Corruption and governance around the world

Seldadyo, H.

Publication date:
2008

Chapter 4

Is Corruption Really Persistent?

"Change does not necessarily assure progress,
but progress implacably requires change."

Henry Steele Commager (1902–1998)

4.1 Introduction

There is strong evidence that corruption has detrimental effects on determinants of economic growth, like investment and trade (see, for instance, Wei (2000b), Anderson and Marcouiller (2002)). It is also widely believed that, despite these negative consequences, corruption is highly persistent.

Theoretical models of corruption generally predict that corruption is immutable. Tirole (1996), for example, argues that once agents are corrupt, they remain so due to reputation effects. Similarly, Mishra (2006) indicates that when there are many corrupt individuals, it is optimal to be corrupt; thus once corruption is widespread, it will be persistent. This is in line with the view of Mauro (2004) that in countries where corruption is pervasive, there are no incentives for individuals to fight corruption, even though

Joint work with Jakob de Haan; the earlier version was presented at the First World Public Choice Meetings 2007. We received enlightening comments from Tom Wansbeek.

they would be better off if corruption were eradicated. Finally, Basu et al. (1992) show that collusion of agents in a hierarchical structure also leads to corruption persistence.[1]

Also on the basis of empirical research, some authors claim that corruption is persistent. Herzfeld and Weiss (2003), for instance, report that the absence of law and order increases corruption, while more corruption lowers enforcement. Thus, once trapped, it is difficult for a country with a weak legal system to escape from corruption. This is in line with the view of Jain (2001: 72), that "Once corrupted, the elite will attempt to reduce the effectiveness of the legal and judicial systems through manipulation of resource allocation and appointments to key positions. Reduced resources will make it difficult for the legal system to combat corruption, thus allowing corruption to spread even more."[2]

However, we argue that corruption changes over time. Using the International Country Risk Guide (ICRG) data for the period 1984-2003, we find strong evidence that corruption is not persistent. Many corrupt countries were able to reduce their level of corruption, while many clean countries have become less clean over the same period. Apart from some standard regressions, we apply non-parametric methods to examine the distributional dynamics of corruption in 101 countries. Specifically, we evaluate the modality of corruption data across countries and over time using a kernel density function. In addition, we apply Markov chain analysis. All our results reject the hypothesis that corruption is persistent.

The remainder of the chapter is organized as follows. Section 4.2 provides evidence on corruption convergence based on correlation. Section 4.3 discusses a Gaussian kernel density function, while section 4.4 presents a Markov chain analysis. Section 4.5 concludes.

[1] Other relevant theoretical papers on the persistence of corruption include Blackburn et al. (2006), Damania et al. (2004) and Ventelou (2002). See also Bardhan (1997, especially p. 1330-1334), for a survey of corruption persistence.

[2] Also Alesina and Weder (2002) and Damania et al. (2004) argue on empirical grounds that corruption appears to be persistent.

4.2 Convergence: Preliminary Evidence

We use data on corruption from ICRG covering 101 countries around the world.[3] These are the countries for which the ICRG corruption indicator is available for all years in the sample 1984-2003. The ICRG data is based on perceived corruption by a panel of experts. The level of corruption is expressed on a scale between zero and six, but for the current analysis the index is rescaled from one to seven where a higher score means less corruption.[4] We use the ICRG data because this is the only indicator of corruption that is available for a long time period on a consistent basis.[5]

The ICRG data suggests that corruption changes over time. In some countries corruption increases, while in others it declines. Table 4.1 displays losers and winners in the world corruption league in which countries are ranked according to their relative change in the level of corruption over 1984-2003. The first panel shows the losers, i.e., countries with an increase in corruption between 1984-2003. Zimbabwe has the highest relative change in this group. It drops from a relatively moderate level (3.92) to the lowest level (1.00). The change in corruption does not only occur in countries categorized as moderately or highly corrupt in 1984, but also in countries that were initially labeled as clean. Canada, Singapore, Switzerland, the United Kingdom, and Iceland fall into this category.

The second panel shows countries with no change in the level of corruption over the same period. There are 12 countries listed in this panel: they come from almost all ranks in the corruption league in 1984. For instance, Indonesia (one of the most corrupt countries in 1984), Jordan, and Colombia (from the mid ranks), as well as Austria (one of the cleanest countries)

[3] `http://www.prsgroup.com/ICRG_Methodology.aspx`.

[4] The rescaling is required as we measure the relative change in corruption as discussed later. The annual ICRG data is constructed on the basis of averaged monthly data. The monthly data is measured on an ordinal scale, but the averaging arguably transforms the data into continuous data. As we will show later, treating the data as ordinal does not affect our main conclusions.

[5] Other measures are problematic. The corruption index of the World Bank, for example, is only available for recent years, while the index of Transparency International is constructed using a non-consistent methodology. See Section 2.2 for further details.

are all in this group. Finland is also in this group, but this country has reached the upper bound thus an increase in the ICRG index is not possible. As in the first panel, both developed and developing countries are in this panel. This part of the table also shows that corruption persistence is not a widespread phenomenon, as corruption remained at the same level in a limited number of countries only.

Finally, the third panel shows the winners, i.e., countries with a declining level of corruption over 20 years. There are 22 countries in this group; most of them were in the low and mid positions in the corruption ranking in 1984. Examples are Iraq, with the smallest positive relative change, as well as the Philippines that has the biggest improvement. Also some African countries saw their level of corruption decline.

