

Anti-corruption Media Coverage and Corruption Perception

Thuy Nguyen

*School of Public and Environmental Affairs, Indiana University Bloomington, USA,
thdnguye@indiana.edu*

Abstract

Despite increasing critique on widely used cross-country perception-based corruption indexes, there is scarce evidence testing the validity of these measurements. Using media content analysis and the instrumental variable method, this paper provides evidence on how the Corruption Perception Index of Transparency International is influenced by media. An exogenous increase in anti-corruption news coverage caused by media infrastructure changes might lead to a higher incidence of perceived corruption. In addition, the Corruption Perception Index positively responds to a country's ratification of global or regional anti-corruption conventions. These findings suggest that many raters of perception-based indexes tend to portray corruption as more serious and pervasive when exposed to news about corruption-crackdown efforts by a country's government institutions.

Keywords: Corruption Perception Index, Perception-based measure, Media coverage, Anti-corruption efforts.

1. Introduction

Corruption is a globally widespread phenomenon that harms local firms and citizens as well global economy. There is tremendous growth in practitioners' attention and campaigns to combat corruption such as the United Nations Convention Against Corruption 2003 and the Organisation for Economic Cooperation and Development (OECD) Anti-Bribery Convention 1997. However, "measurement is the first step that leads to control and eventually to improvement", according to James Harrington, a quality advisor for Ernst and Young. Even with an emerging international consensus of a core definition of corruption as the "misuse of public power for private benefits" (Jensen et al., 2010), corruption is a multi-faceted phenomenon which is difficult to measure accurately.

Despite increasing critique on widely used cross-country perception-based corruption indexes such as the Transparency International's Corruption Perception Index (CPI), the CPI has been the most frequently used index in macroeconomic studies of corruption and widely disseminated among policymakers (Svensson, 2005; Razafindrakoto and Roubaud, 2010). Christiane and Charles (2006); Sampford et al. (2006); Knack (2007) are several examples of using content validation by mainly providing conceptual assessments of perception-based corruption measurement. These studies raise concerns about the corruption measures in terms of underlying theory and aggregation strategies. However, Carmines and Zeller (1979); Adcock (2001) indicate that content validation is harder to use if concepts are abstract and complex, which is the case of corruption.

In recent years a handful of papers have attempted to empirically examine the validity of these perception-based corruption measures using convergent validation procedures.¹ There is mixed evidence on the measurement validity of the perception-based corruption measures (see Table A1). For instance, Lambsdorff (1999) provided the robustness test for the CPI by showing high correlations between the sources of the CPI. Mocan (2008) found high correlations among the author's aggregate corruption index created from the the International Crime Victims Surveys (ICVS), the CPI, and the corruption index by the Business International and International Country Risk Guide (ICRG). Ko and Samajdar (2010) also analyzed high correlations among different perception-based indicators of corruption to show that the validity problems of these indexes are not as serious as critics claims. Nevertheless, Abramo (2008) challenged the CPI by a comparison between the CPI and personal experience of bribery from the Transparency International Global Corruption Barometer 2004 survey within the same-country samples. The author shows that perceptions are not good predictors for experiences of corruption, and by extension the actual incidence of corruption. Hawken and Munck (2011) provided evidence against the CPI by grouping its sources into 5 classes and showing significant variations among these classes. Donchev and Ujhelyi (2014) showed that household and firms' experience with corruption is a weak determinant of the CPI.

¹ Convergent validation addresses whether the scores produced by alternative indicators of a given concept are empirically associated, which heavily relies on correlation analysis (Adcock, 2001).

The main concern about using correlation or regression analyses in the context of corruption is that none of the baseline indicators may be a very good, or “true” measures of corruption. Because corruption is by its nature an illegal activity, relevant objective data are almost impossible to obtain (Kaufmann et al., 2006). Olken (2009) and Razafindrakoto and Roubaud (2010) compared the perception-based measures with more subjective baseline indicators in several specific regions, meanwhile, their tests are unlikely to invalidate the competencies of cross-country perception-based measures.

The various raters of the perception-based indexes may rely on media reports which cover corruption incidents and anti-corruption efforts in countries. Lambsdorff (1999), the developer of the CPI, raised a concern about media influence on raters who might not report their personal experiences but rely on media coverage and reports obtained from others. An example of worsening perceptions of corruption, even though actual corruption fell after the fall of Soeharto in 1998 in Indonesia, was discussed by Pande and Olken (2012). The authors argue that the worsening perceptions might be led by a much freer press which was newly able to report on allegations of corruption. Recently, the international media focused attention on the endeavor of China’s new leadership to combat its rampant corruption in China. However, the CPI of China fell from a score of 39 in 2012 to 36 in 2014 due to worsening perceptions of corruption. These cases illustrate the extent to which the CPI might be greatly influenced by the mass media. Prior empirical evidence on media influence on corruption perception measures remains rare.

The aim of this work is to examine how news coverage influences the most widely used cross-country corruption index, the CPI, and therefore undermines its measurement validity. This study firstly uses search engines of the Lexis-Nexis database to acquire and validate the total number of news stories regarding anti-corruption of 176 countries over the period 1995-2016, which proxies for tangible cross-country anti-corruption efforts. Then, the effects of anti-corruption news coverage on the CPI and other composite indexes of corruption are estimated by both fixed-effects and instrumental variable estimations. Broadband infrastructure is used to instrument endogenous anti-corruption news coverage in order to provide more compelling evidence of media influence on corruption perception.

This study contributes to the literature on the measurement of corruption by

providing some of the first compelling evidence of media influence on the most widely used indicator of corruption. The key evidence is that an exogenous increase in anti-corruption news coverage leads to a considerably higher level of perceived corruption. This positive effect of anti-corruption media suggests that many raters of perception-based indexes tend to portray corruption as more serious and pervasive when exposed to news about corruption-crackdown efforts by a country’s government institutions. The significant effect of media on the CPI also undermines the validity of this perception-based measure.

The rest of the paper is structured as follows. Section 2 describes data acquisition for this paper. Section 3 discusses the empirical strategies used for the analyses. In Section 4, the empirical results are discussed. Finally, the conclusions are given in Section 5.

2. Data

This paper primarily acquires two sources of data, news coverage regarding anti-corruption in countries collected from the Lexis-Nexis database and country-level economic data (including aggregate indexes of corruption). The variables and their sources are summarized in Table A2. This section provides relevant details on the main data sources.

2.1. Anti-corruption Media Coverage

The total number of news stories regarding anti-corruption in a country was acquired by using search engines of the Lexis-Nexis Database and two coders. The Lexis-Nexis provides full-content access to many news sources, which has long been one of standard research tools (Weaver and Bimber, 2008), and is the most widely used digital news archive in social scientific studies (Deacon, 2007). Zhang (2001) provided empirical evidence that these internet-based resources including the Lexis-Nexis database are becoming an important component in scholars’ research with a wide range of purposes and functions. Searches for keywords in the Lexis-Nexis database have been utilized in a numerous number of media analyses such as Moynihan et al. (2000); Fisman (2001); Shoemaker et al. (2001). Moynihan et al. (2000) for example used Lexis-Nexis to obtain stories from US newspapers and analyzed news coverage of medications. Fisman (2001) found the relevant news

about former President Suharto's health in the Lexis-Nexis database to assess the value of political connections in Indonesia. [Shoemaker et al. \(2001\)](#) conducted a quantitative content analysis using the Lexis-Nexis database to examine the prominence of a congressional bill's coverage in the US national newspaper.

