

Corruption and religiosity

What do we know?

Martin Paldam,¹ Aarhus University, Denmark

Abstract

The effect of religiosity on corruption is much debated. The paper develops robust and transparent results using a new theory-near measure of religiosity on a cross-country data set. Corruption and religiosity have a strong and stable positive correlation: Religious countries are more corrupt: However, most of the effect is spurious. When the relation is controlled for income, the positive relation becomes much weaker. The effect of democracy on corruption is negative: Democracies are more honest. However, most of the effect is spurious. The transition of corruption, the religious transition and the democratic transition are all strong and parallel functions of income, giving substantial multicollinearity in the relations between the variables.

Jel: K42, Z12

Keywords: Corruption, religiosity, democracy

¹ Department of Economics and Business, Fuglesangs Allé 4, 8210 Aarhus V, Denmark.
E-mail mpaldam@econ.au.dk. Home page <http://www.martin.paldam.dk>.
The data used are available from <http://www.martin.paldam.dk/GT-Religion.php>.

1. Introduction

A huge literature discusses the effect of religiosity on corruption; see Table 1. The mountain of discussion stands on a small empirical basis, as religiosity is difficult to measure. Hence, much of the discussion deals with issues of moral philosophy and theology. The following is an attempt to strengthen the empirical foundation of the discussion, i.e., to give robust cross-country evidence, using a theory-near measure of religiosity on cross-country data.² An even larger literature discusses the effect of democracy on corruption, but the data used at present are limited by the religiosity data, so the effect of democracy is treated as a sideshow.³

Table 1. Dimensions of the literature: Hits in millions

Search engine	Search word	Hits
Google	Corruption and religion	72
	Corruption and religiosity	83
	Corruption and democracy	110
Google scholar	Corruption and religion	1.6
	Corruption and religiosity	0.2
	Corruption and democracy	2

The searches took place in November 2023

2. Data (R , T , y , V) for religiosity, corruption, income, and democracy

The analysis looks at religiosity not at religions.⁴ While religion is a measure of belonging religiosity measures the intensity of the belief. It is defined as the average importance of religion in all aspects of life. Hence, it is a latent variable that must be polled. The WVS (World Values Survey) is the largest cross-country poll. It has a battery of items trying to catch various aspects of religiosity. The 14 most relevant items are chosen, and the score is calculated for each poll as the fraction of respondents giving a high religiosity answer. R is the average. A factor analysis of the scores shows similar loadings. Thus, R is robust even when most polls miss items.

R , religiosity. It is available for six waves of the WVS, each collected over a 5-year period.⁵

The measure is in per cent. In practice, the range is from 3.8 in China to 93.5 in Qatar.

T , corruption. It is made from TI , the Transparency Internationals corruption perception index.

² The measure was introduced in Gundlach and Paldam (2013). It is updated in Chapter 11 of Paldam (2021a).

³ There is at least ten times more data for a study of democracy and corruption.

⁴ Paldam (2004) is study using a set of dummies or the different religions. The only religion that stands out is the Lutheran version of Christianity that gives low corruption.

⁵ The four WVS waves used were collected: w3 1995-98, w4 1999-04, w5 2005-09, and w6 2010-14.

TI is an honesty index that rises when corruption falls. To change TI into a corruption index it is reversed $T = 10 - TI$, with the range to $[0, 10]$. It rises with corruption. y , income. It is defined as the natural logarithm to gdp , which is the GDP per capita from the Maddison-project database.

V , democracy. It is the polyarchy index from the V-Dem project. It has the range $[0, 1]$, and rises when a country becomes more democratic. This variable is discussed in section 7 only.

The T , y , and V data are averaged over the periods of the WVS waves. The T data starts in 1995, so the first two waves of the WVS cannot be used. The data for (T, R, y, V) covers 240 observations in 101 countries. The average number of countries per wave is 60. Thus, only 59% of the countries are covered in each wave. The overlap of countries in any two waves is 45%. Table 2 shows the variation in the country groups included in the waves, and the skewness of the wave samples. In addition, the WVS is a skew relative to the 188 countries of the T -index.

