



**UNIVERSITA' DEGLI STUDI DI PADOVA**

**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI**

**"M. FANNO"**

**CORSO DI LAUREA IN ECONOMIA**

**PROVA FINALE**

**GENOECONOMICS: COME LA GENETICA STA  
CAMBIANDO L'ECONOMIA**

**RELATORE: Marco Bertoni**


**LAUREANDO: Leonardo Da Ros**

**MATRICOLA N. 2000977**

**ANNO ACCADEMICO 2022 – 2023**

Dichiaro di aver preso visione del “Regolamento antiplagio” approvato dal Consiglio del Dipartimento di Scienze Economiche e Aziendali e, consapevole delle conseguenze derivanti da dichiarazioni mendaci, dichiaro che il presente lavoro non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere. Dichiaro inoltre che tutte le fonti utilizzate per la realizzazione del presente lavoro, inclusi i materiali digitali, sono state correttamente citate nel corpo del testo e nella sezione ‘Riferimenti bibliografici’.

*I hereby declare that I have read and understood the “Anti-plagiarism rules and regulations” approved by the Council of the Department of Economics and Management and I am aware of the consequences of making false statements. I declare that this piece of work has not been previously submitted – either fully or partially – for fulfilling the requirements of an academic degree, whether in Italy or abroad. Furthermore, I declare that the references used for this work – including the digital materials – have been appropriately cited and acknowledged in the text and in the section ‘References’.*

Firma (signature).....

# Indice

<i>Introduzione</i> .....	4
Introduzione alla genetica.....	5
<b><i>CAPITOLO I</i></b> .....	<b>8</b>
<i>L'utilizzo degli indici poligenici</i> .....	8
Applicazione: Misurare la correlazione tra geni e mobilità sociale utilizzando gli indici poligenici	9
Commenti .....	12
<b><i>CAPITOLO II</i></b> .....	<b>13</b>
<i>Imputazione mendeliana del genotipo</i> .....	13
Applicazione: Come la predisposizione genetica dei figli influenza la redistribuzione delle risorse da parte dei genitori .....	14
Commenti .....	16
<b><i>CAPITOLO III</i></b> .....	<b>18</b>
<i>L'interazione Gene-Ambiente o GxE</i> .....	18
Applicazione: La mediazione della condizione familiare di un individuo sull'effetto che il gene MAOA ha nell'apprendimento.....	21
Commenti .....	24
<i>Conclusioni</i> .....	27
<i>Bibliografia</i> .....	29

# Introduzione

La prima apparizione del termine *genoeconomics* è stata nel 2007 nell'articolo scientifico di Daniel J. Benjamin<sup>1</sup>. Come dice il nome stesso, con *genoeconomics* si intende la nuova disciplina che unisce la genetica molecolare all'economia. Dopo che recenti studi hanno portato sempre più nuove scoperte sul funzionamento del genoma umano, si è iniziato a capire come i geni potessero influenzare anche tratti più complessi, come i tratti comportamentali degli individui. Il punto di contatto tra le due materie avviene proprio quando alcuni di questi geni hanno ripercussioni sui comportamenti economici delle persone.

Questa intuizione apre la porta ad un'infinità di nuove applicazioni che possono andare dall'ambito assicurativo, quando si parla di geni che regolano l'avversione al rischio, alle politiche pubbliche, quando si tratta di andare ad assegnare dei sussidi a persone geneticamente svantaggiate che possono essere più suscettibili a certe malattie, oppure ancora nell'educazione, se si trattano geni che riguardano abilità cognitive. Le ricerche in questo campo sono ancora acerbe e, se da una parte i metodi econometrici finora utilizzati devono essere perfezionati, dall'altra il funzionamento del genoma umano nasconde ancora molte incognite.

Data la vastità delle possibili applicazioni, in questo elaborato ci si concentrerà su un aspetto in particolare, ovvero l'andamento scolastico. Nonostante sia un tratto complesso da studiare, la grande quantità di letteratura al riguardo permette di partire da delle basi solide. Ciò facilita anche la raccolta dei dati, grazie a grandi database già esistenti che contengono i risultati scolastici raggiunti o i titoli conseguiti, i parametri di misura più comuni in questi studi. Il primo capitolo di questo lavoro sarà di introduzione alle definizioni e ai termini tecnici della genetica, in modo da facilitare il lettore nella comprensione degli studi genetici analizzati. Successivamente verranno presentati tre paper, tutti riguardanti l'interazione tra i geni e l'andamento scolastico e come questo influenzi un aspetto economico dell'individuo. Sono stati scelti dei lavori che utilizzano metodi diversi perché affrontano problematiche diverse, in modo da dimostrare la grande variabilità che caratterizza gli studi di questo genere.

Il primo si focalizza sull'utilizzo dei cosiddetti "indici poligenici" per misurare l'interazione tra geni e mobilità sociale. Il secondo riguarda l'utilizzo di metodi alternativi, come l'imputazione mendeliana del genotipo, quando si è in mancanza di alcuni dati genetici fondamentali e si concentra sulle differenze di redistribuzione delle risorse dei genitori a figli che hanno diverse dotazioni genetiche. L'ultimo tratta di uno dei maggiori problemi che affligge gli studi sulla

---

<sup>1</sup> BENJAMIN, D., et al., 2007 "*Genoeconomics*"

genoeconomics, ovvero la difficoltà di misurare l'effetto dell'interazione Gene-Ambiente. In particolare, verrà presentato come l'ambiente familiare faccia da mediatore alla manifestazione del gene MAOA. Infine, oltre a riportare i risultati e le difficoltà riscontrate in questi paper, nelle considerazioni finali verranno anche analizzati i problemi di natura etica che scaturiscono da questi studi.

## **Introduzione alla genetica**

La genetica nasce verso la fine dell'Ottocento grazie agli studi di Gregor Mendel, il quale è riuscito a dimostrare che le caratteristiche di un individuo sono determinate da fattori ereditari chiamati geni. Questo non era che l'inizio di una serie di scoperte che avrebbero portato a una visione più ampia della genetica e delle sue implicazioni che ad oggi viene definita come la scienza "che si occupa del materiale ereditario, cioè della sua struttura, del suo modo di funzionare e delle modalità della sua trasmissione"<sup>2</sup>.

Negli ultimi anni, gli studi sulla genetica si sono sviluppati tanto da interessare anche gli economisti, che vedono un grande potenziale nella fusione tra genetica ed economia: la genoeconomics. Nel lavoro di Daniel J. Benjamin (2007) vengono analizzati alcuni dei vantaggi che gli economisti potrebbero trarre se fossero in grado di integrare ai loro studi i concetti di ereditabilità dei geni, la correlazione tra i geni di un individuo, e la loro manifestazione, in quanto si potrebbero ricondurre certi comportamenti economici a un particolare gene o combinazioni di geni. Per portare degli esempi pratici, vi è uno specifico gene che indica la suscettibilità al cancro ai polmoni di un individuo. Il rischio di svilupparlo aumenta di molto se si fa uso di sigarette: una scelta di consumo influenzata dal loro prezzo, dalle accise e dal reddito dell'individuo. In altri casi gli interventi pubblici potrebbero intervenire diminuendo o aumentando le differenze genetiche tra gli individui. Quando si va a sussidiare degli individui che hanno contratto certe malattie si cerca di spianare le differenze prodotte dai geni delle persone e, al contrario, quando si premiano degli individui per meriti scolastici o lavorativi, si incentivano persone con abilità cognitive migliori e quindi già avvantaggiati geneticamente rispetto ad individui con maggiori difficoltà. Gli studi sulla Genoeconomics possono quindi aiutare a riconoscere le differenze genetiche tra le persone e aiutare a rendere le politiche pubbliche più efficienti.

---

<sup>2</sup> Internet, in Treccani.it – Vocabolario Treccani on line, Istituto dell'Enciclopedia Italiana.