Any search of digitalized news archives has to be based on the use of keywords ([Deacon, 2007](#)). This dependence on keywords raised the author's attention to carefully select several keywords and use Boolean searches to obtain relevant news. Searches were conducted using the terms "anti-corruption" and "country name" or "anti-corruption" and "country name" under the segment search category HEADLINE, from all news source types, within a year for 176 countries between 1995 and 2016, in the Lexis-Nexis Academic news database. The selection of these countries and covered period are based on the availability of perception-based corruption indicators. The keywords were searched in the HEADLINE segment because it contains the most important part of any article. The majority of people also use headlines to decide whether or not to read that article. Two independent coders then filtered relevant stories by sorting geography (matched to the searched country) and choosing high-similarity duplicate option (to exclude highly similar stories). The filtered relevant stories about anti-corruption were counted and coded.

There are a number of validation exercises aimed toward validating the measure of anti-corruption news coverage. Firstly, reproductivity or intercoder reliability, the degree to which a process can be replicated by different analysts working under varying conditions ([Krippendorff, 2012](#)), is examined in this study. The average pairwise Percent Agreement, in which the agreements of all possible pairs are calculated and averaged, is 0.922. This simple coefficient indicates that 92.20% of the frequency of relevant news stories are exactly the same among two independent coders. The general rule of thumb for an acceptable percent agreement is a coefficient of 0.90 or greater ([Neuendorf, 2002](#)). Therefore, the coding is accurate and reliable. The average frequency of relevant news stories is used in the quantitative analysis of this paper.

Texts of more than 300 sampled articles from Vietnam, South Korea, Colombia, and Austria have been checked in order to confirm that the number of acquired news stories about anti-corruption of a country indeed reflects tangible anti-corruption efforts. These articles reflect each country's intention and efforts to combating

corruption, including (but not limited to): launching anti-corruption laws and programs, reporting corruption inspections, signing anti-corruption cooperation agreements, and urging more anti-corruption actions by donors. Interestingly, there was a record number of the anti-corruption news stories (48 stories) in 2006 after the enactment of the first Vietnam's anti-corruption law in the end of 2005 (Gainsborough et al., 2009). This validation exercise has confirmed that the number of acquired news stories about anti-corruption of a country indeed reflects tangible anti-corruption efforts.

In addition, an alternative search strategy with the same keywords under the segment search category HEADLINE & LEAD was performed in order to extend searches in lead paragraphs of news and to provide the sensitivity test of this study's content analysis. Because China's recent anti-corruption campaigns are overwhelming in other countries' news, "China" was excluded in the searches for other countries. The alternative news coverage is highly correlated to the data obtained from the main search strategy with a 0.7962 correlation coefficient for the whole sample.

There are a few concerns regarding the search strategy. First, the search strategy was unable to distinguish relevant news stories of a specific country released by a domestic press or foreign publisher. However, the aforementioned validation exercise has confirmed that the number of acquired news stories about anti-corruption of a country indeed reflects tangible anti-corruption efforts by this country. Second, the search strategy covers English-language newspapers, which is similar to other studies such as Fisman and Svensson (2007). Nevertheless, it might lead to a potential systematic difference between English-spoken countries and the others. One of the robustness check of this study is using the sample which excludes countries with an English official language.

The Lexis-Nexis searches for keywords provided us one important set of the data, the total number of news stories regarding anti-corruption in a country. This data represents the frequency of references to anti-corruption in the media and reflect tangible anti-corruption efforts of a country. In the full sample of 176 countries over the period 1995-2016, 2.26 anti-corruption stories of a typical country were annually reported in the media. The range of the counted stories is between 0 and 311.5. Zero anti-corruption news coverage is observed in 63.5 percent of the sample. A

record number of anti-corruption stories were found in Nigeria in 2015. The final sample, used in the analysis, covers 165 countries over the period 1996-2015.

2.2. Composite Indexes of Corruption

This study focuses on examining the empirical link between anti-corruption news coverage and the Corruption Perception Index. The CPI, developed and published by Transparency International since 1995 to capture the “misuse of public power for private benefit”, has been used the most in macroeconomic studies of corruption and widely disseminated among policymakers (Svensson, 2005; Razafindrakoto and Roubaud, 2010). In particular, the CPI is a composite index simply aggregating different sources of corruption-related data that are primarily expert assessments from governmental, non-governmental organizations and commercial institutions such as the World Bank’s Country Policy and Institutional Assessment, World Economic Forum Executive Opinion Surveys, African Development Bank Governance Ratings, Political Risk Services International Country Risk Guide, and Economist Intelligence Country Risk Ratings. For example, the 2009 CPI covers 180 countries and is based on 13 sources from 10 independent institutions. Notably, the first group of sources, e.g. the World Bank’s Country Policy and Institutional Assessment, and Economist Intelligence Unit Risk Ratings, assemble the perceptions of non-residents from developed countries (Lambsdorff, 2007). The second group, e.g. the World Economic Forum Executive Opinion Surveys, gather perceptions made by residents with respect to the performance of their home country. However, these respondents are partly nationals and sometimes expatriates from multinational firms. The CPI in some years also includes information from experience-based surveys of corruption such as the International Crime Victims Surveys, nationally-representative household surveys, and the World Bank Enterprise Surveys (WBES), nationally-representative firm surveys (Donchev and Ujhelyi, 2014). Countries were included in the calculation of the CPI if they were evaluated by at least three sources. A matching percentiles technique was used to place the sources into a common scale before 2012. Since 2012, the CPI has been calculated using a simple average of standardized scores.

Another widely used corruption index, the Control of Corruption estimate from the World Bank’s Worldwide Governance Indicators (WGI), subsequently extended to the Bayesian Corruption index (BCI), is also examined in this paper. The

WGI index has been developed since 1999 in order to measure “the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests” (Kaufmann et al., 2011). The WGI index is a weighted aggregate of the individual indicators of each source using an unobserved components model. For instance, the 2009 WGI covers 211 countries and is based on a total of 22 data sources from various nongovernmental organizations, commercial business information providers, and surveys of individuals or domestic firms. Standaert (2015) extended the WGI methodology to construct the BCI in order to improve aggregation of corruption perception. The BCI is estimated using a Bayesian Gibbs sampling algorithm with the same sources as the WGI are used.

3. Empirical strategy

3.1. Identification

The frequency of references to anti-corruption in the media reflects tangible anti-corruption efforts of a country, which might reduce the actual level of corruption. Di Tella and Schargrodsky (2003), Olken (2007), and Ferraz and Finan (2011) are several examples of empirical evidence of effective anti-corruption efforts and policy interventions in curbing corruption.² Therefore, the first validation exercise of the CPI is to examine how this index responds to anti-corruption news coverage of a country. Any positive associations between perception-based measures and anti-corruption news coverage might shed some light on the potential media bias effect on these measures. That is, many raters tend to portray corruption as more serious and pervasive when exposing news about corruption-crackdown efforts by a country’s government institutions.

The key concern to estimate the effect of anti-corruption news coverage on a

² Di Tella and Schargrodsky (2003) examined the roles of wages and auditing intensity in a crackdown on corruption in the public hospitals of the city of Buenos Aires, Argentina. Olken (2007) used a randomized field experiment in Indonesia and found government audits had a significant impact on reducing corruption while community-based monitoring had little average impact. Additionally, Ferraz and Finan (2011) showed that electoral accountability might reduce corruption practices of incumbent politicians by using audit reports from an anti-corruption program in Brazil.

perception-based measure is their two-way causation. In particular, more anti-corruption efforts might reduce the actual level of corruption, as a result, anti-corruption news coverage might lead to a lower indicator of reliable measures of corruption. However, a country with a higher level of corruption might devote more anti-corruption efforts, therefore more anti-corruption news articles are released. The instrumental variable method was employed to address this endogeneity issue. Anti-corruption news coverage was instrumented with a measure of media infrastructure. Variations in media infrastructure might cause changes in quantity and quality of news stories. It is plausible that telecommunication infrastructure is an important factor shaping the media sector on both audience and news-maker sides. Telecommunication infrastructure developments allow audiences to have a better access to reliable news and information. In addition, from the news-makers' perspective, journalism has always been shaped by technology (Pavlik, 2000; Steensen, 2009, 2011). Steensen (2009, 2011) found that online journalists practice a more audience-driven and source-detached kind of journalism than their print counterparts, and their online newsroom is considerably shaped by new technology. Therefore, it is well established that media or telecommunication infrastructure is a relevant factor influencing news coverage of a general topic including anti-corruption topics.

