

# Economic Complexity Index

Studente: Michele Abruzzese

Matricola: 0001097676

Studente: Elia Cannas

Matricola: 0001097520

Studente: Federico Augelli

Matricola: 0001097518

Implementazione dell'indice di complessità economica  
tramite l'utilizzo di codice python

on 5 gennaio 2024

**Prof:**

Prof. Armando Bazzani

Alma Mater Studiorum  
Bologna

# Abstract

Per Adam Smith, il segreto della ricchezza delle nazioni era legato alla divisione del lavoro. Man mano che le persone e le imprese si specializzano in attività diverse, l'efficienza economica aumenta. Questa divisione del lavoro, tuttavia, è limitata dall'estensione del mercato: più grande è il mercato, più i suoi partecipanti possono specializzarsi e più profonda è la divisione del lavoro che può essere raggiunta. Ciò suggerisce che la ricchezza e lo sviluppo sono legati alla complessità che emerge dalle interazioni tra il numero crescente di attività individuali che compongono un'economia.[3] La complessità economica può aiutarci a comprendere lo sviluppo delle economie nazionali e regionali. I suoi metodi utilizzano dati sulle attività presenti nei luoghi per prevedere le dinamiche di sviluppo di Paesi, città e regioni. Questi metodi possono prevedere le attività economiche che un Paese, una città o una regione intraprenderanno o abbandoneranno in futuro, così come il livello previsto di reddito di un'economia, la crescita economica, la disuguaglianza di reddito e le emissioni. Come campo accademico, la Complessità Economica studia la geografia e la dinamica delle attività economiche utilizzando metodi ispirati alle idee dei sistemi complessi, delle reti e dell'informatica. Ciò che rende il campo della Complessità economica in qualche modo unico è che studia la geografia delle attività utilizzando un approccio basato sui risultati. In altre parole, invece di cercare di capire quali capacità o fattori guidano un'economia, utilizza i dati sulla geografia delle attività economiche per dedurre la presenza di insiemi di capacità. Nel contesto di un mondo sempre più interconnesso e globalizzato, la comprensione della complessità economica riveste un ruolo cruciale per gli economisti, gli uomini d'affari e i responsabili delle politiche pubbliche. Al di fuori del mondo accademico, i metodi della complessità economica sono utilizzati da organizzazioni multilaterali e agenzie di sviluppo nazionali focalizzate sullo sviluppo economico, da agenzie di promozione delle esportazioni e compagnie di navigazione che desiderano prevedere l'evoluzione dei modelli commerciali internazionali e da società finanziarie (per valutare obbligazioni e altri strumenti, prevedendo la crescita economica a lungo termine).

L'Economic Complexity Index (ECI), sviluppato da Cesar A. Hidalgo e Ricardo Hausmann presso il MIT Media Lab, si è rivelato uno strumento fondamentale per analizzare e valutare la diversità e la sofisticatezza delle economie nazionali. Nel corso di questo report, esploreremo il processo di traduzione delle complesse formule dell'ECI in un codice Python, permettendo una più ampia accessibilità e utilità pratica di questo indice. Attraverso l'analisi dettagliata delle equazioni originali di Hidalgo e Hausmann, abbiamo sviluppato un insieme di funzioni Python che consentono il calcolo dell'ECI per una data nazione. Questa trasposizione da teoria a pratica non solo facilita l'implementazione dell'indice, ma offre anche la possibilità di eseguire analisi approfondite sulla struttura economica di vari paesi. Nel seguito di questo report, illustreremo passo dopo passo come le formule dell'ECI sono state adattate in codice Python, evidenziando le scelte metodologiche e gli accorgimenti necessari per garantire una corretta implementazione. Per fornire un contesto pratico, presenteremo un esempio dettagliato di calcolo dell'ECI accompagnato da un ranking dei paesi in base ai loro risultati.[2]

