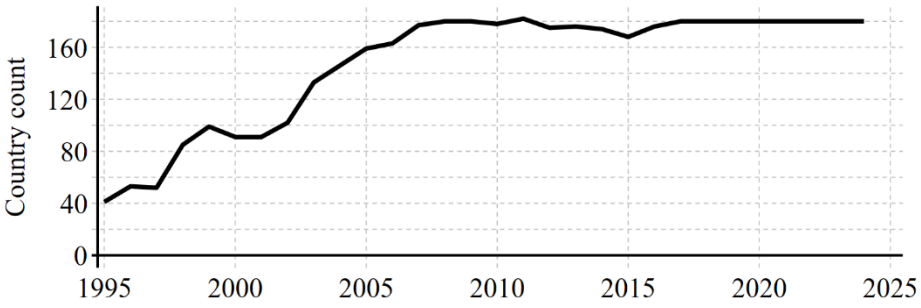
15/11-2025

Martin Paldam

The average of the corruption index – does it mean anything?

This note considers the *TI* ***honesty index*** from Transparency International. It is available from: https://www.transparency.org/en/cpi/2024. It is turned into *T* ***corruption index*** by: $T = 10 - TI$. It is defined on the open interval]0, 10[and given with 1 decimal. The index started in 1995 for 40 countries. The coverage grew to 180 countries in 2007 and have stayed about that since then. There are data for 188 countries, but some countries are covered for a few years only.

Figure 1. The annual number of countries covered by TI's corruption index



1. The development of the average over time – no clear trends

Figures 2 and 4 have lines for two samples. The black line is for the sample of 40 countries that are covered every year. One missing observation has been interpolated. The gray line is the sample for all observations.

Figure 2. The development in the average corruption index over time

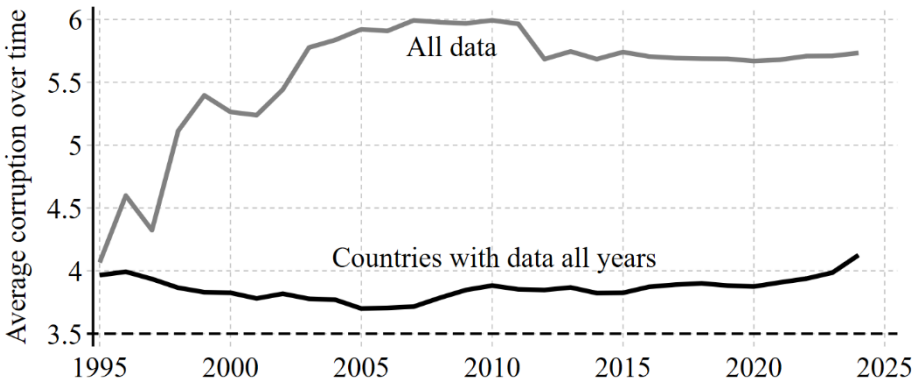


Table 1. Tests for trends in Figure 2

	Years		Constant		
	Estimate	t-ratio	Estimate	t-ratio	R^2 adj
Black sample	0.0035	(1.9)	-3.12	(-0.9)	0.083
Gray sample	0.0321	(3.9)	58.99	(3.5)	0.326

The t-ratio of 1.9 for $N = 30$ has a p-value of 6.7%.

Figure 2 looks at the trends over the 30 years for the two samples. The curve for the black sample is virtually trendless; see Table 1. From 2007 when the number of countries has stabilized the curve for the gray sample becomes trendless as well. In the 30 years covered world GDP per capita has doubled. So, the level of corruption in the world should have fallen substantially, but it increased marginally. The *TI* index is an average of a dozen primary indices. The reason that it has not fallen is the complex calibration that is made to allow the primary indices to be averaged.

2. The main country groups – the great deviation of the West

Table 2 shows the levels of corruption in the main country-groups. The five groups follow the World Bank classification. The Post Socialist countries will soon need another name.

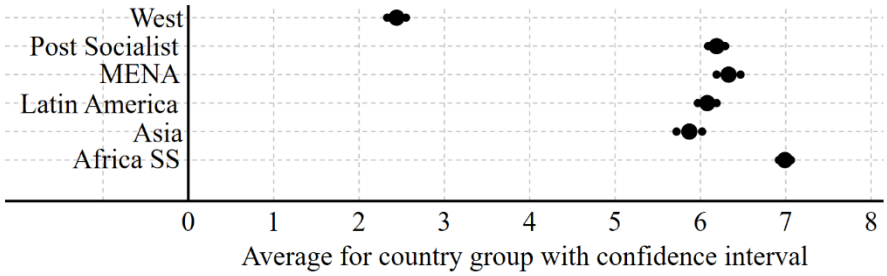
Especially the West stands out. Half of the 40 countries in the black sample are western, so the average is relatively low. The Av (average) and 2se (confidence intervals) columns from Table 2 are depicted in Figure 3. Both the poorest and richest country group – Africa and (especially) West – differ from the other groups. The other four groups overlap.