Broadband infrastructure is chosen to instrument endogenous variations in anti-corruption news coverage. Broadband infrastructure might enable high-speed internet access (Czernich et al., 2011) which allows the generation and distribution of decentralized information and ideas in markets. High-speed internet via broadband infrastructure may accelerate the distribution of ideas and information, as a result, encourage the development of online news media. Therefore, the IV is the 1-year-lagged broadband (high-speed internet services) penetration rate, which is the number of fixed (wired) broadband subscriptions per 100 inhabitants of a population from the World Bank's World Development Indicators database (WDI).

A key underlying assumption of the IV method, the exclusion restriction, is that valid instruments do not have any direct effect on the dependent variable or any effect running through omitted variables. The first concern about the exclusion restriction is any direct link between media infrastructure and corruption. One might argue that corruption leads to wasted resources and distortions in infrastructure investments.

However, there are several reasons plausibly ruling out this argument. The supply of telecommunications infrastructure is generally determined by economic, political, and geographic variables including geographic areas of a country, the government deficit, the waiting list per capita (excess demand at the price), and the price of telephone services given that market structure and the role of governments vary a great deal (Röller and Waverman, 2001). Corruption is likely to directly influence public infrastructure investments but not private infrastructure's ones which are more prevalent in the media sector. For instance, Djankov et al. (2001) examined the patterns of media ownership in 97 countries and found that family controlled newspapers account for 57% of the total while the state controls approximately 29% of newspapers. Furthermore, under the ongoing globalization of media markets and the development of Information Communication technologies (ICT), the corrupt politicians seem not to be able to control the development of media networks. Therefore, it is plausibly no direct link between media infrastructure and media.

The second concern about the exclusion restriction is any link between media infrastructure and corruption running through omitted variables. Röller and Waverman (2001) for instance investigated how telecommunications infrastructure affects economic growth and find a significant positive causal link. In addition, ICT may contribute to building e-governments which facilitates provision of relevant government information to the citizens, empowerment of the people through access to information, and participation in public policy decision-making (Palvia and Sharma, 2007). As a result, media infrastructure might be associated with democracy and transparency. In order to minimize this concern, country fixed-effects and the main time-variant country-level determinants of corruption including economic development, openness, and democracy are controlled for.

3.2. Specifications

To estimate the effect of anti-corruption news coverage on corruption measured by different perception-based indexes, time-series cross-section analysis was applied to a panel data of more than 165 countries over 20 years, or a long time-series panel which is different from longitudinal panels with a short time period. In particular, the rescaled CPI of country k at year t is a function of *news coverage* $_{kt}$, $\log(0.5 \text{ CPI}_{k,t-1} + 0.5 \text{ CPI}_{k,t-2})$, country-level determinants C_{kt} , time trend t , squared time trend t^2 , country fixed-effects, and a normally distributed unobserved

error term ε_{kt} , as the following:

$$\begin{aligned} \log(CPI_{kt}) = & \delta_0 + \delta_1 news\ coverage_{kt} + \log(0.5 CPI_{k,t-1} + 0.5 CPI_{k,t-2}) \\ & + \delta_c C_{kt} + \delta_2 t + \delta_3 t^2 + \delta_{4k} country_k + \delta_{5t} year\ dummies_t + \varepsilon_{kt} \end{aligned} \quad (1)$$

The XTSCC program developed by [Hoechle et al. \(2007\)](#), is used to provide fixed-effects (FE) estimates with Driscoll-Kraay standard errors robust to disturbances being heteroscedastic, autocorrelated of the moving average type with lag length q (up to 4), and cross-sectionally dependent.³ The specification in equation (1) enables to fix a couple of potential estimation issues of a time-series cross-section analysis: omitted population heterogeneity, unit-specific trends, and autocorrelation (see [Worrall and Pratt \(2004\)](#) for discussion and solutions for these issues). The average of the 1-year-lagged and 2-year-lagged CPIs is also controlled in the estimation because some data sources of the CPI are updated only once every two or three years.

Omitted population heterogeneity is unobservable and time-invariant characteristics of the units of analysis, therefore, it is canceled out by a FE method. This method is also able to control for unmeasurable time-invariant differences between units. Social norms regarding corruption such as the degree to which the general public tolerate corruption may be considered as one example of unmeasurable time-invariant difference across countries. This is because, as one of social institutions, social norms regarding corruption is quite stable in a long period of time. In addition, year dummies were added to control for unit-invariant differences between time periods (for example, global shocks and international anti-corruption campaigns). However, there might exist some variations in a unit (country) that departs from the trends captured by year dummies, thus an unit-specific time trend variable coded from 1 to t for each unit and its squared term were included.

The selection of exogenous determinants of corruption was based on the previous studies (for examples, [Treisman \(2000\)](#); [La Porta et al. \(1999\)](#)). The

³ [Driscoll and Kraay \(1998\)](#) propose a nonparametric covariance matrix estimator which produces heteroscedasticity consistent standard errors that are robust to very general forms of spatial and temporal dependence.

main time-variant country-level determinants of corruption were controlled for, including economic development, economic openness, and democracy in the baseline specification. Economic development is measured by per capita income, constant GDP per capita using international \$ from the WDI. Economic openness is measured by imports of goods and services as a share of GDP, from the WDI. Democracy is primarily measured by the index of democratization constructed by Vanhanen and Lundell, which portrays the electoral success of smaller parties and the voting turnout in each election. An alternative measure of democracy is the deliberative democracy index from Varieties of Democracy (V-Dem) Project, a new approach to conceptualizing and measuring democracy. In one alternative specification, transparency was also controlled for, which is proxied by the press freedom index from the Free House. In [Brunetti and Weder \(2003\)](#); [Lindstedt and Naurin \(2010\)](#) for instance, the press freedom indicator was used to proxy for transparency in order to assess the relationship between transparency and corruption. Controlling for transparency in the country-level estimation aims at ruling out the potential link between news coverage and the CPI running through transparency.

To estimate the effect of exogenous variations in anti-corruption news coverage on perception-based measures of corruption, the IV/two-state least squares regressions are implemented. The first-stage regression equation is:

$$\begin{aligned} News\ coverage_{kt} = & \alpha_0 + \alpha_1\ media\ infrastructure_{k,t-1} + \alpha_c C_{kt} + \alpha_2 t + \alpha_3 t^2 \\ & + \alpha_{4k} * country_k + \alpha_{5t} year\ dummies_t + e_{kt} \end{aligned} \quad (2)$$

Where $media\ infrastructure_{k,t-1}$ is the broadband penetration rate of country k at year $t - 1$ and e_{kt} is a normally distributed unobserved error term. The IV/two-state least squares regressions are implemented by the XTIVREG2 package in Stata, which is developed by [Schaffer et al. \(2015\)](#). Standard errors are robust for heteroscedasticity, and autocorrelated error terms (with a lag length up to 4). A similar empirical strategy was applied to estimate the effect of anti-corruption news coverage on the WGI and BCI index.

4. Results and Discussion

4.1. Media Raises More Concern about Corruption

Table [A3](#) presents the time-series cross-section results which suggest significant and positive association between anti-corruption news coverage and the CPI.

Columns 1 to 4 show country fixed-effects (FE) estimates with Driscoll-Kraay standard errors corrected for heteroskedastic, autocorrelated (up to 4 lags), and cross-sectionally dependent disturbances. The Pooled-Ordinary Least Squared (OLS) estimation in column 5 controls for both time-invariant and time-variant determinants of corruption.

