POLITECNICO DI TORINO

Collegio di Ingegneria Gestionale

Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea Magistrale

Principali fattori economici e sociali che determinano il numero di medaglie vinte da un paese alle Olimpiadi: costruzione di un modello per la previsione delle prestazioni dei prossimi Giochi olimpici estivi



Relatore Candidato

Prof. Luigi Buzzacchi

Oberto Carlo Spinola

Anno Accademico 2020/2021

<u>Abstract</u>

Nel corso degli anni sono stati condotti numerosi studi scientifici riguardanti i Giochi olimpici, in particolare quelli estivi. Uno degli argomenti che sono stati maggiormente approfonditi è lo studio delle determinanti economiche, sociali e politiche del numero di medaglie conquistate da un paese alle Olimpiadi. La letteratura di settore stabilisce che i fattori che maggiormente influiscono sul successo di una nazione alle Olimpiadi sono il livello di benessere, espresso come PIL pro capite, e la dimensione della popolazione. Inoltre, le precedenti ricerche in questo ambito hanno affermato che anche la struttura politica ed economica di un paese, così come il fatto di essere la nazione organizzatrice della manifestazione risultano essere dei fattori fortemente rilevanti nel quantificare l'entità del bottino olimpico. Analogamente, alcuni studi hanno affermato che anche la provenienza geografica si dimostra estremamente significativa. L'obbiettivo di questo studio consiste nel costruire dei modelli di regressione per valutare se le ipotesi sviluppate e dimostrate nella letteratura di settore risultino tutt'oggi ancora valide. Per questo motivo, si è scelto di presentare inizialmente un modello tradizionale, che muove direttamente da quelli suggeriti dalla ricerca scientifica, in cui si è testato l'effetto sul totale delle medaglie dato dalla dimensione della popolazione, dal configurarsi come nazione ospitante e dalla struttura politico-economica; a questi fattori si è deciso di aggiungere gli effetti forniti dal numero di neonati morti, dalla crescita demografica e dalla spesa sanitaria pro capite, che oltre ad essere una misura del livello di qualità del sistema sanitario è anche un preciso indicatore del livello di benessere della popolazione. Per definire il secondo modello, sono state incluse nella funzione ottenuta in precedenza le variabili che descrivono l'appartenenza ad una determinata area geografica, utili per catturare l'effetto sulle medaglie dato dalla differente cultura sportiva nelle diverse regioni del globo. Si è poi ripreso il primo modello ed anziché aggiungere le variabili regionali, si è deciso di includere quelle riportanti il tipo di religione professata, in considerazione del fatto che una diversa tradizione religiosa fa da specchio a numerosi aspetti socio-culturali del paese considerato. Infine, per quanto riguarda l'ultimo modello, si è optato per un approccio del tutto infrequente nella letteratura di settore, basato sulla stepwise regression, un algoritmo di regressione utile per determinare quali variabili risultino essere significative, senza escludere a priori alcune variabili sulla base dell'esperienza o di congetture qualitative: il modello ottenuto annovera come variabili indipendenti il numero di atleti inviati da un paese alle Olimpiadi, il numero di articoli scientifici pubblicati in un anno dalla nazione considerata, il tasso di disoccupazione, la spesa sanitaria pro capite, la mortalità neonatale e l'aspettativa di vita. Inoltre, per ognuno dei quattro modelli esposti è stata condotta un'analisi di regressione, verificando che ne fossero rispettati gli assunti, ed è stata svolta un'analisi di robustezza.

<u>Indice</u>

1 Introduzione	5
2 Letteratura	7
2.1 Un breve cenno storico sulle Olimpiadi	7
2.2 Rassegna della letteratura	8
2.2.1 Dagli albori fino al nuovo millennio	8
2.2.2 Gli anni zero del terzo millennio	10
2.2.3 Dal 2010 fino ai giorni nostri	41
2.3 Conclusione	52
3 Metodologia	54
3.1 Dati e fonti	54
3.2 La partecipazione e le performances	54
3.3 Le variabili dipendenti	55
3.3.1 Il numero di medaglie	55
3.3.2 Il numero di atleti	55
3.4 Le variabili indipendenti	56
3.4.1 Le variabili economiche	57
3.4.2 Le variabili geografiche	60
3.4.3 Le variabili politiche	62
3.4.4 Le variabili sociali	63
3.4.5 Le variabili scientifiche	65
3.4.6 Le variabili demografiche	66
4 I fattori che determinano le medaglie alle Olimpiadi dal 2000 al 2016: quattro stime	68
4.1 L'importanza della struttura politico-economica	68

4.1.1 Il modello statistico	68
4.1.2 Verifica degli assunti	71
4.1.3 L'analisi della robustezza	74
4.1.4 Conclusione	74
4.2 Una specifica che tiene conto delle differenze culturali e regionali	75
4.2.1 Il modello statistico	75
4.2.2 Verifica degli assunti	77
4.2.3 L'analisi della robustezza	78
4.2.4 Conclusione	78
4.3 L'importanza della religione	79
4.3.1 Il modello statistico	79
4.3.2 Verifica degli assunti	82
4.3.3 L'analisi della robustezza	82
4.3.4 Conclusione	83
4.4 Un modello innovativo	84
4.4.1 Il modello statistico	84
4.4.2 Verifica degli assunti	86
4.4.3 L'analisi della robustezza	86
4.4.4 Conclusione	87
5 Conclusione	89
Bibliografia	91
Sitografia	93
Ringraziamenti	94
Appendice I	95

1 Introduzione

Sembra banale affermare che un atleta vince una medaglia ai Giochi Olimpici estivi sia grazie alle proprie caratteristiche biometriche, ovvero altezza, peso, massa muscolare, velocità, flessibilità, destrezza, tecnica gestuale, talento, ecc., sia grazie alla propria condizione fisica al momento della competizione, che è funzione della pianificazione e della qualità dell'allenamento in preparazione alla competizione. Talvolta risultano essere determinanti anche le caratteristiche fisiche e tecniche dell'attrezzatura sportiva di cui è dotato l'atleta. Le scienze dure, come fisica, biologia, biomeccanica, ingegneria, medicina, rafforzano questo modo di interpretare e predire il successo olimpico. Tuttavia, un approccio di questo tipo fornisce una spiegazione semplicistica del perché gli atleti di un paese vincono costantemente più medaglie di quelli di un'altra nazione.

Le sole caratteristiche fisiche e tecniche non si dimostrano sufficienti per spiegare perché gli atleti americani sono soliti dominare il medagliere delle Olimpiadi estive o perché i paesi in via di sviluppo vincono complessivamente molte meno medaglie rispetto alla trentina di paesi più sviluppati, benché siano decisamente più numerosi. Tali caratteristiche non riescono nemmeno a spiegare i motivi per i quali la Repubblica Democratica Tedesca e l'Unione Sovietica erano solite vincere un elevato numero di medaglie, di gran lunga superiore alle attese. Inoltre, le scienze dure non riescono a spiegare la grande crescita in termini di medaglie conquistate della Cina, che nelle recenti Olimpiadi di Rio de Janeiro 2016 si è affermata come terza potenza olimpica mondiale, un exploit che i Giochi olimpici estivi di Tokyo 2020 dovrebbero confermare, anche in considerazione del fatto che la Cina si è ormai affermata come preminente potenza economica mondiale, seconda soltanto agli Stati Uniti.

