

Two corruption indices

Martin Paldam, Aarhus University, Denmark

1. Introduction

The two indices are T , Transparency International's corruption *perception* index, and Tw , the World Bank's corruption *control* index. The terms in *italic* indicate a conceptual difference, but as shown in the right-hand panel of Table 1 the difference is small empirically. Only the within-country correlation is moderate in size. The scale of both indices is from corruption to honesty: low scores are for high corruption, and high scores are for honesty. The left side panel of Table 1 shows the scales of the two indices.

Table 1. Descriptive statistics for the two indices

$N = 3,876$	T	Tw	Correlations $\text{cor}(T, Tw)$		
Average	4.286	-0.046	Unified	Between	Within
Standard deviation	2.067	1.034	Data	countries	Countries
Max observation	10	2.48	0.971	0.985	0.448
Min observation	0.7	-1.81	$N = 3,876$	$n = 183$	Avr 183 ^{a)}
Range	9.3	4.29	Discussed in section 4		

T is the Transparency International corruption perception index,

Tw is the World Bank corruption control index. The normalized indices are

a. The within-country correlation is the average of the individual ones for the countries.

Transparency International makes its index from a collection of (all?) primary cross-country indices it can find. Most primary indices are for a limited country group. Thus, they must be calibrated to approximately the same scale, and the average index is reported if it is based on at least three primary indices. Big or especially newsworthy countries may be covered by as many as a dozen primary indices. In addition to the average index the standard deviation of the calibrated primary indices is also reported.¹

¹ This allows an estimate of the substantial measurement error for the T -index, see M. Paldam, Institutional indices: Distributions and measurement errors. Working paper available from <http://martin.paldam.dk>.

The World Bank index starts from the Transparency index, but it includes countries with less than three primary indices, and institutional evidence. It is calibrated to average zero and standard deviation one.² Thanks to the calibration methods used both indices are relative. Section 5 shows that apart from the first few observations they are trendless. It is not possible to study international trends in corruption using the indices.

Table 2 is made for easy reference. The reader should tag this table!

Table 2. Variables and counts used

Variables:	Subscript n indicates a normalization to average zero and standard deviation one ^{a)}
T, T_n	Transparency Internationals corruption perception index, scale [0, 10], scale observed [0.7, 10]
Tw, Tw_n	World Bank corruption control index, scale observed [-1.81, 2.48]
y	Income, $\ln \text{gdp}$, where gdp is real GDP/GNI per capita, from the Maddison project
Counts:	They are connected by: $\sum_{i=1}^n j_i = N$. The average j is $3,876/183 = 21.2$
N	number of observations in the unified data, $N = 3,876$ or less
n	number of countries, $n = 183$ or less
j_i	number of observations for country i . It is 25 or less. Only 8 countries have $n < 10$

a. The normalized indices are $T_{ni} = (T_i - 4.286)/2.067$ and $Tw_{ni} = (Tw_i + 0.046)/1.034$, for country i .

² The overlapping data necessary for comparison misses some observations for both indices. Thus, the average and standard deviation deviate a little from the zero/one in the full data.

2. Two transitions

Before the start of the century, studies of corruption were based on intuitive and anecdotal evidence, but in 1995 the transparency index started for 41 countries. The number of countries rapidly increased (see Figure 7 below), and trust in the index grew. This led to a large empirical literature. The author participated from the start,³ and showed that the T -index has a typical transition like other institutions.⁴ That is, the level is stable but different in traditional and modern society, as shown by Figure 1. It is high in traditional societies and low in modern ones.

