

## Divergenze nei rating ESG: un ostacolo allo sviluppo della finanza sostenibile

LORENZO ESPOSITO, GIULIA CAVALERI\*

### Abstract:

Lo studio analizza la divergenza nei rating ESG assegnati da LSEG Workspace e Bloomberg a circa 900 aziende degli indici STOXX Europe 600 e S&P 500. Emergono differenze significative tra i punteggi, non uniformi tra settori e aree geografiche. Le discrepanze sono influenzate da caratteristiche aziendali come dimensioni, governance, performance finanziaria, livello di disclosure e presenza di controversie ESG. Le imprese più grandi mostrano divergenze più marcate, mentre fattori come efficienza operativa e controversie incidono diversamente tra Europa e Stati Uniti. I risultati mettono in luce limiti metodologici nei rating ESG e mostrano la necessità di maggiore standardizzazione, trasparenza e coerenza tra provider per garantire valutazioni affidabili e favorire mercati finanziari realmente sostenibili.

### Divergences in ESG ratings: an obstacle to the development of sustainable finance

*The study examines the divergence in ESG ratings assigned by LSEG Workspace and Bloomberg to around 900 companies included in the STOXX Europe 600 and S&P 500 indices. Significant differences emerge between the scores, which are not uniform across sectors or geographic areas. These discrepancies are influenced by firm-specific characteristics such as size, governance structure, financial performance, level of disclosure, and ESG-related controversies. Larger companies tend to show more pronounced divergences, while factors like operational efficiency and controversies affect rating consistency differently in Europe and the United States. The findings highlight methodological limitations in ESG ratings and underscore the need for greater standardization, transparency, and alignment across providers to ensure reliable assessments and support truly sustainable financial markets.*

Esposito: Banca d'Italia, Milano, Università Cattolica del Sacro Cuore, Milano, Italia  
email: [lorenzo.esposito@unicatt.it](mailto:lorenzo.esposito@unicatt.it)  
Cavalieri: Università Cattolica del Sacro Cuore, Milano, Italia, email: [giulia.cavalieri01@icatt.it](mailto:giulia.cavalieri01@icatt.it), [giulia.cav@hotmail.com](mailto:giulia.cav@hotmail.com)

### Per citare l'articolo:

Esposito L., Cavalieri G. (2025), "Divergenze nei rating ESG: un ostacolo allo sviluppo della finanza sostenibile", *Moneta e Credito*, 78 (312), pp. 355-377.

### DOI:

<https://doi.org/10.13133/2037-3651/19261>

### JEL codes:

Q56, G24, G15

### Keywords:

ESG ratings, sustainable finance, corporate governance

### Homepage della rivista:

[https://rosa.uniroma1.it/rosa04/moneta\\_e\\_credito](https://rosa.uniroma1.it/rosa04/moneta_e_credito)

L'integrazione dei fattori ESG (Environmental, Social e Governance) è un aspetto sempre più importante nelle decisioni di investimento, e permette la crescita del finanziamento di aziende e progetti sostenibili e le banche e gli altri operatori finanziari si servono ormai correntemente dei rating ESG per valutare la sostenibilità dei progetti finanziati. Sebbene tali rating siano uno strumento utile per orientare gli investitori nelle loro decisioni, presentano talune criticità. Una

\* Le opinioni espresse sono quelle degli autori e non impegnano gli istituti di appartenenza.

delle più significative è la loro eterogeneità: le diverse agenzie di rating adottano metodologie, criteri di ponderazione e basi informative differenti. Nella maggior parte dei casi, ciò porta a risultati divergenti, tanto che lo stesso emittente o strumento finanziario può ricevere valutazioni molto diverse a seconda del provider che le elabora. Diversi studi hanno dimostrato che la correlazione tra questi rating è alquanto ridotta, un risultato preoccupante dato che dovrebbero misurare la stessa realtà. Ciò solleva interrogativi sull'affidabilità e la comparabilità dei punteggi ESG e può generare incertezze per gli investitori. La mancanza di coerenza mette in discussione l'affidabilità complessiva dei rating e alimenta il dibattito sulla necessità di una maggiore trasparenza e standardizzazione nei criteri adottati.

L'obiettivo di questo articolo è contribuire a questo dibattito, documentando e analizzando le discrepanze tra i punteggi ESG forniti da due importanti provider internazionali: LSEG Workspace (Refinitiv – Asset4) e Bloomberg. Nello specifico, utilizziamo un'analisi trasversale basata sui dati più recenti disponibili, per quantificare le discrepanze tra le due fonti e indagarne i fattori determinanti. L'obiettivo è comprendere se vi sia una divergenza sistematica tra i punteggi ESG forniti dai due provider, se questi varino significativamente tra i settori e se siano connesse al diverso contesto (statunitense ed europeo). L'articolo è strutturato come segue. In primo luogo, analizziamo la letteratura relativa all'analisi della coerenza dei rating ESG; presentiamo quindi un'analisi dei dati dello STOXX Europe 600 e dello S&P 500 volta a identificare la presenza e l'entità delle discrepanze, e un'analisi econometrica basata su un modello di regressione multipla volta ad attribuire il divario di punteggio a una serie di fattori. Infine, confrontiamo i dati approfondendo l'origine delle discrepanze. La principale conclusione dello studio concerne l'analisi dei fattori che determinano la divergenza e che rimandano ad aspetti anche strutturali del quadro normativo ed economico europeo.

## 1. Divergenze nei rating ESG in letteratura

I rating ESG sono definiti dall'European Securities and Markets Authority (ESMA) come: “un parere relativo all'impatto o all'esposizione di un'azienda, un emittente o un titolo di debito rispetto ai fattori ESG, all'allineamento con gli accordi internazionali sul clima o alle caratteristiche di sostenibilità, espresso utilizzando un sistema di classificazione definito di categorie di rating” (ESMA, 2021, p. 3, nostra traduzione). A differenza dei rating creditizi, che valutano l'affidabilità finanziaria di un'azienda sulla base di dati ben consolidati che determinano un'indicazione univoca – la probabilità di default – la valutazione della performance ESG è meno oggettiva, in quanto a fonti qualitative e a specifiche metodologie proprietarie adottate dai vari provider. Sotto il profilo della generale procedura di valutazione, il lavoro di Del Giudice et al. (2024) offre un'utile descrizione di come i dati grezzi vengono trasformati in rating ESG: il processo inizia con la raccolta di dati grezzi sull'azienda e le sue attività, con cui vengono costruite le metriche ESG, misure quantitative che consentono di valutare la performance ESG. Queste metriche vengono poi riassunte in indicatori che offrono una rappresentazione sintetica e facilmente comprensibile dei vari aspetti considerati. Gli indicatori, a loro volta, confluiscono in categorie tematiche più ampie, che raggruppano dimensioni correlate come la sostenibilità ambientale e le condizioni dei lavoratori. Il passaggio finale consiste nel collegare queste tematiche ai pilastri E, S e G, che, integrando le informazioni raccolte, consentono la misurazione del punteggio ESG complessivo dell'azienda.

Se l'iter è simile tra tutti i provider, la letteratura ha da tempo evidenziato divergenze nei risultati dell'iter stesso. Il contributo pionieristico di Chatterji et al. (2016) ha messo in luce il

grado di divergenza tra le valutazioni della responsabilità sociale d'impresa, evidenziando una marcata eterogeneità nei giudizi, che, secondo gli autori, può essere ricondotta principalmente a due fattori: la mancanza di un quadro teorico condiviso e la mancanza di commensurabilità tra le diverse misure. L'articolo ebbe il merito di innescare il dibattito sulla divergenza dei rating ESG. Proseguendo il dibattito, Berg et al. (2022a) hanno proposto una ripartizione quantitativa delle differenze riscontrate tra i rating di sei principali agenzie di rating: Kinder, Lydenberg e Domini (KLD), Sustainalytics, Moody's ESG (Vigeo-Eiris), S&P Global (RobecoSAM), Refinitiv (Asset4) e MSCI. Utilizzando un totale di 709 indicatori, organizzati in una tassonomia di 64 categorie, gli autori hanno evidenziato che la correlazione tra i rating considerati varia tra 0,38 e 0,71, con una media complessiva di 0,54. La divergenza è ascritta all'adozione, da parte di ciascun operatore, di una propria metodologia di costruzione. In particolare, la divergenza nella misurazione rappresenta il principale fattore di divergenza (56%), seguita dalla divergenza nello *scope* dell'analisi (38%). Gli autori mostrano inoltre l'esistenza di un "effetto alone", in quanto un'azienda che ottiene un punteggio elevato in una categoria tende a ricevere valutazioni più favorevoli in altre, dato che le agenzie di rating organizzano il lavoro dei loro analisti per azienda piuttosto che per singolo indicatore. Vi è anche il rischio che imprese più grandi ricevano rating ESG migliori perché possono dedicare più risorse a raccogliere e trasmettere dati e che dunque il rating sia l'esito della presenza o assenza di dati più che di un confronto tra dati (Gregory, 2024). Questi fenomeni suggeriscono che la divergenza tra i rating non è dovuta esclusivamente al rumore statistico, ma anche ad altre caratteristiche derivanti dal processo di produzione dei rating, confermando la necessità di una maggiore trasparenza e comparabilità.