Tali domande hanno alimentato un assunto alla base dell'economia dello sport, vale a dire che la dotazione e lo sviluppo delle risorse economiche ed umane sono fattori decisivi nel determinare la prestazione olimpica di una nazione. Nei paper scientifici di settore, emergono relazioni significative se si regredisce il numero di medaglie vinte sul PIL pro capite e sulla popolazione, secondo un approccio macroeconomico delle performance dei paesi ai Giochi olimpici estivi che si è sviluppato già durante la Guerra Fredda.

Nel presente lavoro viene proposta un'analisi economica delle prestazioni olimpiche che tenga conto anche delle caratteristiche regionali e, di conseguenza, della cultura che influenza la pratica sportiva nelle diverse regioni del mondo, delle influenze sociali e religiose e di quelle sanitarie e scientifiche.

L'aumento rispetto all'edizione olimpica precedente nel numero delle medaglie vinte da un paese riflette, in un certo senso, il numero di medaglie perse da altri paesi. Per spiegare le prestazioni olimpiche di un paese, è necessario tenere conto di tutti gli altri paesi partecipanti ai Giochi e dell'interdipendenza tra le loro prestazioni, fermo restando il vincolo generale del numero totale di medaglie assegnate durante una competizione. Negli ultimi anni sono stati compiuti notevoli progressi metodologici nell'analisi statistica delle prestazioni olimpiche, che hanno permesso di raccogliere dati per un campione di paesi sempre più ampio e di introdurre variabili specifiche per alcuni gruppi di paesi.

Una delle finalità di questa tesi consiste nel validare i risultati della letteratura, che tratta come determinanti principali delle vittorie olimpiche le risorse economiche e umane, analizzandone lo sviluppo e la distribuzione tra i paesi, anche in considerazione del diverso regime politico che li caratterizza.

A questo proposito, si è cercato di costruire quattro modelli predittivi, di cui i primi tre appaiono più fedeli alla letteratura di settore mentre il quarto è stato ricavato per mezzo di un algoritmo di regressione.

Tuttavia, il contributo apportato da questa ricerca alla letteratura è limitato dalla scarsità di dati disponibili rispetto a quanto sarebbe stato necessario o quantomeno desiderato.

La tesi risulta strutturata nel seguente modo. La trattazione prende avvio con una rassegna della letteratura di settore, in cui vengono mostrati i risultati raggiunti dai precedenti ricercatori. Vengono poi presentate la metodologia di ricerca utilizzata e le variabili dipendenti e indipendenti che sono state scelte per questa ricerca. Successivamente, vengono descritti i quattro modelli di regressione, che sono stati costruiti a partire da variabili economiche, geografiche, politiche, sociali, scientifiche, e demografiche. Infine, vengono esposte le conclusioni a cui si è giunti attraverso questa ricerca.

2 Letteratura

2.1 Un breve cenno storico sulle Olimpiadi

La culla dei Giochi Olimpici fu Olimpia, in Grecia, dove nel 776 A.C. si svolsero le prime Olimpiadi. Questi giochi erano dedicati agli dei, in particolare a Zeus ed erano celebrazioni atletiche e religiose che si svolgevano ogni quattro anni. Inizialmente consistevano in una gara di corsa poi, nel tempo, furono aggiunti altri sport, come la lotta, il pugilato ed il pentathlon. La superiorità fisica, il desiderio di primeggiare, la sete di gloria così come pure il disprezzo del perdente caratterizzavano gli atleti ellenici. Inizialmente le Olimpiadi si svolgevano nell'arco di un giorno, poi dal 472 a.C. la loro durata si protrasse a cinque giorni, passarono anche da un'organizzazione regionale ad una molto più vasta coinvolgendo Principi, Re e governanti di tutta la Grecia, tanto che durante i giochi venivano sospese le guerre in tutto il territorio in nome della "tregua olimpica" che metteva da parte la bellicosità latente in nome della pace. Tuttavia, con la nascita dell'Impero Romano i giochi iniziarono a perdere pian piano di interesse e di importanza, per arrivare a scomparire definitivamente con l'avvento del Cristianesimo. Si svolsero senza interruzione per undici secoli fino al 394 d.C., anno in cui l'imperatore Teodosio le sospese. Tuttavia, le Olimpiadi non morirono mai completamente, continuando ad esistere prima in Inghilterra e successivamente in Francia ed in Grecia, dove dal XVII secolo in poi si svolsero eventi simili, anche se in scala più ridotta e non a carattere internazionale. L'interesse verso i Giochi Olimpici aumentò notevolmente quando nel XIX secolo gli archeologi inglesi e tedeschi portarono alla luce le rovine dell'antica Olimpia. La grande quantità di reperti testimoniava la grandezza e la magnificenza che aveva Olimpia al tempo dei greci. Si deve al nobile francese Pierre de Frédy barone di Coubertin la nascita delle Olimpiadi dell'era moderna, celebre fu la sua frase "niente della storia dell'antichità mi aveva fatto sognare più di Olimpia". Egli si dedicò alla professione di pedagogista e dirigente sportivo, e fu durante una sua visita ai college britannici che si convinse che attraverso l'educazione sportiva si sarebbe potuto migliorare il sistema educativo e scolastico. De Coubertin ipotizzò che la causa della sconfitta francese nella guerra franco-prussiana fosse dovuta al fatto che i francesi non avevano ricevuto un'educazione fisica adeguata. Voleva inoltre trovare un modo per permettere ai giovani di nazionalità diverse di confrontarsi in una competizione sportiva anziché in guerra, quindi attribuiva allo sport un significato più profondo rispetto a quello di semplice passatempo del tempo dei greci. Nel 1894 alla Sorbona di Parigi, durante quello che sarebbe poi stato definito il primo congresso olimpico della storia, rese pubblica l'idea di voler recuperare gli antichi giochi olimpici. Nei due anni successivi lavorò per fondare il Comitato Olimpico (CIO), del quale divenne prima segretario e successivamente presidente. Il congresso decise inoltre che le prime Olimpiadi dell'era moderna si sarebbero svolte nel 1896 ad Atene, trovando molto suggestivo immaginare i Giochi li dove erano nati. I Giochi olimpici sono arrivati oggi alla loro trentunesima edizione e sono senza dubbio l'evento sportivo più importante e seguito al mondo. Nel corso degli anni il numero dei partecipanti e delle discipline sportive che costituiscono le Olimpiadi è cresciuto enormemente, come si nota dalla Tab. 2.1. Nel 1896 ad Atene parteciparono 14 paesi rappresentati da 285 atleti, in questa occasione non fu data la possibilità alle donne di iscriversi, al fine di rispettare la tradizione greca, che le voleva solo riservate agli uomini. In questa prima Olimpiade gli eventi sportivi furono solamente 43 divisi in 9 discipline diverse. Al contrario, nelle ultime Olimpiadi tenutesi a Rio de Janeiro nel 2016 parteciparono ai 306 eventi, appartenenti a 28 categorie diverse di sport, 11238 atleti provenienti da ben 207 nazioni. Quest'ultima edizione dei giochi olimpici estivi è stata vista da più di 3,6 miliardi di telespettatori in tutto il mondo. I Giochi olimpici invernali si svolsero per la prima volta a Chamonix nel 1924 con il nome di "Settimana internazionale degli sport invernali" e facevano inizialmente parte dei Giochi olimpici estivi. Nel 1925 il CIO sancì ufficialmente la nascita dei Giochi olimpici invernali che fino al 1992 vennero disputati nello stesso anno di quelli estivi e solo a partire dal 1994 si disputarono negli anni pari non divisibili per quattro. A Chamonix presero parte 258 atleti di 16 nazioni che gareggiarono in 16 eventi di 9 discipline diverse. Col passare degli anni il numero di atleti di sesso femminile è notevolmente aumentato, se si pensa che nel 1900 nei Giochi di Parigi la percentuale di quote rosa era pari al 2,2%, oggi le stime per Tokyo 2021 ci indicano che la percentuale di atlete sul totale dei partecipanti sarà del 48,8%.

