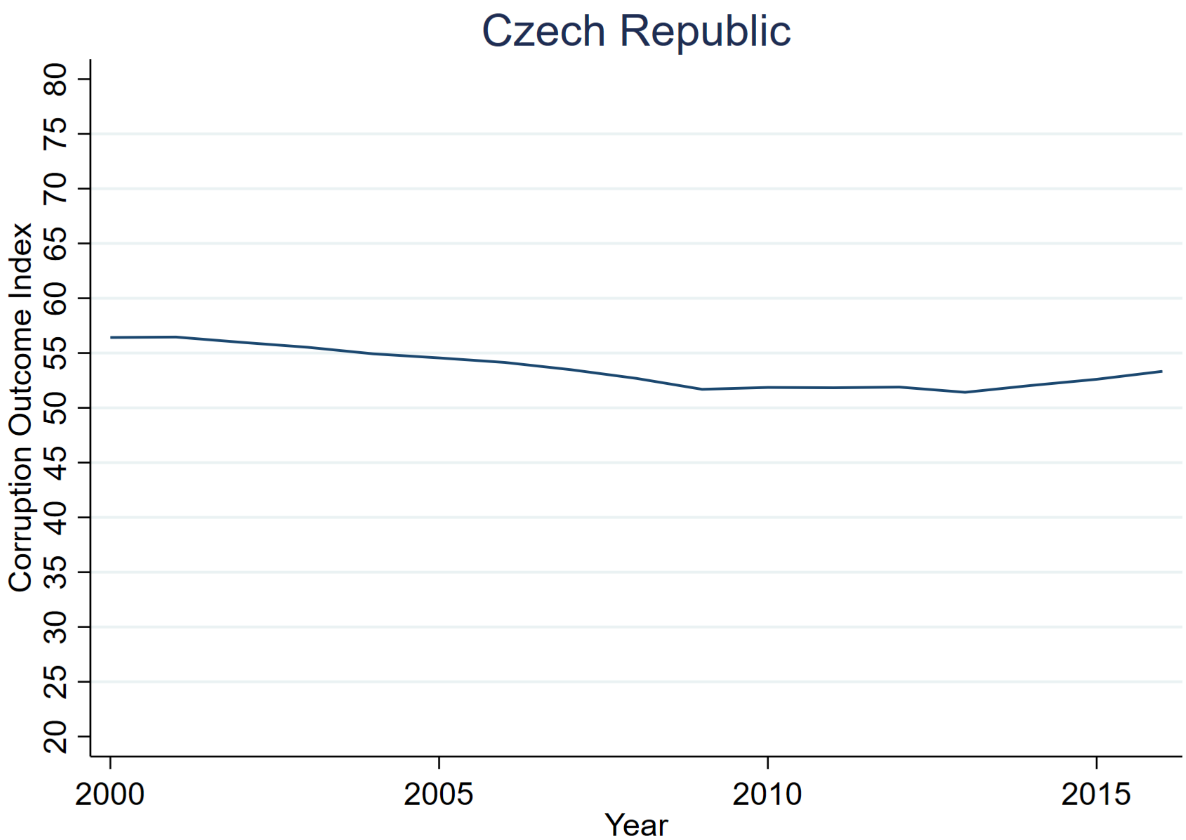

Congo_Democratic Republic of
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Costa Rica
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

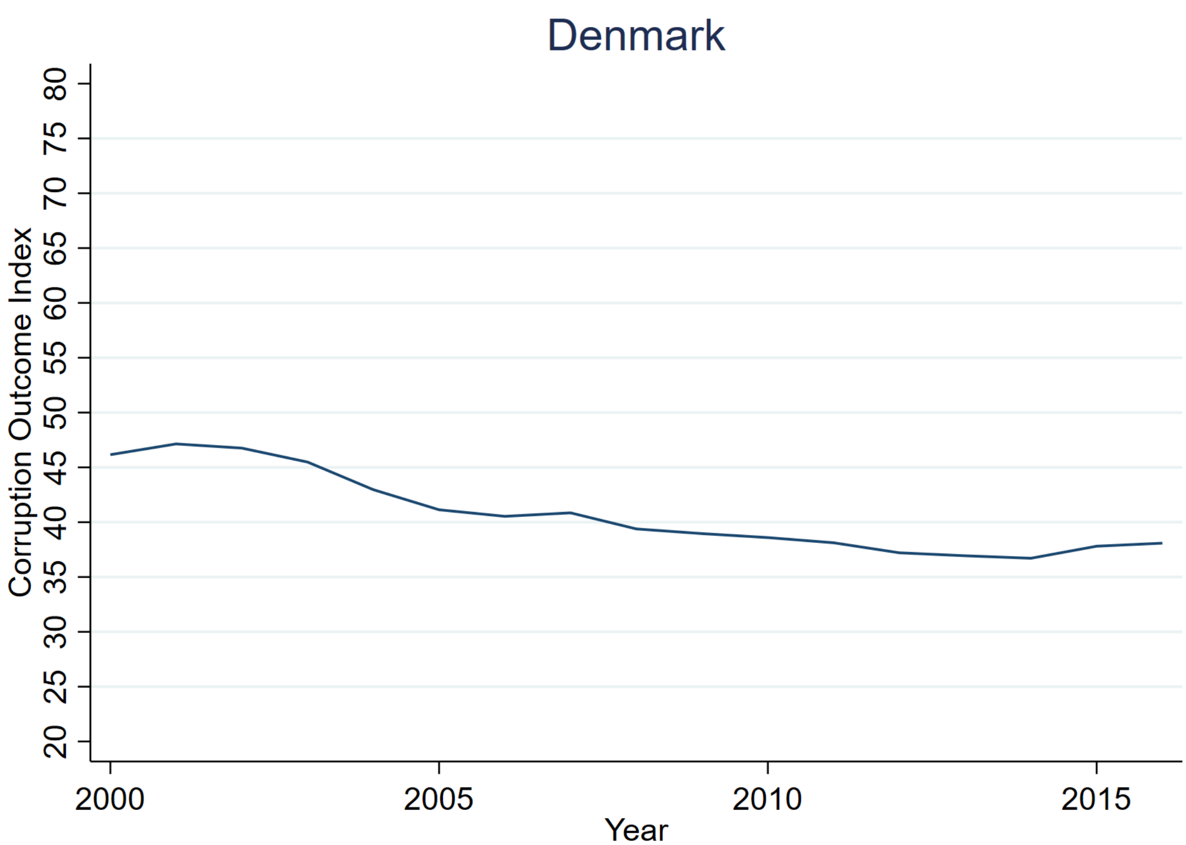

Croatia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Cuba
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

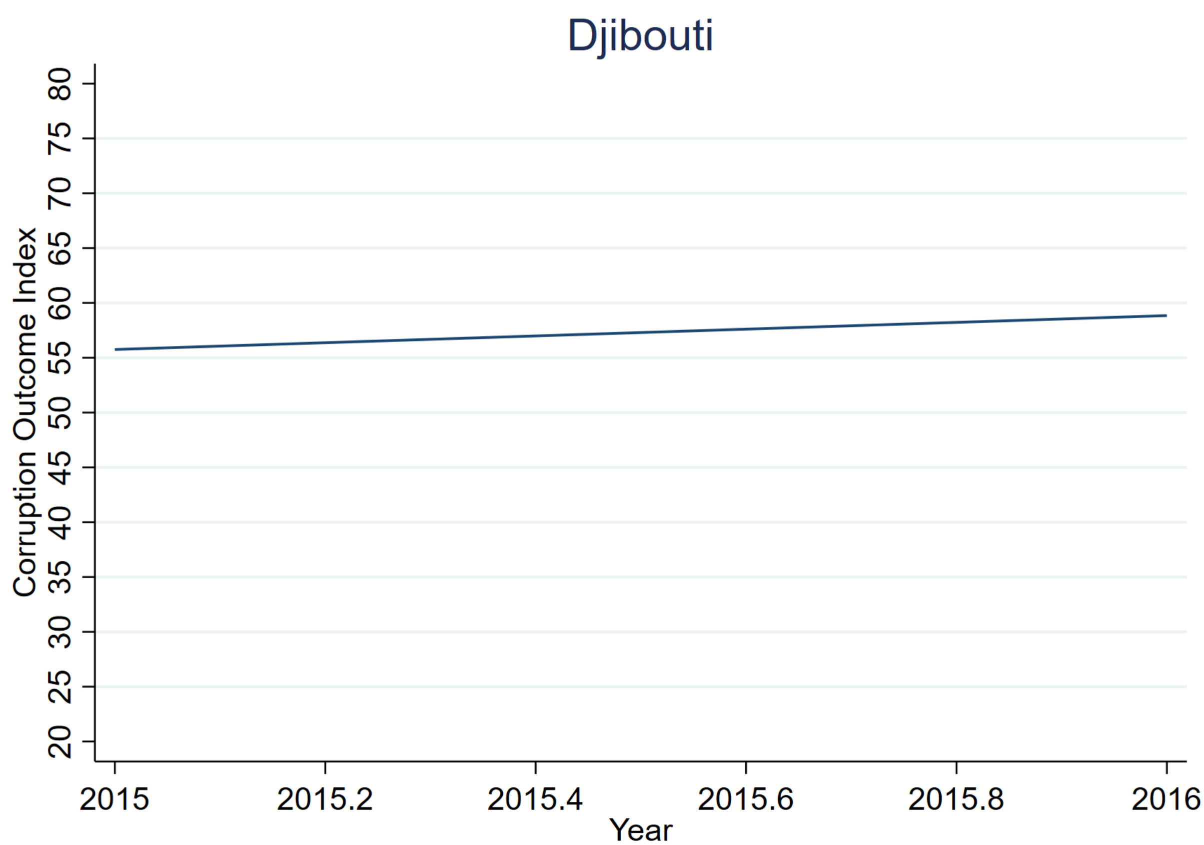

Cyprus
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

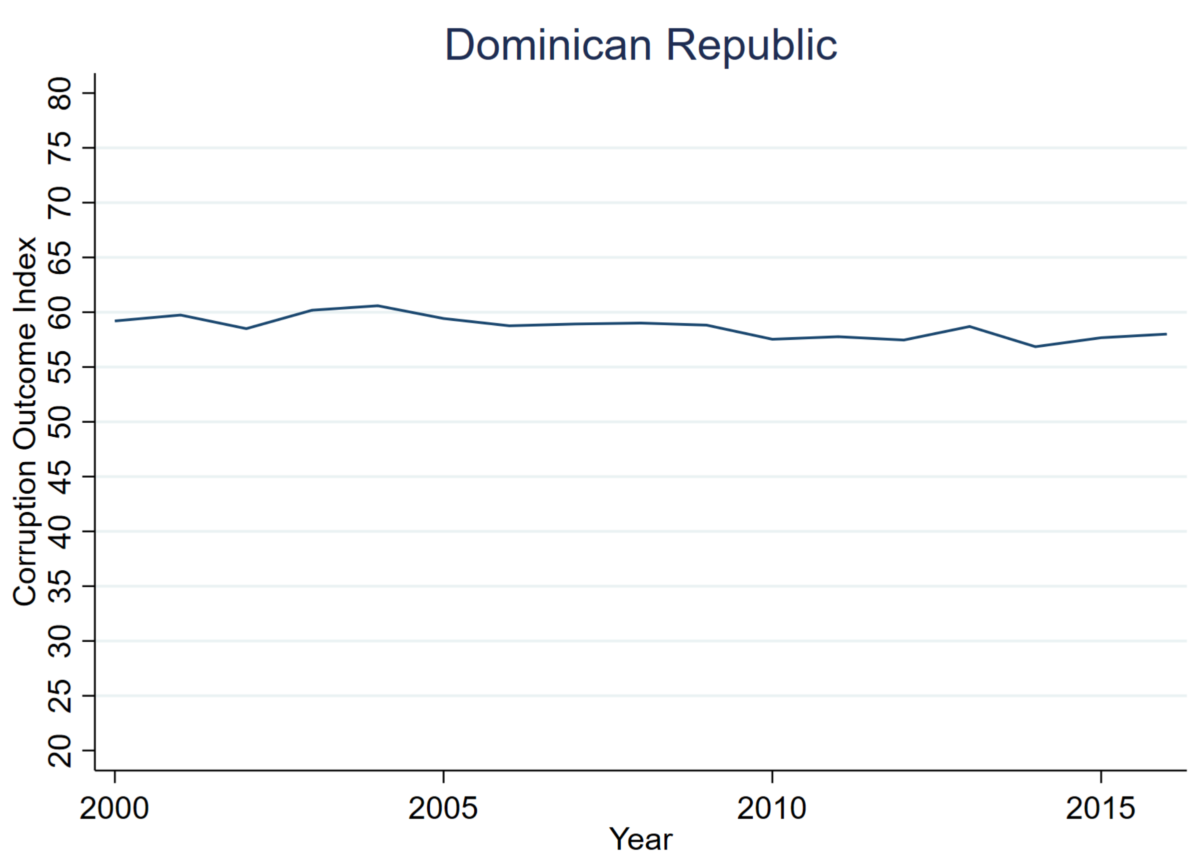


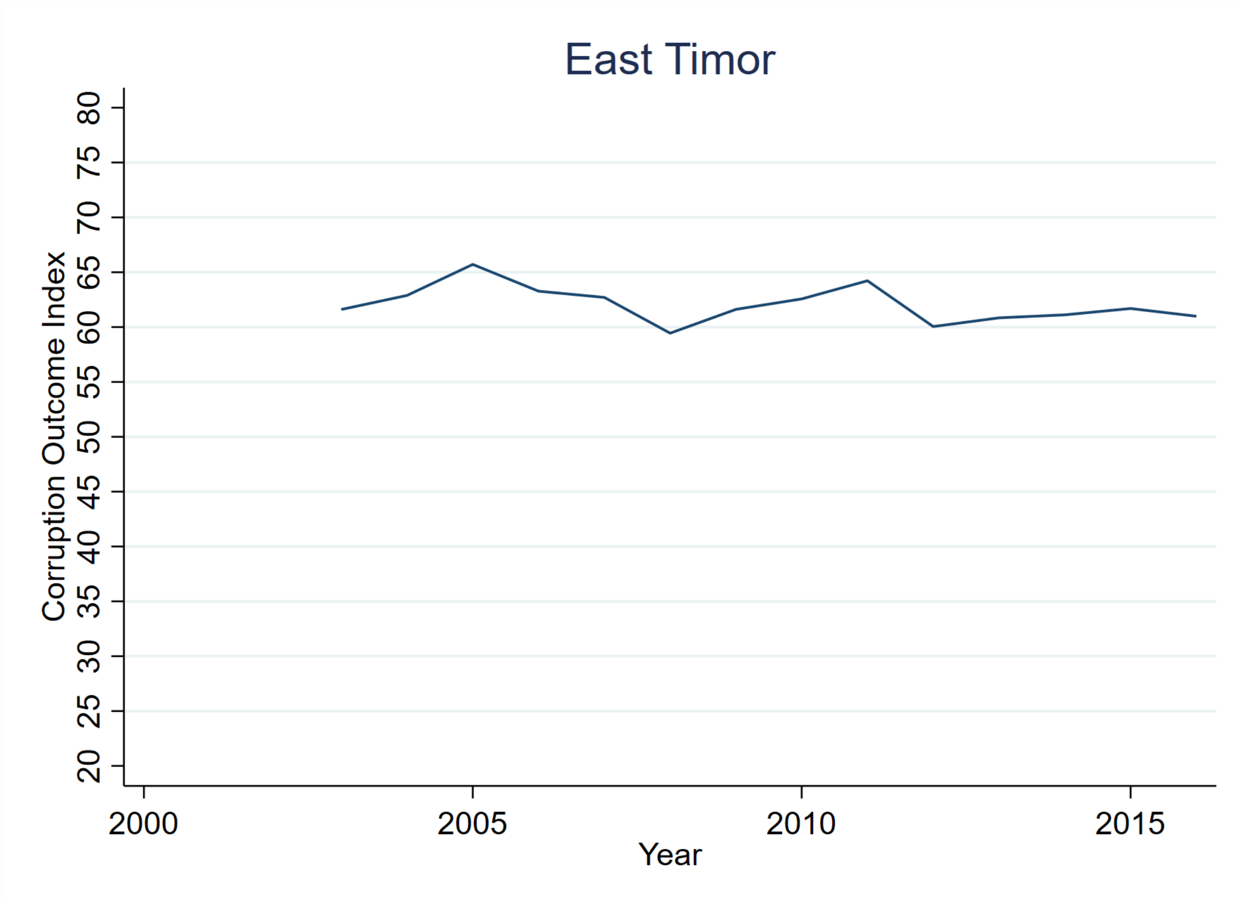
Czech Republic
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

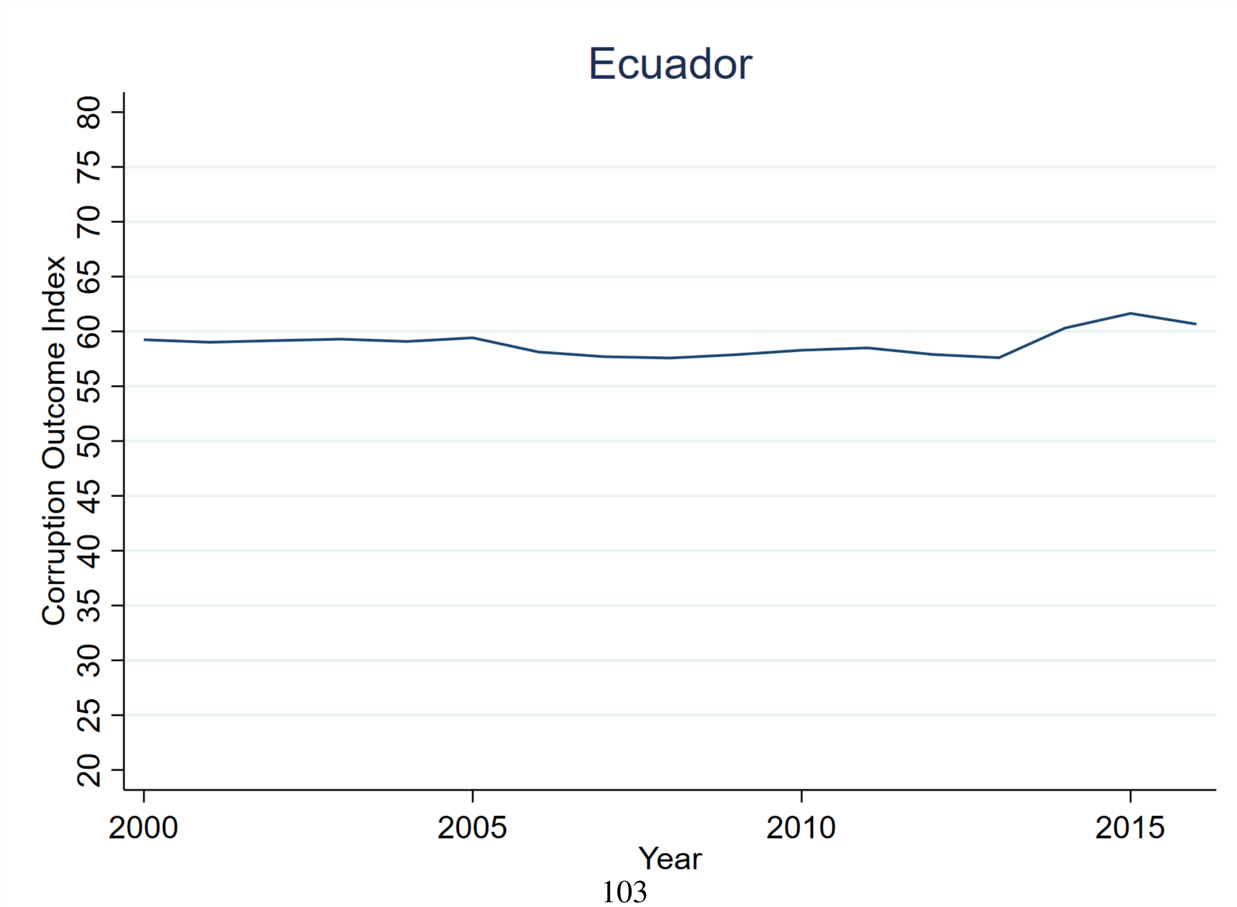


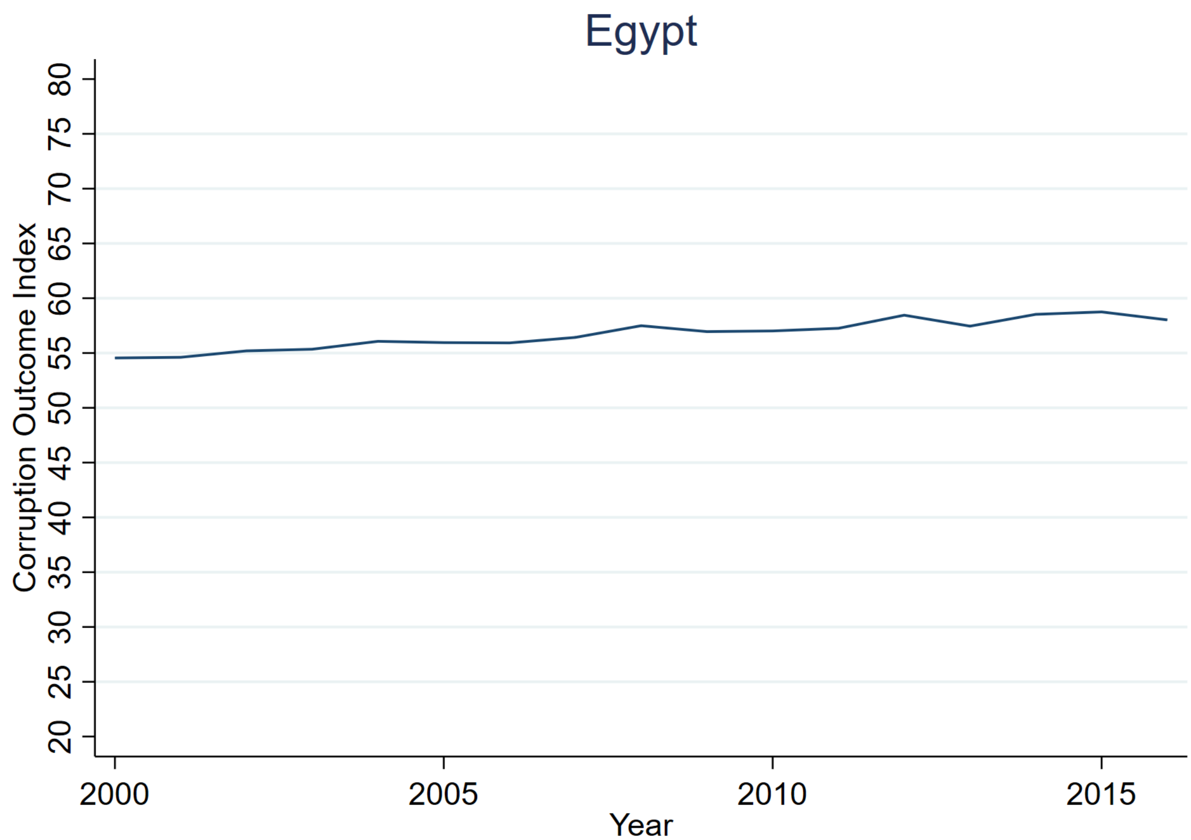
Denmark
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

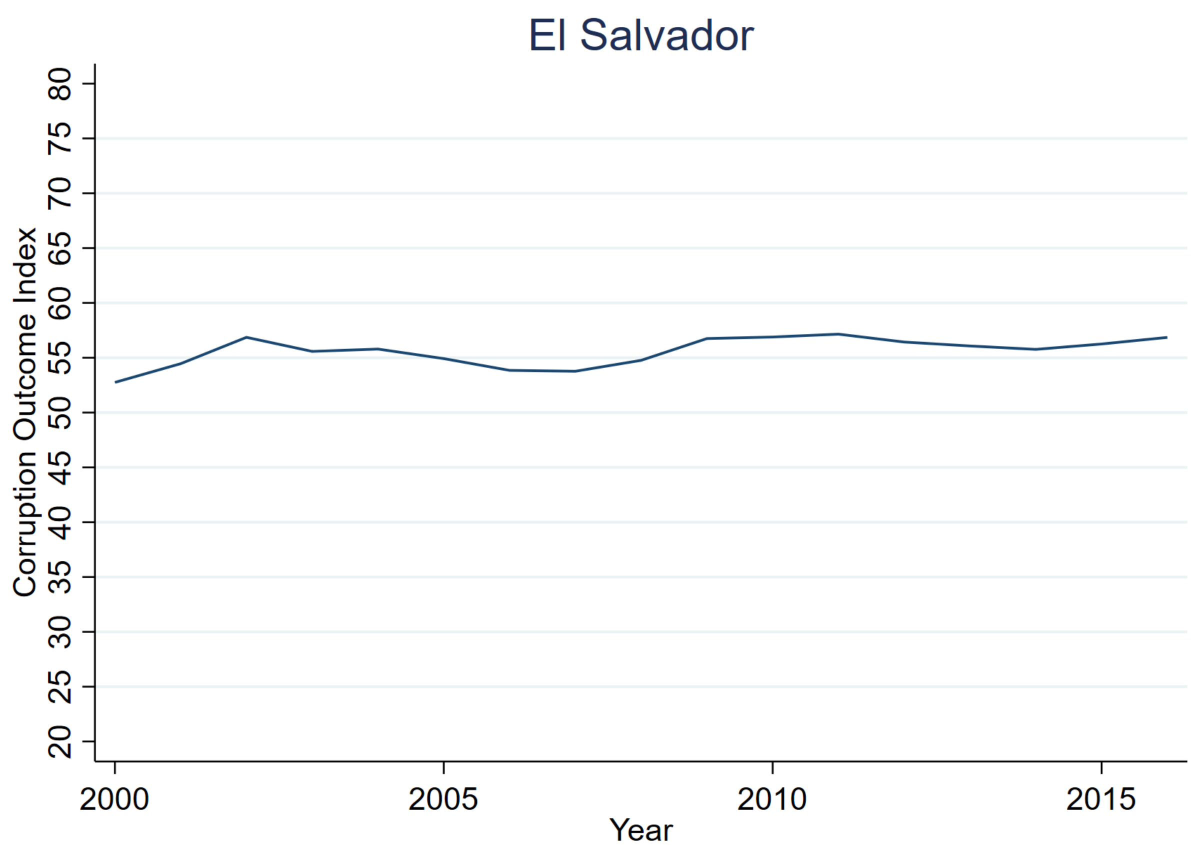

Djibouti
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2015
2015.2
2015.4
2015.6
2015.8
2016
Year

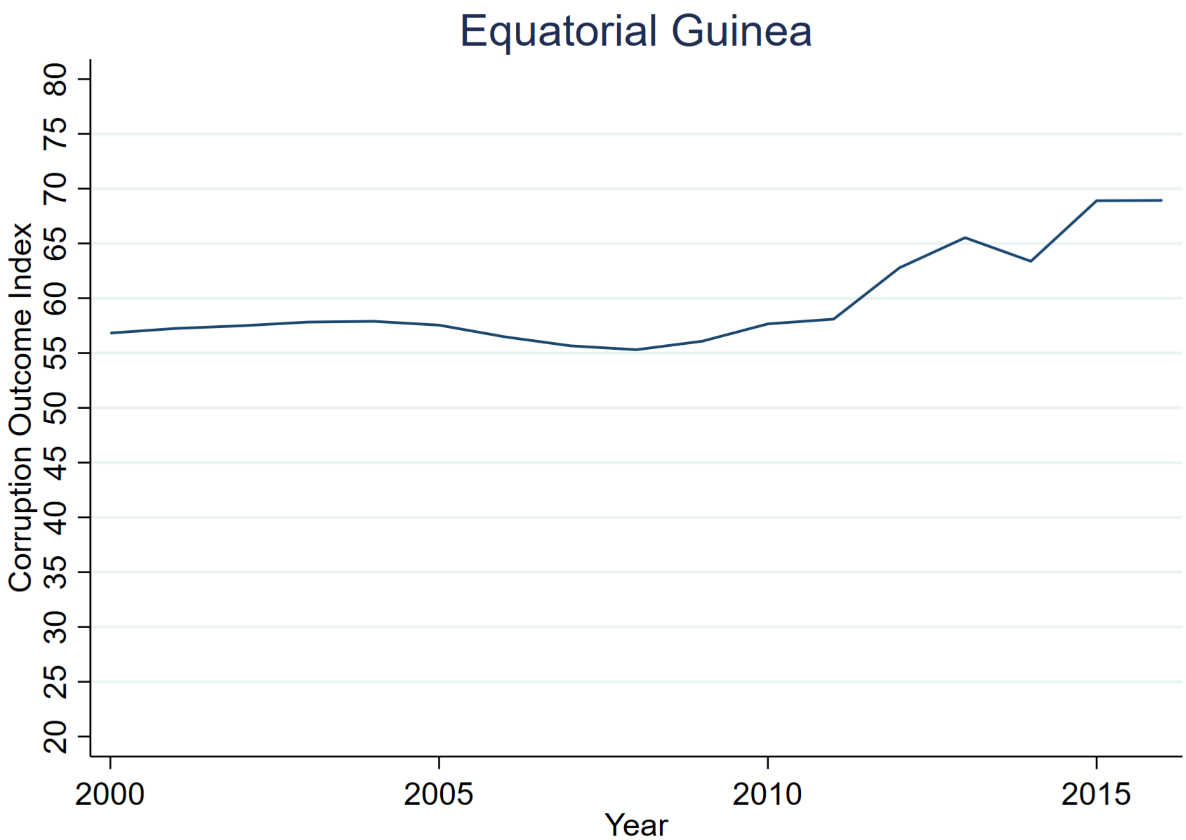

Dominican Republic
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

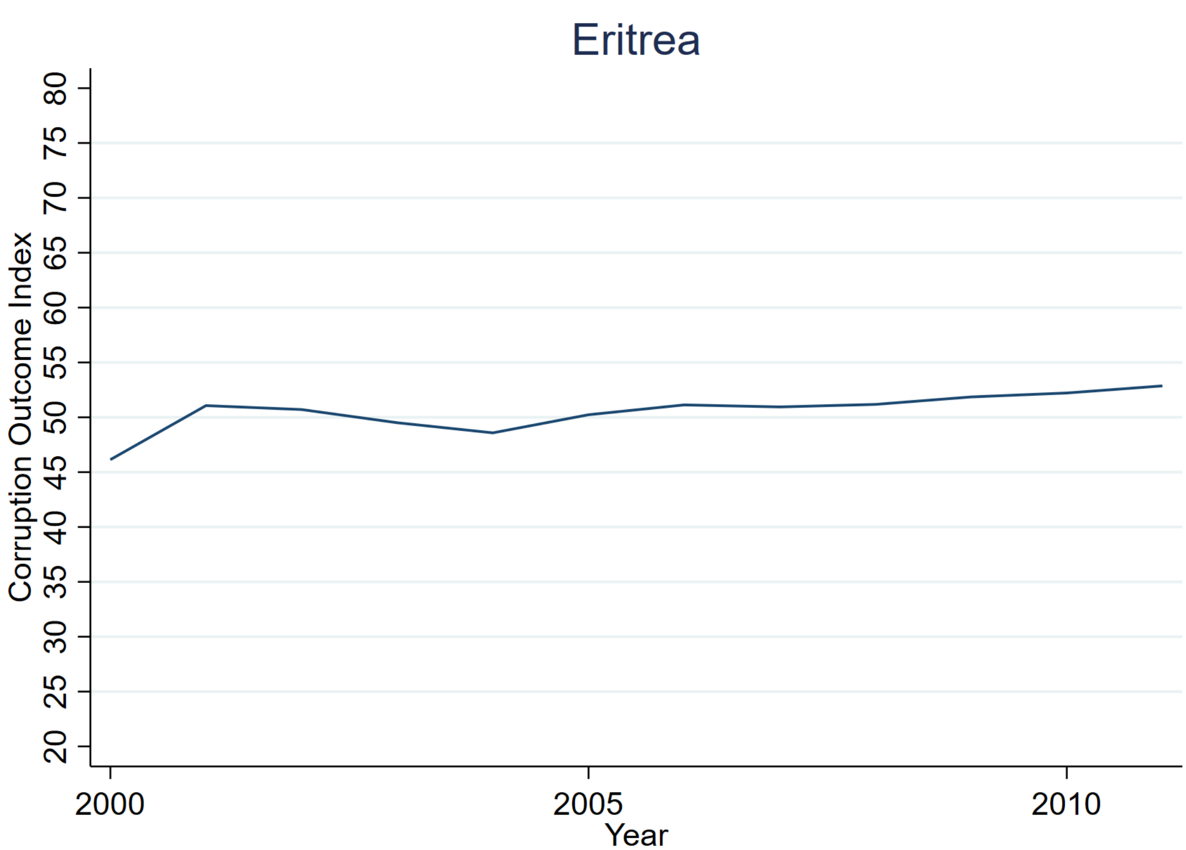

East Timor
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Ecuador
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

