News on Corruption in the
Wall Street Journal and
the Corruption Perception Index (CPI)

by
Runtian Jing
University of Electronic Science and Technology of China, China
Gaoliang Zhang
Zhe Jiang University, China
Tianli Feng
*University of Electronic Science and Technology of China, China*

Based on cross-time-section series data collected from Wall Street Journal (WSJ), this article suggests that the Corruption Perceptions Index (CPI) for 60 countries/regions published by Transparency International is highly correlated with WSJ news about "corruption", especially for eight strong countries. Compared to a slight converging trend of CPI standard deviation, the CPI score presents no significant response to time. Our partial multiple correlation model suggests that WSJ news may be a good indicator of corruption levels in a country, although not as a potential cause influencing CPI.

Introduction

Corruption is an important issue in business, economic, and social science research. Due to differences in research focus and perspective, previous literature has provided varying definitions for corruption. Some studies examine the corrupt behavior of politicians (LaPalombara, 1995; Oldenbury, 1987), while others describe corruption between private parties, such as in commercial bribery (Coase, 1979). Shleifer and Vishny define corruption as "the use of governmental power to create rents via entry controls, regulatory cartel enforcement, or raising rivals' costs" (Shleifer & Vishny, 1993: 599). In Macrae's definition, corruption is "the arrangement that involves an exchange between two parties which (1) has an influence on the allocation of resources either immediately or in the future; and (2) involves the use or abuse of public or collective responsibility for private ends" (Macrae, 1982: 678).

Since Transparency International (TI) began to publish its annual report of *Corruption Perceptions Index* (CPI) in 1995, this index has become a leading indicator of corruption in social science research, and has been widely used for measuring the level of corruption across countries. The definition given by TI focuses on the public sector, and is defined as "the misuse of public power for private benefit such as bribing of public officials, kickbacks on public procurement, or embezzlement of public funds" (Transparency International, 2003). The CPI score reflects the impressions and perceptions of corruption in dozens of countries based on surveys of business people, academics, risk analysts, and the general public, and ranges from "0 = highly corrupt" to "10 = highly clean". As we are interested in the explanation and justification of perceived levels of corruption, TI's definition will be used for the current study.

The other important publication used in our study is the *Wall Street Journal* (WSJ), the flagship publication of the Dow Jones. In addition to the U.S. edition, which is edited in New York, the company publishes the *Asian Wall Street Journal,* edited in Hong Kong, and the Wall Street Journal Europe, edited in Brussels. WSJ and its global editions are some of the most respected sources of worldwide business and financial news. We noticed that the annual number of WSJ reports concerning corruption for a particular country appeared to correlate with the CPI score of that country. For example, despite its economy status, Italy suffers from high levels of corruption. Furthermore, its perceived level of corruption has varied greatly over the past nine years, with CPI values ranging from 2.99 (1995) to 5.5 (2001). After plotting annual amounts of corruption news with the annual CPI score, we found that the changing trend of WSJ news reflects variance of the CPI score[1], as Figure 1 shows. This finding led us to examine the relationship between news coverage and corruption.
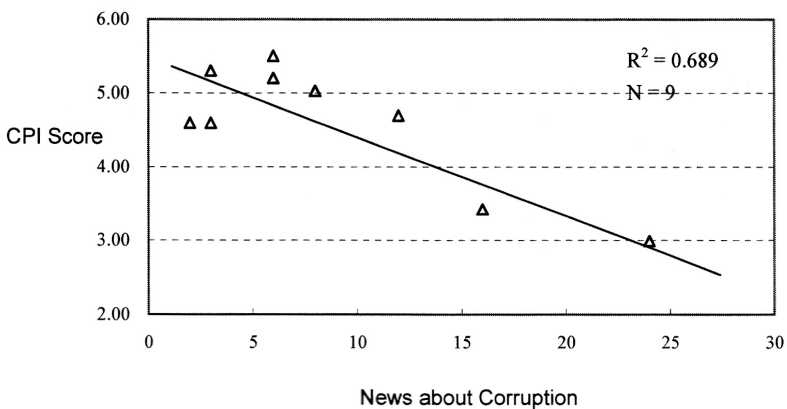
Figure 1. Relationship between CPI score and Corruption news in Italy between 1995-2003.

We have three empirical predications. It must be stated from the outset that we are not concerned with causality issues here, but simply an association between WSJ news and CPI scores.

Hypothesis 1: The CPI of a country is negatively associated with WSJ news items concerning "corruption".

We adopted absolute CPI scores to reflect a country's level of corruption. This is because a country's ranking order can vary annually. Thus, TI recommends that absolute scores be used in year to year comparisons rather than relative ranks. Two factors can account for year to year changes in CPI score: the changing perception of a country's actual performance, and changes in sampling and methodology. Therefore, we predict a time effect, where CPI scores will show shifting trends from year to year.