To investigate more precisely how corruption has changed over time, we calculate the correlation between the level of corruption in 1984—the starting point of our sample period—and corruption in subsequent years. In case of corruption persistence, the correlation coefficients should be stable over time (ρ-persistence). Similarly, we also run a series of autoregressions of the level of corruption on its initial level. If corruption is persistent, the constant and the slope should be zero and one, respectively.

The first columns in Table 4.2 display the relationship between corruption in 1984 and corruption in subsequent years using Pearson, Spearman, and polychoric correlation coefficients.[6] The correlation between corruption in 1984 and corruption 1-4 years later is high, indicating that change in corruption is not a short-term phenomenon. However, as the time-lag i increases, the correlation coefficients gradually decrease to reach the lowest level at time-lag $i = 15$.

[6] Polychoric correlation is mostly used for ordinal variables (Olsson, 1979). Aish and Jöreskog (1990) use this measure to compute the correlation among a set of perception-based indicators of political efficacy and responsiveness at two points in time (1974 and 1981).

Table 4.1: Losers and Winners 1984-2003

Country	Change	Country	Change	Country	Change	Country	Change
Panel 1. Negative Change (Mean: -0.257; N=67)							
Zimbabwe	-0.745	Myanmar	-0.333	Poland	-0.250	Togo	-0.167
S. Africa	-0.524	S. Arabia	-0.308	Tunisia	-0.250	Turkey	-0.167
Lebanon	-0.500	Italy	-0.300	UAE	-0.250	Tanzania	-0.163
PNG	-0.500	Belgium	-0.286	Israel	-0.236	Chile	-0.144
France	-0.429	Paraguay	-0.273	USA	-0.221	Nether.	-0.143
Brunei	-0.417	Venezuela	-0.268	Canada	-0.214	Norway	-0.143
Costa Rica	-0.417	Argentina	-0.268	Singapore	-0.214	Greece	-0.125
Malaysia	-0.417	Angola	-0.250	Switzer.	-0.214	Nicaragua	-0.125
China	-0.417	Bahrain	-0.250	UK	-0.214	Senegal	-0.125
Albania	-0.400	Dom. Rep.	-0.250	Iceland	-0.202	Australia	-0.083
Bulgaria	-0.400	Iran	-0.250	Hungary	-0.200	Denmark	-0.071
Malawi	-0.400	Ireland	-0.250	Taiwan	-0.200	NZ	-0.071
Algeria	-0.375	Japan	-0.250	Spain	-0.182	Sweden	-0.071
India	-0.375	Kuwait	-0.250	Sudan	-0.172	Peru	-0.068
Thailand	-0.375	Libya	-0.250	Guatemala	-0.167	S. Korea	-0.065
Ser.-Mont.	-0.354	Mexico	-0.250	Hong Kong	-0.167	Egypt	-0.063
Gabon	-0.333	Nigeria	-0.250	Jamaica	-0.167		

Continued on next page ...

Country	Change	Country	Change	Country	Change	Country	Change
Panel 2. Zero Change (Mean: 0.000; N=12)							
Austria	0.000	Ecuador	0.000	Indonesia	0.000	Sri Lanka	0.000
Cameroon	0.000	Finland	0.000	Jordan	0.000	Trin.-Tob.	0.000
Colombia	0.000	Guinea	0.000	Panama	0.000	Uruguay	0.000
Panel 3. Positive Change (Mean: 0.596; N=22)							
Iraq	0.045	Ivory Coast	0.231	Bolivia	0.385	Congo DR	1.000
Portugal	0.059	Syria	0.241	Liberia	0.458	Guyana	1.000
Romania	0.097	Zambia	0.241	Kenya	0.486	Haiti	1.000
Brazil	0.132	Pakistan	0.250	Uganda	0.500	Mali	1.000
El Salvador	0.167	Morocco	0.333	Bangladesh	1.000	Philippines	1.917
Honduras	0.167	Ghana	0.355				

The change is measured relative to the 1984 level

The Pearson coefficient reduces from 0.98 to 0.74, the Spearman coefficient drops from more than 0.99 to 0.68, while the polychoric correlation coefficient declines from 0.96 to 0.64, suggesting less than full persistence (no ρ-persistence). Similar results are found for any other initial value of corruption. Likewise, in the autoregressions the R^2 drops from more than 0.96 to only 0.55, suggesting that the initial level of corruption loses much of its power to explain the variation in the current levels of corruption. The table also shows that the constant (τ) and the slope (ω) of the autoregressions move away from zero and one, respectively; ω exhibits the same pattern as the three ρ's and R^2. This conclusion is supported by the result of the joint test that $\tau = 0$ and $\omega = 1$ in the last column Table 4.2. In sum: our evidence does not support corruption persistence.

Table 4.2: Corruption Correlation Over Time

i	ρ_1	ρ_2	ρ_3	$C_{1984+i} = \tau + \omega C_{1984}$			
				R^2	τ	ω	$F^{*]}$
1	0.979	0.996	0.965	0.958	0.272	0.959	6.660
2	0.960	0.970	0.950	0.921	0.400	0.928	5.880
3	0.941	0.960	0.938	0.886	0.586	0.884	8.370
4	0.920	0.952	0.918	0.846	0.757	0.849	10.710
5	0.887	0.912	0.875	0.786	0.819	0.825	8.880
6	0.866	0.892	0.848	0.749	0.944	0.807	9.950
7	0.864	0.872	0.833	0.746	1.078	0.786	13.560
8	0.829	0.851	0.803	0.688	1.388	0.746	20.150
9	0.804	0.813	0.771	0.647	1.690	0.695	29.170
10	0.794	0.796	0.747	0.631	1.888	0.649	37.590
11	0.769	0.777	0.734	0.591	2.018	0.608	39.720
12	0.766	0.746	0.717	0.587	1.912	0.607	34.300
13	0.693	0.669	0.630	0.480	2.039	0.542	33.340
14	0.643	0.641	0.602	0.413	2.001	0.516	30.790
15	0.606	0.587	0.522	0.368	1.989	0.490	32.480
16	0.641	0.622	0.567	0.411	1.851	0.510	34.040
17	0.665	0.620	0.573	0.442	1.635	0.527	38.600
18	0.749	0.685	0.630	0.561	1.122	0.568	72.540
19	0.740	0.680	0.639	0.548	1.266	0.551	68.830