Nel lavoro di Berg et al. (2022b) si sottolinea la difficoltà di misure accurate della performance ESG rendendo difficile lo sviluppo di analisi econometriche per stimare una relazione realistica con variabili finanziarie. Se i rating ESG, che soffrono dei problemi che abbiamo visto, vengono utilizzati come regressori, le stime dei coefficienti saranno influenzate da un *bias* di attenuazione, con una distorsione sistematica verso lo zero. Le diverse fonti utilizzate per generare i rating ESG differiscono per caratteristiche chiave, come il grado di accessibilità al pubblico, la natura obbligatoria o volontaria dell'informativa e la conformità agli standard di reporting. Di conseguenza, gli indicatori alla base dei rating ESG incorporano inevitabilmente una quantità significativa di rumore. L'analisi di 1.297 aziende e sette rating ESG (ISS, Moody's, MSCI, Refinitiv, Sustainalytics, S&P Global e Truvalue Labs) evidenzia una bassa correlazione media (0,36), confermando un grado considerevole di divergenza tra le valutazioni fornite dai diversi provider. Per affrontare il problema dell'attenuazione dei coefficienti, gli autori sostituiscono le stime OLS con un approccio a due stadi, le cui regressioni rivelano un effetto molto più forte della performance ESG sui rendimenti azionari: quasi tutti i coefficienti crescono significativamente, confermando l'ipotesi che i rating ESG soffrano di un *bias* di attenuazione. Un risultato particolarmente interessante è che, nonostante le correlazioni relativamente basse tra loro, i rating ESG si dimostrano robusti nel contesto della stima strumentale: ogni rating è ben previsto da una combinazione dei punteggi forniti da altri provider. Tuttavia, gli autori stessi osservano che l'arco temporale dello studio è limitato e vi è anche un problema di continuità nelle serie storiche poiché il settore dei rating ESG è in fase di consolidamento e fusioni e acquisizioni tra *raters* determinano spesso una revisione delle metodologie e, in alcuni casi, la ricostruzione retrospettiva dei punteggi passati sulla base di nuove procedure. Queste stesse considerazioni motivano la scelta di focalizzare la nostra analisi su una prospettiva trasversale, ritenuta più robusta date le attuali caratteristiche dei dati disponibili.

Da un punto di vista metodologico, riteniamo che il punto focale sia trovare una soluzione pratica al problema del *bias* di attenuazione. Ogni volta che un punteggio ESG viene utilizzato

come regressore, è possibile utilizzare un secondo rating per ridurre il *bias*, producendo risultati più robusti e statisticamente significativi. Il tema della limitata disponibilità di dati ESG è affrontato da Gibson et al. (2021) analizzando un panel di aziende appartenenti all'indice S&P 500 nel periodo 2010-2017 e i rating ESG forniti da sette provider (Asset4/Refinitiv, Sustainalytics, Inrate, Bloomberg, FTSE, KLD e MSCI IVA). La correlazione media è piuttosto bassa (0,45). I risultati delle regressioni aggregate del panel mostrano che i rendimenti azionari sono positivamente associati alla discordanza tra i rating, suggerendo l'esistenza di un premio al rischio per le aziende caratterizzate da una maggiore eterogeneità nelle valutazioni ESG. Inoltre, gli autori sottolineano che il livello di discordanza tende a essere maggiore per le aziende più grandi, a causa della loro complessità organizzativa, mentre le aziende più redditizie mostrano livelli di discordanza inferiori, probabilmente per la possibilità di allocare maggiori risorse a politiche e pratiche di informativa ESG. Christensen et al. (2021) concentrano la loro analisi sul ruolo della *disclosure* aziendale come fattore determinante della divergenza, esaminando tre provider internazionali (MSCI, Thomson Reuters e Sustainalytics). Gli autori evidenziano come, contrariamente alle aspettative, una maggiore *disclosure* da parte delle aziende stesse sia associata a un aumento del disaccordo tra le agenzie di rating ESG. Per giungere a queste conclusioni, gli autori utilizzano un approccio *difference-in-differences* con effetti fissi a livello aziendale, introducendo shock normativi come i requisiti di *disclosure* ESG. Nel complesso, lo studio dimostra che il semplice aumento della *disclosure* non è una soluzione efficace.

Capizzi et al. (2021) hanno condotto un'analisi empirica su un campione di 210 società italiane quotate nel 2019 e nel 2020, valutando la coerenza dei rating ESG forniti da sei provider internazionali (MSCI, Refinitiv, S&P Global, Arabesque, Truvalue Labs e Inrate). Solo 22 società del campione erano contemporaneamente coperte da tutti e sei i provider considerati, indicando una significativa incoerenza nella disponibilità dei rating ESG. Lo studio mostra che, anche nel contesto italiano, il livello di correlazione tra i diversi provider è generalmente basso. Per l'intero campione, le correlazioni variano da molto basse (0,03) a moderate (0,64), con una media complessiva di 0,32. Nel sottoinsieme limitato alle società coperte da tutti i provider, la media sale leggermente (0,41). Gli autori osservano che la divergenza tra i rating ESG riflette profonde incongruenze metodologiche, in particolare nella definizione delle metriche e nell'assegnazione dei pesi ai vari indicatori. La principale raccomandazione degli autori è quella di standardizzare il processo di costruzione del rating attraverso l'introduzione di requisiti e normative comuni da parte dei *regulators* come la citata ESMA.

Abhayawansa e Tyagi (2021), nella loro revisione della letteratura, confermano i risultati di Chatterji et al. (2016): le divergenze tra i rating derivano dall'adozione di diverse definizioni concettuali (aspetti teorici) e dall'utilizzo di metodologie eterogenee nella misurazione delle performance ESG delle aziende (aspetti di misurazione). A ciò si aggiunge la mancanza di trasparenza da parte delle agenzie di rating in merito alle metodologie e alle fonti utilizzate. Gli autori ritengono che il problema della trasparenza persisterà poiché le metodologie di rating e le fonti di dati (ad esempio, le interviste e i questionari) sono proprietà intellettuale delle agenzie di rating. Inoltre, a loro avviso, l'utilità dei rating ESG è ulteriormente compromessa da potenziali conflitti di interesse.

Nel già citato lavoro di Del Giudice et al. (2024) si analizzano i rating ESG forniti da Moody's, Refinitiv e Bloomberg, considerando un campione di 1.776 aziende per il 2021. Gli autori osservano un problema strutturale nella disponibilità e completezza dei dati ESG, che risultano frammentati e incoerenti. Ad ogni modo, l'analisi rivela un livello di correlazione piuttosto basso (0,43) per il punteggio complessivo. Tra i tre provider, i rating di Moody's e Refinitiv sono più simili, mentre Bloomberg tende a mostrare deviazioni più marcate rispetto agli altri. Un altro

punto di interesse è la diversa coerenza riscontrata nei diversi contesti geografici: i rating delle aziende europee mostrano una correlazione media più elevata (0,49) rispetto a quelli delle aziende americane (0,35). Questo divario è in parte spiegato, secondo gli autori, dal quadro normativo europeo, caratterizzato da direttive e standard di rendicontazione come gli European Sustainability Reporting Standards.

Agosto e Tanda (2025), nel confermare le divergenze tra i rating ESG, osservano che la loro proliferazione deriva dalla forte domanda degli investitori per queste valutazioni, che deve essere soddisfatta anche se non sempre dati e metodi sono ben testati. Rilevano, inoltre, che risultati divergenti possono essere fonte di perplessità per il management aziendale, che non è in grado di individuare quali strategie possano aiutare l'azienda a diventare più sostenibile. Charlin et al. (2024), analizzando i rating forniti da quattro principali agenzie di rating (ISS, MSCI, S&P e Sustainalytics) alle società dell'indice S&P 500, hanno riscontrato un basso livello di affidabilità e concordanza. Infine, Stewart (2025), analizzando 13 diverse agenzie di rating ESG (Bloomberg, CDP, EcoVadis, FTSE Russell's ESG Ratings, ISS ESG, Just Capital, Moody's Investor Services, MSCI, Refinitiv, RepRisk, S&P Global, Sustainable Fitch e Sustainalytics) osserva difformità soprattutto nella valutazione della governance e una generale eterogeneità di dati che può generare posizioni contrastanti sulle performance ESG, riducendo la credibilità di questi rating.

Nel complesso, l'analisi della letteratura sul tema conferma che le divergenze rimangono significative, rendendo importante analizzarne e attenuarne le cause perché diventino strumenti sempre più efficaci nella scelta degli investimenti.

## 2. Analisi descrittiva

Abbiamo predisposto un set di dati ricavato dai database Bloomberg e LSEG Workspace che include le società degli indici STOXX Europe 600 e S&P 500 a maggio del 2025, per le quali è ragionevole aspettarsi una maggiore disponibilità di informazioni e una più ampia copertura dei dati ESG. Per questo motivo, come nel caso di Berg et al. (2022b), il campione è costituito da aziende di grandi dimensioni. La disamina della letteratura ha chiaramente evidenziato come l'arco temporale e la frequenza dei dati ESG siano ancora piuttosto limitati: molte agenzie hanno iniziato a pubblicare i propri punteggi solo di recente e gli aggiornamenti sono prevalentemente annuali. Inoltre, il consolidamento del settore e i cambiamenti metodologici comportano una ricostruzione retrospettiva dei punteggi. Queste considerazioni suggeriscono l'adozione di un approccio trasversale, che cattura la situazione più recente disponibile a quella data. Tutte le osservazioni con valori mancanti sono state escluse per garantire coerenza e comparabilità dei dati. Dopo aver eliminato le osservazioni con dati mancanti, il numero totale si riduce a 504.