Hypothesis 2: The CPI score is sensitive to the year that has been surveyed.

CPI is a composite index. For example, the 2003 CPI report consisted of survey years between 2001 to 2003 (information for each country was drawn on 17 different polls/surveys from 13 independent institutions, including the World Economic Forum (WEF), Institute of Management Development (IMD), World Bank (WB), and Gallup International (GI)).

Hypothesis 3: Countries display different sensitivity with respect to the relationship between CPI score and WSJ news.

In addition to the time difference, the sampling countries in these sources also vary greatly. For example, the *Risk Ratings Survey* conducted by World Markets Research Center (WMRC) covers 186 countries, while the *Asian Intelligence Issue* conducted by Political & Economic Risk Consulting (PERC) covers only 14 countries. During the combination process, the minimal required number of surveys is three before one country can be included in the final CPI report. Thus, in the 2003 report, the number of surveys used to generate country scores varied from 3 to 17. Such a large variance affects the reliability of the CPI score for each country. In general, a strong country will enjoy a large number of data sources, and thus convey high reliability as to the perceived score. This fact may affect the significance of correlations between WSJ news and CPI score.

Methods

We developed a dataset from WSJ (including its global editions) describing news items concerning corruption for 60 countries/regions from 1995 to 2003. The data is from searched editions using the exact matching method (i.e. the "citation or text" of news contained the word "corruption" for a given country in a given year). While we also searched the database using associated

words such as "bribery", the results showed that news items with these words are far less numerous than for "corruption". Additionally, search results for the former overlapped in records. Thus, we have limited the dataset in our study to items found by the word "corruption".

We denote news items about corruption as n_{it} for country i ($i = 1...60$) in year t ($t = 1...9$, refers to year 1995-2003 respectively). The CPI score for that year is c_{it}.

We searched and calculated the total items of news items covering all topics (besides corruption) for each country over the past nine years, which is as r_i. Therefore, the frequency of news (f_i) for country i can be calculated as:

$$f_i = \frac{r_i}{\sum\limits_{i=1}^{n} r_i} \times 100\% \ ,$$

Here, $n = 60$, means the 60 countries/regions in our study.

Countries differ greatly in the frequency of news in WSJ[2]. From Table 1 the U.S. has the highest coverage, and was involved in 35.04% of all the items. Eighty percent of all the countries are below the average level of coverage (1.667%).

In most of cases, corruption news in a strong country will have a high probability of being reported, even if its degree of seriousness is similar or less than that of other countries. However, once reported by WSJ, one additional piece of corruption news in a less strong country will cause much deeper impression toward the minds of its readers. Hence, we take the following formula to estimate the real corruption events by the reported news in WSJ.

$$\tilde{n}_{it} = 1n \left(1 + \frac{n_{it}}{f_i}\right)$$

$\tilde{n}_{it}$ reflects the amount of news events that are newsworthy to the same degree across all countries. The estimated amount provides a reasonable base to predict the perceptual index for each country, and to compare the level of corruption across all countries.

Since changes in corruption for a country usually evolve slowly, while public perception may change more quickly and is influenced by short-term events, TI has based the CPI on a three-year rolling average. Thus, the 2003 CPI is based on the survey data provided between 2001 and 2003. However, once a serious corruption event happens, WSJ can report it very quickly. Thus, we assume a time lag for correlations between $\tilde{n}_{it}$ and c_{it}. To catch such a time lag we have conducted further model competition analysis using seven different time lagging or rolling methods.

Using the 2003 CPI data as an example, Figure 2 shows that the correlation between WSJ news and CPI score changes very gradually under different time lag situations. Comparatively speaking, the three-year rolling data of news has the strongest correlation coefficient to the 2003 CPI value. In reality, 2003 CPI report was released on October 7 2003, and its 17 survey sources were carried out between 2001 to 2003. Therefore, this fact has explained the result of our model competition.
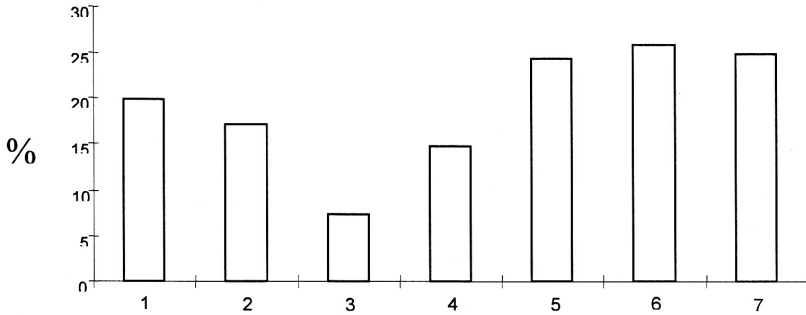
Method 1: Using the news data of 2003 (Without time lag)

Method 2: Using the news data of 2002 (One year's time lag)

Method 3: Using the news data of 2001 (Two years' time lag)

Method 4: Using the news data of 2000 (Three years' time lag)

Method 5: Using the averaged data of 2003 and 2002 (Two years' rolling)

Method 6: Using the averaged data of 2003, 2002 and 2001 (Three years' rolling)

Method 7: Using the averaged data of 2003, 2002, 2001 and 2000 (Four years' rolling)

Figure 2. Time lag between CPI (2003) and WSJ News.