ρ_1: Pearson, ρ_2: Spearman; ρ_3: Polychoric

*] Joint test F: τ=0 and ω=1

As an alternative to the OLS regressions, we also estimate ordered logit and probit regressions as perhaps not everyone might be convinced that our corruption data can be treated as cardinal data. Tables 4.3 and 4.4 show the results of the ordered logit and order probit estimations respectively. It is clear that the longer the time lag, the lower is the influence of the initial level of corruption (ω) on its current level. The estimate of ω reduces quickly until $i = 15$, from 6.09 and 0.75 in the logit models and from 3.22 to 0.44 in the probit models. A similar pattern is found for the pseudo-R^2. The ability of the initial level to explain the current level of corruption reduces if the time lag increases. So, these results are fully in line with those reported in Table 4.2.

Next, we adapt some terminology from the economic growth literature (see, for instance, Barro and Sala-i-Martin, 1992, 1995). First, there is corruption convergence if a corrupt country improves more than a clean country, so that the corrupt country catches up with the clean one. This type of convergence corresponds to the so-called β-convergence in the economic growth literature.

The second type of convergence occurs if the cross-sectional dispersion (measured by the cross-country standard deviation and the coefficient of variation of the level of corruption) declines over time; this type of convergence is akin to σ-convergence in the economic growth literature.

To examine β-convergence, we estimate a simple linear relation between the relative change in corruption over the period of 1984-2003—i.e., the change in the level of corruption over 20 years corrected by its initial value—and the initial level of corruption. We prefer the relative instead of the absolute change in the level of corruption because one point improvement (deterioration) for corrupt countries differs substantively from the same improvement (deterioration) for clean countries. In absolute terms, an increase of the corruption index from 1 to 2, for example, is exactly the same as an increase from 5 to 6, but it is completely different if the initial level of corruption is taken into account. Still, we also report estimates with the absolute change in corruption as dependent variable and other possible measures.[7]

[7] The correlations among these measures varies between 0.86 and 0.95.

Table 4.3: Ordered Logit Regression

i	Ordered Logit							Pseudo
	ω	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	R^2
1	6.089	8.700	13.023	20.102	26.105	33.418	39.428	0.765
	(0.890)	(1.935)	(1.919)	(2.895)	(3.601)	(5.014)	(5.856)	
2	3.963	5.766	7.905	13.134	17.627	21.697	25.594	0.626
	(0.465)	(1.163)	(1.114)	(1.627)	(1.984)	(2.615)	(3.065)	
3	3.494	4.945	6.743	11.380	15.654	19.199	23.490	0.579
	(0.411)	(1.063)	(1.030)	(1.432)	(1.776)	(2.301)	(2.820)	
4	3.019	4.035	5.431	9.637	13.644	16.184	20.775	0.521
	(0.356)	(0.956)	(0.923)	(1.244)	(1.577)	(1.890)	(2.464)	
5	2.322	3.168	4.223	7.500	10.484	12.711	15.992	0.423
	(0.266)	(0.806)	(0.779)	(0.971)	(1.187)	-1.447	(1.803)	
6	2.055	2.605	3.388	6.508	9.136	10.980	14.216	0.382
	(0.238)	(0.758)	(0.727)	(0.879)	(1.057)	(1.251)	(1.607)	
7	1.921	1.151	3.322	5.997	8.358	10.263	13.147	0.358
	(0.221)	(0.892)	(0.704)	(0.818)	(0.971)	(1.164)	(1.475)	
8	1.729	0.798	2.796	4.980	7.175	8.864	11.883	0.320
	(0.205)	(0.877)	(0.699)	(0.745)	(0.873)	(1.020)	(1.357)	
9	1.557	0.617	2.043	3.895	6.404	8.006	10.801	0.288
	(0.193)	(0.880)	(0.700)	(0.693)	(0.825)	(0.957)	(1.264)	
10	1.451	-0.356	1.586	3.360	5.990	7.823	10.152	0.273
	(0.183)	(1.117)	(0.708)	(0.659)	(0.793)	(0.951)	(1.193)	

continued on next page ...

i	Ordered Logit							Pseudo
	ω	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	R^2
11	1.315	-0.617	0.919	2.871	5.594	7.355	9.389	0.249
	(0.172)	(1.113)	(0.713)	(0.621)	(0.760)	(0.912)	(1.102)	
12	1.247	-0.669	0.873	3.091	5.672	6.979	9.610	0.232
	(0.170)	(1.120)	(0.728)	(0.629)	(0.776)	(0.899)	(1.140)	
13	0.970	0.738	2.674	4.463	6.244	7.887	–	0.165
	(0.146)	(0.623)	(0.580)	(0.669)	(0.823)	(0.966)		
14	0.846	0.610	2.599	4.124	5.628	7.107	–	0.136
	(0.136)	(0.572)	(0.561)	(0.637)	(0.751)	(0.884)		
15	0.754	0.490	2.581	4.027	5.259	6.542	–	0.108
	(0.134)	(0.568)	(0.575)	(0.641)	(0.730)	(0.858)		
16	0.819	0.605	2.701	4.664	5.636	6.939	–	0.133
	(0.135)	(0.553)	(0.560)	(0.676)	(0.749)	(0.874)		
17	0.877	-1.732	1.117	3.033	5.010	6.323	7.545	0.144
	(0.137)	(1.088)	(0.544)	(0.574)	(0.698)	(0.811)	(0.931)	
18	1.086	-1.265	2.499	5.243	6.493	8.222	10.659	0.202
	(0.157)	(1.091)	(0.595)	(0.734)	(0.849)	(1.011)	(1.409)	
19	1.058	-1.260	2.449	4.935	6.217	8.046	10.479	0.191
	(0.156)	(1.099)	(0.600)	(0.715)	(0.822)	(1.001)	(1.402)	