Column 1 of Table A3 shows the baseline specification which covers a sample of 165 countries over the period 1996-2015. The positive and significant coefficient on news coverage implies that a standard-deviation increase in the number of anti-corruption news stories per one thousand internet users (0.05 news stories) is associated with a 0.42-percent increase in the rescaled CPI (a 0.24-point increase at mean). This result confirms [Lambsdorff \(1999\)](#)'s concern about media influence on raters of the CPI who might not report their personal experiences instead relying upon media coverage and reports obtained from others. News coverage of anti-corruption (the frequency of references to anti-corruption in the online media) plausibly reflects cross-country anti-corruption efforts. Anti-corruption efforts might reduce the actual level of corruption.⁴ An accurate measure of corruption is thereby expected to decline with more intensive news coverage of anti-corruption. As a result, the positive association between the CPI and news coverage of anti-corruption might shed some light on the potential media bias of the CPI.

The coefficients on the control variables in the baseline specification are generally statistically significant with signs consistent with theories and previous analyses. The significant and negative coefficients on GDP per capita and imports share are consistent with [Treisman \(2000\)](#)'s findings. For instance, countries with more developed economies and higher imports were perceived as less corrupt. The degree of democracy is less significant in the baseline specification while it was insignificant in [Treisman \(2000\)](#)'s analysis. An alternative measure of democracy is the deliberative democracy index obtained from Varieties of Democracy Project, which indicates the extent to which the ideal of deliberative democracy is achieved in a country. The alternative specification with this alternative measure of democracy is reported in column 2, Table A3. Though the number of observations dropped

⁴ [Di Tella and Schargrodsky \(2003\)](#), [Olken \(2007\)](#), and [Ferraz and Finan \(2011\)](#) are several examples of empirical evidence of effective anti-corruption efforts and policy interventions in curbing corruption.

by 12%, this alternative specification provides a more significant and slightly larger estimate on the coefficient of news coverage. The second alternative specification, which controls for transparency (proxied by the press freedom index), yields results highly similar to those reported in the baseline specification. The degree of press freedom entered insignificantly into the model (column 3, Table A3). The third estimation includes quadratic news coverage, which is reported in column 4, Table A3. The significant and negative coefficient on this term suggests a diminishing influence of media on the CPI. Additionally controlling for Protestant traditions, histories of colonial rules, legal regimes, and region dummies, the Pooled OLS estimation yields a slightly smaller point estimate for news coverage, but it is statistically significant only at the 85 percent confidence level.

4.2. Anti-corruption Efforts Raise More Concern about Corruption

Table A4 reports the time-series cross-section estimations on the effects of the most important anti-corruption global and regional conventions on the CPI, including: the United Nations Convention against Corruption (UNCAC), the OECD Anti-Bribery Convention, the Group of States against Corruption - European Anti-corruption Convention (GRECO), the Inter-American Convention against Corruption (IACAC), and the African Union (AU) Convention on Preventing and Combating Corruption. The results suggest that, in a couple of following years after ratifying these conventions, corruption is perceived as more pervasive and serious in member states of these conventions.

Columns 1 and 2 show the estimates on Conventions, the combination indicator of anti-corruption conventions which equals 1 if a country ratifies one of the above conventions and equals 2 if it ratifies two of the anti-corruption conventions in a specific year. The 1-year-lagged and 2-year-lagged Conventions are positively associated with the CPI. Column 3-4 show the estimates on anti-corruption conventions separately. Ratifications of the UNCAC, GRECO, or IACAC raise more concern about corruption in the following year. Ratifications of the AU Convention on Preventing and Combating Corruption is statistically insignificant. Ratifications of the OECD Anti-Bribery Convention is negatively associated with the CPI with only the 90 percent confidence level. Evidence on positive association between anti-corruption efforts and the CPI is consistent to the finding of positive association between news coverage of anti-corruption and the CPI in Section 4.1 .

4.3. Exogenous Changes in New Coverage and Corruption Perception

Positive association between anti-corruption new coverage (or efforts) and the CPI might be a result of the reverse causation between anti-corruption efforts and corruption level. In particular, a country with a higher level of corruption might devote more anti-corruption efforts or call more attention to combating corruption in media. The IV estimations, reported in Table A5, aim to address endogeneity issues and provide more robust estimates of the effects of anti-corruption news coverage on the CPI. The IV estimation results imply that anti-corruption news coverage has a statistically significant and positive impact at the 95 percent confidence level on the CPI. The IV estimates on news coverage are significant within a range of 6.58 to 8.15 across the different specifications utilized.

In the baseline specification in column 1 of Table A5, the point estimate indicates that one standard deviation increase in the number of anti-corruption news stories per one thousand internet users is associated with a 39.19-percent increase in the rescaled CPI (a 22-point increase at mean - one standard deviation). Therefore, an exogenous increase in anti-corruption news coverage leads to a considerably higher incidence of perceived corruption. This finding is consistent with the argument of the potential upward bias of anti-corruption news coverage on the CPI. Particularly, exposing more news about corruption-crackdown efforts by a country leads selective raters to likely portray corruption as more serious and pervasive. An example of worsening perceptions of corruption, even though actual corruption fell after the fall of Soeharto in 1998 in Indonesia, was discussed by Pande and Olken (2012). The authors argue that the worsening perceptions might be led by a much freer press which was newly able to report on allegations of corruption.

4.3.1. Robustness Checks

Columns 1 to 4 of Table A5 present the IV estimation results using the 1-year-lagged broadband penetration rate to instrument anti-corruption news coverage. Standard errors are clustered within country-industry groups and robust to heteroscedasticity and serial autocorrelation up to 4 lags. The IV estimates are consistent across a couple of specifications. Both the Cragg-Donald Wald F statistic and the Kleibergen-Paap rk Wald F statistics are reported and used to test the

relevance of the IV.⁵ Referring to the Stock-Yogo critical values for the relevance test, the F statistics exceed 10, which confirms the relevance of the telephone penetration rate (a measure of media infrastructure) in capturing exogenous changes in anti-corruption news coverage. The first-stage regressions, reported in Table A5, provide additional evidence on the relevance of the broadband penetration rate.

Column 5 of Table A5 shows the IV estimation results combining the 1-year-lagged broadband penetration rate, the 1-year-lagged ratification, and the 2-year-lagged ratification of the anti-corruption conventions. This alternative identification strategy produces results fairly similar to the baseline specification in column 1. However, neither the 1-year-lagged ratification nor 2-year-lagged ratification of the anti-corruption conventions are statistically significant in the first stage regressions.

Although IV estimations are considered to be dependent on the data at hand, the IV estimates are robust across different sub-samples. Table A6 reports IV estimation results using the baseline specification on 8 sub-samples. The news coverage data of this study is collected from LexisNexis’s English news database; therefore, news coverage of English-speaking countries (official language) might be different from those of their counterparts. Using the sample of countries where English is not their official language marginally modified the result of the IV regression in column 1. The regressions, presented in columns 2 and 3, cover observations after 2003 and after 2004 - the period of explosively increasing use of information and communication technology as well as the number of internet users (Wunnava and Leiter, 2009). Using these samples produces results highly similar to those of the whole sample, though the IV estimate on news coverage is smaller. This result might be explained by the role of explosively expanded use of the internet and online media, as well as the diminishing effect of news coverage (which is discussed in section 4.1). Using either the sample excluding G20 or BRICs countries, or the Europe Union only marginally changed the IV estimates on news coverage (columns 4 and 6). Similarly, columns 7 and 8 report the IV results using samples excluding countries with a military political regime and a one-party regime, respectively.