Figure 1a. Transition of corruption by the T -index

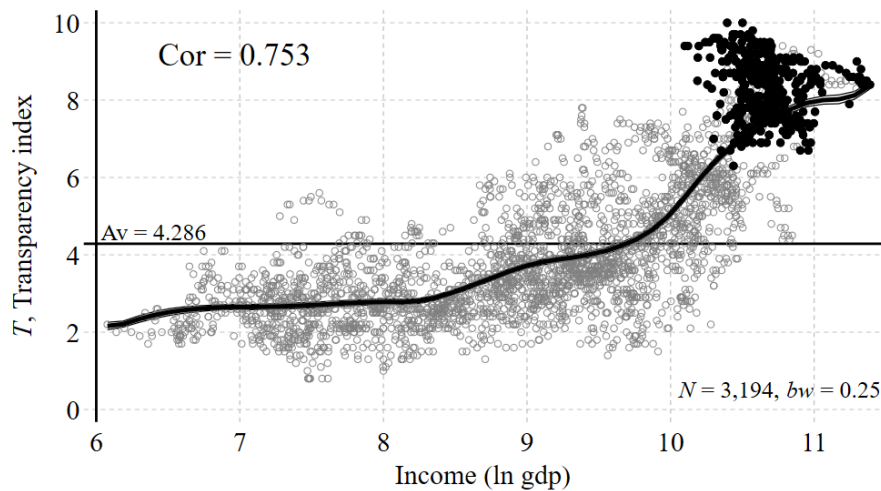
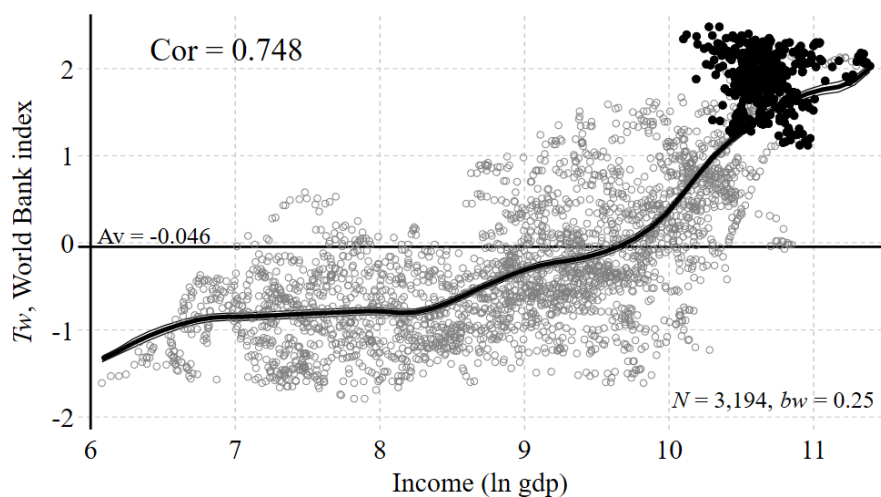


Figure 1b. Transition of corruption by the Tw -index



³ See Paldam, M., 2002. The cross-country pattern of corruption: Economics, culture, and the seesaw dynamics. *European Journal of Political Economy* 18, 215-40. And *The Grand pattern of Development and the transition of Institutions*. Cambridge UP, NY 2021.

⁴ The term institution is woolly. It is assumed that corruption is an institution, like democracy, religion etc.

Transitions of institutions are relatively late. It requires that countries are wealthy and have been so for some time. Figure 1a shows the transition in the T -index. Income is defined as the (natural) logarithm to the real GDP/GNP per capita (in PPP) prices. The curve is a moving average estimated by kernel regression, with the bandwidth of 0.25 income points. While corruption is high in traditional societies it is low in modern ones. Figure 1a shows the transition in the T -index.

The present version of the (robust) curve uses data for the years 1996 to 2022, a few countries are missing income data, and the OPEC countries are omitted as they have a different transition. Still there are 3,194 observations, where the data are available.

The observations for the old West are reported in black in the figures. They are all on the right side of the graphs. A dozen countries are converging to the West.⁵ They are not singled out to prevent cluttering the graphs, but they are on the way to low corruption. This confirms that convergence to the modern steady state needs high income and some time as well.

Figure 1b uses the same format as Figure 1a, but for the Tw index. It is similar, showing precisely the same transition. The convergence to the modern steady state is perhaps slightly slower but the difference is marginal.

⁵ The Old West is the following 18 countries: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Iceland, Ireland, Luxembourg, Netherlands, New Zealand, Norway, Sweden, Switzerland, UK, USA. They are joined by two groups of convergers: The new west Cyprus, Greece, Israel, Malta, Portugal, and Spain, and the new rich Asian countries Japan, Hong Kong, Singapore, South Korea and Taiwan. One converger has joined the top-ten list: Singapore.