Note: Figures in brackets are the standard errors

The annual data is rescaled into an ordinal score (C is ICRG index):

1: $C \leq 1.5$; 2: $1.5 < C \leq 2.5$; 3: $2.5 < C \leq 3.5$; 4: $3.5 < C \leq 4.5$;

5: $4.5 < C \leq 5.5$; 6: $5.5 < C \leq 6.5$; 7: C > 6.5

Table 4.4: Ordered Probit Regression

i	Ordered Probit							Pseudo
	ω	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	R^2
1	3.217	4.623	7.130	10.749	13.902	17.646	20.826	0.761
	(0.360)	(0.877)	(0.872)	(1.240)	(1.475)	(2.050)	(2.385)	
2	2.057	3.099	4.077	6.790	9.222	11.098	13.255	0.605
	(0.205)	(0.539)	(0.525)	(0.749)	(0.895)	(1.117)	(1.366)	
3	1.823	2.655	3.491	5.852	8.225	9.913	12.247	0.558
	(0.184)	(0.510)	(0.496)	(0.654)	(0.807)	(0.997)	(1.259)	
4	1.605	2.212	2.883	5.059	7.264	8.481	11.060	0.507
	(0.166)	(0.484)	(0.468)	(0.596)	(0.742)	(0.842)	(1.156)	
5	1.235	1.726	2.273	3.937	5.589	6.733	8.554	0.410
	(0.124)	(0.421)	(0.406)	(0.467)	(0.559)	(0.662)	(0.849)	
6	1.122	1.485	1.896	3.512	4.985	5.975	7.805	0.375
	(0.116)	(0.408)	(0.392)	(0.439)	(0.515)	(0.597)	(0.787)	
7	1.055	0.728	1.892	3.263	4.587	5.626	7.263	0.351
	(0.109)	(0.478)	(0.382)	(0.415)	(0.480)	(0.564)	(0.731)	
8	0.969	0.480	1.576	2.762	4.019	4.987	6.706	0.322
	(0.104)	(0.464)	(0.380)	(0.391)	(0.443)	(0.514)	(0.687)	
9	0.870	0.369	1.131	2.126	3.562	4.491	6.095	0.292
	(0.097)	(0.454)	(0.373)	(0.363)	(0.414)	(0.479)	(0.636)	
10	0.821	-0.135	0.913	1.865	3.366	4.425	5.788	0.273
	(0.094)	(0.564)	(0.376)	(0.353)	(0.404)	(0.479)	(0.612)	

continued on next page ...

i	Ordered Probit							Pseudo
	ω	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	R^2
11	0.757	-0.186	0.608	1.644	3.219	4.245	5.453	0.253
	(0.090)	(0.538)	(0.372)	(0.339)	(0.398)	(0.469)	(0.575)	
12	0.715	-0.235	0.572	1.767	3.262	4.013	5.564	0.238
	(0.089)	(0.542)	(0.377)	(0.339)	(0.404)	(0.460)	(0.597)	
13	0.570	0.489	1.572	2.637	3.678	4.652	–	0.172
	(0.079)	(0.338)	(0.322)	(0.361)	(0.432)	(0.518)		
14	0.510	0.443	1.579	2.499	3.397	4.274	–	0.145
	(0.075)	(0.320)	(0.319)	(0.353)	(0.407)	(0.486)		
15	0.442	0.303	1.516	2.387	3.121	3.876	–	0.117
	(0.073)	(0.315)	(0.321)	(0.344)	(0.388)	(0.464)		
16	0.488	0.447	1.628	2.803	3.382	4.152	–	0.138
	(0.076)	(0.318)	(0.324)	(0.374)	(0.411)	(0.486)		
17	0.509	-0.613	0.683	1.774	2.942	3.712	4.426	0.145
	(0.076)	(0.455)	(0.315)	(0.327)	(0.379)	(0.437)	(0.510)	
18	0.615	-0.368	1.419	3.004	3.722	4.736	6.012	0.200
	(0.084)	(0.462)	(0.334)	(0.390)	(0.442)	(0.535)	(0.696)	
19	0.603	-0.391	1.375	2.836	3.587	4.662	5.934	0.193
(0.082)	(0.458)	(0.329)	(0.377)	(0.427)	(0.527)	(0.689)		

Note: Figures in brackets are the standard errors

The annual data is rescaled into an ordinal score (C is ICRG index):

1: $C \leq 1.5$; 2: $1.5 < C \leq 2.5$; 3: $2.5 < C \leq 3.5$; 4: $3.5 < C \leq 4.5$;

5: $4.5 < C \leq 5.5$; 6: $5.5 < C \leq 6.5$; 7: $C > 6.5$

In all regressions the coefficients of the initial level of corruption are negative and highly significant ($p < 0.01$, the t-values are in parentheses). The initial level of corruption explains around 29-44% of the variation in the change in corruption. Clearly, corrupt countries improve faster than clean countries, supporting β-convergence. However, various clean countries have also become less clean over the same period. In other words, Table 4.5 shows that there is not only a 'bottom-up' process, but also a 'top-down' process. This result is in line with our previous finding that there are both losers and winners in the world corruption league (see Table 4.1).