An alternative search strategy within Lexis-Nexis Database was conducted with

⁵ The Kleibergen-Paap rk Wald F statistics are preferred when the independent and identical distributed disturbances are not assumed in the estimations.

the aim to test the sensitivity of this study’s content analysis. The same keywords, “anti-corruption” and “country name”, were searched throughout news stories’ *headline & lead* instead of only in the title. Because China’s recent anti-corruption campaigns are overwhelming in other countries’ news, “China” was excluded in the searches for other countries. The alternative archived news coverage (news coverage 2) is highly correlated with the baseline archived news coverage with a 0.8612 correlation coefficient. The average number of anti-corruption news per one thousand internet users using the alternative search is 0.06 for a typical country which is 10 times that of the base search. Applying the empirical strategy of the FE and IV baseline estimations on the alternative news coverage, the results are presented in Table A7. The IV estimates on news coverage are significant within a range of 0.65 to 0.77 which are considerably smaller than those of the baseline news coverage. The results suggest that anti-corruption news coverage had a statistically significant and positive impact on the CPI despite using different search strategies in this paper’s content analysis.

4.3.2. Countries Observed More Media Bias

The composite indexes of corruption compile from various data sources, primarily given by international businessmen and country experts, however, they also contain information from domestic experience-based surveys. Surveys of bribe-payers are considered as the most direct way of measuring bribery and one of the only areas providing consistent measurement of corruption over time and across countries (Olken, 2006). Unlike to the perception-based measures which mostly rely on questions about perceptions of corruption, surveys of bribe-payers place a greater emphasis on residents and domestic firms’ experience. The ICVS, one key dataset of cross-country representative households, were compiled by the United Nations Interregional Crime and Justice Research Institute. The WBES are nationally-representative surveys of firms assessing corruption and other problems related to their businesses. In general, the WBES survey asks senior firm executives comparable questions about firms’ informal payments in order to “get things done”, to secure government contracts, and to have access to public services.

Particularly, the CPI includes the frequency of bribery from the ICVS and WBES surveys for a number of countries in certain years. Notably, these surveys are not available in every country and every year. Countries without these surveys might

observe larger effects of media coverage on their scores in the CPI because the CPI would plausibly rely more on the perception of international businessmen and country experts. Columns 1 and 2 of Table A8 report the FE and IV estimations on the sample of observations with the availability of the ICVS and WBES surveys (the first and second following years of the survey year are counted for). Columns 3 and 4 report the estimations on the sample of observations without surveys.

The IV estimate on news coverage of the former sample is statistically insignificant and negative. In contrast, the IV estimate of the latter sample mostly doubles the IV estimate of the whole sample, and it is significant at the 99 percent confidence level. As the aforementioned expectation, countries without these experience-based surveys tend to observe larger media bias on their scores in the CPI. These results are consistent with the argument that media is most effective in changing the attitudes of those who are the least well informed (Schmidt, 1993; Iyengar et al., 1982). If the CPI is solely based on perception-based ratings by international business and country experts, its validity might be greatly undermined because of the media bias.

4.3.3. Media Influence on Other Perception-based Measures of Corruption

Estimations on another widely used perception-based index, the Worldwide Governance Indicator corruption index, are reported in columns 1 and 3, Table A9. The FE results imply positive association between anti-corruption news coverage and the WGI corruption index at the 99 percent confidence level. Nevertheless, the WGI corruption index is not significantly associated with a country's ratification of anti-corruption conventions. The IV estimate is positive, however, it is statistically insignificant. Columns 4 and 6 show the results of the Bayesian Corruption Index (BCI). The IV results strongly confirm that an exogenous increase in anti-corruption news coverage leads to a higher incidence of perceived corruption measured by the BCI. However, the FE estimations provide weaker evidence on positive association between news coverage and the BCI.

5. Conclusion

This paper has conducted various validation exercises on the most widely used measure of corruption, the Transparency International's perception corruption

index. The key validation exercise is using the frequency of references to anti-corruption in the online media (anti-corruption news coverage) of 165 countries over the period 1996-2015 to examine whether anti-corruption coverage increases the index of corruption.

The fixed-effects results suggest that the CPI is positively associated with anti-corruption news coverage. The IV estimations were carried out to address endogeneity issues due to the potential reverse causation between the CPI and anti-corruption news coverage. Broadband infrastructure, which enables high-speed internet access and accelerate the distribution of ideas and information, was employed to instrument the endogenous news coverage. The IV results show that an exogenous increase in anti-corruption news coverage leads to a considerably higher incidence of perceived corruption. For instance, on average, one standard deviation increase in the number of anti-corruption news stories per one thousand internet users raises the CPI by 22 points (almost one standard error). The IV estimates were found robust and consistent across different sub-samples, specifications, and an alternative measure of news coverage utilized. Notably, in countries where experience-based surveys of corruption are not available, media coverage has more influence on the CPI. The large effects of media coverage might be explained by the fact that the CPI tends to solely rely on inputs of international businessmen and country experts. In addition, the paper found positive effects of anti-corruption news coverage on the Bayesian Corruption index which is based on a new approach for aggregating corruption perceptions by [Standaert \(2015\)](#). Evidence of media influence on another widely used composite index of corruption, the Worldwide Governance Indicator corruption index, was only found in the fixed-effects estimations.

This work has several important implications. Firstly, a new empirical protocol is disclosed which assesses measurement validity of corruption measures in particular as well as measures of other abstract and complex constructs by using media content analysis and the instrumental variable method. Media content analysis is a sub-set of content analysis which is a widely used and well-established research method in different fields such as health, media, and communication studies ([Macnamara et al., 2005](#)). Nevertheless, [Fisman \(2001\)](#), and [Palau and Davesa \(2013\)](#) are notable examples of the scarce empirical studies of corruption applying this research tool.

Secondly, this paper provides one of the first empirical evidence of media influence on corruption perception. In particular, exogenous variations in anti-corruption news coverage causes changes in the aggregate index of corruption perceptions.

There are a couple of limitations in this study. Focusing on counting the frequency of specific words (anti-corruption) might neglect considerable variations in stories regarding corruptions across different countries due to different types of articles, tones, and languages. Future work using more advanced media content analysis which goes beyond word counts will help us better measure anti-corruption efforts. Using latent content analysis which focuses on discovering underlying meanings of the words is an interesting avenue for future research ([Hsieh and Shannon, 2005](#)). In addition, there are several cross-country measures of corruptions which have not been examined in this study. For example, the corruption index of the Political Risk Services International Country Risk Guide is produced by a team of country experts. The ICRG is a component of both the CPI and WGI, however, it is not publicly available.

Appendices

Table A1: A survey of studies on assessing for measurement validity of corruption indexes

Indexes assessed	Validation procedure	Conclusion	Authors
Reported corruption perceptions by Indonesian villagers	Comparing with “missing expenditure” by correlation and regression analysis	The reported corruption perceptions do contain real information of corruption, but are insufficient	Olken (2009)
CPI, WGI, ICRG, World Value Survey, Bribe Payers Index, ICVS	Testing convergent validity by correlation analysis	High correlations among CPI, WGI and ICRG by year	Ko and Samajdar (2010)
CPI and WGI	Grouping, comparing among sources of CPI and WGI	Distinguishable sources of data except the data generated by expert ratings by multilateral development banks and surveys of the mass public	Hawken and Munck (2011)
CPI, WGI, ICRG	Compare with ICVS and WBES by regression analysis	corruption experience is a weak determinant of all three corruption perception indexes	Donchev and Ujhelyi (2014)
Experts’ opinions about petty corruption in 8 African countries	Comparing with of population’s experience over bureaucratic corruption	No correlation; perception systematically overestimate corruption	Razafindrakoto and Roubaud (2010)
CPI	Global Corruption Barometer 2004 survey by regression analysis within the same samples	Personal or household experience of bribery is not a good predictor of perceptions	Abramo (2008)
1998 CPI	correlation between the sources of CPI	The sources do not differ considerably, indicating the overall reliability of CPI	Lambsdorff (1999)
CPI, Business International, ICRG	Comparing with an created aggregate index from ICVS by correlation analysis	High correlations	Mocan (2008)