Olimpiade	Atene 1896	Chamonix 1924	Rio 2016
Numero di paesi partecipanti	14	16	207
Numero di atleti partecipanti	285	258	11238
Numero di eventi sportivi	43	16	306
Numero di discipline	9	9	28
Numero di atleti uomini	285	245	6179
Numero di atleti donne	0	13	5059

Tab. 2.1: La variazione della partecipazione ai Giochi olimpici

Nonostante le Olimpiadi siano un evento caratterizzato dallo spirito di competizione internazionale e dall'eccellenza atletica e sportiva globale, non tutti i paesi hanno le stesse possibilità di partecipare ai Giochi, né hanno eguali probabilità di vincere una medaglia.

2.2 Rassegna della letteratura

Considerata l'importanza ed il carattere internazionale dell'evento le Olimpiadi sono da sempre un argomento di studio e ricerca scientifica per moltissimi ricercatori negli ambiti più disparati: da quello economico a quello medico, dagli studi di psicologia a quelli di natura matematica. Un rilevante filone di ricerca si propone di identificare i fattori socio-economici che determinano il successo sportivo ai Giochi olimpici alla scala del Paese partecipante. Buona parte di questi studi definiscono il "successo sportivo" sulla base del numero di medaglie conquistate; le relative ricerche si propongono quindi di stimare quali variabili demografiche, culturali, economiche e geografiche spiegano, e in che misura, il numero di medaglie effettivamente conquistate in uno o più eventi olimpici.

2.2.1 Dagli albori fino al nuovo millennio

Una delle prime ricerche riguardante i fattori determinanti del successo è quella del ricercatore canadese Donald W. Ball del 1972 intitolata "Olympic Games Competition: Structural Correlates of National Success". Sempre nello stesso anno, uscì il lavoro di un altro ricercatore N. Levine titolato "Why do Countries Win Olympic Medals? Some Structural Correlates of Olympic Games Success: 1972", basato anch'esso sulla ricerca dei fattori che determinano il successo di una nazione alle Olimpiadi. Nel 1974 i ricercatori Grimes, Kelly e Rubin pubblicarono un testo sul medesimo argomento, dal titolo "A Socioeconomic Model of National Olympic Success". Sia queste ultime due pubblicazioni che quella di Ball mostrano come le performance dei paesi socialisti e delle regioni che ospitano i Giochi surclassino di gran lunga quelle degli altri paesi. A queste tre ricerche importanti seguirono 20 anni di "buio", durante i quali non furono più approfondite queste tematiche. Questo periodo di silenzio durò fino al 1993 quando i ricercatori Shughart and Tollison pubblicarono "Going for the Gold: Property Rights and Athletic Effort in Transitional Economies", nelle cui pagine cercarono di spiegare come le conseguenze della caduta dell'Unione Sovietica si ripercossero sui

risultati ottenuti alle Olimpiadi. Shughart and Tollison affermarono che il livello di impegno che veniva messo dagli atleti nella competizione olimpica dipendeva anche dal valore della ricompensa, che non era solo l'entità simbolica della medaglia olimpica ma anche un ritorno in termini economici e sociali. I partecipanti a questo tipo di competizioni devono gestire con diligenza la qualità dei loro allenamenti per ottenere un successo finale in modo da massimizzare il guadagno netto che ci si aspetta dalla partecipazione alla gara, sapendo che la classifica finale dipenderà dalla qualità del lavoro svolto dai singoli in confronto a quello degli altri partecipanti. Il livello ottimale di impegno è determinato uguagliando il beneficio marginale ed il costo marginale attesi. Sotto il regime Sovietico vincere una medaglia garantiva all'atleta di affrancarsi sia a livello sociale che economico tanto che queste ricompense fecero sì che gli atleti incrementassero il proprio livello di impegno. Con il collasso dell'Unione Sovietica tutti i privilegi riservati agli atleti olimpici svanirono, essi non poterono più pensare di capitalizzare le proprie medaglie in cambio di fama e fortuna commerciale e sociale. Il crollo del blocco Sovietico nel 1991 portò quindi ad una diminuzione dei benefici ricevuti per le vittorie Olimpiche e di conseguenza portò alla riduzione del livello di impegno fornito dagli atleti. Shughart and Tollison hanno confrontato le medaglie vinte dai paesi del blocco Sovietico nelle Olimpiadi del 1992, del 1988 e del 1976.

			_	
Olympic	Medal '	Cotals (Summer	Games)

	1992 – Barcelona (815 Total Medals)			1988 – Seoul (739 Total Medals)		1976 – Montreal (613 Total Medals)	
Team	Medals	Percent	Medals	Percent	Medals	Percent	
EUN/SOV	119ª	14.6	132	17.9	125	20.4	
GER/GDR	82	10.1	40/102	5.4/13.8	39/90	6.4/14.7	
BUL	16	2.0	35	4.7	22	3.6	
CZE	7	0.9	8	1.1	8	1.3	
HUN	30	3.7	23	3.1	22	3.6	
POL	19	2.3	16	2.2	26	4.2	
ROM	18	2.2	24	3.2	27	4.4	
YUG	8 ^b	1.0	12	1.6	8	1.3	
USA	108	13.3	94	12.7	94	15.7	
CHN	54	6.6	28	3.4	c		

a Includes medals awarded to Latvia (3), Lithuania (2), and Estonia (2).

Tab. 2.2: L'impatto dello scioglimento dell'URSS sui Giochi olimpici estivi

La Tab. 2.2 evidenzia chiaramente come il collasso del socialismo abbia avuto un impatto significativo sulla performance dei paesi appartenenti al Blocco orientale nelle Olimpiadi del 1992. Al contrario, la Cina ha adottato i metodi sovietici di sovvenzione e ricompensa per i suoi atleti olimpici ed ha quindi registrato un notevole incremento delle sue performance alle Olimpiadi di Barcellona dello stesso anno. I dati contenuti in questa tabella supportano l'ipotesi a cui sono giunti Shughart and Tollison che hanno anche analizzato nel loro studio i risultati dei Giochi olimpici invernali, come si osserva in Tab. 2.3.