# Indice

<b>Abstract</b>	<b>i</b>
<b>Tabella</b>	<b>ii</b>
<b>1 Introduzione</b>	<b>1</b>
1.1 Obiettivi . . . . .	1
<b>2 Metodologia</b>	<b>2</b>
<b>3 Codice e Risultati</b>	<b>5</b>
3.0.1 RCA . . . . .	5
3.0.2 ECI . . . . .	7
3.0.3 Risultati . . . . .	9
<b>4 Conclusioni</b>	<b>10</b>
<b>Bibliografia</b>	<b>11</b>

# Capitolo 1

## Introduzione

### 1.1 Obiettivi

L'obiettivo del nostro lavoro è quello di riprodurre la classifica ECI (Economic Complexity Index) utilizzando i dati delle esportazioni comprese tra l'anno 2016 e il 2018, classificandole secondo il Sistema Armonizzato (HS92) con una profondità di 4 cifre. Utilizzando per lo studio solo i paesi con una popolazione pari o superiore a 1 milione di abitanti, per un numero di esportazioni di almeno 1 miliardo e per un valore dei prodotti complessivo di almeno 500 milioni di dollari.

## Capitolo 2

# Metodologia

L'Indice di Complessità Economica (ECI) è una misura olistica delle capacità produttive di grandi sistemi economici, solitamente città, regioni o Paesi. In particolare, l'ECI cerca di spiegare la conoscenza accumulata in una popolazione che si esprime nelle attività economiche presenti in una città, un Paese o una regione. Per raggiungere questo obiettivo, l'ECI definisce la conoscenza disponibile in un luogo, come la conoscenza media delle attività presenti in esso, e la conoscenza di un'attività come la conoscenza media dei luoghi in cui viene svolta quell'attività economica. L'equivalente del prodotto dell'Indice di complessità economica è l'Indice di complessità del prodotto o PCI.

Il problema della stima della complessità economica è che bisogna stimare sia la complessità dei luoghi (ad esempio paesi, città, regioni) sia quella delle attività in essi presenti (ad esempio prodotti, industrie, tecnologie). L'idea generale è che le attività presenti, prodotte o esportate da un luogo trasportano informazioni sulla complessità di quel luogo, mentre i luoghi in cui è presente un'attività trasportano informazioni sulla complessità richiesta per svolgere un'attività. Ad esempio, possiamo dire che città come San Francisco, Boston e New York sono complesse perché ospitano attività complesse. Allo stesso modo, possiamo dire che un'attività come quella biotecnologica o aerospaziale è complessa, se si trova prevalentemente in economie complesse, come quelle di Boston e San Francisco. Questo argomento circolare può essere tradotto in un insieme generale di equazioni che possono essere utilizzate per stimare la complessità delle economie.

Formalmente, consideriamo la complessità  $K$  di una posizione  $c$  (ad esempio paese o città) essere  $K_c$  e la complessità  $K$  di un'attività  $p$  (ad esempio prodotto o settore) essere  $K_p$ . Inoltre, lasciamo  $M_{cp}$  sia una matrice riassuntiva delle attività ( $p$ ) presenti nel luogo ( $c$ ). Di solito  $M_{cp}$  è definito come  $M_{cp} = 1$  quando la produzione di una località in un'attività è maggiore di quanto previsto per una località della stessa dimensione e un'attività con la stessa produzione totale. Questo può essere fatto utilizzando un indicatore come il vantaggio comparativo rivelato ( $RCA$ ).

Definiamo:

$$M_{cp} = 1 \quad \text{if} \quad RCA_{cp} \geq 1$$

dove:

$$RCA_{cp} = \frac{E_{cp}/\sum_{p' \in P} E_{cp'}}{\sum_{c' \in C} E_{c'p}/\sum_{c' \in C, p' \in P} E_{c'p'}}$$

E	Esportazioni
c, c'	Indice del paese
C	Insieme di paesi
p, p'	Indice delle materie prime
P	Insieme di merci

Figura 2.1: Elementi di RCA

Un **vantaggio** comparativo può essere rilevato se il valore del RCA è superiore a 1, in caso contrario parliamo di **svantaggio** comparativo.