Prima di proseguire oltre, è utile che vengano chiariti alcuni concetti che stanno alla base della genetica. Il gene è l'unità fondamentale, definita come il fattore ereditario che regola una particolare caratteristica del soggetto. L'insieme di tutte queste informazioni, cioè il cosiddetto patrimonio genetico, è immagazzinato nel DNA: il gene non è altro che un tratto di DNA che codifica per una specifica informazione. I lunghi filamenti di DNA delle nostre cellule vengono chiamati cromosomi e possono contenere fino a migliaia di geni l'uno. Nel caso dell'essere umano, ogni cellula possiede un totale di 46 cromosomi divisi per coppie, uno ereditato dalla madre e l'altro dal padre. Nonostante una coppia di cromosomi codifichi per gli stessi caratteri, vi possono essere differenze all'interno dello stesso gene chiamate alleli: la combinazione di due alleli che codifica per un carattere viene detta genotipo, mentre la manifestazione direttamente osservabile del genotipo viene detta fenotipo. Le modalità in cui il genotipo si manifesta nel fenotipo dipendono dagli alleli che lo codificano: se essi sono uguali (omozigosi) il fenotipo è uno solo ed è quello indicato dal genotipo, ma se gli alleli sono differenti (eterozigosi) vi sono diverse possibili combinazioni che portano a diverse manifestazioni del carattere osservato.



Una prima difficoltà riguarda proprio la variabilità degli alleli: nel processo di copia e formazione di nuovo DNA possono avvenire piccole variazioni del codice genetico che però non influenzano il fenotipo riscontrato. Per gli scienziati diventa quindi difficile identificare una specifica sequenza per un determinato carattere in quanto ci possono essere numerosissime variazioni anche di un solo nucleotide che si manifestano con lo stesso fenotipo. Se sono presenti nella popolazione in proporzione maggiore dell'1% queste variazioni vengono chiamate Polimorfismo a Singolo Nucleotide (o SNP)<sup>3</sup>. Le difficoltà aumentano se si vanno a studiare dei caratteri che sono influenzati da più geni o da alcune combinazioni di essi. Un'altra complessità è dovuta all'interazione tra gene e ambiente cioè come l'ambiente è in grado di influenzare la manifestazione di alcuni geni e viceversa. Una determinata caratteristica

---

<sup>3</sup> BROOKES, A, 1999, "The essence of SNPs"

ambientale può inibire o accentuare la manifestazione di un gene e, al contrario, certi geni possono contribuire a cambiare l'ambiente circostante. Non essendo due fattori indipendenti tra loro è necessario sviluppare dei metodi che tengano conto delle diverse forme di correlazione che si possono presentare. Nei prossimi capitoli verranno difatti presentati diverse metodologie per ovviare a questi problemi ed isolare l'effetto che hanno i geni nel condizionare alcune caratteristiche di un individuo.

# CAPITOLO I

## L'utilizzo degli indici poligenici

Come detto in precedenza, uno dei maggiori problemi in questo tipo di ricerche è dovuto alla variabilità del DNA che per lo stesso gene può arrivare a formare tanti diversi SNP. Considerare un singolo SNP per la propria analisi è controproducente in quanto il suo potere esplicativo è davvero molto piccolo. Gli studiosi hanno quindi cominciato a raggruppare gli SNP in quelli che vengono chiamati indici poligenici. Un indice poligenico in pratica è la somma ponderata dei diversi SNP che ha la funzione di prevedere la propensione di un individuo a manifestare un certo carattere (Genç, et al, 2021)<sup>4</sup>. La formula è la seguente:

$$PGI_i = \sum_{j=1}^J \beta_j^{GWAS} x_{ij},$$

Dove  $x_{ij}$  è il numero di copie dell'allele dell'individuo  $i$  con SNP  $j$ , mentre  $\beta_j^{GWAS}$  è il coefficiente  $\beta$  dell' SNP  $j$ . Assegnando un peso ( $\beta_j^{GWAS}$ ) ad ogni SNP (con  $x_{ij}$  che assume i valori 0,1 o 2), gli SNP con effetti maggiori vengono pesati di più di quelli con effetti minori. La precisione di questi indici dipende dall'ampiezza del campione di riferimento e quindi, in questo caso, dalla raccolta di dati sul patrimonio genetico di un maggior numero di individui. Nascono così gli studi di associazione genome-wide (GWAS) che hanno l'obiettivo di mappare il patrimonio genetico di un gran numero di soggetti e individuare e catalogare le differenze con cui uno stesso gene si presenta in diverse persone. L'uso combinato degli indici poligenici e dei GWAS permette di stimare in maniera più precisa la dipendenza di un particolare fenotipo anche se esso può dipendere da diversi SNP.

---

<sup>4</sup> GENÇ E., et al., 2021, "Polygenic Scores for Cognitive Abilities and Their Association with Different Aspects of General Intelligence—A Deep Phenotyping Approach"



## **Applicazione: Misurare la correlazione tra geni e mobilità sociale utilizzando gli indici poligenici**

Viene ripreso questo procedimento nel lavoro di Belsky (2018)<sup>5</sup>, nel quale vengono utilizzati dei “poligenic scores” (o PGS) per misurare la mobilità sociale dei figli rispetto ai genitori. In questo studio sono stati utilizzati dei PGS legati all’istruzione di più di 20000 individui per scoprire se a PGS più alti, e quindi una maggiore predisposizione all’apprendimento, corrispondessero maggiori successi nella carriera scolastica e lavorativa e infine un avanzamento di classe sociale. Vengono formulate tre ipotesi e quindi tre diversi modi per valutare come il genoma di un individuo sia correlato alla mobilità sociale. Per prima cosa viene specificato che sarebbe impossibile trovare dei geni che influenzino direttamente il livello di educazione di un individuo, ma, al contrario, si possono trovare dei geni che codificano dei tratti o comportamenti che a loro volta hanno un effetto positivo sull’educazione (ad esempio un maggior sviluppo di certe aree del cervello). Una seconda spiegazione prevede che il patrimonio genetico di un individuo, in quanto ereditato dai suoi genitori, sia influenzato dai geni che hanno permesso ai suoi genitori di arrivare ad una certa classe sociale. Inoltre, quest’individuo erediterebbe anche l’ambiente della classe sociale dei genitori. Se l’effetto dell’ambiente fosse quello preponderante sarebbe quindi l’ambiente ereditato dai genitori l’elemento di connessione tra genetica e mobilità sociale. Una terza teoria prevede invece che la correlazione tra geni e istruzione sia spuria perché influenzata da un’“eredità sociale”. Ad esempio, perché posizioni sociali privilegiate sono state tramandate di generazione in generazione o perché le persone tendono a sposarsi con altre persone provenienti dallo stesso ceto sociale.

Il primo campione preso come riferimento è quello dell’Environmental Risk Longitudinal Twin Study (E-Risk) che contiene 1860 gemelli, omozigoti ed eterozigoti, nati tra il 1994-95 in Inghilterra e Galles. La raccolta dei dati si divide in due fasi: viene fatto prima un test all’età di 5 anni e successivamente un altro al raggiungimento dei 18 anni di età per valutare l’andamento scolastico. Il livello di istruzione viene misurato sia ai figli che ai genitori tramite il General Certificate of Education Examination, un indice che va da 0 a 4 a seconda dei risultati raggiunti. Per la costruzione dei PGS è stato utilizzato il software PRsice basato sui dati del Social Science Genetic Association Consortium, un GWAS che raccoglie i geni che influenzano il livello di istruzione.