^b Includes medals awarded to Croatia (2) and Slovenia (2).

^c The People's Republic of China did not field an Olympic Team in 1976.

Table 2
Olympic Medal Totals (Winter Games)

	1992 – A (171 Tota					Sarajevo al Medals)
Team	Medals	Percent	Medals	Percent	Medals	Percen
EUN/SOV	23	13.4	29	21.0	25	22.1
GER/GDR	26	15.2	8/25	5.8/18.1	4/24	3.5/17.4
BUL	0	0	0	0	0	(
CZE	3	1.8	3	2.2	6	5.3
HUN	0	0	0	0	0	(
POL	0	0	0	0	0	(
ROM	0	0	0	0	0	(
YUG	0	0_	3	2.2	1	0.8
USA	11	6.4	6	4.3	8	7.1

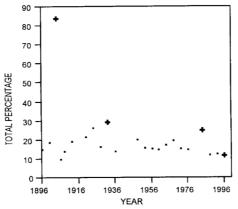
Tab. 2.3: L'impatto dello scioglimento dell'URSS sui Giochi olimpici invernali

A partire dalle Olimpiadi di Albertville, la squadra della Germania Orientale ha ottenuto un numero inferiore di medaglie rispetto alle Olimpiadi precedenti, così come la Squadra Unificata, quella che raccoglieva gli atleti dei paesi dell'ex Unione Sovietica, ottenne in quella occasione scarsi risultati. Queste evidenze dimostrano come gli incentivi siano uno stimolo molto forte. Shughart and Tollison concludono dicendo che i paesi del Blocco orientale vinsero meno medaglie in occasione delle Olimpiadi di Albertville e Barcellona rispetto alle edizioni precedenti perché la vittoria olimpica aveva assunto agli occhi degli atleti, in questo determinato contesto storico e politico, un minor ritorno sociale ed economico, e quindi si è verificata una riduzione del livello di sforzo da parte degli atleti. Tuttavia, a causa del limitato numero di riscontri disponibili questo lavoro è caratterizzato da un elevato tasso di casualità delle prestazioni individuali e di squadra, che gli autori non sono stati in grado di spiegare. I risultati di Shughart and Tollison mettono in evidenza come i benefici economico-sociali che si sommano al puro ritorno agonistico della vittoria abbiano un ruolo assai rilevante, e che il collasso del socialismo e l'incompleta transizione verso un'economia di mercato ha ridotto le opportunità per gli atleti del Blocco orientale di appropriarsi dei benefici derivanti dal vincere una medaglia.

2.2.2 Gli anni zero del terzo millennio

Lo studio dei fattori socio-economici nella performance sportiva ai Giochi olimpici a nostra conoscenza non è più oggetto di pubblicazioni sulla letteratura scientifica fino al 2000, quando Stephen Clarke, Professore di Scienze matematiche della Swinburne University of Technology di Melbourne, condusse una ricerca per quantificare l'entità del vantaggio dato dall'essere il paese ospitante ai Giochi Olimpici dal titolo "Home advantage in the Olympic". Clarke raccolse i dati riguardanti le Olimpiadi dal 1896 al 2000; diciassette paesi hanno ospitato in questi anni i Giochi olimpici estivi, quattro stati lo hanno fatto per due volte e gli USA per quattro volte. Su diciassette paesi che hanno ospitato le Olimpiadi gli autori hanno evidenziato che quattordici hanno vinto la maggior percentuale di medaglie disponibili proprio nell'edizione in cui erano il paese ospitante. Gli atleti del paese organizzatore delle Olimpiadi sono avvantaggiati dalla familiarità con le infrastrutture e dalla presenza del pubblico, mentre per la squadra ospite il viaggio rappresenta, insieme alla mancanza di sostenitori, un effetto negativo. Per esempio, alle Olimpiadi di Sidney del 2000 gli atleti provenienti da altri paesi hanno sofferto il cambio di fuso orario e la differenza

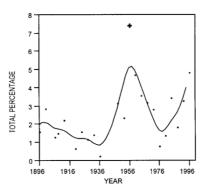
climatica, mentre gli atleti australiani hanno potuto beneficiare del sostegno del proprio pubblico. Vi sono poi altre ragioni secondo Clarke per le quali i paesi ospitanti outperformano alle Olimpiadi di casa; essi, infatti, presentano un maggior numero di atleti e competono in più discipline rispetto a quanto farebbero di solito. La ricerca di Clarke si basa essenzialmente sull'analisi del numero di medaglie, d'oro, d'argento e di bronzo, vinte dai paesi nelle diverse edizioni dei Giochi. Se si esclude l'Olimpiade di Atlanta del 1996, le altre tre edizioni ospitate dagli USA hanno visto un risultato eccezionale in termini di medaglie vinte.



The number of medals won by the USA each year as a percentage of medals available.

Fig. 2.1

Stesso discorso vale per l'Australia, che nelle Olimpiadi di Melbourne del 1956, ha vinto oltre il 50% di medaglie in più rispetto ai Giochi olimpici di maggior successo.



Percentage of available medals Australia won each year.

Fig. 2.2

Per comparare e valutare le differenze tra le performances casalinghe e quelle in trasferta dei vari paesi Clarke si è concentrato sull'analisi della percentuale delle medaglie disponibili vinte dai diciassette paesi che hanno ospitato le Olimpiadi estive. Clarke ha raffigurato nella Tab. 2.4 la performance media domestica e quella fuori casa; si osserva che c'è un chiaro vantaggio per il paese ospitante.

	Perc	entage of	
Country	available	medals won	Ratio
	Away	Home	Home: Away
Australia	2.3	7.4	3.2
Belgium	1.3	8.1	6.0
Canada	2.1	1.8	0.9
Finland	3.3	4.8	1.4
France	5.0	23.7	4.8
Great Britain	5.3	25.3	4.8
Germany	6.6	14.8	2.2
Greece	0.4	26.8	62.3
The Netherlands	1.7	5.8	3.4
Italy	4.3	7.8	1.8
Japan	3.1	5.8	1.9
Korea	1.2	4.5	3.6
Mexico	0.4	1.7	3.9
Soviet Union	16.9	30.9	1.8
Spain	0.5	2.7	5.1
Sweden	4.5	21.0	4.7
USA	17.0	37.9	2.2

Tab. 2.4: la performance media domestica e quella fuori casa

Il 50% dei paesi organizzatori ha vinto un numero di medaglie superiore a 3,5 volte la propria media storica di medaglie vinte per Olimpiade. Clarke ha anche confrontato la performance del paese ospitante con la media dei piazzamenti dell'edizione precedente e di quella successiva a quella ospitata.