Un **esempio pratico** lo si può fare per quanto avvenuto nel 2010, in cui la soia rappresentava lo 0,35% del commercio mondiale con esportazioni di 42 miliardi di dollari. Di questo totale, il Brasile ha esportato quasi 11 miliardi di dollari, e poiché le esportazioni totali del Brasile per quell'anno erano di 140 miliardi di dollari, la soia rappresentava il 7,9% delle esportazioni del Brasile. Dato che  $7,9/0,35 = 22$ , il Brasile esporta 22 volte la sua "giusta quota" delle esportazioni di soia, e quindi possiamo dire che il Brasile ha un alto vantaggio comparativo rivelato nella soia.<sup>[1]</sup>

Seguendo la notazione utilizzata sopra, l'ipotesi generale fatta dagli indicatori di complessità economica è che:

- La complessità di una posizione ( $K_c$ ) è una funzione ( $f$ ) della complessità ( $K_p$ ) delle attività in esso presenti ( $M_{cp}$ )
- La complessità di un'attività ( $K_p$ ) è una funzione ( $g$ ) della complessità ( $K_c$ ) dei luoghi in cui tale attività è presente ( $M_{cp}$ ).

Questa logica circolare è equivalente alla seguente mappatura:

$$\begin{aligned} K_c &= f(M_{cp}, K_p) \\ K_p &= g(M_{cp}, K_c) \end{aligned}$$

Queste mappature implicano che le misure della complessità delle economie, o delle attività economiche, sono soluzioni di equazioni autoconsistenti della forma:

$$\begin{aligned} K_c &= f(M_{cp}, g(M_{cp}, K_c)) \\ K_p &= g(M_{cp}, f(M_{cp}, K_p)) \end{aligned}$$

può essere ridotto, o approssimato, da un'equazione lineare della forma:

$$\begin{aligned} K_c &= \widetilde{M}_{cc'} K_c \\ K_p &= \widetilde{M}_{pp'} K_p \end{aligned}$$

Queste equazioni implicano che le metriche della complessità delle economie, o delle attività in esse presenti, sono rispettivamente autovettori di matrici che collegano paesi correlati ( $M_{cc'}$ ) o prodotti correlati ( $M_{pp'}$ ).

Utilizzando il quadro di cui sopra, definiamo l'Indice di Complessità Economica di un luogo, o **ECI**, come la media dell'Indice di Complessità del Prodotto, o PCI, delle attività in esso presenti. Allo stesso modo, definiamo l'Indice di Complessità del Prodotto di un'attività, o **PCI**, come l'Indice di Complessità Economica, o ECI, medio dei luoghi in cui è presente quell'attività. Definiamo cioè la complessità di un luogo come la complessità media delle sue attività, e la complessità di un'attività, come la complessità media dei luoghi in cui quell'attività è presente. Formalmente, la formula ECI è la soluzione del sistema di equazioni:

$$\begin{aligned} K_c &= \frac{1}{M_c} \sum_p M_{cp} k_p \\ K_p &= \frac{1}{M_p} \sum_c M_{cp} k_c \end{aligned}$$

Il che, inserendo la seconda equazione nella prima, equivale a diagonalizzare la seguente matrice:

$$\widetilde{M}_{cc'} = \sum_p \frac{M_{cp} M_{c'p}}{M_c M_p}$$

Qui  $M_c = \sum_p M_{cp}$  è il numero di attività (o diversità) di un luogo e  $M_p = \sum_c M_{cp}$  è l'ubiquità di un'attività (numero di luoghi in cui è presente). Poiché la complessità economica è una metrica relativa, i risultati vengono solitamente normalizzati utilizzando una trasformata  $Z$ . Cioè:

$$\begin{aligned} ECI &= \frac{K_c - \bar{K}_c}{\sigma(K_c)} \\ PCI &= \frac{K_p - \bar{K}_p}{\sigma(K_p)} \end{aligned}$$

Dove  $\bar{K}_c$  è la media di  $K_c$  e  $\sigma(K_c)$  è la deviazione standard di  $K_c$ . [4]

## Capitolo 3

# Codice e Risultati

La preparazione dei dati utili è stato il primo passo che abbiamo compiuto per poter effettuare lo studio, effettuando dunque un filtraggio sul datasets per i valori indicati nel capitolo precedente.