Le prime due variabili da considerare sono i rating ESG forniti da Refinitiv (Asset4)-LSEG Workspace, che vanno da 0 a 100 (*lseg\_esg\_score*) e quelli di Bloomberg, che vanno da 0 a 10 (*b\_esg\_score*) questi ultimi moltiplicati per 10 per renderli comparabili. Da queste variabili otteniamo la variabile differenza (*difference*), definita come la distanza algebrica tra i due punteggi per ciascuna azienda. Le altre variabili incluse nel dataset includono: *name*: il nome dell'azienda; *GICS sector*, classifica l'appartenenza al settore secondo lo standard GICS con le seguenti categorie: servizi di comunicazione, beni di consumo discrezionali, beni di consumo primari, energia, finanza, sanità, industria, tecnologia dell'informazione, materiali, immobiliare, servizi di pubblica utilità; *employees*: numero totale di dipendenti (part-time e full-time); *board size*: dimensione del CdA; *board diversity*: diversità del CdA come percentuale di donne nel CdA; *board independent members*: percentuale di membri indipendenti nel CdA; *CO2 emissions*: emissioni totali di anidride

carbonica; *ESG CONT score*: il livello di esposizione alle controversie fornito dal punteggio LSEG; *ESG DISC score*: il grado di trasparenza delle informazioni fornito dal punteggio Bloomberg. A questi, aggiungiamo alcuni usuali indicatori di mercato: *mkt\_cap*, *return\_on\_assets*, *return\_on\_equity*, *debt\_on\_equity*, *p\_e\_ratio*, *ev\_on\_ebitda*, *capex\_on\_assets*, *total\_assets*, *beta*, *volatility*, per controllare le caratteristiche dell'azienda.

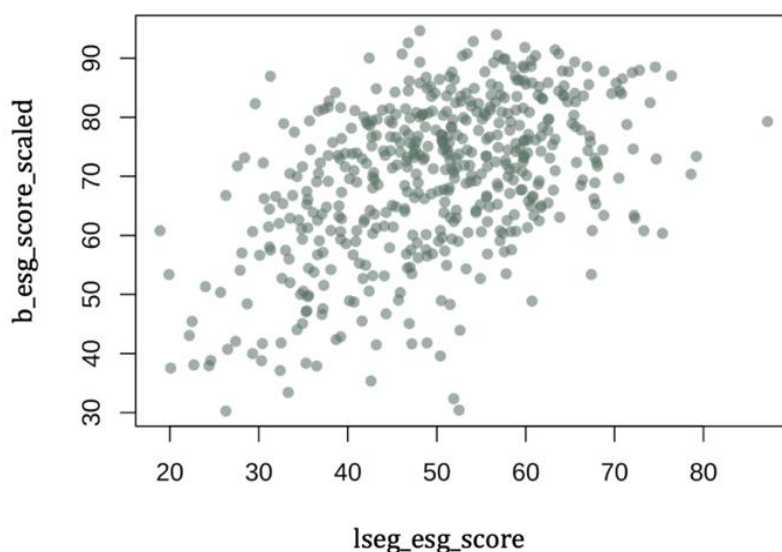
Per evidenziare le caratteristiche principali e identificare somiglianze o discrepanze significative tra lo STOXX Europe 600 e lo S&P 500, i due campioni sono stati trattati separatamente e poi confrontati.

### 2.1. Analisi descrittiva per lo STOXX Europe 600

Le statistiche descrittive presentate in questa sezione ci consentono di delineare le principali caratteristiche dei dati.

Un'analisi grafica iniziale (figura 1), effettuata utilizzando il diagramma di dispersione dei due punteggi ESG, rivela una relazione positiva, sebbene non perfettamente lineare, tra le due misure. La correlazione tra i due punteggi è pari a 0,47, con significativi disallineamenti tra le due fonti, meritevoli di ulteriori analisi.

Figura 1 – Diagramma di dispersione del punteggio ESG (STOXX Europe 600)



La tabella 1 mostra che i punteggi assegnati da LSEG sono in media più elevati (69,81 contro 49,92). Questa differenza è significativa anche considerando la mediana (71,84 contro 50,65), suggerendo una distribuzione sistematicamente maggiore dei valori per LSEG. Entrambi i punteggi presentano una variabilità simile, ma la distribuzione dei valori di LSEG è spostata verso l'estremità superiore della scala, mentre i punteggi Bloomberg sono distribuiti in modo più simmetrico. Dopo aver analizzato separatamente i due punteggi, passiamo ora alla variabile

*difference*. I dati suggeriscono che la sua distribuzione è prossima alla normalità. La media e la mediana si attestano intorno ai 20 punti; la dispersione (deviazione standard pari a 12,76) mostra inoltre che la coerenza tra i due indici non è uniforme e che esistono divergenze significative. Parte della divergenza deriva dalla composizione settoriale delle due fonti. Un utile riepilogo numerico è fornito dalla tabella 2, che mostra i valori medi e mediani della variabile differenza per ciascun settore GICS, evidenziando chiaramente come le divergenze tra i due punteggi ESG varino a seconda del settore considerato. Settori come sanità, beni di consumo di base e beni di consumo discrezionali presentano differenze particolarmente elevate. Di conseguenza, l'entità della variabile varia significativamente tra i settori, riflettendo specificità settoriali.

Tabella 1 – Sintesi dei rating ESG e differenze (Stoxx 600)

Variabile	Min	1° Q	Mediana	Media	3° Q	Max	DevStand
<i>lseg_esg_score</i>	30,26	61,44	71,84	69,81	79,66	94,65	13,09
<i>b_esg_score_scaled</i>	18,9	41,88	50,65	49,92	58,33	87,2	11,66
<i>difference</i>	-22,08	10,98	20,46	19,89	28,95	55,64	12,76

Tabella 2 – Media e mediana delle differenze tra settori (Stoxx 600)

Settori GICS	Media	Mediana
Servizi di comunicazione	16,47	17,58
Beni di consumo discrezionali	21,74	23,5
Beni di consumo primari	24,01	25,48
Energia	12,88	15,16
Finanza	20,14	20,79
Sanità	26,32	28,05
Industria	18,06	17,31
ICT	17,52	17,91
Materiali	20,16	20,33
Real Estate	14,1	15,4
Utilities	20,19	20,14
<b>Complessivo</b>	<b>19,89</b>	<b>20,46</b>

Passiamo ora alle altre variabili, introdotte nella descrizione del dataset. L'obiettivo è duplice: in primo luogo, fornire una panoramica generale delle caratteristiche delle aziende incluse nel campione e valutare eventuali relazioni preliminari sia con la variabile *difference* (dipendente) sia con le altre variabili (regressori). Per semplicità di presentazione, la tabella 3 riporta solo la media e la mediana delle variabili, accompagnate da un breve commento che evidenzia i risultati principali.

Tabella 3 – Media e mediana delle variabili esplicative (Stoxx 600)

<b>Variabile</b>	<b>Media</b>	<b>Mediana</b>
<i>esg_contr_score</i>	74,92	100
<i>CO2_emissions</i>	2.181.578	95.596
<i>employees</i>	43.976	16.752
<i>board_size</i>	10,91	10
<i>board_diversity</i>	39,8	40
<i>board_indep_members</i>	64,22	66,67
<i>mkt_cap</i>	24.400.000	9.470.000
<i>return_on_assets</i>	7,09	5,62
<i>return_on_equity</i>	18,34	13,55
<i>debt_on_equity</i>	154,34	71,5
<i>dividend_yield</i>	3,16	2,71
<i>p_e_ratio</i>	33	19,45
<i>ev_on_ebitda</i>	13,06	10,57
<i>capex_on_assets</i>	3,51	2,58
<i>total_assets</i>	81.500.000	14.100.000
<i>beta</i>	1	0,98
<i>volatility</i>	0,3	0,28
<i>esg_disc_score</i>	58,43	58,23

Le società europee del campione si distinguono per un elevato livello di controversie su temi ESG, presentando punteggi relativamente elevati con una mediana pari al punteggio massimo. Le caratteristiche di governance relative al CdA sono più stabili, con valori medi e mediani simili. Gli indicatori di performance mostrano una moderata variabilità. Le variabili beta e volatilità assumono valori in linea con quelli tipici delle società quotate. Anche i multipli di mercato sono dispersi.

## 2.2. Analisi descrittiva per lo S&P 500

Discutiamo ora i risultati nel contesto statunitense. Come nella sezione precedente, l'analisi inizia con un confronto tra i punteggi ESG forniti da LSEG e Bloomberg, opportunamente riscalati.

La relazione tra i due punteggi ESG (figura 2) mostra una correlazione positiva tra le due metriche, sebbene meno marcata rispetto al campione europeo (0,31). Questo risultato suggerisce che le aziende valutate più favorevolmente da LSEG tendono, in media, a ottenere punteggi relativamente elevati anche da Bloomberg, ma la relazione è debole.

Osserviamo (tabella 4) che media e mediana confermano uno spostamento sistematico verso l'alto nei punteggi forniti da LSEG, mentre la variabilità è simile. Notiamo inoltre che le aziende del campione si distinguono per un elevato livello di controversie ESG, presentando punteggi relativamente elevati. Le variabili *CO2\_emissions*, *employees*, *mkt\_cap* e *total\_assets* mostrano differenze tra media e mediana, indicative di distribuzioni asimmetriche e della presenza di valori anomali. Tuttavia, le caratteristiche di governance relative al CdA risultano più stabili, con valori di media e mediana molto simili. Gli indicatori di performance mostrano una moderata variabilità.

Le variabili *beta* e *volatility* assumono valori in linea con quelli tipici delle società quotate. Il rapporto *debt\_on\_equity*, con una media elevata influenzata da valori anomali, evidenzia la presenza di aziende fortemente indebitate. Anche i multipli di mercato risultano differenziati. Infine, l'indicatore *ev\_on\_ebitda* risulta in linea con gli standard di mercato, seppur con una media più elevata, riflettendo la presenza di aziende con valutazioni particolarmente generose. Per quanto riguarda l'analisi di correlazione, la variabile *difference* mostra deboli correlazioni con la maggior parte dei regressori e non vi è alcuna evidenza di forti relazioni lineari. Le associazioni più significative riguardano la dimensione aziendale, suggerendo che le aziende più grandi tendono a mostrare maggiori deviazioni tra i rating ESG. Il punteggio di controversia (*esg\_contr\_score*) è negativamente correlato alla differenza, mentre l'*esg\_disc\_score* mostra una correlazione positiva: le aziende più esposte a controversie mostrano minori differenze tra i rating, mentre le aziende più trasparenti tendono a mostrare deviazioni maggiori. Gli indicatori di performance finanziaria mostrano relazioni deboli e negative con la differenza. Nel complesso, i dati, confermano la consistenza delle differenze tra i rating.