In pratica gli SNP presenti nel database E-Risk sono stati abbinati ai risultati pubblicati dal

---

<sup>5</sup> BELSKY, D., et al., 2018, “*Genetic analysis of social-class mobility in five longitudinal studies*”

GWAS: per ogni SNP è stato assegnato un valore (ad esempio per il numero di alleli legati all'istruzione che contiene) che viene moltiplicato per la dimensione dell'effetto stimato dal GWAS. Infine, si ottengono i PGS calcolando la media ponderata e utilizzando la dimensione dell'effetto come peso. Viene calcolato il coefficiente di correlazione tra PGS e la carriera scolastica e i risultati indicano che c'è una correlazione positiva ( $r=0.27$ ,  $P<0,001$ ).

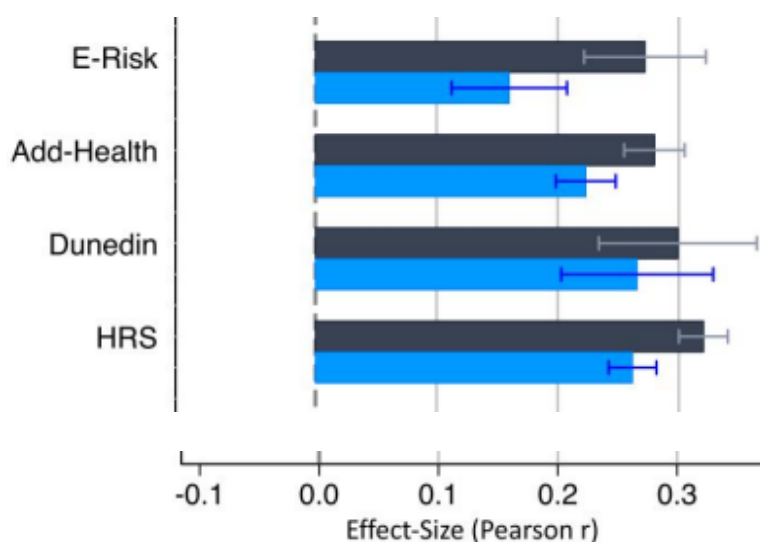
Come detto in precedenza, questo risultato potrebbe essere distorto in quanto influenzato da altri fattori esterni di cui non si è tenuto conto. Infatti, molti dei soggetti con PGS alti provenivano da famiglie con un'educazione migliore: questo potrebbe significare che i figli ottengono migliori risultati anche a causa del livello di educazione dei genitori e non solamente grazie alla propria dotazione genetica. Quando infatti si calcola la correlazione tra PGS del figlio e i loro successi, tenendo conto però anche dell'istruzione dei genitori, il coefficiente di correlazione indica sempre una correlazione positiva ma molto inferiore al precedente ( $r=0.16$ ). Il passo successivo è stato quello di studiare le differenze tra gemelli eterozigoti. Nei gemelli omozigoti il patrimonio genetico è lo stesso, ma negli eterozigoti è diverso e quindi possiedono differenti PGS. Essendo gemelli, condividono lo stesso background familiare e scolastico e ciò permette di tenere già conto di influenze esterne che nei calcoli precedenti sarebbe stato impossibile considerare. Viene quindi stimato un nuovo coefficiente chiamato SD (Siblings-Difference) su un minor gruppo di gemelli, pari a 388 coppie, che testava se il fratello con un polygenic score più alto ottenesse risultati migliori rispetto a quello con il punteggio più basso. Il risultato viene poi standardizzato per renderlo confrontabile con i valori dell'intero campione e risulta essere in linea con quello già ottenuto con un  $r=0.13$ .

Come si può notare il coefficiente di correlazione, tenuto conto di diverse variabili di controllo, è quasi dimezzato da quello iniziale: ciò indica che la relazione tra geni e mobilità sociale è in realtà mediata da altri fattori e potrebbero essercene altri da tenere in considerazione. In particolare, si ipotizza che il patrimonio genetico dei genitori possa contribuire a formare l'ambiente in cui i figli crescono e quindi che possa influenzarne l'andamento scolastico.

Per tenere conto di ciò, viene condotta un'analisi sociogenetica ovvero un'analisi sugli effetti che il genotipo di una persona ha sul fenotipo di un'altra. In questo caso si cerca di capire se i geni di una madre sono in grado di prevedere l'educazione del figlio indipendentemente dai geni del figlio. Se il patrimonio genetico di una madre fosse in grado di prevedere il livello di educazione del figlio vorrebbe dire che l'ambiente creato dalla madre influenza esternamente l'andamento scolastico del figlio. Al contrario, se non fosse in grado di prevederlo, indicherebbe che l'unico effetto che ha la madre sul figlio sono i geni che esso ha ereditato, senza altre influenze esterne. Dai risultati ottenuti si trova che vi è correlazione tra l'ambiente esterno influenzato dalla madre e il livello di educazione del figlio: i figli con madri che avevano un

PGS più alto ottenevano migliori risultati indipendentemente dalla loro dotazione genetica. In questa analisi si ottiene un  $r$  aggiustato di 0.12, un valore che è molto inferiore al primo test, ma comunque in linea con i risultati ottenuti dai successivi in cui si andava a controllare per altri fattori.

Per cercare di generalizzare e confrontare i risultati ottenuti su questo campione, tutti i test precedentemente esposti vennero ripetuti anche per altri 4 gruppi di figli e genitori ovvero per l'Add Health Study (5526 soggetti negli USA), Dunedin (831 soggetti in Nuova Zelanda), WLS (7111 soggetti in USA, dal diploma fino ai 60 anni) e HRS (8546 soggetti in USA, fino ai 70 anni). Come si può vedere nella figura riportata sotto, anche negli altri campioni si ottengono indici simili: nel grafico il grigio indica la correlazione tra poligenic score dell'individuo e conseguimento di titoli scolastici, mentre in azzurro la correlazione tra poligenic score dell'individuo e mobilità sociale.



I risultati ottenuti sono coerenti con l'ipotesi che i geni legati all'istruzione siano dei predittori del successo sociale. Infatti, nei diversi campioni si sono ottenuti indici simili che indicano come i bambini con PGS più alti ottengono maggior successo nella carriera scolastica e lavorativa, nonostante rimanga il fatto che questi risultati sono molto influenzati anche dall'ambiente in cui crescono. Ci sono però dei limiti di cui bisogna tenere conto. I geni selezionati per realizzare i PGS sono solo una parte dell'influenza che i geni hanno sull'istruzione e ciò è dovuto alle innumerevoli complessità che sorgono quando si va a studiare il genoma di un essere umano, soprattutto quando si parla di caratteri complessi come il livello di istruzione che possono derivare da moltissimi geni differenti e da più combinazioni di essi. Inoltre, i dati utilizzati sono cronologicamente molto diversi tra loro, con differenze di decenni tra gli anni di nascita dei partecipanti di un campione e l'altro: vi è il rischio che nel corso degli

anni le condizioni sociali per raggiungere migliori risultati a scuola e nel lavoro non siano rimaste le stesse.

## **Commenti**

I risultati mostrati aiutano a chiarire l'importanza che il genoma umano ha per gli economisti quando questo influenza alcuni fattori che hanno un impatto economico. Dal punto di vista dell'economia politica lo studio dei poligenic score come proxy dell'apprendimento potrebbe fornire indicazioni sul livello di istruzione della popolazione. Ciò faciliterebbe la creazione di interventi pubblici mirati e quindi più efficaci ed efficienti. Questi interventi non possono chiaramente andare a modificare le caratteristiche genetiche dei bambini, ma possono fornire ambienti stimolanti ed educativi agli individui che ne hanno maggiormente bisogno, contribuendo a spianare il divario tra chi è geneticamente "avvantaggiato" e chi non lo è. Inoltre, è probabile che questi geni siano correlati ad altri fattori oltre che all'educazione, come l'andamento della carriera lavorativa o la qualità della vita: questo studio apre la via a numerosi altri approfondimenti.