Year	Country	Average of performance in games before and after being host	Home performance	Ratio
1896	Greece	11.4	38.5	3.4
1900	France	13.4	37.0	2.8
1904	USA	14.7	84.4	5.7
1906	Greece	19.9	15.0	0.8
1908	Great Britain	12.2	44.9	3.7
1912	Sweden	11.1	21.0	1.9
1920	Belgium	1.4	8.1	5.6
1924	France	7.9	10.4	1.3
1928	Netherlands	2.4	5.8	2.4
1932	USA	15.8	29.9	1.9
1936	Germany	5.6	22.9	4.1
1948	Great Britain	3.0	5.6	1.9
1952	Finland	4.0	4.8	1.2
1956	Australia	3.6	7.5	2.1
1960	Italy	5.4	7.8	1.4
1964	Japan	4.3	5.8	1.3
1968	Mexico	0.2	1.7	9.3
1972	Germany	5.7	6.7	1.2
1976	Canada	3.6	1.8	0.5
1980	USSR	19.1	30.9	1.6
1984	USA	14.0	25.3	1.8
1988	Korea	3.6	4.5	1.3
1992	Spain	1.3	2.7	2.1
1996	USA	13.3	12.0	0.9

Tab. 2.5: il confronto tra le performance

Clarke conclude affermando che il vantaggio di giocare in casa è maggiore per gli atleti più forti, e che porta ad un aumento della percentuale di medaglie d'oro sul totale delle medaglie vinte. Se si escludono le prime edizioni dei Giochi, gran parte dei paesi ospitanti ha vinto un maggior numero di medaglie d'oro grazie al fatto di giocare in casa.

Sempre nel 2000 venne prodotto un altro lavoro molto interessante, quello di Daniel K.N. Johnson e di Ayfer Ali. Questa ricerca si occupa di definire quali siano i fattori economici e politici che determinano la partecipazione degli atleti, ed in particolare di quelli femminili, di un paese alle Olimpiadi estive. Essi hanno analizzato il successo di una nazione ai Giochi sia dal punto di vista individuale che da quello nazionale, e hanno visto che i risultati sono simili. Lo scopo del loro lavoro era quello di comprendere quali fossero le variabili che incoraggiano o scoraggiano un paese dal mandare i suoi atleti alle Olimpiadi, nello stesso modo hanno analizzato il successo di una nazione in questa manifestazione. Johnson e Ali affermano che un alto PIL pro capite può essere associato ad una migliore qualità di allenamento e anche ad una maggiore possibilità da parte del paese di sostenere le spese per mandare gli atleti ai Giochi. Secondo gli autori una popolazione numerosa indica che i costi di allenamento, delle infrastrutture e dei trasporti possono essere divisi tra un

maggior numero di persone, rappresentando quindi una maggiore probabilità di coltivare dei talenti. Essere la nazione ospitante, o una delle nazioni confinante, presenta costi inferiori di trasporto ed il vantaggio di allenarsi nelle condizioni climatiche e ambientali in cui si svolgeranno le gare. I due ricercatori americani ipotizzano che anche la struttura politica ed il colonialismo influenzano la partecipazione ed il successo alle Olimpiadi. La ricerca si basa sui dati raccolti alle Olimpiadi a partire dal 1952. Dai dati si evince che la partecipazione globale alla manifestazione è un fenomeno recente, i primi Giochi avevano una partecipazione limitata ai soli paesi altamente sviluppati. La situazione cambiò radicalmente dopo la Seconda Guerra Mondiale, si è avuto un incremento del numero di paesi partecipanti ma anche un aumento della concentrazione del numero di atleti provenienti dai paesi sviluppati. Nel 1952 gli atleti rappresentanti i primi dieci paesi, classificati per PIL pro capite, costituivano il 19,1% degli atleti totali, nel 1996 questa percentuale era salita al 37,9%.

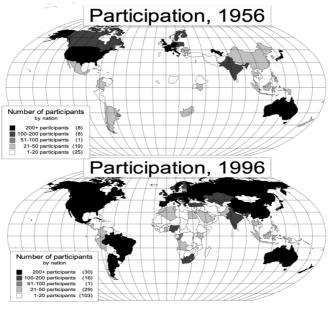


Fig. 2.3

I dati analizzati da Johnson e Ali mostrano anche un notevole aumento della partecipazione femminile, passando da una donna ogni 8 uomini della prima edizione successiva al conflitto mondiale, a meno di una donna ogni 2,5 uomini. Per quanto riguarda il successo in termini di medaglie vinte su 241 paesi partecipanti solo 119 nazioni hanno vinto una medaglia, e tutti i paesi africani sommati hanno vinto meno del 2% delle medaglie disponibili. La Tab. 2.6 sviluppata dai due ricercatori americani mostra una chiara relazione tra la dimensione della popolazione di un paese ed il suo successo alle Olimpiadi. Ai Giochi estivi del 1956 in media una nazione che vinceva almeno una medaglia aveva una popolazione sei volte superiore ed un PIL pro capite doppio rispetto a quelli di una nazione che non riusciva ad aggiudicarsi nemmeno una medaglia. Nei Giochi del 1996 queste statistiche sono leggermente diminuite rimanendo comunque significative.

Year	Medal status	Population		GDP pe	r capita
		average	st. dev.	average	st. dev.
1956	no	6873.54	7332.28	1766.27	1479.51
	yes	41011.00	78294.98	4015.04	2352.94
1996	no	15921.07	27816.92	4697.35	4827.85
	yes	78615.46	209508.50	7767.25	5695.97

Tab. 2.6: relazione tra la dimensione della popolazione ed il successo olimpico

Johnson e Ali hanno quindi sviluppato un semplice modello di regressione lineare per stimare i fattori che determinano la partecipazione di un paese alle Olimpiadi:

$$part_{i} = \alpha + \beta_{1}GDP + \beta_{2}GDP^{2} + \beta_{3}POP + \beta_{4}POP^{2} + \beta_{5}HOME + \beta_{6}NEIGH$$
$$+ \sum_{i=7}^{10} \beta_{j}POL_{j} + \beta_{11}COLGDP + \beta_{12}COLPOP + \beta_{13}COLMED + u_{i} + e$$

Johnson e Ali hanno scelto di non inserire nella funzione le "lagged performance variables" poiché nonostante siano significative ed altamente esaustive volevano concentrarsi unicamente sui fattori economici e politici. I risultati ottenuti mostrano che le nazioni con il PIL pro capite più alto mandano un numero maggiore di atleti e soprattutto di atlete alle Olimpiadi. Per ogni 1000\$ in più di PIL pro capite i paesi sviluppati tendono a mandare 8 atleti in più. Lo stesso fanno i paesi con una popolazione più numerosa inviando un maggior numero di atleti ai Giochi, un aumento di 10000 cittadini porta a 4 atleti in più; mentre sono necessari 30000 cittadini in più per aumentare di un atleta donna la squadra. L'effetto paese ospitante è molto significativo, essere una nazione geograficamente vicina a quella dove si terranno i Giochi rappresenta un boost per la partecipazione. La nazione ospitante presenta in media 209 atleti in più, di cui 64 donne. I paesi confinanti inviano 48 partecipanti più del previsto, di cui 14 sono donne. Mentre il sistema politico ha un effetto ridotto sui livelli di partecipazione, l'abilità economica e olimpica delle ex potenze coloniali sono anche potenti indicatori sul numero dei loro partecipanti. Le ex colonie di nazioni ricche partecipano meno di quanto sarebbe previsto da altre variabili, ma questo è compensato da una tendenza a partecipare di più se l'ex potenza coloniale è una nazione con tendenza a vincere molte medaglie. Johnson e Ali hanno ottenuto risultati simili e coerenti anche utilizzando una funzione Cobb-Douglas, anziché una semplice funzione lineare. Confrontando le previsioni di partecipazione per le Olimpiadi del 2000 a Sidney con i dati reali si evince che il lavoro di Johnson e Ali, eccezion fatta per l'Australia, risulta abbastanza preciso ed accurato. Per quanto riguarda la quantificazione del successo da parte di una nazione, i due ricercatori si sono concentrati su due metodologie di analisi: quella a livello individuale e quella a livello nazionale. Per analizzare il punto di vista nazionale essi hanno utilizzato, anche in questo caso, una funzione lineare:

$$\begin{split} medals_i &= \alpha + \beta_1 GDP + \beta_2 GDP^2 + \beta_3 POP + \beta_4 POP^2 + \beta_5 HOME + \beta_6 NEIGH \\ &+ \sum_{j=7}^{10} \beta_j POL_j + \beta_{11} COLGDP + \beta_{12} COLPOP + \beta_{13} COLMED \\ &+ \beta_{14} MED + \beta_{15} TIME + u_i + e \end{split}$$

Anche in questa analisi il PIL pro capite e la popolazione hanno un impatto positivo, anche se con rendimento di scala decrescente. La nazione ospitante gode di un enorme vantaggio, vincendo in media 25 medaglie in più del previsto, di cui 12 d'oro. Minore risulta essere l'effetto per i paesi confinanti che si aggiudicano in media un extra di 3,5 medaglie, questo risultato positivo spinge anche i paesi confinanti con la nazione ospitante a supportarla nella gara per aggiudicarsi i Giochi. Per quanto riguarda il sistema politico Johnson e Ali mostrano che i paesi in cui esiste ancora la monarchia performano peggio rispetto ai paesi in cui vige il regime Comunista, si tratta di una differenza di circa 12 medaglie a favore di questi ultimi. Le ex potenze coloniali denotano effetti simili a quelli già visionati sul tema della partecipazione, con le ex colonie di grandi nazioni che presentano minori speranze di vincere una medaglia mentre quelle di nazioni vincitrici di medaglie hanno una maggior probabilità di successo. Johnson e Ali concludono la loro ricerca con la previsione

delle medaglie vinte dalle varie nazioni alle Olimpiadi di Sidney 2000, che non si rivela essere così precisa, soprattutto per quanto riguarda la Russia e la Cina.

Nel 2001 il lavoro di Johnson e Ali viene ripreso da Gerard Kuper e Elmer Sterken, i quali però si concentrano sulle Olimpiadi invernali. L'obiettivo principale della loro ricerca è determinare i fattori geografici, economici e demografici alla base della partecipazione e del successo ai Giochi Olimpici invernali, in modo da essere in grado di produrre delle previsioni attendibili. Mentre i Giochi estivi sono un argomento trattato e studiato numerose volte, sono pochi i ricercatori che hanno analizzato le Olimpiadi invernali, questo poiché sono eventi più ridotti sia in termini di partecipazione che di gare: la partecipazione è quasi esclusivamente limitata a paesi con determinate condizioni climatiche, riducendo così il numero di competizioni sportive. Lo studio si basa sui dati di tutte le Olimpiadi invernali a partire dal 1924. Kuper e Sterken lavorano calcolando sia la partecipazione che il successo in percentuale invece che in valore assoluto. Mentre Johnson e Ali modellano la decisione di partecipazione e la probabilità di successo separatamente, Kuper e Sterken stimano il numero di partecipanti e, su questo dato, prevedono il successo. Quindi prima stimano la partecipazione e basandosi su questa stimano il successo. Per calcolare la partecipazione usano la variabile PSH_{it}, che rappresenta la percentuale di atleti inviata da un paese sul totale degli atleti olimpici. Il numero di atleti inviati da una nazione ai Giochi è dipendente dalla popolazione totale del paese, seppure in modo non proporzionale. La variabile PSH sarà quindi determinata dalla radice quadrata del logaritmo della popolazione. Similarmente il PIL pro capite influenza le condizioni di allenamento e di salute, influenzando così la partecipazione. La variabile dipendente di partecipazione è inoltre condizionata dall'effetto di essere la nazione ospitante e dalla distanza tra il paese partecipante e quello ospitante.

$$\log PSH_{it} = a \sqrt{\log(POPSH_{it})} + b \log(YSH_{it}) + c*HOME_{it} + d*DIST_{it} + C_i + e_{it}$$

Kuper e Sterken raccolgono all'interno della variabile C_i i dati geografici, come la latitudine e la presenza di montagne, il sistema legale del paese e l'emancipazione femminile, misurata come partecipazione al lavoro femminile. La variabile C_i comprende anche una variabile dummy hockey su ghiaccio, creata poiché l'hockey è l'unico sport di squadra dei Giochi invernali e da solo porta più di 20 atleti. Quasi tutte le variabili hanno un impatto significativo sulla partecipazione. Il vantaggio del giocare in casa non risulta essere un determinante convincente, in opposizione ai risultati trovati negli studi sulle Olimpiadi estive condotti dagli stessi Kuper e Sterken. Un secondo modello proposto dai due studiosi olandesi, simile a quello appena visto, asserisce che ci vogliono più di quattro anni di allenamento ed esperienza per aumentare significativamente il numero di atleti che un paese invia ai Giochi. Dai risultati risulta quindi evidente che la quota di partecipazione precedente spieghi quella attuale in modo significativo. I due ricercatori olandesi hanno analizzato il successo alle Olimpiadi in termini di medaglie vinte, MSH, come percentuale del totale di medaglie a disposizione. L'innovazione portata da Sterken e Kuper sta nel fatto che modellizzano la quota di medaglie di un paese come funzione della sua quota di partecipazione, e poiché la partecipazione è una variabile endogena utilizzano quella calcolata in precedenza. I due accademici definiscono la percentuale di medaglie vinte come funzione della quota di partecipazione, del sistema legale, del vantaggio dell'essere il paese ospitante e nuovamente del reddito pro capite. Sterken e Kuper utilizzano anche due dummy variable per gli USA e l'URSS, poiché questi due paesi mostrano dei risultati straordinari non essendo stati esclusi dall'analisi i Giochi affetti da boicottaggio. Analizzando i risultati, visibili nella Tab. 2.7, sembra che l'essere la nazione ospitante influenzi soprattutto la vittoria di medaglie d'argento, questo trend contrasta con quello emerso dall'analisi delle Olimpiadi estive, dove l'home advantage aveva come incremento principale quello delle medaglie d'oro vinte. I paesi che adottano un sistema legale di tipo Scandinavo o Tedesco normalmente vincono più medaglie, lo stesso vale per gli USA e l'Unione Sovietica che tendono a vincere più medaglie di quanto previsto. Infine, la stima della quota di partecipazione risulta essere un fattore determinante significativo per il successo olimpico. La seconda tabella in basso modifica lievemente il modello aggiungendo una variabile dummy anche per la Russia, e evidenzia come nelle ultime cinque edizioni dei Giochi invernali l'essere il paese ospitante non sia più un fattore determinante; inoltre la Russia sembra aver finalmente preso il posto dell'URSS come numero di vittorie. I paesi con sistema legale tedesco si confermano quelli con le migliori performance.