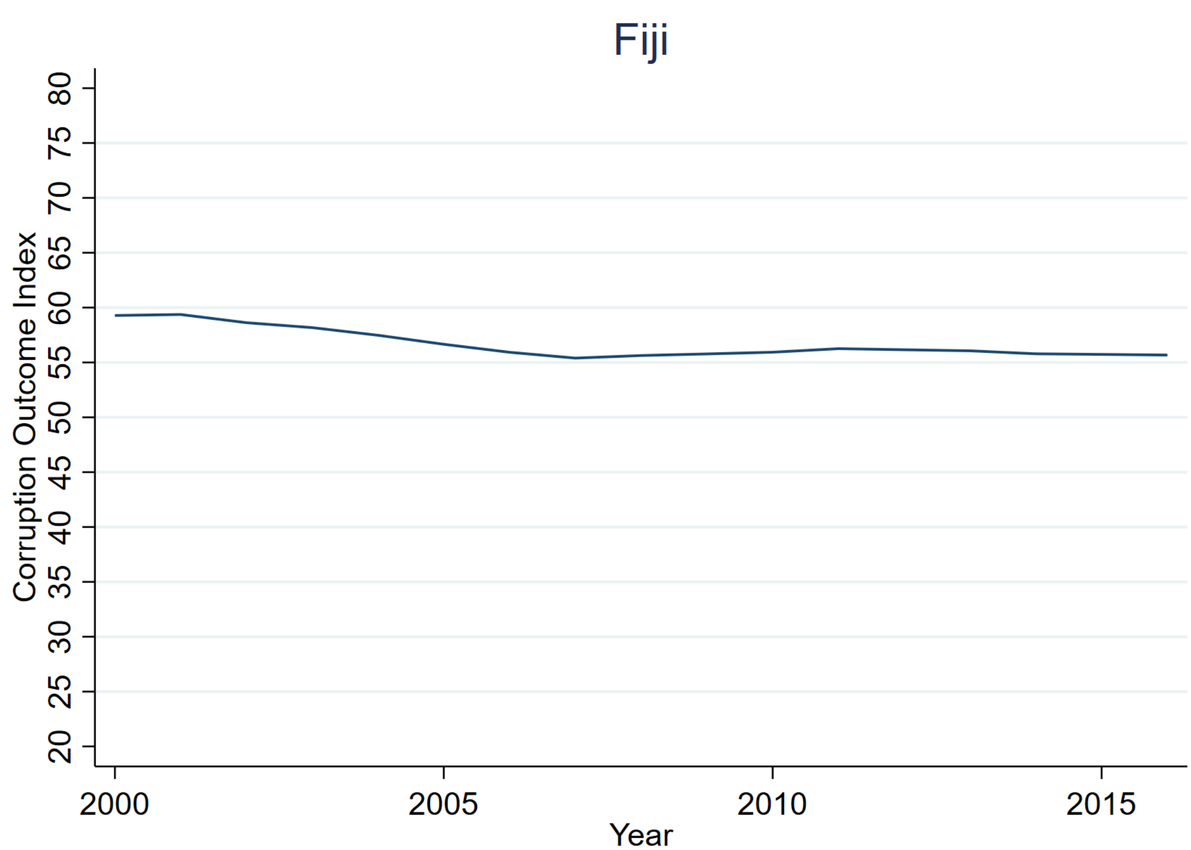

Egypt
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

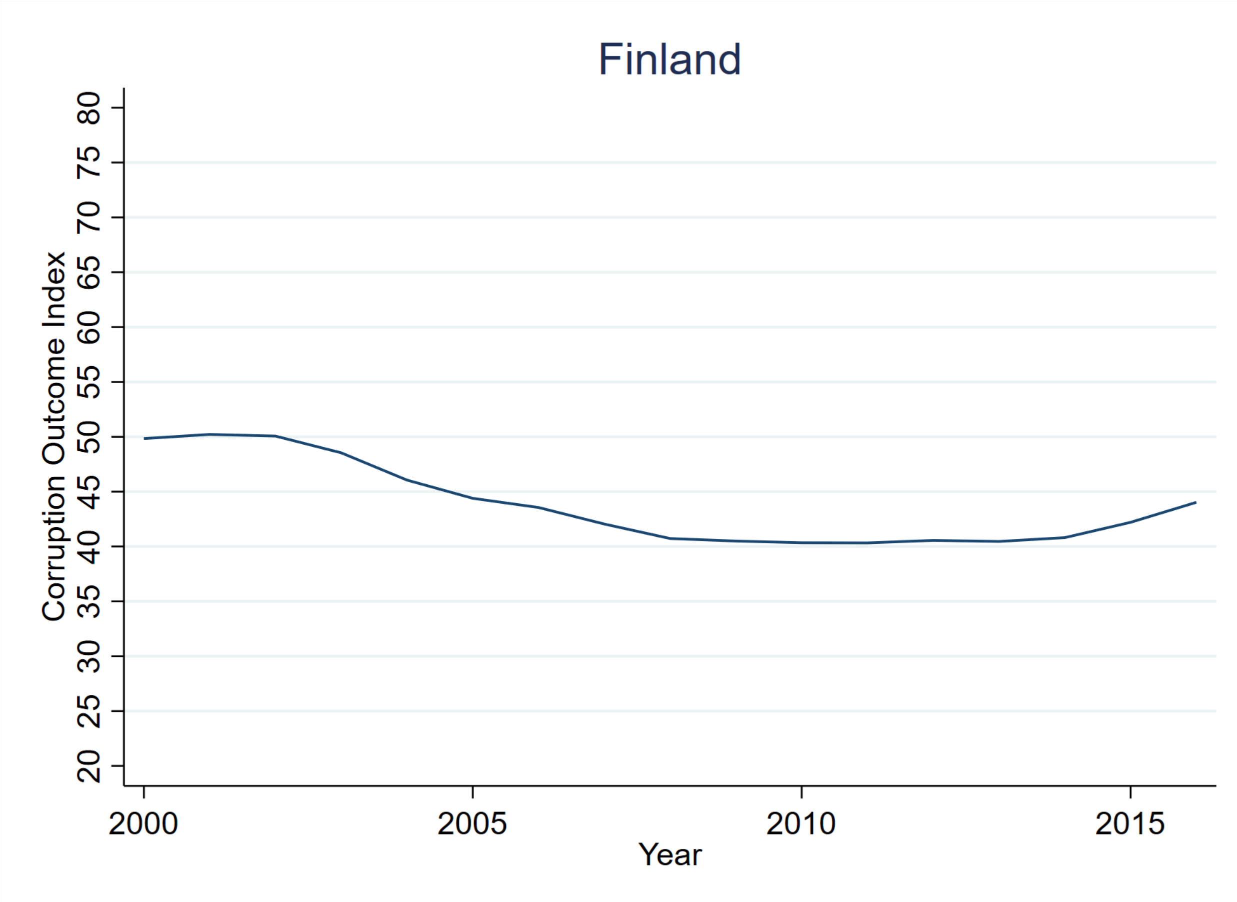

El Salvador
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

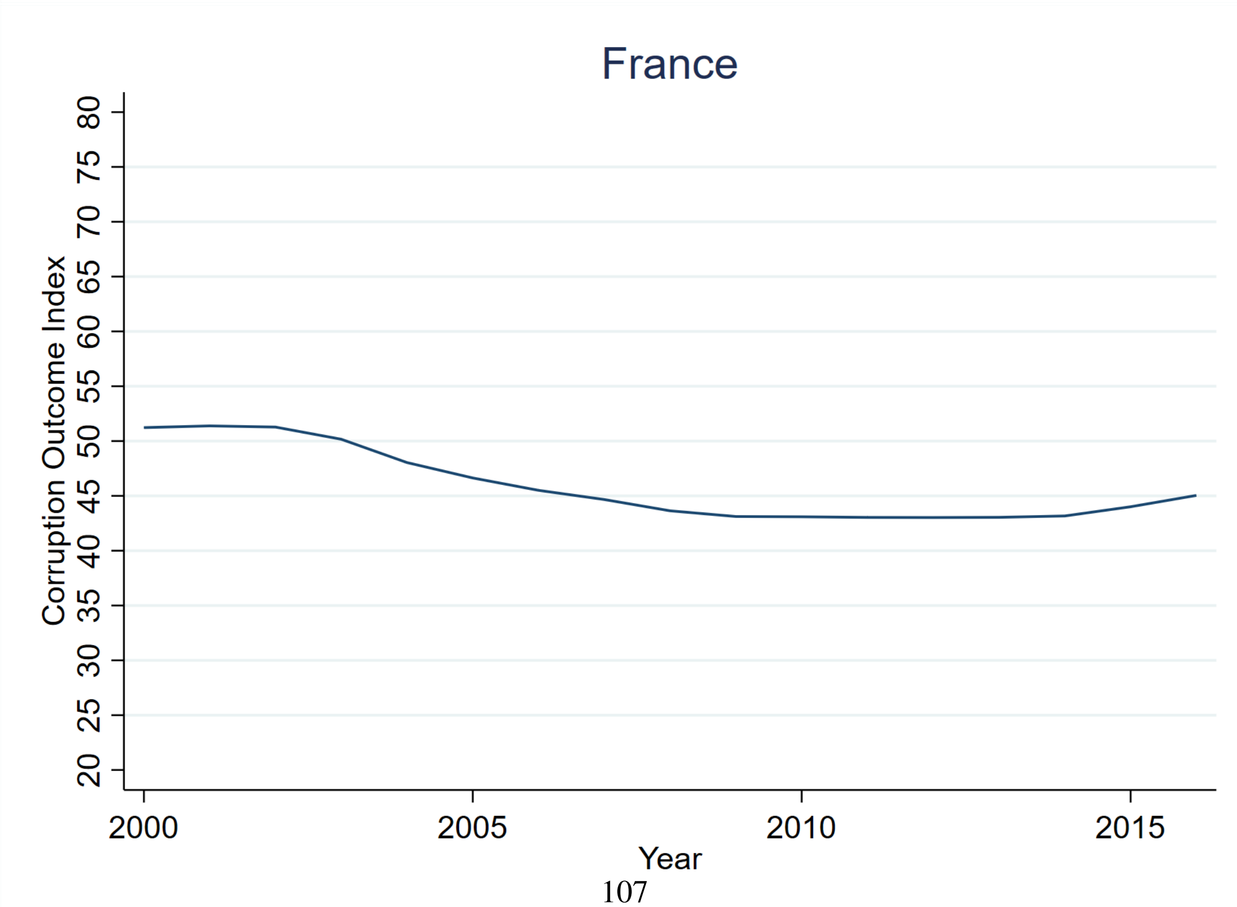

Equatorial Guinea
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

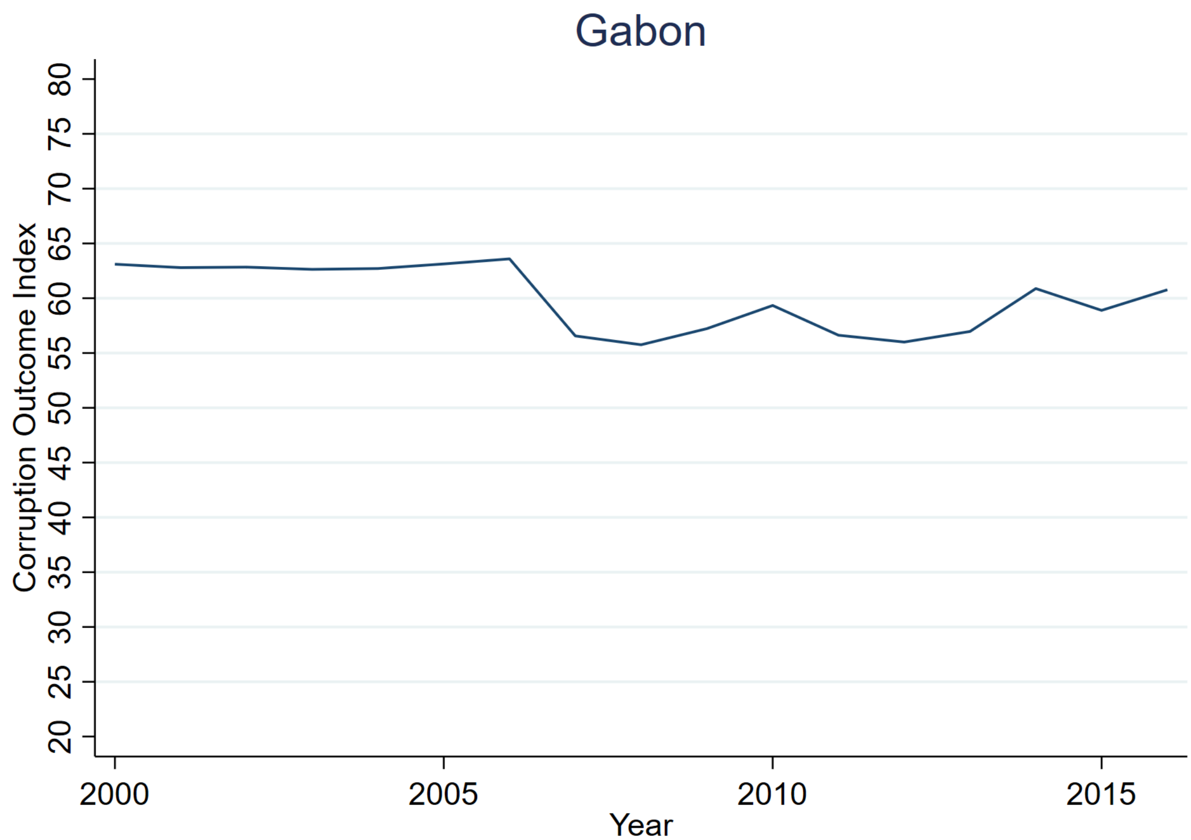

Eritrea
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
Year

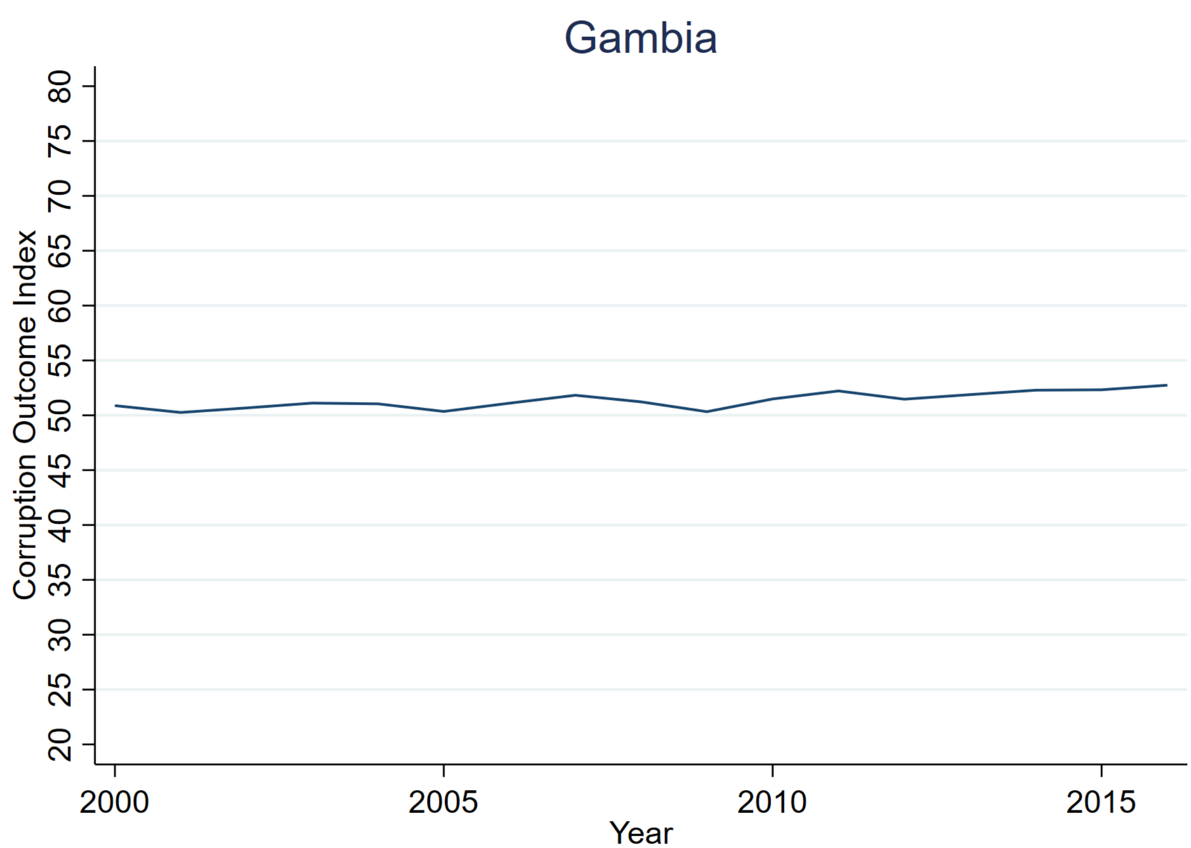

Estonia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

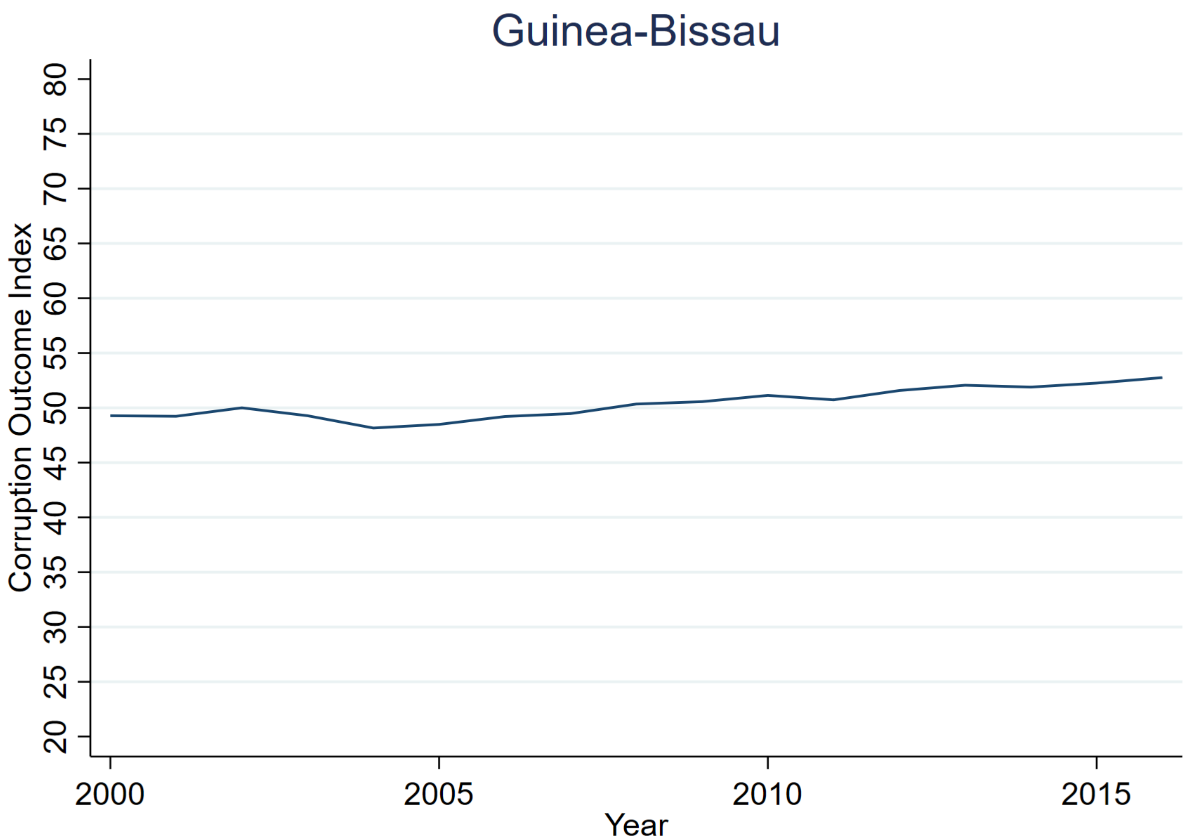

Fiji
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

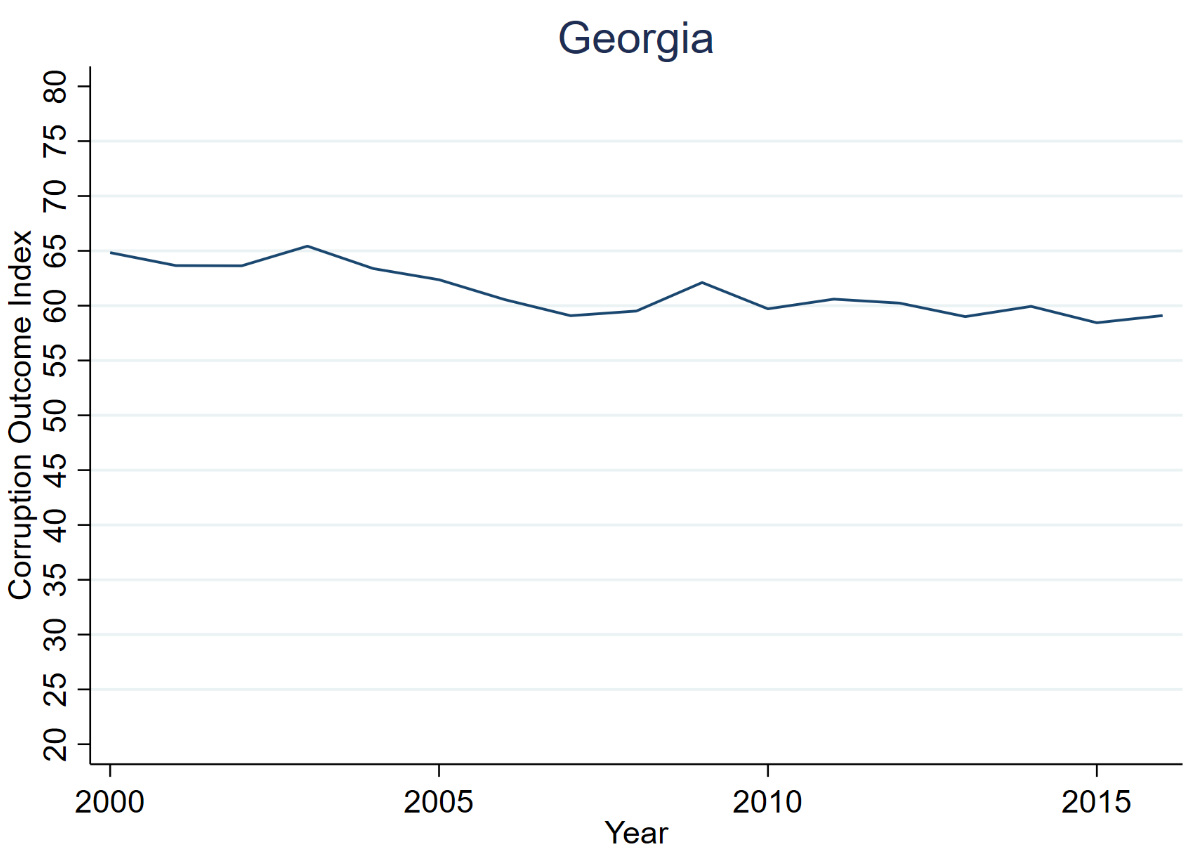

Finland
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

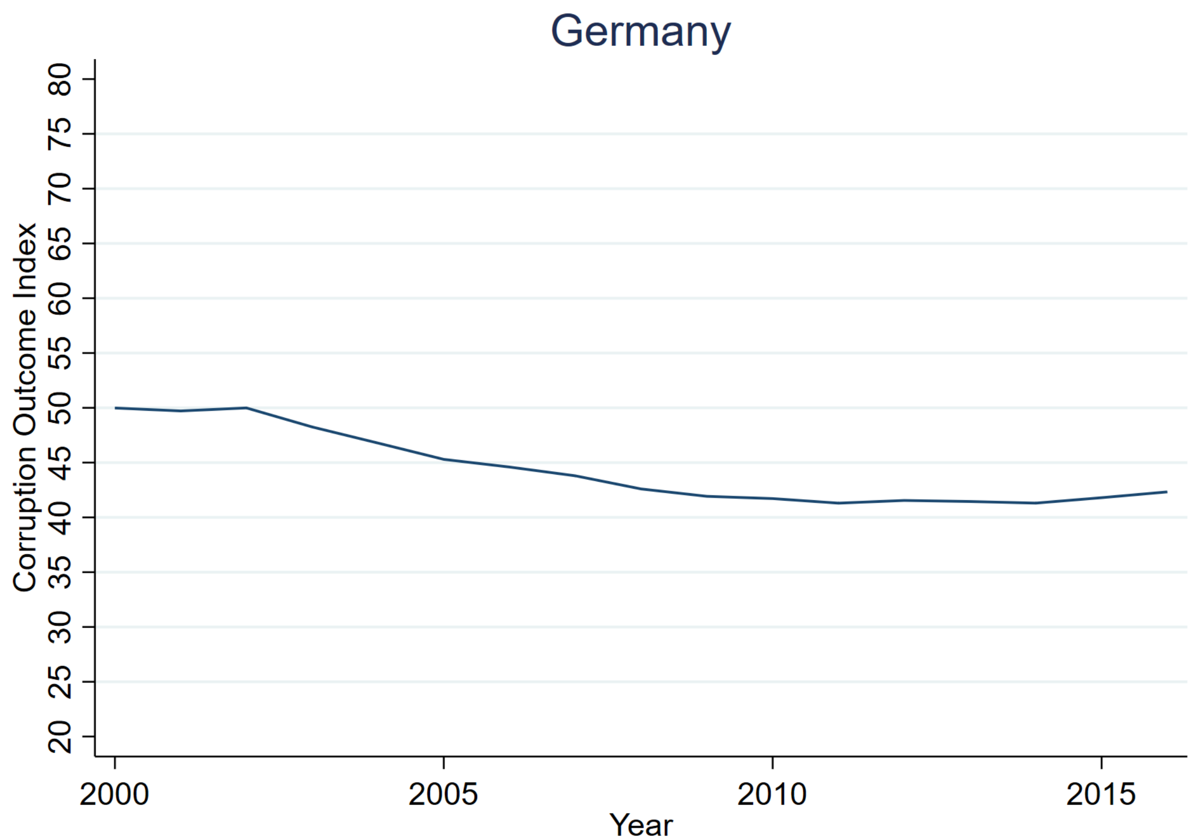

France
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

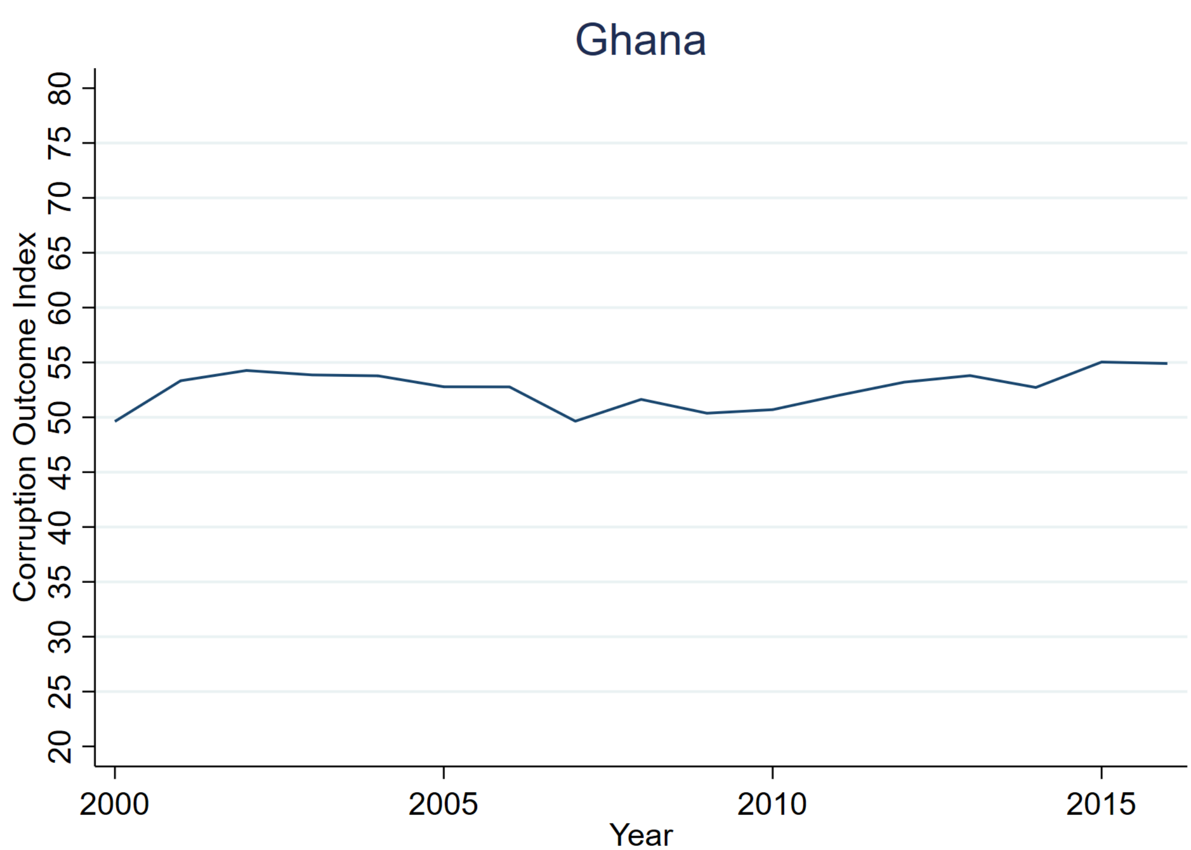

Gabon
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

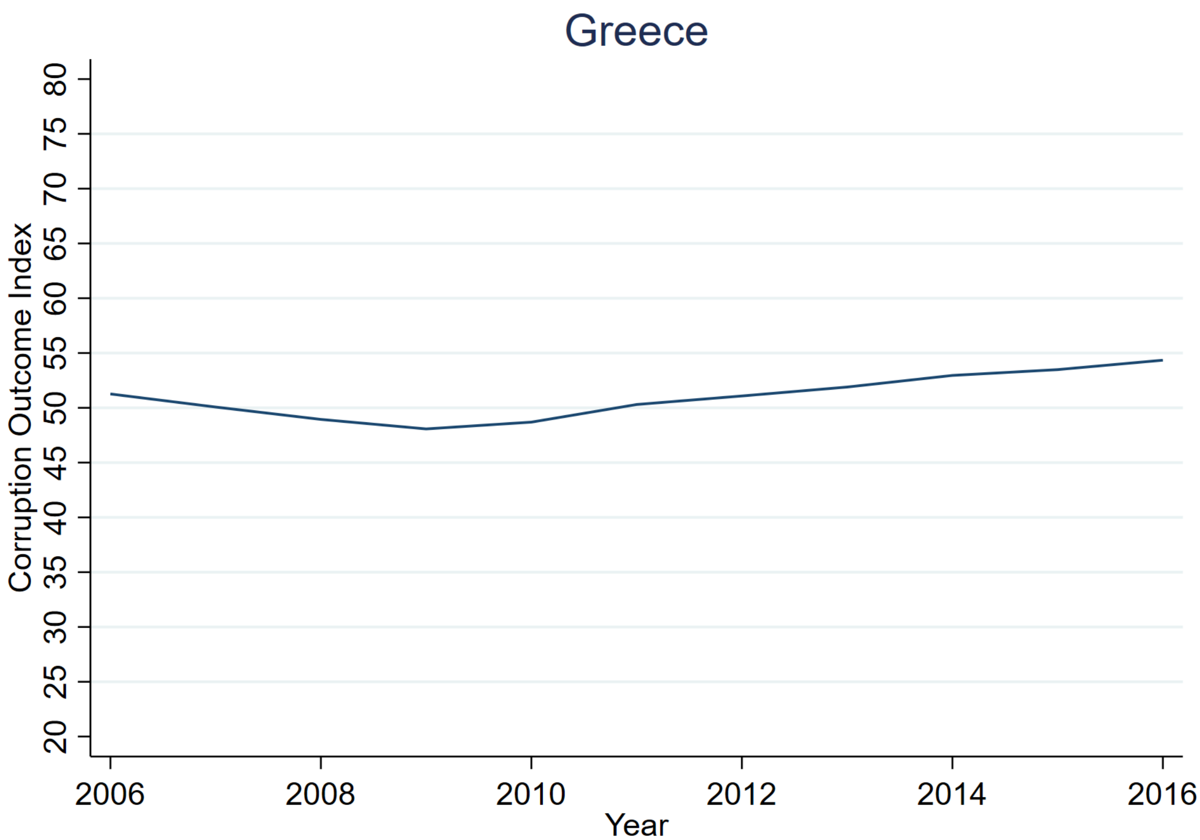

Gambia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Guinea-Bissau
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

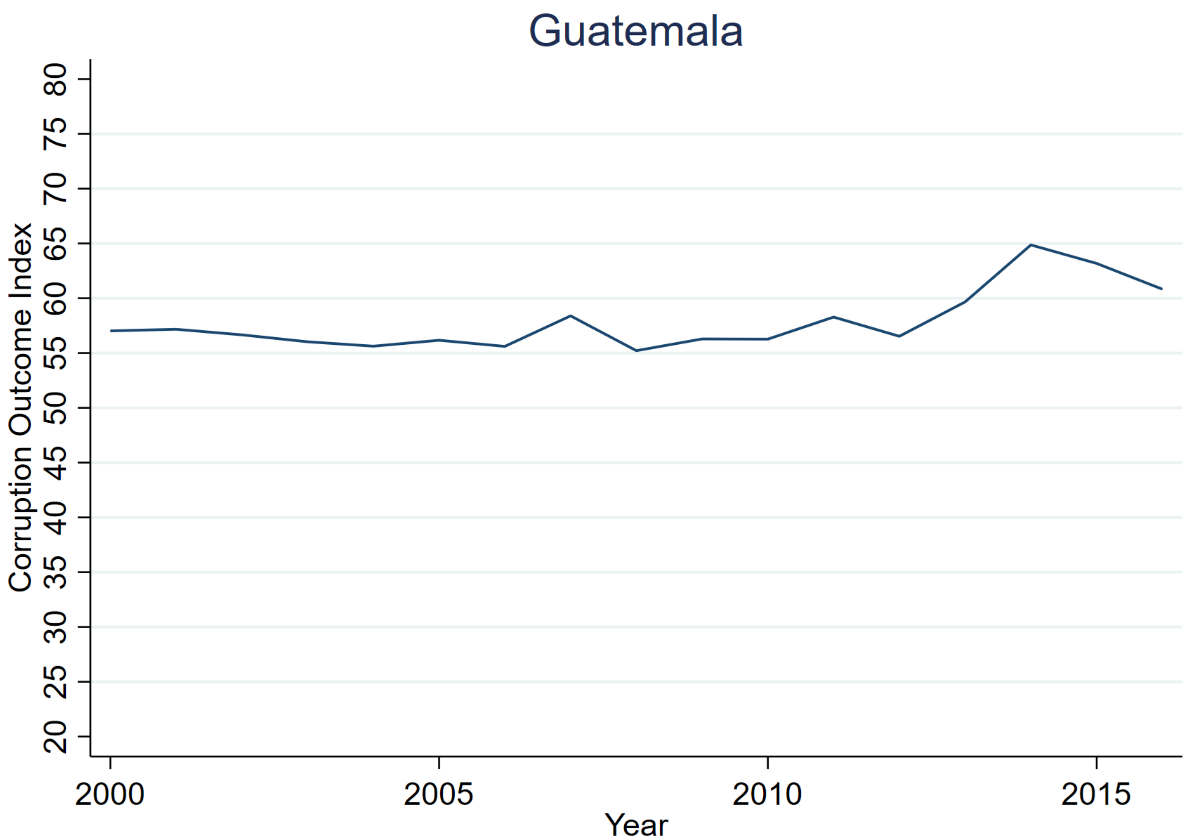

Georgia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Germany
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

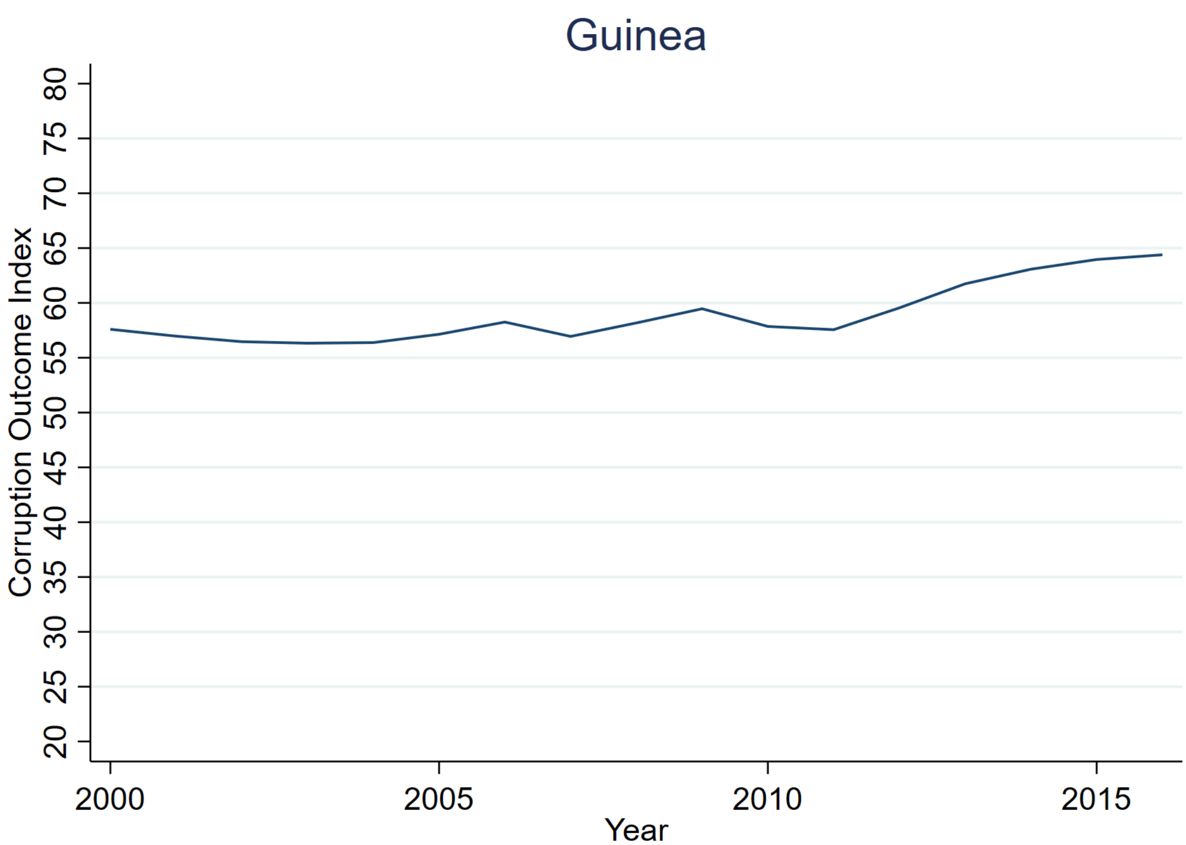

Ghana
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Greece
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2006
2008
2010
2012
2014
2016
Year

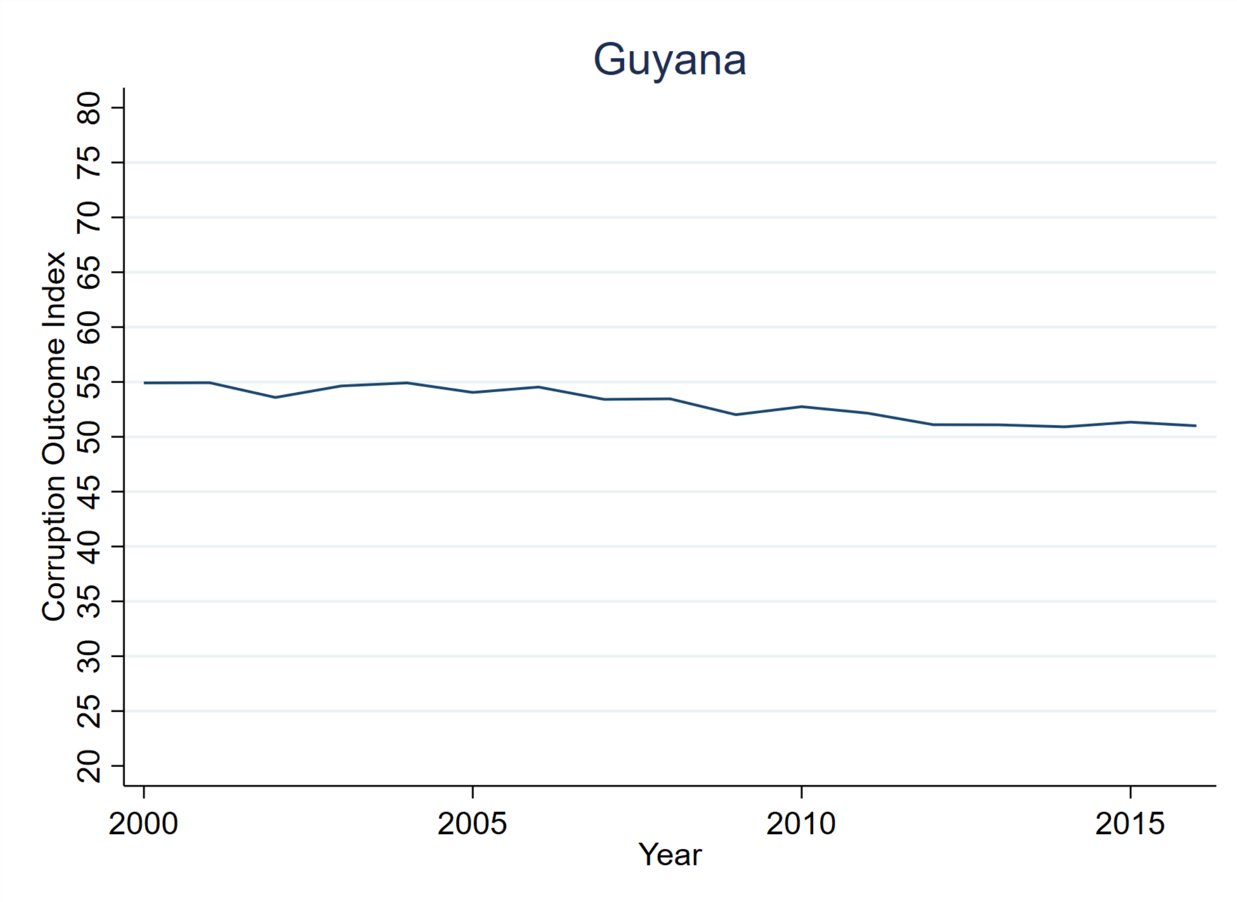

Grenada
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

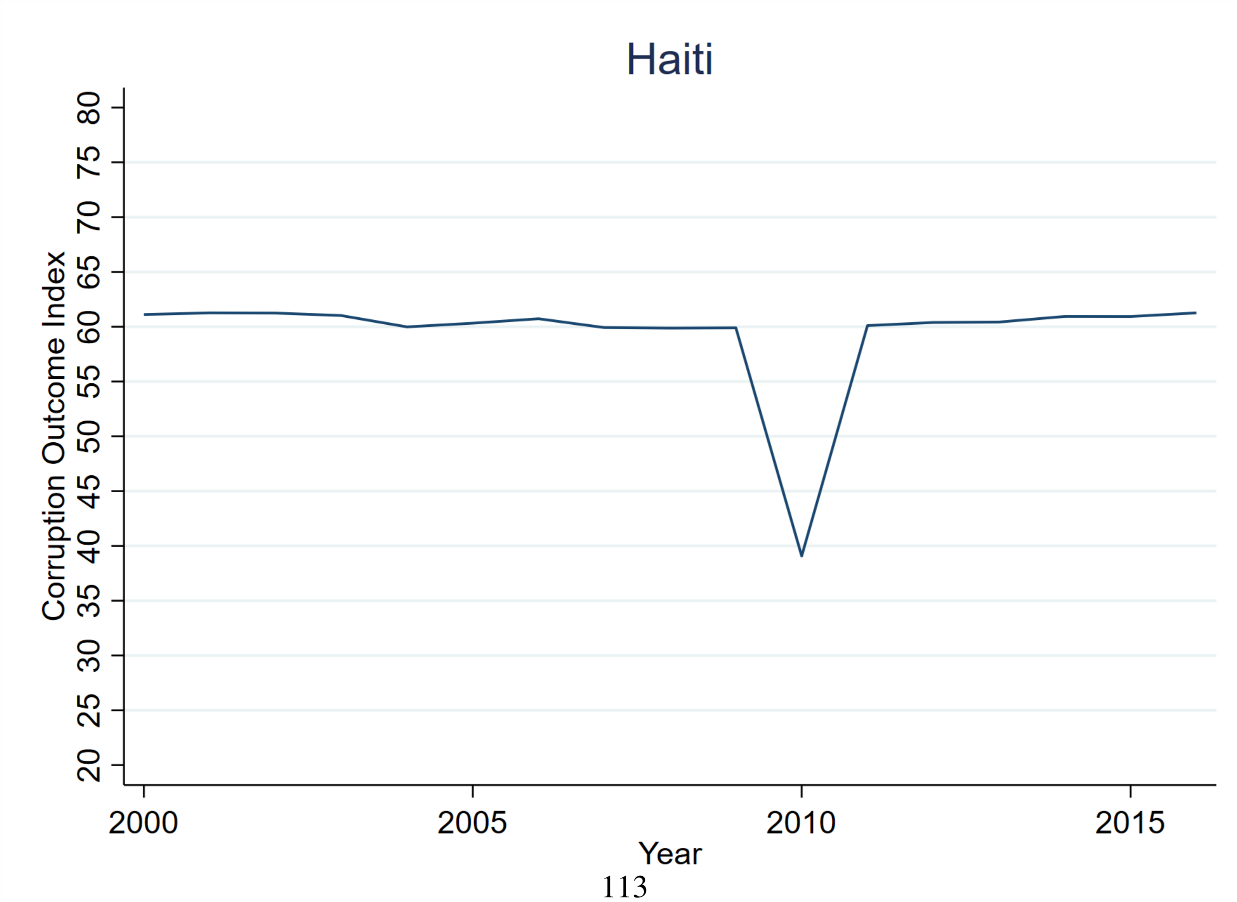


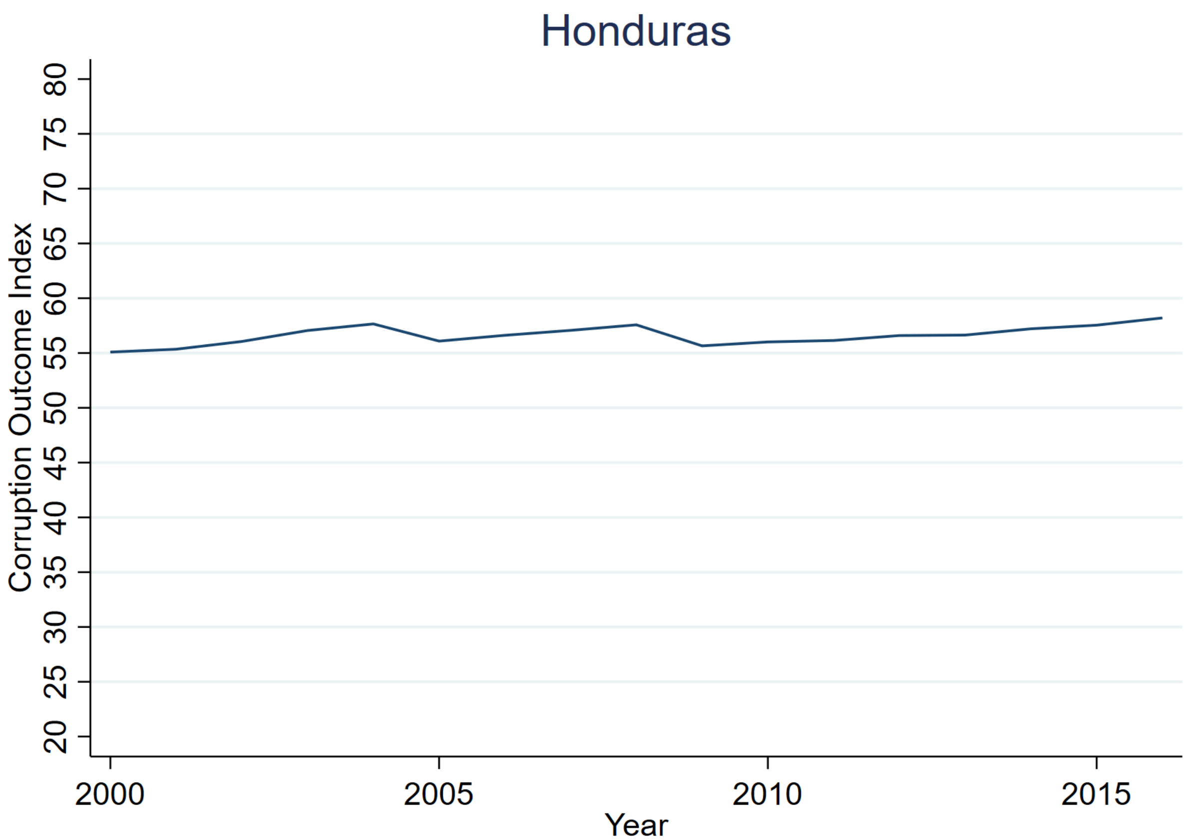
Guatemala
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

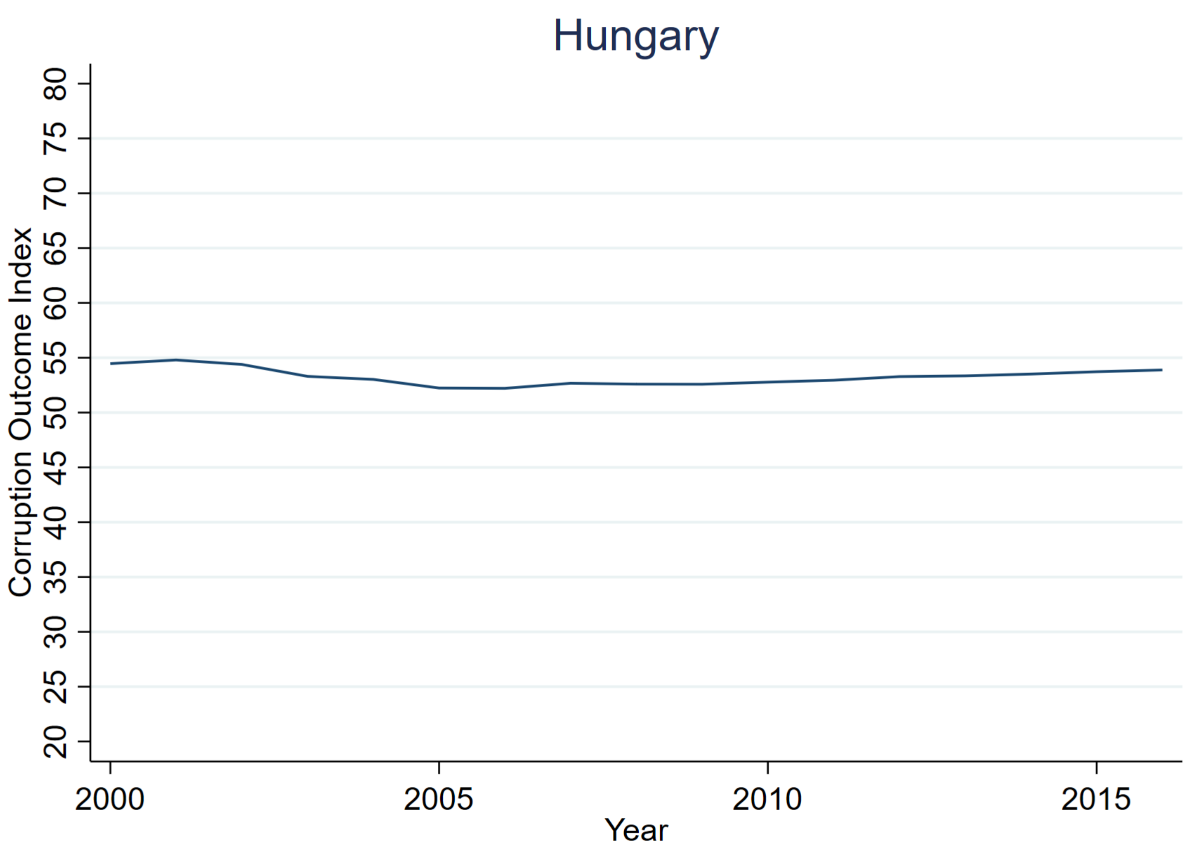


Guinea
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

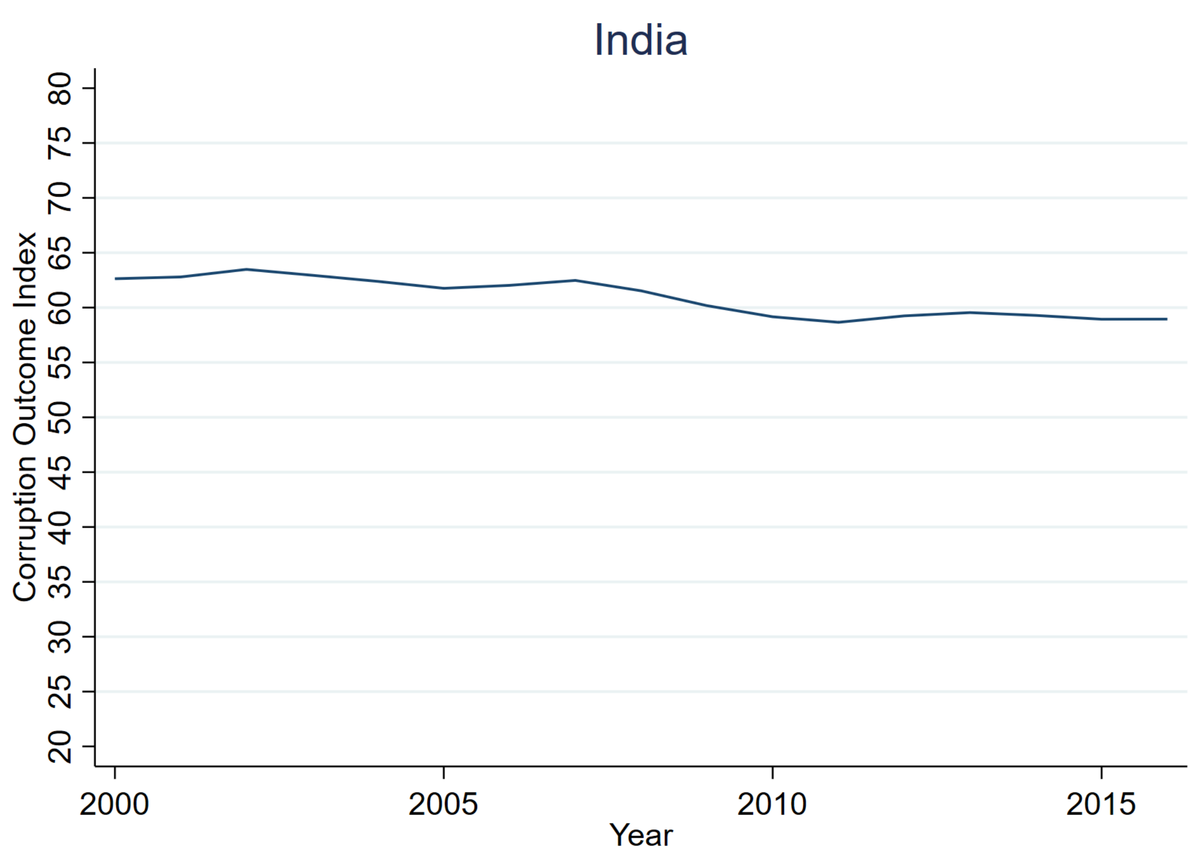

Guyana
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Haiti
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

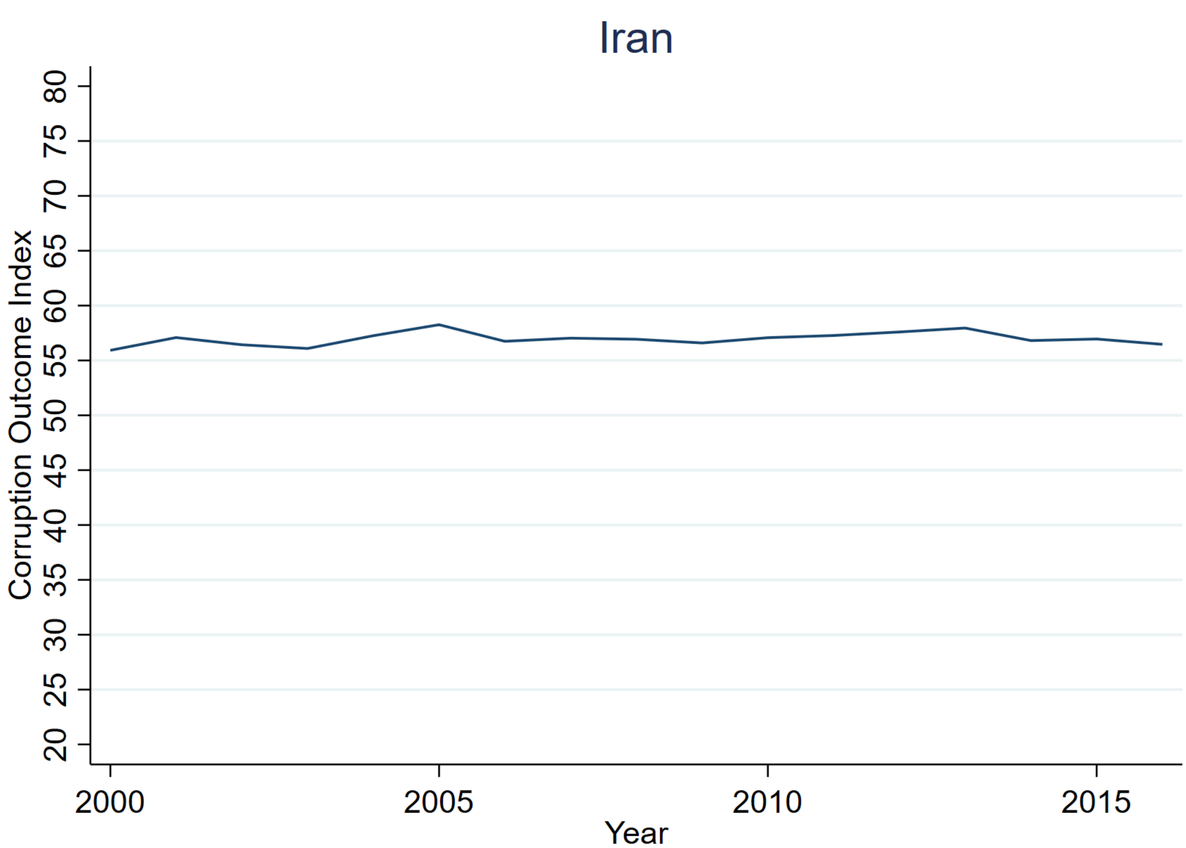

Honduras
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

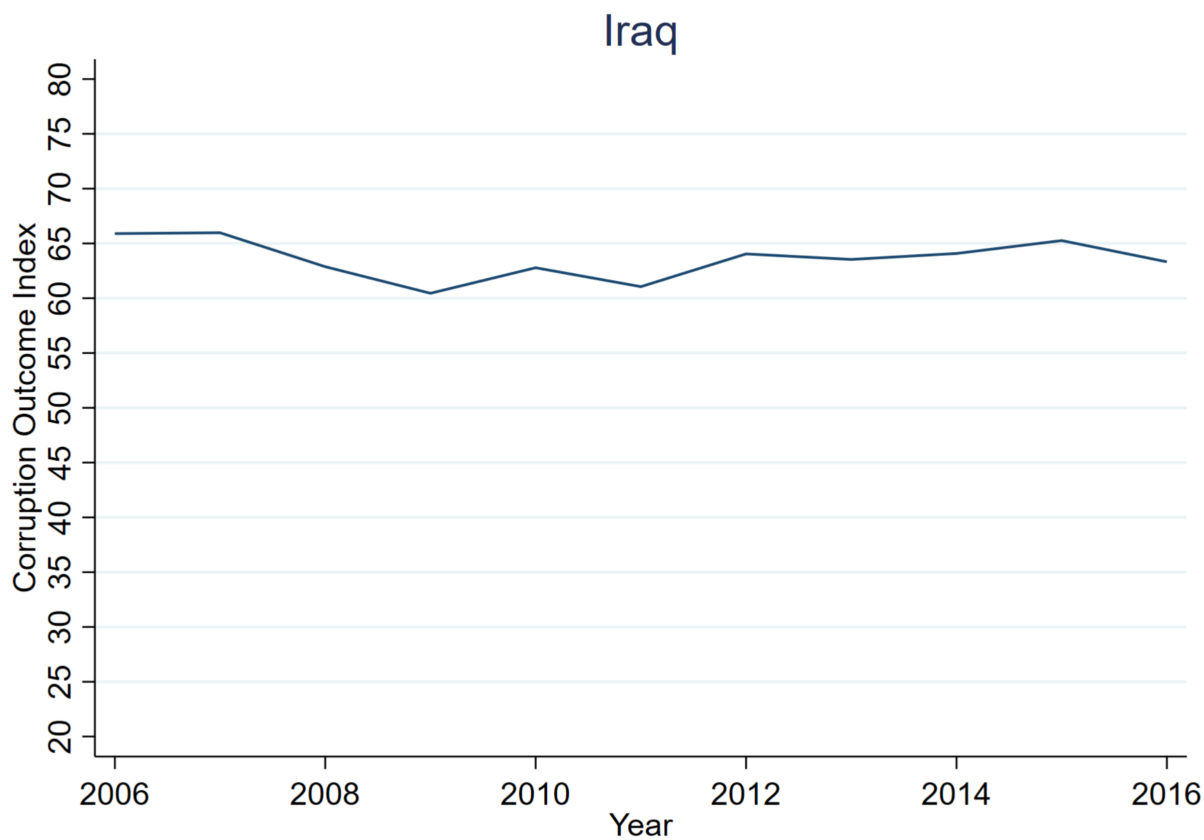

Hungary
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Iceland
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


India
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Indonesia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

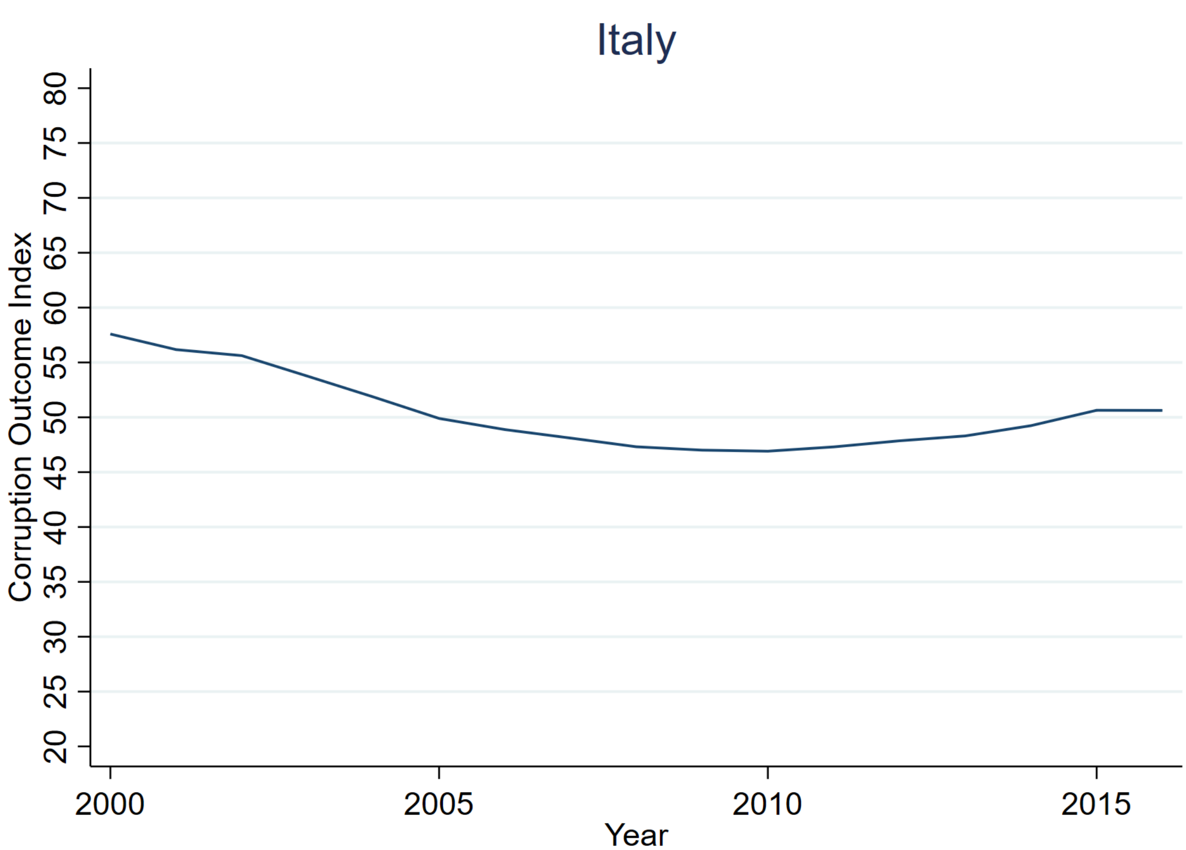

Iran
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

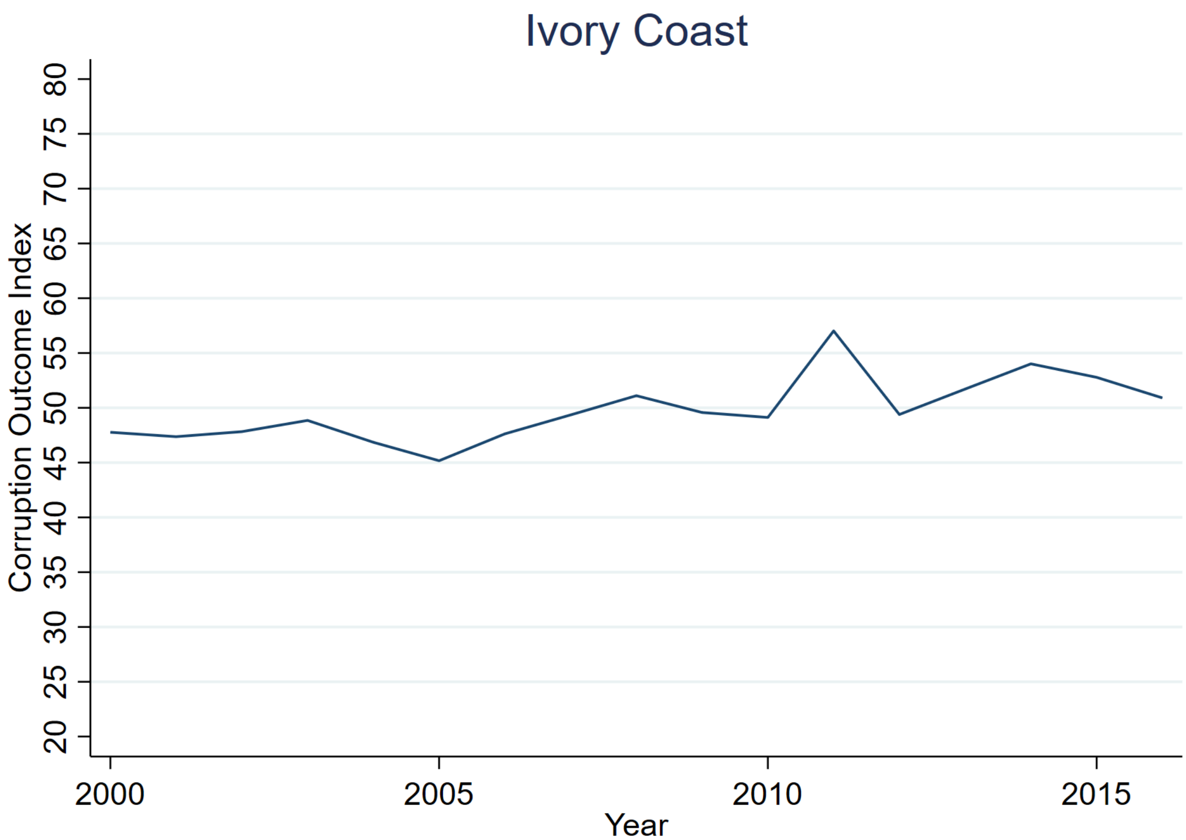


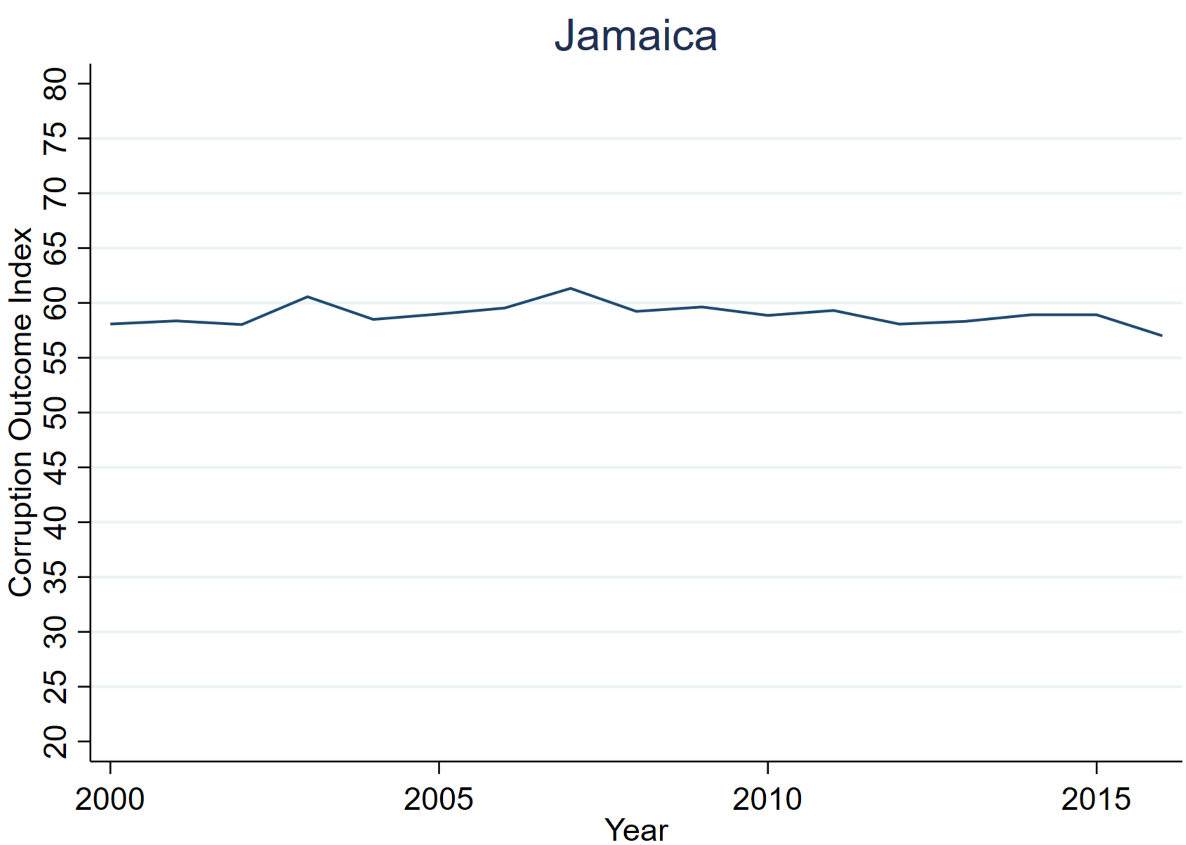
Iraq
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2006
2008
2010
2012
2014
2016
Year

