

Indicele de percepție a Corupției 2016

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, în același an fiind aduse unele schimbări importante în metodologia IPC. Metoda folosită până în 2012 pentru a agrega diferite surse de date a fost simplificată, incluzând acum doar datele pe un singur an de la fiecare sursă de date. În mod vital, această metodă ne permite să comparăm punctajele 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.

1. Selectarea surselor de date

IPC se bazează pe o serie de surse disponibile care captează percepții asupra corupției. Fiecare sursă este evaluată în raport cu criteriile enumerate mai jos. Fiecare instituție care oferă date a fost contactată pentru a verifica metodologia utilizată pentru generarea punctajelor și pentru a obține permisiunea de a publica punctajele rescalate din fiecare sursă, alături de punctajul compozit al Indexului.

A) Colectarea de date și metodologie de încredere provenite de la o instituție credibilă:

Este necesar să avem încredere în validitatea datelor pe care le folosim. Ca atare, fiecare sursă trebuie să provină de la o instituție de specialitate care își documentează în mod clar metodele de colectare a datelor. Aceste metode trebuie să fie sigure din punct de vedere al metodologiei. De exemplu, în cazul în care se oferă un "aviz de specialitate", căutăm o asigurare privind calificările expertului sau, în cazul în care se desfășoară un studiu de afaceri, ne asigurăm că eșantionul sondajului este reprezentativ.

B) Datele au în vedere corupția din sectorul public: Întrebarea sau analiza trebuie să se refere la percepția asupra nivelului de corupție în mod explicit în sectorul public. Întrebarea se poate referi la un "tip" definit de corupție (de exemplu, corupție mică) și, după caz, la eficacitatea măsurilor de prevenire a corupției, informațiile obținute putând fi folosite pentru stabilirea nivelului de corupție perceput la nivel național.

¹ Metodologia utilizată pentru a calcula IPC 2016 se bazează pe activitatea de examinare a abordărilor alternative pentru construirea IPC efectuate de Prof. Andrew Gelman: Profesor, Departamentul de Statistică și Departamentul de Științe Politice, Universitatea Columbia, și Dr. Piero Stanig: Asociat, Institutul de Metodologie, London School of Economics and Political Science. Acest lucru a fost prezentat Transparency International într-un raport disponibil la cerere. Persoană de contact: Santhosh Srinivasan, e-mail: ssrinivasan@transparency.org.

- C) Varietate cantitativă:** Scalele utilizate de către sursele de date trebuie să permită suficientă diferențiere în date (cel puțin o scală de la zero la patru) cu privire la nivelele de corupție percepute în cadrul fiecărei țări, astfel încât datele să poată fi rescalate pe scala de la 0 la 100 a IPC.
- D) Comparabilitate între țări:** Din moment ce IPC clasează țările în comparație una cu cealaltă, și datele direct de la sursă trebuie să fie comparabile între țări și nu specifice fiecărei țări. Sursa trebuie să măsoare același lucru în fiecare țară căreia îi acordă un punctaj, utilizând aceeași scală.
- E) Set de date multianual:** Vrem să putem compara punctajul unei țări și Indicele în general de la un an la altul. De aceea, sunt excluse sursele care captează percepțiile asupra corupției doar într-un singur moment în timp și nu sunt proiectate să fie repetate în timp.

2. Standardizarea surselor de date

Fiecare sursă este apoi standardizată pentru a fi compatibilă cu alte surse disponibile, pentru aderarea la scala IPC. Standardizarea convertește toate sursele de date la o scală de la 0 la 100 în care 0 = cel mai înalt nivel de corupție perceput și 100 = cel mai scăzut nivel al corupției percepute

Orice sursă care este scalată astfel încât punctajele mici să reprezinte un nivel scăzut de corupție trebuie să fie mai întâi inversată. Acest lucru se face prin înmulțirea fiecărui punctaj din setul de date cu -1.

Fiecare punctaj este apoi standardizat (la un punctaj z), prin scăderea mediei datelor și împărțirea la abaterea standard. Acest lucru duce la un set de date centrat în jurul valorii 0 și cu o abatere standard de 1.

Pentru ca aceste punctaje z să fie comparabile între seturile de date, trebuie să definim media și parametrii de abaterestandard ca parametri la nivel global. Prin urmare, acolo unde un set de date acoperă o gamă limitată de țări, atribuim punctaje pentru toate țările care lipsesc din respectivul set de date. Valorile lipsă pentru țările lipsă din fiecare sunt atribuitecu ajutorul pachetului de programe statistice STATA și, mai specific, comanda de input a programului. Această comandă regresează fiecare set de date la sursele de date IPC care sunt complete cel puțin în proporție de 50%, pentru a estima valorile pentru fiecare țară căreia îi lipsesc date în fiecare set de date în parte. De la aceste reguli fac excepție datele Indicelui de Transformare al Fundației Bertelsmann, care nu sunt utilizate pentru input-ul Indicatorilor de Guvernare Sustenabilă ai Fundației Bertelsmann, deoarece nu există nicio suprapunere în cazul celor două surse de date pentru țările monitorizate. Media și abaterea standard pentru setul de date sunt calculate ca o medie a seturilor de date complete și sunt utilizate ca parametri pentru standardizarea datelor brute. Foarte important, setul de date complet cu valorile introduse este folosit doar pentru a genera acești parametri, iar valorile introduse nu sunt utilizate ca date pentru punctajele IPC ale țărilor.

În mod critic, punctajele z sunt calculate folosind parametri de medie și de abatere standard de la punctajele introduse în 2012. Regula este stabilită astfel pentru ca 2012 să fie efectiv anul de referință pentru datele de calcul, iar punctajele rescalate să poată fi comparabile de la un an la altul. Când sunt introduse noi surse în Indice, pentru a reflecta în mod adecvat schimbările în timp, calculul de rescalare permite ca acestea să fie în concordanță cu parametri de bază 2012. În prima etapă se estimează dacă a existat o schimbare la nivel global în media și abaterea standard din 2012, iar apoi se utilizează noile valori, care ar putea să fi deviat de la 50 și 20 pentru a rescala noul set de date².

Punctajele z sunt apoi rescalate pentru a se potrivi scalei IPC între 0-100. Se utilizează o formulă simplă de rescalare, care stabilește valoarea medie a setului de date standardizate la aproximativ 45, iar deviația standard la aproximativ 20. Orice punctaj care depășește limitele 0-100 va fi plafonat.

² Din moment ce o nouă sursă de date s-a adăugat la IPC, procedura de mai sus a fost utilizată pentru a verifica dacă a existat o schimbare în medie și abaterea standard, începând cu 2012. Am stabilit că media și abaterea standard nu s-au schimbat, menținând în acest fel comparația de la an la an din scorurile IPC

3. Agregarea datelor rescalate

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 ăm niciuna dintre valorile introduse 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%, care reflectă variația în valorile datelor sursă care alcătuiesc punctajul IPC.

Marjă standard de eroare se calculează ca abaterea standard a datelor sursă rescalate, împărțit la rădăcina pătrată a numărului de surse. Folosind această eroare standard se poate calcula intervalul de încredere de 90%, presupunând o distribuție normală.