Figura 2 – Diagramma di dispersione del punteggio ESG (S&P 500)

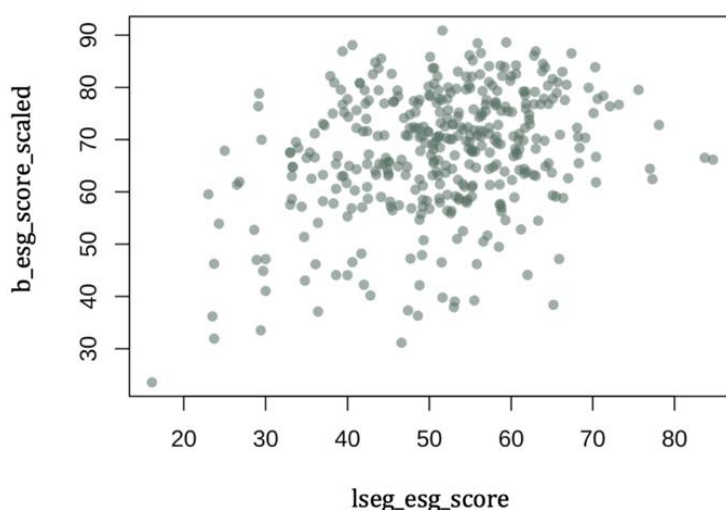


Tabella 4 – Sintesi dei rating ESG e differenze (S&P 500)

Variabile	Min	1° Q	Mediana	Media	3° Q	Max	DevStand
<i>lseg_esg_score</i>	23,58	61,3	69,23	67,62	76,78	90,87	11,81
<i>b_esg_score_scaled</i>	16,1	44,5	52,2	51,5	58,7	84,7	10,92
<i>difference</i>	-26,8	7,1	16,85	16,12	24,71	49,62	13,34

Tabella 5 – *Media e mediana delle differenze tra settori (S&P 500)*

<b>Settori GICS</b>	<b>Media</b>	<b>Mediana</b>
Servizi di comunicazione	21,46	26,21
Beni di consumo discrezionali	13,28	15
Beni di consumo primari	21,34	21,8
Energia	11,43	12,53
Finanza	19,41	21,07
Sanità	19,51	20,61
Industria	13,81	16,18
ICT	16,32	16,33
Materiali	14,22	15,37
Real Estate	18,25	20,74
Utilities	9,78	11,24
<b>Complessivo</b>	<b>16,12</b>	<b>16,85</b>

Come si vede in tabella 5, la differenza media e mediana è positiva in tutti i settori. Tuttavia, l'entità della deviazione non è uniforme. Nel complesso, le statistiche confermano una distorsione positiva a favore di LSEG, ma evidenziano anche come l'entità della differenza vari significativamente a seconda del settore, suggerendo che il disallineamento tra i due fornitori sia influenzato da caratteristiche specifiche delle diverse industrie.

Tabella 6 – *Media e mediana delle variabili esplicative (S&P 500)*

<b>Variabile</b>	<b>Media</b>	<b>Mediana</b>
<i>esg_contr_score</i>	57,38	57,27
<i>CO2_emissions</i>	4.402.196	320.134
<i>employees</i>	59569	21698
<i>board_size</i>	11,08	11
<i>board_diversity</i>	33,34	33,33
<i>board_indep_members</i>	86,38	90
<i>mkt_cap</i>	124.000.000	38.700.000
<i>return_on_asstes</i>	9,47	7,32
<i>return_on_equity</i>	90,13	17,6
<i>debt_on_equity</i>	155,22	74,41
<i>dividend_yield</i>	1,99	1,64
<i>p_e_ratio</i>	34,14	25,35
<i>ev_on_ebitda</i>	16,35	14,66
<i>capex_on_assets</i>	4,39	3
<i>total_assets</i>	98.400.000	29.000.000
<i>beta</i>	0,95	0,92
<i>volatility</i>	0.29	0,27
<i>esg_disc_score</i>	58,94	58,88

Per quanto riguarda le altre variabili (tabella 6), osserviamo che *esg\_contr\_score* assume valori medi prossimi alla mediana, indicando una distribuzione relativamente bilanciata, mentre le variabili *CO2\_emissions*, *mkt\_cap* e *total\_assets* mostrano una distanza significativa tra media e mediana, indicando la presenza di distribuzioni con valori estremi, inevitabile considerando la presenza di grandissimi gruppi. Le variabili di governance, invece, mostrano una maggiore omogeneità. Particolarmente degno di nota è il caso del *return\_on\_equity*, caratterizzato da una media molto elevata rispetto alla mediana, a testimonianza della presenza di valori anomali nel campione. Per quanto riguarda le valutazioni di mercato, il *p\_e\_ratio* mostra una media più elevata rispetto alla mediana, indicatore di alcune aziende con multipli molto elevati. Un comportamento simile si riscontra nell'indicatore *ev\_on\_ebitda*, coerente con i multipli generalmente osservati sui mercati. I dati mostrano inoltre che non vi sono evidenti problemi di multicollinearità tra le variabili esplicative. In primo luogo, le variabili relative alla dimensione aziendale mostrano, come previsto, correlazioni positive: ad esempio, la variabile sui dipendenti è associata a *mkt\_cap* e *total\_assets*, mentre *CO2\_emissions* è correlata a *capex\_on\_assets*, le diverse misure catturano dimensioni complementari. Per quanto riguarda le relazioni con la variabile dipendente *difference*, emergono correlazioni positive con *employees*, *total\_assets* e, in misura minore, con *mkt\_cap*. Ciò suggerisce che le aziende più grandi tendono a mostrare deviazioni maggiori. La variabile relativa alle controversie *esg\_disc\_score* mostra una correlazione molto debole ma positiva: le aziende con informative più trasparenti tendono ad avere rating più divergenti; mentre quelle coinvolte in controversie *esg\_contr\_score* mostrano deviazioni minori. Emergono anche alcune relazioni interessanti con le variabili di governance: *difference* è positivamente correlata a *board\_indep\_members* e *employees*, mentre la relazione tra *board\_size* e *difference* è quasi nulla. Infine, il *return\_on\_assets* è moderatamente correlato con le variabili *mkt\_cap* ed *ev\_on\_ebitda*, mentre il *dividend\_yield* è negativamente associato alla redditività e al rischio di mercato, evidenziando l'effetto di un minore autofinanziamento. In linea con la teoria, *beta* e *volatility* mostrano un'elevata correlazione.

### 2.3. Confronti

Per fornire una panoramica completa, è utile confrontare i risultati delle analisi descrittive per identificare somiglianze e potenziali discrepanze. In entrambi i casi, si osserva una correlazione positiva tra i punteggi ESG, ma la sua entità è diversa (del 50% più elevata per lo STOXX Europe 600). Questi valori sono coerenti con studi precedenti. In particolare, Del Giudice et al. (2024) riscontrano una correlazione media di 0,49 per le aziende europee e di 0,35 per le aziende statunitensi. Analogamente, Berg et al. (2022a) riportano una correlazione media di 0,54, con un intervallo di valori per le correlazioni compreso tra 0,38 e 0,71; Berg et al. (2022b) riscontrano un valore medio di 0,36; Gibson et al. (2021) di 0,45; infine, Capizzi et al. (2021) riportano un valore di 0,32 per le aziende italiane.

Tabella 7 – Comparazione delle statistiche descrittive (*difference*)

Differenza	Min	Mediana	Media	Max	DevStand
STOX Europe 600	-22,08	20,46	19,89	55,64	12,76
S&P 500	-26,8	16,85	16,12	49,62	13,34

La tabella 7 riassume il confronto delle discrepanze di punteggio nei due campioni. In entrambi i casi, la variabile mostra valori medi e mediani positivi, confermando la tendenza di LSEG ad assegnare punteggi più elevati. Tuttavia, l'entità della discrepanza varia significativamente tra settori e mercati. Per lo STOXX Europe 600, le differenze appaiono in media più elevate in settori come sanità e beni di consumo di base, che mostrano valori ben al di sopra della media complessiva, mentre in settori come energia e *real estate*, le deviazioni sono più contenute. Per lo S&P 500, tuttavia, i settori con le divergenze più marcate sono servizi di comunicazione e beni di consumo di base, mentre gli altri settori si concentrano attorno a valori prossimi alla media complessiva. Questi risultati sono utili per il confronto con i risultati dello studio di Del Giudice et al. (2024), che ha identificato i settori oil & gas e utilities come i meno soggetti a divergenze.

Un confronto tra le caratteristiche dei due campioni rivela alcune differenze rilevanti. Le aziende dell'indice S&P 500 presentano dimensioni maggiori e le emissioni di CO<sub>2</sub> sono significativamente più elevate tra le aziende statunitensi, in linea con le loro maggiori dimensioni; in entrambi i campioni, la variabile è caratterizzata da asimmetria dovuta alla presenza di grandi aziende inquinanti. Le aziende dell'indice STOXX Europe 600 mostrano punteggi *esg\_contr\_score* medi più elevati, indicando una maggiore concentrazione di aziende con un elevato numero di controversie. Al contrario, *esg\_disc\_score* assume valori simili nei due campioni, indicando che i livelli di trasparenza sono relativamente simili. Le variabili di governance appaiono simili nei due campioni, sebbene con una maggiore indipendenza tra i membri del CdA nel campione S&P 500. Dal punto di vista della performance, gli indicatori sono in media più elevati tra le aziende statunitensi, così come i multipli di mercato, riflettendo il clima più positivo dei mercati finanziari statunitensi. Inoltre, le aziende dell'indice S&P 500, in media, dedicano una quota maggiore di *capex\_on\_assets*, indicando una maggiore propensione a reinvestire mentre il rendimento da dividendi è in media più elevato per le società europee, riflettendo una diversa politica di remunerazione degli azionisti. Infine, i valori di beta e volatilità sono simili nei due campioni, indicando un profilo di rischio di mercato comparabile tra le società quotate americane ed europee.