3. Distributions

Figure 2 shows the frequency distributions over the range of each index divided into almost the same number of bins.

Figure 2a. Distribution of T , the Transparency index (20 bins)

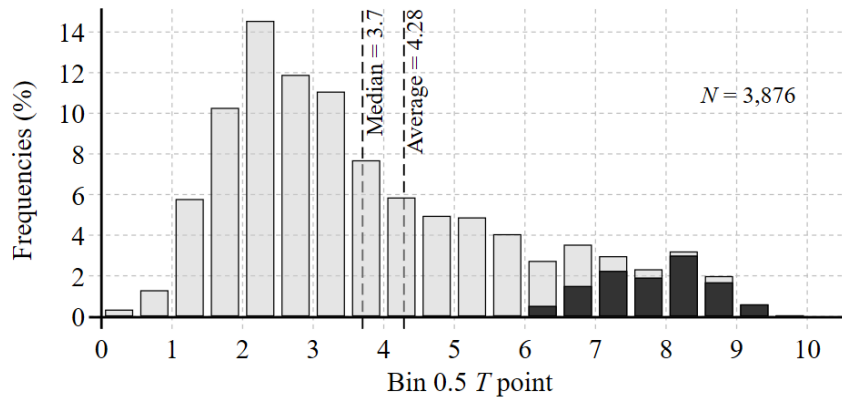
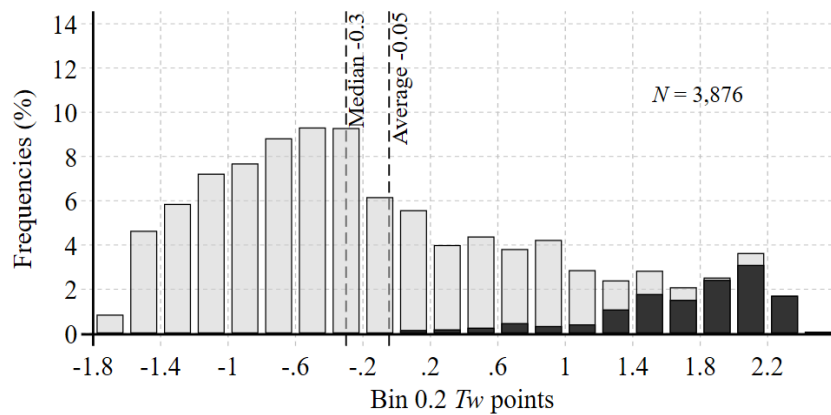


Figure 2b. Distribution of T_w , the World Bank index (22 bins)



Both frequency distributions are two-topped with a large top to the left, and a small one to the right that consists mainly of western countries. This is much as predicted from Figure 1. As more countries are reaching high incomes the peak to the right will become stronger.

To compare the statistics for the two curves they are normalized, as described in Table 2. Table 3 give the standard statistics, which look as expected from Figures 2. It is interesting to note that the Dif series ($Dif = T_n - Tw_n$) is too peaked and have a tail to the left, as seen on Figure 3, which shows a distribution that is close to normal, but a little skewed to the left, indicating a small difference in the skewness of the two indices. Thus, the two peaks are the same in T and T_w . The countries of the old west have much the same distribution, but with a

slightly smaller standard deviation. Figure 4 is the (Tw, T) -scatter, with a kernel average. The 95% confidence intervals are so narrow that even small deviations from linearity are detected. The ones observed are negligible.

The standard deviation of the Dif series is a measure of the uncertainty of the measure, but it is a dubious measure as the T index is a part of the Tw index. But it makes it clear that corruption measures have a substantial measurement error.