Table 2. Countries in the WVS-waves

Group	(1) W3	(2) W4	(3) W5	(4) W6	(5) All	(6) T -index
Africa, Sub Saharan	2	5	7	5	11	47
Asia (+ Oceania)	8	10	9	11	14	35
Latin America	9	5	9	10	13	32
MENA	1	8	6	12	14	19
Post socialist	12	21	10	13	24	30
West	21	23	15	8	25	25
All	53	72	56	59	101	188
χ^2 -test (i) for waves	2.0%	29.3%	48.4%	0.8%	Na	Na
χ^2 -test (ii) for world	0	0	4.2%	1.6%	0.01%	Na

The χ^2 -test: (i) compares within waves. (1) is the number of countries in in (1) a random draw of the countries in (2) to (4). (2) is the number of countries in in (1) a random draw of the countries in (1), (3), and (4). Etc. It is rejected for (1) and (4). (ii) Compare (6) to all other columns. It is rejected for all columns.

3. Notes on the huge literature

Even when the scholarly literature is a small fraction of the discussion, it is still very large. The paper studies corruption and religiosity, where Google Scholar gives ‘only’ 190,000 hits. When looking at the lists of papers a great many are irrelevant to this paper, but there is still many papers of some relevance.

Many studies deplore the lack of data for religiosity, and it appears that all macro studies work with a poorer data-sample than the present paper. A dozen papers use one or two items from one wave of the WVS as their measure of religiosity, see e.g., Gokcekus and Ecki (2020). Most papers use the TI data as the measure of corruption. However, some papers have found a

dataset that allows a study of one aspect of the problem in one country; see e.g., Flavin and Ledet (2013).

Most papers find a positive relation between the two variables, but a minority finds a negative relation. Some papers manage to find one aspect of religions that gives a negative effect on corruption; see e.g., Zakeria (2018). Nearly all papers control for income and a couple of ad hoc controls. Such variables are often correlated with income, so this gives a wide spectrum of results. An example is the use of a democracy index as a measure of accountability.⁶

4. What should we expect about the (R, T) -relation?

The huge discussion of the relation referred to boils down to a claim about the sign on the correlation between the two variables.

Think of a religious and an irreligious discussant. The religious claims that religion is a force for the good, so increasing religiosity should lead to less corruption, i.e., he claims that the correlation is negative. The irreligious claims that there is no connection, or she may suspect that the connection goes the other way, i.e., she claims that the correlation is zero or positive. The two claims have been elaborated into many versions, hence the 83 million hits in google.

The claims can also be expressed as a disagreement about hypocrisy. The religious claims that strong religiosity induces people to behave better, while the irreligious claims that religiosity may serve as an excuse for behaving badly. Maybe she recalls Blaise Pascal,⁷ who noted: “Men never do evil so completely and cheerfully as when they do it from religious conviction.”

Religion has a strong effect on people’s priors, and priors affect research results.⁸ This is a key part of the replication crisis in the social sciences. The classical research strategy (of theory/model/regression) is so flexible that it permits a wide range of results, due to the many ad hoc control variables that come and go in estimates. To reduce the replication problem, control variables should be strongly justified.

⁶ A set of papers study the corruption within church organizations and confirm that monopoly of Churches increases corruption; see e.g., Gutmann (2015).

⁷ Blaise Pascal (1623-62) was a French mathematician and writer. He was deeply religious, but he lived most of his life during the Thirty Years’ War. Hence, he knew that the relation of moral and religion is complex.

⁸ In November 2023 Google scholar gave 600,000 hits to ‘replication crisis’. One third of the hits was to economics.