Table A2: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N	Definition & source
CPI	56.14	22.05	0.6	92	2332	Rescaled Corruption Perceptions Index [0-100], Transparency International
WGI	2.64	1	0.04	4.43	2063	Rescaled Control of Corruption, 2.59-WGI, Worldwide Governance Indicators
BCI	47.73	13.2	14.52	69.01	2928	Bayesian Corruption Index, Standaert 2015
News	3.12	12.52	0	311.5	2332	Number of news, Author's collection from LexisNexis
News coverage per internet user	0.006	0.0484	0	1.3148	2332	Number of news per one thousand internet users
News coverage 2	0.06	0.39	0	10.9	2198	Number of news using an alternative search, Author's collection from LexisNexis
GDPPC	18251.52	19073.99	555.56	132937.67	2332	Constant GDP per capita, PPP, WDI
Import share	45.01	25.49	0.06	246.81	2332	Imports of goods and services (% of GDP), WDI
Democratization	19.04	12.31	0	47.72	2332	Index of Democratization, Vanhanen and Lundell, 2014
Democracy	0.49	0.28	0	0.93	2020	Deliberative democracy index, Varieties of Democracy (V-Dem) project
Press freedom	54.3	23.57	3	95	2332	Freedom of the press (scores), Freedom House
Broadband penetration	7.51	10.39	0	42.56	1791	Fixed (wired) broadband subscriptions (per 100 people) , WDI
Conventions	0.11	0.33	0	2	2332	Combination indicator of anti-corruption conventions
UNCAC	0.06	0.24	0	1	2332	1 indicating the country ratified the United Nations Convention against Corruption
OECD	0.02	0.13	0	1	2332	1 indicating the country ratified the OECD Anti-Bribery Convention

GRECO	0.01	0.12	0	1	2332	1 indicating the country joined the Group of States against Corruption (GRECO)
IACAC	0.01	0.07	0	1	2332	1 indicating the country ratified the Inter-American Convention against Corruption
AU Convention	0.01	0.11	0	1	2332	1 indicating the country ratified the African Union Convention on Preventing and Combating Corruption
Latitude	0.28	0.19	0.01	0.72	1934	The absolute value of the latitude of the capital city, divided by 90, the QoG dataset
Protestant	12.65	21.76	0	97.8	1934	Protestants as percentage of population in 1980, the QoG dataset
Colonial origin 1	0.01	0.09	0	1	2332	Dutch, the QoG dataset
Colonial origin 2	0.13	0.34	0	1	2332	Spanish, the QoG dataset
Colonial origin 3	0	0.06	0	1	2332	Italia, the QoG dataset
Colonial origin 4	0.01	0.09	0	1	2332	US, the QoG dataset
Colonial origin 5	0.25	0.44	0	1	2332	British, the QoG dataset
Colonial origin 6	0.12	0.33	0	1	2332	French, the QoG dataset
Legal origin 1	0.32	0.47	0	1	1934	English Common Law, the QoG dataset
Legal origin 2	0.51	0.5	0	1	1934	French Commercial Code, the QoG dataset
Legal origin 3	0.08	0.27	0	1	1934	Communist Laws, the QoG dataset

Table A3: Effect of anti-corruption news on the CPI using FE and Pooled OLS regressions

	(1)	(2)	(3)	(4)	(5)
Method	FE	FE	FE	FE	Pooled OLS
News coverage	0.0876** (0.035) [0.012]	0.0932*** (0.034) [0.006]	0.0886** (0.035) [0.011]	0.2112*** (0.035) [0.000]	0.0606 (0.038) [0.115]
<i>News coverage</i> ²				-0.1296*** (0.028) [0.000]	
Log(previous CPI)	0.7044***	0.6763***	0.7044***	0.7041***	0.9317***
Log(GDPPC)	-0.0863***	-0.1116***	-0.0867***	-0.0855***	-0.0117***
Import share	-0.0003*	-0.0003	-0.0003*	-0.0003*	-0.0002**
Democratization	0.0008*		0.0008	0.0008*	0.0000
Time trend	0.1904***	0.2872***	0.1896***	0.1897***	-0.0388***
<i>Time trend</i> ²	-0.0058***	-0.0095***	-0.0057***	-0.0057***	0.0013***
Democracy		0.0018			
Press freedom			0.0002		
Protestant					-0.0003
Latitude					-0.0319
LEGAL 2					0.0117**
LEGAL 3					0.0149
LEGAL 4					-0.0149
LEGAL 5					-0.0340
Year fix-effects	Y	Y	Y	Y	Y
Country fix-effects	Y	Y	Y	Y	
Region fix-effects					Y
Colonial origin dummies					Y
Observations	2,332	2,053	2,332	2,332	1,934
Number of countries	165	154	165	165	134
R-squared	0.5044	0.4554	0.5044	0.5046	0.9739

Driscoll-Kraay standard errors in parentheses, p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. Constant coefficient is not reported; *** p<0.01, ** p<0.05, * p<0.1

Table A4: Effect of anti-corruption conventions on the CPI

	(1)	(2)	(3)	(4)
Method	FE	FE	FE	FE
Lag1_Conventions	0.0069*	0.0097**		
	[0.096]	[0.036]		
Lag2_Conventions	0.0051*	0.0061**		
	[0.097]	[0.029]		
Lag1_UNCAC			0.0064**	0.0094**
			[0.021]	[0.015]
Lag1_OECD Convention			-0.0280*	-0.0207
			[0.057]	[0.189]
Lag1_GRECO			0.0430***	0.0395***
			[0.000]	[0.001]
Lag1_IACAC			0.0343***	0.0311***
			[0.002]	[0.009]
Lag1_AU Convention			-0.0022	-0.0009
			[0.545]	[0.839]
Log(previous CPI)	0.7102***	0.6787***	0.7049***	0.6767***
Log(GDPPC)	-0.0909***	-0.1203***	-0.0898***	-0.1145***
Import share	-0.0003**	-0.0003***	-0.0003	-0.0002
Democratization	0.0007		0.0008*	
Democracy		-0.0044		0.0021
Time trend	0.1868***	0.2602***	0.1882***	0.2903***
<i>Time trend</i> ²	-0.0056***	-0.0081***	-0.0056***	-0.0096***
Year fix-effects	Y	Y	Y	Y
Country fix-effects	Y	Y	Y	Y
Observations	2,314	2,036	2,352	2,073
Number of countries	165	154	154	165
R-squared	0.5057	0.4523	0.5064	0.4571

Driscoll-Kraay standard errors in parentheses, p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. Constant coefficient is not reported; *** p<0.01, ** p<0.05, * p<0.1

Table A5: Effect of anti-corruption news coverage on the CPI using IV regressions

	(1)	(2)	(3)	(4)	(5)
Method	IV	IV	IV	IV	IV
News coverage	8.0968** (3.294) [0.014]	6.5792** (3.030) [0.030]	7.7006** (3.151) [0.015]	8.1482** (3.295) [0.013]	7.5971** (3.179) [0.017]
Log(previous CPI)	0.7017***	0.6639***	0.6946***	0.7010***	0.7025***
Log(GDPPC)	-0.0631**	-0.1629***	-0.0903**	-0.0647**	-0.0661**
Import share	-0.0004	-0.0004	-0.0007	-0.0005	-0.0004
Democratization	0.0002			0.0001	0.0003
Democracy		0.0167			
Press freedom				0.0010*	
Time trend	-0.2933	-0.3400	-0.2042	-0.2817	-0.2943
<i>Time trend</i> ²	0.0071	0.0083	0.0049	0.0068	0.0071
Country, year dummies	Y	Y	Y	Y	Y
Observations	1,790	1,513	1,814	1,790	1,790
Number of countries	162	149	165	162	162
R-squared	0.4528	0.0977	0.1105	0.4515	0.4738
Relevance test (F-stat 1)	45.47	13.08	14.25	45.13	15.47
Relevance test (F-stat 2)	10.58	10.65	10.96	10.75	3.76
Over-identification test					0.1866
First-stage of IV regression - only excluded instruments are reported					
Lag1_Broadband	0.0003*** [0.000]	0.0004*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]	0.0003*** [0.000]
Lag1_Conventions					-0.0005 [0.319]
Lag2_Conventions					-0.0000 [0.973]
Observations	1,805	1,519	1,830	1,804	1,805
R-squared	0.4271	0.2733	0.2850	0.4285	0.4274