## CAPITOLO II

### Imputazione mendeliana del genotipo

Nel caso precedente, oltre al genoma dei figli, era stato catalogato anche quello dei genitori. In molti database quest'ultimo manca, e non potendo andare a campionare il DNA di ogni genitore, bisogna stimarlo. Il lavoro di Young<sup>6</sup> introduce un metodo in grado di stimare gli alleli trasmessi dai genitori al figlio, isolando i possibili effetti esterni che potrebbero condizionare l'espressione del gene del figlio e senza aver campionato il genoma dei genitori. In pratica viene ripreso il metodo di imputazione utilizzato da Mendel nei suoi esperimenti con cui, grazie a una semplice tabella, si è in grado di calcolare i possibili esiti delle combinazioni dei geni dei genitori. Ad esempio, se i genitori per lo stesso carattere presentano gli alleli AA e AB, il figlio erediterà sicuramente A dal primo genitore e A o B dal secondo genitore con una probabilità del 50%. Vengono riportati in questa tabella tutti i possibili risultati: con il 25% di probabilità potrà presentare la forma allelica omozigote AA o BB, mentre al 50% potrà presentare la forma eterozigote AB o BA.

Padre/Madre	A	B
A	AA	AB
B	BA	BB

Il metodo di Mendel serve per prevedere i possibili risultati della combinazione dei geni dei genitori nel figlio, ma in questo caso verrà utilizzato all'inverso: conoscendo il genoma del figlio si risale a quello dei genitori. Ad esempio, se il figlio presenta gli alleli AA entrambi i genitori hanno dato al figlio l'allele A, se presenta gli alleli AB invece avrà preso dal padre l'allele A e dalla madre l'allele B.

Questo metodo diventa fondamentale nei casi in cui i database forniti dai GWAS non dispongono di informazioni sul genoma dei genitori, come nello studio di cui si andrà a discutere di seguito.

---

<sup>6</sup> YOUNG, A.I., et al., 2022, "Mendelian imputation of parental genotypes improves estimates of direct genetic effects."

## Applicazione: Come la predisposizione genetica dei figli influenza la redistribuzione delle risorse da parte dei genitori

L'obiettivo della ricerca condotta da Sanz de Galdeano e Terskaya (2022)<sup>7</sup> è quello di analizzare le scelte di investimento dei genitori rispetto ai figli che possiedono una maggiore o minore predisposizione genetica all'apprendimento: si tratta quindi di stimare l'effetto che le diverse dotazioni genetiche dei figli hanno sulle scelte redistributive dei genitori. Per calcolare ciò si specificano le preferenze di investimento dei genitori con la seguente funzione di utilità:

$$U = \{V_1^p + V_2^p\}^{\frac{1}{p}}$$

Con il vincolo di bilancio:

$$p_1PI_1 + p_2PI_2 = I$$

Per semplicità viene considerato il caso con soli due figli:  $V_1$  indica il capitale umano per il primo figlio, mentre  $V_2$  quello per il secondo. La variabile  $p$  può essere interpretata con il grado di avversione alla disuguaglianza dei genitori. Se  $0 < p < 1$  i genitori non sono avversi alla disuguaglianza e attueranno una strategia di rinforzamento verso il figlio con il capitale umano più alto: preferiscono quindi l'efficienza all'uguaglianza. Invece, quando  $p < 0$  i genitori terranno maggiormente conto dell'uguaglianza e decideranno quindi di investire maggiormente nel figlio con il capitale umano minore. Con  $p=0$  i genitori sono indifferenti nelle loro scelte di investimento. Queste conclusioni sono importanti perché permettono di introdurre una legge matematica per scelte di trade off tra uguaglianza ed efficienza che i genitori possono compiere.

Il modello econometrico utilizzato per stimare il coefficiente  $p$  è il seguente

$$PI_{if} = \beta_0 + \beta_1(g_{if} - g_{jf}) + \beta_2g_{if} + \beta_3g_{pf} + X'_{if}\alpha + S'_{jf}\delta + F'_{fy} + u_{if}$$

Dove  $PI_{if}$  è l'indicatore dell'investimento dei genitori per il figlio  $i$  nella famiglia  $f$ ,  $g_{if}$  misura la dotazione genetica del figlio  $i$ ,  $g_{jf}$  la stessa misura per il figlio  $j$ , mentre  $g_{pf}$  indica la dotazione genetica dei genitori. Le caratteristiche individuali del figlio  $i$  e del figlio  $j$  sono misurate rispettivamente da  $X'_{if}$  e  $S'_{jf}$ ,  $F'_{fy}$  misura invece le caratteristiche della famiglia che sono quindi condivise dai figli. Dato che  $g_{if} - g_{jf}$  non è altro che la dotazione genetica relativa del figlio  $i$

---

<sup>7</sup> SANZ DE GALDEANO, Anna, TERSKAYA, Anastasia, 2019. "Sibling Differences in Educational Polygenic Scores: How Do Parents React?".

rispetto al figlio  $j$ , il coefficiente  $\beta_1$  misura quindi l'effetto delle preferenze d'investimento dei genitori a causa della differente dotazione genetica dei figli. Con  $\beta_1 < 0$  i genitori saranno avversi all'ineguaglianza ( $p < 0$ ). Con  $\beta_1 > 0$  si avranno dei genitori che scelgono di investire in maniera efficiente ( $0 < p < 1$ ), mentre con  $\beta_1 = 0$  i genitori saranno indifferenti tra efficienza e uguaglianza ( $p = 0$ ).

Il coefficiente  $\beta_2$  rappresenta l'effetto del costo dell'investimento: con  $\beta_2 > 0$  i genitori investiranno di più nel figlio con la dotazione genetica migliore perché è meno costoso investire in essi piuttosto che nei figli con una dotazione genetica peggiore.

Il campione di riferimento utilizzato è quello contenuto nel National Longitudinal Study of Adolescent to Adult Health. Nel 1994 sono state intervistate 3139 coppie di fratelli di età compresa tra i 12 e i 20 anni e l'intervista è stata ripetuta più volte nel '96, 2001/02, nel 2008 e nel 2016/18. Solo a partire dal 2008 sono stati raccolti i campioni di saliva di 1092 coppie di soggetti per mapparne il DNA. Di queste sono state prese 619 coppie di cui 192 gemelli e 412 non gemelli. Per ognuno di essi viene calcolato un poligenic score legato all'istruzione per avere una misura della dotazione genetica dei ragazzi (la "g<sub>i</sub>" o "g<sub>j</sub>" nel modello). Non è stata effettuata nessuna raccolta di DNA per i genitori e quindi non si è in grado di ottenere direttamente la "g<sub>p</sub>" del modello: per colmare questa mancanza ecco che viene utilizzato il metodo dell'imputazione mendeliana del genotipo. Come spiegato in precedenza, andando a vedere il codice genetico del figlio si è in grado di risalire a quello dei genitori, isolando i fattori esterni che potrebbero aver influenzato la manifestazione del carattere studiato. Una volta ottenuto il codice genetico dei genitori, viene calcolato anche il loro poligenic score. Calcolate le dotazioni genetiche di figli e genitori rimane da stimare quanto i genitori investono nei propri figli. La strategia utilizzata è stata quella di porre delle domande ai figli riguardo al comportamento dei genitori, ad esempio se nell'ultimo mese fossero andati al cinema, teatro, museo, concerto o evento sportivo con la madre/padre, se la madre/padre avesse dato loro una mano con i compiti per casa o con dei progetti scolastici, quante volte alla settimana cenano con la propria madre/padre ed altri quesiti di questo tipo. Assegnando un punteggio ad ogni risposta è stato possibile costruire un indice per la madre ed uno per il padre, la cui somma ha dato una stima dell'investimento dei genitori verso i propri figli.