5. First view of the data: the univariate relation

The (T, R) -relation is univariate. The bold correlations in row (1) of Table 3 analyze this relation. They are positive, stable, rather large, and not due to extreme observations. However, rows (2) to (6) in the table show that all four variables are similarly correlated numerically.

Table 3. The correlation of religiosity, R , corruption, T , income, y , and democracy, V

Correlation Sample period	Wave 3 1995-98		Wave 4 1999-04		Wave 5 2005-09		Wave 6 2010-14		All 1995-14	
	r	ρ	r	ρ	r	ρ	r	ρ	r	ρ
(1) R, T	0.48	0.52	0.52	0.53	0.63	0.58	0.51	0.44	0.54	0.53
(2) R, y	-0.56	-0.56	-0.55	-0.56	-0.63	-0.68	-0.44	-0.53	-0.53	-0.57
(3) T, y	-0.82	-0.85	-0.85	-0.89	-0.75	-0.81	-0.67	-0.70	-0.76	-0.82
(4) V, T	-0.67	-0.78	-0.64	-0.72	-0.67	-0.80	-0.56	-0.68	-0.63	-0.73
(5) V, y	0.73	0.76	0.70	0.76	0.63	0.63	0.30	0.38	0.56	0.65
(6) R, V	-0.32	-0.44	-0.50	-0.59	-0.43	-0.55	-0.33	-0.38	-0.42	-0.51
(7) N	53		72		56		59		240	

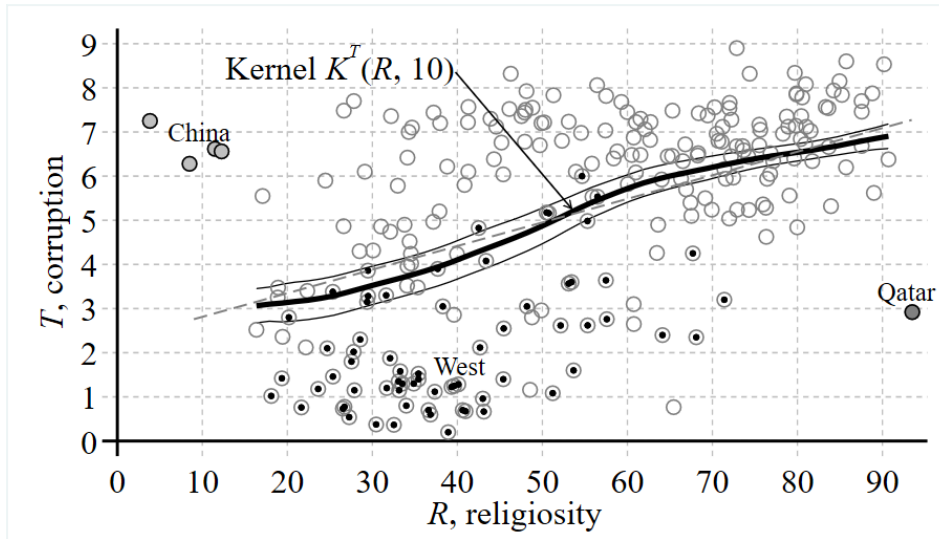
The correlations are r the standard Pearson coefficient, and ρ is Spearman's rank correlation. The latter is robust to extreme observations, but for normally distributed data r is more efficient. Rows (1) to (3) are modelled in section 6, while the gray shaded rows (4) to (6) are modeled in section 7.

Table 4 is a factor analysis demonstrating that the strong intercorrelation between the four variables is due to one and only one common factor. Other work by the author (notably Paldam 2021a) show that this factor in the Grand Transition: R , T , and V have parallel *transitions*, which are systematic movements when the country goes from poor to rich. A full transition normally takes more than a century. Transitions are roughly equivalent in wide cross-country samples and long time series. They are functions of income; see Gundlach and Paldam (2009) and Paldam (2021a) for causality tests for the (T, y) , (R, y) , and (V, y) relations. In these relations the main causal direction is from y to the other variable.