### 3. Analisi empirica

Come per l'analisi descrittiva, i due dataset sono stati trattati separatamente per valutare se e come i modelli stimati differiscono tra lo STOXX Europe 600 e lo S&P 500. La scelta di sviluppare l'analisi separatamente mira a evidenziare come i due insiemi di variabili selezionati dalla selezione del "miglior sottoinsieme" differiscano nei due contesti, mentre una specificazione unica del modello avrebbe contrastato con l'obiettivo primario di evidenziare le potenziali differenze valutando quali variabili sono incluse in questo miglior sottoinsieme. La sezione dedicata ai confronti presenta le variabili risultate significative nei modelli, evidenziando in particolare i segni dei coefficienti stimati. A questo stadio di sviluppo del dibattito, il nostro obiettivo non è predittivo, ma esplorativo: i modelli sono stati utilizzati per comprendere le relazioni statistiche tra le variabili. Per questo motivo, non sono stati utilizzati metodi di validazione. La selezione dei possibili regressori è stata effettuata testando tutti i possibili regressori e identificando quelli che meglio descrivono il fenomeno. Abbiamo scelto l'Akaike Information Criterion (AIC) perché consente di selezionare il modello che offre un buon compromesso tra accuratezza e parsimonia, dimostrandosi particolarmente adatto ad analisi esplorative come quella qui condotta, assieme al Bayesian Information Criterion (BIC) con lo stesso scopo. Per quanto riguarda la correlazione, la

variabile *difference* mostra deboli correlazioni con la maggior parte dei regressori. Le associazioni più significative riguardano la dimensione aziendale, a conferma che le aziende più grandi tendono a mostrare maggiori deviazioni. Il punteggio di controversia (*esg\_contr\_score*) è negativamente correlato con la differenza (-0,13), mentre l'*esg\_disc\_score* mostra una correlazione positiva (0,12): le aziende più esposte a controversie mostrano minori differenze tra i rating, mentre le aziende più trasparenti tendono a mostrare deviazioni maggiori. Gli indicatori di performance finanziaria mostrano relazioni deboli e negative con la differenza. È anche ovvio perché le variabili dimensionali siano positivamente correlate tra loro. Tra le variabili di governance, si osserva una correlazione negativa tra la dimensione del CdA e la percentuale di membri indipendenti. Infine, *esg\_contr\_score* è associato negativamente a *esg\_disc\_score*, a indicare che le aziende più trasparenti tendono ad avere meno controversie; tra gli indicatori di rischio, emerge una moderata correlazione tra beta e volatilità.

### 3.1. Specifiche del modello e test diagnostico per STOXX Europe 600

Figura 3 – Numero ottimale di regressori (STOXX Europe 600)

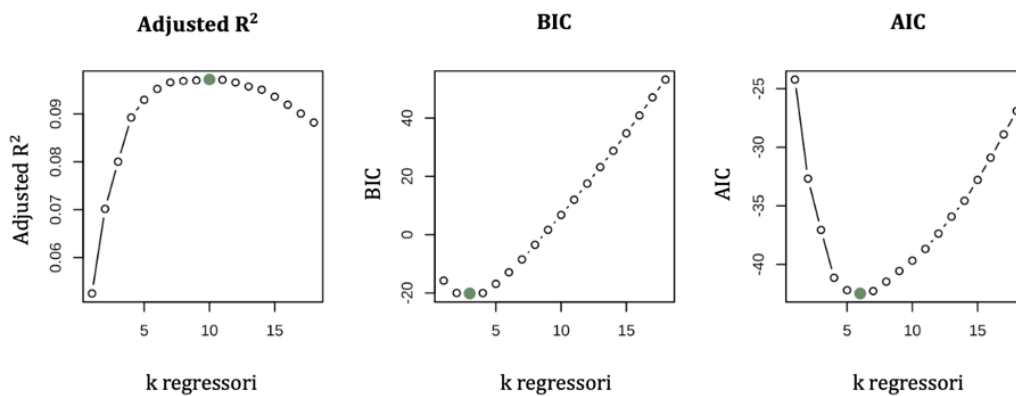


Tabella 8 – Variabili selezionate in base ai tre criteri indicati (Stoxx Europe 600)

Adj R <sup>2</sup>	BIC	AIC
<i>CO2_emissions</i>	<i>employees</i>	<i>employees</i>
<i>employees</i>	<i>board_size</i>	<i>board_size</i>
<i>board_size</i>	<i>debt_on_equity</i>	<i>return_on_assets</i>
<i>board_diversity</i>		<i>debt_on_equity</i>
<i>return_on_assets</i>		<i>total_assets</i>
<i>return_on_equity</i>		<i>beta</i>
<i>debt_on_equity</i>		
<i>p_e_ratio</i>		
<i>total_assets</i>		
<i>beta</i>		

La figura 3 mostra il comportamento di tre criteri: adjusted R<sup>2</sup>, BIC e AIC al variare del numero di regressori. Il punto verde rappresenta il massimo per il primo criterio e il minimo per gli altri due, indicando il numero ottimale di regressori da utilizzare, numero oltre il quale l'aggiunta di altre variabili non migliora più la qualità dell'adattamento in base ai criteri (le variabili selezionate sono elencate in tabella 8).

Come emerge dai dati, il valore ottimale è 10 per l'R<sup>2</sup> adjusted (il modello più ampio), 3 per il BIC (il più restrittivo) e 6 per l'AIC. Avendo selezionato il criterio AIC, si riporta la seguente specifica del modello, dove i coefficienti beta quantificano l'effetto marginale di ciascuna variabile esplicativa sulla variabile dipendente, *difference*:

$$\begin{aligned} difference_i = & \beta_0 + \beta_1 employees_i + \beta_2 board\_size_i + \beta_3 return\_on\_assets_i + \\ & \beta_4 debt\_on\_equity_i + \beta_5 total\_assets_i + \beta_6 beta_i + \varepsilon_i \end{aligned} \quad (1)$$

La tabella 9 mostra l'output del modello.

Tabella 9 – Risultati della regressione OLS (Stoxx Europe 600)

	Coefficiente	Standard error	t_ratio	p_value
Costante	15,37	2,249	6,84	0 ***
<i>employees</i>	0,00002092	0,000007491	2,79	0,005 **
<i>board_size</i>	0,569	0,173	3,29	0,001 **
<i>return_on_assets</i>	-0,137	0,059	-2,35	0,019 *
<i>debt_on_equity</i>	-0,002	0,001	-2,65	0,008 **
<i>total_assets</i>	0,0000000040	0,0000000020	1,95	0,052 .
<i>beta</i>	-1,692	1,128	-1,5	0,134
R <sup>2</sup>	0,106			
F (6, 497)	9,82			
Adjusted R <sup>2</sup>	0,095			
P-value	0,001			
Livelli di significatività: "." 10%; * p<0,05; ** p<0,01; *** p<0,001				

L'intercetta stimata è 15,37 ed è altamente significativa ( $p < 0,001$ ). La significatività dell'intercetta suggerisce l'esistenza di una componente non spiegata dalle variabili indipendenti incluse nel modello. Il numero di dipendenti mostra un coefficiente positivo e significativo al livello dell'1%. Nello specifico, mantenendo costanti le altre variabili, un aumento di un'unità nel numero di dipendenti porta a un aumento medio della variabile di risposta minimo. Sebbene l'effetto appaia piccolo in termini assoluti, considerando variazioni più ampie, l'impatto diventa significativo. Ciò significa che le aziende con più dipendenti tendono a presentare deviazioni maggiori nei loro punteggi ESG. Anche la dimensione del CdA è positivamente associata alla differenza tra i punteggi, con un coefficiente stimato di 0,57 e significativo al livello dell'1%. Ciò significa che, a parità di altri fattori, un membro aggiuntivo del CdA determina, in media, un

aumento di circa mezzo punto per la variabile *difference*. Al contrario, il rendimento delle attività presenta un coefficiente negativo ed è significativo al livello del 5%. Esiste una relazione inversa: le aziende più redditizie tendono, in media, ad avere una differenza minore tra i rating ESG. Analogamente, il rapporto debito/equity presenta un coefficiente negativo ed è significativo al livello dell'1%. Le aziende con una leva maggiore presentano, in media, differenze minori nei rating ESG. La variabile *total\_assets* mostra un effetto positivo marginalmente significativo al livello del 10%, coerente con l'idea che la dimensione degli asset possa avere un impatto positivo, seppur limitato. L'effetto stimato suggerisce che, sebbene con evidenze statistiche limitate, le aziende più grandi tendono, in media, ad avere valori leggermente più elevati della variabile dipendente.

Nel complesso, il modello mostra ridotti  $R^2$  e adjusted  $R^2$ , indicando che le variabili considerate spiegano circa il 10% della variabilità della variabile dipendente. Nonostante il potere esplicativo relativamente limitato, il test F è altamente significativo ( $p < 0,001$ ), confermando che almeno una delle variabili esplicative contribuisce in modo non casuale a spiegare la variabilità della variabile dipendente. In sintesi, il modello suggerisce che la varianza tra i punteggi ESG non è casuale, ma è legata a specifiche caratteristiche aziendali. In particolare, la complessità organizzativa e dimensionale (più dipendenti, un CdA più numeroso, maggiori asset) tende a favorire valutazioni ESG meno omogenee, mentre la solidità finanziaria e un debito più elevato sono associati a una percezione esterna più uniforme delle performance ESG.