Dopo aver estrapolato i dati utili sono stati formattati per la creazione della matrice RCA.

### 3.0.1 RCA

L'indice di Vantaggio Comparativo Rivelato (RCA) è stato introdotto da Balassa nel 1965 e viene utilizzato per valutare i principali prodotti esportati da un paese e i loro vantaggi comparativi rispetto al livello delle esportazioni mondiali. L'indice viene utilizzato in seguito per il calcolo del economic complexity come discusso da Hidalgo et al. nel 2007. [3]

Il calcolo dell'RCA avviene tramite la seguente formula:

$$RCA_{cp} = \frac{E_{cp} / \sum_{p' \in P} E_{cp'}}{\sum_{c' \in C} E_{c'p} / \sum_{c' \in C, p' \in P} E_{c'p'}}$$

In cui:

- $E$  rappresentano le esportazioni
- $c$  e  $c'$  gli indici dei paesi
- $C$  è il set di paesi
- $p$  e  $p'$  gli indici dei prodotti
- $P$  il set di prodotti

In altri termini l' $RCA$  può essere espressa come la percentuale delle esportazioni di un paese che sono nella classe in esame, diviso la percentuale di esportazioni mondiali che sono di quella classe in esame.

A livello di codice implementato è stata passata in input alla funzione `rca` la tabella pivot contenente al suo interno: *CountryID*, *HS4ID* e *TradeValue*. Calcoliamo la **quota delle esportazioni**



per categoria dividendo il **valore** delle esportazioni per categoria con il **totale** delle esportazioni del paese. In seguito è stata calcolata la **quota delle esportazioni mondiali per categoria**, dividendo il **totale** delle esportazioni mondiali per categoria con il **totale** delle esportazioni mondiali.

A questo punto l'*RCA* è stata calcolata facendo la divisione tra questi due valori.

$$RCA = \frac{\text{Quota delle esportazioni per categoria}}{\text{Quota delle esportazioni Mondiali per categoria}}$$

Il codice utilizzato per il calcolo è il seguente:

```

1  #fill missing values with zeros
2  tb1 = tb1.fillna(value=0)
3
4  col_sums = tb1.sum(axis=1) # 1:columns
5  col_sums = col_sums.to_numpy().reshape((len(col_sums), 1))
6
7  rca:numerator = np.divide(tb1, col_sums)
8  row_sums = tb1.sum(axis=0) # 0:index
9
10 total_sums = tb1.sum().sum()
11 rca_denominator = row_sums / total_sum
12 rcas = rca_numerator / rca_denominator
13
14 return rcas

```

HS4 ID	10101	10102	10103	10104	10105	\
Country						
Afghanistan	98.0	725494.0	0.0	3317969.0	0.0	
Albania	965.0	18655.0	0.0	40350.0	87581.0	
Algeria	359367.0	0.0	0.0	0.0	263.0	
Angola	615.0	23491.0	593.0	3264.0	9879.0	
Argentina	103130398.0	2823338.0	428327.0	100035.0	3564490.0	
...	...	...	...	...	...	
Venezuela	473153.0	80672.0	0.0	0.0	1530.0	
Vietnam	1239500.0	9621365.0	135860396.0	0.0	2436708.0	
Yemen	4804.0	0.0	0.0	519121.0	0.0	
Zambia	52153.0	167489.0	102844.0	0.0	12152582.0	
Zimbabwe	207552.0	95098.0	11813.0	2011.0	2586678.0	