Sostituendo i risultati nel modello si ottengono le seguenti stime per la differenza tra livello d'istruzione tra i due figli (EA PGI- Sibling's EA PGI in tabella), per il livello d'istruzione del primo figlio (EA PGI) e per il livello di istruzione dei genitori (Parental PGI).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Pooled Sample			Non-Twins			Twins		
<u>EA PGI-Sibling's EA PGI</u>	-0.232	-0.233	-0.226	-0.265	-0.270	-0.256	-0.050	-0.043	-0.052
SE	(0.084)	(0.081)	(0.082)	(0.101)	(0.100)	(0.102)	(0.134)	(0.129)	(0.138)
<i>p</i> – value	[0.006]	[0.004]	[0.006]	[0.009]	[0.007]	[0.012]	[0.708]	[0.737]	[0.710]
<u>EA PGI</u>	0.205	0.212	0.201	0.168	0.174	0.148	0.245	0.266	0.273
SE	(0.080)	(0.077)	(0.086)	(0.093)	(0.095)	(0.117)	(0.135)	(0.123)	(0.130)
<i>p</i> – value	[0.011]	[0.006]	[0.019]	[0.074]	[0.067]	[0.206]	[0.071]	[0.032]	[0.037]
<u>Parental EA PGI</u>	-0.001	-0.001	-0.002	0.011	0.023	0.008	-0.011	-0.014	-0.023
SE	(0.073)	(0.074)	(0.089)	(0.089)	(0.093)	(0.123)	(0.110)	(0.107)	(0.116)
<i>p</i> – value	[0.984]	[0.994]	[0.978]	[0.903]	[0.800]	[0.945]	[0.921]	[0.893]	[0.846]
<i>N</i>	604	604	604	412	412	412	192	192	192

Le colonne dall'1 al 3 riportano le stime dei coefficienti  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  del campione studiato, dapprima senza controlli (colonna 1), poi aggiungendo dei controlli di base (colonna 2) e poi aggiungendo ulteriori controlli (colonna 3). Le colonne dal 4 al 6 e dal 7 al 9 ripetono lo stesso procedimento rispettivamente per i sottocampioni di coppie di fratelli che non sono gemelli e poi per i gemelli. Nella colonna 3 il coefficiente  $\beta_1$ , ovvero la differenza della dotazione genetica legata all'istruzione tra un figlio rispetto all'altro, è significativo e minore di zero.

Come anticipato questo risultato è conforme con l'ipotesi di genitori avversi alla disuguaglianza e in particolare, con l'aumento di una unità della differenza tra la dotazione genetica dei due figli, l'investimento dei genitori verso il figlio più dotato diminuirà di 0.233. Il risultato che si ottiene è simile nel sottocampione dei non gemelli, ma non risulta significativo in quello dei gemelli. L'effetto prezzo è rappresentato dal coefficiente  $\beta_2$  ovvero la dotazione genetica del figlio preso in considerazione. Sempre in riferimento alla colonna 3 si nota come questo sia positivo: ciò sta a significare che è più conveniente investire nel figlio maggiormente dotato e se questo effetto è abbastanza grande può indirizzare anche i genitori avversi alla disuguaglianza a cambiare le loro preferenze.

## Commenti

Dai risultati ottenuti si può affermare che i genitori delle famiglie americane presi in considerazione siano avversi all'ineguaglianza tra i figli, preferendo investire nel figlio con un livello di educazione più basso rispetto all'altro e scegliendo così una strategia incentrata sull'uguaglianza. Quest'affermazione è vera nel campione completo e nel sottocampione dei figli non gemelli, ma non vale invece per i figli gemelli. Questo può essere spiegato dal fatto



che i genitori di due gemelli, o più in generale di figli con poca differenza di età, facciano fatica ad investire solo su un figlio ed escludere l'altro. Ad esempio, nei momenti in cui un genitore gioca, fa un'attività o aiuta con i compiti il proprio figlio è difficile non includere anche l'altro figlio, rendendo arduo distinguere l'investimento posto in uno rispetto all'altro. Queste considerazioni sono importanti per aumentare l'efficacia delle politiche pubbliche che, investendo maggiormente nei figli riconosciuti come geneticamente meno dotati, riducono le inefficienze dovute alla redistribuzione delle risorse attuata dai genitori avversi alla disuguaglianza. Nei casi invece di genitori che preferiscono l'efficienza o con un effetto prezzo che supera l'avversione dei genitori alla disuguaglianza, queste politiche otterranno un effetto ancora maggiore incoraggiando i genitori ad investire ancora di più nei figli con una dotazione genetica peggiore. Anche in questo caso c'è da ribadire che i risultati ottenuti si basano su dei PGS che, per come sono costruiti, sono delle stime e quindi conducono a dei risultati che possono essere imprecisi. Solo ulteriori studi e nuove tecniche saranno in grado di ottenere stime più precise.

## CAPITOLO III

### L'interazione Gene-Ambiente o GxE

La difficoltà maggiore riscontrata negli studi precedentemente esposti e, in generale nella maggioranza dei lavori di questo genere, è quella di isolare l'influenza che l'ambiente ha sulla manifestazione dei geni per ottenere dei risultati che siano meno distorti possibili. Il primo passo è quindi quello di classificare in maniera empirica i possibili scenari che si possono creare a seconda delle caratteristiche di G (Gene) ed E (Environment).

Le ricerche di Biroli<sup>8</sup> (2022) hanno portato alla compilazione di una tabella riassuntiva in cui, nelle righe, si prendono in considerazione tre diversi tipi di dati sul genotipo: la prima in cui G è esogeno perché sono disponibili dati anche per la famiglia e con un poligenic score basato sia sul genotipo del figlio che quello del genitore, la seconda in cui G è esogeno, ma il PGS è costruito solo sul genotipo del figlio e la terza in cui G è endogeno perché non si hanno i dati dei genitori. Le colonne invece riguardano le tre tipologie di E: esogeno, endogeno e predeterminato. Con “predeterminato” si intende un ambiente non correlato con il genotipo G, ma correlato con altre caratteristiche ambientali o con il genotipo dei genitori (per esempio l'inquinamento atmosferico o il reddito familiare).

Di seguito viene riportata la tabella presa in considerazione e successivamente verrà spiegato cosa comporta, ai fini della distorsione dei risultati, ogni tipologia di G ed E.

---

<sup>8</sup> BIROLI, Pietro, et al., 2022, “*The Economics and Econometrics of Gene-Environment Interplay*”

	E esogeno	E endogeno	E predeterminato
G esogeno e PGS di genitori e figlio	<ul style="list-style-type: none"> <li>- G non distorto</li> <li>- E non distorto</li> </ul>	<ul style="list-style-type: none"> <li>- G non distorto</li> <li>- E potrebbe essere influenzato da altre caratteristiche ambientali o dall'interazione GxE</li> </ul>	<ul style="list-style-type: none"> <li>- G non distorto</li> <li>- E potrebbe essere influenzato da altre caratteristiche ambientali</li> </ul>
G esogeno e PGS del figlio	<ul style="list-style-type: none"> <li>- G distorto verso il basso: errore di misurazione dell'influenza genetica intrafamiliare che porta a una sovrastima di G dei genitori</li> <li>- E non distorto</li> </ul>	<ul style="list-style-type: none"> <li>- G distorto verso il basso: errore di misurazione dell'influenza genetica intrafamiliare che porta a una sovrastima di G</li> <li>- E potrebbe essere influenzato da altre caratteristiche ambientali o dall'interazione GxE</li> </ul>	<ul style="list-style-type: none"> <li>- G distorto verso il basso: errore di misurazione dell'influenza genetica intrafamiliare che porta a una sovrastima di G</li> <li>- E potrebbe essere influenzato da altre caratteristiche ambientali</li> </ul>
G endogeno e PGS del figlio	<ul style="list-style-type: none"> <li>- G distorto verso l'alto: potrebbe essere influenzato da altre caratteristiche ambientali o dal genoma dei genitori</li> <li>- E non distorto</li> </ul>	<ul style="list-style-type: none"> <li>- G distorto verso l'alto: potrebbe essere influenzato da E, altre caratteristiche ambientali o dal genoma dei genitori</li> <li>- E potrebbe essere influenzato da altre caratteristiche ambientali, dal genoma dei genitori o dall'interazione GxE</li> </ul>	<ul style="list-style-type: none"> <li>- G distorto verso l'alto: potrebbe essere influenzato da altre caratteristiche ambientali (predeterminate) o dal genoma dei genitori</li> <li>- E potrebbe essere influenzato da altre caratteristiche ambientali o dal genoma dei genitori</li> </ul>