Table 4. A factor analysis of the four variables

	Factor1	Factor2
Eigenvalue	2.265	-0.037
Factor loadings		
On R	-0.611	Na
On T	-0.864	Na
On y	0.823	Na
On V	0.684	Na

Figure 1. The scatter of all 240 (T , R) datapoints



The kernel regression is a smoothed moving average with a fixed bandwidth. The four NW-points in light gray are for China. The SE-point in dark gray is for Qatar. They are not included in the kernel regression. It is surrounded by (thin) 95% confidence intervals. The points for the West have a black dot in the observation. The dashed straight line is fully within the confidence intervals. Hence, it cannot be rejected that the kernel is linear.

Figure 1 shows what a correlation of 0.54 means. It is a highly significant relation, and it is almost linear. This means that the transitions in the T and the R -data are parallel.

6. Controlling the (R , T)-relation for transitions: the income variable

The estimates of Model (1) in Table 5 correspond to row (1) in Table 3, and the estimates of α_1 are also highly significant and stable. Correlations are independent of the scales of the variables, but regressions are not. The coefficients to R in Model (1) of the table should be about ten times smaller than the correlations. It is reassuring that they are.

Model (2) shows what happens when the relation is controlled for y . Both T and R , have a strong transition as demonstrated in Chapters 10 and 11 in Paldam (2021a). While the religious and the democratic transitions are fairly parallel the transition of corruption is late. The two transitions give a spurious component in the (R , T)-relation. Hence, the control variable of y is justified, and all coefficients on y in relation (2) are indeed highly significant.

The inclusion of y has three effects: (i) The coefficient on R decreases three times. The effect of R is even insignificant in the data for the two first waves. However, the effect is always positive, and significant in the 'All' row. (ii) the CV for the coefficient on R is 14 for model (1) while it is 62 for model (2), so the key coefficient is almost five times more variable in (2). (iii) The R^2 -score increases 2-3 times, so relation (2) explains much more of the variation.

Table 5. Regressions to sort out the size of the effect of religiosity on corruption

Sample Wave <i>N</i>	Model (1) $T = \alpha_1 R + \mu$			Model (2) $T = \alpha_2 R + \beta_1 y + \mu$				α_2/α_1 Ratio
	α_1	μ	aR ²	α_2	β_1	μ	aR ²	
W3 53	0.060 (3.9)	1.73 (2.2)	0.22	0.004 (0.3)	-2.28 (-8.1)	26.06 (8.5)	0.65	0.06
W4 72	0.067 (5.0)	1.49 (1.9)	0.26	0.009 (0.9)	-1.92 (-10.7)	22.60 (11.1)	0.72	0.13
W5 56	0.069 (5.9)	1.54 (2.4)	0.38	0.028 (2.2)	-1.27 (-5.3)	15.61 (5.8)	0.59	0.40
W6 59	0.049 (4.5)	2.61 (3.9)	0.25	0.026 (2.6)	-1.25 (-5.2)	15.93 (6.1)	0.49	0.53
All 240	0.061 (9.6)	1.83 (5.1)	0.28	0.021 (3.8)	-1.58(-13.9)	18.93 (15.0)	0.60	0.34
<i>CV</i>	12.8	24.5	21.0	62.2	-26.1	22.2	13.8	4.9

CV is the coefficient of variation, which is 100 times the standard deviation over the average. It is calculated for the four waves. The aR² is the R² adjusted for the degrees of freedom.

As both *R* and *T* have transitions, the three variables *T*, *R*, and *y* have a great deal of confluence giving multicollinearity. Thus, some randomness enters the way *y* and *R* share the common part of the coefficient, giving the rise in the coefficient of variation for the effect of *R*.