Di seguito sono presentati i test diagnostici eseguiti sui residui del modello stimato (tabella 11). Poiché abbiamo eseguito un'analisi trasversale, non è stato necessario controllare i problemi di autocorrelazione.

L'assenza di problemi di multicollinearità è stata verificata calcolando il VIF per ciascun regressore (tabella 10). Tutti i valori riportati sono al di sotto della soglia convenzionale ( $VIF < 5$ ), indicando che non vi è ridondanza tra i regressori. Inoltre, l'analisi non ha rilevato dipendenze lineari tra i regressori. Si può quindi concludere che il modello non è affetto da problemi di multicollinearità.

Tabella 10 – *Variance Inflation Factor (VIF)*

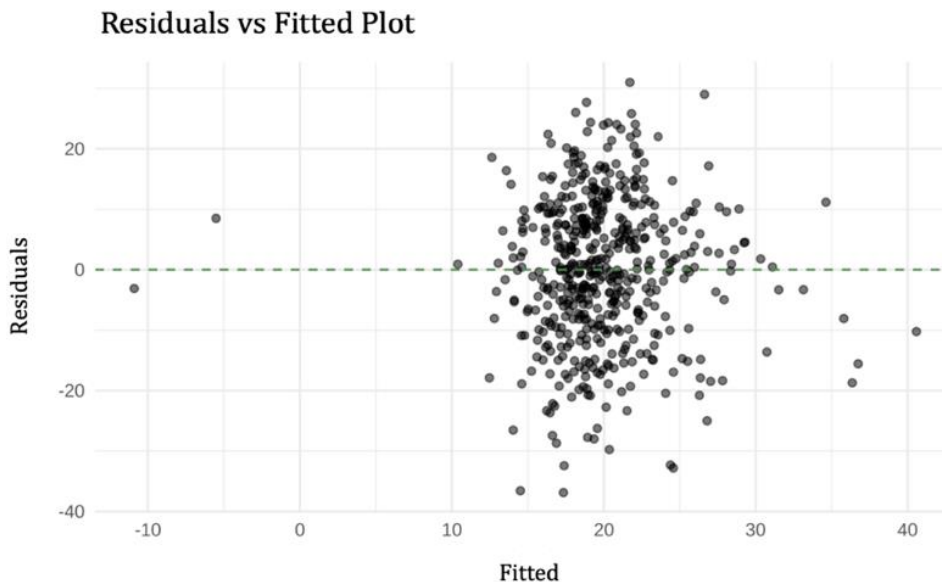
Variabile	VIF
<i>employees</i>	1,177
<i>board_size</i>	1,239
<i>return_on_assets</i>	1,056
<i>debt_on_equity</i>	1,016
<i>total_assets</i>	1,128
<i>beta</i>	1,031

Tabella 11 – *Test diagnostici*

Test	Statistica	DoF	p_value	Risultato (5%)
Shapiro-Wilk	0,9947	–	0,0813	non rifiutata
Breusch-Pagan	1,0418	6	0,984	non rifiutata
RESET	5,213	2 – 495	0,0057	rifiutata

La normalità dei residui è stata testata utilizzando il test di Shapiro-Wilk. Il valore p ( $W = 0,0813$ ) superiore al 5% porta a non rifiutare l'ipotesi nulla; sebbene il valore sia vicino alla soglia, suggerisce che non vi siano deviazioni significative dalla normalità. Il test di Breusch-Pagan non rivela problemi di eteroschedasticità. Questo risultato è importante perché supporta la validità delle inferenze basate sul modello OLS che stiamo stimando. Infine, il risultato del test Ramsey RESET indica che la specificazione del modello potrebbe non essere del tutto corretta. La statistica test è 5,213, con un p-value di 0,0057, inferiore alla soglia convenzionale del 5%. Di conseguenza, l'ipotesi nulla di specificazione corretta viene rifiutata, suggerendo la possibile presenza di variabili esplicative omesse o di una forma funzionale non lineare non catturata dal modello stimato. Questo risultato non invalida necessariamente le stime ottenute, ma indica che l'analisi potrebbe trarre beneficio dall'inclusione di ulteriori regressori o trasformazioni funzionali delle variabili esistenti, per migliorare la rappresentazione del fenomeno. Un ultimo strumento diagnostico, utile da considerare, è il grafico dei residui rispetto ai valori stimati (figura 4): una distribuzione casuale e simmetrica dei residui attorno allo zero (linea tratteggiata rossa) indica che la relazione lineare è appropriata e che non emergono evidenti problemi di specificazione. L'assenza di andamenti regolari o un progressivo allargamento della dispersione suggerisce inoltre che la varianza degli errori è costante (omoschedasticità). Infine, la presenza di residui molto elevati può segnalare osservazioni atipiche o influenti.

Figura 4 – Grafico dei residui rispetto ai valori stimati (STOXX Europe 600)



Nel complesso, i test confermano che le ipotesi di normalità e omoschedasticità sono soddisfatte, sebbene emergano indicazioni di potenziali limitazioni nella specificazione funzionale del modello. Ciò implica che vi sia spazio per futuri miglioramenti del modello, ad esempio attraverso l'inclusione di ulteriori variabili esplicative o forme funzionali non lineari.

### 3.2. Specificazione del modello e test diagnostico per lo S&P 500

Procediamo in modo analogo per le aziende incluse nello S&P 500 (figura 5). La tabella 12 elenca i regressori selezionati per ciascun criterio.

Figura 5 – Numero ottimale di regressori (S&P 500)

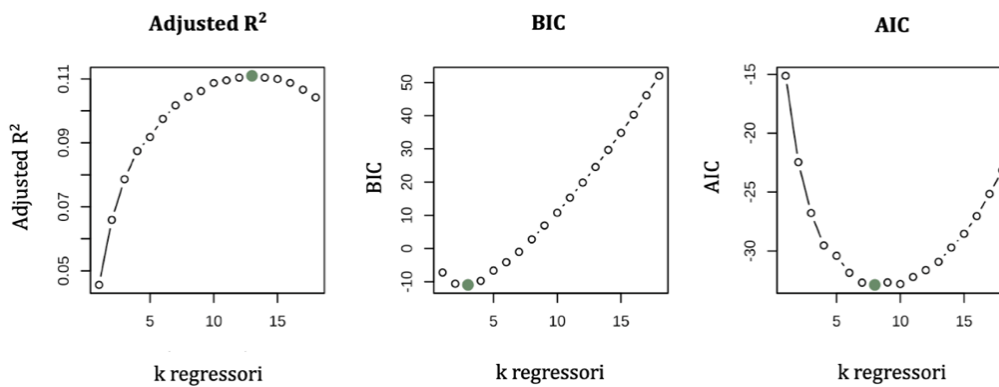


Tabella 12 – Variabili selezionate in base ai tre criteri indicati (S&P 500)

Adj R <sup>2</sup>	BIC	AIC
<i>esg_contr_score</i>	<i>employees</i>	<i>esg_contr_score</i>
<i>CO2_emissions</i>	<i>board_indep_members</i>	<i>CO2_emissions</i>
<i>employees</i>	<i>total_assets</i>	<i>employees</i>
<i>board_size</i>		<i>board_size</i>
<i>board_indep_members</i>		<i>board_indep_members</i>
<i>debt_on_equity</i>		<i>debt_on_equity</i>
<i>dividend_yield</i>		<i>total_assets</i>
<i>p_e_ratio</i>		<i>volatility</i>
<i>ev_on_ebitda</i>		
<i>total_assets</i>		
<i>beta</i>		
<i>volatility</i>		
<i>esg_disc_score</i>		

Anche in questo caso adottiamo il criterio AIC da cui segue la specifica del modello:

$$\begin{aligned} difference_i = & \beta_0 + \beta_1 esg\_contr\_score_i + \beta_2 CO2\_emissions_i + \beta_3 employees_i + \\ & \beta_4 board\_size_i + \beta_5 board\_indep\_members_i + \beta_6 debt\_on\_equity_i + \beta_7 total\_assets_i + \\ & \beta_8 volatility_i + \varepsilon_i \end{aligned} \quad (2)$$

Nella tabella 13 vediamo i risultati.

Tabella 13 – Risultati della regressione OLS (S&P 500)

	Coefficiente	Deviazione standard	t_ratio	p_value
Costante	8,246	9,255	0,891	0,374
<i>esg_contr_score</i>	-0,03167	0,01838	-1,723	0,086 .
<i>CO2_emissions</i>	-8,209E-08	0,000000056	-1,466	0,144
<i>employees</i>	0,00001635	0,000004445	3,679	0,000268 ***
<i>board_size</i>	-0,6966	0,3597	-1,937	0,054 .
<i>board_indep_members</i>	0,2385	0,08865	2,69	0,007463 **
<i>debt_on_equity</i>	0,002593	0,001254	2,067	0,039 *
<i>total_assets</i>	0,0000000048	0,0000000021	2,312	0,021 *
<i>volatility</i>	-16,11	7,667	-2,102	0,036 *
R2	0,1229			
F (8, 379)	6,638			
Adjusted R2	0,1044			
P-value	3,937E-08			

Livelli di significatività: "." 10%; \* p<0,05; \*\* p<0,01; \*\*\* p<0,001

L'intercetta stimata, che non è statisticamente significativa, non fornisce informazioni rilevanti per l'interpretazione economica del modello. La variabile *esg\_contr\_score* ha un coefficiente negativo (-0,032) ed è marginalmente significativa al livello del 10%. La variabile *CO2\_emissions* mostra un coefficiente non significativo. Pertanto, non vi è alcuna evidenza statistica che i livelli di emissioni influenzino sistematicamente la varianza nelle valutazioni ESG. Il numero di dipendenti è associato positivamente alla variabile dipendente. Anche in questo caso, considerando cambiamenti più ampi, le aziende più grandi presentano valutazioni ESG più divergenti. La dimensione del CdA mostra un coefficiente negativo marginalmente significativo al livello del 10%. Al contrario, *board\_indep\_members* mostra un coefficiente positivo (0,239) significativo al livello dell'1%. In media, un amministratore indipendente aggiuntivo è associato a un aumento del divario di circa 0,24. Il coefficiente positivo del rapporto debito/equity indica che un aumento del debito è associato, in media, a un aumento del divario tra i due punteggi ESG, a parità di tutti gli altri fattori. La variabile *total\_assets* è associata positivamente alla variabile dipendente e significativa al livello del 5%. A parità di tutte le altre variabili, un aumento della

dimensione degli asset è associato, in media, a un aumento della differenza, seppur piccolo. La variabile di volatilità mostra un coefficiente negativo (-16,11) ed è significativa al livello del 5%. In media, un aumento dell'indicatore di volatilità è associato a una riduzione della varianza tra i punteggi, al netto delle altre variabili.