In altro a sinistra nella tabella appare un caso particolare, quello in cui G ed E sono entrambi esogeni. Quasi tutti gli ambienti osservati sono endogeni a causa di una miriade di fattori che potrebbero influenzare l'ambiente di un individuo, ma che non vengono considerati: questa condizione viene difatti definita come "l'esperimento ideale". Invece, per ottenere un G esogeno sono necessari i dati del genoma dei genitori oltre che quello dei figli. Come già visto, il genoma di un figlio è la combinazione casuale dei genomi dei genitori. Tenendo conto dell'influenza che il genoma parentale ha sui figli nel calcolo degli indici poligenici, si possono così ottenere dei risultati che non sono distorti. Nel caso di un G esogeno (quindi si ha a disposizione il genoma dei genitori), ma con un poligenic index che non è calcolato tenendo conto dell'influenza genetica parentale si ottengono dei risultati che non sono corretti. Se in un secondo momento si va a correggere l'indice, tenendo conto del genoma dei genitori, si va a sovrastimare il termine di errore perché il genotipo dei genitori contiene l'influenza che esso ha sul genoma del figlio, ma anche una parte degli effetti genetici diretti che eredita il figlio. Un termine di errore troppo grande determina infine un risultato di G che è sottostimato.

A maggior ragione si ottengono misure distorte quando G è endogeno e non si hanno informazioni sul genoma dei genitori. Generalmente i ricercatori cercano di stimare l'influenza dei genitori includendo nei controlli un gran numero di componenti genetiche che sono comuni nel tipo di popolazione utilizzata come campione. Nonostante ciò, le stime di G continuano ad essere distorte verso l'alto in quanto non si riescono ad isolare correttamente l'influenza che l'ambiente e il genoma dei genitori hanno su G.

Nella seconda colonna vi sono i casi in cui E è endogeno cioè E è correlato con il termine di errore nel modello che si sta regredendo. Ciò può essere dovuto da causalità inversa, distorsione da variabile omessa, errore di misurazione e correlazione tra il campione del GWAS scelto e l'ambiente. In pratica quello che succede è che E potrebbe contenere anche gli effetti di altri ambienti che sono correlati ad E oppure che E venga modellato da G tramite l'interazione GxE. E funge quindi da elemento di mediazione tra G e il fenotipo che si vuole studiare, diventando quindi una variabile correlata con G stesso e il termine di errore. Diverso è il caso di E predeterminato cioè di caratteristiche ambientali che sono correlate con l'ambiente di studio. Solitamente vengono prese in considerazione le caratteristiche dell'ambiente infantile del figlio che possono influenzare la manifestazione di determinati caratteri, come il tasso di mortalità nella zona di interesse, il livello di infrastrutture e servizi o il reddito familiare. La difficoltà nell'isolare il loro effetto sta nel fatto che solitamente sono raggruppate e correlate fra di loro. Ad esempio un alto tasso di disoccupazione tende a presentarsi in aree con meno strutture e quindi in ambienti più poveri, mentre a più alti livelli di istruzione dei genitori corrisponde un maggior reddito familiare e quindi un'area più ricca. Considerare solo una di queste

caratteristiche senza tener conto di tutte le altre ad essa correlate significherebbe trascurare una parte di ambiente E che influisce sulla manifestazione G dell'individuo.

### **Applicazione: La mediazione della condizione familiare di un individuo sull'effetto che il gene MAOA ha nell'apprendimento**

Thompson<sup>9</sup>(2014) nel suo lavoro ha voluto andare a studiare l'effetto di mediazione che l'ambiente ha nella manifestazione di un gene e gli effetti dell'interazione GxE. Come visto in precedenza, i risultati scolastici di un individuo dipendono dalla sua predisposizione genetica, ma anche dall'ambiente in cui cresce: statisticamente a migliori condizioni familiari e a genitori istruiti corrispondono figli con migliori risultati scolastici. Thompson si concentra su uno specifico gene, il MAOA, e le sue varianti e nota che per una variante di MAOA la correlazione tra condizione economica e risultati accademici è molto forte e positiva, mentre per un'altra variante dello stesso gene questa relazione è molto più debole. Il MAOA, o monoamminoossidasi-A, è un gene che influenza il funzionamento di neurotrasmettitori come la dopamina e la serotonina, neurotrasmettitori che sono collegati a comportamenti come impulsività, ricerca di novità, ansia, aggressività e altri tratti comportamentali che possono influire sul livello di istruzione.

L'MAOA si presenta in 5 varianti a seconda che una certa sequenza di questo gene si ripeta per 2, 3, 3.5, 4 o 5 volte. Le varianti con 3.5 o 4 ripetizioni sono in grado di legarsi e bloccare meglio i neurotrasmettitori come la serotonina, la noradrenalina e la dopamina e sembra che questo permetta di ottenere migliori risultati in ambito scolastico, come dimostrato da Guo, Roettger, e Cai<sup>10</sup> (2008). Ai fini di questa ricerca, gli individui con le varianti 3.5 e 4 sono stati catalogati con MAOA positivo (1), mentre tutti gli altri con MAOA neutro (0). I dati utilizzati sono quelli del già citato Add Health che contiene il genoma e i dati sulla condizione familiare di 27000 adolescenti che avevano tra gli 11 e 19 anni nell'anno scolastico 1994-95 e che sono stati intervistati di nuovo nel 2008. Più precisamente, è stato preso in considerazione un sottocampione contenente 2612 gemelli (sia omozigoti che eterozigoti). Il gene MAOA si trova sul cromosoma X: negli individui di sesso maschile è unico, mentre in quelli di sesso femminile si presenta in entrambi i cromosomi anche in combinazioni diverse. Per motivi di

---

<sup>9</sup> THOMPSON, Owen, 2014. *“Economic Background and Educational Attainment: The Role of Gene-Environment Interactions.”*

<sup>10</sup> GUANG, Guo, ROETTGER, Michael E., CAI Tianji. 2008. *“The Integration of Genetic Propensities into Social-Control Models of Delinquency and Violence Among Male Youths*

semplificazione sono quindi stati selezionati 1213 individui maschi e ad ognuno di essi è stato assegnato un punteggio in base agli anni di college frequentati e all'ottenimento o meno del diploma.

Nella tabella sottostante sono stati riportati i dati riassuntivi del campione, tra cui la presenza di MAOA positivo, il livello di reddito familiare, gli anni di educazione frequentati e l'ottenimento del diploma del college e altre variabili sulla condizione familiare dell'individuo.

Variable	Mean	SD
Positive MAOA Status	0.57	0.50
Household Income (\$10k 1994 dollars)	4.50	4.50
Attended College	0.73	0.44
Graduated from College	0.41	0.49
Total Years of Education	13.96	2.21
White	0.73	0.44
African American	0.17	0.38
Hispanic	0.14	0.35
Other Race	0.16	0.37
English Spoken at Home	0.92	0.27
Number of Siblings	3.11	1.40
Birth Order	2.22	1.40
Observations	931	

È da notare come la condizione di MAOA positivo sia abbastanza diffusa nel campione, con il 57% di presenza degli individui. Inoltre, il campione si può considerare rappresentativo in quanto per il 27% include individui non bianchi e l'elevata deviazione standard del reddito familiare indica un ampio range di condizioni lavorative.