Table 3. Statistics for the normalized data

	T_n	Tw_n	Dif
Average	Zero per definition		
Standard deviation	One per definition		0.240
Median	-0.284	-0.246	0.030
Skewness	0.808	0.614	-0.770
Tail	right	right	left
Kurtosis	-0.282	-0.561	1.980
Form	flat	flat	peaked

Figure 3. Difference between the normalized indices, see note to Table 2

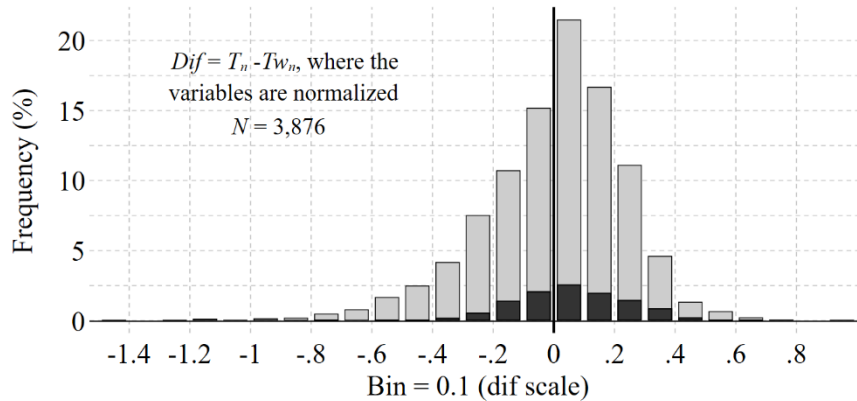
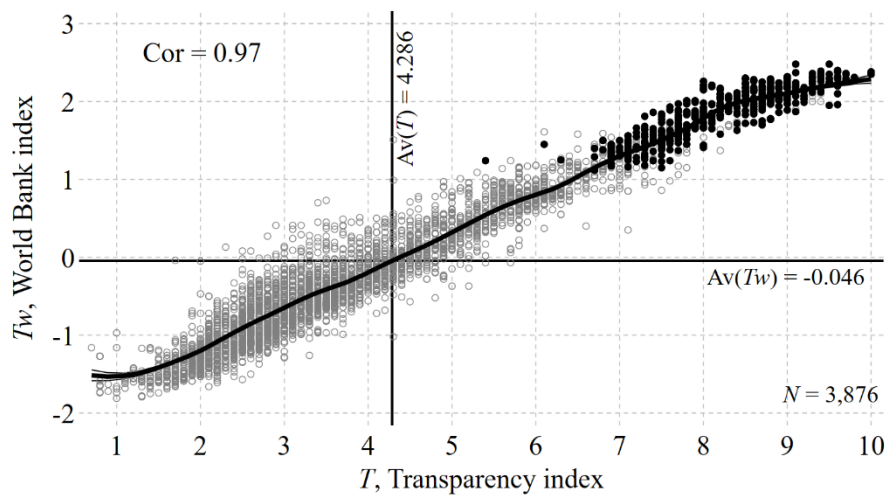


Figure 4. Scatter of the (Tw, T) pairs, with a kernel curve



4. Correlations: between and within countries

The right-hand panel of Table 1 showed three correlations. For the $N = 3,876$ unified data $\text{cor}_u = 0.971$ (also reported on Figure 4). It is broken into the between-country correlation, $\text{cor}_b = 0.985$, and the (average) within-country correlation $\text{cor}_w = 0.448$.⁶ While cor_u and cor_b are very high and virtually the same, cor_w is about half only.

Figure 5. The distribution of cor_w , the within-countries correlations (20 bins)

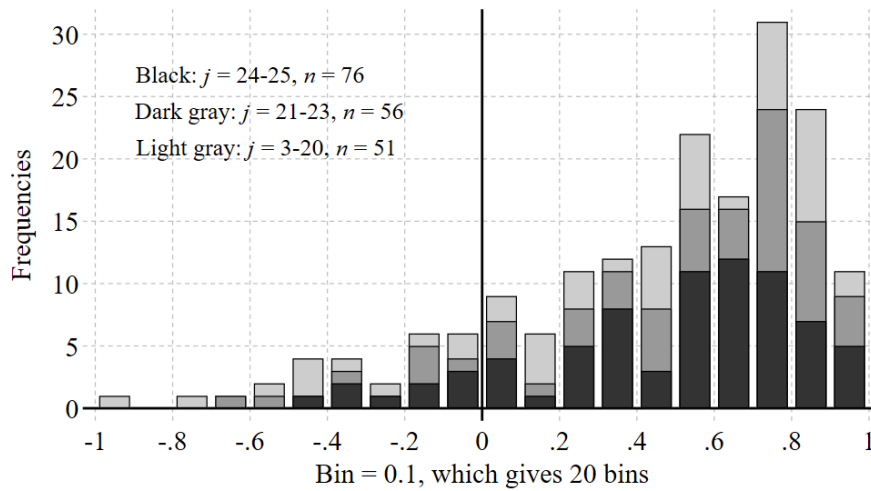


Table 4. Regressions explaining cor_w , the within country correlation, brackets holds t-ratios