As panel one in Table 1 shows, we calculated the three-year rolling data of news for 60 countries/regions according to the following formula:

$$\bar{\bar{n}}_{it} = \sum_{k=t-2}^{t} n_{ik}$$

Here, k means the starting year of the three-year rolling method, and ranged from 1993 to 2001.

One concern of this study was whether the CPI score changed significantly over time. Based on a period of nine years (1995~2003), we define eight dummy variables (denoted as Y96, Y97, Y98, Y99, Y00, Y01, Y02 and Y03) to capture changes that took place over the period of the study.

Table 1. News Reported about Corruption
Panel One. Raw data

Country/Region	News Fraction	Reports concerning corruption										TI index (CPI)	
		2003	2002	2001	2000	1999	1998	1997	1996	1995	9-yr Ave.	2003	9-yr Ave.
Group 1:													
1 United States	35.046	1.063	1.107	1.165	1.165	1.251	1.138	1.017	0.887	0.951	1.083	7.5	7.629
2 Japan	6.745	0.909	0.909	1.112	1.206	1.264	1.325	1.083	1.298	1.657	1.196	7.0	6.638
3 China	5.223	2.031	2.110	2.128	2.158	2.087	2.095	1.995	2.091	2.260	2.106	3.4	3.097
4 United Kingdom	5.130	0.213	0.119	0.318	0.471	0.412	0.213	0.059	0.286	0.227	0.257	8.7	8.548
5 Germany	4.492	1.016	1.034	1.151	1.088	0.983	0.828	1.101	1.168	1.445	1.090	7.7	7.838
6 Canada	3.285	0.700	0.877	0.690	0.513	0.463	0.463	0.533	0.583	0.583	0.601	8.7	9.014
7 France	3.246	1.499	1.316	1.129	1.102	1.175	1.363	1.576	1.958	2.176	1.477	6.9	6.724
8 Russia	3.133	1.730	1.992	2.440	2.363	2.102	1.653	1.596	1.596	1.738	1.912	2.7	2.431
Group 2:													
9 Mexico	2.958	1.870	2.022	2.189	2.292	2.624	2.943	2.791	2.371	1.971	2.341	3.6	3.338
10 Hong Kong (China)	2.341	1.308	1.451	1.451	1.303	1.279	1.523	1.523	1.523	1.409	1.419	8.0	7.634
11 Korea, South	2.083	1.395	1.541	1.541	1.386	1.362	1.611	1.611	1.611	1.496	1.506	4.3	4.300
12 Israel	1.680	1.776	1.743	1.550	1.496	1.795	1.694	1.629	0.903	0.957	1.505	7.0	7.260
13 Brazil	1.509	2.022	2.268	2.177	1.723	1.477	1.510	1.248	1.772	1.557	1.750	3.9	3.680
14 India	1.423	1.730	1.903	1.967	1.993	2.214	2.562	2.407	1.812	1.400	1.999	2.8	2.773
15 Italy	1.307	1.545	1.545	1.745	1.480	1.737	1.825	2.503	3.226	3.955	2.173	5.3	4.593
16 Indonesia	1.283	2.661	2.953	3.120	3.058	2.435	1.916	0.874	0.874	0.472	2.040	1.9	2.046
17 Taiwan (China)	1.152	3.104	3.234	3.263	3.263	3.120	3.194	3.129	3.264	3.440	3.224	5.7	5.409

18	Singapore	1.100	1.165	0.870	0.870	0.942	0.942	1.298	0.787	0.787	0.561	0.913	9.4	9.102
19	Argentina	0.995	1.966	2.027	1.565	0.831	0.882	1.188	1.787	1.137	0.599	1.331	2.5	3.307
20	Australia	0.978	0.705	1.125	1.358	1.665	1.245	1.012	0.705	0.841	1.074	1.081	8.8	8.651
21	Spain	0.954	1.708	1.992	2.150	1.978	1.558	1.227	1.513	1.933	2.370	1.826	6.9	6.140

Group 3:

22	Thailand	0.831	0.850	0.850	1.437	1.258	1.743	1.666	1.257	0.773	0.672	1.167	3.3	3.142
23	Turkey	0.822	0.590	1.180	1.591	1.514	1.576	1.430	0.918	0.531	0.677	1.112	3.1	3.506
24	Pakistan	0.801	1.253	1.401	1.889	1.836	1.938	2.194	1.534	1.015	0.519	1.509	2.5	2.260
25	Jordan	0.760	0.963	1.242	1.242	0.989	0.989	1.410	1.410	1.131	0.280	1.073	4.6	4.656
26	Switzerland	0.745	1.968	2.298	2.848	3.004	3.029	2.718	2.263	1.292	1.406	2.314	8.8	8.692
27	Netherlands	0.674	0.606	1.252	0.948	0.948	0.303	0.303	0.303	0.303	0.948	0.657	8.9	8.892
28	Malaysia	0.648	1.848	1.702	1.822	2.368	2.449	1.984	1.279	1.199	0.887	1.727	5.2	5.101
29	Vietnam	0.608	1.820	1.673	1.755	1.755	2.003	2.281	2.120	2.018	2.076	1.945	2.4	2.541
30	Philippines	0.568	3.055	3.433	3.285	2.467	1.986	2.134	2.134	1.876	1.728	2.455	2.5	2.912
31	South Africa	0.532	1.586	1.586	1.419	1.225	0.872	0.872	0.983	1.335	0.983	1.207	4.4	5.050
32	Venezuela	0.525	2.493	2.572	2.572	2.572	2.129	2.026	1.943	2.349	2.693	2.372	2.4	2.581
33	Ireland	0.515	0.529	0.889	1.248	1.079	1.079	1.079	0.720	0.360	0.000	0.776	7.5	7.811
34	Chile	0.476	2.234	2.027	2.027	1.762	1.649	1.649	1.477	0.927	0.755	1.612	7.4	7.143
35	Sweden	0.473	0.664	1.043	1.043	1.043	1.216	1.216	0.930	0.379	0.757	0.921	9.3	9.244
36	Belgium	0.431	1.092	0.692	1.092	1.092	1.993	2.542	3.121	2.996	2.624	1.916	7.6	6.338
37	Poland	0.416	1.697	1.812	1.966	1.966	1.850	1.581	1.759	1.759	1.581	1.775	3.6	4.406
38	Egypt	0.359	1.401	2.147	1.816	1.373	0.627	0.000	0.000	0.627	1.373	1.040	3.3	3.206
39	New Zealand	0.351	1.473	1.473	1.473	1.083	1.083	0.449	0.000	0.000	0.000	0.782	9.5	9.423
40	Luxembourg	0.308	0.482	0.482	0.482	0.000	0.482	0.482	0.482	0.000	0.482	0.375	8.7	8.495
41	Czech Republic	0.289	0.690	0.690	0.690	0.690	1.381	0.690	0.690	0.690	0.690	0.767	3.9	4.471

42	Austria	0.259	1.054	1.581	1.581	1.054	0.527	0.000	0.000	0.000	0.000	0.644	8.0	7.637
Group 4:														
43	Greece	0.240	1.824	2.569	2.159	1.839	1.095	1.095	0.547	1.292	1.292	1.523	4.3	4.644
44	Colombia	0.233	3.078	3.256	2.988	3.248	3.198	3.692	3.837	3.708	3.303	3.368	3.7	3.089
45	Norway	0.229	0.759	0.000	0.000	0.000	0.560	1.121	1.121	0.560	0.000	0.458	8.8	8.811
46	Finland	0.218	0.773	0.773	0.773	1.669	1.669	0.897	0.000	0.000	0.000	0.728	9.7	9.594
47	Hungary	0.217	1.351	0.775	0.899	0.899	0.899	0.000	0.000	0.775	0.775	0.708	4.8	4.951
48	Portugal	0.213	0.580	0.580	0.580	0.580	0.580	0.580	0.580	0.000	0.580	0.516	6.6	6.429
49	Denmark	0.186	0.947	0.947	0.618	1.439	2.477	1.859	1.038	0.000	0.000	1.036	9.5	9.654
50	Nigeria	0.159	3.018	3.061	3.192	3.618	3.524	3.394	3.302	2.083	1.865	3.006	1.4	1.394
51	Ecuador	0.144	1.928	1.928	1.928	2.956	3.357	3.230	2.893	1.591	1.591	2.378	2.2	2.456
52	Romania	0.127	1.877	2.005	1.794	2.521	1.454	1.665	0.938	1.665	1.665	1.732	2.8	2.977
53	Kenya	0.123	2.718	2.930	2.637	3.252	3.833	3.621	2.268	0.737	0.737	2.526	1.9	2.087
54	Bolivia	0.082	3.014	1.937	0.859	0.000	1.077	1.077	1.937	0.859	2.068	1.425	2.3	2.494
55	Uruguay	0.081	2.296	2.296	1.213	0.000	0.864	0.864	0.864	0.000	0.000	0.933	5.5	4.757
56	Costa Rica	0.071	0.000	0.000	0.000	0.903	0.903	1.806	0.903	0.903	0.903	0.702	4.3	5.121
57	Bangladesh	0.070	3.166	2.037	0.909	0.000	1.129	2.483	3.910	3.690	3.245	2.285	1.3	1.298
58	Uganda	0.060	1.179	1.179	0.000	0.000	1.311	1.311	1.311	0.000	0.958	0.806	2.2	2.287
59	El Salvador	0.059	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.318	0.146	3.7	3.717
60	Iceland	0.036	1.117	1.117	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.248	9.6	9.300
Total		100.000	90.093	93.556	89.853	87.480	91.245	89.989	82.239	71.343	73.728	85.503	324.200	322.721
Average		1.667	1.502	1.559	1.498	1.458	1.521	1.500	1.371	1.189	1.229	1.425	5.403	5.379