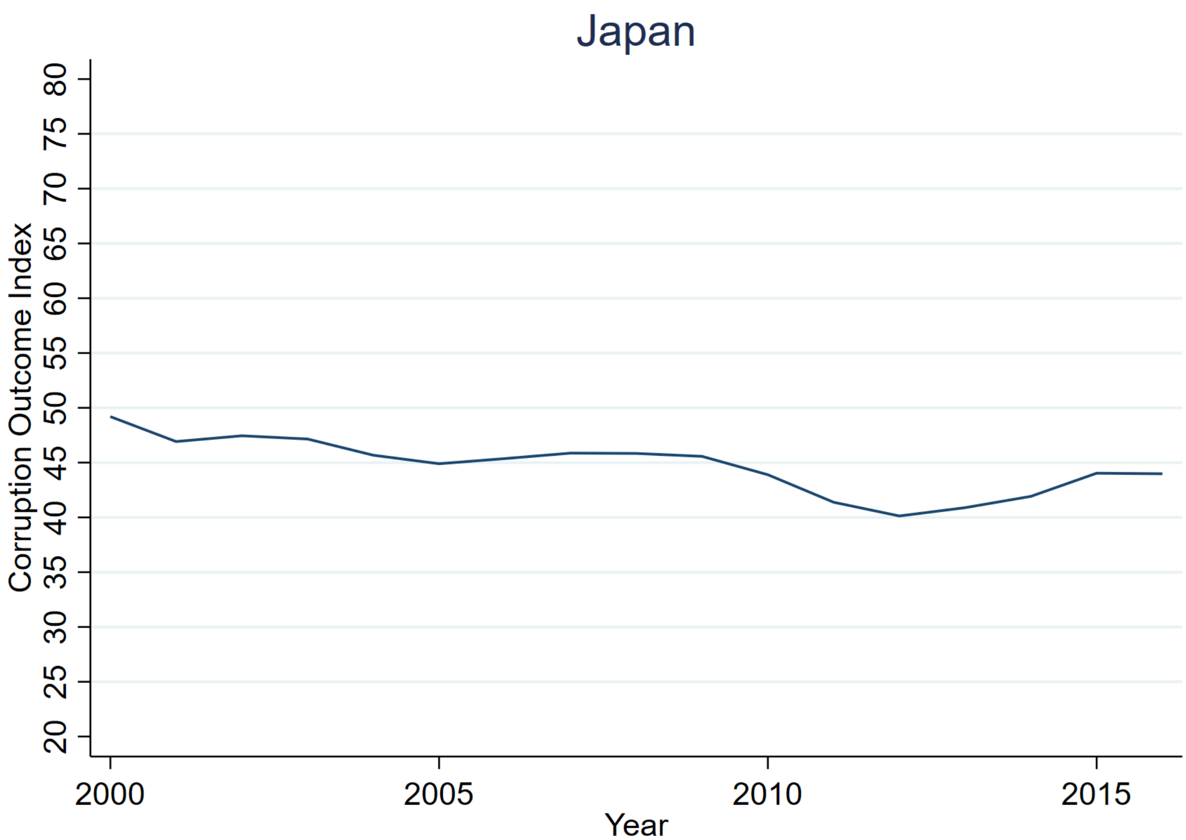
Ireland
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Israel
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

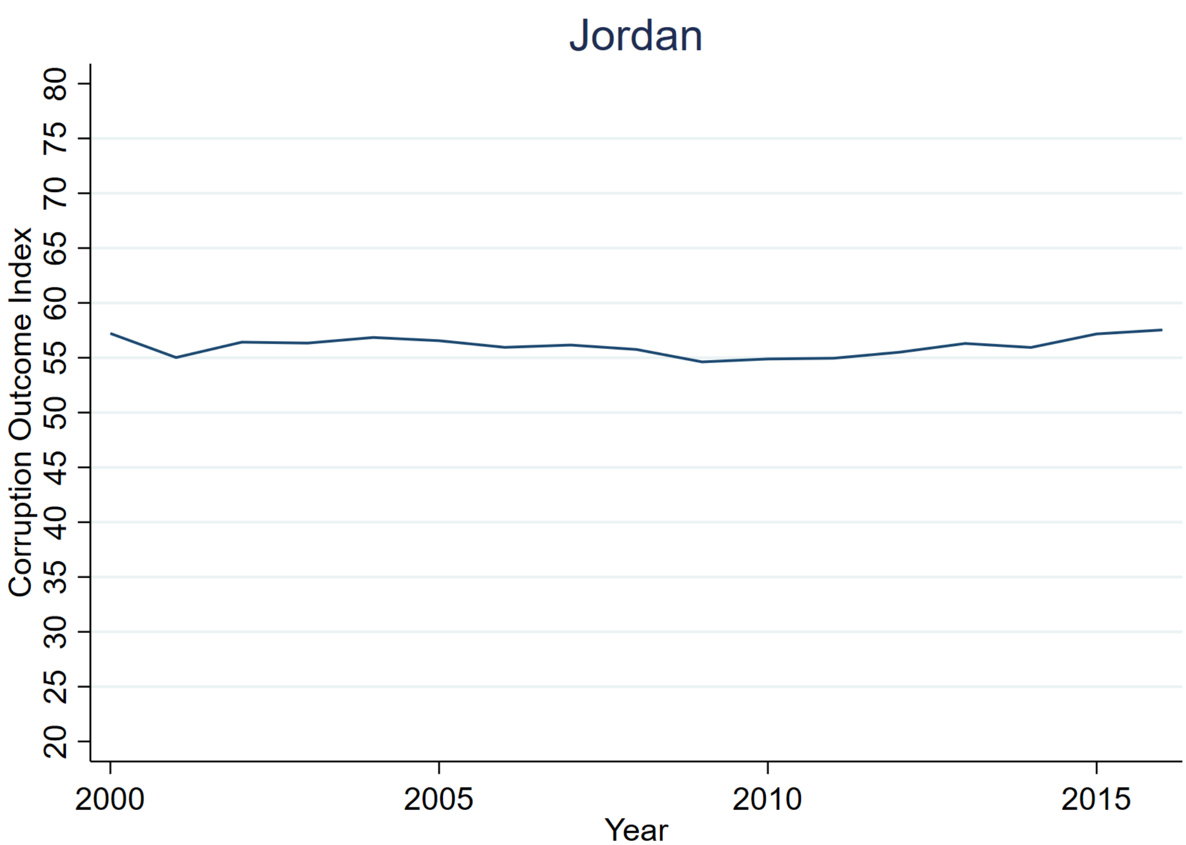

Italy
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Ivory Coast
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

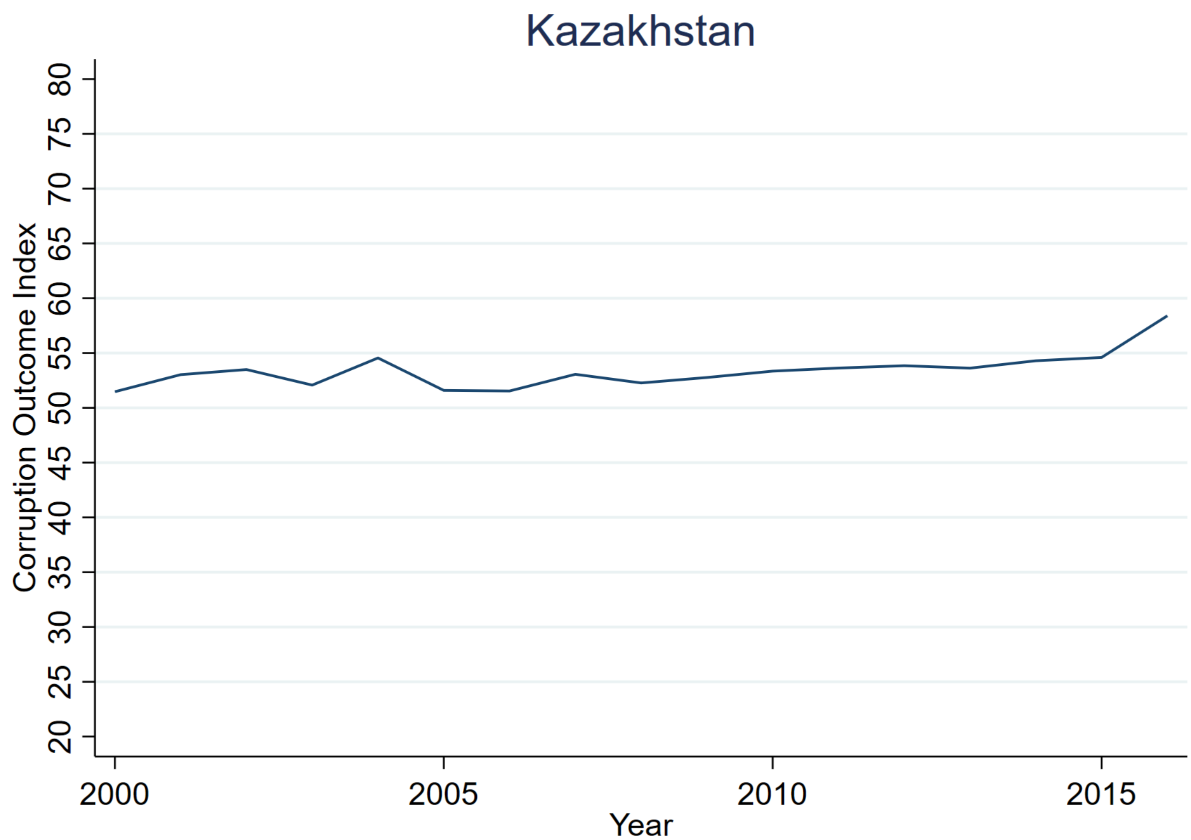

Jamaica
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year



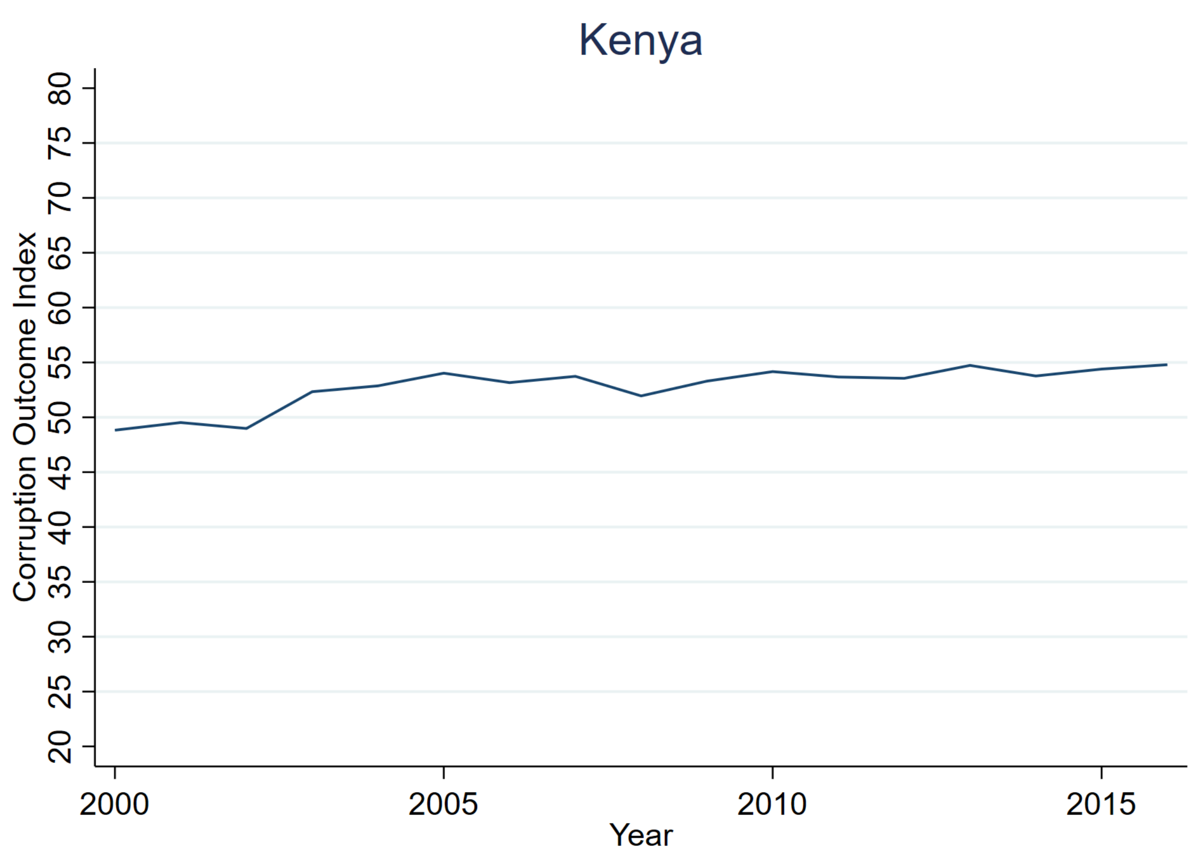
Japan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year



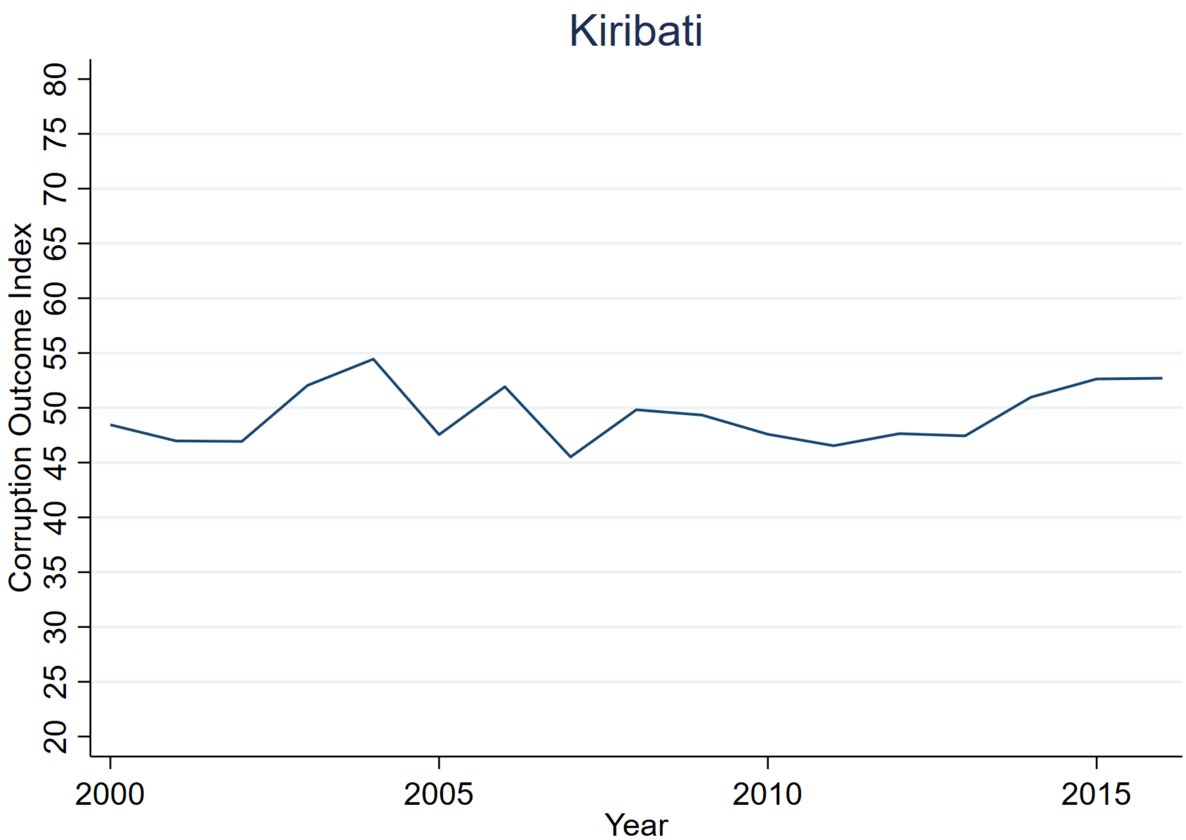
Jordan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

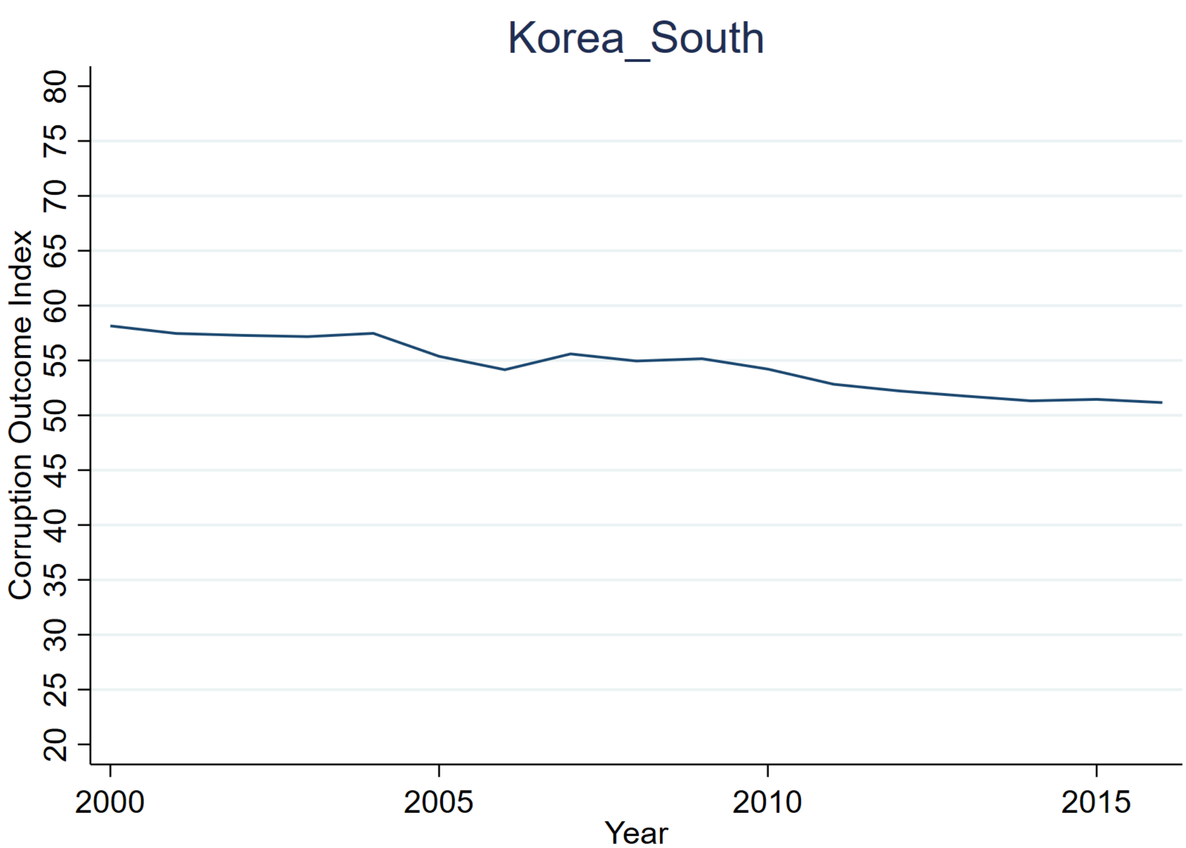


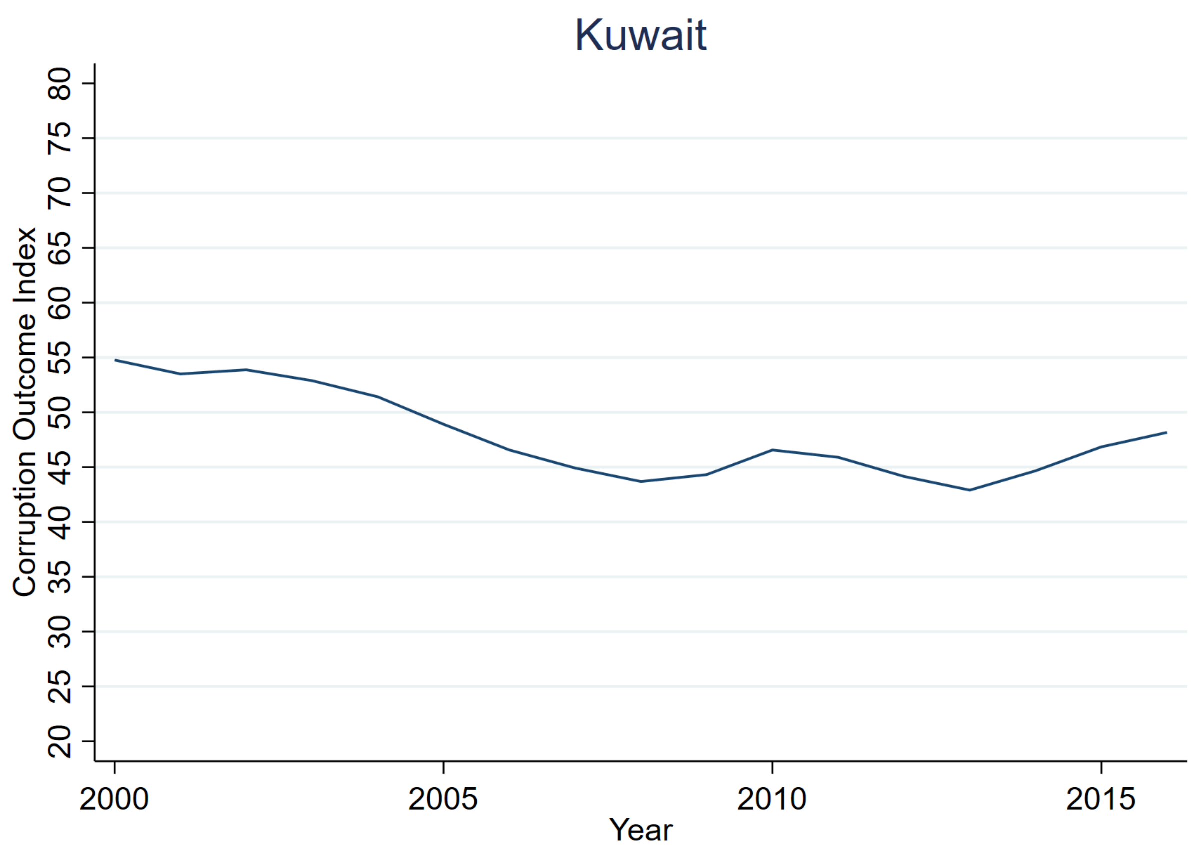
Kazakhstan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

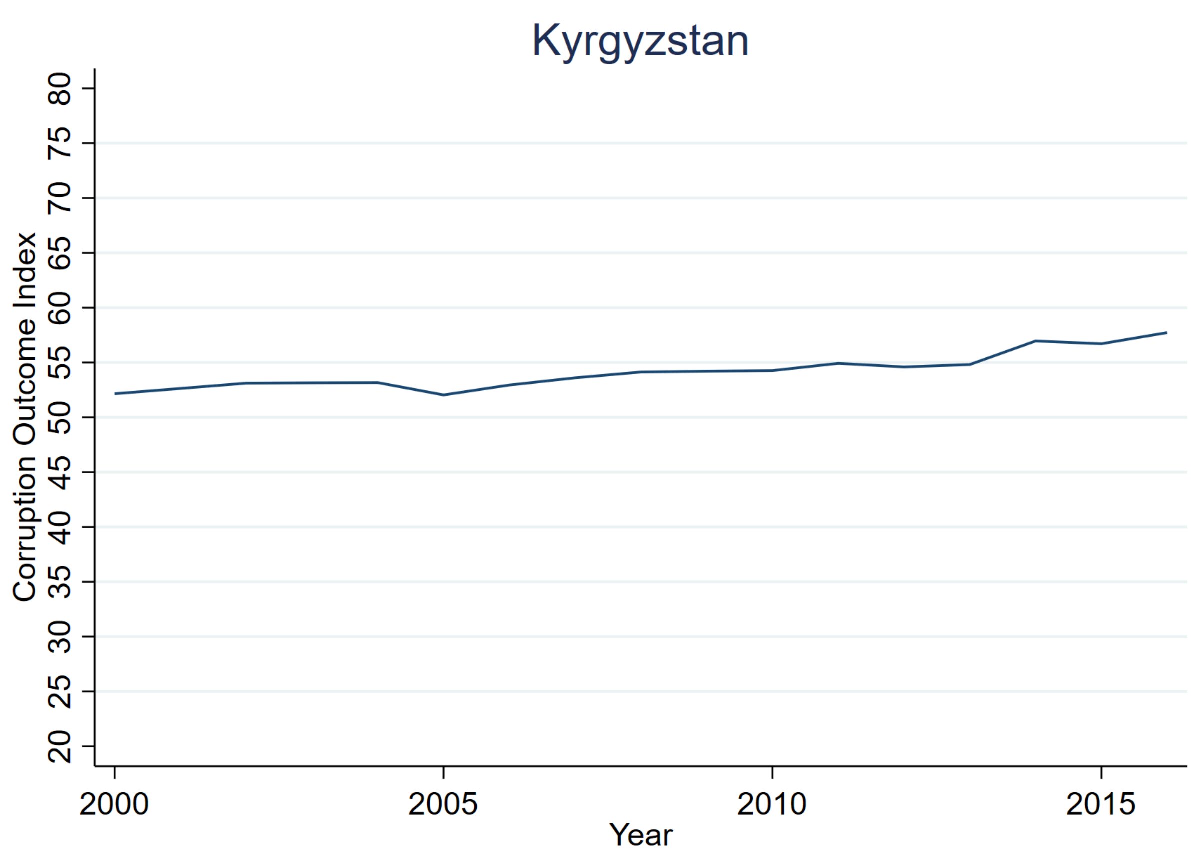


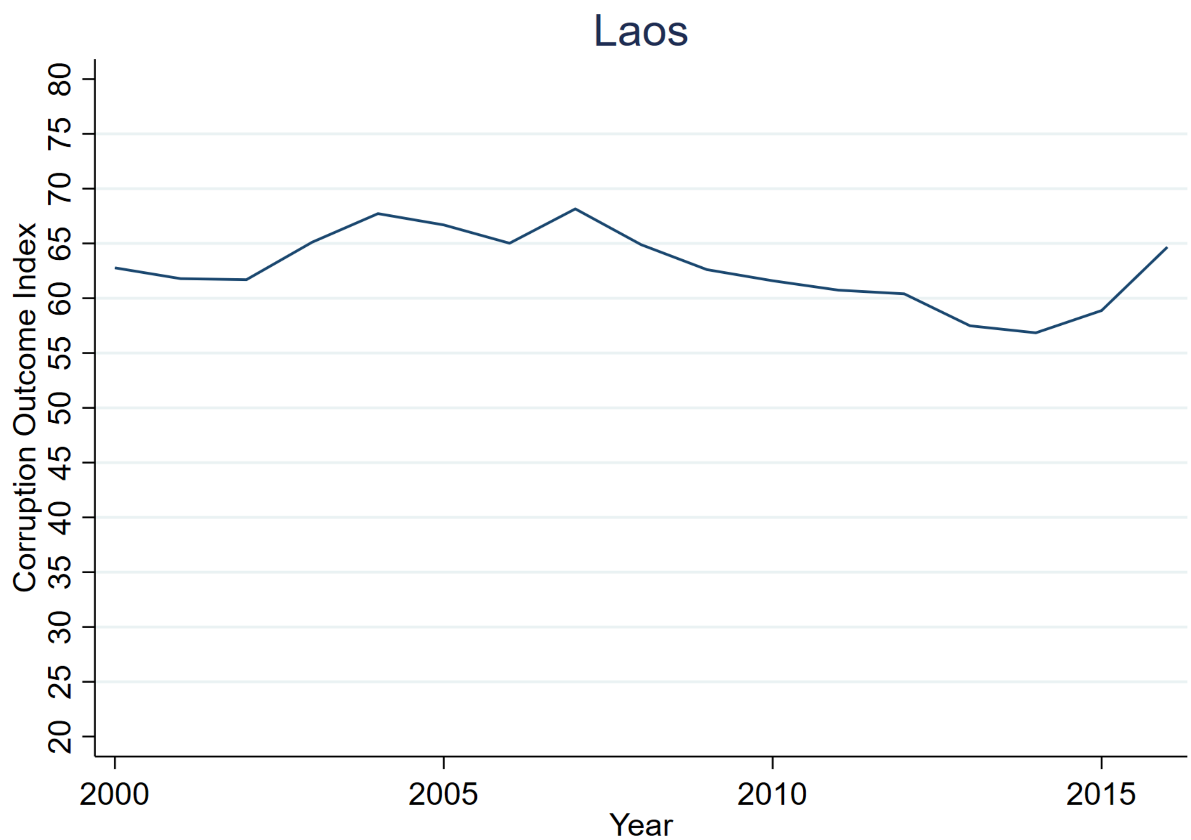
Kenya
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

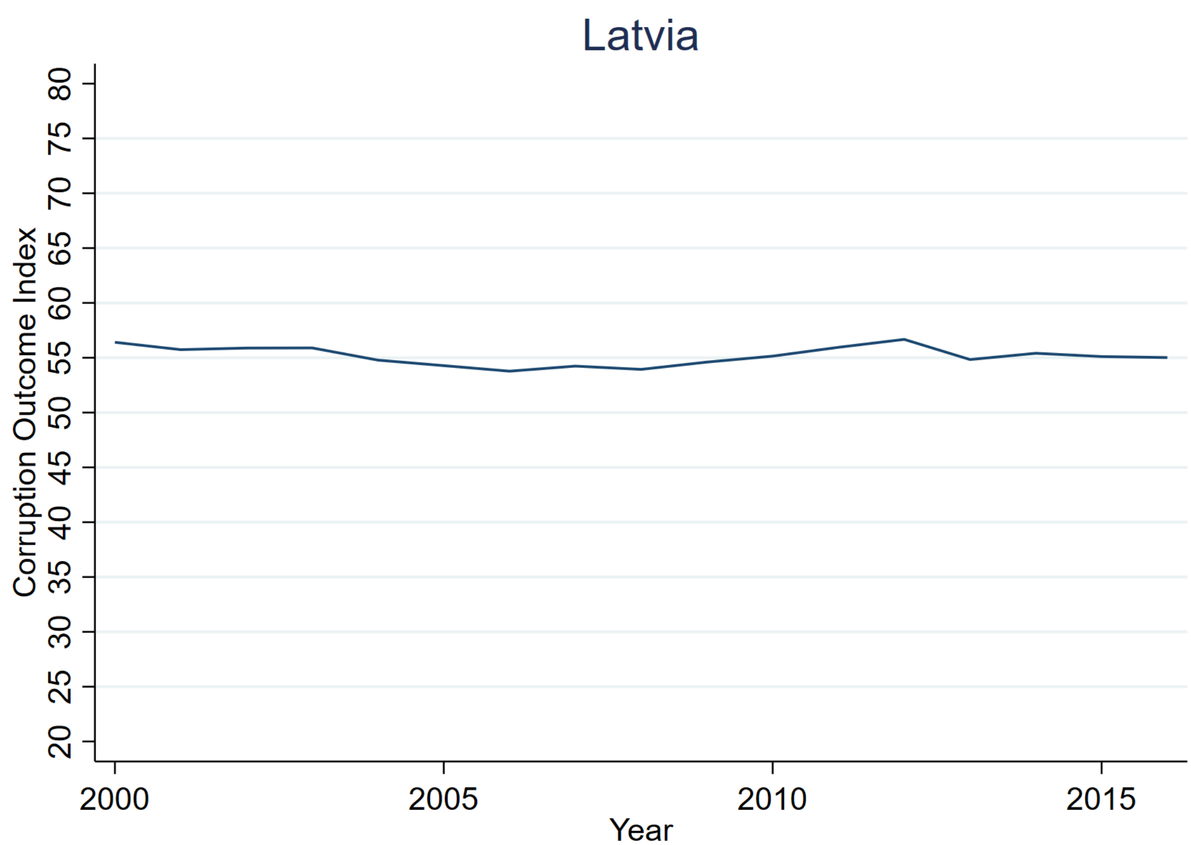

Kiribati
Corruption Outcome Index
80
75
70
65
60
55
50
45
40
35
30
25
20
2000
2005
2010
2015
Year

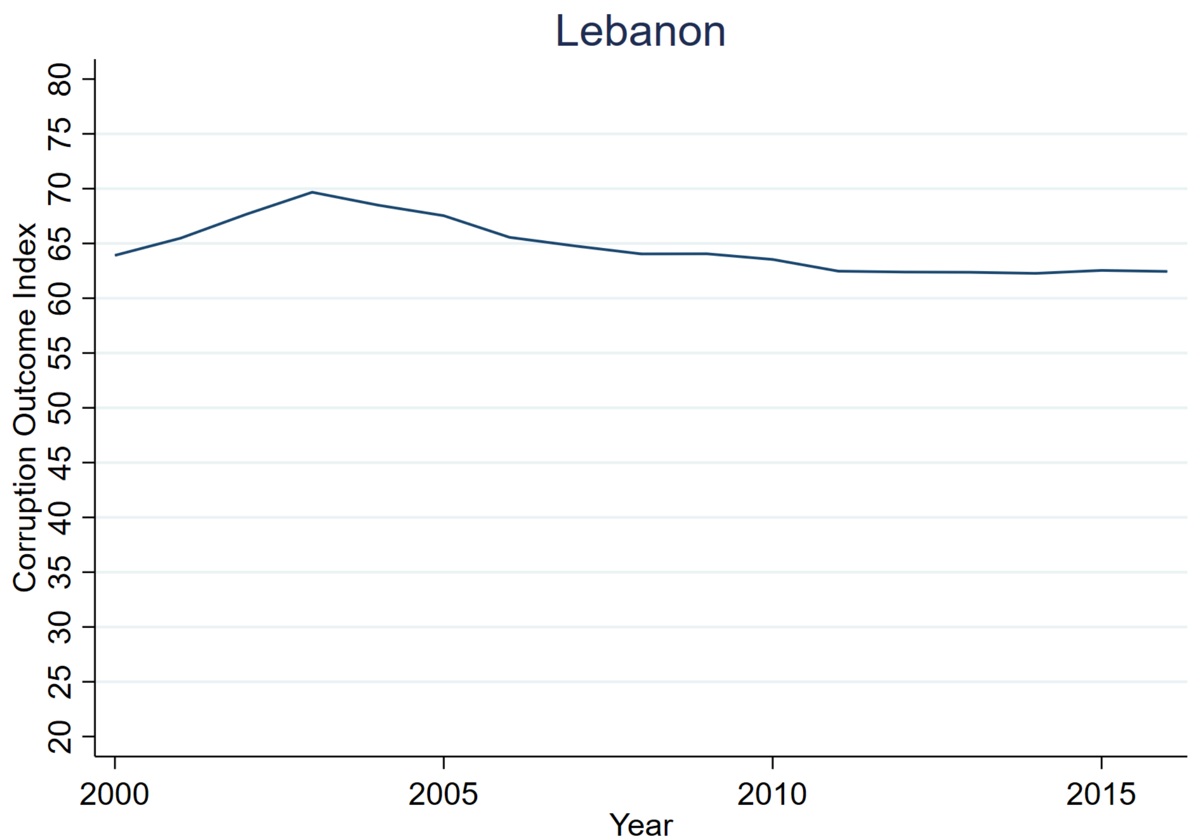

Korea_South
Corruption Outcome Index
80
75
70
65
60
55
50
45
40
35
30
25
20
2000
2005
2010
2015
Year

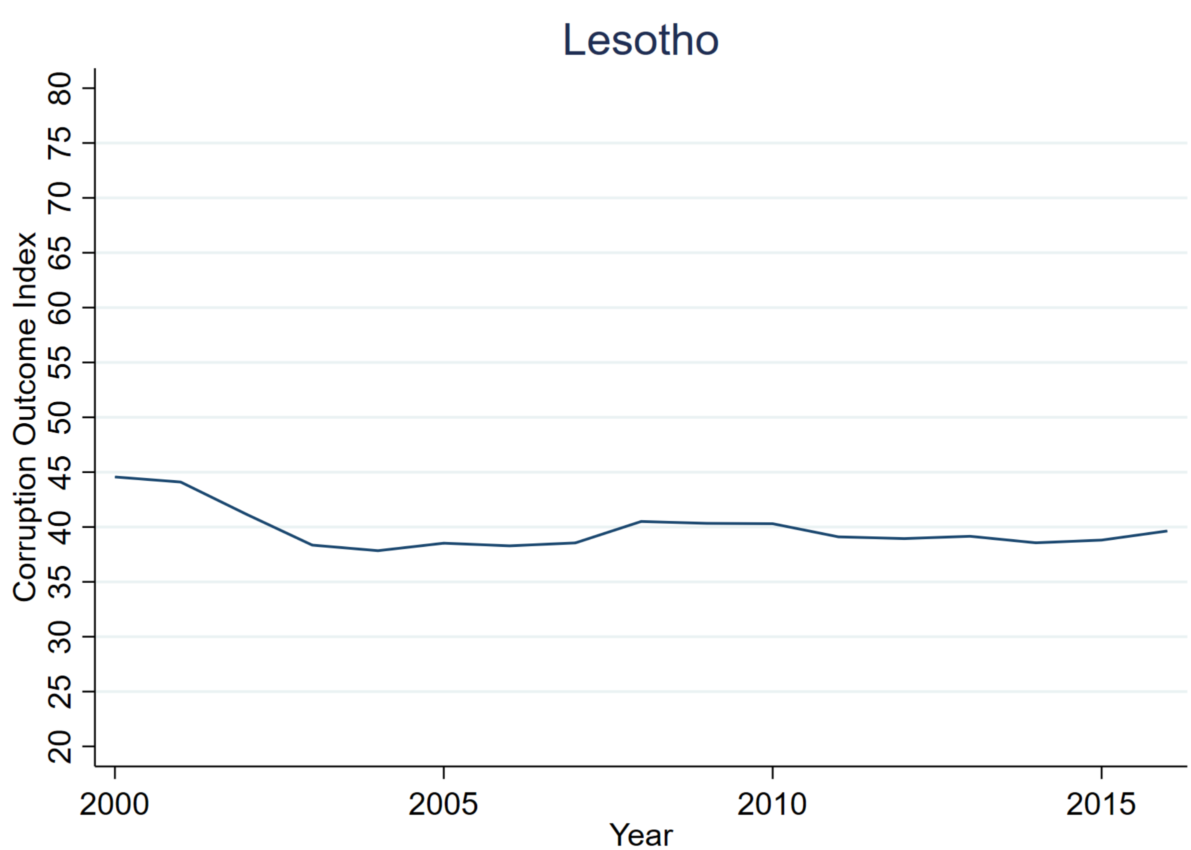

Kuwait
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

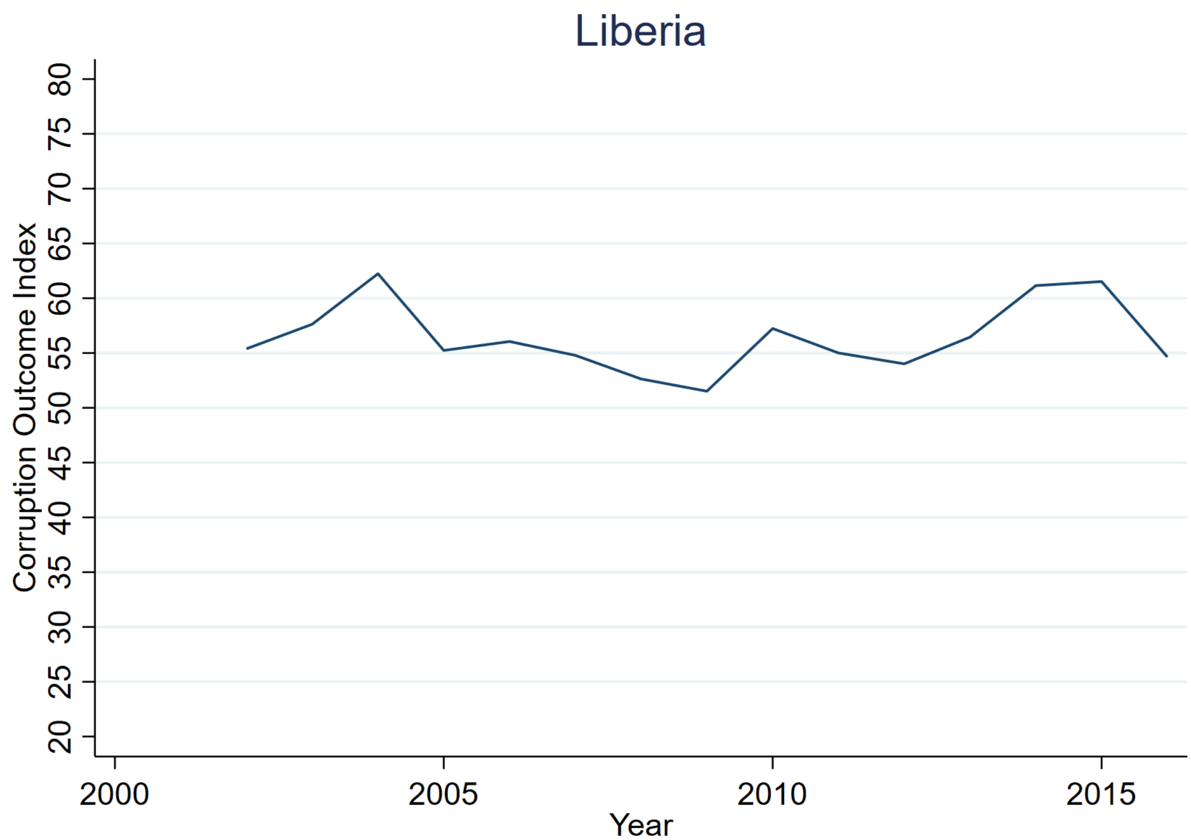

Kyrgyzstan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

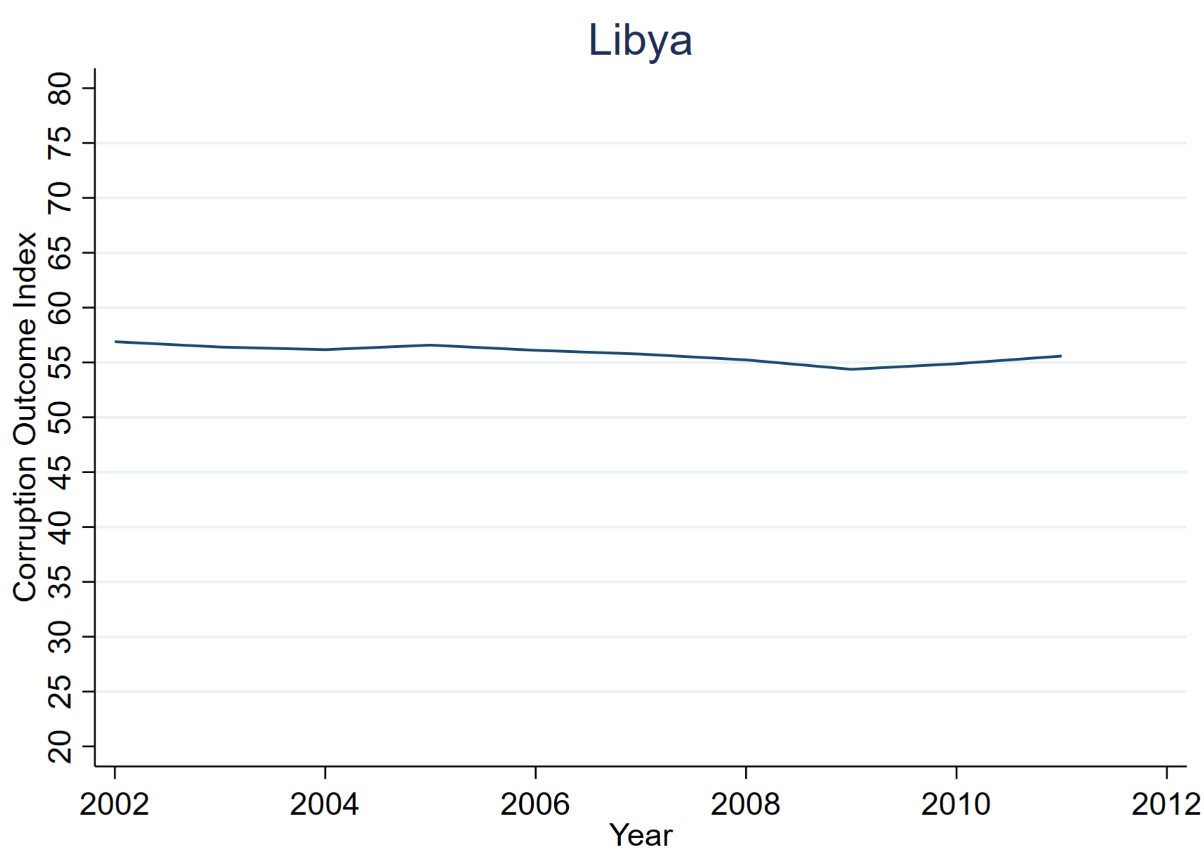

Laos
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

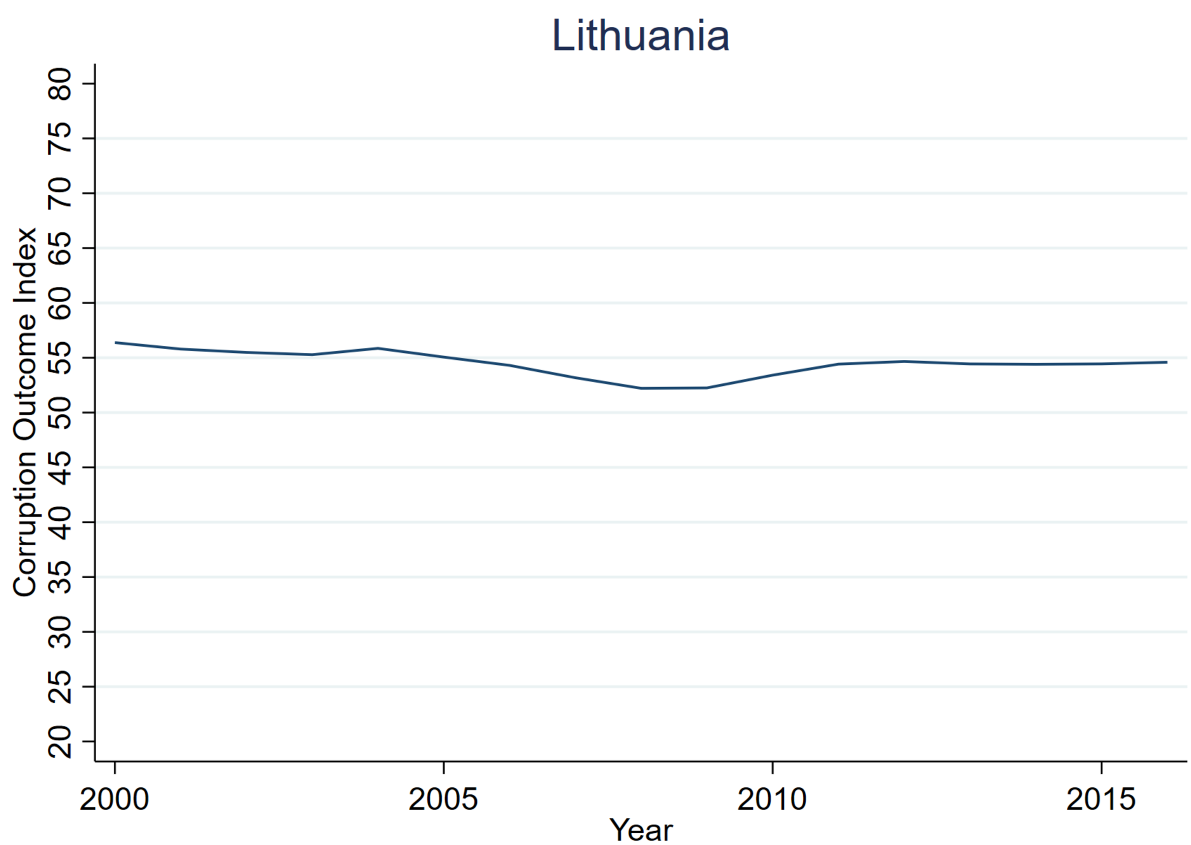

Latvia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

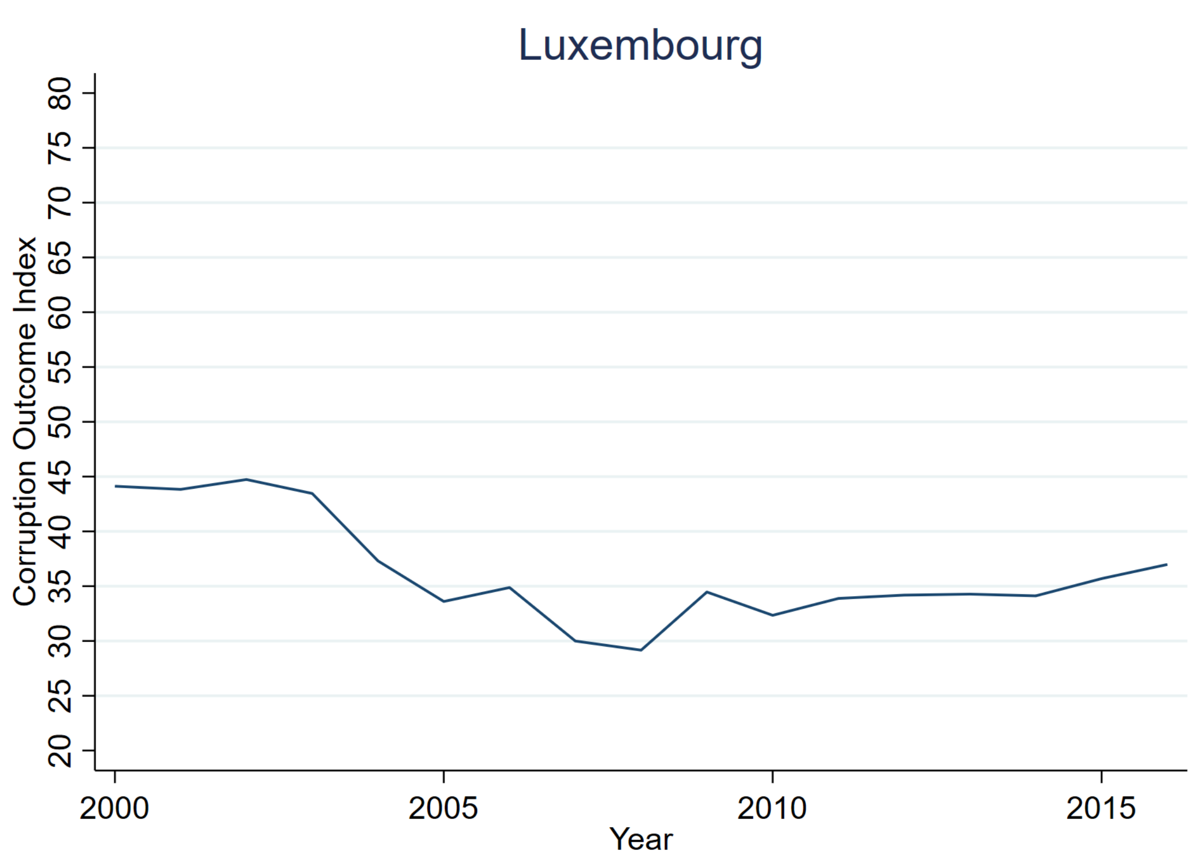

Lebanon
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

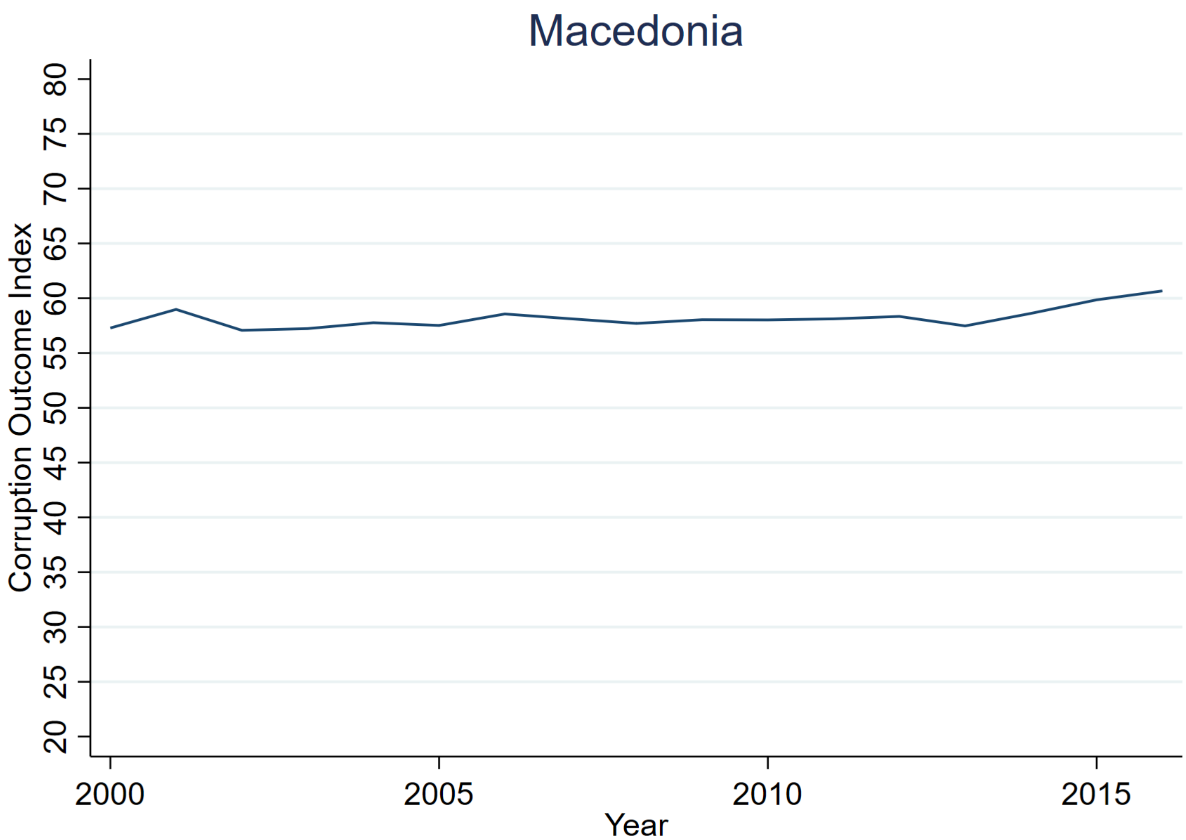

Lesotho
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Liberia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