Reg	Con	j (t)	Av T (t)	Av Tw	Av y	T trend	Tw trend	Dif	R^2 adj	n
1	0.194 (1.4)	0.012 (1.9)							0.014	183
2	0.395 (5.5)		0.013 (0.8)						-0.002	183
3	0.449 (15)			0.020 (0.7)					-0.003	183
4	0.073 (0.3)				0.039 (1.4)				0.007	151
5	0.432 (14)					0.009 (2.1)			0.019	183
6	0.446 (16)						0.021 (4.8)		0.108	183
7	0.466 (15)							-0.011 (2.3)	0.024	183

⁶ The within-country correlation is the average of the correlations calculated for each of the $n = 183$ counties.

Figure 5 demonstrates that the 183 cor_w are very variable, as they cover the interval $[-0.961, 0.967]$. The three levels of the j -variable indicated by the color of the bars suggest that the dependency of j is weak.

Table 4 provides seven regressions trying to explain the within-country correlations by seven variables, one at a time. Reg 1 shows that j , as suggested by Figure 5, has a weak relation to the correlation. Reg 2, 3 and 4 show that the averages of T , Tw , and y have no effect on the correlation. The last three regressions analyze if the correlations are due to (relative) trends in T and Tw and the difference between the trends.⁷ Here the coefficients are significant, but rather small. The largest R2-score is 0.1 to the Tw trend.

Consequently, the table demonstrates that none of the seven variables tried can explain the great variation in cor_w . The seven variables are surely not all variables, but perhaps they are the most likely. It is tempting to say that the random element in cor_w is large.

⁷ The trends are estimated as the slope β in the regression T or $Tw = \alpha + \beta \text{Year}$.

5. Time paths of the two indices

Figure 6 shows the (average) path over time of the 3,876 observations of the two indices, and Figure 7 shows the number of countries, n , each year. The path over time is strongly affected by the calibration methods used. Figure 6 reports the 95% confidence intervals (calculated as twice the standard error) as the two thin lines. It shows that the two series are trendless from 2002, when the data has reached 130 countries.