Source: Wall Street Journal (1993–2003)

Because of its close geographical and economic relation with U.S., Mexico also has a high faction.

Panel Two. Results of Multivariable Regression

Independent Variables:	*Dependent Variables*			
	Model 1: CPI Score		Model 2: CPI SD*	
	Estimate	T-value	Estimate	T-value
News	-1.409***	-12.290	0.067	1.590
Y96	0.131	0.290	0.678***	4.020
Y97	-0.340	-0.790	0.861***	5.450
Y98	0.056	0.130	0.238	1.500
Y99	0.045	0.110	0.177	1.150
Y00	0.094	0.230	0.147	0.960
Y01	0.121	0.290	0.138	0.890
Y02	-0.003	-0.010	-0.007	-0.050
Y03	0.051	0.120	-0.057	-0.370
Constant:	7.519***	22.140	0.648***	5.170
Model Parameter:				
R^2	0.238		0.109	
F-value	17.050		6.640	
Number of Observations:	501		501	

Notes:

* Here, CPI SD refers to the difference in the value of the source. The greater CPI SD, the greater the difference of perceptions of a country among the sources.

The score and standard deviation of the CPI is specified as:

$$CPI(S,SD) = \beta_0 + \beta_1\,(NEWS) + \beta_2\,(Y96) + \beta_3\,(Y97) + \beta_4\,(Y98) + \beta_5\,(Y99)$$
$$+ \beta_6\,(Y00) + \beta_7\,(Y01) + \beta_8\,(Y02) + \beta_9\,(Y03) + \epsilon_i$$

Here, CPI (S, SD) refers to two dependent variables (Score and Standard Deviation of CPI), NEWS is the cross-section-time variable based on the three-year rolling data ($\tilde{n}_{it}$), Y96~Y03 refer to the time dummy variable indicating the change for year 1996~2003 respectively from the baseline (1995), β_1 is the regression coefficient to be estimated, and ε_i refers to the error terms.

While Table 1 only shows the 2003 CPI score, data from other years can be obtained from the TI website (www.transparency.org).

Panel Three. Results of Regression for Four Groups

	Group 1	*Group 2*	*Group 3*	*Group 4*
Independent Variables:	3.156***	-1.367***	-1.199***	-1.271***
News	(0.230)	-(0.264)	-(0.227)	-(0.180)
Constant	10.359***	7.681***	7.304***	6.716***
	(0.312)	(0.510)	(0.353)	(0.328)
Number of Observations	71	115	179	136
Number of Countries	8	13	21	18
R2	0.732	0.192	0.136	0.271
Adjusted R2	0.728	0.185	0.131	0.265

* = Significant at 10 percent; ** = Significant at 5 percent; *** = Significant at 1 percent.

The dependent variable is the CPI score, and the independent variable is the WSJ news about corruption. Entries are regression coefficients without standardization, and standard errors are in parentheses.

Results

Hypothesis 1 is supported. Panel two in Table 1 shows that there is a strong relationship between the CPI score and the news reported by WSJ, with an $R^2 \sim 23.4\%$ overall. A negative coefficient (-1.401) refers to countries with heavy corruption (indicated by smaller CPI values) that are more frequently reported by WSJ.

Hypothesis 2 is rejected. There is no time-oriented change in the perceived corruption score, and insufficient evidence to suggest that methodological changes would mislead yearly comparisons. However, Table 1 shows a slightly converging trend of annual published standard deviation of the CPI score from each respondent. Year-to-year learning behavior and the cognitive schema (Fiske & Taylor, 1991) may account for this trend over the past nine years.