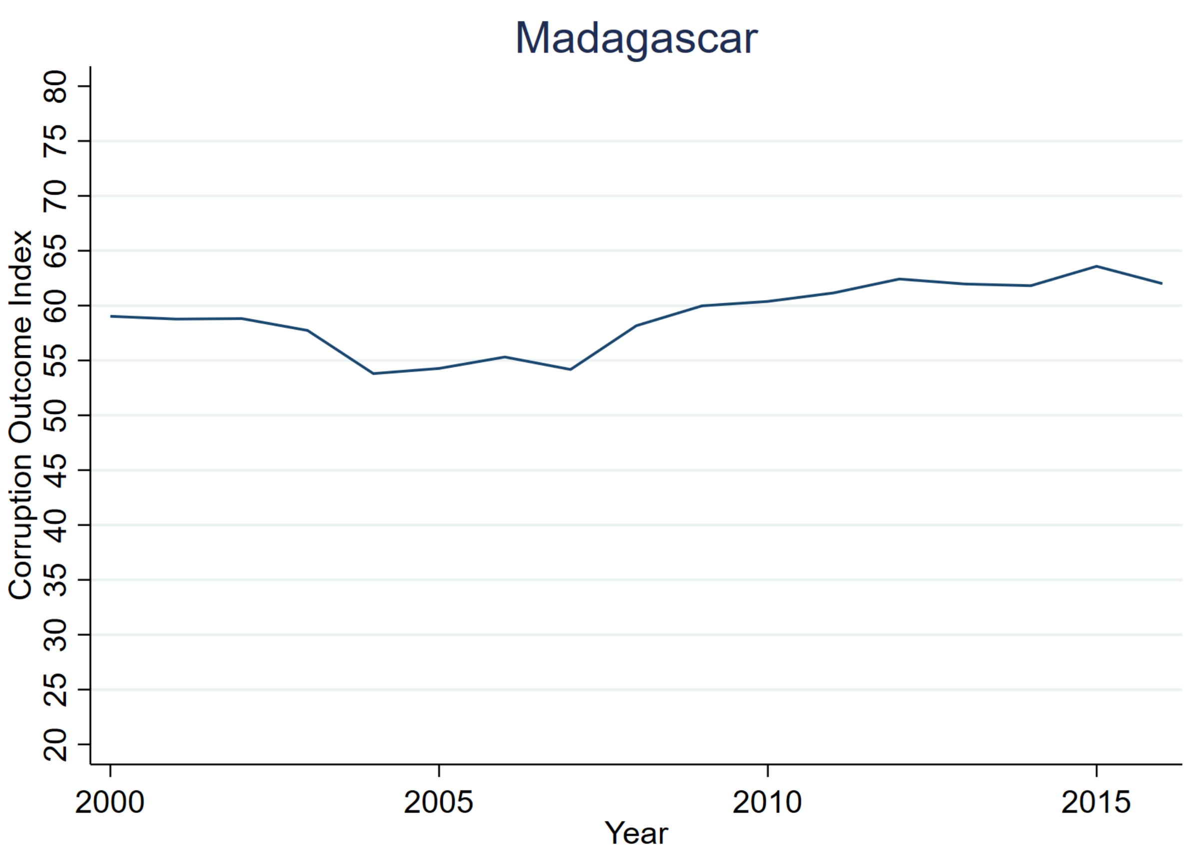

Libya
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2002
2004
2006
2008
2010
2012
Year


Lithuania
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

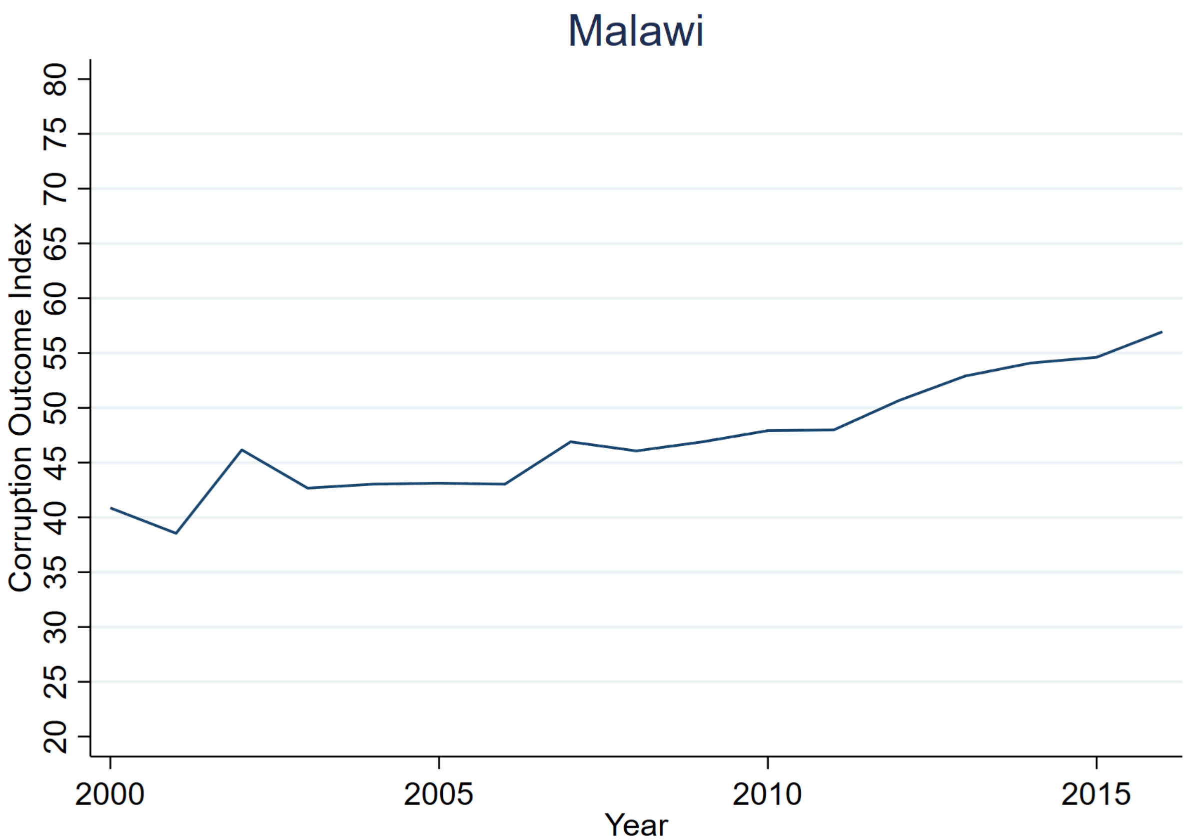

Luxembourg
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

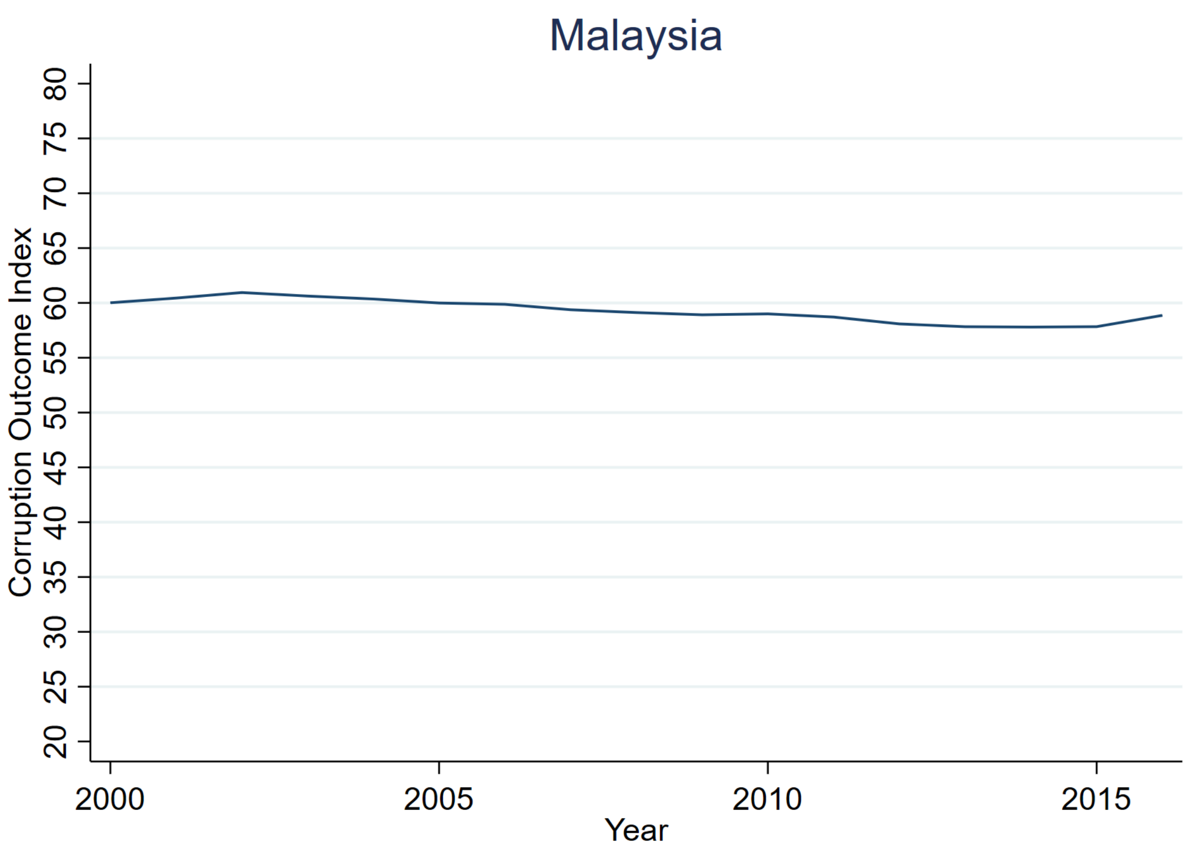


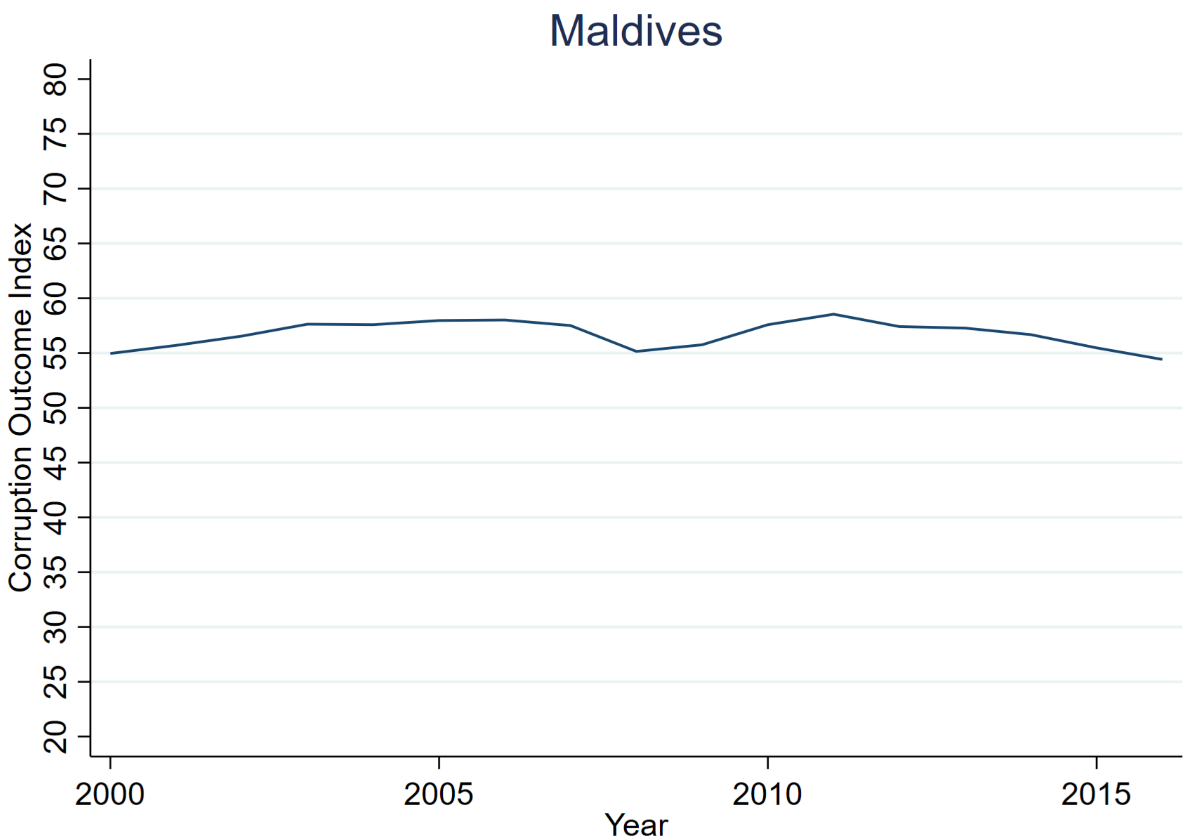
Macedonia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year



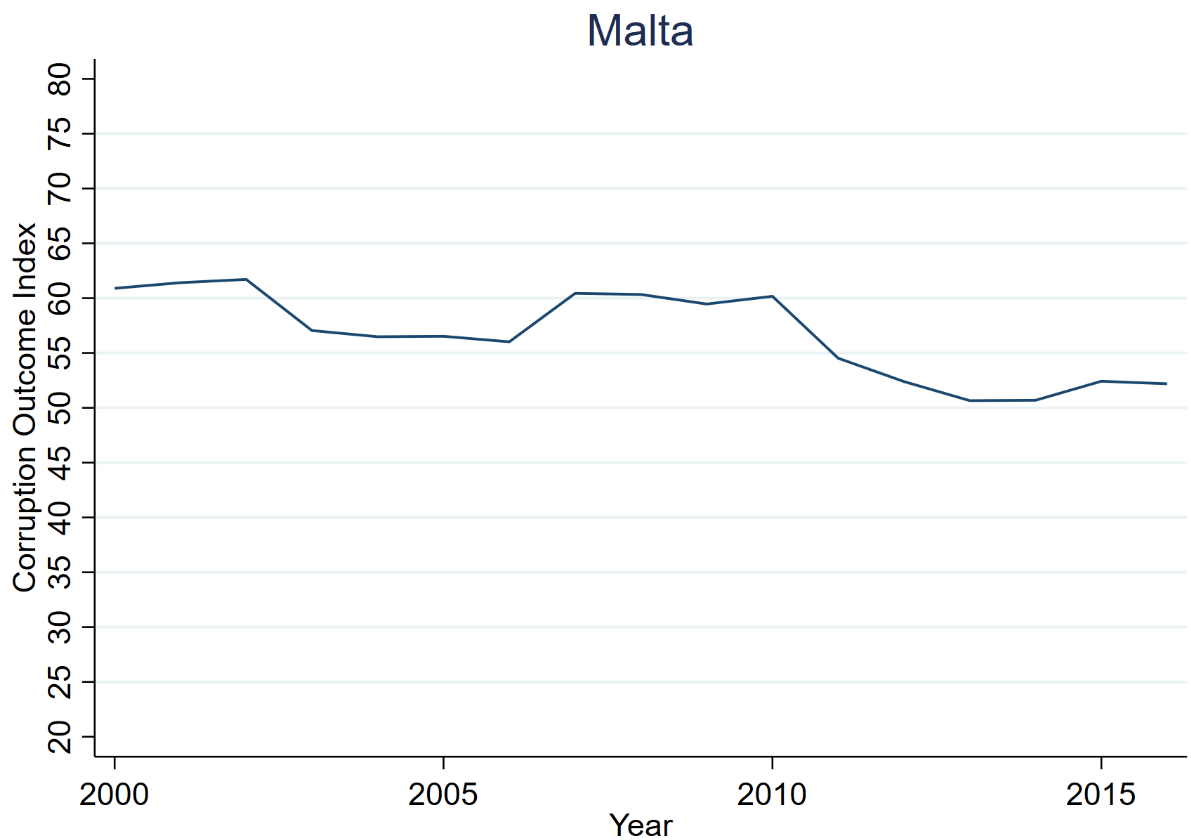
Madagascar
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

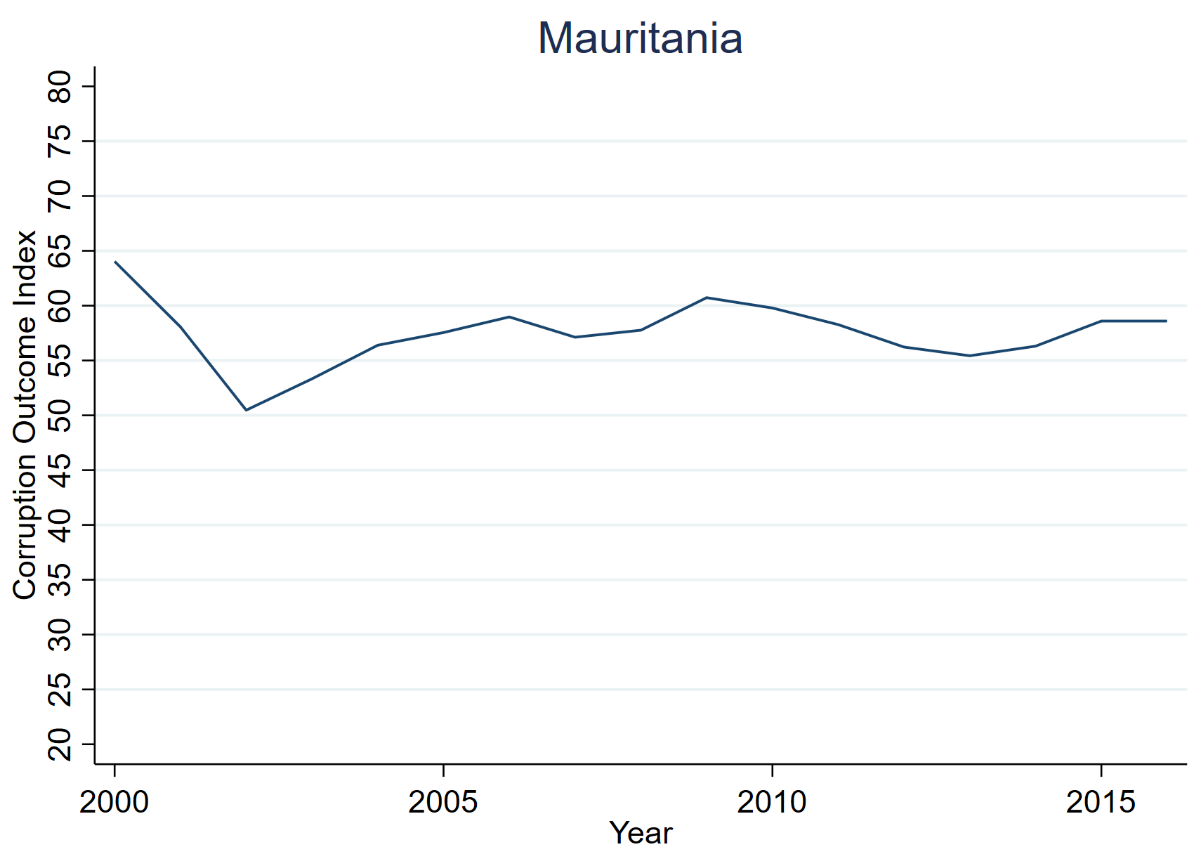

Malawi
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

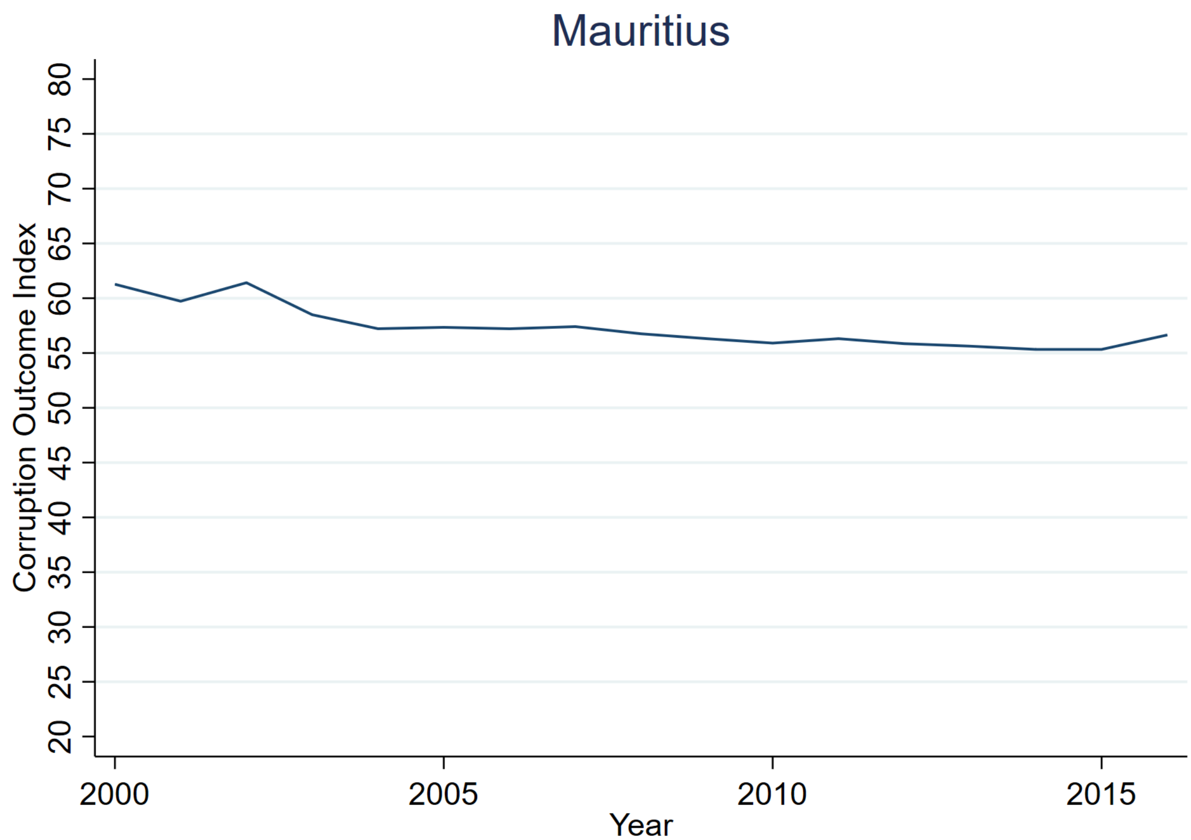

Malaysia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

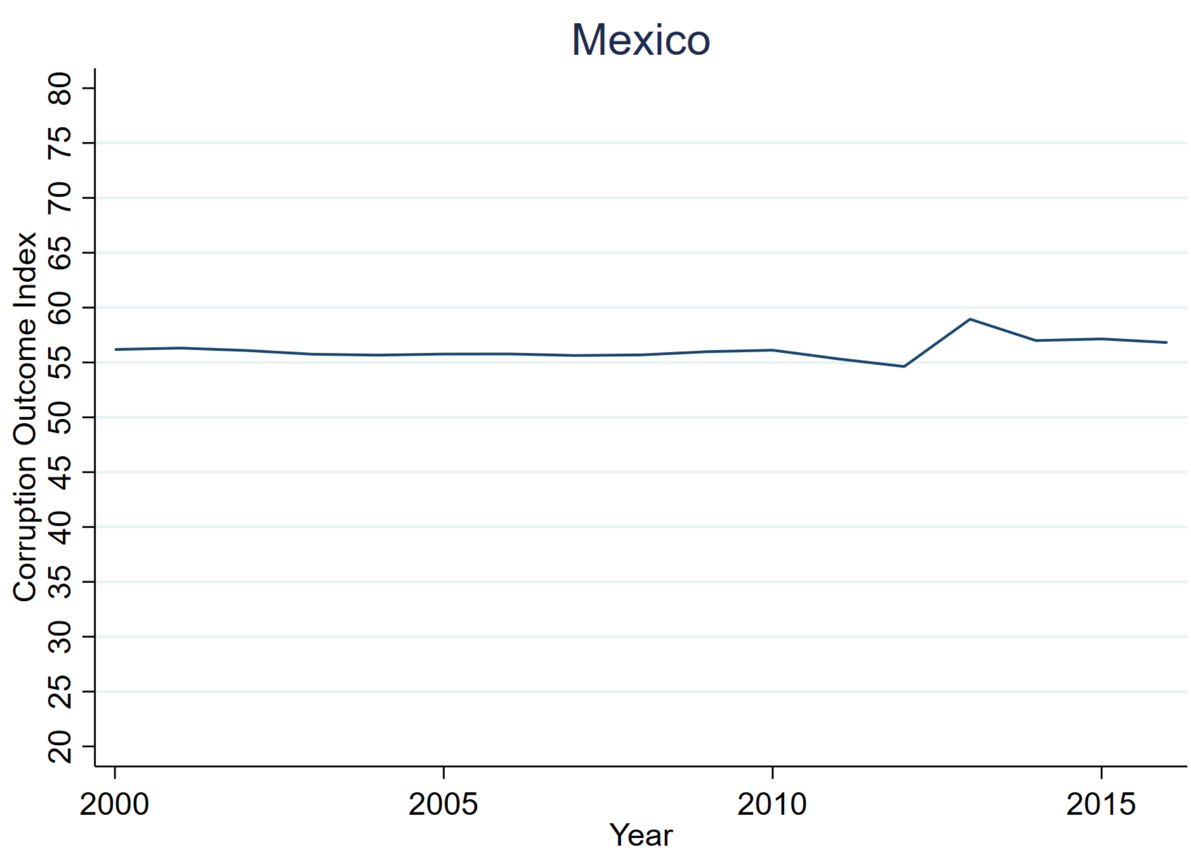

Maldives
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

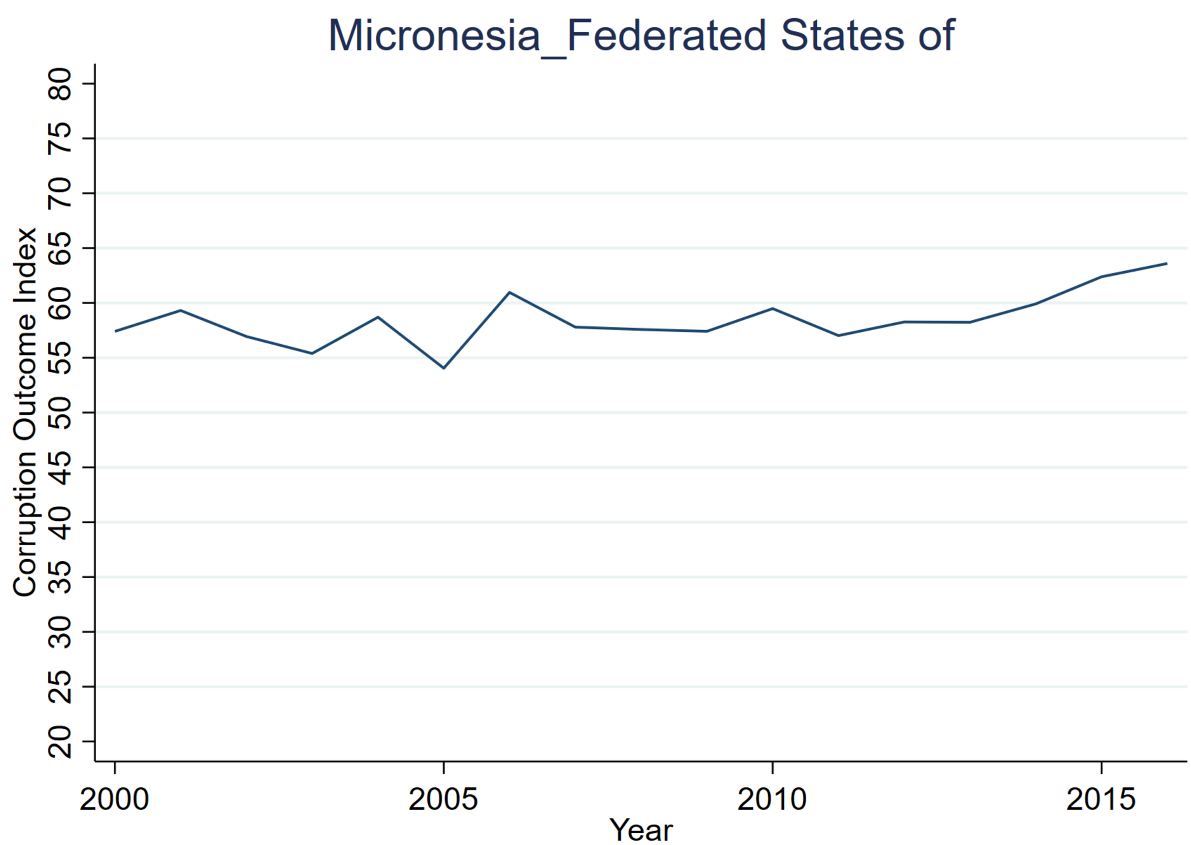

Mali
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Malta
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

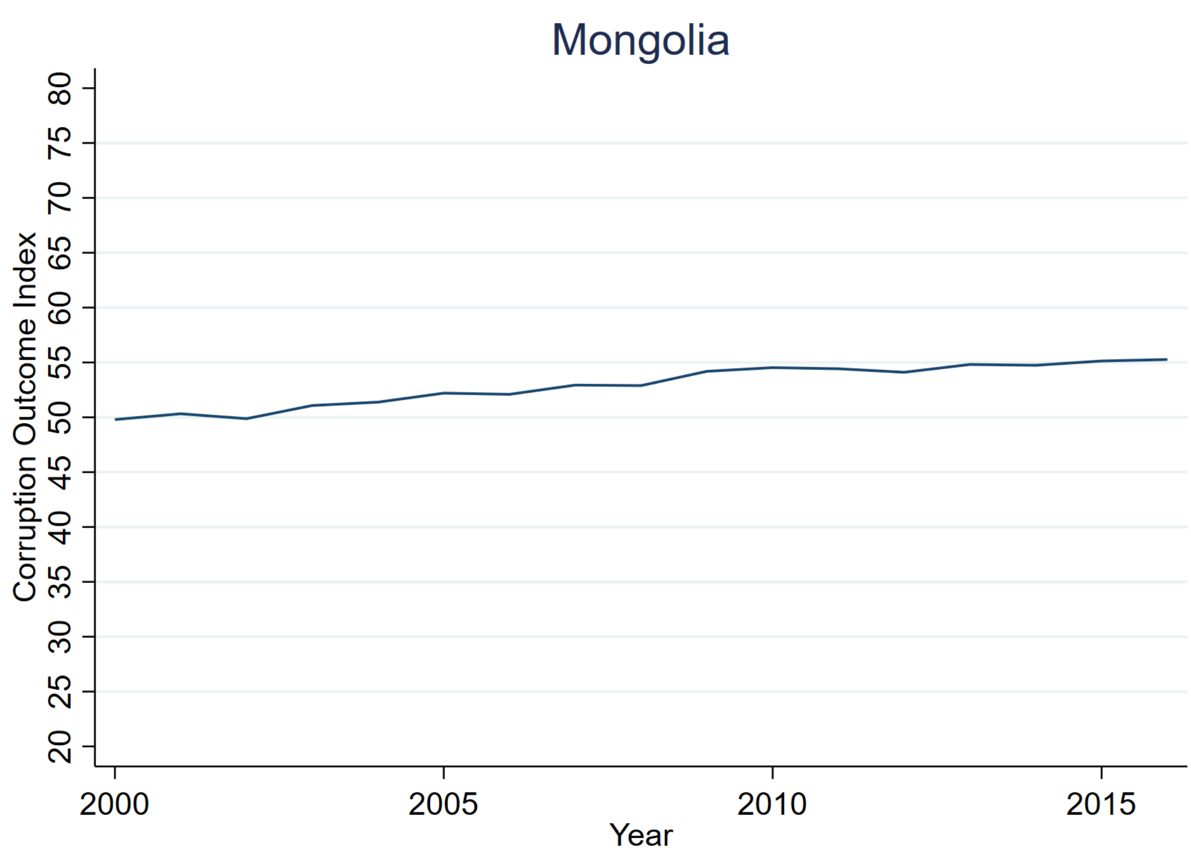

Mauritania
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Mauritius
Corruption Outcome Index
20 25 30 35 40 45 50 55 60 65 70 75 80
2000 2005 2010 2015
Year

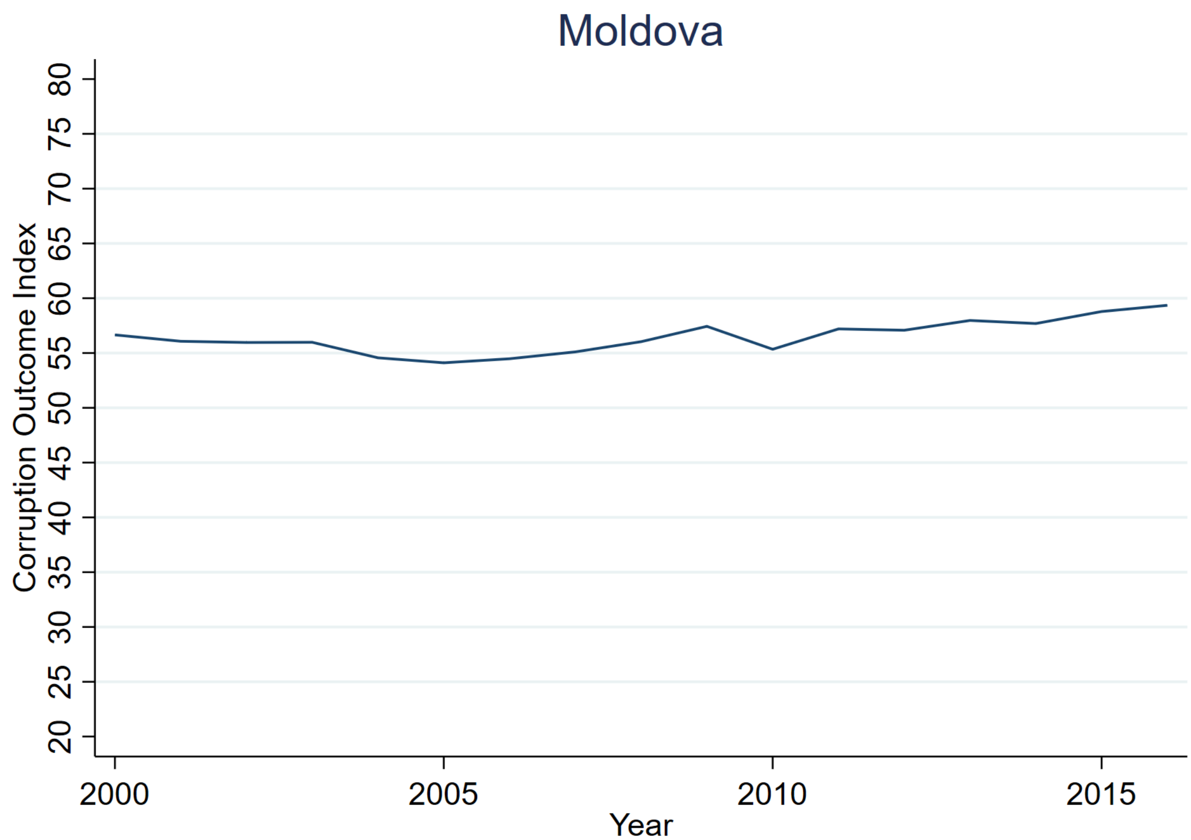

Mexico
Corruption Outcome Index
20 25 30 35 40 45 50 55 60 65 70 75 80
2000 2005 2010 2015
Year

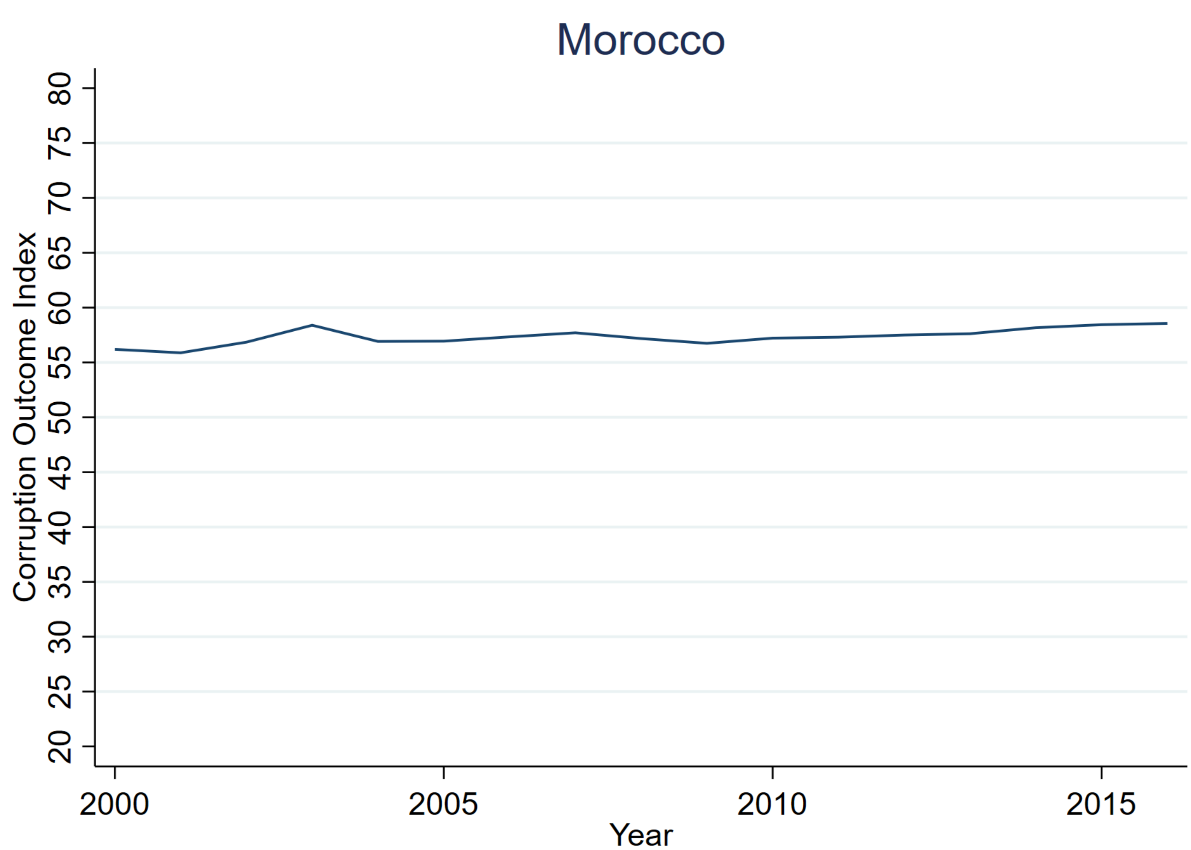


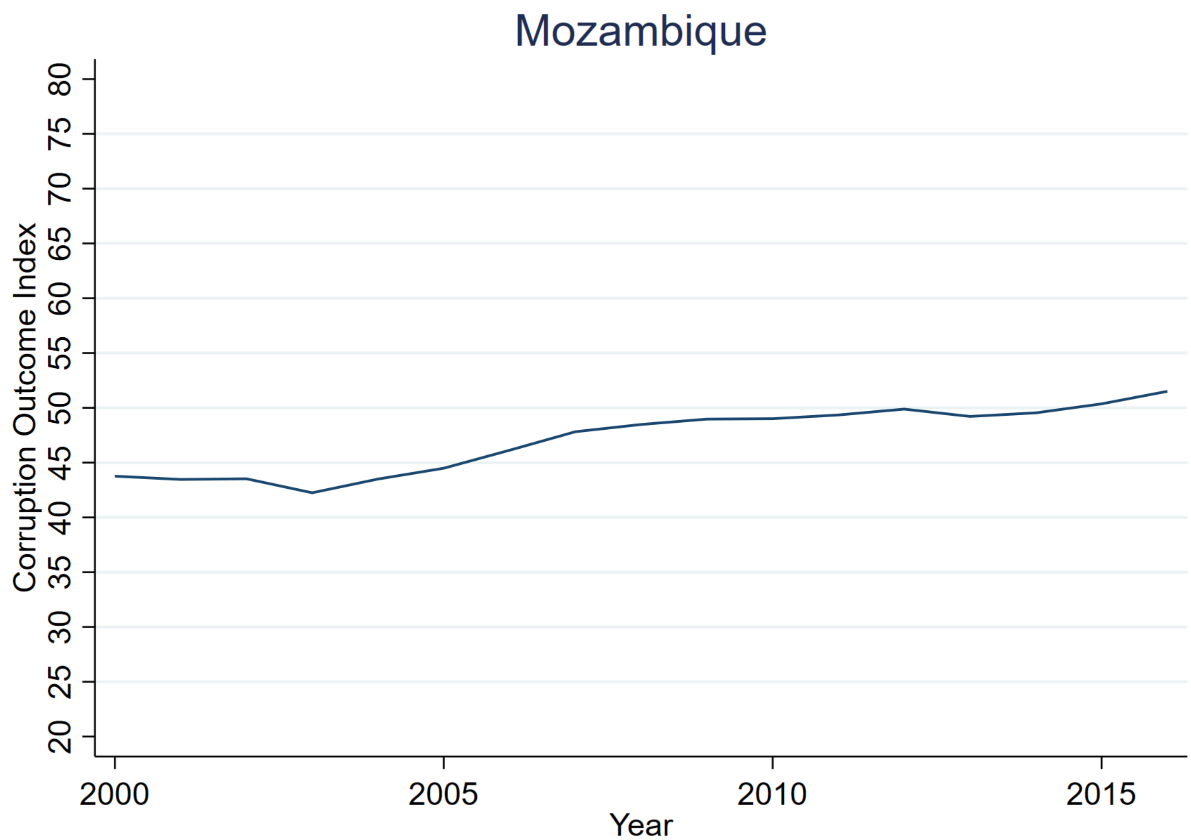
Micronesia_Federated States of
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

