

Indicele de Percepție a Corupției 2017:

Notă Metodologică Tehnică

Context

Indicele de Percepție a Corupției (IPC) a fost înființat în anul 1995 ca un indicator compozit utilizat pentru a măsura percepția corupției în sectorul public din diferite țări din întreaga lume. În ultimii 20 de ani, atât sursele utilizate pentru elaborarea Indexului cât și metodologia au fost ajustate și rafinate. Cel mai recent proces de revizuire a avut loc în 2012¹, când au fost aduse unele schimbări importante în metodologia IPC, pentru a permite compararea punctajelor obținute de la un an la altul, ceea ce nu a fost posibil din punct de vedere metodologic înainte de 2012.

Metodologie

Metodologia urmează 4 etape de bază: selectarea surselor de date, rescalarea surselor de date, agregarea datelor rescalate și apoi raportarea unei măsuri de incertitudine. Procesul de calcul cuprinde, de asemenea, un mecanism strict de control al calității care constă în colectarea și calcularea datelor într-un proces paralel independent, efectuat de doi cercetători TI și doi consilieri universitari din afara Transparency International.

1. Selectarea surselor de date

IPC se bazează pe 13 surse de date care captează percepția experților și a executivilor din mediul de afaceri asupra unei serii de comportamente corupte din sectorul public, incluzând aici:

- Mita;
- delapidarea de fonduri publice;
- folosirea poziției publice pentru câștiguri private;
- nepotism în serviciul public;
- acapararea statului.

Unele dintre sursele utilizate verifică de asemenea mecanismele disponibile la nivel de țară pentru prevenirea corupției, precum:

- Abilitatea guvernului de a aplica mecanisme de integritate
- Urmărirea eficientă a funcționarilor corupți
- Povara birocratică excesivă
- Existența legislației adecvate cu privire la transparența financiară, prevenirea conflictelor de interese și accesul la informații de interes public.
- Protecția legală oferită avertizorilor de integritate, jurnaliștilor și anchetatorilor.

¹ Metodologia utilizată pentru a calcula IPC 2018 are la bază analiza de lucru a prof. Andrew Gelman, Departamentul de Statistică și Departamentul de Științe Politice, Universitatea Columbia, respectiv Dr. Piero Stanig, Profesor asociat Institute of Methodology, London School of Economics and Political Science. Lucrarea a fost prezentată Transparency International într-un raport care este disponibil la cerere, printr-un e-mail transmis la cpi@transparency.org.

Fiecare sursă folosită pentru a calcula Indicele de Percepție a Corupției este evaluată în raport cu criteriile enumerate mai jos.

- A) Fiabilitatea metodologiei și reputația instituțională:** Pentru ca o sursă să fie inclusă în IPC, este necesar să asigurăm calitatea abordării sale metodologice. Ca atare, fiecare sursă trebuie să provină de la o instituție de specialitate care își documentează în mod clar metodele de colectare a datelor și metoda de măsurare. Ulterior, Transparency International evaluează soliditatea metodologiei. Dacă, de exemplu, se oferă un „aviz de specialitate”, este necesar să se solicite o asigurare privind calificările expertului. În cazul sondajelor din mediul de afaceri, Transparency International verifică dacă eșantionul utilizat este reprezentativ.
- B) Alinierea conceptuală a datelor:** Având în vedere că este vorba de o măsurare a corupției în sectorul public, toate sursele de date folosite la alcătuirea IPC trebuie să fie legate de gradul de corupție din sectorul public. Întrebările se pot referi la un anumit „tip” de corupție (de exemplu, corupția mică) sau la eficacitatea mecanismelor de prevenire a corupției, care, de asemenea, pot fi folosite ca indicatori pentru perceperea nivelului de corupție dintr-o țară.
- C) Varietate cantitativă:** Scalele utilizate de către sursele de date trebuie să permită suficientă diferențiere în date (cel puțin o scală de patru puncte) cu privire la gradul de corupție perceput în cadrul fiecărei țări, astfel încât datele să poată fi redimensionate pe baza scalei de la 0 la 100 a IPC.
- D) Comparația între țări:** Având în vedere că IPC clasează țările în comparație una cu cealaltă, și datele direct de la sursă trebuie să fie comparabile între țări și să nu fie specifice fiecărei țări. În plus, sursa trebuie să măsoare același concept în toate țările, utilizând aceeași scală.
- E) Disponibilitatea datelor pe mai mulți ani:** Deoarece IPC are rolul să măsoare corupția în mai multe țări și perioade diferite, sunt excluse sursele care captează percepțiile asupra corupției doar într-un singur moment în timp.