Table 4.5: Tests of β-Convergence

Dep. Var	Constant	β	β-std	R^2-Adj
Relative Change	0.525	-0.139	-0.582	0.330
[1984 and 2003]	(0.088)	(0.020)		
Absolute Change	1.266	-0.449	-0.667	0.439
[1984 and 2003]	(0.227)	(0.050)		
Relative Change	0.039	-0.009	-0.635	0.398
[Averaged-Annual]	(0.005)	(0.001)		
Absolute Change	0.067	-0.024	-0.667	0.440
[Averaged-Annual]	(0.012)	(0.003)		
Log $(\frac{C_{03}}{C_{84}})$	0.364	-0.117	-0.543	0.287
	(0.082)	(0.018)		

Note: Figures in brackets are the standard errors

Absence of σ-convergence would imply that the standard deviation and the coefficient of variation of cross-country corruption would be constant over time. However, we find that the standard deviation (SD) shrinks by more than 30%, from 1.63 to 1.22 (Table 4.6), suggesting σ-convergence. Also the coefficient of variation (CV) demonstrates a long-run decreasing trend over 1984-2003, although it reverses slightly after 1996. So also if the mean level of corruption is taken into account, corruption dispersion declines.

Table 4.6: Tests of σ-Convergence

Year	Mean	SD	CV	Year	Mean	SD	CV
1984	4.211	1.630	0.387	1994	4.622	1.332	0.288
1985	4.310	1.599	0.371	1995	4.577	1.288	0.281
1986	4.309	1.577	0.366	1996	4.471	1.292	0.289
1987	4.308	1.531	0.355	1997	4.321	1.275	0.295
1988	4.331	1.504	0.347	1998	4.173	1.308	0.313
1989	4.292	1.516	0.353	1999	4.051	1.316	0.325
1990	4.342	1.520	0.350	2000	4.000	1.298	0.324
1991	4.387	1.483	0.338	2001	3.853	1.291	0.335
1992	4.529	1.466	0.324	2002	3.515	1.237	0.352
1993	4.616	1.408	0.305	2003	3.588	1.215	0.339

4.3 A Closer Look

In the previous section we have presented some evidence on the absence of corruption persistence using some simple approaches borrowed from the economic growth literature. However, these approaches ignore the cross-country distribution of corruption and its dynamics over time. Using non-parametric approaches that are also borrowed from the economic growth literature (Quah, 1993a, b, and c; and Fingleton, 1997 and 1999), we provide some further evidence on corruption convergence.

We apply a kernel density estimation that provides information about the modality of the distribution over time, i.e., the bump(s) in a kernel density.[8] We are therefore able to examine whether a particular distribution with, say, two classes of countries (bimodal)—e.g., corrupt and clean countries—is transformed into a unimodal density.

For the dataset $X_1, X_2, ..., X_n$, a kernel density estimator is drawn from

[8] According to three tests on normality of the data (D'Agostino et al. (1990), Shapiro-Wilk (1965), and Shapiro-Francia (1972)), normality of the distribution should be rejected for years at the end of the sample period (1999-2003). The test on skewness captures a positive skewness (1984-1989) that becomes negative in 1990-1994, and then turns positive but close to zero. The kurtosis test detects that the distribution initially has thinner tails compared to the normal distribution, but at the end of the sample period it tends to have heavier tails.

an unknown probability density function, $f(x)$:

$$\hat{f}(x) = \frac{1}{nh_n} \sum_{i=1}^{n} k\left(\frac{x - X_i}{h_n}\right) \tag{4.1}$$

where k is a bounded non-negative kernel function satisfying $\int k(x)dx = 1$ and h is a sequence of positive numbers called the bandwidth. There are several kernels satisfying the condition $\int k(x)dx = 1$. However, the choice of a kernel function does not really matter, since the difference in efficiency among kernel functions is very small (Silverman, 1986; Table 3.1). What does matter is the choice of bandwidth h, because it determines the bias-variance trade-off in the density estimation.

To find an optimal bandwidth, one usually estimates the closeness of the estimator $\hat{f}$ to the true density f, where the estimate $\hat{f}$ depends on the data, the kernel, and the bandwidth. The most widely used approach is to minimize the mean integrated square error, $\mathrm{MISE}(\hat{f}) = E\int[\hat{f}(x) - f(x)]^2 dx$. Silverman (1986) argues that the optimal window width of a Gaussian kernel can be computed via $h_{opt} = 1.06\sigma n^{-\frac{1}{5}}$. We apply this to evaluate the 1984-2003 distributions of the level of corruption.

Figure 4.1 shows the evolution of the distribution of the level of corruption over 20 years. The vertical axis represents the estimated kernel density, while the horizontal axis corresponds to the corruption level (i.e., the rescaled ICRG index). Several conclusions can be drawn. First, the evolution is marked by a slow change in modality of the distribution from bimodal to unimodal as it starts with two bumps in 1984 and ends with a single bump in 2003, indicating convergence.[9] In 1984, the first bump is below a score of four of the ICRG index and around 0.25 of the kernel estimate. It moves slowly to the right to an ICRG score above four by 1993 and about 0.30 of the kernel density (the very right graph on the second row in Figure 1), but goes back to the left to a score around three for the ICRG index and 0.40 of the density by 2003. This suggests that there has been an

[9] As an experiment, we also checked the evolution of the distribution using mean-adjusted data and a set of optimum bandwidths. We find similar results (available on request).

improvement in world corruption in the first part of our sample period, but a worsening in the second half. There is a decrease in the number of clean countries over the same period as depicted by the dynamics of the second bump. In 1984, the second bump starts with a very high level of the ICRG index (about six), but it has dropped ten years later and vanishes in the next ten years.