Figura 3.1: Elementi di RCA

per calcolare il numeratore e il denominatore della formula RCA procediamo nel seguente modo:

- calcoliamo il **totale delle esportazioni per paese** tramite la somma delle colonne della tabella *tbl* nella figura precedente
- a *rca\_numerator* gli assegniamo la divisione di ogni elemento della tabella per il valore delle esportazioni per paese (*col\_sums*)
- calcoliamo il **totale delle esportazioni mondiali per categoria** tramite la somma delle righe della tabella *tbl*

- calcoliamo il totale delle esportazioni mondiali per tutte le categorie tramite la somma di tutti gli elementi della tabella e lo assegniamo a *total<sub>sum</sub>*
- a *rca<sub>denominator</sub>* gli assegniamo la divisione del totale delle esportazioni per categoria (*row<sub>sums</sub>*) per il totale delle esportazioni mondiali di tutte le categorie (*total<sub>sum</sub>*)

HS4 ID	10101	10102	10103	10104	10105	10106	\
Country							
Afghanistan	0.000164	0.369302	0.000000	9.026971	0.000000	9.139085	
Albania	0.000804	0.004732	0.000000	0.054699	0.063779	5.388141	
Algeria	0.020650	0.000000	0.000000	0.000000	0.000013	0.000156	
Angola	0.000034	0.000390	0.000021	0.000289	0.000471	0.078006	
Argentina	3.600635	0.030018	0.009875	0.005685	0.108810	0.238170	
...	...	...	...	...	...	...	
Venezuela	0.034015	0.001766	0.000000	0.000000	0.000096	0.098860	
Vietnam	0.011456	0.027080	0.829185	0.000000	0.019691	0.596998	
Yemen	0.006406	0.000000	0.000000	1.126684	0.000000	0.000000	
Zambia	0.010761	0.010524	0.014013	0.000000	2.192407	0.391491	
Zimbabwe	0.117857	0.016445	0.004430	0.001859	1.284257	12.021238	

Figura 3.2: Elementi di RCA

### 3.0.2 ECI

La funzione all'interno del codice che permette di calcolare l'ECI prende in input la matrice RCA, binarizzata in questo contesto, dove i valori maggiori o uguali a 1 diventano 1 e i valori inferiori a 1 diventano 0. Successivamente vengono calcolate la somma delle colonne e delle righe della matrice per ottenere i valori  $K_p$  (Complessità del prodotto, sommando le colonne) il numero di paesi dove il prodotto p viene esportato e  $K_c$  (Complessità geografica, sommando le righe) il numero di prodotti che il paese c esporta.

$$K_c = \frac{1}{M_c} \sum_p M_{cp} K_p,$$

$$K_p = \frac{1}{M_p} \sum_c M_{cp} K_c,$$

Nel for presente nel codice andiamo a risolvere il sistema di equazioni di cui sopra. Ad ogni iterazione aggiorniamo i valori  $k_p$  e  $k_c$  in modo tale da raffinare i valori facendoli convergere a un valore stabile. Abbiamo utilizzato il valore 20 per le iterazioni basandoci sugli esperimenti condotti dall'osservatorio OEC.