The estimate (2) for all is $T = 0.021R - 1.58y + 18.93$. Over a 20 years period *R* falls by 3 pp, and *y* increases by 0.57 pp in the 19 countries that are included in all waves. This gives an effect on *T* of -0.07 pp and -0.90 respectively. Corruption does fall when religiosity falls, but not very much for realistic orders of magnitudes.

7. Another confluent variable: *V*, the polyarchy democracy index

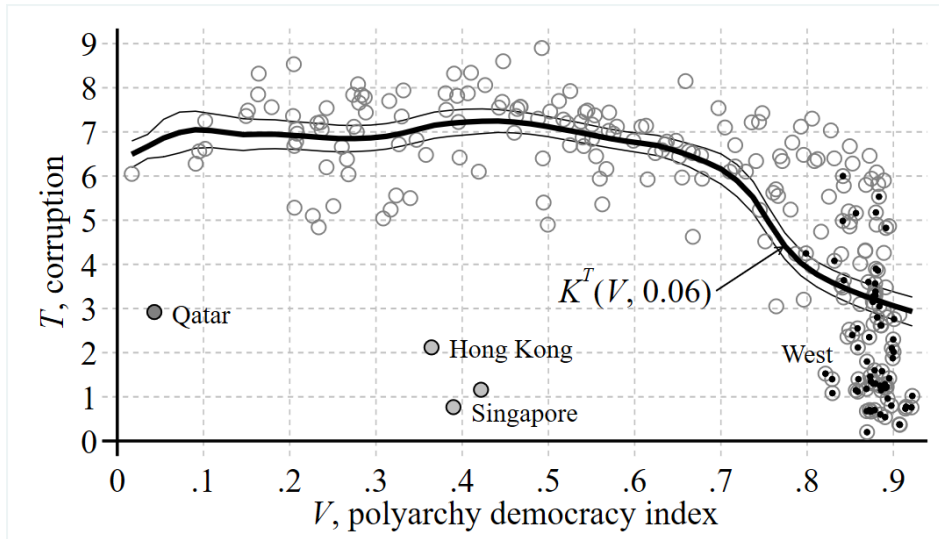
Row (4) in Table 3 showed that the *V*-index has a strong and stable negative correlation to corruption. However, the *V*-index also has a strong transition, so most of the correlation is spurious. Table 6 adds *V* to the relation (2). This greatly reduces the effect of *V*-index, but it remains negative. The effect of *R* is also reduced but it is still positive.

Table 6. Adding the political regime, *V*, polyarchy

Sample Wave <i>N</i>	Model (3) $T = \alpha_3 R + \beta_2 y + \gamma V + \mu$					(a)	
	α_3	β_2	γ	<i>M</i>	aR ²	ΔR^2	
W3 53	0.007 (0.5)	-1.87 (-4.8)	-2.07 (-1.5)	23.60 (6.8)	0.66	0.01	
W4 72	0.008 (0.8)	-1.83 (-8.2)	-0.59(-0.7)	22.21 (10.5)	0.71	-0.00	
W5 56	0.026 (2.2)	-0.87 (-3.4)	-2.89 (-3.0)	13.89 (5.4)	0.64	0.05	
W6 59	0.017 (1.9)	-1.09 (-5.1)	-2.89 (-4.0)	16.55 (7.1)	0.59	0.11	
All 240	0.016 (3.0)	-1.28 (-10.8)	-2.58 (-5.9)	17.99 (15.1)	0.65	0.05	
<i>CV</i>	53.1	-31.2	-44.5	20.9	6.6		

(a): The ΔR^2 column show the increase in the adjusted R² from model (2) to Model (3).

Figure 2. The scatter of all 240 (T , V) datapoints



See Figure 1. The SW-point in dark gray is for Qatar. The three light gray points are for Hong Kong and Singapore. They are not included in the kernel regression. The points for the West have a black dot in the observation.

The strong non-linearity is explained by the lateness of the transition of corruption relative to the democratic transition as mentioned.