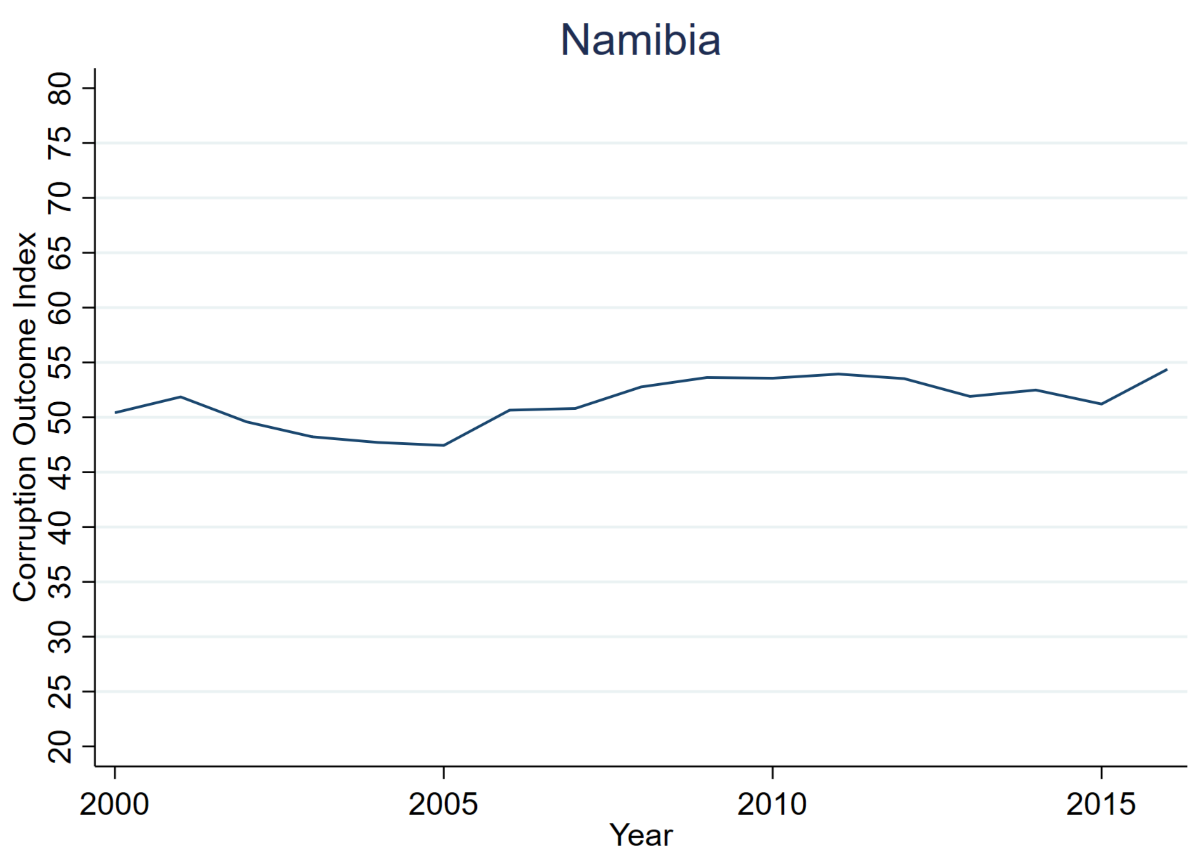


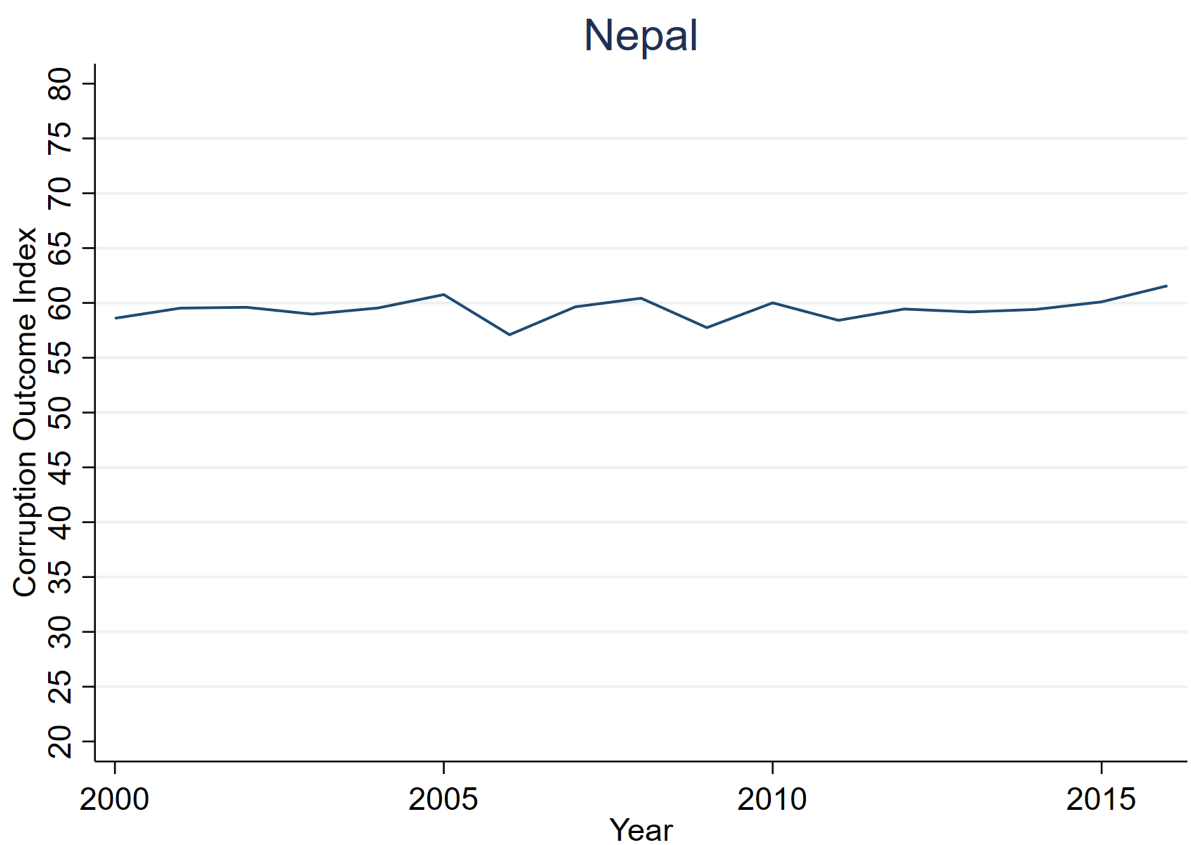
Mongolia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

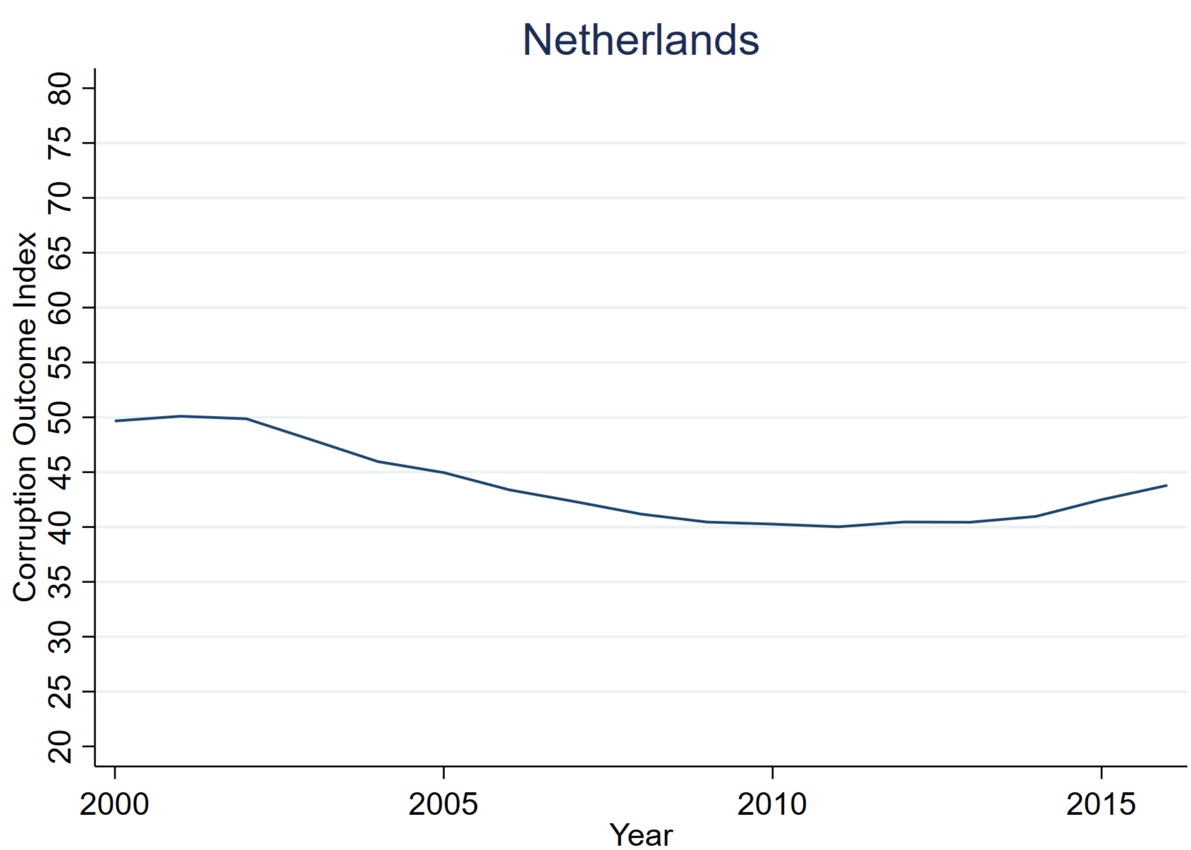

Moldova
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

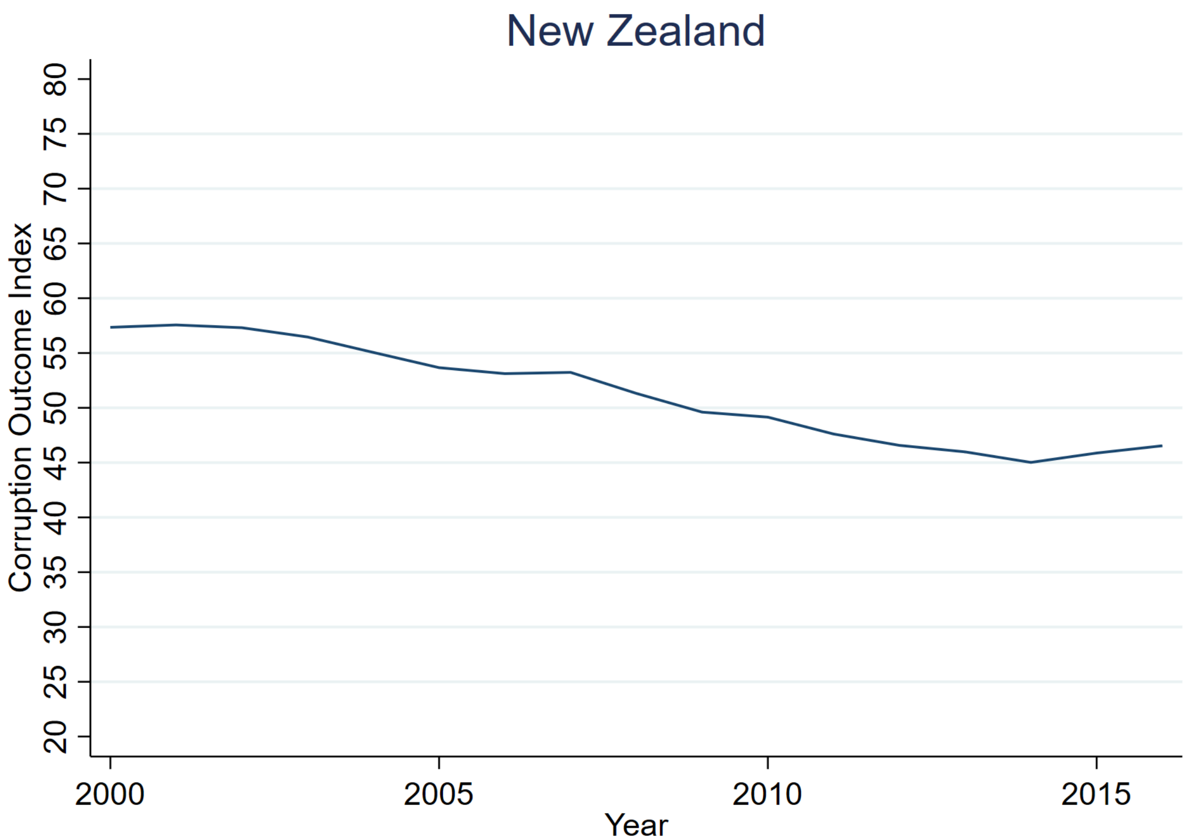

Morocco
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

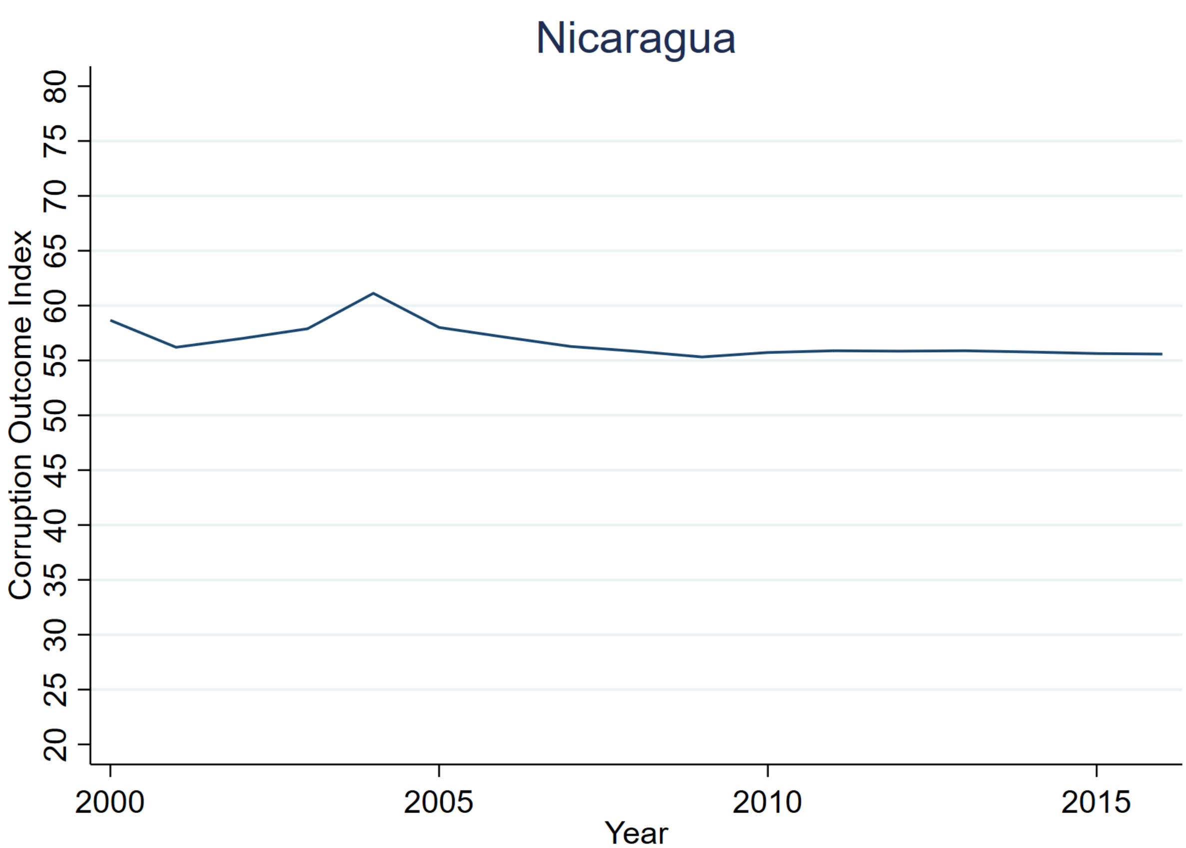

Mozambique
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

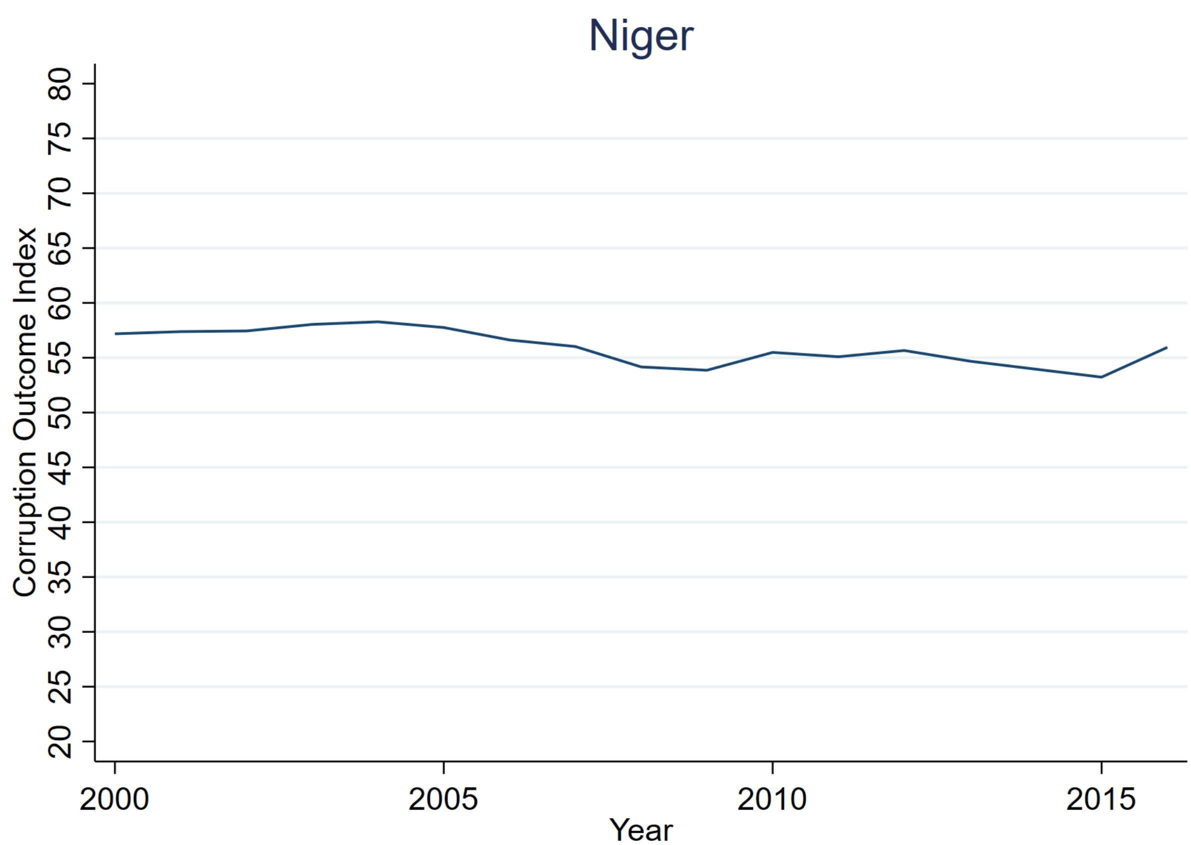

Namibia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

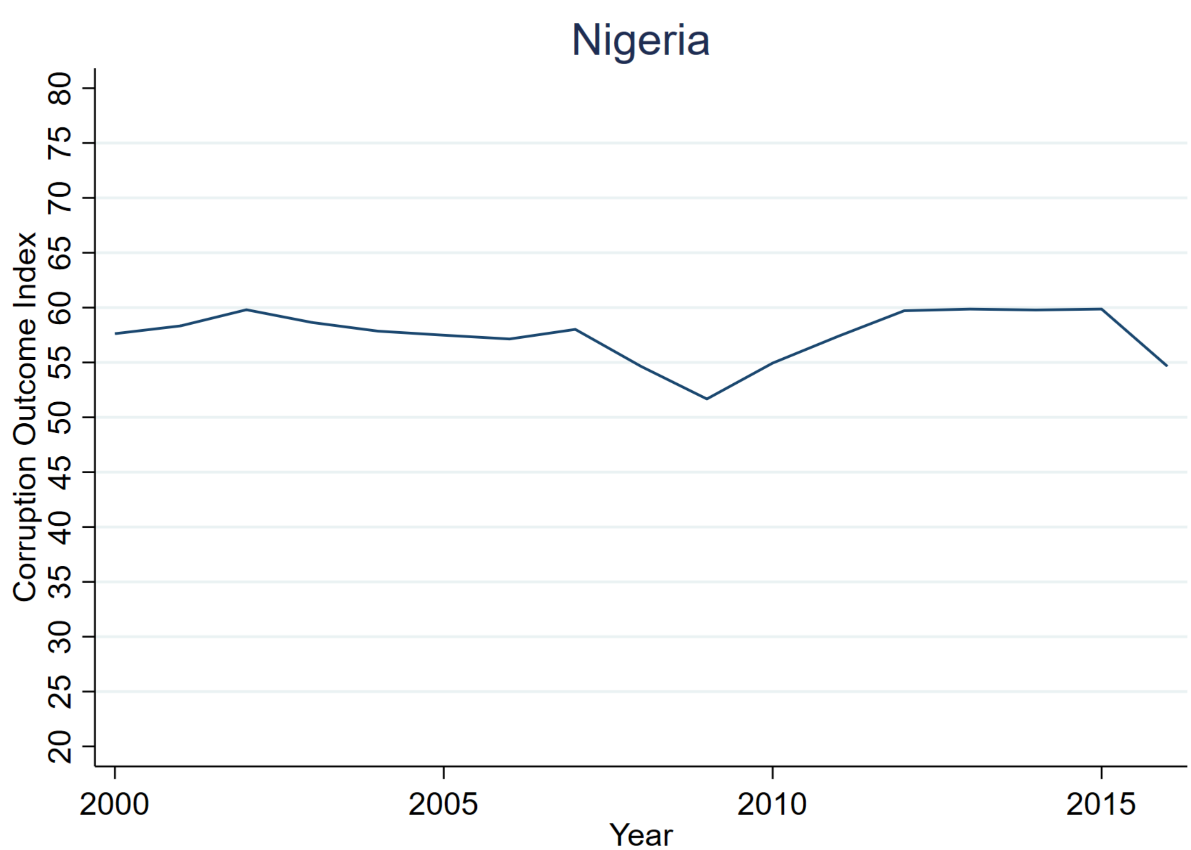

Nepal
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

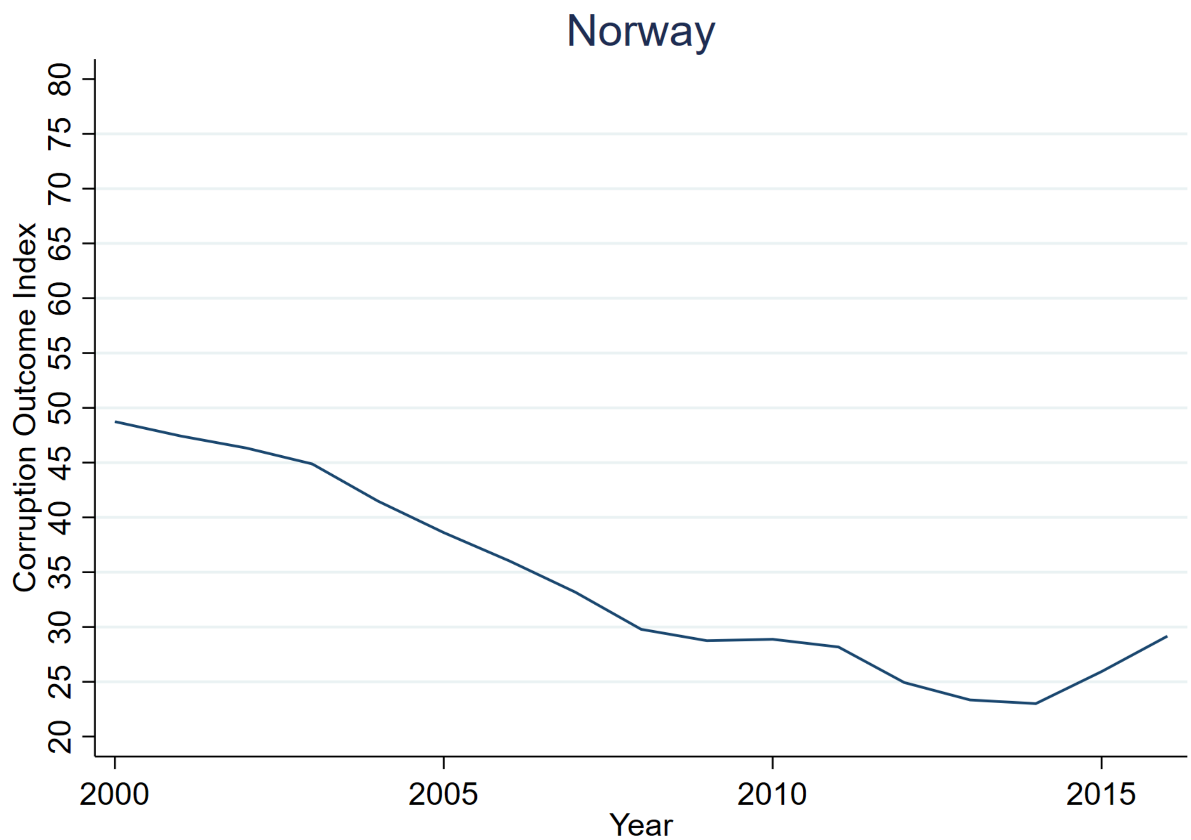

Netherlands
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

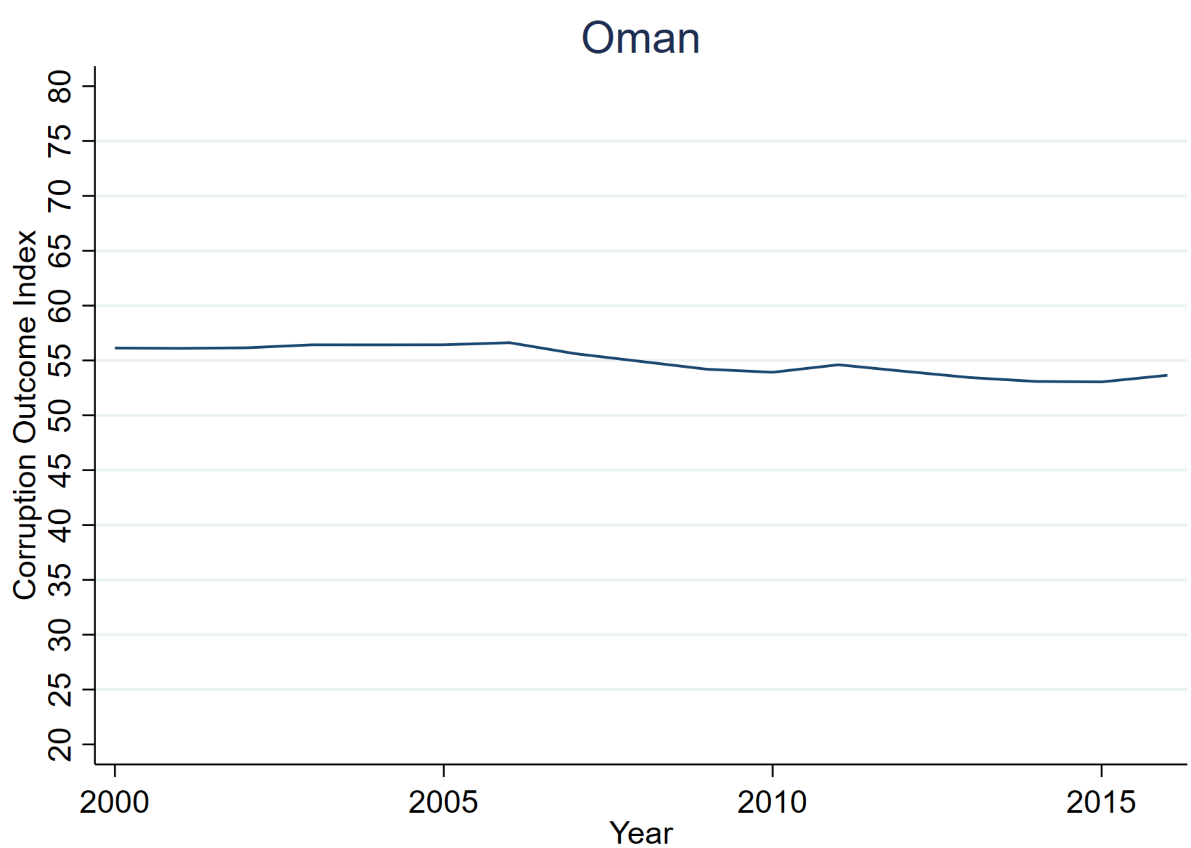

New Zealand
80
75
70
65
60
55
50
45
40
35
30
25
20
Corruption Outcome Index
2000
2005
2010
2015
Year

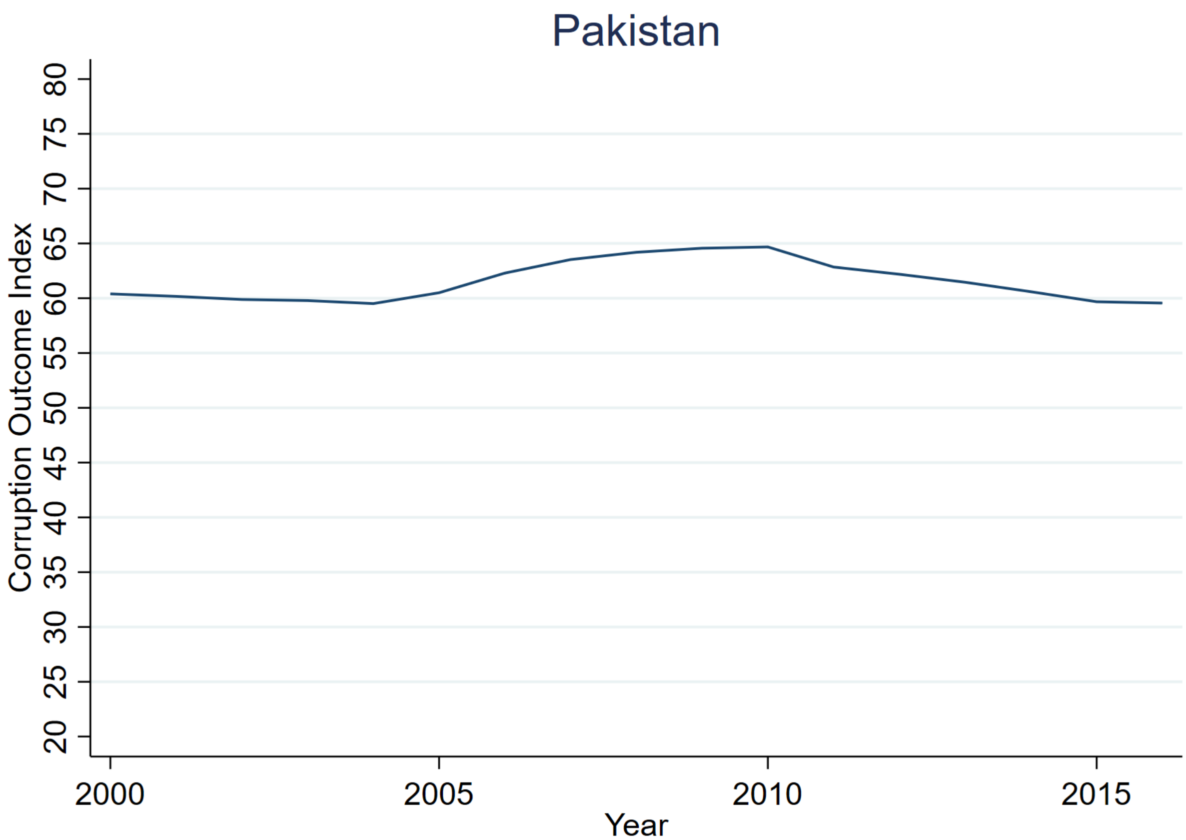

Nicaragua
80
75
70
65
60
55
50
45
40
35
30
25
20
Corruption Outcome Index
2000
2005
2010
2015
Year

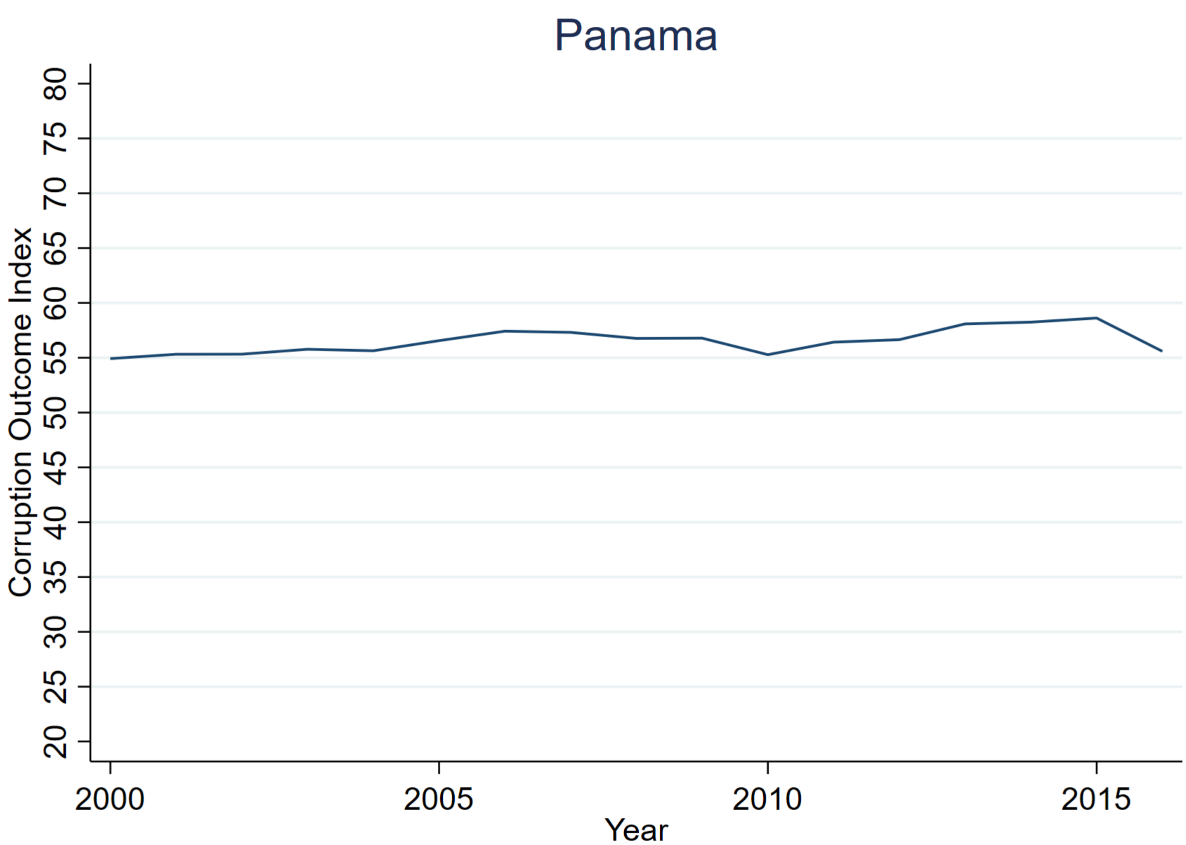

Niger
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

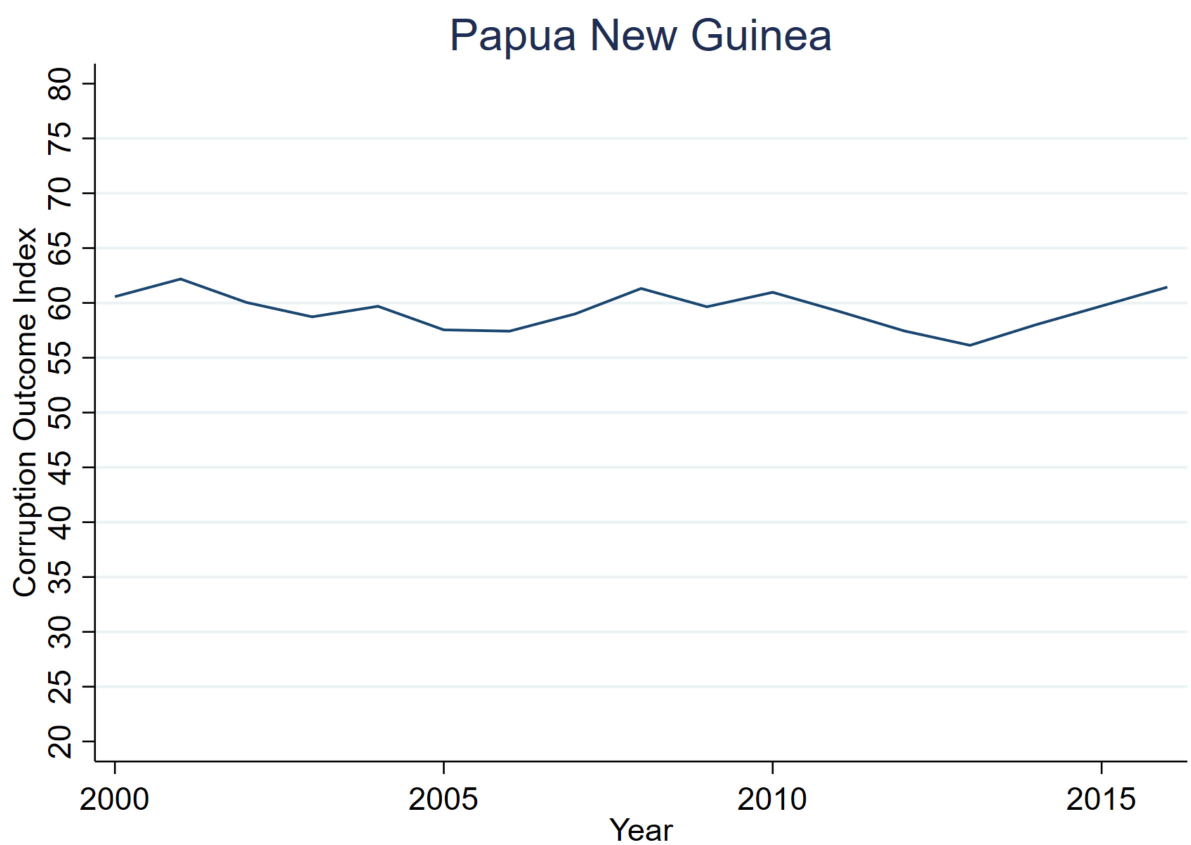

Nigeria
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

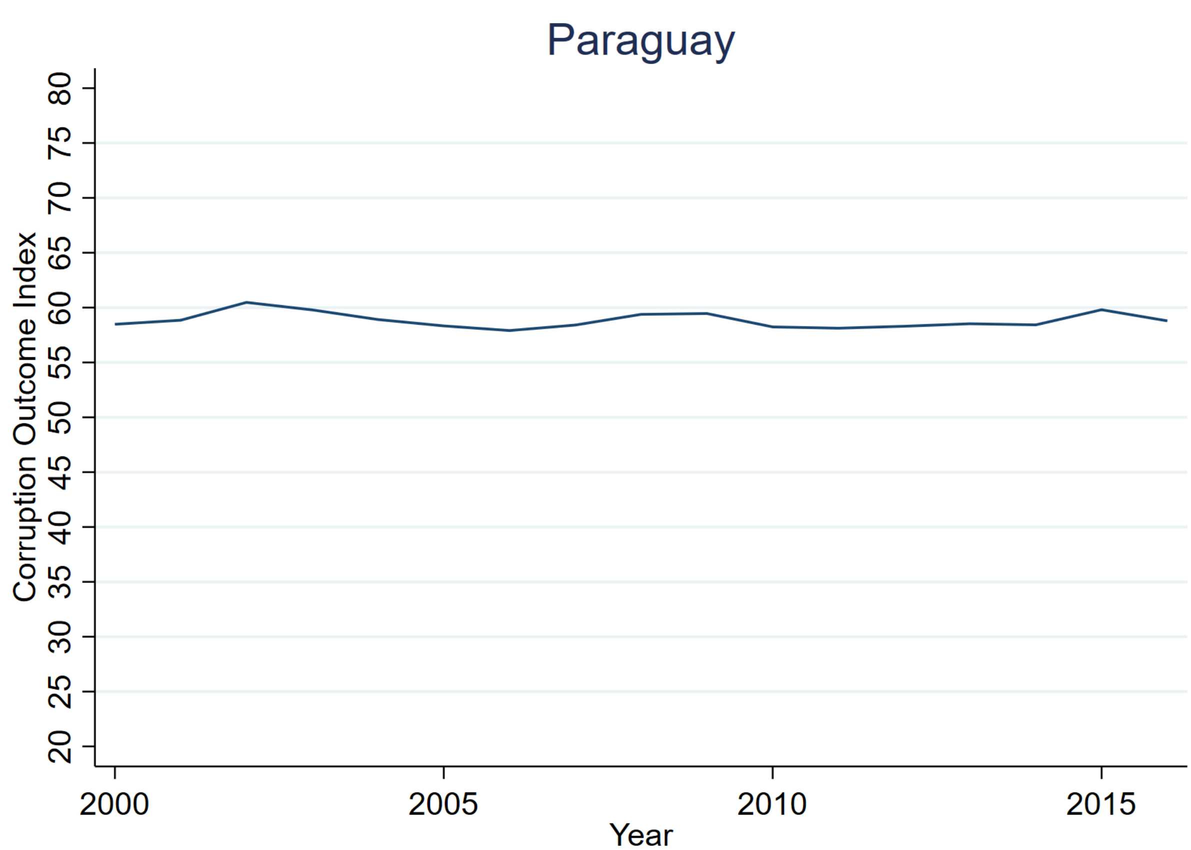

Norway
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

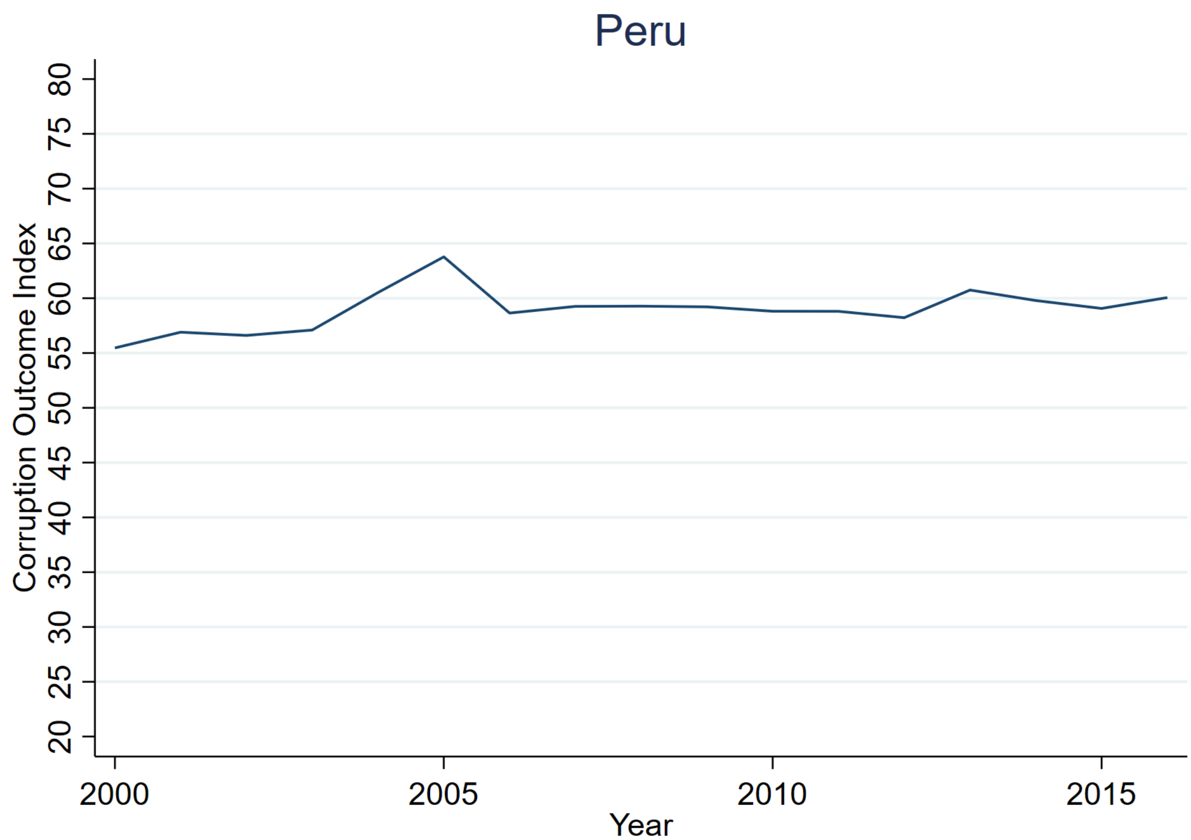

Oman
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

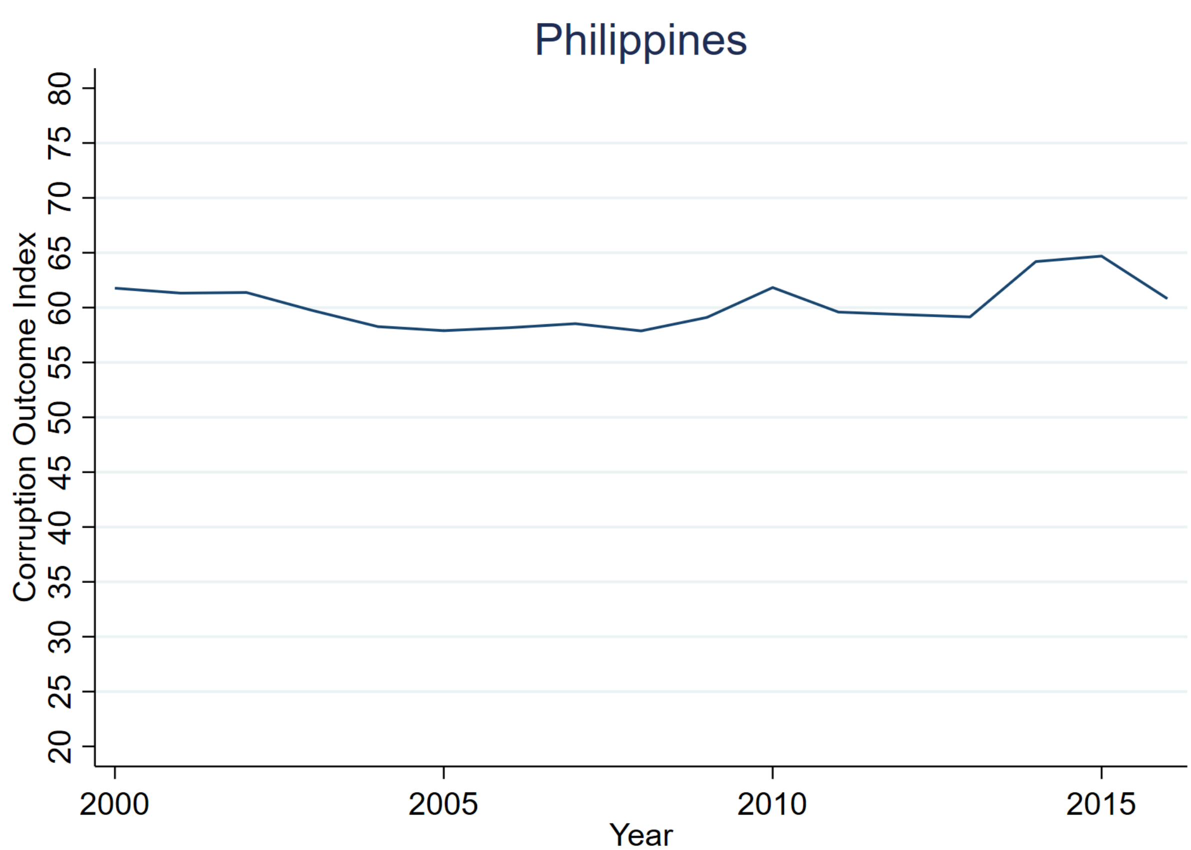

Pakistan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

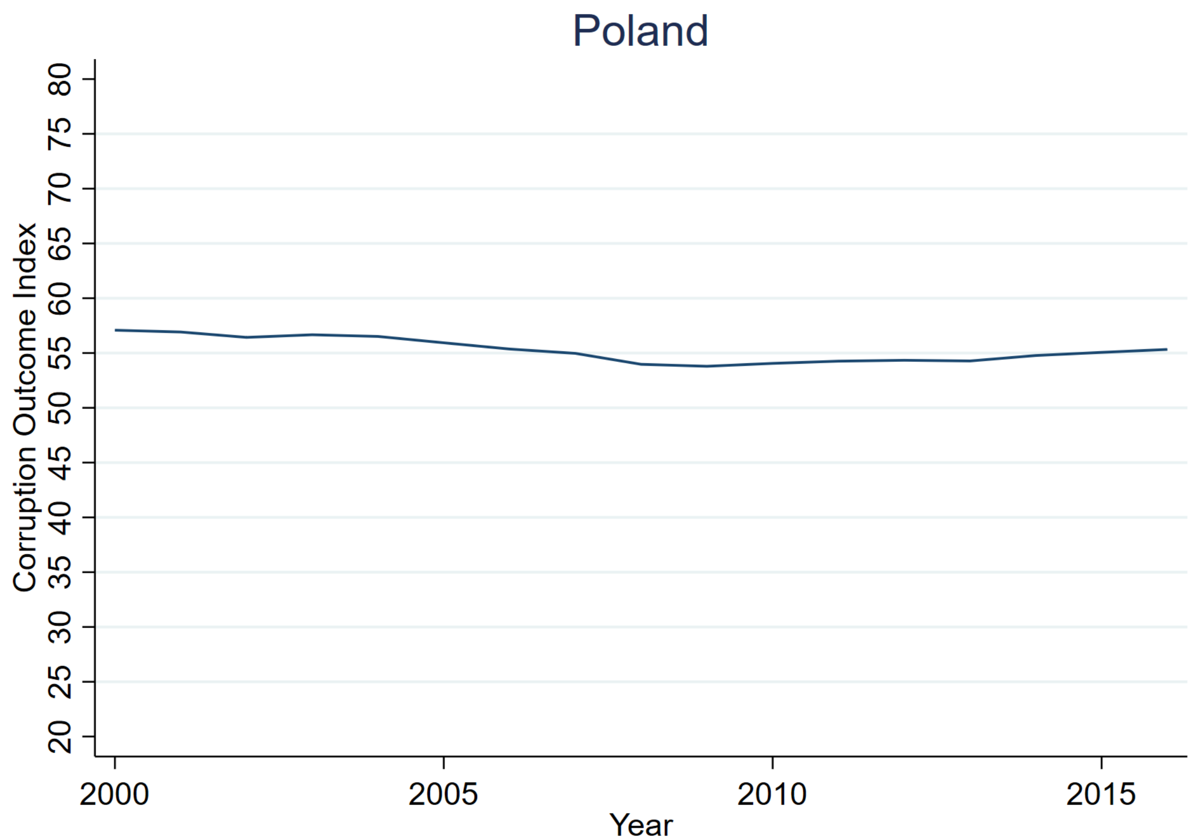

Panama
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

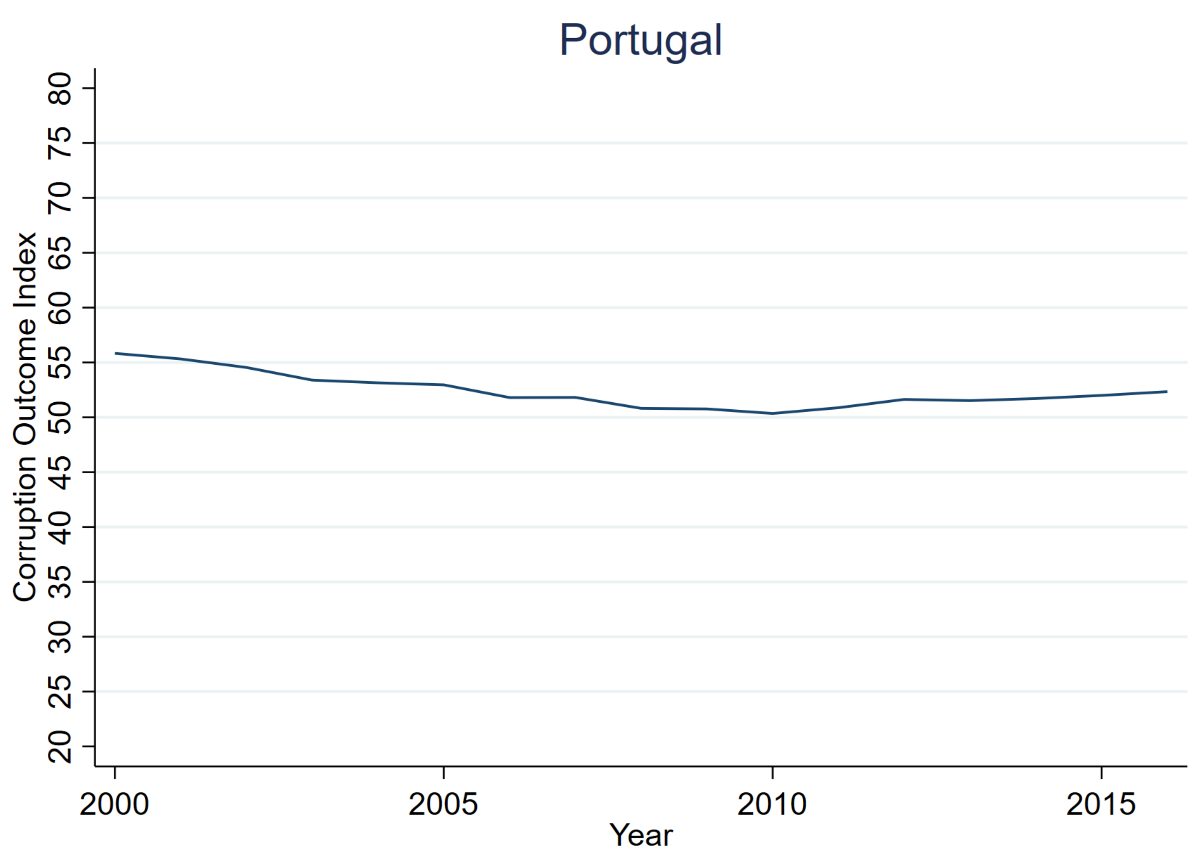

Papua New Guinea
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

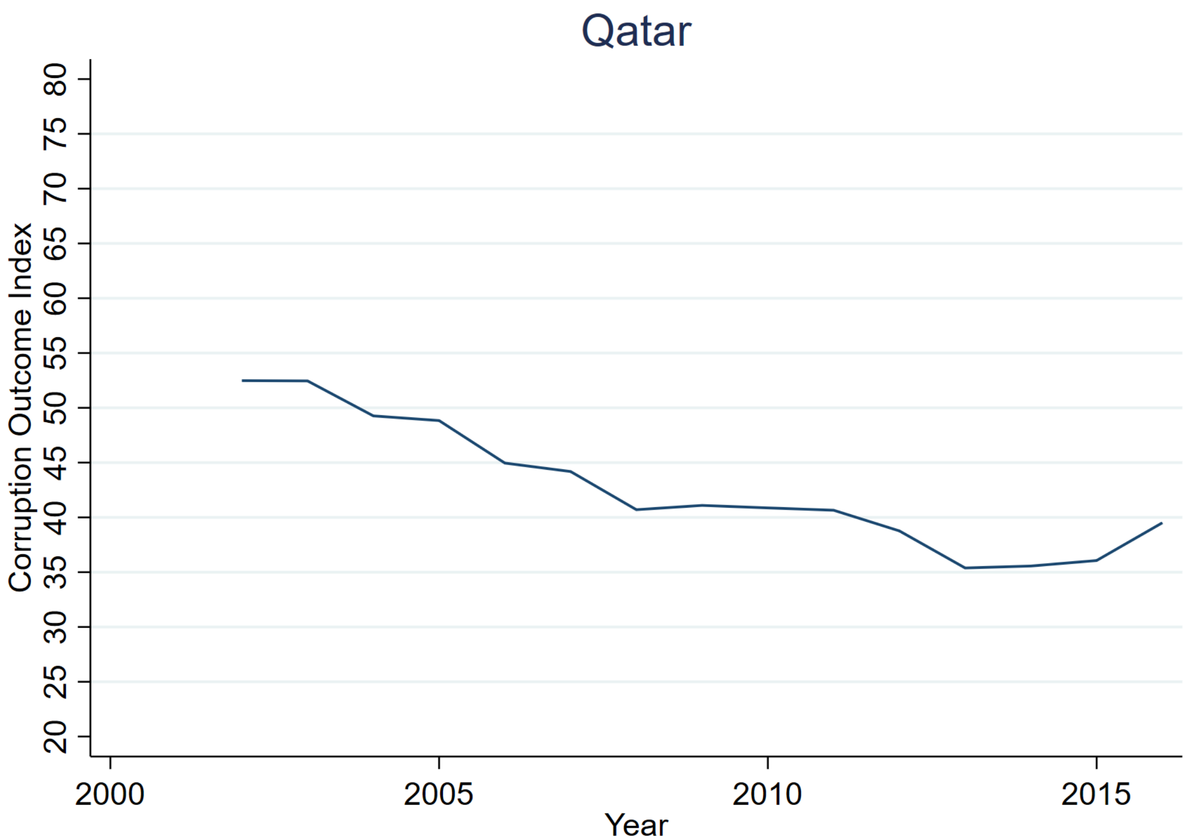

Paraguay
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

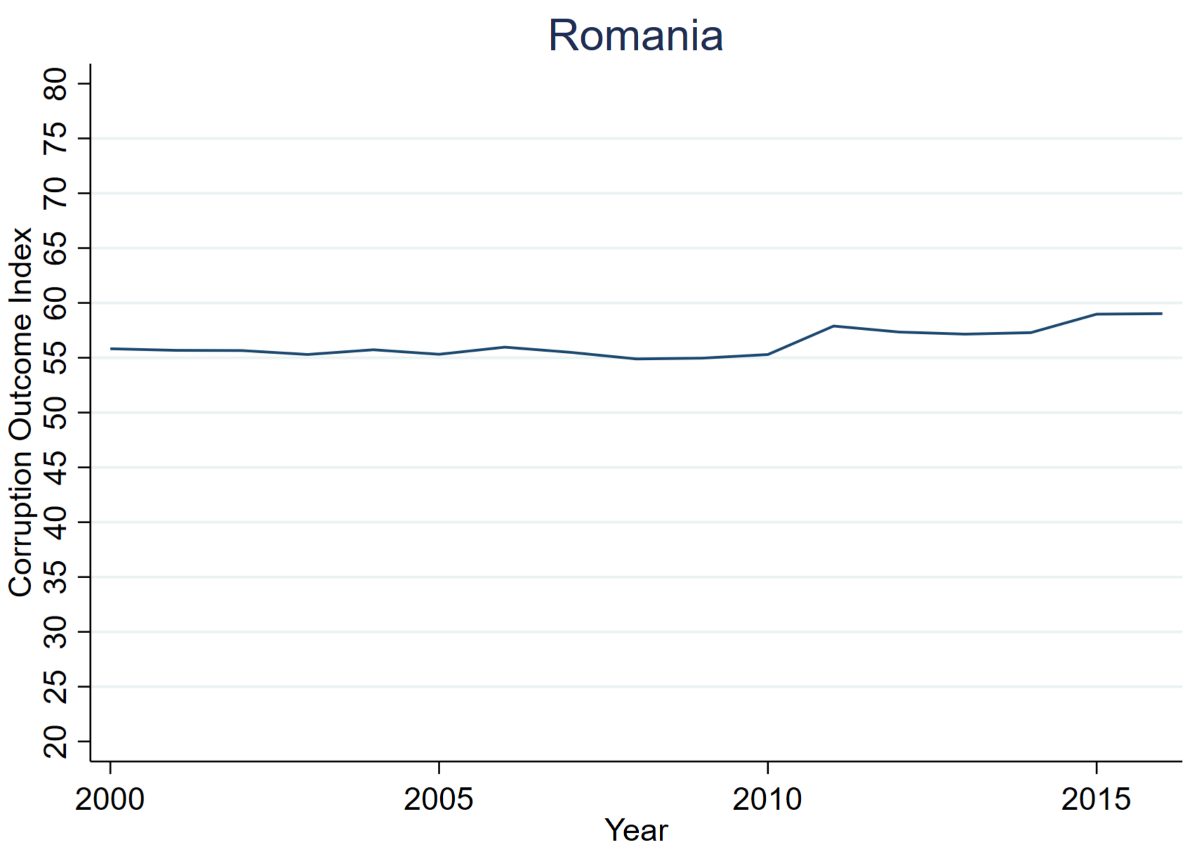

Peru
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Philippines
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