Pentru a efectua acest proces de asigurare a calității, Transparency International se adresează fiecărei instituții care asigură date pentru a verifica metodologia de generare a scorurilor. Deoarece unele surse nu sunt disponibile public, Transparency International solicită, de asemenea, permisiunea de a publica scorurile redimensionate de la fiecare sursă împreună cu scorul IPC compozit. Cu toate acestea, Transparency International nu are permisiunea să facă publice scorurile originale ale surselor private.

2. Standardizarea surselor de date

Fiecare dintre surse incluse în IPC este standardizată pentru a permite agregarea în scorul IPC. Standardizarea transferă toate datele pe o scală de la 0 la 100 unde 0 reprezintă cel mai înalt nivel de corupție perceput și 100 cel mai scăzut nivel al corupției percepute. În timp ce majoritatea surselor IPC sunt, de asemenea, codificate în aceeași direcție (cu scoruri mici indicând niveluri înalte de corupție perceput), patru surse sunt scalate opus, cu scoruri mici indicând niveluri scăzute de corupție perceput. Pentru a putea fi comparate, aceste patru surse sunt inversate multiplicând fiecare scor cu -1.

Sursele care trebuie să fie inversate includ:

- Indicatorul de corupție al Economist Intelligence Unit
- Scorul corupției națiunilor în tranzit elaborat de Freedom House
- Scorul percepției corupției al Political and Economic Risk Consultancy Asian Intelligence
- Indexul corupției politice al Varieties of Democracy Project

Deoarece multe din sursele folosite pentru IPC nu au acoperire globală, valorile lipsă pentru aceste surse sunt imputate pentru anul de referință.² Acest proces se efectuează utilizând pachetul de software de statistică STATA și, mai exact, comanda de „impute” (atribuire) a programului. Comanda de „impute” (atribuire) estimează o valoare pentru fiecare punct de date lipsă, folosind doar acele surse de date cu cel puțin 50% din numărul total de țări acoperite de IPC anual³. Excepție fac datele din indicele de transformare al „Bertelsmann Foundation”, care nu sunt utilizate pentru imputarea indicatorilor de guvernare durabilă ai „Bertelsmann Foundation”.

După procesul de imputare, media și abaterea standard pentru fiecare sursă de date este calculată și folosită ca parametri pentru standardizarea datelor originale în scoruri z. Scorurile z standardizate sunt calculate prin scăderea mediei fiecărei surse din scorul fiecărei țări și împărțind-o cu abaterea standard a respectivei surse. Aceasta are ca rezultat un set de date centrat în jurul valorii de 0 și cu o abatere standard de 1. Scorurile z sunt calculate folosind parametrii de medie și abatere standard din sursele imputate pentru anul de referință. Acest lucru asigură că scorurile IPC sunt comparabile an de an față de anul de referință.

Este important de observat că setul complet de date cu valori imputate este utilizat doar pentru a genera parametrii globali de referință, iar valorile atribuite în sine nu sunt utilizate în agregarea finală care produce scorurile CPI.

Punctele z sunt apoi transformate pentru a se potrivi scării CPI de 0-100. Aceasta folosește o formulă de rescaling, care urmărește stabilirea valorii medii a setului de date standardizat la 45 și abaterea standard la 20. Următoarea formulă este utilizată în acest scop:

² Până în 2016, parametrii globali din 2012 au fost utilizați ca bază de referință. O dată cu introducerea datelor VDEM în IPC 2016, valorile lipsă pentru anul 2016 au fost imputate din nou. Cu toate acestea, o comparație a parametrilor globali din 2012 și 2016 s-a dovedit a fi nesemnificativă din punct de vedere statistic, ceea ce înseamnă că modificarea valorii inițiale nu a afectat comparabilitatea în timp. În 2017, VDEM și-a extins gradul de acoperire la peste 50% din țările IPC. Din acest motiv, parametrii globali au trebuit să fie re-calculați din nou. Din nou s-a observat că modificările din media și abaterea standard nu au fost semnificative din punct de vedere statistic. Ca urmare, anul de referință rămâne anul 2012.