Nel complesso, il modello mostra un  $R^2$  di 0,123 e un adjusted  $R^2$  di 0,104, indicando che le variabili considerate spiegano circa il 10% della variabilità della variabile dipendente. Nonostante il potere esplicativo relativamente limitato, il test F è altamente significativo ( $p < 0,001$ ), confermando che almeno una delle variabili esplicative contribuisce in modo non casuale a spiegare la variabilità della differenza. In sintesi, il modello per il campione S&P 500 suggerisce che la varianza nei punteggi ESG può essere ricondotta ad alcune caratteristiche specifiche. La complessità organizzativa (più dipendenti, asset più consistenti) sembra ostacolare valutazioni coerenti, mentre fattori di governance e struttura finanziaria (dimensioni del CdA, indipendenza, debito) incidono in modo diverso sulla coerenza delle valutazioni ESG. Infine, la volatilità del mercato sembra essere associata a una minore divergenza nei punteggi, il che suggerisce che il rischio percepito potrebbe anche influenzare le modalità con cui le agenzie valutano le pratiche di sostenibilità. Avendo già presentato i test nel caso europeo, li riassumiamo in una sola tabella.

Tabella 14 – Test diagnostici (S&P 500)

#### Multicollinearità (VIF)

Variabile	VIF
<i>esg_contr_score</i>	1,208
<i>CO2_emissions</i>	1,041
<i>employees</i>	1,124
<i>board_size</i>	1,137
<i>board_indep_members</i>	1,071
<i>debt_on_equity</i>	1,006
<i>total_assets</i>	1,12
<i>volatility</i>	1,08

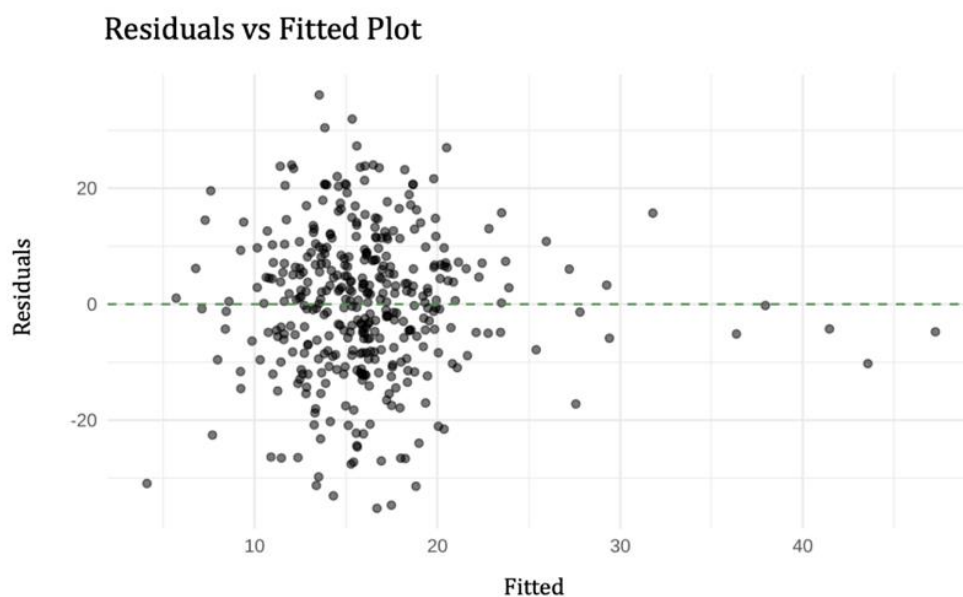
#### Test diagnostici

Test	Statistica	DoF	p_value	Risultato (5%)
Shapiro-Wilk	W=0,094	-	0,125	non rifiutata
Breusch-Pagan	BP=13,993	8	0,082	non rifiutata
RESET	F=2,185	2 - 377	0,114	non rifiutata

La tabella 14 riassume i test diagnostici eseguiti sul modello stimato. La prima parte riporta i valori del VIF per ciascun regressore: tutti sono ben al di sotto della soglia critica, indicando l'assenza di problemi di multicollinearità. Per quanto riguarda i test diagnostici, il test di Shapiro-Wilk non mostra deviazioni significative dalla normalità dei residui, il test di Ramsey RESET non fornisce alcuna indicazione di errata specificazione funzionale. Il test di Breusch-Pagan, tuttavia, restituisce un p-value di 0,082: sebbene non significativo al livello convenzionale del 5%, il valore prossimo alla soglia suggerisce una possibile debole presenza di eteroschedasticità.

L'analisi dei residui rispetto ai valori stimati (figura 6) mostra che i punti sono distribuiti in modo sostanzialmente irregolare attorno all'asse orizzontale, senza indicazione di strutture ricorrenti o andamenti curvilinei marcati. L'assenza di un chiaro ventaglio lungo l'asse delle ascisse suggerisce che la varianza degli errori rimane relativamente costante, supportando l'ipotesi di omoschedasticità. Tuttavia, si osservano alcune osservazioni isolate, caratterizzate da valori anomali. Nel complesso, i controlli diagnostici confermano l'assenza di criticità significative.

Figura 6 – Grafico dei residui rispetto ai valori stimati (S&P 500)



### 3.3. Risultati principali

Tabella 15 – Comparazione delle variabili statisticamente significative

<b>STOXX Europe 600</b>		<b>S&amp;P 500</b>	
<i>employees</i>	+ **	<i>employees</i>	+ ***
<i>board_size</i>	+ **	<i>board_size</i>	- .
<i>debt_on_equity</i>	- **	<i>debt_on_equity</i>	+ *
<i>total_assets</i>	+ .	<i>total_assets</i>	+ *
<i>return_on_assets</i>	- *	<i>esg_contr_score</i>	- .
		<i>board_indep_members</i>	+ **
		<i>volatility</i>	- *

Il confronto tra i due modelli OLS evidenzia somiglianze ma anche differenze significative tra il campione europeo e quello statunitense. Per facilitare il confronto delle variabili significative nei due campioni, la tabella 15 riporta il segno dell'effetto marginale sulla variabile dipendente (*difference*) e il livello di significatività statistica.

In entrambi i campioni, le variabili relative alla dimensione aziendale, come dipendenti e totale attivo, sono significative: l'effetto stimato è positivo, sebbene con diverse entità e livelli di significatività. Ciò dimostra che le aziende più grandi e complesse tendono a mostrare rating meno omogenei e potrebbe riflettere la difficoltà per i provider nel valutare in modo univoco aziende molto grandi e diversificate. Questo risultato conferma il lavoro di Gibson et al. (2021), che riscontra un maggiore disaccordo per le aziende più grandi.

Le differenze più marcate emergono nella governance e nella struttura finanziaria. La variabile *board\_size* mostra segni opposti nei due modelli: positiva per lo STOXX Europe 600 e negativa per lo S&P 500, suggerendo che l'impatto della composizione del CdA può variare in base al contesto di riferimento. Ciò potrebbe dimostrare che il top management in Europa ha un atteggiamento proattivo rispetto a un atteggiamento più correttivo ex post negli Stati Uniti per quanto concerne le iniziative connesse a temi ESG (Augusto et al., 2020). Inoltre, poiché i CdA statunitensi hanno una dimensione media più ampia e meno variabile, ciò potrebbe dipendere da una minore importanza marginale dell'ampliamento del consiglio. In terzo luogo, il risultato potrebbe anche implicare che, nel contesto europeo, CdA più grandi portino a una governance più complessa (Gantenbein e Volonté, 2019), il che potrebbe riflettersi in strategie ESG percepite in modo non uniforme, mentre nel contesto statunitense, CdA più grandi favoriscono una maggiore convergenza tra i punteggi ESG.

Anche il debito presenta risultati contrastanti: mentre per il campione europeo il rapporto debito/equity ha un effetto negativo sulla variabile *difference*, nel caso statunitense l'effetto stimato è positivo. Questo risultato riflette le differenze strutturali nel finanziamento delle aziende e la diversa percezione del rischio associato al debito. Innanzitutto, nei diversi settori il rapporto è diverso, quindi anche la diversa composizione dell'economia statunitense ed europea determina un rapporto diverso. Inoltre, un accesso più facile ai mercati dei capitali per le aziende statunitensi, ad esempio per l'assenza di un mercato dei capitali paneuropeo, rende il rapporto meno rilevante, producendo così strategie di gestione finanziaria molto diverse (Gaud et al., 2007). Anche diversi incentivi fiscali possono avere un impatto. Gibson et al. (2021) riscontrano un minor disaccordo per le aziende più redditizie che tendono ad allocare maggiori fondi in aree legate alla sostenibilità, riducendo così il margine di discrezionalità nella valutazione da parte delle agenzie di rating.