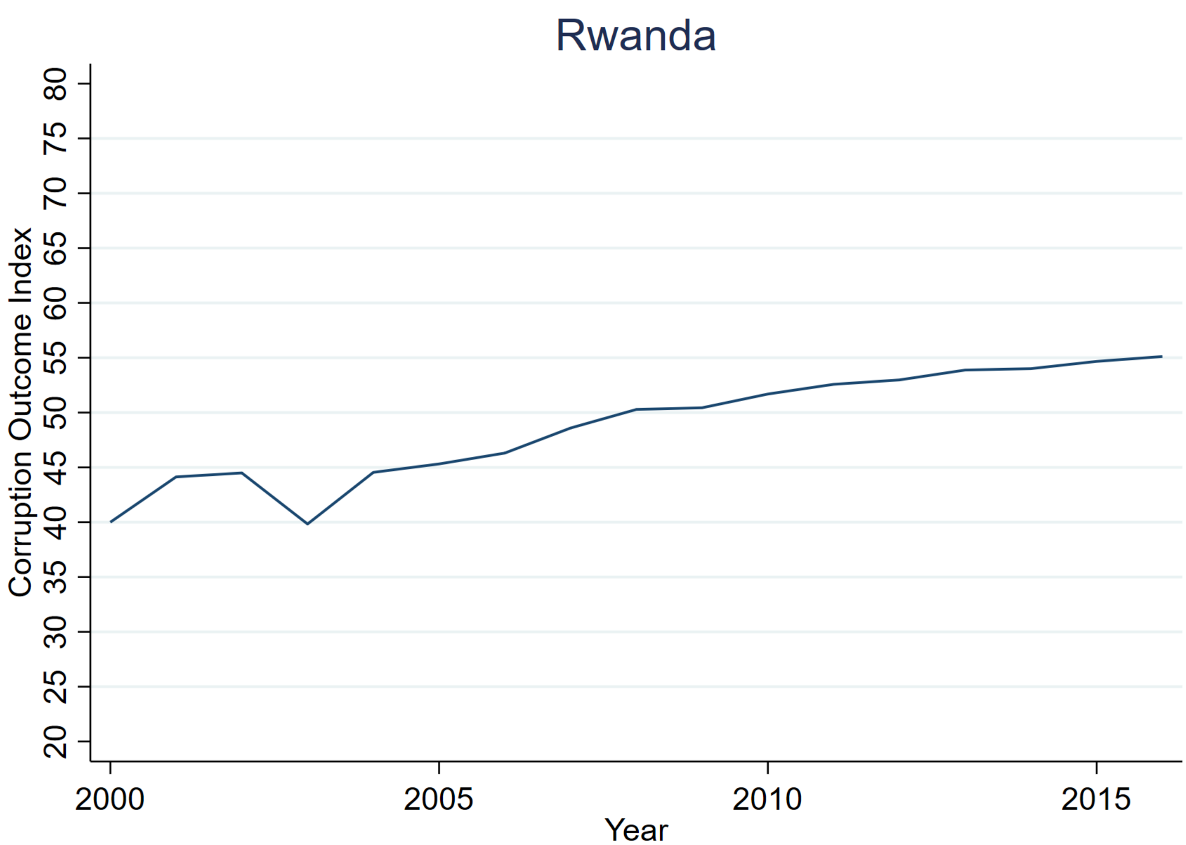

Poland
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Portugal
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

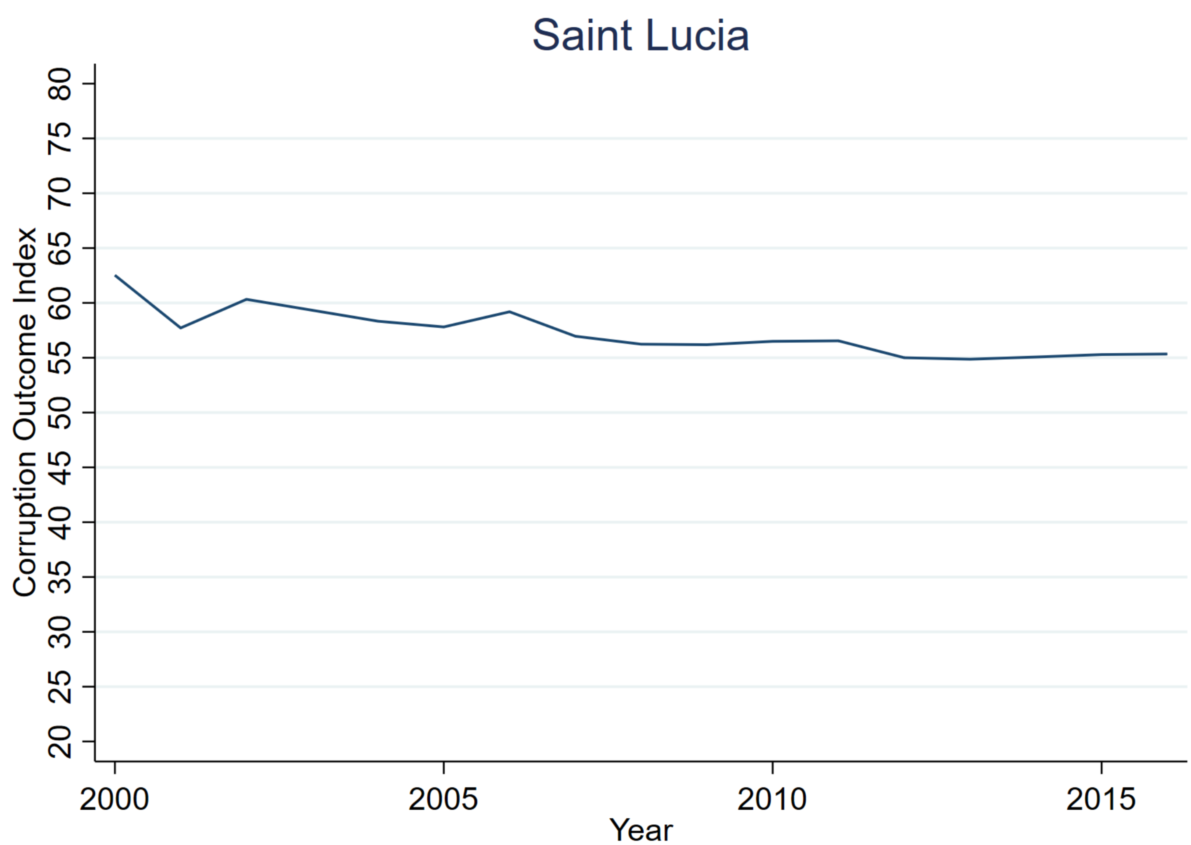

Qatar
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

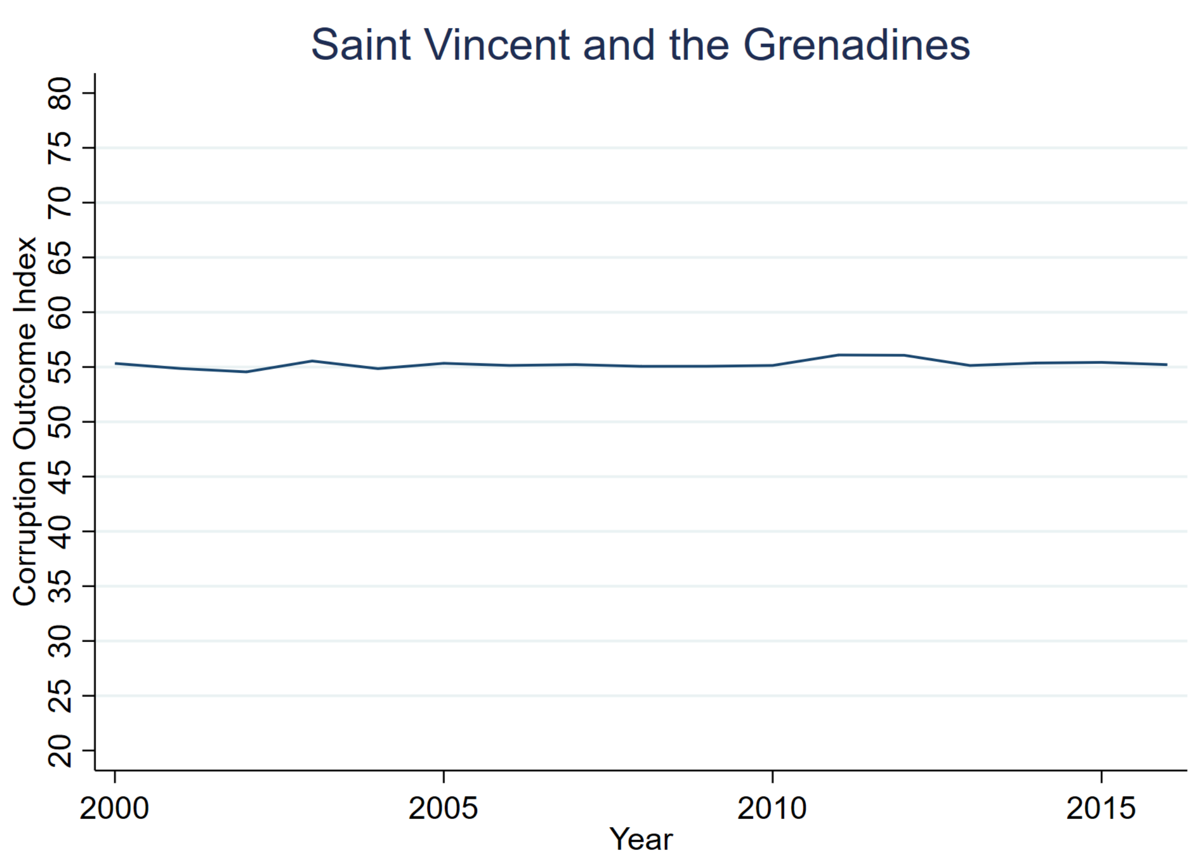

Romania
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

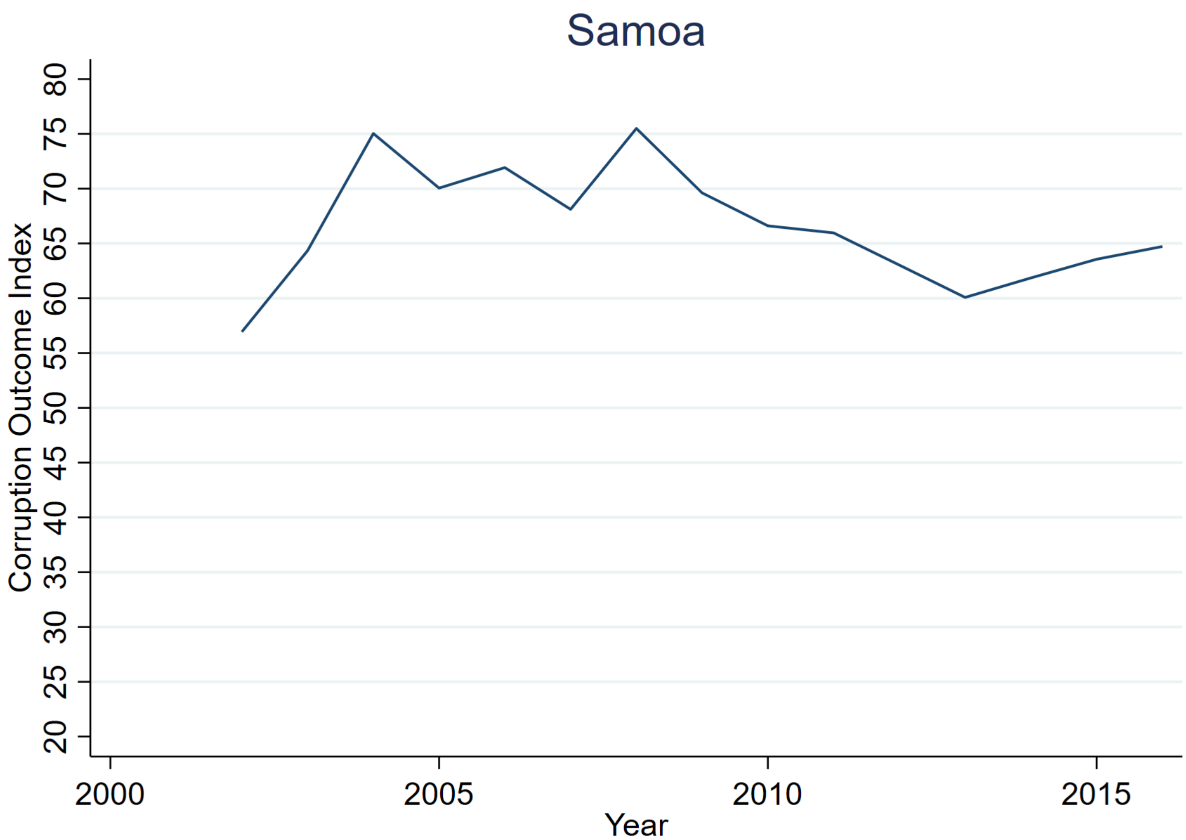
Russia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

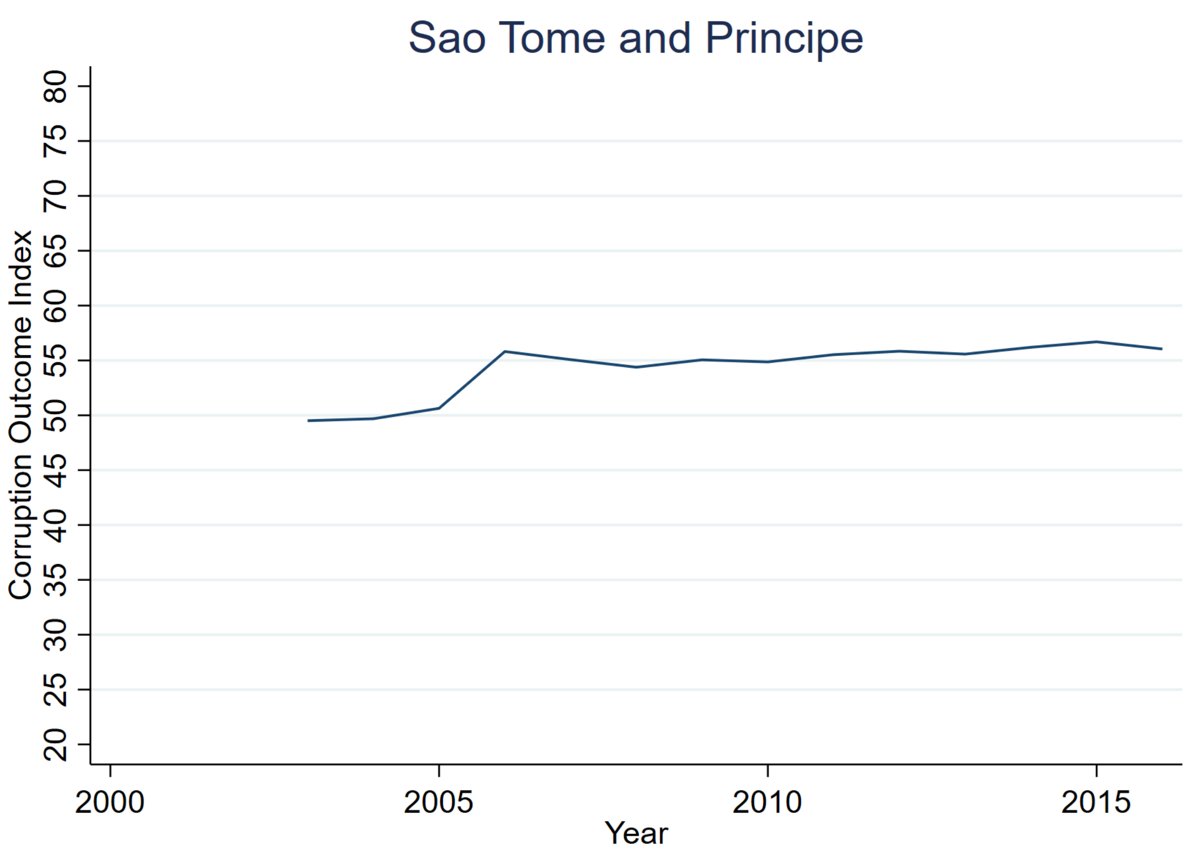


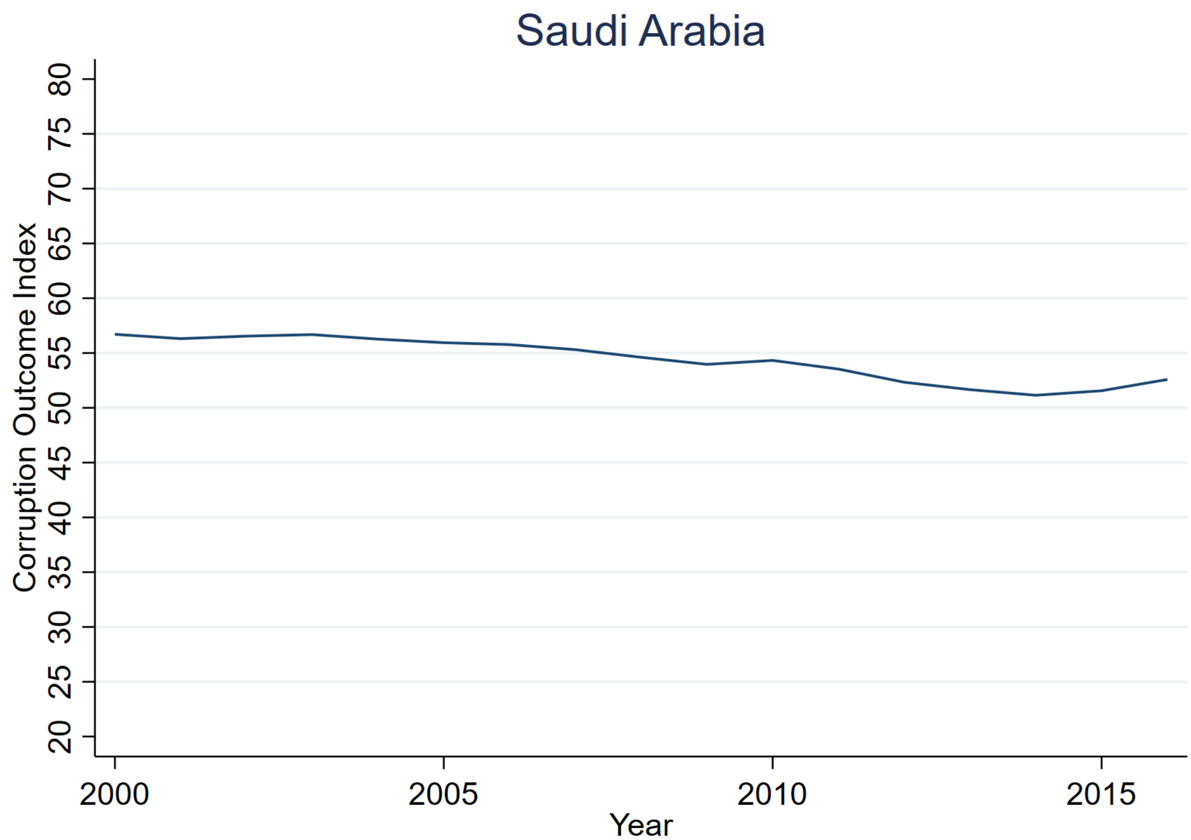
Rwanda
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

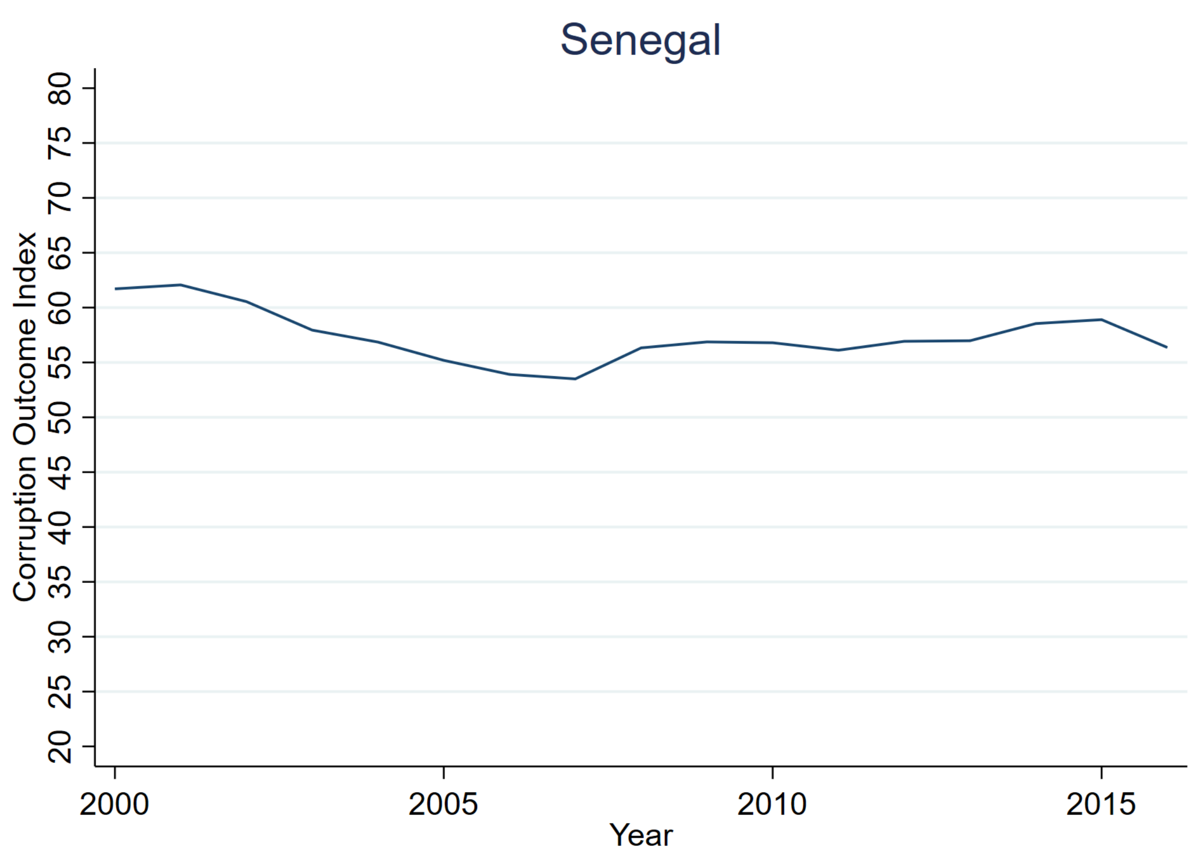

Saint Lucia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

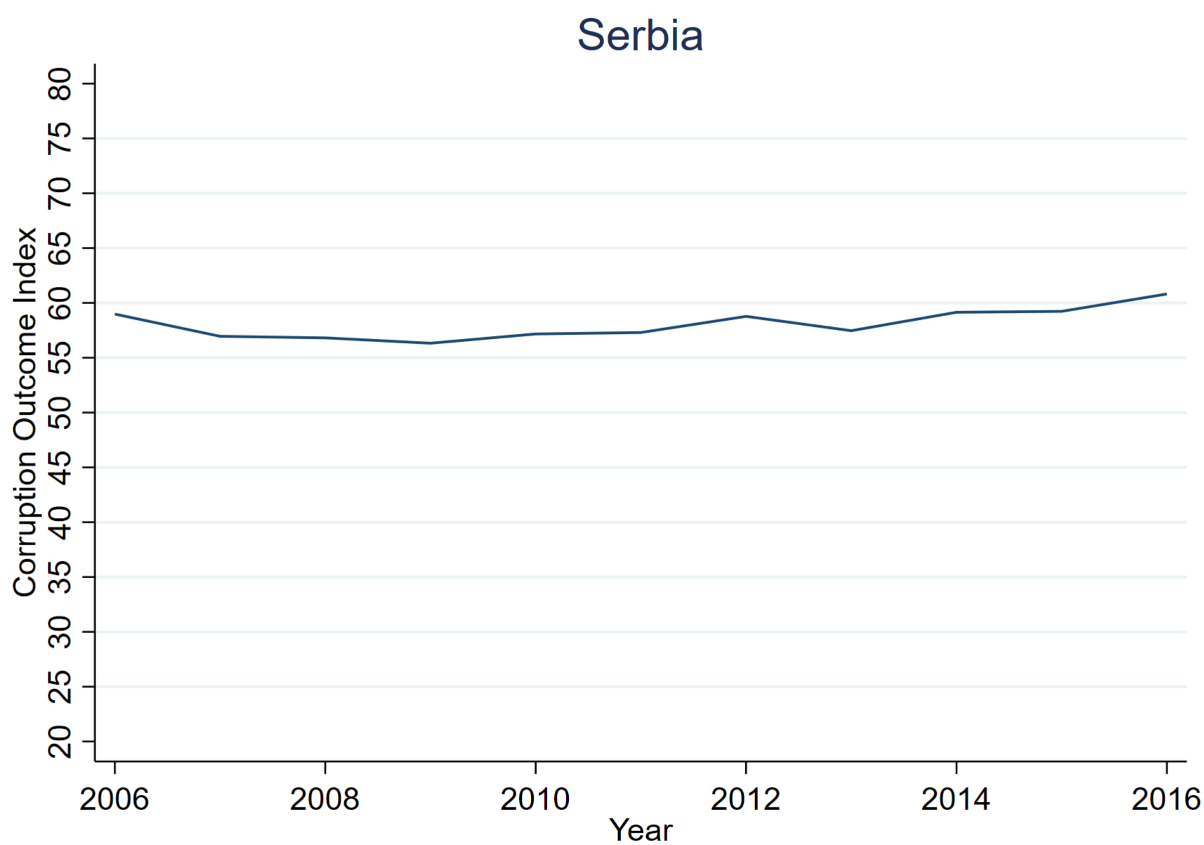

Saint Vincent and the Grenadines
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

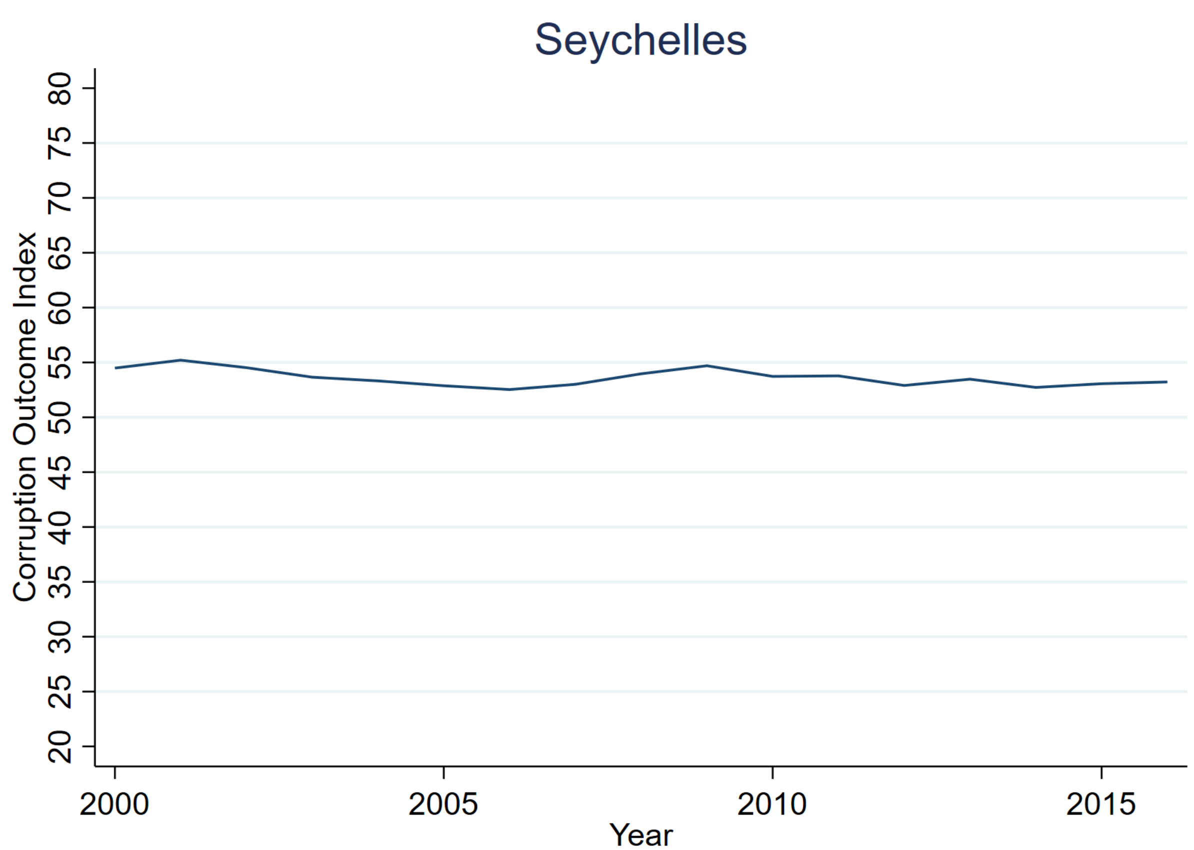

Samoa
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

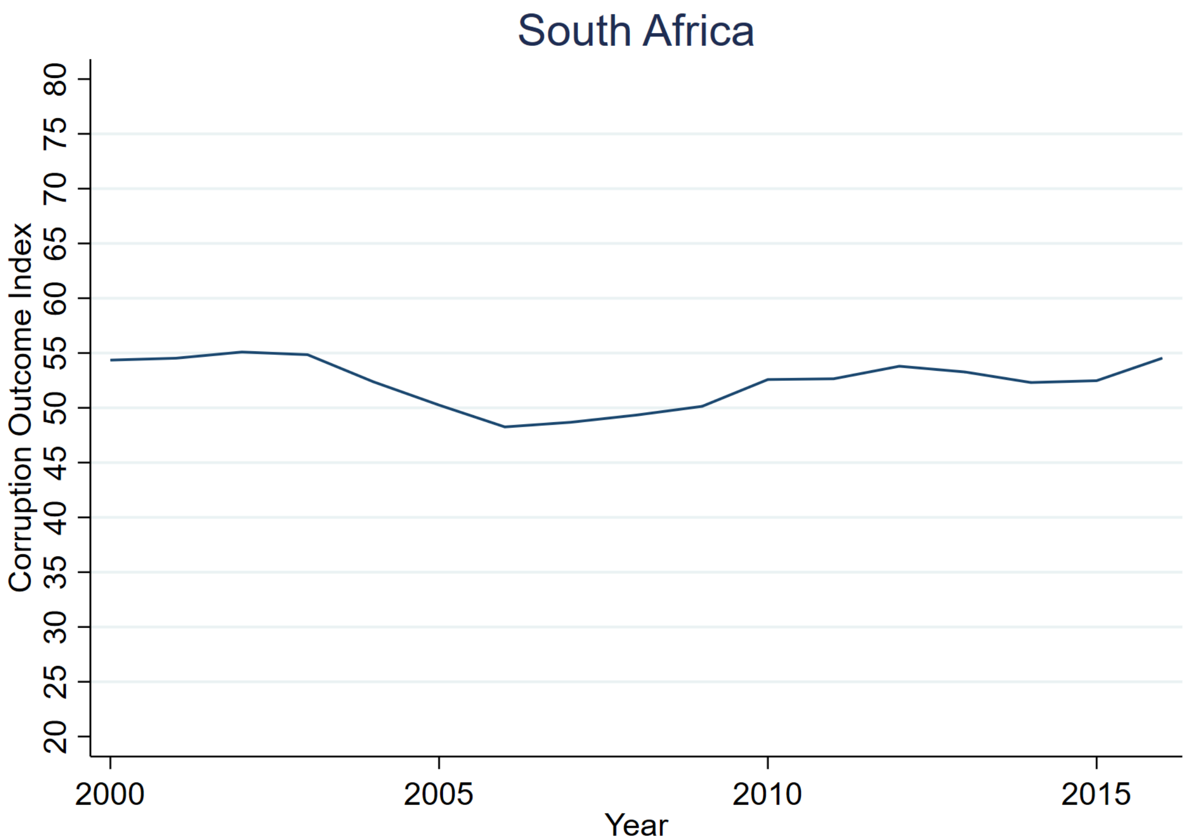

Sao Tome and Principe
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

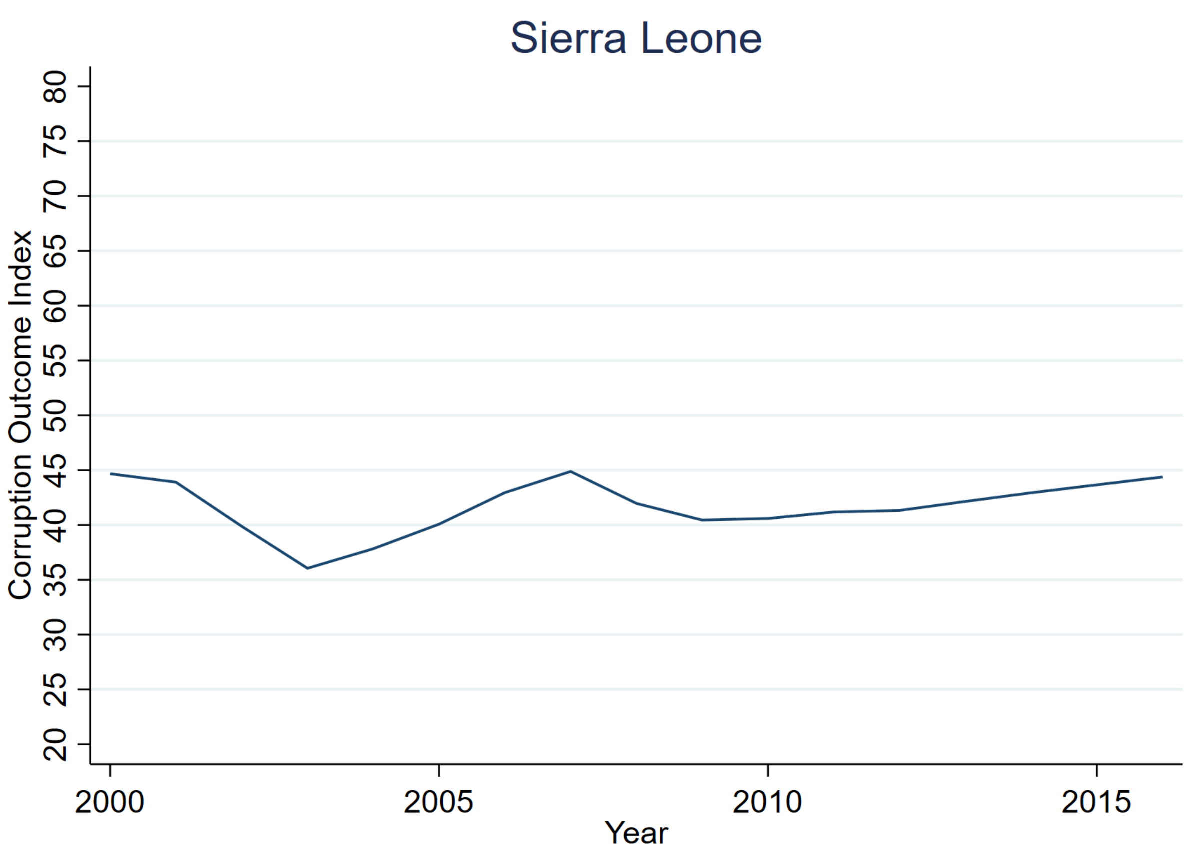

Saudi Arabia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

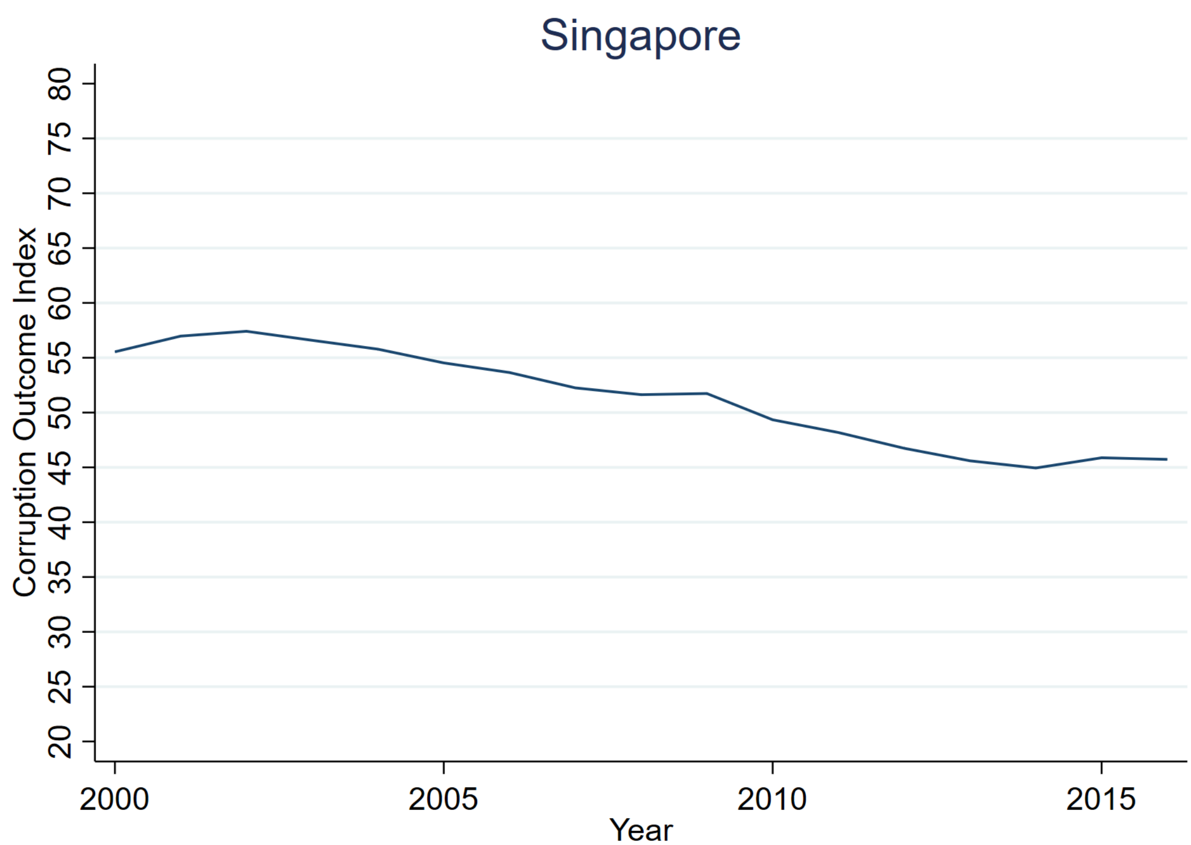

Senegal
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

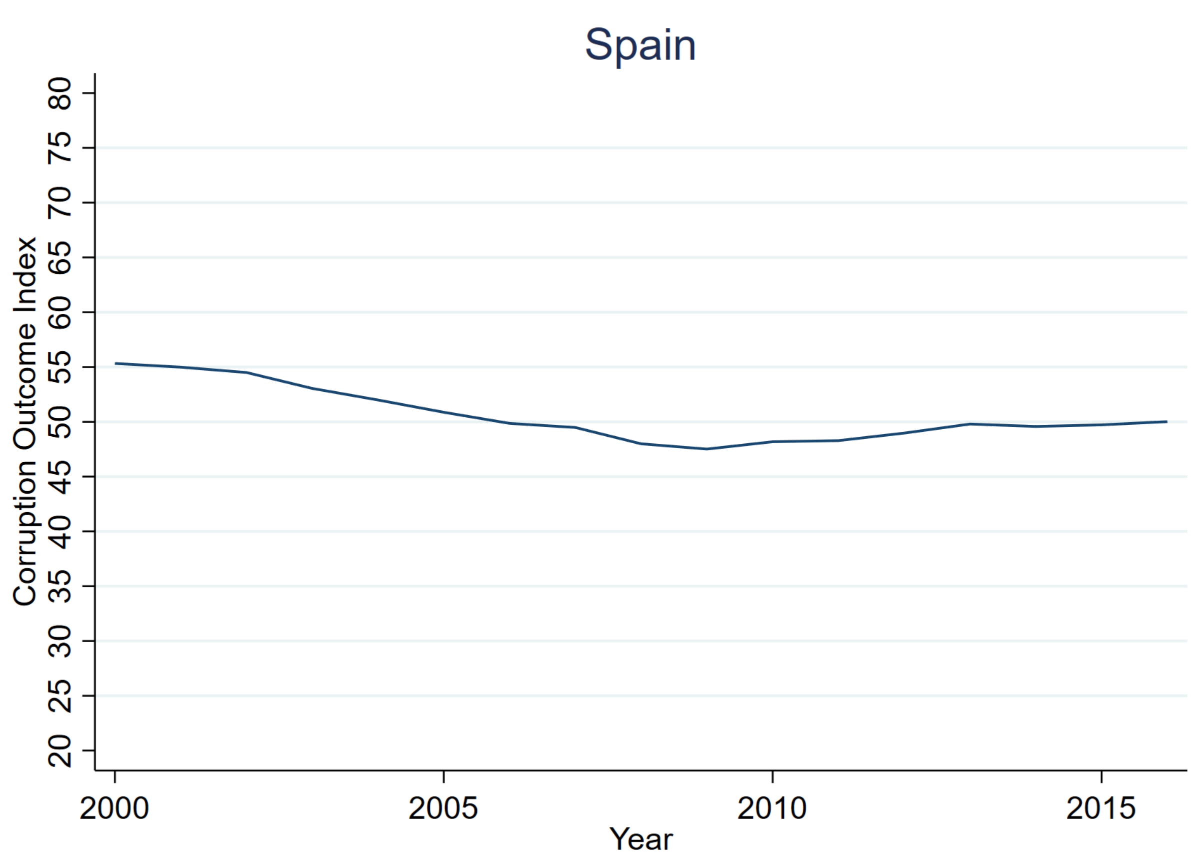

Serbia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2006
2008
2010
2012
2014
2016
Year