Second, a similar pattern can be discerned in the development of the tails. During 1984-1990, the score of the ICRG index in the left tail of the distribution is close to one, suggesting that there are a few highly corrupt countries. Although the ICRG score of the left tail stays more or less at the same level, there is a clear tendency for the kernel estimate to move to zero during the years of 1991-1996. In the years 1997-2000, the kernel density increases to 0.1, while the ICRG index moves to two.

Third, the right tail of the distribution lies above 0.1 of the kernel (1984) and reaches its highest point in 1992 at about 0.15. However, it gradually goes down to close to zero. This confirms our finding that there is an increase in the number of clean countries over the first half of the period and a decrease in the second half.

Table 4.7 shows that in 1984, the world corruption distribution is clustered around two peaks (3.915 and 6.462). This polarization continues to be visible in most years until 1996; the distribution is marked by two convergence clubs—to borrow the terminology from the economic growth literature. In the first club, the corruption level moves from below to above four (the 'bottom-up' process) over 1984-1995. The second club shows the opposite pattern: the mode shrinks from above to below six (the 'top-down' process). It is interesting also to see that, within this period, the distance between the two midpoints of the bumps reduces from 2.56 (1984) to 1.15 (1995). Before these bumps vanish, three bumps emerge in 1996. Between 1997 and the end of the sample period, only one bump is detected that confirms the σ-convergence process.

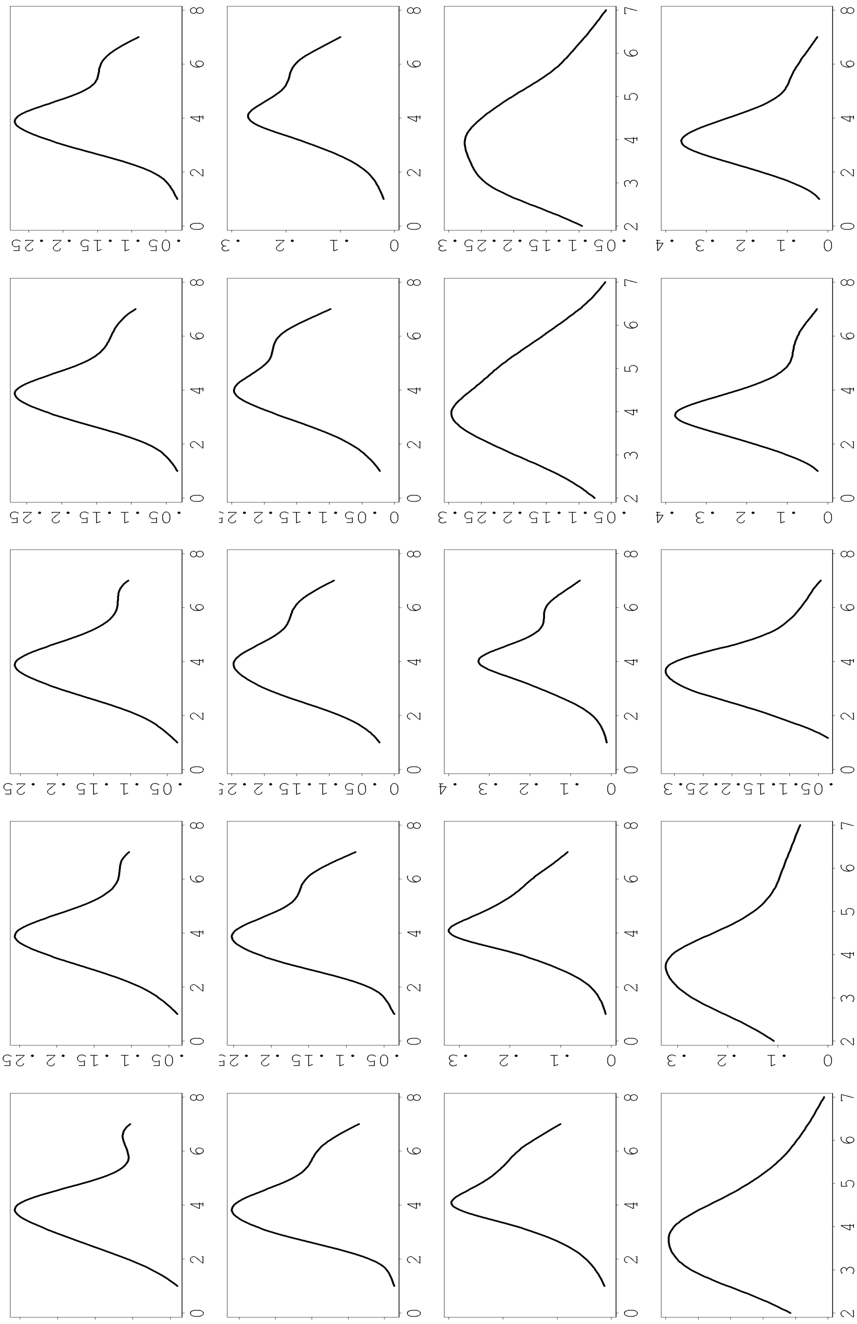

Figure 4.1: Annual Kernel Density, 1984 (top-left) to 2003 (bottom-right)