To test hypothesis 3, we clustered the 60 countries/regions into four groups using the Euclidian distance based on the variable Frequency of news. Due to its large distance from other countries, the U.S. was classified as one single group. The other 59 countries were classified into four other groups using the k-mean method. Many reasons beside economic influence may account for the extremely high value of Frequency of news[2] for the U.S. Therefore, we subsequently combined the U.S. into the nearest group (Group

1) to make correlation analysis more reasonable in its sample size. The four-group partition reflects well the influential position of each country. Group 1 contains eight strong countries (U.S., Japan, China, United Kingdom, Germany, Canada, France, and Russia), all of which have a powerful impact on global business. Group 2 has less influence and power than Group 1, and contains the 13 countries. Countries in Group 3 (21 countries) and Group 4 (18 countries) make up the remainder of those in our study.

As panel three in Table 1 shows, different groups report different sensitivity for CPI value with respect to WSJ news. The correlation coefficient for Group 1 is extremely high (R^2=0.732). Although also significant, Group 3 is much lower (R^2=0.136). Therefore, hypothesis 3 is supported.

Discussion

Ballantine Illustration

Previous literature shows that cultural value is one important latent variable in determining differences in corruption for countries. For example, Husted's comparative study (1999) of 44 countries has shown that corruption is significantly correlated to three cultural dimensions. Getz and Volkema confirmed the positive relation between power distance and corruption level (Getz & Volkema, 2001), and Houston and Graham (2001) suggest that relationship-oriented countries (higher power distance and lower individualism) tend to be more corrupted. Hence, we introduce cultural as another independent variable influencing corruption. The improved multivariable regression model of CPI score is as follows (due to their insignificant effect, the time dummy variables are omitted):

$$CPI = \beta_0 + \beta_1 \ (NEWS) + \beta_2 \ (CULTURE) + \epsilon_i$$

Our data agrees with the previously mentioned studies. As Table 2 shows, exceedingly high correlation exists between three cultural dimensions (power distance, uncertainty avoidance, and individualism) and the nine-year averaged CPI score (Model 1, simple regression model with CULTURE as the single independent variable). Two other regression models were also conducted: regression of the nine-year averaged news data on the nine-year averaged CPI score (Model 2, simple regression model with NEWS as the single independent variable), and the regression of cultural values and WSJ news on the nine-year averaged CPI score (Model 3, multiple regression model with NEWS and CULTURE as the independent variables). Both these models have shown significant correlations.

We adopted partial correlation analysis to examine exact causality among these three variables: CPI (denoted as Y), NEWS (denoted as X1) and CUL-

Table 2. Correlation between Culture and Corruption

	Model 1	Model 2	Model 3
Independent Variables:			
Culture - Power Distance	-0.057*** (0.013)		-0.047*** (0.013)
Culture - Uncertainty Avoidance	-0.022** (0.009)		-0.023** (0.009)
Culture - Individualism	0.032*** (0.011)		0.029** (0.011)
News	-1.964*** (0.371)		-0.655* (0.343)
Constant	8.793*** (1.290)	8.177*** (0.596)	9.368*** (1.293)
N			
R2	0.653	0.326	0.677
Adjusted R2	0.632	0.315	0.650

* = Significant at 10 percent; ** = Significant at 5 percent; *** = Significant at 1 percent.

The dependent variable is the CPI score, and the independent variable is the WSJ news about corruption. Entries are regression coefficients without standardization, and standard errors are in parentheses.
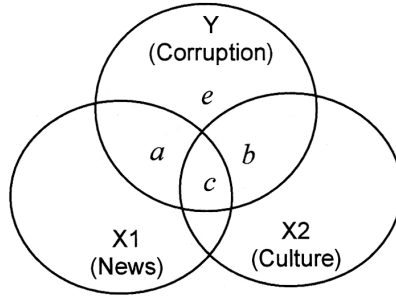
Figure 3. Ballantine illustration

TURE (denoted as X2). The Ballantine method suggested by Cohen (1983) provides a good visual illustration of this causality (Figure 3).

According to above, we have:

Correlation coefficient in Model 1: $r^2{}_{Y2} = b + c = 0.632$

Correlation coefficient in Model 2: $r^2{}_{Y1} = a + c = 0.315$

Correlation coefficient in Model 3: $R^2{}_{Y \cdot 12} = a + b + c = 0.650$

Thus, the four parameters reflecting different parts of variance in Figure 3 can be calculated as:

$a = 0.018$

$b = 0.335$

$c = 0.297$

$e = 1 - (a+b+c) = 0.297$

When X2 (CULTURE) is partial from both Y (CPI) and X1 (NEWS), the correlation coefficient is:

$$pr_1{}^2 = r^2{}_{Y1 \cdot 2} = \frac{R^2{}_{Y \cdot 12} - r^2{}_{Y2}}{1 - r^2{}_{Y2}} = \frac{a}{a + e} = 0.049$$

When X1 (NEWS) is partial from both Y (CPI) and X2 (CULTURE), the correlation coefficient is:

$$pr_2{}^2 = r^2{}_{Y2 \cdot 1} = \frac{R^2{}_{Y \cdot 12} - r^2{}_{Y1}}{1 - r^2{}_{Y1}} = \frac{b}{b + e} = 0.489$$

According to the above result, cultural impact upon corruption still remains after removing the effect of WSJ news. Conversely, WSJ news can hardly contribute any further reduction of variance after removing the proportion associated with cultural value. This suggests that WSJ news cannot be regarded as one exogenous variable, and its effect has been mostly explained

by cultural values. Thus, the assumed causality between WSJ news and CPI score has disappeared in this situation, which implies that we cannot say that WSJ news will affect the perception of corruption.