In seguito viene effettuata una standardizzazione dei valori ottenuti  $K_p$  e  $K_c$  sottraendo la media e dividendo per la loro deviazione standard.

```

1  # Binarize input RCA
2  rcas = pd.DataFrame(rca.copy())
3  rcas[rcas >= 1] = 1
4  rcas[rcas <= 1] = 0
5
6  # drop columns / rows only if completely nan
7  rcas_clone = rcas.copy(deep=True)
8  rcas_clone = rcas_clone.dropna(how="all")
9  rcas_clone = rcas_clone.dropna(how="all", axis=1)
10
11 if rcas_clone.shape != rcas.shape:
12     logger.warning(
13         "RCAs contain columns or rows that are entirely comprised of
14         NaN values."
15     )
16 if drop:
17     rcas = rcas_clone
18
19 kp = rcas.sum(axis=0) # sum columns
20 kc = rcas.sum(axis=1) # sum rows
21 kp0 = kp.copy()
22 kc0 = kc.copy()
23
24 for i in range(1, iterations):
25     kc_temp = kc.copy()
26     kp_temp = kp.copy()
27     kp = rcas.T.dot(kc_temp) / kp0
28     if i < (iterations - 1):
29         kc = rcas.dot(kp_temp) / kc0
30
31 geo_complexity = (kc - kc.mean()) / kc.std()
32 prod_complexity = (kp - kp.mean()) / kp.std()
33
34 return geo_complexity, prod_complexity

```

### 3.0.3 Risultati

Tramite il codice riportato nei paragrafi precedenti e le formule discusse, abbiamo ottenuto i seguenti risultati:

```
Country
Japan      2.389178
Chinese Taipei  2.183193
Switzerland  2.085713
South Korea  2.019607
Germany     1.984514
...
Nigeria     -1.648391
Papua New Guinea -1.738730
Iraq         -1.870009
South Sudan  -2.240257
Chad         -2.547961
Length: 148, dtype: float64
```

Figura 3.3: Risultati

Osservando in questo modo che il Giappone risulta essere il paese con un indice di ECI (Economic Complexity) maggiore, seguito da altri paesi altrettanto rilevanti come Taiwan, Svizzera, Corea del Sud e Germania. I risultati ci portano a definire in maniera concreta i paesi che hanno una maggiore influenza sull'economia globale e ad assegnare loro un valore che ne rappresenta l'importanza.

I valori ottenuti sono conformi a quanto illustrato dall'osservatorio OEC (Observatory of Economic Complexity) in quanto viene utilizzata per il calcolo la stessa teoria.

## Capitolo 4

# Conclusioni

L'analisi condotta nel corso di questo progetto ha permesso di mettere in luce l'importanza dell'Economic Complexity Index (ECI) come strumento di valutazione della diversità e sofisticatezza delle economie nazionali. Sviluppato da Cesar A. Hidalgo e Ricardo Hausmann presso il MIT Media Lab, l'ECI si è rivelato essenziale per comprendere la complessità dell'economia globale. L'ECI, utilizzando una metodologia basata su dati e algoritmi avanzati, è in grado di fornire una rappresentazione accurata della capacità produttiva di una nazione e del suo potenziale di crescita economica. I risultati ottenuti nel corso della nostra analisi sono in linea con quelli presentati dall'Observatory of Economic Complexity (OEC), confermando la validità e l'efficacia dell'ECI come strumento di analisi economica. Tuttavia, come ogni strumento, l'ECI ha le sue limitazioni e non dovrebbe essere utilizzato da solo per prendere decisioni economiche. Invece, dovrebbe essere integrato con altre metriche e considerazioni per fornire una visione più completa e accurata della salute economica di una nazione. In conclusione, l'Economic Complexity Index è uno strumento prezioso per comprendere la complessità economica. Attraverso la sua applicazione, è possibile identificare le forze trainanti dietro la crescita economica e fornire indicazioni preziose per le politiche economiche future.

# Bibliografia

- [1] B. Balassa. “Trade Liberalization and Revealed Comparative Advantage, 33, 99-123”. In: (). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9957.1965.tb00050.x>.
- [2] Luciano Pietronero. “Economic Complexity”. In: (). URL: <http://www.lucianopietronero.it/wp-content/uploads/2017/04/EC4U.pdf>.
- [3] “The building blocks of economic complexity di César A. Hidalgo and Ricardo Hausmann”. In: (). URL: <https://www.pnas.org/doi/full/10.1073/pnas.0900943106>.
- [4] “The Observatory of Economic Complexity”. In: (). URL: <https://oec.world/en/resources/methods#economic-complexity>.