Figure 6a. Path of the T index

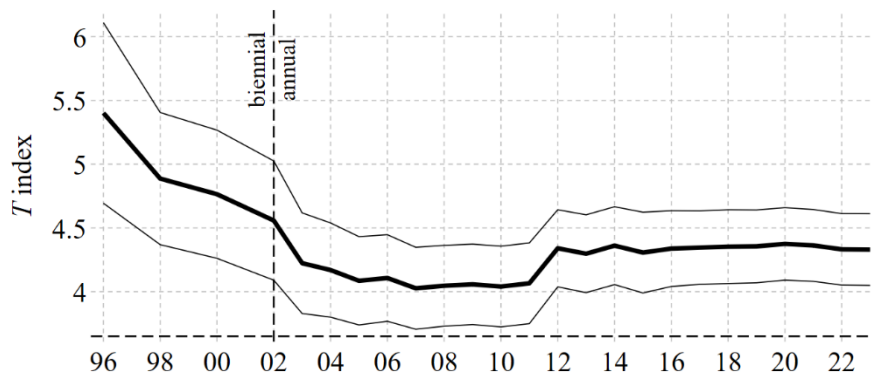


Figure 6b. Path of the Tw index

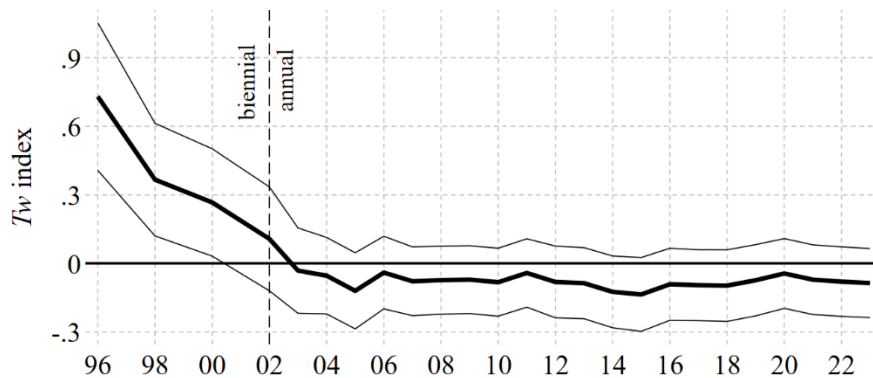
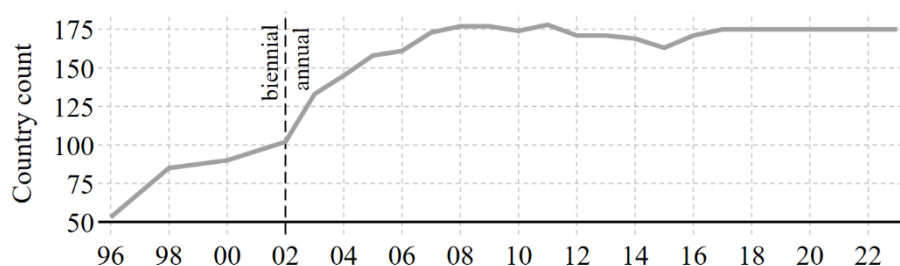


Figure 7. Number of countries included each year



The 25 years covered by the analysis have seen a doubling of real gdp in the world, so income has grown with 0.7 log points. In the perspective of Figure 1 there has surely been a fall in corruption in the world, but thanks to the calibration of the indices it does not show. The indices are relative, not absolute. The growth literature has shown that income grows almost linearly. Thus, T and T_w are relative variables while y is an absolute one.

6. Conclusions including a note on the politics of corruption indices

The first paragraph in the paper mentioned that the two indices are conceptually different. It is not problematic that T is a *perception* index, which is true. If people on average have rational perceptions, it measures corruption.⁸

However, it is a bit strange that Tw is termed a corruption *control* index. It is unclear what it means. It seems to imply that corruption would be the same in all countries, but for the different systems controlling corruption. This is surely nonsense. Thus, my interpretation of the term control in the Tw index is that it serves as an expression of a wish. But perhaps it is a vestigial element left from the old politics of corruption research.

In the 1970s and 80s a left-wing part of development thinking was prominent: The theory of imperialism claimed that the (relative?) world income was a zero-sum game, where the wealth of the DCs was caused by the exploitation of the LDCs, making them poor. Many leaders of poor countries liked that theory. They thought that it was to add insults to injury to claim that poor countries were more corrupt. The World Bank works in most countries, so it became bad taste to observe that poor countries are more corrupt than rich countries. The problem of corruption was swept under the carpet in discussions of development.⁹

It was probably a reaction to this suppression that the high World Bank official Peter Eigen took early retirement and founded the NGO Transparency International in 1993. It is devoted to fighting corruption, notably in the poor countries, and the organization claims that corruption is a main development problem. The T index was developed by the research chief of the organization, Johan Graf Lambsdorff. It has been a huge success, which has done much to allow the discussion of corruption, and to make it an issue in development that cannot be swept under the carpet. In periods the World Bank has joined anti-corruption, and as discussed, it has also made a corruption index.

Fortunately, the two indices measure almost the same, so the conceptual difference does not matter in practice.

⁸ This should be the case also by Condorcet's Jury theorem.

⁹ The thinking of the development establishment around the World Bank is expressed in the 32 chapters of the 1773 pages of the first two volumes of the North Holland Handbook of Development Economics issued in the mid-1980s. No chapter dealt with corruption, or the related subject of rent seeking, and these subjects were not mentioned in the comprehensive index.