³ În conformitate cu regula gradului de acoperire a 50% din țări menționată mai sus, s-au utilizat următoarele șapte surse de date pentru CPI 2017:

- Bertelsmann Foundation’s Transformation Index 2018 (62.32 %)
- Economist Intelligence Unit Country Risk Ratings 2017 (63.29 %)
- Global Insights Country Risk Ratings (98.55 %)
- Political Risk Services International Country Risk Guide (67.63 %)
- World Economic Forum Executive Opinion Survey 2017 (64.42 %)
- World Justice Project rule of Law Index 2017-2018 (54.59 %)
- Varieties of Democracy Project 2017 (83.58 %)

$$\text{Rescaled Indicator } X_{\text{Country}} = \frac{\text{Original Indicator } X_{\text{Country}} - \text{mean imputed Indicator } X^{t=\text{baseline}}}{\text{standard deviation imputed Indicator } X^{t=\text{baseline}}} * 20 + 45$$

Orice scoruri revizuite care iau valori mai mici de 0 sunt egale cu 0, iar scorurile redimensionate care depășesc 100 sunt limitate la 100.

3. Agregarea datelor redimensionate

Punctajul IPC al fiecărei țări se calculează ca o medie simplă a tuturor punctajelor rescalate disponibile pentru acea țară (De menționat că noi nu utilizăm niciuna dintre valorile imputate ca punctaj pentru IPC agregat). Unei țări i se va acorda un punctaj doar în cazul în care există cel puțin trei surse de date disponibile din care să se calculeze această medie.

4. Raportarea unei măsuri de incertitudine

Punctajul IPC va fi raportat alături de o marjă standard de eroare și de o marjă de încredere de 90 la sută, care reflectă variația în valorile datelor sursă care alcătuiesc punctajul IPC. Marjă standard de eroare se calculează ca abaterea standard a datelor sursă redimensionate, împărțit la rădăcina pătrată a numărului de surse. Folosind această eroare standard se poate calcula intervalul de încredere de 90 la sută, presupunând o distribuție normală.

În urma recomandării de audit efectuată de Centrul Comun de Cercetare al Comisiei Europene asupra coerenței conceptuale și statistice a IPC, formula de calcul al erorilor standard a fost modificată începând de la CPI2018. Formula anterioară a calculat termenii de eroare standard ca deviația standard a datelor sursă modificate, împărțită la rădăcina pătrată a numărului de surse.

Noua formulă exprimată în ecuația de mai jos ia în considerare un număr mic de surse:

$$\Sigma = \sqrt{\frac{N - n \sigma}{N - 1 \sqrt{n}}}$$

unde Σ este termenul de eroare standard, N este numărul total de surse utilizate pentru calculul IPC, n este numărul de surse pentru scorul CPI din orice țară dată, și σ este deviația standard a scorului IPC pentru acea țară dată.⁴ Folosind această eroare standard, putem calcula intervalul de încredere de 90% și raportăm limitele superioare și inferioare ale scorului CPI pentru fiecare țară, presupunând o distribuție normală.

După calcularea erorilor standard, stabilim dacă schimbarea punctajului CPI pentru fiecare țară este semnificativă din punct de vedere statistic sau nu. În acest scop, calculăm mai întâi dimensiunea efectului diferenței dintre scorul unei țări în cei doi ani comparați folosind formula de mai jos:

$$\text{effect size} = \frac{M_1 - M_2}{\sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2}{N_1 + N_2 - 2}}}$$

unde M_1 este scorul CPI pentru o anumită țară într-un an dat, M_2 este scorul CPI pentru țara respectivă în anul de comparație, N_1 este numărul de surse din anul respectiv, N_2 este numărul de surse din anul de comparație, $SD_1/2$ este pătratul abaterii standard a scorului IPC în anul dat și $SD_2/2$ este pătratul deviației standard a scorului IPC în anul de comparație.

⁴ În 2018, suma totală de surse (N) a fost de 13.

După obținerea dimensiunii efectelor, vom calcula deviația standard (σ) a mărimii efectului utilizând formula de mai jos:

$$\sigma(d) = \sqrt{\frac{N1 + N2}{N1 \times N2} + \frac{d^2}{2 \times (N1 + N2)}}$$

unde d este dimensiunea efectului diferențelor în scorurile IPC, iar $N1$ și $N2$ reprezintă numărul de surse disponibile pentru fiecare țară. Apoi, intervalul de încredere de 90% se calculează utilizând formula de mai jos:

$$CI_{0,90} = d \pm 1.96 \times \sigma(d)$$

Dacă acest interval de încredere include valoarea zero, atunci nu există diferențe semnificative statistic între țări. Pe de altă parte, dacă zero este în afara intervalului, atunci diferența este "semnificativă statistic la nivelul de 10%".