Table 2. The average corruption for country groups

Country group	Av.	2se	Nc	N	Missing
Africa	6.99	±0.07	47	1050	25.5%
Asia	5.87	±0.14	34	728	28.4%
Latin America	6.08	±0.11	33	738	25.5%
MENA	6.20	±0.14	19	427	25.1
Post-Socialist	6.19	±0.10	30	750	16.7
West	2.44	±0.11	25	727	3.1%
All	5.69	±0.06	188	4419	21.6%

The two numbers Nc and N are for countries and observations. Note that Nc adds to 188, while the high value in Figure 1 is only 180, as some countries are covered for a few years only.

Figure 3. Illustration of Table 2



3. The development in the standard deviations in the two samples -- convergence

While the *T*-index is virtually trendless it has strong trends in its standard deviation, and it is similar in the two samples as shown in Figure 4. Table 3 confirms that the trends are very significant.

Figure 4. The standard deviations of corruption in the two samples

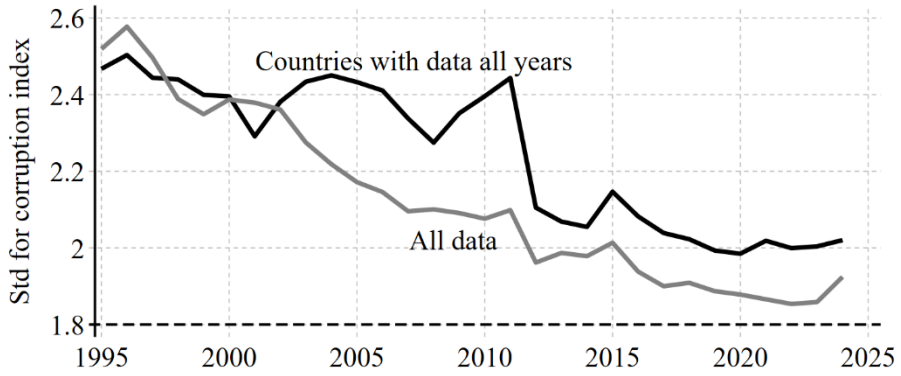


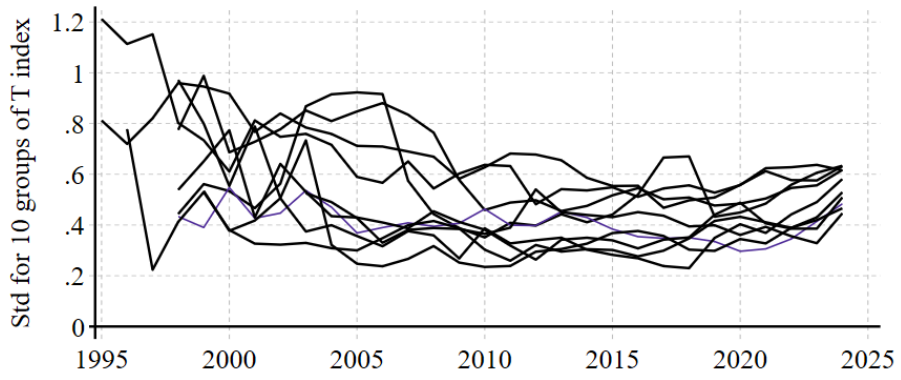
Table 3. Tests for trends in Figure 4

	Years		Constant		
	Estimate	t-ratio	Estimate	t-ratio	R^2 adj
Black sample	-0.0195	(-11)	41.42	(11)	0.806
Gray sample	-0.0243	(-19)	50.96	(20)	0.927

As the average of the two samples have no trends, the std can be seen as the σ-measure of convergence. Thus, Figure 4 says that corruption in the world has converged.

Figure 5 illustrates the convergence. The figure is made in four steps: (i) the T-matrix is sorted by the average corruption, (ii) the 188 country-rows in the sorted matrix is divided into 10 groups. (iii) The standard deviation every year for each group every year with more than 5 observations (iv) The resulting 10 rows are drawn on Figure 5.

Figure 5. The convergence of the standard deviations of the 10 groups, see text



It is obvious that the lines get closer over time. The vertical range of the lines are about 0.6 before 2005 and only 0.3 after 2020.

I think that we must interpret Figure 2 as saying the *T*-index is relative to a constant level. That makes it tricky to give an economic interpretation to Figure 4 and 5. However, the story told by Figure 3 is surely true, especially as the pattern found is so strong.