F stat 1: Cragg-Donald Wald F statistics; F stat 2: Kleibergen-Paap rk Wald F statistic. P-value of Hansen J statistic is reported for the over-identification test. Standard errors in parentheses; p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Estimated media bias on the CPI: a sensitivity analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	No	After	After	Exclude	Exclude	Exclude	Exclude	Exclude
	English	2003	2004	G20	BRICS	EU	single-party	military-party
Method	IV	IV	IV	IV	IV	IV	IV	IV
News coverage	6.5736*	6.7654**	5.3731**	8.8027**	9.4035**	10.2167**	10.3890*	10.0851*
	(3.712)	(2.935)	(2.588)	(3.642)	(3.722)	(4.769)	(5.455)	(5.300)
	[0.077]	[0.021]	[0.038]	[0.016]	[0.012]	[0.032]	[0.057]	[0.057]
Log(previous CPI)	0.6897***	0.7389***	0.7239***	0.6843***	0.6973***	0.7682***	0.6045***	0.6026***
Log(GDPPC)	-0.0759***	-0.0649**	-0.0665**	-0.0597	-0.0662	-0.0277	-0.0600	-0.0675
Import share	0.0001	-0.0003	-0.0004	-0.0002	-0.0003	0.0002	-0.0012	-0.0012
Democratization	0.0001	0.0000	0.0001	0.0003	0.0002	-0.0004	-0.0006	-0.0006
Time trend	-0.3057	-0.3307*	-0.3295*	0.6794**	0.6638**	1.0253***	-0.0753	-0.0936
<i>Time trend</i> ²	0.0075	0.0081*	0.0080*	-0.0178**	-0.0174**	-0.0267***	0.0024	0.0030
Country, year dummies	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,465	1,620	1,530	1,397	1,593	1,312	1,016	1,016
Number of countries	131	162	162	143	157	136	147	148
R-squared	0.5330	0.5069	0.5262	0.3583	0.3666	0.1722	0.0496	0.0766
Relevance test (F-stat 1)	36.36	41.07	37.71	37.75	40.77	28.66	16.00	16.74
Relevance test (F-stat 2)	7.73	10.85	8.39	10.24	10.25	9.17	7.27	7.41

F stat 1: Cragg-Donald Wald F statistics; F stat 2: Kleibergen-Paap rk Wald F statistic. P-value of Hansen J statistic is reported for the over-identification test. Standard errors in parentheses; p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Estimated media bias on perception based measures using another search strategy

	(1)	(2)	(3)	(4)
Method	FE	IV	IV	IV
News coverage	0.0178*** (0.005) [0.001]	0.7678** (0.353) [0.029]	0.6466* (0.335) [0.053]	0.7338** (0.335) [0.028]
<i>News coverage</i> ²	-0.0006 (0.001) [0.277]			
Log(previous CPI)	0.6967***	0.7050***	0.6733***	0.7036***
Log(GDPPC)	-0.0933***	-0.1072**	-0.1473***	-0.1158***
Import share	-0.0002	-0.0000	-0.0006	-0.0002
Democratization	0.0008	0.0006		
Time trend	0.2472***	0.5731	1.0033***	0.6064*
<i>Time trend</i> ²	-0.0085***	-0.0151	-0.0262***	-0.0159*
Country, year dummies	Y	Y	Y	Y
Observations	2,198	1,660	1,457	1,683
Number of countries	165	162	149	165
R-squared	0.4824	0.1732	0.3821	0.1802
Number of groups	165			
Relevance test (F-stat 1)		23.12	36.33	21.59
Relevance test (F-stat 2)		6.55	6.18	6.84

F stat 1: Cragg-Donald Wald F statistics; F stat 2: Kleibergen-Paap rk Wald F statistic. P-value of Hansen J statistic is reported for the over-identification test. Standard errors in parentheses; p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Estimated media bias on the CPI of samples with and without surveys

	(1)	(2)	(3)	(4)
Sample	Surveys	Surveys	No surveys	No surveys
Method	FE	IV	FE	IV
News coverage	0.6872** (0.297) [0.022]	-7.7079 (7.203) [0.285]	0.0766** (0.033) [0.022]	16.7434*** (6.315) [0.008]
Log(previous CPI)	0.4139***	0.4814***	0.7668***	0.7453***
Log(GDPPC)	-0.0812*	-0.2737***	-0.0720***	-0.0065
Import share	-0.0007*	-0.0027**	-0.0002	-0.0005
Democratization	0.0003	0.0011	0.0008	-0.0005
Time trend	0.4223***	0.0840	0.1342***	-0.7705**
<i>Time trend</i> ²	-0.0137***	-0.0023	-0.0041***	0.0190**
Year dummies	Y	Y	Y	Y
Country dummies	Y	Y	Y	Y
Observations	590	472	1,742	1,308
Number of countries	135	122	165	160
R-squared	0.5014	0.2315	0.5208	0.1056
Relevance test (F-stat 1)		4.63		23.91
Relevance test (F-stat 2)		1.86		10.65

F stat 1: Cragg-Donald Wald F statistics; F stat 2: Kleibergen-Paap rk Wald F statistic. P-value of Hansen J statistic is reported for the over-identification test. Standard errors in parentheses; p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Estimated media bias the WGI and BCI

	(1)	(2)	(3)	(4)	(5)	(6)
Method	FE	FE	IV	FE	FE	IV
Index	WGI	WGI	WGI	BCI	BCI	BCI
News coverage	0.0602*** (0.022) [0.006]		0.5189 (3.389) [0.878]	0.0006 (0.000) [0.179]		1.7357*** (0.663) [0.009]
Lag1_Conventions		-0.0035			-0.0013**	
Lag2_Conventions		0.0034			-0.0007	
Log(previous WGI)	0.5055***	0.5052***	0.4932***			
Log(previous BCI)				1.0637***	1.0641***	1.0336***
Log(GDPPC)	-0.1048***	-0.1039***	-0.1143***	-0.0028*	-0.0033**	0.0069
Import share	-0.0003*	-0.0003**	-0.0007	0.0000**	0.0000**	0.0001
Democratization	0.0007**	0.0006**	0.0011*	0.0000	0.0000	0.0001
Time trend	0.1572**	0.1559**	0.4352	-0.1345***	-0.1341***	-0.0789**
<i>Time trend</i> ²	-0.0051**	-0.0051**	-0.0109	0.0043***	0.0043***	0.0020**
Country, year dummies	Y	Y	Y	Y	Y	Y
Observations	2,063	2,091	1,789	2,928	2,871	1,768
Number of countries	166	166	163	165	165	162
R-squared	0.2879	0.2883	0.2834	0.9333	0.9322	0.7469
Relevance test (F-stat 1)			16.44			15.77
Relevance test (F-stat 2)			10.89			10.58

F stat 1: Cragg-Donald Wald F statistics; F stat 2: Kleibergen-Paap rk Wald F statistic. P-value of Hansen J statistic is reported for the over-identification test. Standard errors in parentheses; p-values in brackets; Statistics robust to heteroscedasticity and autocorrelation. *** p<0.01, ** p<0.05, * p<0.1.