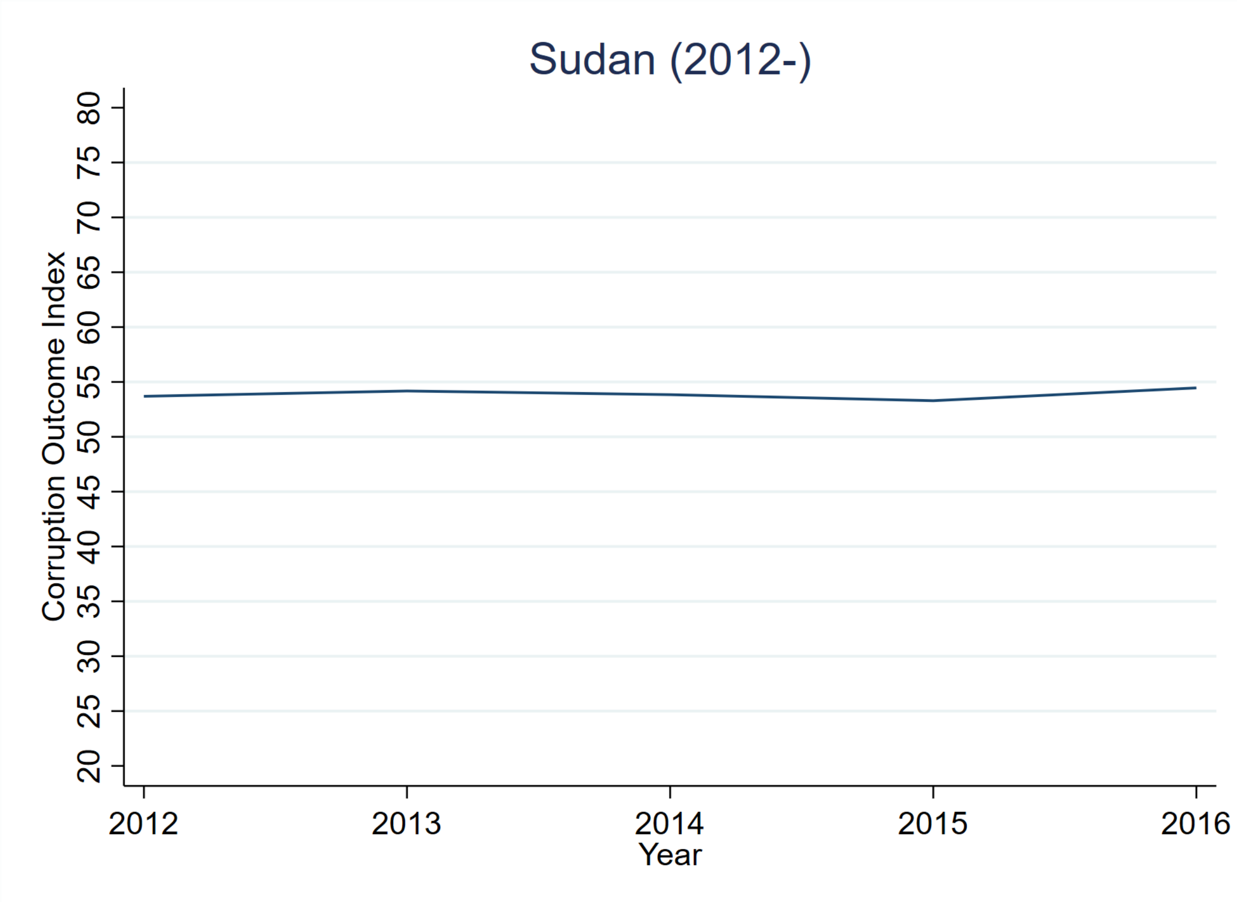

Seychelles
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

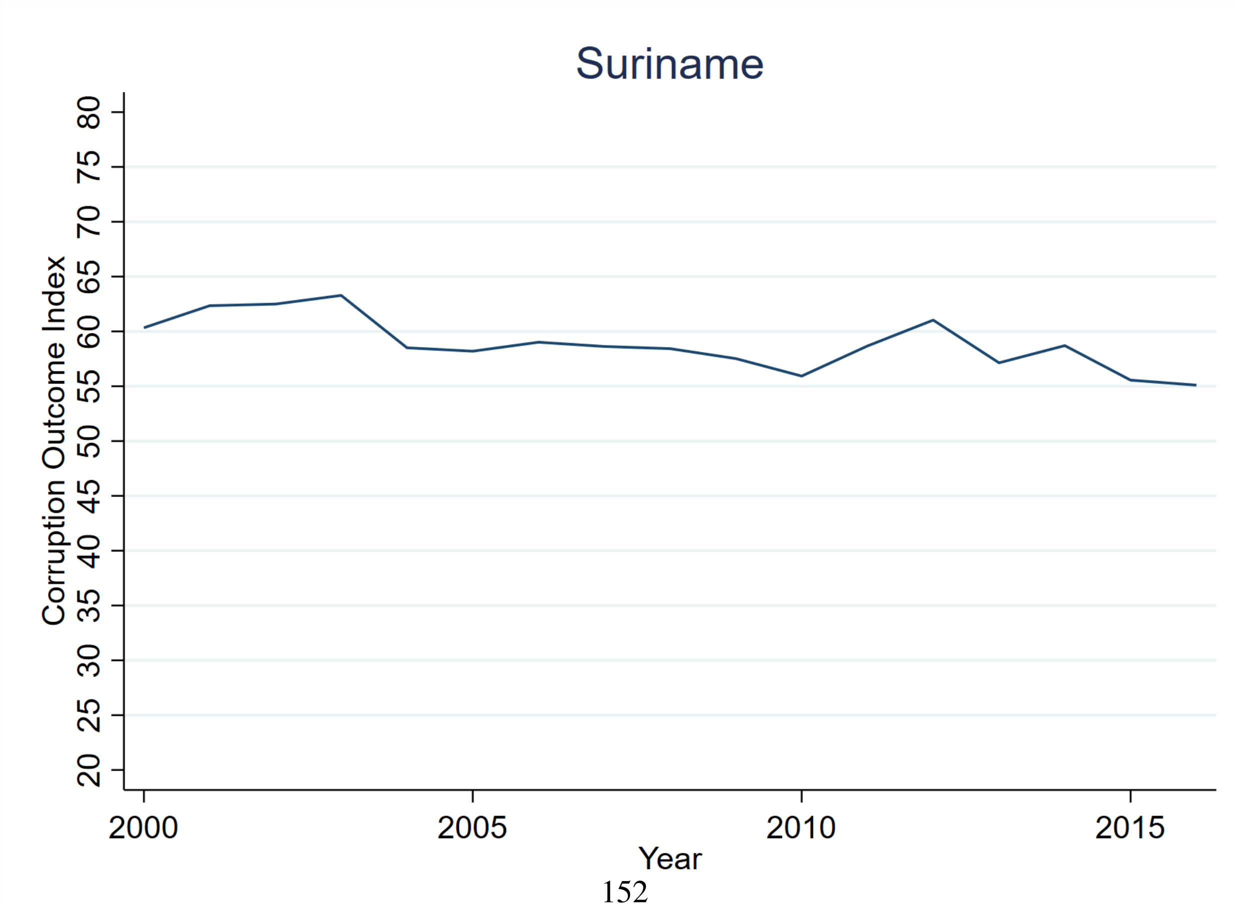

South Africa
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

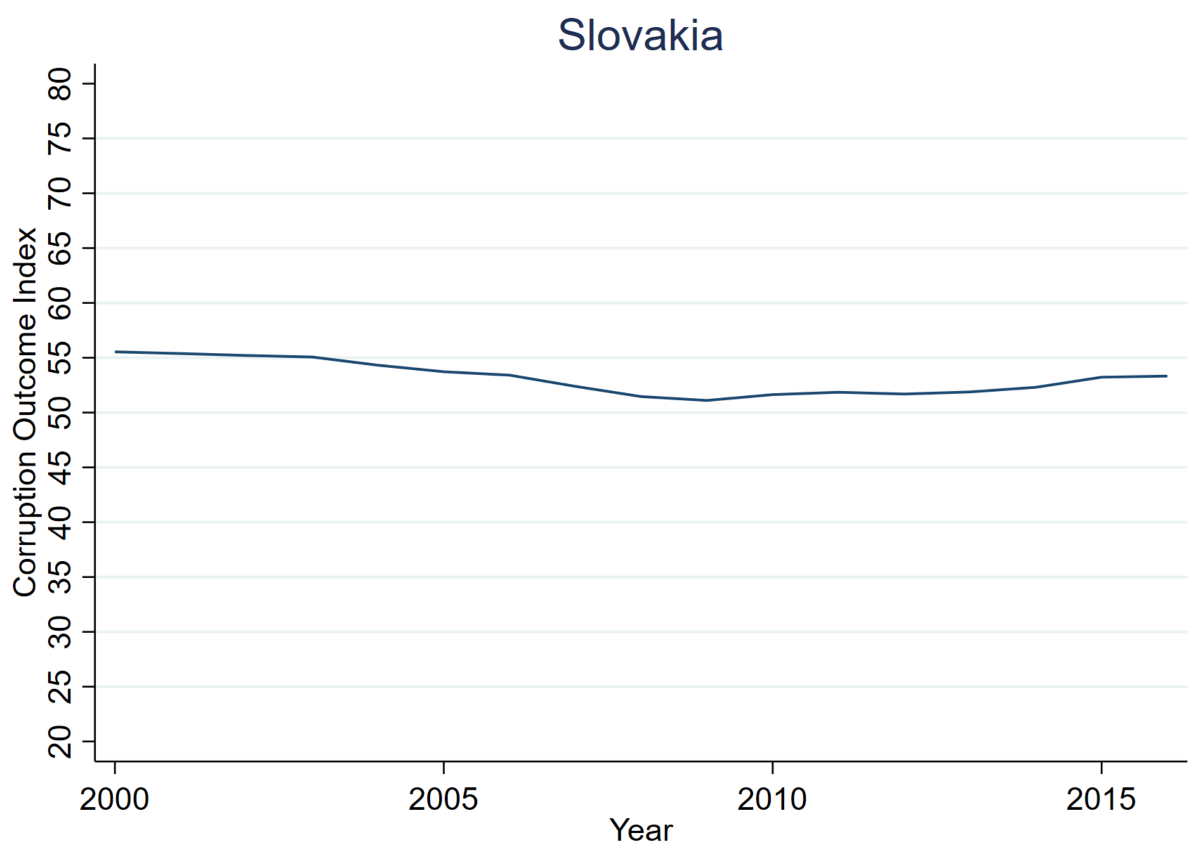

Sierra Leone
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

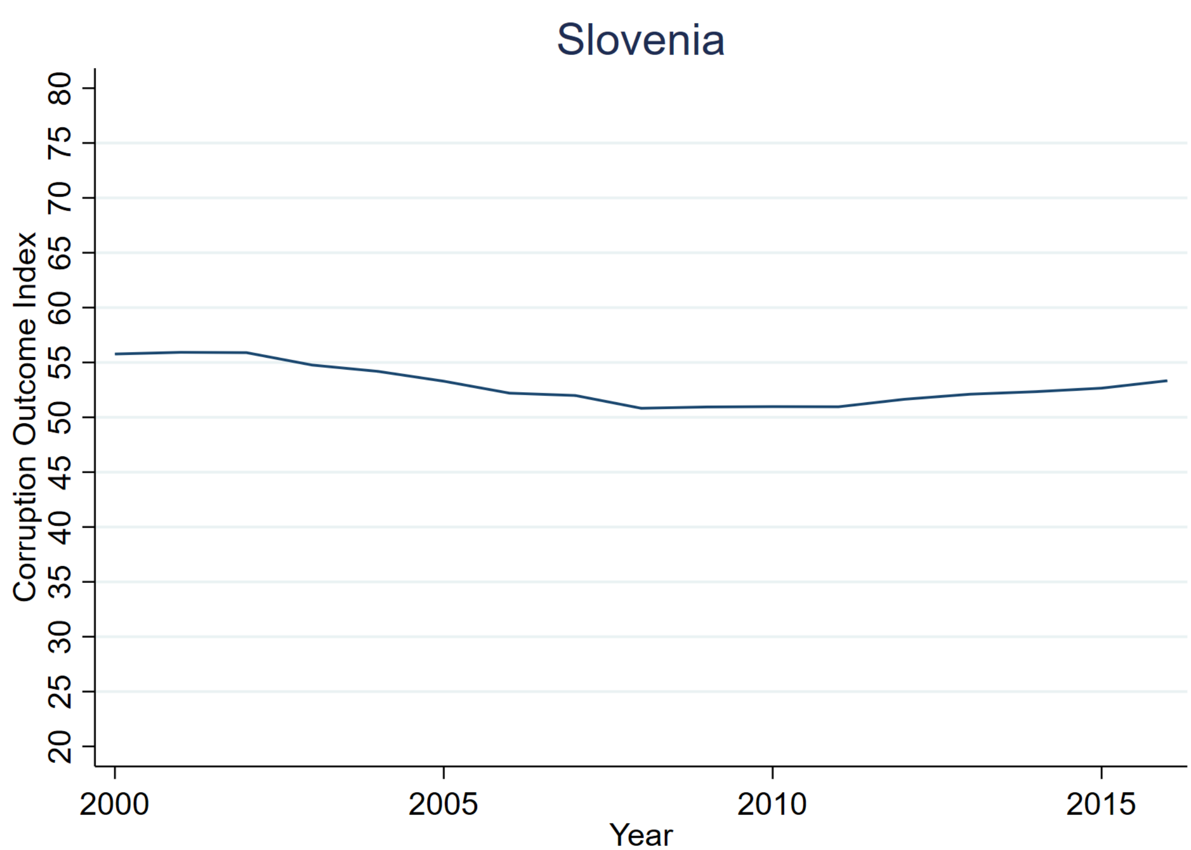

Singapore
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

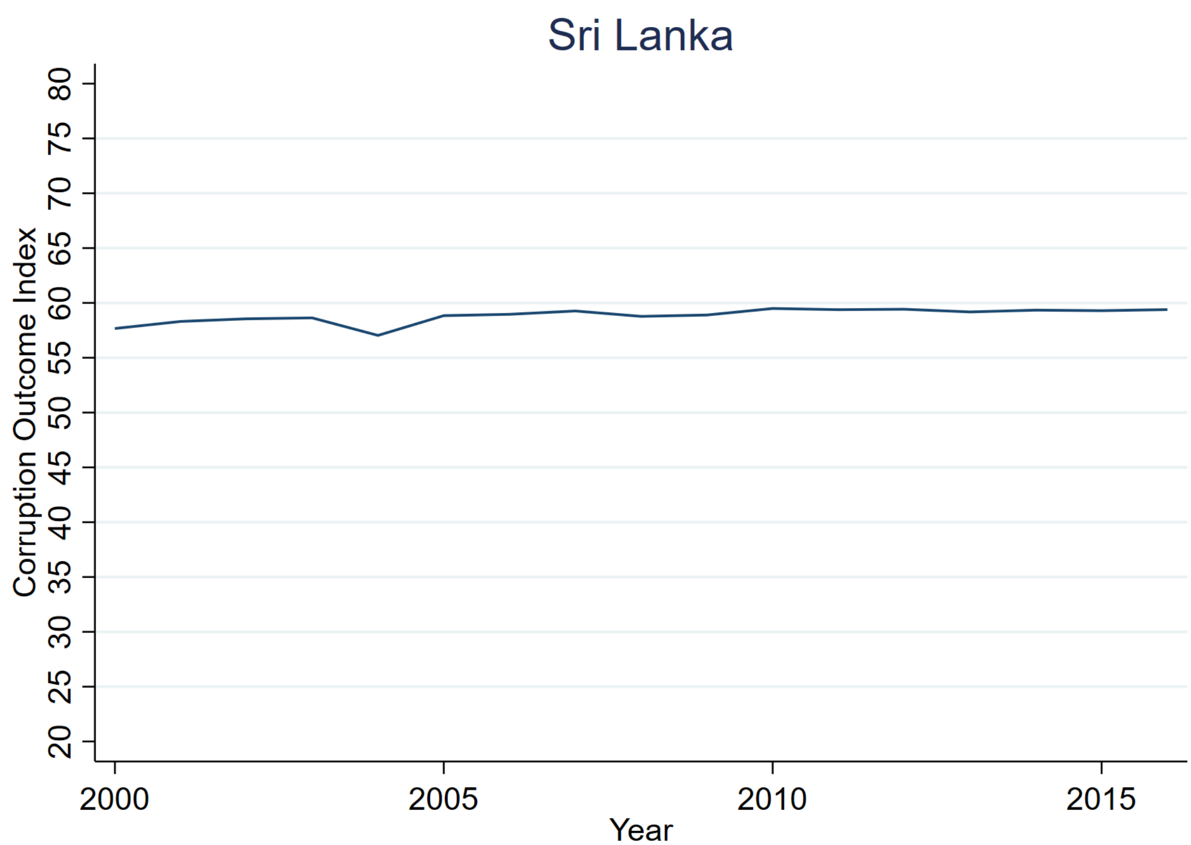

Spain
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

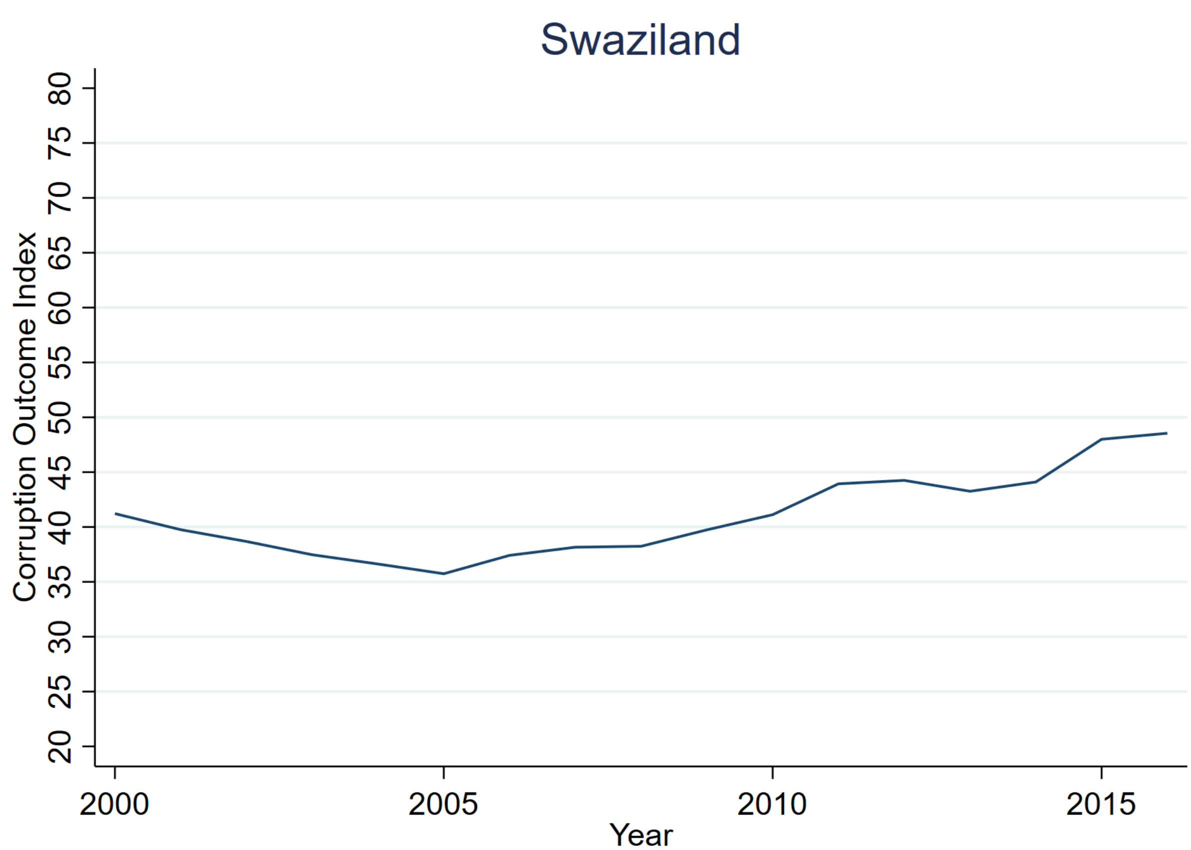

Sudan (2012-)
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2012
2013
2014
2015
2016
Year

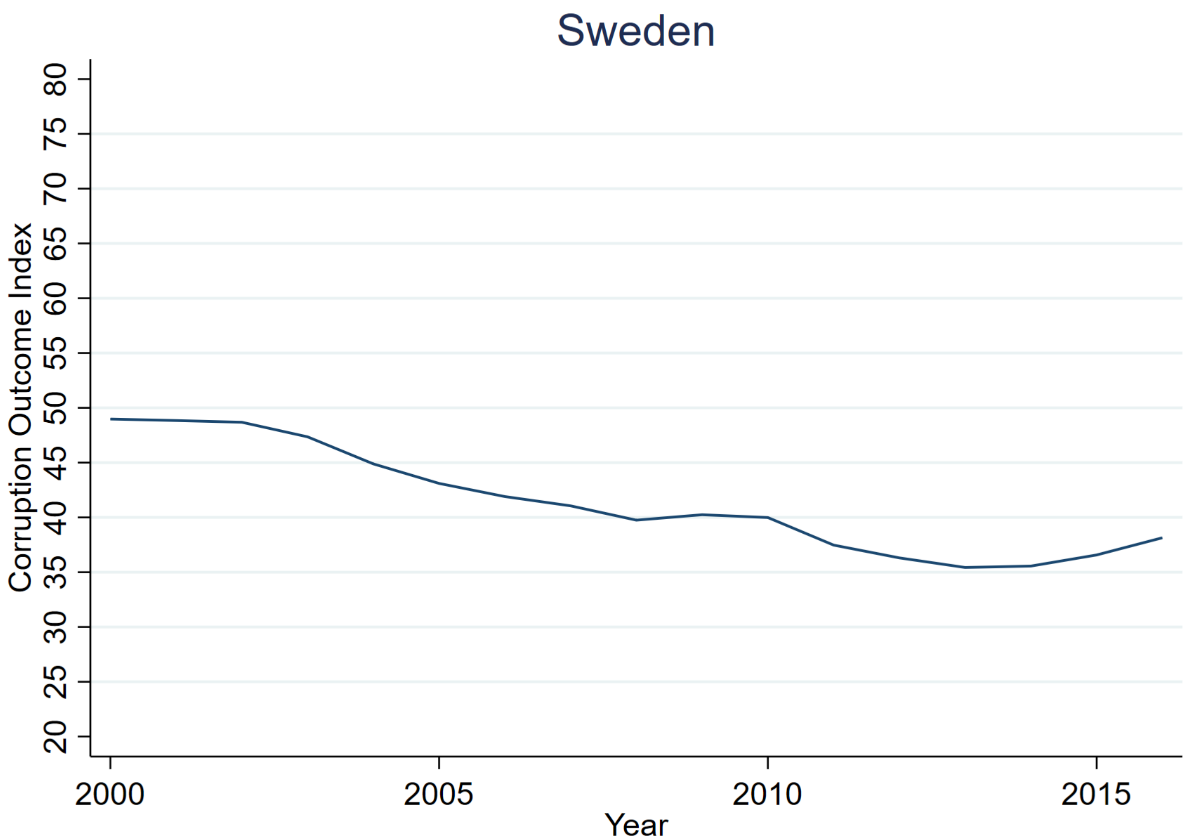

Suriname
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Slovakia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

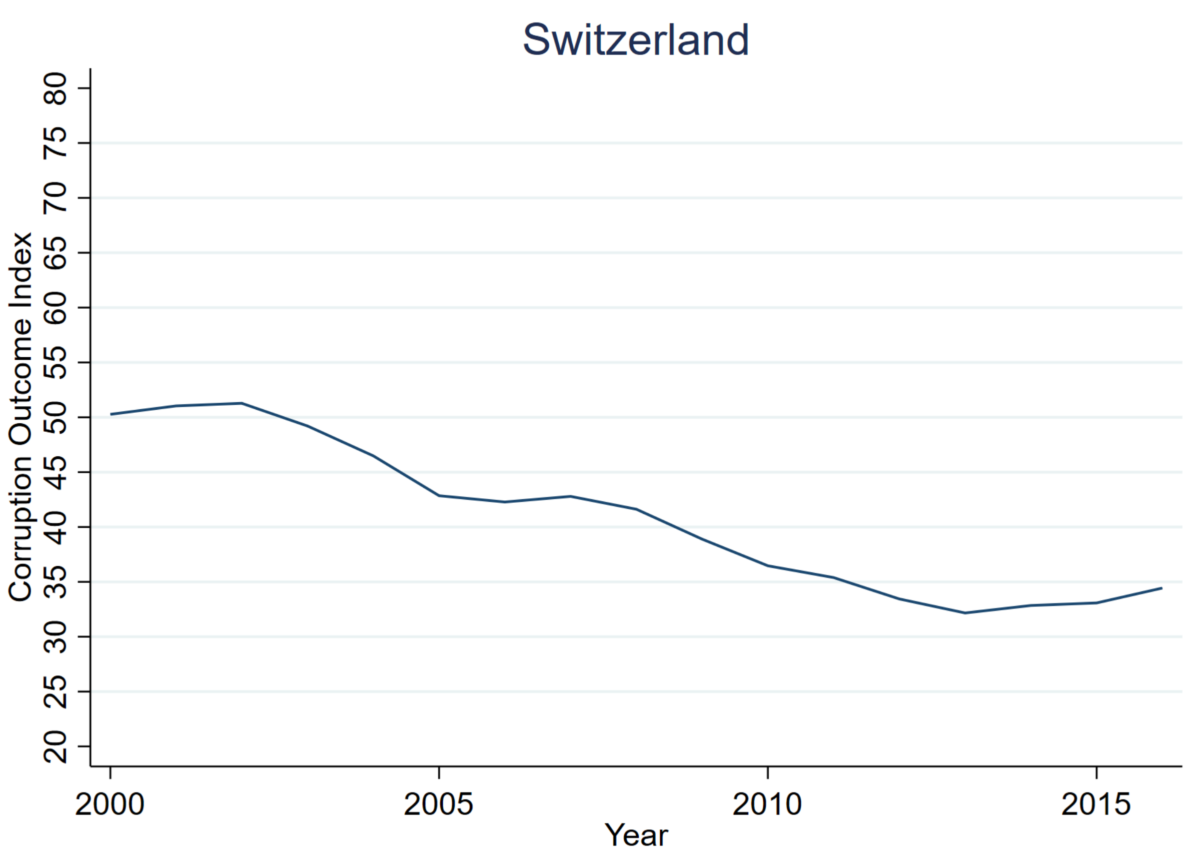

Slovenia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Sri Lanka
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

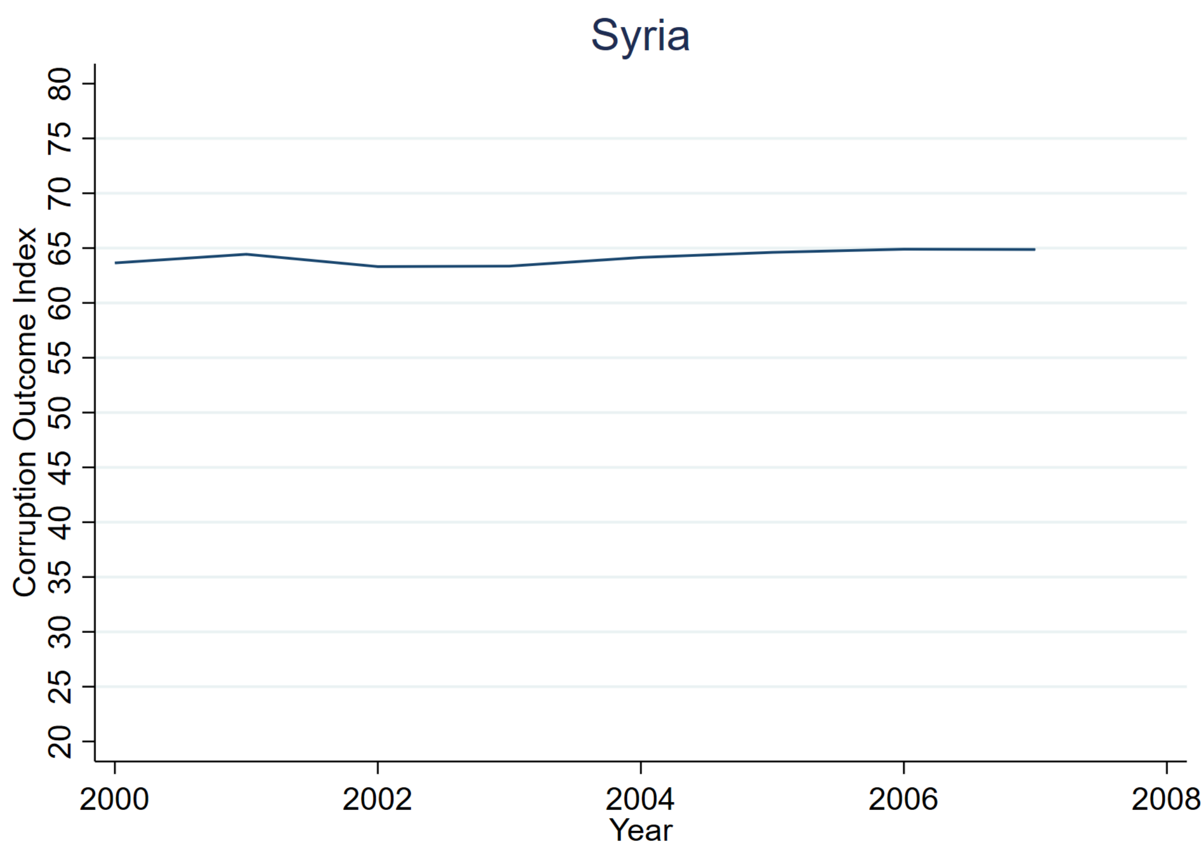

Swaziland
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

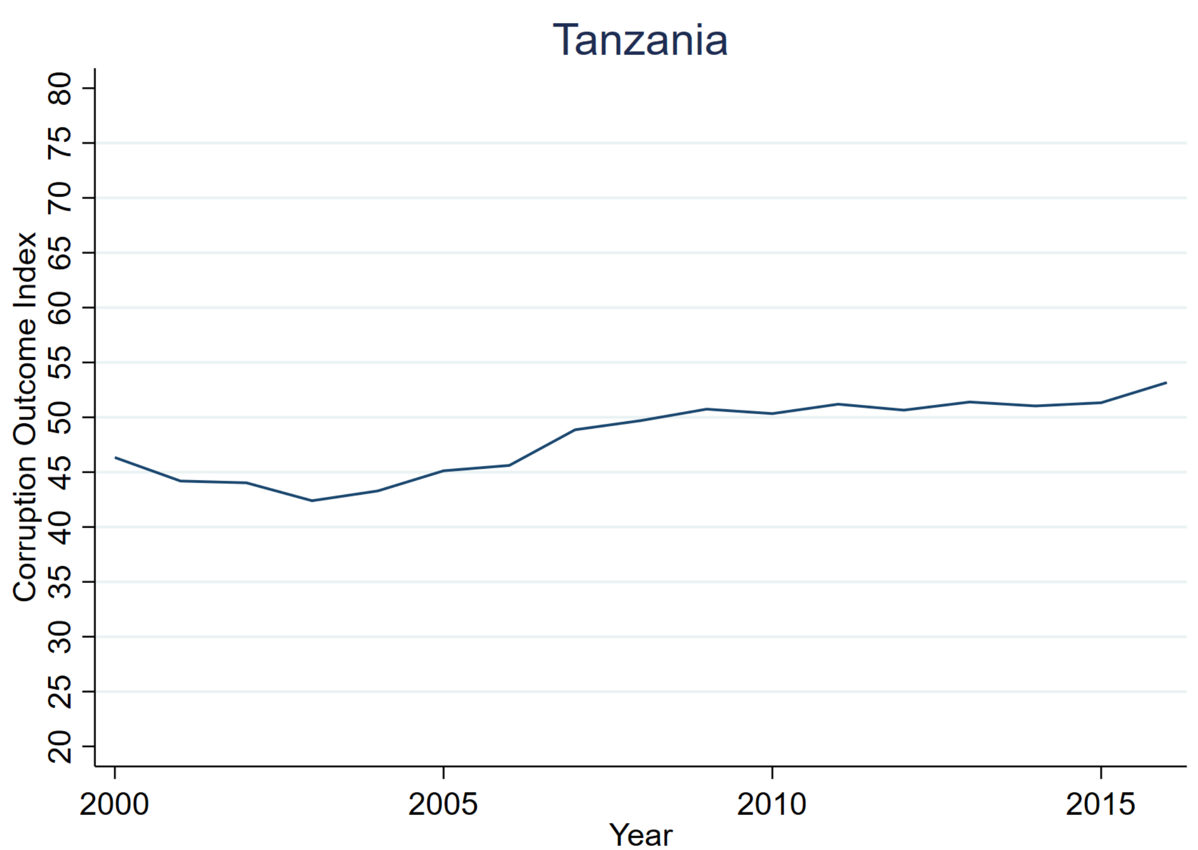


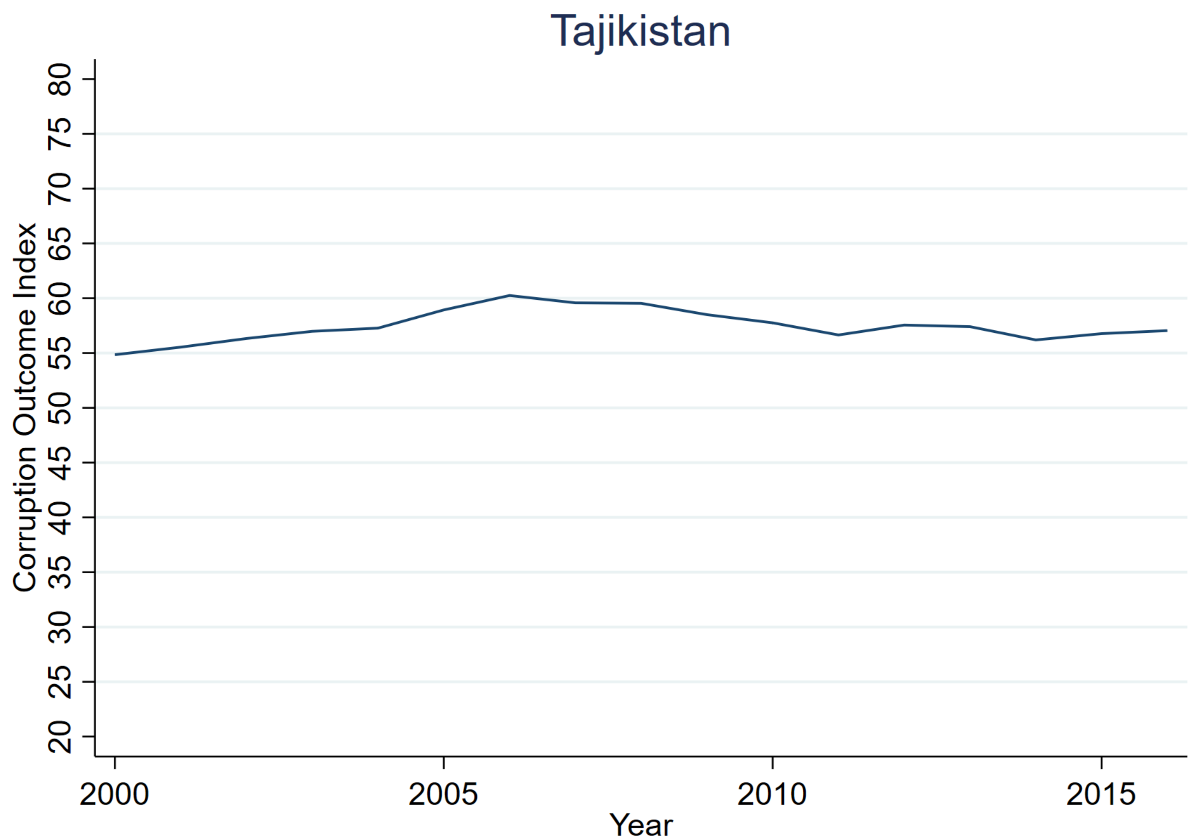
Sweden
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

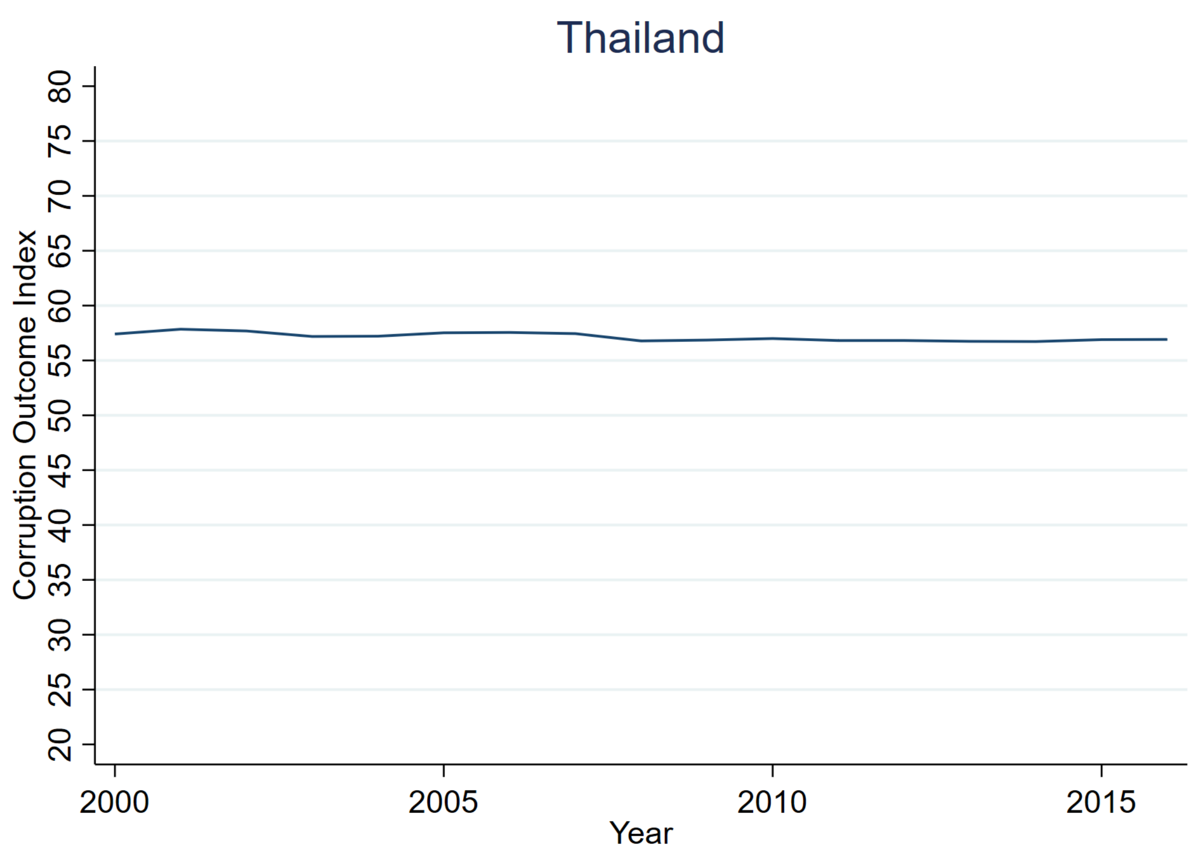


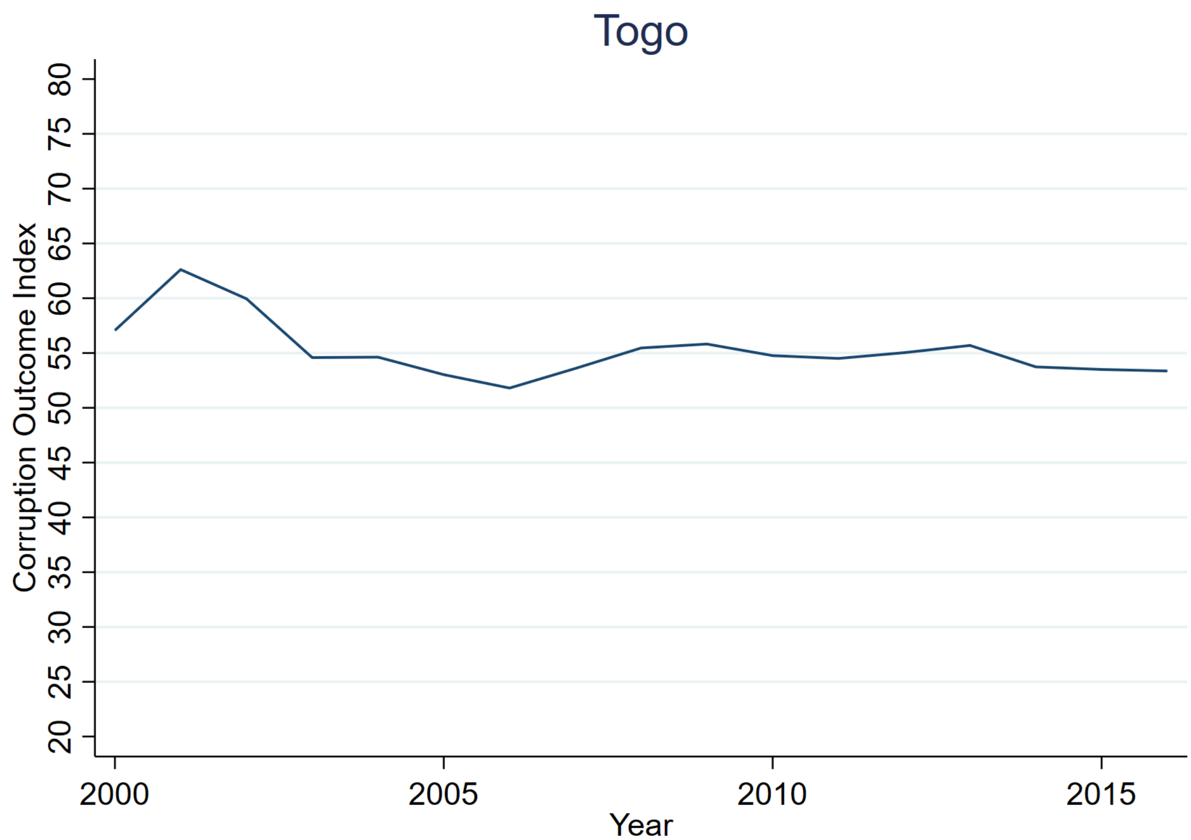
Switzerland
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

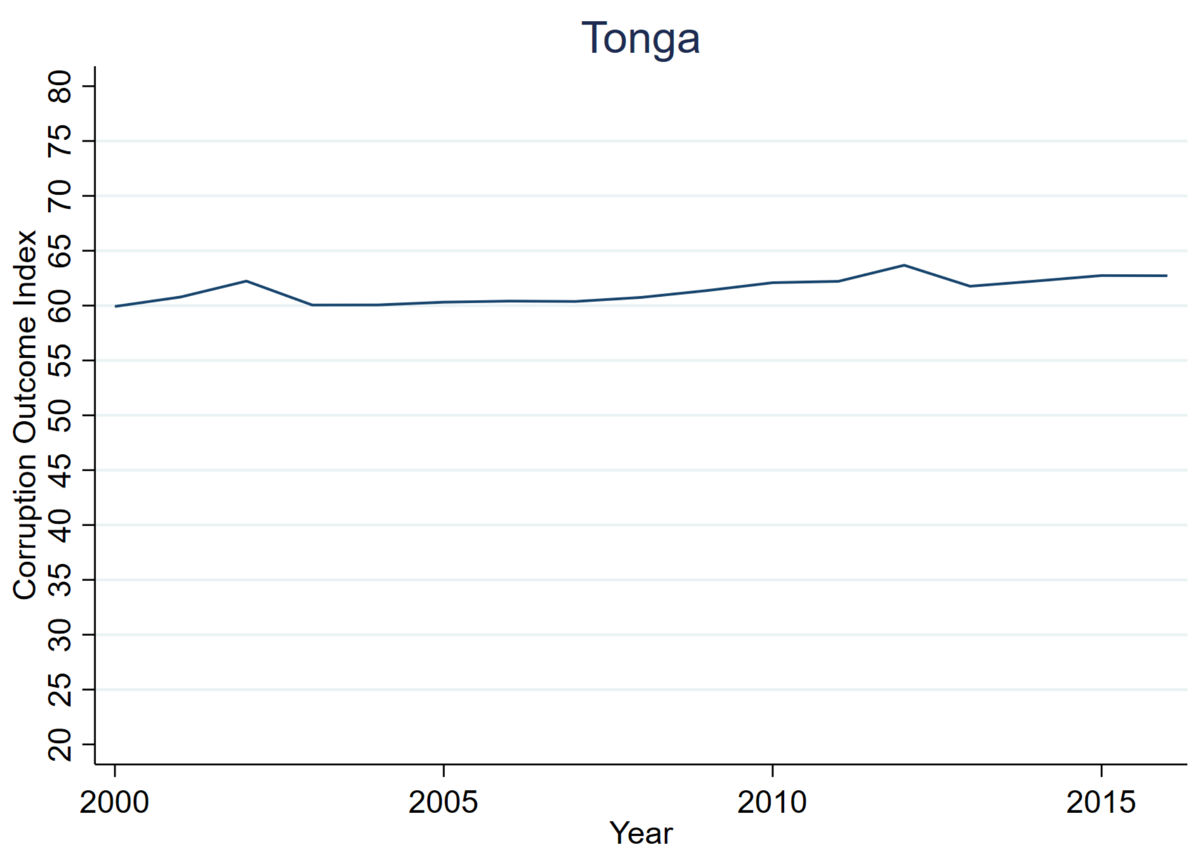

Syria
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2002
2004
2006
2008
Year

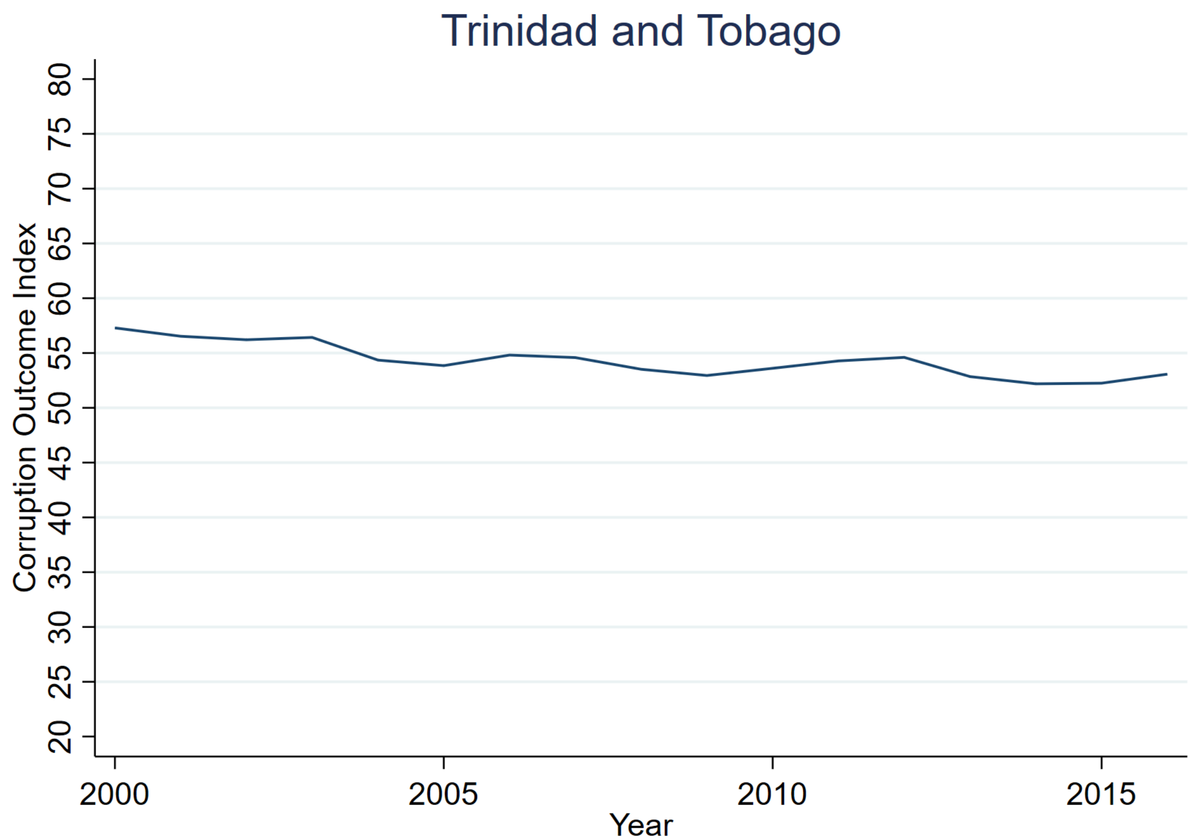

Tanzania
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

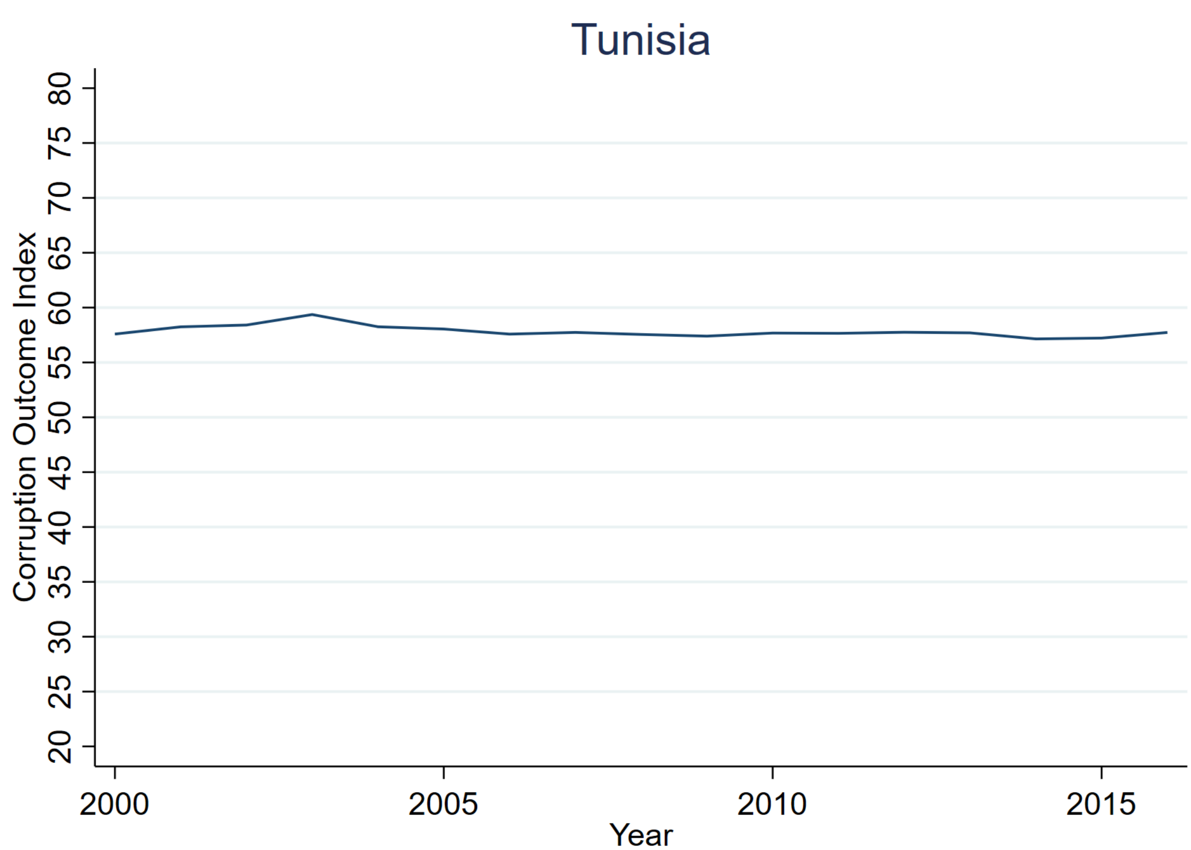

Tajikistan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

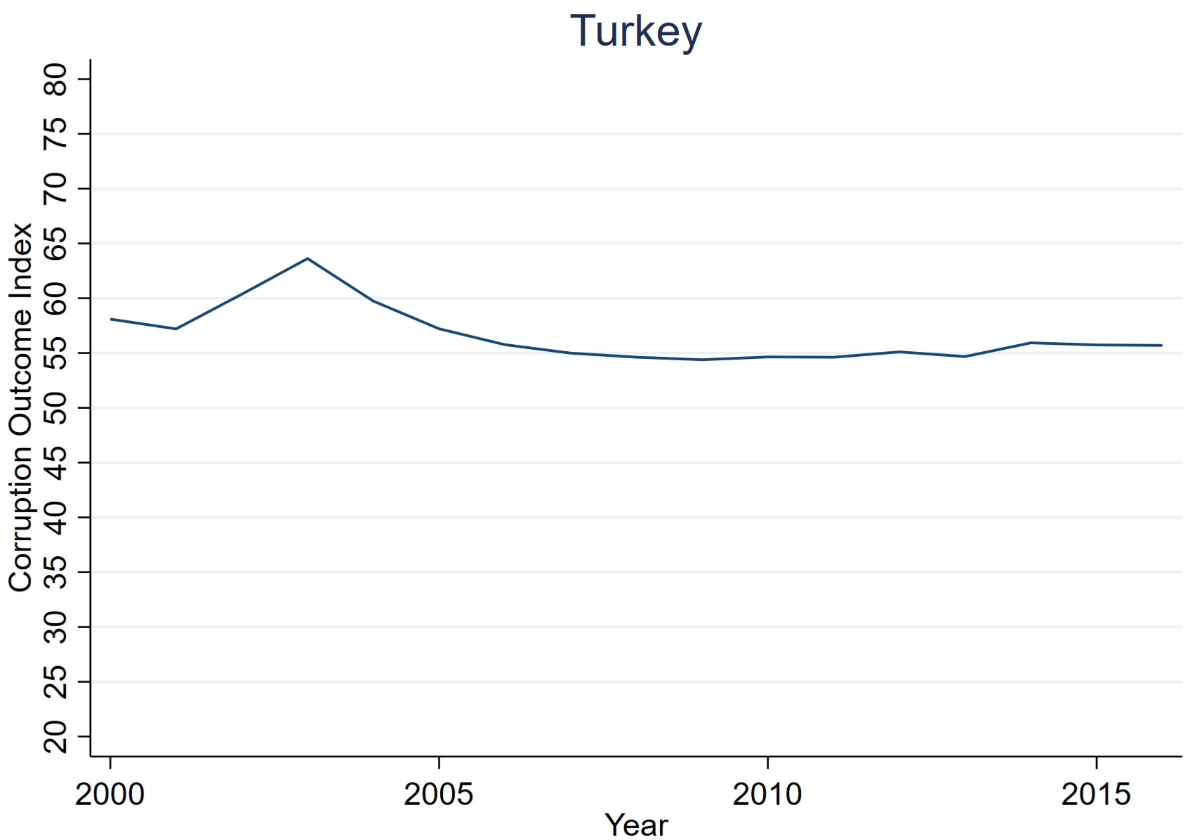

Thailand
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

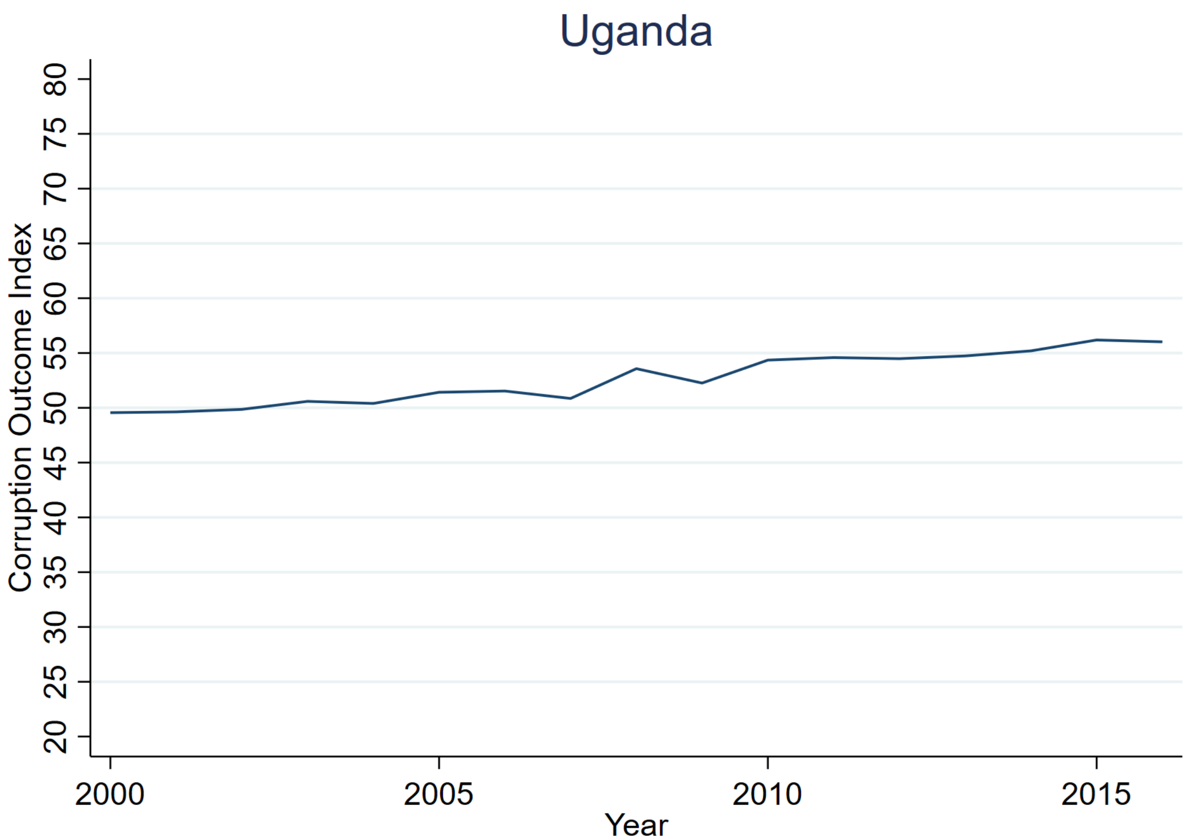

Togo
Corruption Outcome Index
20 25 30 35 40 45 50 55 60 65 70 75 80
2000 2005 2010 2015
Year