Table 4.7: Modality (Rescaled) ICRG data

Year	Bandwidth	Midpoint		
		Mode 1	Mode 2	Mode 3
1984	0.687	3.915	6.462	-
1985	0.672	3.915	-	-
1986	0.664	3.916	-	-
1987	0.645	3.916	5.429	-
1988	0.633	3.916	5.760	-
1989	0.639	3.916	5.511	-
1990	0.640	3.916	5.679	-
1991	0.625	3.916	5.590	-
1992	0.618	4.000	5.253	5.671
1993	0.593	4.085	5.530	-
1994	0.561	4.000	5.236	-
1995	0.542	4.000	5.148	-
1996	0.544	4.000	5.395	5.641
1997	0.537	3.940	-	-
1998	0.551	3.938	-	-
1999	0.554	3.937	-	-
2000	0.547	3.868	-	-
2001	0.544	3.845	-	-
2002	0.521	3.016	-	-
2003	0.512	3.104	-	-

4.4 Markov Chain Analysis

Now we turn to Markov chain analysis as an alternative way to examine the dynamics of corruption over time. We may describe a Markov chain as follows. In a set of states $N = \{1, 2, \ldots, N\}$, the Markov process starts in one of these states and moves from one state to another. The chain moves from the current state i to the next step j with a (transition) probability p_{ij}, where the probability does not depend on which states the chain was in before the current state.[10] Adopting Bickenbach and Bode (2003) who investigate the dynamics of income per capita, here we also assume that corruption level follows a finite first-order Markov chain with stationary transi-

[10] However, the process may remain in the state it is in, with probability p_{ii}.

tion probabilities. A finite first-order Markov process with stationary transition probabilities is a stochastic process $\{X_t, t \in T\}$ with $T = \{0, 1, 2, \ldots\}$ satisfying the Markov property (Bickenbach and Bode, 2003):

$$\begin{aligned} &P[X(t) = j|X(t-1) = i, X(t-2) = i_{t-2}, \ldots, X(0) = i_0] = \\ &P[X(t) = j|X(t-1) = i] \end{aligned} \tag{4.2}$$

for all t and all possible states $j, i, i_k (k = 0, 1, \ldots, t-2)$.

If the transition probabilities $p_{ij}(t) := P\{X(t) = j|X(t-1) = i\}$ do not depend on t ($p_{ij}(t) = p_{ij}$ for all $t \in T$), the Markov chain is said to be time stationary. Here, a time-stationary Markov chain is determined by the Markov transition matrix $\mathcal{P}$,

$$\mathcal{P}_{t,t+s} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{bmatrix}$$

$\forall_i$, $0 \leq p_{ij} \leq 1$, and the initial distribution $F(0) = (F_1(0), F_2(0), \ldots, F_N(0))$ where $\sum F_i(0) = 1$. Since $\mathcal{P}$ is a right-row stochastic matrix, then $\sum_j p_{ij} = 1$.[11] This transition matrix can be estimated via Maximum Likelihood (ML) in which case the estimate is consistent with the relative frequency of the transitions observed over the period of analysis.

The question now is what is the probability that after a certain transition period the Markov chain is in a given state? For a given law of motion, as determined by $\mathcal{P}$, and given initial distribution $F(0)$, the distribution after the first transition period (t=1) is

$$F(1) \quad = \quad F(0) \times \mathcal{P}_{t,t+S}, \tag{4.3}$$

[11] We may say that p_{ij} is the probability for a country to move from corruption level i to level j during the time span s.

or, in general,

$$F_{x,t+S} \quad = \quad F_{x,t} \times \mathcal{P}_{t,t+S}. \tag{4.4}$$

The distribution after the m-th transition period is:

$$F_{x,t+(m \times s)} = F_{x,t} \times \mathcal{P}^{m}_{t,t+s}. \tag{4.5}$$

Provided the Markov chain is regular, the distribution converges toward the limiting (ergodic) distribution is given by

$$F^{*} = \lim_{m \to \infty} F(0)\mathcal{P}^{m}_{t,t+s}. \tag{4.6}$$

This can be obtained via a $1 \times N$ row vector π such that

$$\pi = \pi \mathcal{P} \tag{4.7}$$

as m tends to infinity. Comparing this limiting distribution to the initial distribution of corruption can tell us whether countries converge or diverge. Higher probabilities in the median-states of the limiting distribution, as compared to the initial distribution, indicate convergence.

A critical issue here is to define the number of states as well as the length of the interval between states. This is somewhat arbitrary and researchers usually consider the trade-off between the number of observations in every state and the inter-state dynamics. Some researchers first divide the observed countries evenly and determine the length of the interval later (see, for instance, Bickenbach and Bode (2003), Hammond and Thompson (2006)). Others determine the length of the interval first and then decide upon the distribution—which is equally plausible (see, for example, Quah, 1993b; Fingleton, 1997)). In our case, due to the nature of the data, it is difficult to generate a particular interval that results in an equal distribution. We, therefore, split the countries into six states according to their level of corruption, where the distance from a state to another is the same—consequently, the distribution is unequal.