References

- Abramo, C. W., 2008. How much do perceptions of corruption really tell us? *Economics: The Open-Access, Open-Assessment E-Journal* 2, 3.
- Adcock, R., 2001. Measurement validity: A shared standard for qualitative and quantitative research. In: *American Political Science Association*. Vol. 95. Cambridge Univ Press, pp. 529–546.
- Brunetti, A., Weder, B., 2003. A free press is bad news for corruption. *Journal of Public economics* 87 (7), 1801–1824.
- Carmines, E. G., Zeller, R. A., 1979. *Reliability and validity assessment*. Vol. 17. Sage publications.
- Christiane, A., Charles, O., 2006. *Development Centre Studies Uses and Abuses of Governance Indicators*. OECD Publishing.
- Czernich, N., Falck, O., Kretschmer, T., Woessmann, L., 2011. Broadband infrastructure and economic growth. *The Economic Journal* 121 (552), 505–532.
- Deacon, D., 2007. Yesterday's papers and today's technology digital newspaper archives and 'push button' content analysis. *European Journal of Communication* 22 (1), 5–25.
- Di Tella, R., Schargrotsky, E., 2003. The role of wages and auditing during a crackdown on corruption in the city of buenos aires. *Journal of Law and Economics* 46 (1), 269–292.
- Djankov, S., McLiesh, C., Nenova, T., Shleifer, A., 2001. *Who owns the media?* Tech. rep., National Bureau of Economic Research.
- Donchev, D., Ujhelyi, G., 2014. What do corruption indices measure? *Economics & Politics* 26 (2), 309–331.
- Driscoll, J. C., Kraay, A. C., 1998. Consistent covariance matrix estimation with spatially dependent panel data. *Review of economics and statistics* 80 (4), 549–560.

- Ferraz, C., Finan, F., 2011. Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review* 101, 1274–1311.
- Fisman, R., 2001. Estimating the value of political connections. *American Economic Review*, 1095–1102.
- Fisman, R., Svensson, J., 2007. Are corruption and taxation really harmful to growth? firm level evidence. *Journal of Development Economics* 83 (1), 63–75.
- Gainsborough, M., Dinh, D. N., Phuong, T. T., 2009. Corruption, public administration reform and development: Challenges and opportunities as viet nam moves towards middle-income status. *Public Administration Reform and Anti-corruption*, UNDP, 379.
- Hawken, A., Munck, G. L., 2011. Does the evaluator make a difference? measurement validity in corruption research. *IPSA Committee on Concepts and Methods Working Paper Series*.
- Hoechle, D., et al., 2007. Robust standard errors for panel regressions with cross-sectional dependence. *Stata Journal* 7 (3), 281.
- Hsieh, H.-F., Shannon, S. E., 2005. Three approaches to qualitative content analysis. *Qualitative health research* 15 (9), 1277–1288.
- Iyengar, S., Peters, M. D., Kinder, D. R., 1982. Experimental demonstrations of the “not-so-minimal” consequences of television news programs. *American Political Science Review* 76 (04), 848–858.
- Jensen, N. M., Li, Q., Rahman, A., 2010. Understanding corruption and firm responses in cross-national firm-level surveys. *Journal of International Business Studies* 41 (9), 1481–1504.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2006. Measuring governance using cross-country perceptions data. *International handbook on the economics of corruption* 52.
- Kaufmann, D., Kraay, A., Mastruzzi, M., 2011. The worldwide governance indicators: methodology and analytical issues. *Hague Journal on the Rule of Law* 3 (2), 220–246.

- Knack, S., 2007. Measuring corruption: A critique of indicators in eastern europe and central asia. *Journal of Public Policy* 27 (03), 255–291.
- Ko, K., Samajdar, A., 2010. Evaluation of international corruption indexes: Should we believe them or not? *The Social Science Journal* 47 (3), 508–540.
- Krippendorff, K., 2012. *Content analysis: An introduction to its methodology*. Sage.
- La Porta, R., Lopez-de Silanes, F., Shleifer, A., Vishny, R., 1999. The quality of government. *Journal of Law, Economics, and Organization* 15 (1), 222–279.
- Lambsdorff, J. G., 1999. The transparency international corruption perceptions index 1999: Framework document.
- Lambsdorff, J. G., 2007. The methodology of the corruption perceptions index 2007.
- Lindstedt, C., Naurin, D., 2010. Transparency is not enough: Making transparency effective in reducing corruption. *International political science review* 31 (3), 301–322.
- Macnamara, J. R., et al., 2005. Media content analysis: Its uses, benefits and best practice methodology. *Asia-Pacific Public Relations Journal* 6 (1), 1.
- Mocan, N., 2008. What determines corruption? international evidence from microdata. *Economic Inquiry* 46 (4), 493–510.
- Moynihan, R., Bero, L., Ross-Degnan, D., Henry, D., Lee, K., Watkins, J., Mah, C., Soumerai, S. B., 2000. Coverage by the news media of the benefits and risks of medications. *New England Journal of Medicine* 342 (22), 1645–1650.
- Neuendorf, K. A., 2002. *The content analysis guidebook*. Sage publications.
- Olken, B. A., 2006. Corruption and the costs of redistribution: Micro evidence from indonesia. *Journal of public economics* 90 (4), 853–870.
- Olken, B. A., 2007. Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy* 115 (2).
- Olken, B. A., 2009. Corruption perceptions vs. corruption reality. *Journal of Public economics* 93 (7), 950–964.

- Palau, A. M., Davesa, F., 2013. The impact of media coverage of corruption on spanish public opinion. *Revista Española de Investigaciones Sociológicas* 144 (1), 97–124.
- Palvia, S. C. J., Sharma, S. S., 2007. E-government and e-governance: definitions/domain framework and status around the world. In: *International Conference on E-governance*.
- Pande, R., Olken, B., 2012. Corruption in developing countries. *Annual Review of Economics* 4 (1), 479–509.
- Pavlik, J., 2000. The impact of technology on journalism. *Journalism Studies* 1 (2), 229–237.
- Razafindrakoto, M., Roubaud, F., 2010. Are international databases on corruption reliable? a comparison of expert opinion surveys and household surveys in sub-saharan africa. *World development* 38 (8), 1057–1069.
- Röller, L.-H., Waverman, L., 2001. Telecommunications infrastructure and economic development: A simultaneous approach. *American economic review*, 909–923.
- Sampford, C. J., Shacklock, A. H., Connors, C., Galtung, F., 2006. *Measuring corruption*. Ashgate.
- Schaffer, M. E., et al., 2015. xtvreg2: Stata module to perform extended iv/2sls, gmm and ac/hac, liml and k-class regression for panel data models. *Statistical Software Components*.
- Schmidt, D. E., 1993. Public opinion and media coverage of labor unions. *Journal of Labor Research* 14 (2), 151–164.
- Shoemaker, P. J., Eichholz, M., Kim, E., Wrigley, B., 2001. Individual and routine forces in gatekeeping. *Journalism & mass communication quarterly* 78 (2), 233–246.
- Standaert, S., 2015. Divining the level of corruption: a bayesian state-space approach. *Journal of Comparative Economics* 43 (3), 782–803.

- Steensen, S., 2009. The shaping of an online feature journalist. *Journalism* 10 (5), 702–718.
- Steensen, S., 2011. Online journalism and the promises of new technology: A critical review and look ahead. *Journalism Studies* 12 (3), 311–327.
- Svensson, J., 2005. Eight questions about corruption. *The Journal of Economic Perspectives* 19 (3), 19–42.
- Treisman, D., 2000. The causes of corruption: a cross-national study. *Journal of public Economics* 76 (3), 399–457.
- Weaver, D. A., Bimber, B., 2008. Finding news stories: a comparison of searches using lexisnexis and google news. *Journalism & Mass Communication Quarterly* 85 (3), 515–530.
- Worrall, J. L., Pratt, T. C., 2004. Estimation issues associated with time-series-cross-section analysis in criminology. *W. Criminology Rev.* 5, 35.
- Wunnava, P. V., Leiter, D. B., 2009. Determinants of intercountry internet diffusion rates. *American Journal of Economics and Sociology* 68 (2), 413–426.
- Zhang, Y., 2001. Scholarly use of internet-based electronic resources. *Journal of the American society for information science and technology* 52 (8), 628–654.