Panel A - All Games: 1924-1998

	Gold	Silver	Bronze
log(Home+1)	0.678	0.816	0.041
	(0.291)	(0.297)	(0.282)
log(Socialist+1)	-0.395	-0.212	-0.481
	(0.138)	(0.141)	(0.134)
log(Scandinavian+1)	1.531	1.659	1.093
	(0.189)	(0.193)	(0.183)
log(British+1)	-0.071	-0.144	-0.184
	(0.163)	(0.169)	(0.158)
log(German+1)	0.573	0.770	0.381
	(0.182)	(0.185)	(0.176)
$\log(USA+1)$	1.949	2.052	1.078
	(0.317)	(0.322)	(0.307)
$\log(USSR+1)$	3.594	2.999	2.657
,	(0.391)	(0.399)	(0.379)
$\log(PSH^e)$	0.248	0.252	0.443
	(0.041)	(0.042)	(0.040)

Tab. 2.7: I risultati della regressione

Panel C - Post-World War II Games: 1984-1998

	Gold	Silver	Bronze
log(Home+1)	0.234	0.817	-0.202
	(0.577)	(0.618)	(0.407)
log(Socialist+1)	-0.454	-0.263	-0.526
	(0.180)	(0.185)	(0.168)
log(Scandinavian+1)	0.800	0.601	0.453
	(0.408)	(0.356)	(0.336)
log(British+1)	-0.204	-0.166	-0.245
	(0.168)	(0.167)	(0.160)
log(German+1)	0.965	0.808	0.544
	(0.319)	(0.328)	(0.272)
$\log(USA+1)$	1.907	1.349	0.052
	(0.317)	(0.420)	(0.440)
$\log(USSR+1)$	3.201	2.876	2.632
	(0.290)	(0.390)	(0.307)
$\log(RUS+1)$	3.132	2.376	1.157
	(0.235)	(0.266)	(0.254)
$\log(PSH^e)$	0.275	0.299	0.422
	(0.054)	0.0)	55)

Tab. 2.8: I risultati della regressione

I due studiosi olandesi concludono il lavoro con una previsione della partecipazione e del successo per le Olimpiadi di Salt Lake City del 2002 utilizzando i modelli basati sulle ultime cinque edizioni. Essi chiudono la ricerca affermando che la partecipazione è determinata da fattori economici e geografici, mentre il successo è fortemente legato all'home advantage e nelle ultime edizioni anche al PIL pro capite.

Country	Participants	Country	Gold	Silver	Bronze	
USA	373	RUS		18	13	7
CAN	195	USA		12	14	4
GER	171	GER		6	6	8
RUS	160	AUT		4	4	5
SWE	115	SUI		4	4	5
FRA	103	SWE		4	4	5
AUT	102	FIN		4	3	5
SUI	102	JPN		3	3	3
FIN	90	KOR		3	2	3
UKR	90	NOR		2	1	2
ITA	83	CAN		2	2	4
JPN	82	DEN		2	1	1
CZE	77	LIE		2	1	1
NOR	64	FRA		1	1	2
BLR	62	ICE		1	1	1
KAZ	57	CHN		1	1	2
GBR	54	CZE		1	1	2
CHN	46	SLO		1	1	1
SLO	46	ITA		1	1	1
LAT	45	BLR		1	1	1

Tab. 2.9: Le previsioni Sterken e Kuper

Come si può osservare dalla Tab. 2.9, confrontando le previsioni di Sterken e Kuper con i dati reali si nota che mentre non vi è un grosso scostamento nella partecipazione lo stesso non si può dire per il successo. Il paese che ha vinto il maggior numero di medaglie d'oro è infatti la Norvegia, seguita dalla Germania che risulta essere il paese con il maggior numero di medaglie vinte. Il medagliere finale le vede entrambe davanti a USA e Russia. Il Canada che secondo le previsioni avrebbe dovuto vincere solamente 8 medaglie, ne ha vinte in realtà 17, lo stesso vale per la Svizzera, a cui erano state pronosticate 13 medaglie ed in realtà non è entrata neanche nelle prime dieci nazioni del medagliere.

Nel 2001 viene pubblicato dal Professore francese Wladimir Andreff uno studio che ha dato una grossa spinta a tutta la letteratura dei Giochi olimpici, soprattutto a quella dei paesi francofoni. Andreff è stato uno dei primi ricercatori a dimostrare una correlazione tra il livello di sviluppo economico di un paese ed il suo livello di sviluppo sportivo con il lavoro dal titolo "The Correlation between Economic Underdevelopment and Sport". Questa ricerca afferma che qualunque analisi economica sulle relazioni tra sport e sviluppo economico non può fare a meno di dimostrare che empiricamente nei paesi meno sviluppati dal punto di vista economico non si ha una crescita delle prestazioni sportive, del numero d'impianti sportivi, di coloro che praticano attività sportiva e dei finanziamenti per lo sport. Infatti, la stragrande maggioranza dei paesi del Terzo Mondo dipende per quanto riguarda lo sport dagli aiuti dall'estero, dall'importazione di articoli e attrezzature sportive e dagli sponsor multinazionali. Allo stesso tempo però questi paesi subiscono una drammatica fuga dei loro atleti più talentuosi per competere e allenarsi nei paesi più sviluppati. Risulta inoltre molto difficoltoso raccogliere dati statistici sullo sport nei paesi meno sviluppati. Per molti bambini il primo contatto con lo sport avviene nelle scuole primarie, la ricerca mostra quindi che, definendo il rapporto netto di iscrizione nelle scuole primarie come il rapporto tra il numero di bambini iscritti e l'intera fascia di età pertinente, il rapporto netto di iscrizione è superiore al 95% in tutti i paesi sviluppati, mentre la situazione risulta drammatica nei paesi meno sviluppati dove il tasso di iscrizione scende al 30% in Eritrea, al 28% in Etiopia e Mali e al 25% in Niger. In altre parole, un bambino che vive in Niger ha una probabilità di poter praticare educazione fisica in una scuola primaria pari ad un quarto della probabilità di un bambino che vive in Cina, dove il tasso di iscrizione raggiunge 100%. La durata e la frequenza delle lezioni di educazione fisica avranno una profonda influenza nell'incentivare bambini e adolescenti a praticare sport. Nei paesi in via di sviluppo le ore dedicate all'educazione fisica spesso non vengono mai effettivamente svolte. Nelle scuole primarie