Pavarala has summarized the causes of corruption found in the literature as "administrative/bureaucratic, political, economic, and cultural"(Pavarala, 1996: 81). According to statistical principles, a more complex model with more independent variables will not reject our above finding, a fact that we found after conducting further analysis.

Validity of Finding

There are other indicators of corruption beside the TI CPI. Kaufmann, Kraay, and Zoido-Lobaton (2001) have developed a *Graft Corruption Index* that includes scores for 155 countries. Despite some substantial differences in data collection method, their dataset is almost identical to that of CPI, correlated at 0.98 (Knack and Azfar, 2001). In the annual *Global Competitiveness Report* (GCR), two other indicators of corruption (Irregular payments and Burden of corruption) are also included, and both are scaled from 1 to 7, with smaller scores implying higher levels of corruption (Schwab, Porter, Sachs & World Economic Forum, 1999; 2000; 2002). The *World Business Environment Survey* (WBES) is a World Bank Group initiative that assesses the enabling environment for private enterprise in a large number of countries. This survey has an important indicator of corruption called General corruption constraint.

Table 3 presents the statistics of these frequently cited indicators of corruption, and their correlations to the WSJ news. All the correlations are strong (i.e. all p-values are less than 1%). As in the CPI case, the cross-section-time correlation for the eight strong countries is still extremely high (R = -0.910, p-value < 0.0001). The rejection of the causal assumption between CPI score and WSJ news does not reduce the value of the current study. Instead, we believe that WSJ news can be regarded as another indicator of corruption for each country.

Conclusion

The current study examined the relationship between corruption news in the WSJ and the CPI. Based on our dataset of news for 80 countries, the cross-section-time regression analysis suggests a strong correlation between these two variables. Time does not matter to the CPI score across the past nine years, with only a slight converging trend of the dispersion of CPI value given by the different respondents. After clustering the 80 countries into four groups according to their appearing frequency in WSJ, we found out that correlation in the four groups varies greatly. Group 1 includes eight strong countries (United States, Japan, China, United Kingdom, Germany, Canada, France, and Russia) and shows extremely high correlation.

Table 3. Regression on Some Other Corruption Indicators

Indicators of Corruption	Mean	Std Dev	Min	Max	Correlation on News		
					N	R	p-value
Global Competitiveness Report:							
Irregular payment (2002) [a]	5.133	1.117	2.350	7.000	56	-0.470	0.0003
Irregular payment (2000 -2002, 60 countries/regions) [b]	5.002	1.266	1.940	7.000	162	-0.414	<.0001
Irregular payment (2000 -2002, 8 countries) [c]	5.487	1.242	2.200	6.500	24	-0.910	<.0001
Burden of corruption (2002) [d]	4.896	1.027	2.900	6.900	56	-0.496	<.0001
World Bank Research:							
General Corruption constraint (2001) [e]	2.494	0.721	1.180	3.550	40	0.463	0.0026
Graft Corruption Index (1997-1998) [f]	3.819	1.966	0.740	6.910	59	0.577	<.0001

Notes:

a. The cross-section data of "irregular payment (2002)". Here, the irregular payment is the averaged measure across five aspects, that is, irregular payment in exports and imports, in government procurement, in tax collection, in public contracts and in loan applications (*Global Competitiveness Report*, 2002: 7.01-7.05).

b. The cross-section-time data of "irregular payment (2000-2002)" for all the 60 countries/regions (*Global Competitiveness Report*, 2002: 7.01-7.05; 2001: 4.03; 2000: 8.03).

c. The cross-section-time of "irregular payment (2000-2002)" for the top 8 countries frequently reported by WSJ (*Global Competitiveness Report*, 2002: 7.01-7.05; 2001: 4.03; 2000: 8.03).

d. The cross-section data of "business costs of corruption (2002) " . (*Global Competitiveness Report*, 2002: 7.06).

e. The cross-section data of "general Corruption constraint (2001)" (Research datasets of World Bank (http://www.worldbank.org/data, accessed 18 November 2003).

f. The cross-section data of *Graft Corruption Index* (1997-1998)" (Research datasets of World Bank (http://www.worldbank.org/data, accessed 18 November 2003).|

The practical application of our findings is important. WSJ news concerning corruption can be viewed as a one valuable data source to add to the corruption index, especially for the eight strong countries. Additionally, when compared to costly worldwide surveys, this indicator has the advantage of being low-cost and high-convenience.