Nella successiva tabella vengono riportati gli stessi risultati in media, ma divisi in righe per la condizione di MAOA neutro o positivo e in colonne per il reddito familiare se sotto o sopra la mediana, che, per il campione considerato, corrispondeva a 38,000\$ nel 1994.

	Household Income Below Median		Household Income Above Median	
MAOA = 0	College attendance:	0.58	College attendance:	0.84
	College graduation:	0.25	College graduation:	0.52
	Years of schooling:	13.12	Years of schooling:	14.89
	Observations:	220	Observations:	169
MAOA = 1	College attendance:	0.64	College attendance:	0.85
	College graduation:	0.35	College graduation:	0.52
	Years of schooling:	13.32	Years of schooling:	14.62
	Observations:	246	Observations:	299

Nella prima colonna, cioè quella di famiglie con un reddito inferiore alla mediana, si può notare come ci sia una sostanziale differenza tra gli individui con MAOA uguale a zero e quelli con MAOA pari a uno: in media gli individui con MAOA positivo frequentano il college il 6% in più, con il 10% in più che ottiene il diploma e frequentano la scuola 0,2 anni in più rispetto agli individui con MAOA neutro. Invece, tra le famiglie con reddito superiore alla mediana, queste differenze sono pressoché nulle: la percentuale di individui che frequentano il college e si diplomano è praticamente la stessa e anzi, gli individui con MAOA positivo frequentano la scuola in media 0.27 anni in meno. Se si legge la tabella in orizzontale invece, i dati suggeriscono che ci sia una più marcata differenza, negli individui con MAOA nullo, tra le famiglie con reddito inferiore alla mediana e quelle con reddito superiore. Questi risultati suggeriscono come l'andamento accademico dipenda sia dai geni che dall'ambiente familiare, ma anche da un'interazione tra essi. Per tener conto di questi effetti Thompson stima il seguente modello di regressione che include una serie di variabili di controllo:

$$Educational\ Attainment_i = \beta_0 + \beta_1 Log\ Income_i + X'_i \beta + MAOA_i + \varepsilon_i$$

Dove  $X$  è un vettore di controllo che contiene delle dummy riguardo le condizioni della scuola, il numero di fratelli, indicatori del gruppo etnico di appartenenza, la lingua parlata, mese e anno di nascita del figlio e dei genitori. Includendo questi fattori si cerca di isolare l'effetto dovuto all'ambiente familiare: un maggior reddito generalmente implica anche un livello d'istruzione migliore per i figli.

	College Attendance		College Graduation		Years of Education	
	MAOA = 0 (1)	MAOA = 1 (2)	MAOA = 0 (3)	MAOA = 1 (4)	MAOA = 0 (5)	MAOA = 1 (6)
Log Income	0.136*** (0.051)	0.042** (0.020)	0.117** (0.051)	0.017 (0.024)	0.913*** (0.282)	0.287** (0.136)
Observations	394	537	394	537	354	471
R-squared	0.460	0.320	0.451	0.295	0.587	0.347
P-value from test of whether income coefficients are equal across MAOA groups	0.0256		0.0379		0.0157	

I risultati ribadiscono ancora una volta come l'effetto del reddito familiare sulle variabili iscrizione, diploma al college e anni scolastici frequentati è molto minore negli individui con MAOA positivo. Secondo questo modello predittivo, un aumento del 100% del reddito familiare porterebbe ad un aumento dell'11.7% di studenti con MAOA pari a zero che potrebbero ottenere un diploma al college, mentre la stessa variazione per gli studenti con MAOA uno comporterebbe un aumento del solo 1.7%. Stesso ragionamento lo si può fare con la prima colonna, dove un aumento del 100% del reddito familiare porterebbe ad un aumento del 13.6% ragazzi iscritti al college nel caso del MAOA nullo, mentre sarebbe solo del 4.2% negli individui con MAOA positivo.

## Commenti

Nonostante i risultati siano significativi e coerenti con l'ipotesi che l'interazione tra il gene MAOA e il reddito familiare influenzi la carriera scolastica di un individuo, sono necessarie alcune precisazioni sul campione studiato e le variabili di controllo utilizzate. Il reddito familiare potrebbe essere un fattore ambientale non esogeno: ci potrebbero essere altri fattori ambientali correlati con esso che sono stati omessi nel modello e distorcerebbero così la misura dell'interazione GxE. Per dimostrare se c'è o meno questa distorsione è utile inserire nel modello altre caratteristiche socioeconomiche che potrebbero influenzare il reddito familiare. Il dato più plausibile disponibile nel database dell'Add Health è quello del livello di istruzione della madre. Per studiarne l'effetto viene riformulato il modello in due versioni: nella prima (A) viene inserito solo un indicatore del livello di istruzione della madre, mentre la seconda (B) contiene sia questo indicatore che quello del reddito familiare.



	College Attendance		College Graduation		Years of Education	
	MAOA = 0 (1)	MAOA = 1 (2)	MAOA = 0 (3)	MAOA = 1 (4)	MAOA = 0 (5)	MAOA = 1 (6)
<i>Panel A: Maternal Education Only</i>						
Mom High School Graduate	0.163 (0.125)	0.071 (0.070)	-0.029 (0.131)	-0.042 (0.089)	0.797 (0.643)	0.752 (0.481)
Mom College Graduate	0.334** (0.129)	0.264*** (0.080)	0.328*** (0.116)	0.231** (0.089)	2.206*** (0.586)	1.874*** (0.461)
Observations	397	543	397	543	356	474
<i>Panel B: Both Income and Maternal Education</i>						
Log Income	0.091* (0.047)	0.032 (0.020)	0.069 (0.043)	0.014 (0.025)	0.748*** (0.278)	0.224 (0.141)
Mom High School Graduate	0.382*** (0.119)	0.019 (0.085)	0.062 (0.127)	-0.101 (0.095)	1.180* (0.623)	0.244 (0.513)
Mom College Graduate	0.424*** (0.125)	0.219** (0.101)	0.357*** (0.119)	0.195** (0.097)	1.821*** (0.617)	1.322*** (0.488)
Observations	362	495	362	495	328	432

La tabella A con solo l'educazione della madre come variabile di controllo presenta sempre lo stesso trend: la sua correlazione con i migliori risultati scolastici dei figli è maggiore negli individui con MAOA zero rispetto a quelli con MAOA uno. Gli stessi risultati si ottengono aggiungendo il reddito familiare, ma la differenza tra i due MAOA è più pronunciata. Questi numeri indicano come altri fattori socioeconomici potrebbero aiutare a rendere la regressione più precisa, ma purtroppo questi dati non sono disponibili. Il fatto che si ottengano risultati simili anche quando il reddito non è considerato, dimostra comunque come il reddito familiare sia una buona proxy di altri fattori ambientali che non posso essere considerati nel modello.

Un altro problema che può emergere quando si studiano caratteri complessi come l'apprendimento è l'interazione gene-gene: è molto probabile, e anzi quasi certo, che le abilità cognitive di un individuo non siano regolate da un solo gene, ma da una combinazione di essi. In parte si cerca di risolverlo utilizzando come campione coppie di fratelli della stessa famiglia: avendo gli stessi genitori e quindi lo stesso genoma di partenza i geni dei figli sono ereditati in maniera casuale ed indipendente, permettendo così di evitare alcuni dei bias dovuti appunto all'interazione gene-gene. Altro fattore da considerare è che la presenza o meno del gene MAOA potrebbe condizionare l'ambiente dell'individuo, ad esempio, per i suoi comportamenti indotti dal gene o perché i genitori (come nella ricerca precedente) potrebbero decidere di investire di meno o di più. Per come è stato strutturato questo modello la variabile dummy MAOA comprende in sé sia gli effetti diretti che indiretti. Nel caso in cui un individuo con MAOA positivo venga trattato diversamente e quindi l'ambiente di riferimento venga modificato, questo cambio di condizione ambientale è già assorbito dallo stato di MAOA positivo. Con questa specificazione non risulta problematico, ma se si volessero separare gli effetti diretti da quelli indiretti andrebbe formulato un modello diverso.