Following Shorrocks (1978), we also measure the degree of mobility as:

$$\text{Mobility} = \frac{N - \text{tr}(\mathcal{P}_{t,t+s})}{N - 1}, \tag{4.8}$$

where N is the number of states. This index ranges between zero and $\frac{N}{N-1}$ where a higher value means more mobility across corruption classes.[12]

Using the same data as in the ordered logit and probit regressions, Table 4.8 gives the results of the Markov chain analysis: the transition matrix, the initial distribution, the limiting distribution, and the mobility index. The first panel displays the short-run dynamics in which the transition matrix is measured as the average of one-year transition over every year from 1984-1985 to 2002-2003. Not surprisingly, as countries are concentrated in the main diagonal of this matrix over a one-year horizon, there is a high degree of corruption persistence. The first row shows that over the full sample (i.e., 101 countries and 20 years) 41 observations fall in state 1. Of these, about 83 per cent remain in that state in the following year. Similarly, of the 165 observations in state 7, 91 per cent stay in that state in the following year. In line with our previous findings, it turns out that few countries exhibit a substantial change over this one-year horizon. For example, only 17 per cent of the countries move from state 1 to states 2 and 3, while 11 per cent jump from state 4 to states 2 and 3. The mobility index is very low (0.18), confirming that one-year transitions do not indicate convergence.

However, comparing the limiting distribution with the initial one suggests that there is convergence. First, the share of the very upper tail (state 7) greatly reduces from nine per cent in the initial year to less than one per cent in the limiting distribution. Second, the fraction of countries in state 2 almost doubles from seven to 14 per cent. Third, while no substantial change is found for the shares of states 1 and 4, and a slight decrease occurs

[12] As indicated by the trace of the matrix, an index of zero simply means that all observations lie on the main diagonal. A value of $\frac{N}{N-1}$ implies that none of the observations lies on the diagonal. Thus, in our case the index ranges between 0.00 and 1.20.

for that of state 5, there is a substantial rise in the share of state 3, from 22 to 33 per cent. In short, these findings show that some countries become less corrupt while others become less clean.

Table 4.8: Markov Chain Estimates, 1984-2003

Transition	State 1	State 2	State 3	State 4	State 5	State 6	State 7
	Annual Transition (1984-2003)						
State 1 (41)	0.829	0.122	0.049	0.000	0.000	0.000	0.000
State 2 (134)	0.030	0.813	0.119	0.037	0.000	0.000	0.000
State 3 (419)	0.000	0.069	0.828	0.086	0.017	0.000	0.000
State 4 (594)	0.000	0.002	0.109	0.832	0.054	0.003	0.000
State 5 (307)	0.000	0.000	0.033	0.127	0.811	0.029	0.000
State 6 (259)	0.000	0.000	0.004	0.008	0.085	0.900	0.004
State 7 (165)	0.000	0.000	0.000	0.000	0.012	0.073	0.915
Initial Dist.	0.021	0.070	0.218	0.310	0.160	0.135	0.086
1st Iteration	0.020	0.075	0.230	0.300	0.163	0.133	0.079
19th Iteration	0.020	0.120	0.298	0.297	0.157	0.089	0.020
Limiting Dist.	0.025	0.141	0.329	0.309	0.142	0.052	0.002
Mobility Idx.	0.179						
	20-Year Transition (1984 and 2003)						
State 1 (4)	0.000	0.750	0.250	0.000	0.000	0.000	0.000
State 2 (10)	0.000	0.300	0.500	0.200	0.000	0.000	0.000
State 3 (19)	0.000	0.474	0.368	0.158	0.000	0.000	0.000
State 4 (35)	0.029	0.143	0.514	0.257	0.057	0.000	0.000
State 5 (9)	0.000	0.000	0.667	0.222	0.111	0.000	0.000
State 6 (10)	0.000	0.000	0.300	0.000	0.600	0.100	0.000
State 7 (14)	0.000	0.000	0.071	0.071	0.357	0.429	0.071
Initial Dist.	0.040	0.099	0.188	0.347	0.089	0.099	0.139
1st Iteration	0.010	0.198	0.406	0.168	0.139	0.069	0.010
Limiting Dist.	0.005	0.346	0.445	0.191	0.012	0.000	0.000
Mobility Idx.	0.965						

The convergence phenomenon is much more visible when we use a 20-year horizon of the transition matrix as shown in the second panel of Table 4.8. A direct indication of convergence is provided by the main diagonal: only 37 per cent of the countries stay in the same state between two distant years. Compared to the annual transition, the number of entries on the main diagonal reduces sharply. For example, no single country is found to remain in state 1. Rather, all countries that were in state 1 (in 1984) improve their level to states 2 and 3 (in 2003). On the contrary, all countries that were in state 4 drop down to states 1 (3 per cent), 2 (14 per cent), 3 (51 per cent), or jump to state 5 (6 per cent). These upward and downward inter-state shift dynamics are reflected in the very high value of the Shorrocks index of 0.96.

4.5 Conclusion

Is corruption really persistent? Using ICRG data over the period 1984-2003 we argue that it is not. We find that many countries see their level of corruption decrease or increase. There are more countries with a decline than countries with a rise in their level of corruption. The correlation between corruption in two years is high, but the larger the distance between the two years becomes, the lower is the correlation, so there is no ρ-persistence. In a very simple autoregressive model of corruption, the coefficient of the lagged dependent variable and the R^2 reduce over time. This applies also to the similar ordered logit and probit models.

We find strong evidence for β-convergence and σ-convergence. Not only do corrupt countries become less corrupt (catching-up), but clean countries tend to become less clean over time. This pattern is confirmed in our analysis of the dynamics of the distribution of the data. There is a significant modality shift in the corruption distribution over time. Using a Gaussian kernel function we detect a transformation in the corruption distribution from a bimodal to a unimodal distribution. Finally, on the basis of a Markov chain analysis, we also find interclass upward and downward shifts of countries.

Our analysis shows that research should not focus on the explanation of corruption persistence, but should be redirected towards explaining the process of corruption convergence.