From a theory-development standpoint, our study enhances understanding of the perceived corruption index because WSJ news can be regarded as the filtered result of corruption information after the perception of worldwide journalists and editors. Interestingly, irrespective of the different background of respondents, their taste seems to correlate soundly. This finding shows the robustness of perceptual figures, and contributes to understanding of real levels of corruption. Another implication of this study to theory is its illustration for exploring causality. The strong association between WSJ news and CPI score may lead to incorrect assumptions, but after introducing culture into our model, the partial correlation convinced us that there is no causal relation between these two variables. The Ballantine Illustration has provided a further and clear explanation about causality in general.

As always, one must exercise caution when generalizing from one study. An important limitation of our study was potential error in the data collection process. In searching for news items for each country, we simply used items with an exact match for the word "corruption", even though this may have caused some miscounts (for example, some positive reports about anti-corruption actions may have been mistaken as negative). However, we expect that the uniform collection method across all the countries may reduce this bias to an acceptable range. Another limitation of this study was its small dataset. Although the cross-time-section collection method considerably expanded our sample size, the number of observations for each country is relatively small, which made some more rigid methods of analysis impossible.

Endnotes

1. In order to expand the sample size from nine observations we used cross-time-section data in our actual analysis. The Italian case in the Introduction is only a preliminary example.

2. The U.S. WSJ is different from other editions, and as the case of Mexico shows, factors such as geographical distance influences news reports here.

Acknowledgments

References

Ali, A. M. and Isse, H.S. (2003). Determinants of Economic Corruption: A Cross-country Comparison. *Cato Journal*. 22 (3): 449-466.

Brunetti, A., Kisunko, G. and Weder, B. (1997). *How Businesses See Government: Responses from Private Sector Surveys in 69 Countries*. The World Bank: Washington, D.C.

Coase, R.H. (1979). Payola in Radio and Television Broadcasting. *Journal of Law and Economics*. 22: 269-328.

Cohen, J. and Cohen, P. (1983). *Applied Multiple Regression / Correlation Analysis for the Behavioral Sciences, 2nd Ed*. Lawrence Erlbaum Associates: New Jersey.

Fiske, S.T. and Taylor, S.E. (1991). *Social Cognition, 2nd* Ed. McGraw-Hill: Toronto.

Getz, K.A. and Volkema, R.J. (2001). Culture, Perceived Corruption, and Economics. *Business and Society*. 40 (1): 7-31.

Gwartney, J. and Lawson R. (2003). *Economic Freedom of the World: 2003 Annual Report*. The Fraser Institute: Vancouver.

Houston, R. and Graham, J.L. (2001). Culture and Corruption in International Markets: Implications for Policy Makers and Managers. *Consumption, Markets and Culture*. 315 (3): 207-343.

Husted, B.W. (1999). Wealth, Culture and Corruption. *Journal of International Business Studies*, 30 (2): 339-59.

Kaufmann, D., Kraay, A. and Zoido-Lobaton, P. (1999). *Governance Matters*. http://info.worldbank.org/governance (accessed 12 November 2003).

Knack, S. and Azfar, O. (2001). Trade Intensity, Country Size and Corruption. *Economics of Governance*. 4 (1): 1-18.

LaPalombara, J. (1994). Structural and Institutional Aspects of Corruption. *Social Research*. 61 (2): 325-350.

Macrae, J. (1982). Underdevelopment and the Economics of Corruption: A Game Theory Approach. *World Development*. 10 (8): 677-687.

Oldenbury, P. (1987). Middlemen in third-world corruption: Implication of an India Case. *World Politics*. 39: 508-535.

Pavarala, V. (1996). *Interpreting Corruption, Elite Perspectives in India*. Sage Publications: Thousand Oaks.

Schwab, K., Porter, M.E., Sachs, J.D. and World Economic Forum (2000). *The Global Competitiveness Report 1999*. Oxford University Press: Oxford.

Schwab, K., Porter, M.E., Sachs, J.D. and World Economic Forum (2001). *The Global Competitiveness Report 2000*. Oxford University Press: Oxford.

Schwab, K., Porter, M.E., Sachs, J.D. and World Economic Forum (2002). *The Global Competitiveness Report 2001-2002*. Oxford University Press: Oxford.

Shleifer, A. and Vishny, R.W. (1993). Corruption. Q*uarterly Journal of Economics*. 108 (3): 599-617.

Transparency International (2003). *Corruption Perceptions Index 1995-2003*. http://www.transparency.org (accessed 10 November 2003).