If a few more variables with transitions, such as human capital and urbanization are added, the coefficients on R will be (even) more variable, with a bit of luck and enough experiments the range of estimates of the coefficient on R will extend into the negative. Then both the religious and the irreligious may choose a result they like. This is why there is a replication crisis in the social sciences.

8. Confluence in the explanation of corruption

The analyses have shown a great deal of confluence due to the common transition in the variables. A fine tool to see the confluence is to start from model (3) and calculate the marginal reduction in R^2 when each of the variables are omitted. Table 7 shows that the three variables explain a good deal of the variation in corruption, but neither adds much to the explanation. In addition, the table shows that the marginal effects are unstable.

The explanation for the small and unstable marginal effects of the three explanatory variables is that they contain the common element of the transition. The strongest of the three is income, y , in the sense that income is the best proxy for development and hence the best variable explaining the transition. Both R and V add a little, but not much.

Table 7. Marginal effects of variables explaining corruption

Contribution	The effect of all and each variable	Wave N	W3 53	W4 72	W5 56	W6 59	All 240	CV Waves
Total effect	(3) from Table 6	aR^2	0.661	0.714	0.644	0.593	0.653	6.6
R marginal	Deleting R from (3)	$-\Delta aR^2$	-0.005	-0.002	0.027	0.019	0.012	139
y marginal	Deleting y from (3)	$-\Delta aR^2$	0.149	0.274	0.071	0.179	0.168	43
V marginal	Deleting V from (3)	$-\Delta aR^2$	0.008	-0.002	0.054	0.106	0.050	103
Sum of the three marginal effects			0.152	0.270	0.152	0.304	0.230	20

It is a problem that the analysis covers 240 observations only. The robustness is analyzed by looking across waves. It is clear from the table that neither the effect of religiosity nor of democracy is stable across waves. This is due to the non-random variation of the four samples; see Table 2.

9. Conclusion

Table 2 showed that the samples of the WVS waves have big problems as regards their representativity both compared to each other and to the wider world. However, a few results do stand out:

There is a strong and stable positive correlation between religiosity and corruption. The more religious the more corrupt, but most of this effect is spurious. The story of the effect of democracy is parallel, though negative. The more democracy the less corrupt, but, once again, most of the large and stable effect is spurious.

The strongest explanation of both corruption and religiosity is income – both variables have strong transitions, and so has democracy. Once the corruption/religiosity-relation is controlled for income, the effect of religiosity becomes quite modest, but it remains positive. And once the corruption/democracy-relation is controlled for income, the effect of democracy also becomes modest, but it is still negative.

The note illustrates a main problem for cross-country regressions. Most macro variables have transitions that give confluence in wide cross-country samples. This makes datamining a tempting possibility. In fields with strong priors as the one considered – this allows researchers to find results they like. Thus, control variables should be strongly justified.

References:

- Flavin, P., and R. Ledet. 2013. Religion and government corruption in the American states. *Public Integrity* 15, 329–43
- Gokcekus, O., Ecki, T., 2020. Religion, religiosity, and corruption. *Review of Religious Research* 63, 563-81
- Gundlach, E., Paldam, M., 2009. The transition of corruption: From poverty to honesty. *Economic Letters* 103, 146-48
- Gundlach, E., Paldam, M., 2013. The religious transition. A long run perspective. *Public Choice* 156, 105-23
- Gutmann, J. 2015. Believe, but verify? The effect of market structure on corruption in religious organizations. *Kyklos* 68, 153–64
- Paldam, M., 2001. Corruption and religion. Adding to the economic model. *Kyklos* 54, 383-414
- Paldam, M., 2021a. *The Grand Pattern of Development and the Transition of Institutions*. Cambridge UP, New York and Cambridge UK (with nine pages of references)
- Zakeria, P., 2018. Religiosity and corruption, pp 69-90 in Kubbe, I., Engelbert, A., eds. *Corruption and Norms*. Palgrave Macmillan