Tonga
Corruption Outcome Index
20 25 30 35 40 45 50 55 60 65 70 75 80
2000 2005 2010 2015
Year

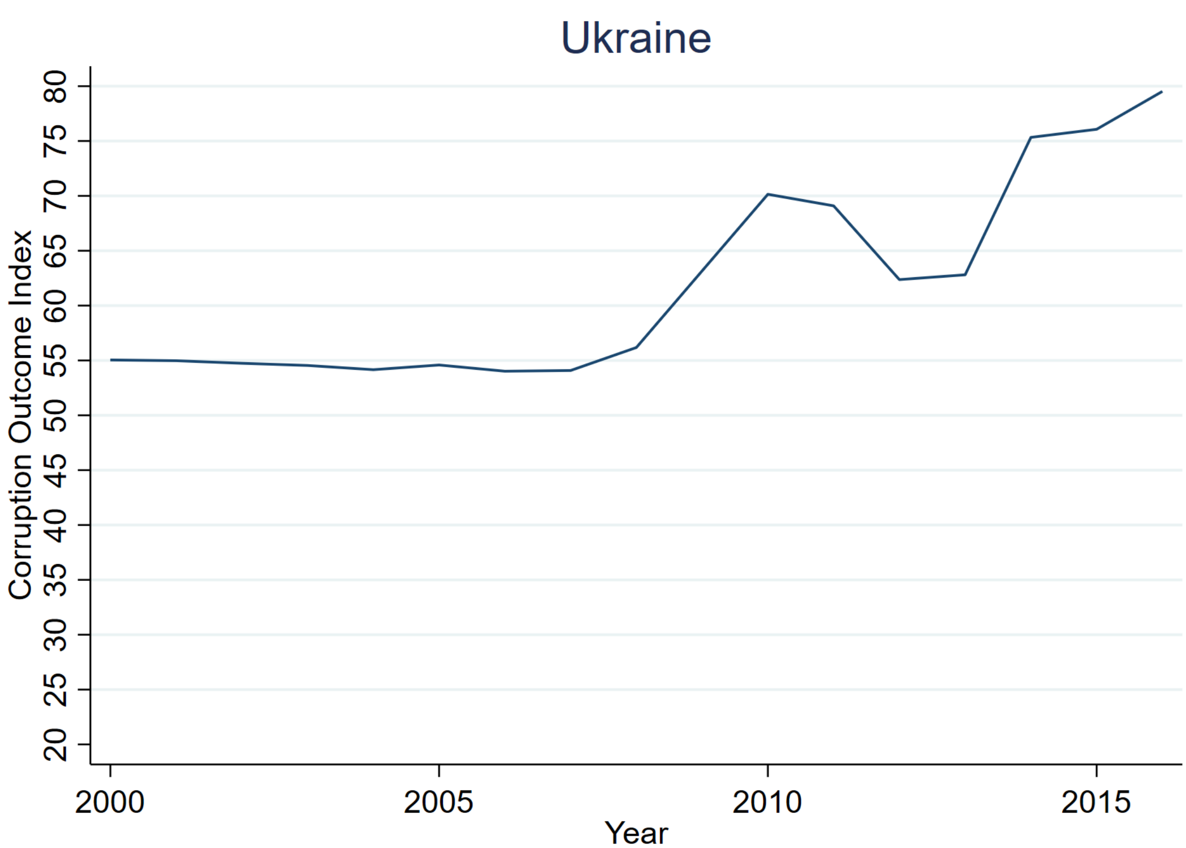

Trinidad and Tobago
Corruption Outcome Index
20 25 30 35 40 45 50 55 60 65 70 75 80
2000 2005 2010 2015
Year


Tunisia
Corruption Outcome Index
20 25 30 35 40 45 50 55 60 65 70 75 80
2000 2005 2010 2015
Year

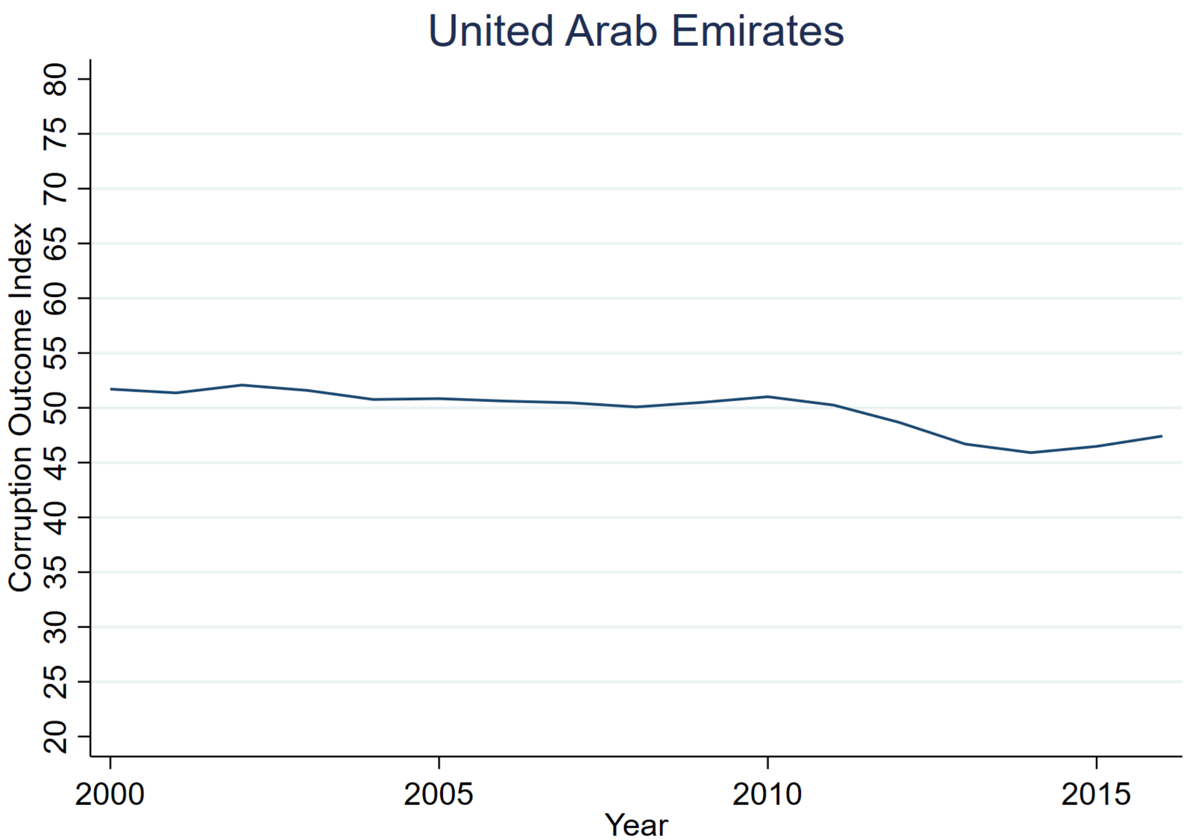

Turkey
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

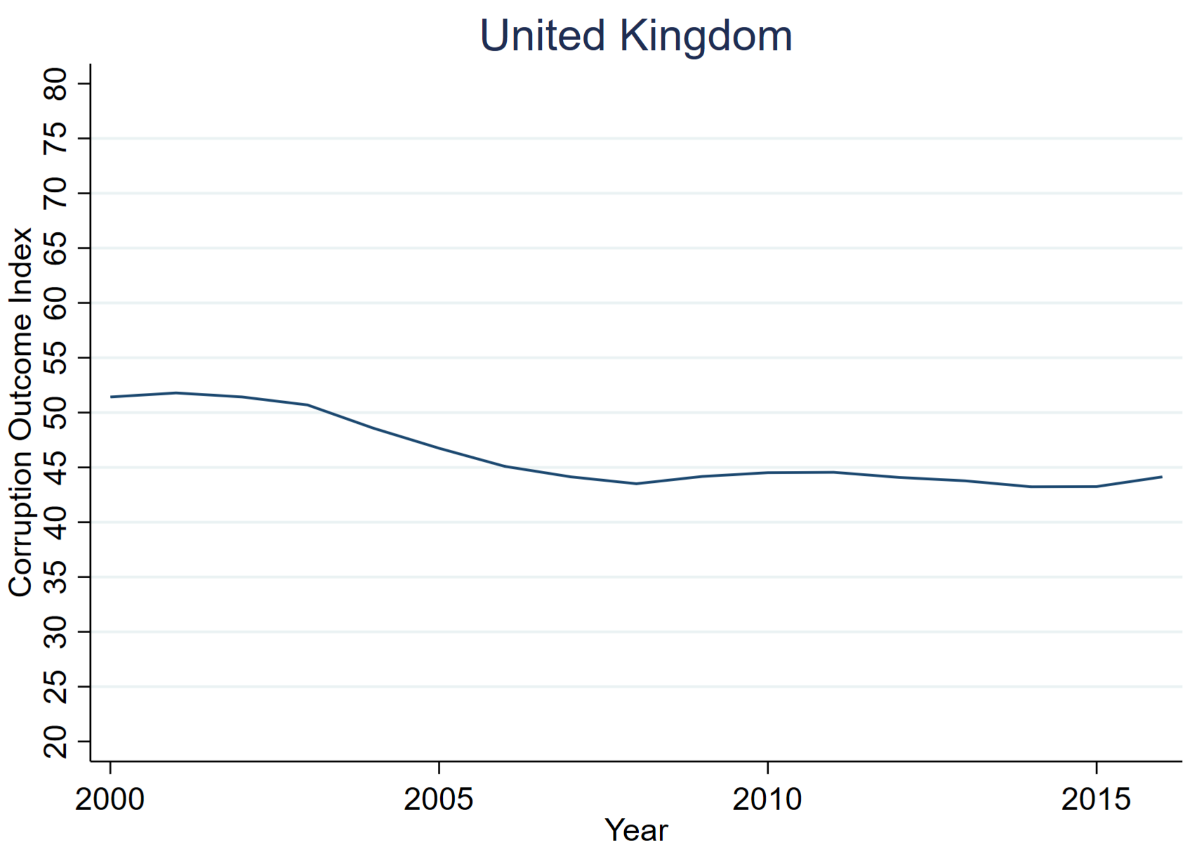


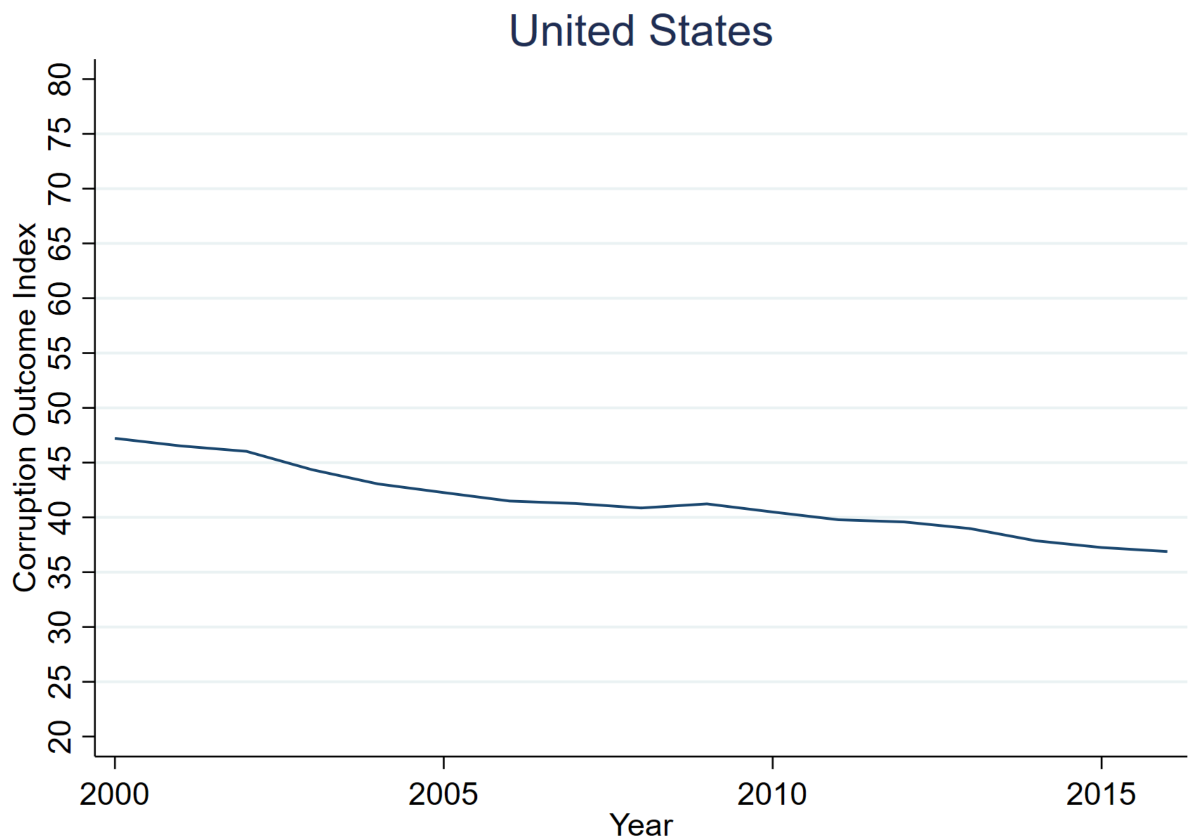
Uganda
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

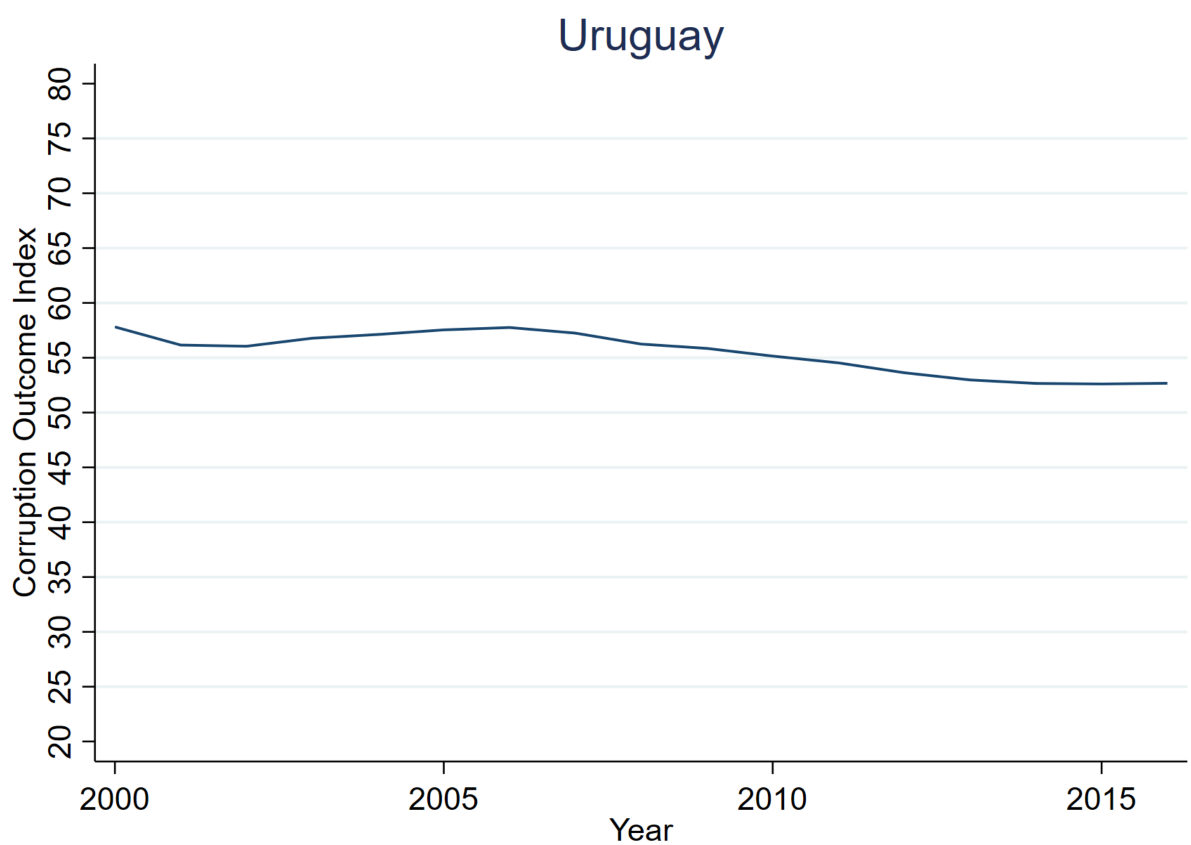


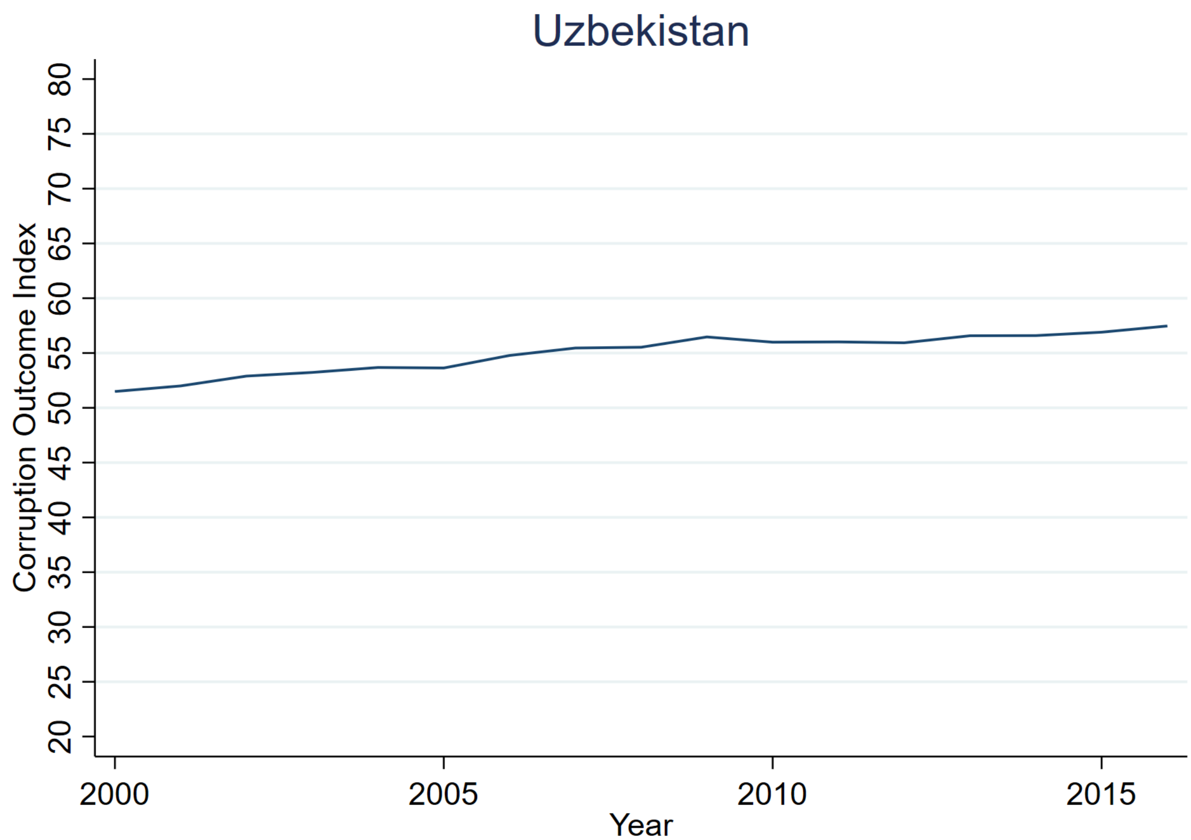
Ukraine
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

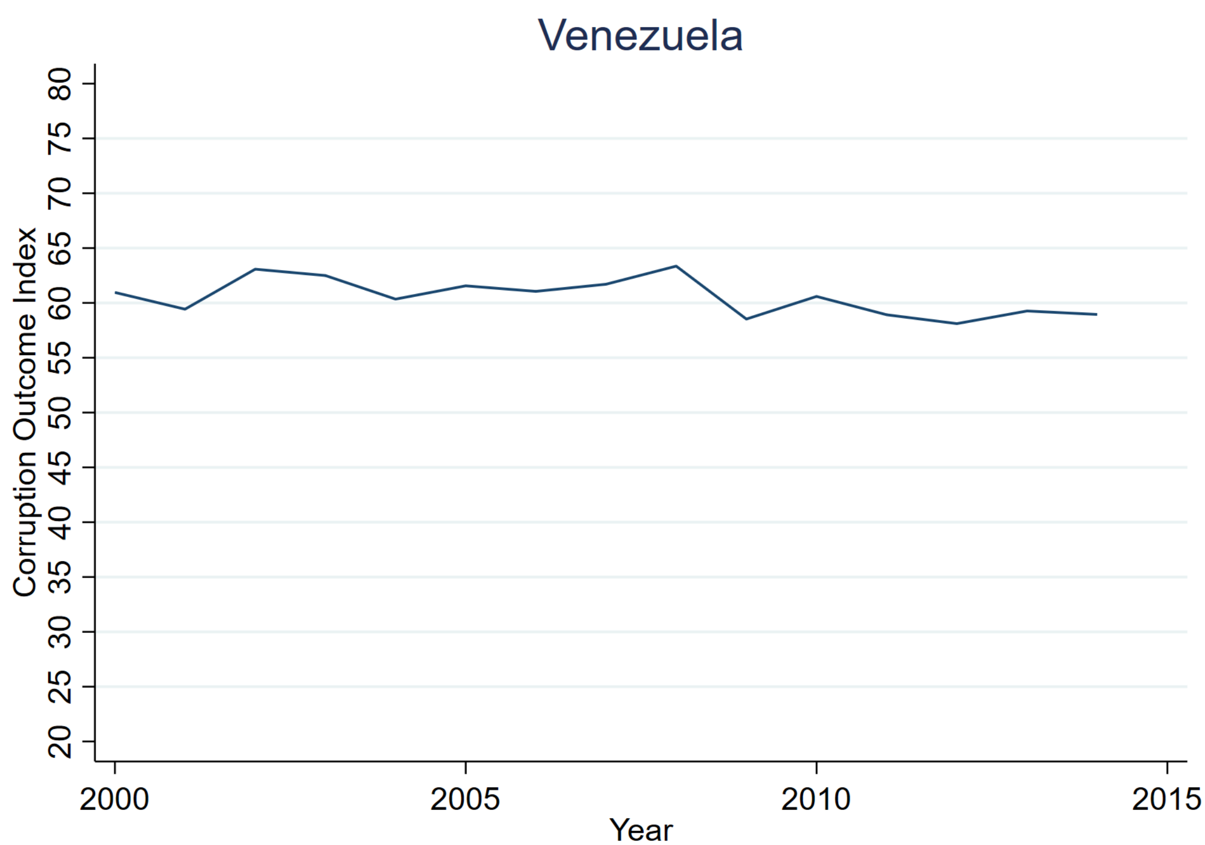

United Arab Emirates
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

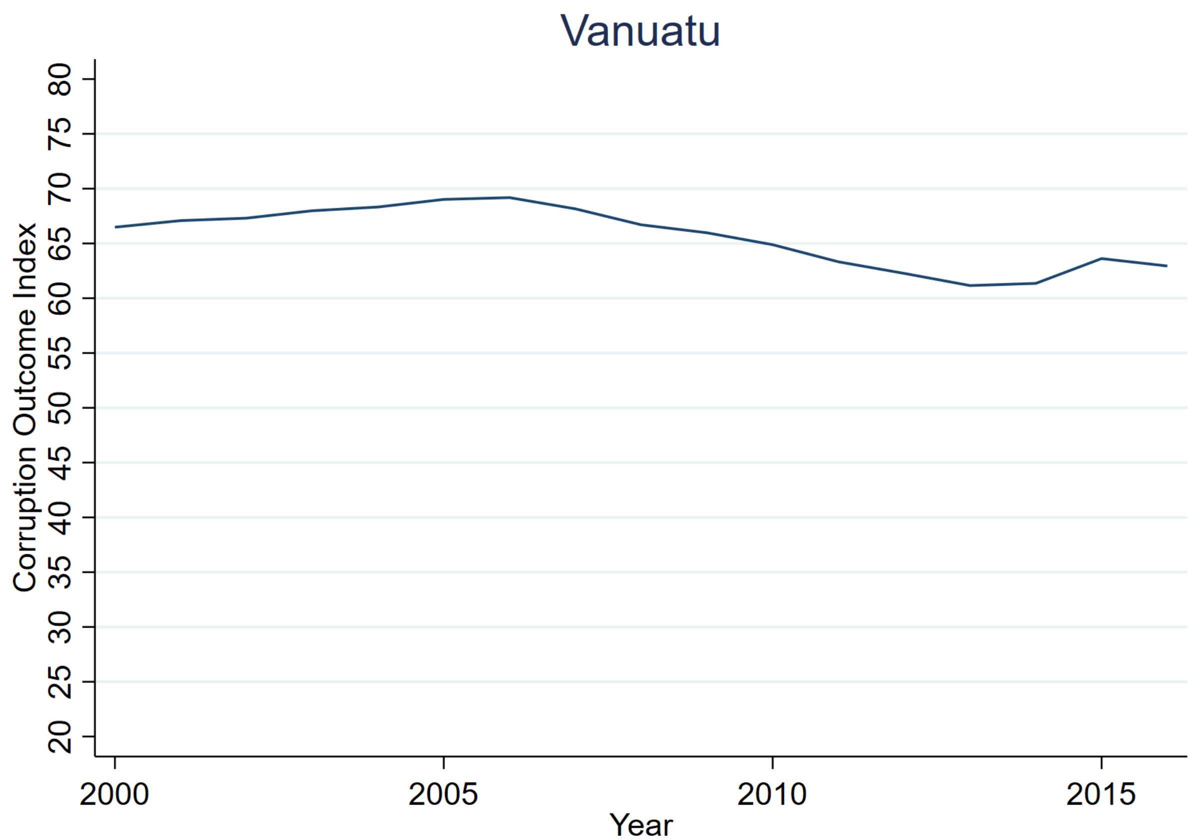

United Kingdom
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

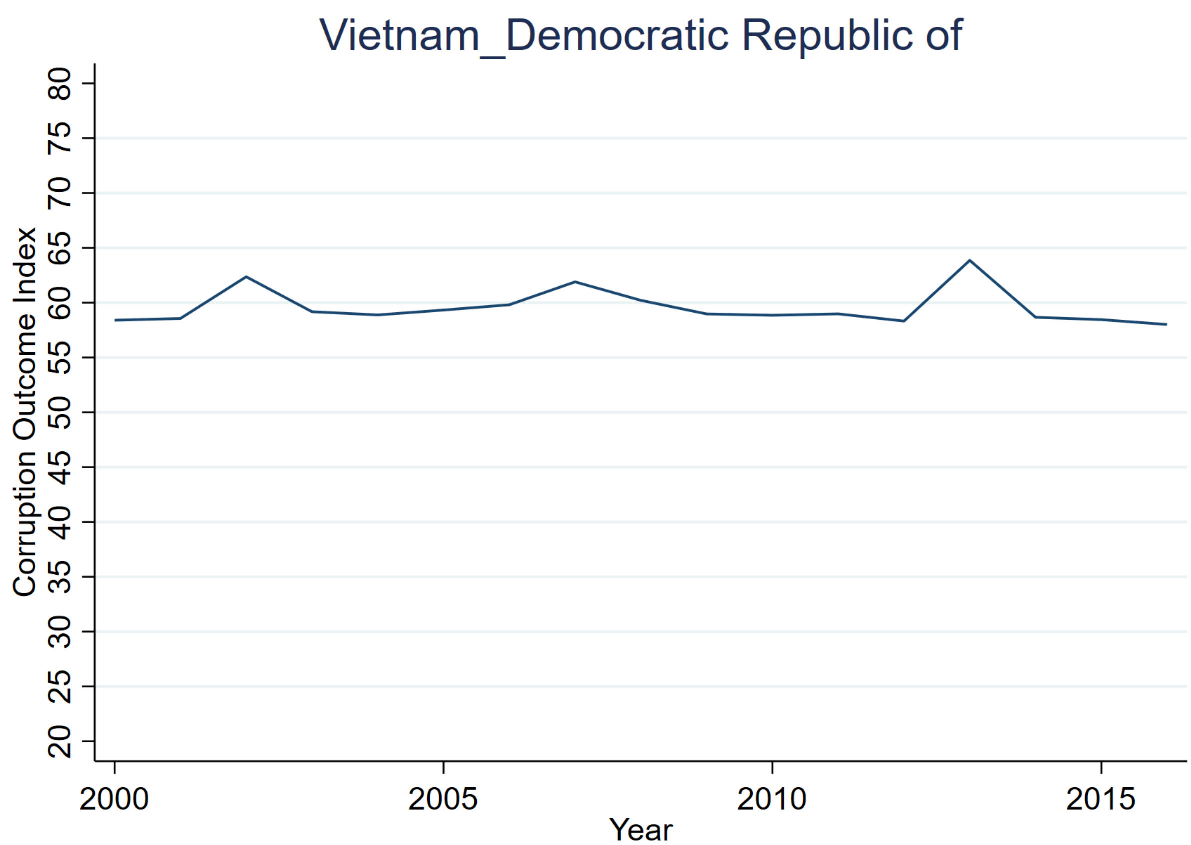

United States
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

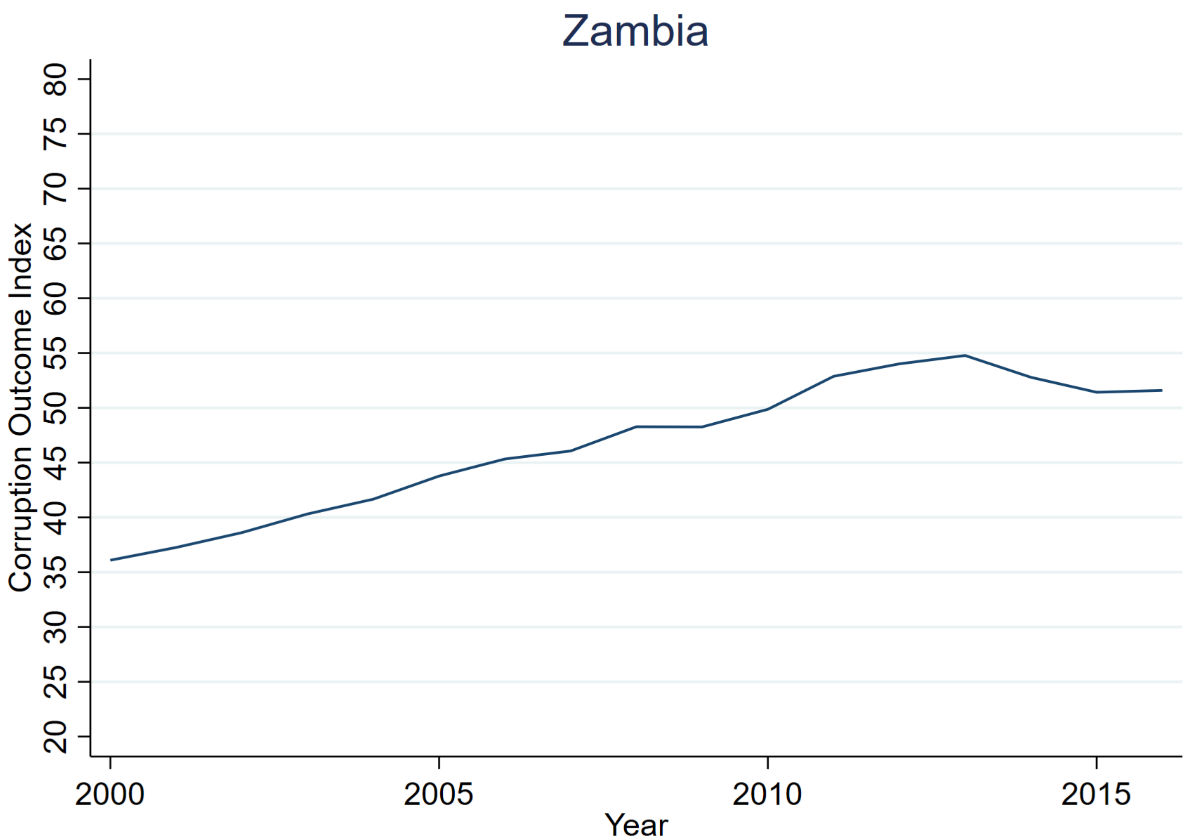

Uruguay
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year

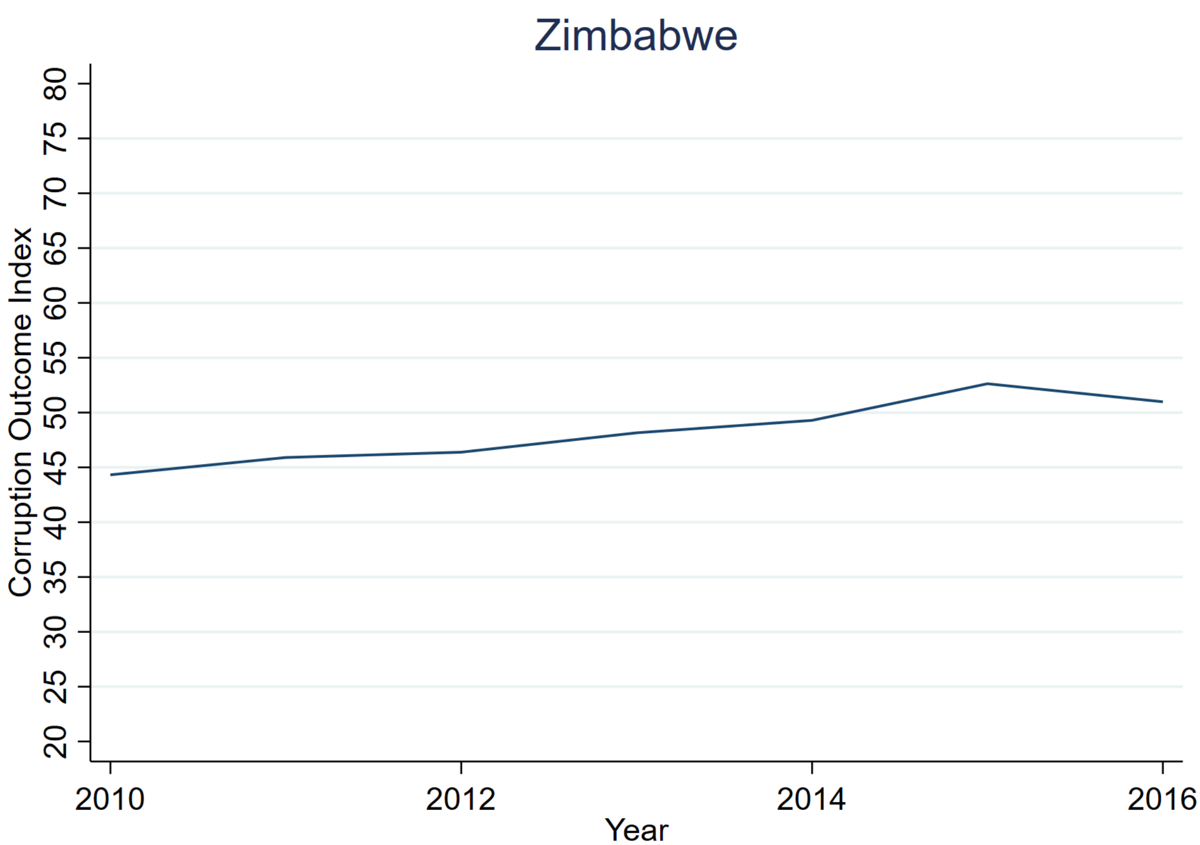

Uzbekistan
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Venezuela
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Vanuatu
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Vietnam_Democratic Republic of
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Zambia
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2000
2005
2010
2015
Year


Zimbabwe
Corruption Outcome Index
20
25
30
35
40
45
50
55
60
65
70
75
80
2010
2012
2014
2016
Year

References

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A Robinson. 2019. "Democracy does cause growth." *Journal of political economy* 127(1):47–100.

Akçay, Selçuk. 2006. "Corruption and human development." *Cato J.* 26:29.

Anbarci, Nejat, Monica Escaleras and Charles Register. 2006. "Traffic fatalities and public sector corruption." *Kyklos* 59(3):327–344.

Auxier, Brooke and Monica Anderson. 2021. "Social media use in 2021." *Pew Research Center* .

Becker, Dan. N.d. "XGBoost Advanced Regression Techniques.".
URL: *https://www.kaggle.com/dansbecker/xgboost*

Chayes, Sarah. 2015. *Thieves of state: Why corruption threatens global security*. WW Norton & Company.

Chen, Tianqi and Carlos Guestrin. 2016*a*. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16 New York, NY, USA: ACM pp. 785–794.
URL: *http://doi.acm.org/10.1145/2939672.2939785*

Chen, Tianqi and Carlos Guestrin. 2016*b*. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. pp. 785–794.

Chen, Tianqi, Michaël Benesty, Tong He and Y Tang. 2018. "Understand your dataset with xgboost." *R Document* .

Cieślik, Andrzej and Łukasz Goczek. 2018. "Control of corruption, international investment, and economic growth–Evidence from panel data." *World Development* 103:323–335.

Closser, Svea. 2019. "The Corruption Game: Health Systems, International Agencies, and the State in South Asia." *Medical Anthropology Quarterly* .

De Maria, Bill. 2008. "Neo-colonialism through measurement: a critique of the corruption perception index." *Critical perspectives on international business* .

Dell'Anno, Roberto and Désirée Teobaldelli. 2015. "Keeping both corruption and the shadow economy in check: the role of decentralization." *International Tax and Public Finance* 22(1):1–40.

Dincer, Oguzhan and Ozgur Teoman. 2019. "Does corruption kill? Evidence from half a century infant mortality data." *Social Science & Medicine* 232:332–339.

d'Agostino, Giorgio, J Paul Dunne and Luca Pieroni. 2016. "Government spending, corruption and economic growth." *World Development* 84:190–205.

Escresa, Laarni and Lucio Picci. 2017. "A new cross-national measure of corruption." *The World Bank Economic Review* 31(1):196–219.

Factor, Roni and Minah Kang. 2015. "Corruption and population health outcomes: an analysis of data from 133 countries using structural equation modeling." *International journal of public health* 60(6):633–641.

Friedman, Jerome H. 2001. "Greedy function approximation: a gradient boosting machine." *Annals of statistics* pp. 1189–1232.

Gründler, Klaus and Niklas Potrafke. 2019. "Corruption and economic growth: New empirical evidence." *European Journal of Political Economy* 60:101810.

Gründler, Klaus and Tommy Krieger. 2018. "Machine learning indices, political institutions, and economic development.".

Gupta S, Davoodi HR, Tiongson E. 2002. *Corruption and the provision of health care and education services. In: Amed G, Gupta S (eds) Governance, corruption, and economic performance.* Washington, DC: International Monetary Fund p. 245–279.

Harrison, Elizabeth. 2007. "Corruption." *Development in practice* 17(4-5):672–678.

Henderson, J Vernon, Adam Storeygard and David N Weil. 2012. "Measuring economic growth from outer space." *American economic review* 102(2):994–1028.

Heß, Simon. 2017. "Randomization inference with Stata: A guide and software." *Stata Journal* 17(3):630–651.

URL: *http://www.stata-journal.com/article.html?article=st0489*

House, Freedom. 2021. “Freedom House: Freedom in the world Lesotho Country Report.”.
URL: *https://freedomhouse.org/country/lesotho/freedom-world/2021*

Hu, Yingyao and Jiaxiong Yao. 2019. *Illuminating economic growth*. International Monetary Fund.

Hua, Law Teik, Robert B Noland and Andrew W Evans. 2010. “The direct and indirect effects of corruption on motor vehicle crash deaths.” *Accident Analysis & Prevention* 42(6):1934–1942.

Ibrahem Ahmed Osman, A, AA Najah, M Chow, HY Feng and A El-Shafie. 2021. “Extreme gradient boosting (Xgboost) model to predict thegroundwater levels in Selangor Malaysia.” *Ain Shams Eng. J* .

Index of public integrity. 2021.
URL: *https://integrity-index.org/methodology/*

Indicators, Global Integrity. 2021. “Africa Integrity Indicators – Project Summary Methodology.”. data retrieved from Global Integrity, Africa integrity.

Integrity, Global. 2018. “Africa integrity indicators.”. data retrieved from Global Integrity, Africa integrity indicators `https://www.globalintegrity.org/`.

Justesen, Mogens K and Christian Bjørnskov. 2014. “Exploiting the poor: Bureaucratic corruption and poverty in Africa.” *World Development* 58:106–115.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2007. “Measuring corruption: myths and realities.”.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi. 2010. “The worldwide governance indicators: A summary of methodology.” *Data and Analytical Issues, World Bank Policy Research Working Paper* 5430.

Kirigia, Joses Muthuri and Alimata J Diarra-Nama. 2008. “Can countries of the WHO African Region wean themselves off donor funding for health?” *Bulletin of the World Health Organization* 86:889–892.

Klitgaard, Robert. 1998. “International cooperation against corruption.” *Finance & Development* 35(001).

Klitgaard, Robert. 2013. "Tropical Gangsters II: Adventures in Development in the World's Poorest Places." *Amazon: KDP Books* .

Klitgaard, Robert. 2017. On culture and corruption. In *This paper was presented at the Public Integrity and Anti-Corruption workshop at Nuffield College, Oxford June*. Vol. 13 p. 2017.

Lindberg, Staffan I, Michael Coppedge, John Gerring and Jan Teorell. 2014. "V-Dem: A new way to measure democracy." *Journal of Democracy* 25(3):159–169.

Machen, Ronald C. etal. 2021. "Investigation of Data Irregularities in Doing Business 2018 and Doing Business 2020.".

URL: *https://thedocs.worldbank.org/en/doc/84a922cc9273b7b120d49ad3b9e9d3f9-0090012021/original/DB-Investigation-Findings-and-Report-to-the-Board-of-Executive-Directors-September-15-2021.pdf*

Mahoney, Melissa and Robert Klitgaard. 2019. "Revitalizing Mandaue city: obstacles in implementing a performance governance system." *Policy Design and Practice* 2(4):383–399.

Moiseeva, IA. 2018. "Corruption Perception Index as an Indicator of the Level of Corruption in an Individual Country and the World." *Actual Probs. Econ. & L.* p. 89.

Mondo, Bianca Vaz. 2016. "Measuring political corruption from audit results: a new panel of Brazilian municipalities." *European Journal on Criminal Policy and Research* 22(3):477–498.

Mungiu-Pippidi. 2021. "The Doing Business Gate: Let's not throw the baby out with the bathwater.".

URL: *https://www.againstcorruption.eu/articles/the-doing-business-gate-lets-not-throw-the-baby-out-with-the-bathwater/*

Mungiu-Pippidi, Alina and Ramin Dadašov. 2016. "Measuring control of corruption by a new index of public integrity." *European Journal on Criminal Policy and Research* 22(3):415–438.

Nikolov, Dimitar, Diego FM Oliveira, Alessandro Flammini and Filippo Menczer. 2015. "Measuring online social bubbles." *PeerJ computer science* 1:e38.

Noh, Byungjoo, Changhong Youm, Eunkyoung Goh, Myeounggon Lee, Hwayoung Park, Hyojeong Jeon and Oh Yoen Kim. 2021. "XGBoost based machine learning approach to predict the risk of fall in older adults using gait outcomes." *Scientific reports* 11(1):1–9.

Razafindrakoto, Mireille and François Roubaud. 2010. "Are international databases on corruption reliable? A comparison of expert opinion surveys and household surveys in sub-Saharan Africa." *World development* 38(8):1057–1069.

Rose-Ackerman, Susan and Rory Truex. 2012. "Corruption and policy reform." *Yale Law & Economics Research Paper* (444).

Rothstein, Bo. 2012. Defining and Measuring Quality of Government. In *In S. Holmberg and B. Rothstein (eds). Good Government: The Relevance of Political Science*. Citeseer.

Savedoff, William D and Karen Hussmann. 2006. "Why are health systems prone to corruption." *Global Corruption Report* 2006:4–16.

Spohr, Dominic. 2017. "Fake news and ideological polarization: Filter bubbles and selective exposure on social media." *Business Information Review* 34(3):150–160.

Standaert, Samuel. 2015. "Divining the level of corruption: a bayesian state-space approach." *Journal of Comparative Economics* 43(3):782–803.

Steinwart, Ingo and Andreas Christmann. 2008. *Support vector machines*. Springer Science & Business Media.

Stiftung, Bertelsmann. 2018. *Transformation Index BTI 2018: Governance in International Comparison*. Verlag Bertelsmann Stiftung.

Stinson, Philip M, John Liederbach, Steven L Brewer, Hans D Schmalzried, Brooke E Mathna and Krista L Long. 2013. "A study of drug-related police corruption arrests." *Policing: An International Journal of Police Strategies & Management* .

Teorell, Jan, Staffan Kumlin, Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon and Richard Svensson. 2019. "The Quality of Government OECD Dataset, Version Jan19." *University of Gothenburg: The Quality of Government Institute* .

Transparency International. 2020*a*. "Corruption Perception Global Scores." *Transparency International, Berlin* .

Transparency International. 2020*b*. "Corruption Perceptions Index 2020: Technical Methodology

Note." *Transparency International, Berlin* .

Treisman, Daniel. 2000. "The causes of corruption: a cross-national study." *Journal of public economics* 76(3):399–457.

Treisman, Daniel. 2007. "What have we learned about the causes of corruption from ten years of cross-national empirical research?" *Annu. Rev. Polit. Sci.* 10:211–244.

World Bank. 2018. "World Development Indicators.".
URL: *http://datatopics.worldbank.org/world-development-indicators/*

World Health Organization. 2018. "World Health Organization, Global Health Observatory.".
URL: *https://www.who.int/gho/database/en/*

Yoo, Seung Han. 2008. "Petty corruption." *Economic Theory* 37(2):267–280.

Zhang, Yan, Lin Chen et al. 2021. "A Study on Forecasting the Default Risk of Bond Based on XGboost Algorithm and Over-Sampling Method." *Theoretical Economics Letters* 11(02):258.