Per lo S&P 500, *esg\_contr\_score* ha un effetto negativo, suggerendo che un livello più elevato di controversia riduce le discrepanze tra le valutazioni ESG. La percentuale di amministratori indipendenti mostra un effetto positivo, suggerendo che una maggiore indipendenza del CdA può aumentare la variabilità delle valutazioni ESG. Infine, la volatilità è associata negativamente alla differenza solo nel campione statunitense. Ciò potrebbe indicare che il margine di discrezionalità nella valutazione è ridotto per le aziende percepite come più rischiose.

In sintesi, i due modelli confermano alcuni risultati comuni, in particolare il ruolo delle dimensioni aziendali nel determinare le deviazioni ESG, ma rivelano anche caratteristiche legate al contesto specifico. Ciò evidenzia come le dinamiche alla base della coerenza (o incoerenza) delle valutazioni ESG non siano universali, ma piuttosto influenzate dalle caratteristiche in termini di struttura dell'economia, canali di finanziamento, governance e situazione finanziaria delle aziende nei diversi contesti nazionali e settoriali. Conferma inoltre che le aziende europee sono negativamente condizionate dalla frammentazione del mercato dei capitali.

#### 4. Conclusioni: per lo sviluppo della finanza sostenibile

*Ciascun confusamente un bene apprende*  
Dante

I rating ESG rappresentano un punto di riferimento fondamentale per interpretare i comportamenti delle aziende in tema di sostenibilità e transizione. Tuttavia, l'incoerenza tra le diverse valutazioni rappresenta un ostacolo significativo al loro sviluppo. Non solo, infatti, questa divergenza solleva questioni metodologiche, ma pone anche problemi concreti per operatori finanziari e investitori, che si confrontano con giudizi difficilmente comparabili e talvolta contraddittori. L'analisi empirica evidenzia che la ragione principale di queste discrepanze risiede nella mancanza di standardizzazione nei metodi e i nostri risultati confermano tali discrepanze metodologiche come anche limiti nella disponibilità e qualità dei dati. Il nostro studio ha anche evidenziato che i fattori che determinano l'incoerenza dei rating variano tra le diverse aree economiche, con l'emergere di fattori specifici, come la governance e la struttura finanziaria. Ad esempio, nel caso europeo, una maggiore efficienza economica e finanziaria sembra ridurre l'eterogeneità delle valutazioni, mentre nel contesto statunitense, le controversie ESG e la volatilità del mercato sono fattori significativi nel ridurre la discrepanza, mentre l'indipendenza del CdA accresce l'eterogeneità. Vi sono poi fattori comuni come la dimensione. In base a queste evidenze, l'incoerenza nelle valutazioni ESG deve essere interpretata non solo come il risultato di differenze metodologiche, ma anche come espressione delle specifiche caratteristiche nazionali, e dei settori in cui operano le aziende. Incidono quindi sulle differenze anche la diversa composizione dell'output produttivo tra Stati Uniti ed Europa. La produzione industriale tradizionale, più forte in Europa, determina maggiore eterogeneità nei rating ESG rispetto ai settori digitali in cui gli Stati Uniti primeggiano. Questa conclusione è particolarmente allarmante in Europa che, nel nuovo contesto della politica internazionale, evidenzia tutta la debolezza della propria struttura di governance messa in evidenza da numerosi rapporti promossi proprio dalle istituzioni europee (Draghi, 2024, Letta, 2024).

Ora, sarebbe prematuro attendersi dai rating ESG, che hanno pochi anni e una metodologia alquanto complessa, oltre che non ancora omogenea, la stessa attendibilità dei rating ordinari che esistono da decenni (Charlin et al., 2024). È però necessario che questa attendibilità cresca perché l'incoerenza costa in termini di costo del *funding* delle imprese più sostenibili, ritardando la necessaria transizione verso un'economia sostenibile (He et al., 2025). È dunque importante che i diversi provider sviluppino definizioni condivise con l'aiuto del coordinamento pubblico. In questo senso è quindi anche decisivo lo sviluppo del dibattito scientifico su questi temi che aiuterà anche una migliore e più granulare raccolta di dati e lo sviluppo di modelli più sofisticati che, per ora, è a uno stadio iniziale, come anche mostrato nelle nostre elaborazioni.

Infine, è d'uopo un'osservazione conclusiva sul piano più politico. Le divergenze politiche tra le due sponde dell'Atlantico si stanno accentuando. Innanzitutto, ciò riguarda il quadro giuridico. Sarà particolarmente interessante analizzare se l'entrata in vigore del Regolamento UE 2024/3005 aggraverà le diverse traiettorie delle imprese statunitensi ed europee ed è stato già citato il ruolo degli standard del reporting di sostenibilità approvati nell'UE. In secondo luogo, le relazioni politiche sempre più difficili tra Stati Uniti ed Europa stanno producendo un atteggiamento divergente nei confronti della finanza sostenibile e della transizione verde. Ciò è legato alle convinzioni dei cittadini (Pouster et al., 2025), al diverso ruolo delle due aree in termini di import/export di energia (essendo gli Stati Uniti un forte esportatore e l'UE un forte

importatore di petrolio e gas) e, di conseguenza, alle capacità industriali e tecnologiche nei processi produttivi green. Nel complesso, è probabile che la divergenza dei rispettivi quadri regolamentari e di mercati aumenti, rendendo più difficile la comparabilità tra le imprese statunitensi ed europee. Questo contribuirà al *retrenching* nazionale dei mercati finanziari che caratterizza quest'epoca (Esposito et al., 2024).

## Riferimenti bibliografici

- Abhayawansa S. e Tyagi S. (2021), "Sustainable Investing: The Black Box of Environmental, Social and Governance (ESG) Ratings", *The Journal of Wealth Management*, 24(1), pp. 49-54.
- Agosto A. e Tanda A. (2025), "Divergence and Aggregation of ESG Ratings: A Survey", *Open Research Europe*, 5(28), pp. 1-13.
- Augusto M., Pascoal R. e Reis P. (2020), "Firms' Performance and Board Size: a Simultaneous Approach in the European and American Contexts", *Applied Economics Letters*, 27(13), pp. 1039-1043.
- Berg F., Kölbel J. F. e Rigobon R. (2022a), "Aggregate Confusion: The Divergence of ESG Ratings", *Review of Finance*, 26(6), pp. 1315-1344.
- Berg F., Kölbel J. F., Pavlova A. e Rigobon R. (2022b), "ESG Confusion and Stock Returns: Tackling the Problem of Noise", *NBER Working Paper*, n. 30562, Cambridge (MA): National Bureau of Economic Research.
- Capizzi V., Gioia E., Giudici G. e Tenca F. (2021), "The Divergence of ESG Ratings: An Analysis of Italian Listed Companies", *Journal of Financial Management, Markets and Institutions*, 9(02), art. 2150006.
- Charlin V., Cifuentes A. e Alfaro J. (2024), "ESG Ratings: An Industry in Need of a Major Overhaul", *Journal of Sustainable Finance & Investment*, 14(4), pp. 1037-1055.
- Chatterji A.K., Durand R., Levine D.I. e Touboul S. (2016), "Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers", *Strategic Management Journal*, 37(8), pp. 1597-1614.
- Christensen D.M., Serafeim G. e Sikochi A. (2021), "Why is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings", *The Accounting Review*, 97(1) pp. 47-175.
- Del Giudice A., Gallucci C. e Santulli R. (2024), "I rating ESG: un confronto internazionale", *FIN-GOV, Centro di ricerche finanziarie sulla corporate governance*, Milano: Università Cattolica del Sacro Cuore. [Disponibile online](#).
- Draghi M. (2024), *The Future of European Competitiveness. Part A. A Competitiveness Strategy for Europe*, Bruxelles: Commissione Europea. [Disponibile online](#).
- ESMA – European Securities and Markets Authority (2021), "Letter to the European Commission on the regulation of ESG ratings and supervision of ESG rating providers". [Disponibile online](#).
- Esposito L., Gatti G. e Mastromatteo G. (2024), "Quo Vadis Terra? The Future of Globalization between Trade and War", *Working Paper Dipartimento di Politica Economica*, n. 40, Milano: Università Cattolica del Sacro Cuore. [Disponibile online](#).
- Gantenbein P. e Volonté C. (2019), "Corporate Governance in the European Context", *Die Unternehmung*, 73(2), pp. 122-145.
- Gaud P., Hoesli M. e Bender A. (2007), "Debt-Equity Choice in Europe", *International Review of Financial Analysis*, 16(3), pp. 201-222.
- Gibson R., Krueger P. e Schmidt P. S. (2021), "ESG Rating Disagreement and Stock Returns", *ECGI Working Paper Series in Finance*, n. 65, Bruxelles: European Corporate Governance Institute. [Disponibile online](#).
- Gregory R.P. (2024), "The Influence of Firm Size on ESG Score Controlling for Ratings Agency and Industrial Sector", *Journal of Sustainable Finance & Investment*, 14(1), pp. 86-99.
- He Y., Pan Y., Shan T. e Zhou Y. (2025), "ESG Rating Disagreement and the Cost of Equity Financing", *International Review of Financial Analysis*, 107, art. 104565.
- Letta E. (2024), "Much More Than a Market", Bruxelles: Consiglio Europeo. [Disponibile online](#).
- Pouster J., Fagan M., Smerkovich M. e Preozorovsky A. (2025), "International Opinion on Global Threats", Washington (DC): Pew Research Center, 19 agosto. [Disponibile online](#).
- Stewart R. (2025), "Towards a Better Understanding of ESG Ratings", *Journal of Sustainable Finance & Investment*, 15(3), pp. 624-643.