Si può infine affermare che l'obiettivo di questa ricerca sia stato raggiunto. Tutti i modelli utilizzati e le verifiche fatte indicano come la relazione tra condizione economica familiare e risultati scolastici vari a seconda della presenza o meno di alcune varianti del gene MAOA. I risultati ottenuti rendono chiaro come tratti socioeconomici, in questo caso l'andamento scolastico, siano regolati dall'interazione tra geni e ambiente. Questo studio apre la strada a una serie di altri lavori dello stesso tipo che potranno contare su risultati più precisi grazie a nuovi database sempre più completi oppure fornire un modello di partenza nella ricerca di altri tratti di natura socioeconomica che possono essere utili agli economisti in diversi ambiti di studio.

## Conclusioni

Dopo l'esposizione di questi tre lavori, pare evidente come l'unione tra genetica ed economia sia un binomio vincente. L'ampliamento delle ricerche socioeconomiche mediante l'utilizzo delle conoscenze sulla genetica ha permesso agli economisti di arrivare a conclusioni che mai si sarebbero potute formulare prima. La potenza esplicativa dei geni è immensa e con studi sempre più avanzati si potrà arrivare a prevedere con maggiore precisione la gran parte dei tratti di un individuo. Già in questo elaborato si sono viste diverse applicazioni solamente per i geni che riguardano l'apprendimento. Si è visto come assegnare un punteggio genetico alle persone in base alle loro capacità cognitive può aiutare i governi a sviluppare interventi pubblici più mirati. Oppure ancora, il prevedere come i genitori distribuiscono le risorse a seconda della diversa dotazione genetica dei figli può suggerire agli economisti come strutturare delle politiche pubbliche più efficaci. Ciò non sarebbe possibile se non fosse appunto supportato da evidenze scientifiche che sono in grado di dimostrare che le differenze genetiche tra le persone comportano diversi comportamenti di natura economica.

Le complicazioni nascono proprio dall'affidabilità di questi modelli e dalla loro precisione. Essendo una scienza relativamente "giovane" gli studi che la riguardano sono ancora in fase di sviluppo e necessitano di ulteriori e più precise ricerche. Quello che risalta dai lavori analizzati è che pur raggiungendo dei risultati apprezzabili e verosimili, nascondono nei loro procedimenti criticità e grandi approssimazioni. Come già detto, i PGS sono per definizione delle stime e vengono costruiti come dei predittori lineari, ma molte volte il legame tra genotipo e fenotipo risulta non lineare. Altro problema è la scarsa disponibilità di dati genetici dovuta sia al tempo materiale che è richiesto per raccogliarli, che a volte implica l'attesa di decenni, sia ai costi dovuti al campionamento completo del genoma di centinaia di migliaia di individui, costi che solo ora si stanno abbassando. Inoltre, tutti i principali GWAS si basano su DNA di individui con discendenti europei: le conclusioni fatte potrebbero quindi non essere valide per un campione più omogeneo che comprende anche individui non europei. Le altre difficoltà tecniche nascono dal fatto che ancora molti dei meccanismi di regolazione genica e ambientale sono sconosciuti agli scienziati e quindi potrebbero distorcere i risultati ottenuti.

In aggiunta a ciò è da tenere in considerazione anche l'aspetto etico di questo tipo di studi. L'ottenimento di grandi quantità di dati genetici sugli individui deve essere regolamentato e protetto da specifiche leggi sulla privacy. E ancora, parlare di "differenze genetiche" tra individui maggiormente o meno dotati può portare a conclusioni sbagliate che deviano dagli intenti della geno-economics. Per evitare ciò, alcuni scienziati appartenenti al Social Science

Genetic Association Consortium (SSGAC) hanno infatti stabilito degli standard per una comunicazione responsabile negli studi genetici (Lee, 2018)<sup>11</sup>.

Con queste giuste premesse si possono quindi vedere appieno le potenzialità di future ricerche riguardo la *genoeconomics*. L'obiettivo dei paper esposti non era tanto quello di ottenere dei risultati precisi e corretti, quanto più quello di mostrare la via ai prossimi lavori, dare una base di partenza per migliorare sempre di più i metodi utilizzati. Questi sono gli inizi di una disciplina che ha le potenzialità di rivoluzionare le scienze sociali: seppur ancora lontani dal rappresentare la complessità del genoma umano, i risultati ottenuti fanno ben sperare che nei prossimi anni si possano compiere grandi passi avanti nello studio della genetica. È importante quindi che gli economisti partecipino attivamente a questi processi di innovazione perché, una volta perfezionati, cambieranno il modo di approcciare le scienze socioeconomiche.

---

<sup>11</sup> LEE, James J., et al., 2018. “*Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals.*”

## Bibliografia

- BELSKY, Daniel W., et al., 2018, “*Genetic analysis of social-class mobility in five longitudinal studies*”, Proceedings of the National Academy of Sciences of the United States of America, disponibile su <https://www.pnas.org/doi/10.1073/pnas.1801238115>
- BENJAMIN, Daniel J., et al., 2007 “*Genoeconomics*” In: National Research Council (US) Committee on Advances in Collecting and Utilizing Biological Indicators and Genetic Information in Social Science Surveys; Weinstein M, Vaupel JW, Wachter KW, editors. Biosocial Surveys. Washington (DC): National Academies Press (US); 2008, pp. 1-5
- BIROLI, Pietro, et al., 2022, “*The Economics and Econometrics of Gene-Environment Interplay*”, Tinbergen Institute Discussion Paper 2022-019/V, disponibile su SSRN: <http://dx.doi.org/10.2139/ssrn.4048082>
- BROOKES, Aj, 1999, “*The essence of SNPs*”, Gene, Volume 234, Issue 2, 1999, Pa 177-186
- GENÇ Erhan, et al., 2021, “*Polygenic Scores for Cognitive Abilities and Their Association with Different Aspects of General Intelligence—A Deep Phenotyping Approach*”, Mol Neurobiol. 2021 Aug;58(8):4145-4156. doi: 10.1007/s12035-021-02398-7, disponibile su <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8280022/>
- GUANG, Guo, ROETTGER, Michael E. and CAI Tianji. 2008. “*The Integration of Genetic Propensities into Social-Control Models of Delinquency and Violence Among Male Youths.*” American Sociological Review 73(4)
- Internet, in Treccani.it – Vocabolario Treccani on line, Istituto dell'Enciclopedia Italiana. <https://www.treccani.it/enciclopedia/genetica> consultato il 22 luglio 2023.
- LEE, James J., et al., 2018. “*Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals.*” Nat Genet 50, 1112–1121. Disponibile su: <https://doi.org/10.1038/s41588-018-0147-3>
- SANZ DE GALDEANO, Anna, TERSKAYA, Anastasia, 2019. “*Sibling Differences in Educational Polygenic Scores: How Do Parents React?*” IZA Discussion Papers 12375, Institute of Labor Economics (IZA).
- THOMPSON, Owen, 2014. “*Economic Background and Educational Attainment: The Role of Gene-Environment Interactions.*” The Journal of Human Resources, 49(2), 263–294. Disponibile su <http://www.jstor.org/stable/23799085>
- YOUNG, Alexander.I., et al., 2022, “*Mendelian imputation of parental genotypes improves estimates of direct genetic effects.*” Nat Genet 54, 897–905 disponibile su <https://doi.org/10.1038/s41588-022-01085-0>