questa mancanza è dovuta principalmente alla carenza di insegnanti ed al numero troppo elevato di alunni, da 40 a 100 per docente, mentre nelle scuole secondarie è attribuibile all'assenza di impianti ed attrezzature sportive ed alla carenza di docenti qualificati. Andreff afferma che il modo più semplice per ottenere delle informazioni riguardo il livello di partecipazione sportiva in un paese è confrontare il numero di membri affiliati alle federazioni sportive con il numero di abitanti, e sottolinea come nel 1995 in Africa un solo abitante su 800 praticasse sport, uno su 707 in Asia e uno su 100 in America Latina, il dato è impietoso se confrontato con l'Europa, dove un cittadino su quattro risulta essere iscritto ad una federazione sportiva. Inoltre, nei paesi in via di sviluppo anche lo sport agonistico risulta essere influenzato dalla carenza di allenatori e dall'assenza, in molti stati, di un sistema medico sanitario dedicato. Lo studio rivela che il governo è spesso il principale sponsor dell'attività sportiva nei paesi meno sviluppati; da questi dati emerge che circa l'80% di questi paesi sovvenziona le federazioni sportive, ma che, ad eccezione del calcio, l'importo dei sussidi è troppo basso per garantire lo svolgimento delle attività sportive per tutta la durata dell'anno. Nei paesi in via di sviluppo presi in considerazione dallo studio le infrastrutture e gli impianti sportivi sono numericamente insufficienti e talvolta addirittura assenti, e spesso non vengono nemmeno fatti i lavori di manutenzione e mantenimento. La manutenzione degli impianti sportivi comporta costi molto elevati così le autorità riducono il loro utilizzo a pochi eventi, il che fa a sua volta diminuire il numero delle competizioni sportive locali e quindi l'attività sportiva delle società e delle federazioni sportive. Diminuisce di conseguenza anche il numero di partecipanti allo sport e la quantità di tempo dedicata all'attività fisica. La conseguenza è che gli impianti sportivi sono sottoutilizzati, la loro manutenzione non è ritenuta né utile né necessaria, creando quindi un pericoloso circolo vizioso. Andreff raccoglie nella Tab. 2.10 le medaglie vinte in tutti i Giochi olimpici invernali fino al 1998, suddividendole in base ai diversi livelli di sviluppo economico dei paesi:

Country	Gold	Silver	Bronze	Total
Developed countries*	575	585	566	1726
Emerging countries*	38	31	36	105
Middle-income countries*	4	4	6	14
Least developed countries*	0	0	. 0	0
Total	617	620	608	1845

Tab. 2.10: Distribuzione delle medaglie per livello di sviluppo economico

Osservando questa tabella il celebre ricercatore francese si rese conto che il numero di medaglie vinte da un paese potrebbe essere correlato al livello di sviluppo economico ed alla dimensione della popolazione dei due anni precedenti all'anno delle Olimpiadi, poiché un paese che intende partecipare ai Giochi deve preparare gli atleti almeno con due anni di anticipo. Andreff utilizza un modello di regressione logit ordinario, dove la variabile dipendente da spiegare è il numero di medaglie vinto da un paese alle Olimpiadi. Il modello suddivide i 181 paesi dello studio in quattro classi in base alle loro performance:

- Classe M1: zero medaglie
- Classe M2: 1 o 2 medaglie, con almeno un oro, o più di 2 medaglie senza oro
- Classe M3: da 3 a 9 medaglie, con almeno un oro
- Classe M4: 10 medaglie e oltre, con almeno un oro.

Allo stesso modo anche livello di sviluppo economico e la dimensione della popolazione, che rappresentano le variabili indipendenti, sono suddivise in quattro classi di appartenenza. Per il reddito:

- Classe G1: Paesi meno sviluppati con un PIL pro capite inferiore ai \$600,
- Classe G2: MIC (middle income countries) con un PIL pro capite superiore ai \$600 ma inferiore ai \$ 2.000

- Classe G3: EMC (emerging countries) con un PIL pro capite superiore ai \$2.000 ma inferiore ai \$9.000
- Classe G4: Paesi sviluppati un con PIL pro capite superiore ai \$9.000.

Per la popolazione:

- Classe P1: paesi con meno di 2 milioni di abitanti
- Classe P2: paesi di medie dimensioni con più di 2 milioni di abitanti ma meno di 29 milioni di abitanti
- Classe P3: grandi paesi con più di 29 milioni di abitanti ma meno di 100 milioni di abitanti
- Classe P4: paesi molto grandi con una popolazione superiore a 100 milioni di abitanti.

Dai risultati della regressione si osserva che per le Olimpiadi del 1996, entrambe le variabili indipendenti sono estremamente significative. Entrambi i coefficienti sono positivi dimostrando che un paese ha un'alta probabilità di passare a una classe di medaglie superiore quando il suo PIL pro capite o la sua popolazione sono in aumento. Tuttavia, il passaggio è più probabile con un incremento della dimensione della popolazione che del PIL pro capite. Stessa cosa può essere osservata per i Giochi di Sidney 2000. Altrettanto significativi, ma leggermente differenti, sono i risultati per i Giochi invernali fino al 1998: il coefficiente del reddito è ancora positivo ed è maggiore di quello della popolazione, essendo la prima variabile più influente della seconda sulla probabilità di passare ad una classe di medaglie superiore. Infatti, il livello di sviluppo economico determina in modo significativo la probabilità di un paese di vincere delle medaglie ai Giochi Olimpici; lo sviluppo economico è quindi un fattore determinante per le prestazioni sportive. Tuttavia, questo effetto è mitigato dalla dimensione della popolazione di un paese. Inoltre, Andreff mette in evidenza il fatto che le medaglie vinte dai paesi del Terzo Mondo alle Olimpiadi dal 1896 in poi, sono concentrate in pochi sport: l'80% delle medaglie vinte dai paesi africani sono legate all'atletica ed al pugilato. I paesi asiatici si sono specializzati nelle arti marziali e nella boxe. L'America Latina, invece, sarebbe praticamente inesistente sulla scena internazionale senza il calcio ed il tennis. Per ovvie ragioni geografiche e climatiche, la maggior parte dei paesi in via di sviluppo è ancora meno capace di conquistare medaglie alle Olimpiadi invernali. Andreff pone l'accento sul fatto che circa trenta paesi sviluppati gestiscano l'organizzazione del 95% di tutti gli eventi sportivi internazionali e mondiali. Il resto del mondo, indicativamente 170 paesi in via di sviluppo, ospita non più del 5% dei grandi eventi sportivi. Normalmente le competizioni internazionali localizzate nel Terzo Mondo sono in parte finanziate dall'estero o sono eventi sportivi internazionali a carattere regionale. Spesso poi, questi ultimi vengono rinviati o annullati a causa di problemi finanziari. Ospitare un evento sportivo internazionale di solito è una spesa enorme per il governo di un paese in via di sviluppo, e la maggior parte dei paesi africani, che non sono nemmeno in grado di partecipare ai Giochi africani o ai Mondiali di calcio a causa delle finanze inadeguate, non hanno certo le possibilità di finanziare l'organizzazione di un evento sportivo internazionale. Inoltre, quando i governi di questi paesi decidono di affrontare questa spesa, di solito è a carico dell'intera popolazione che, in quanto contribuenti, alla fine coprirà il deficit. Tuttavia, secondo Andreff, ospitare le Olimpiadi è l'unica opportunità per sviluppare gli impianti sportivi e le infrastrutture di un paese. Il celebre ricercatore parigino affronta poi il tema della cosiddetta "fuga di muscoli", ovvero gli atleti di livello internazionale e gli sportivi di talento che non appena raggiungono un certo standard di prestazioni e risultati sportivi, lasciano il loro paese per andare ad allenarsi nei paesi più sviluppati; ad esempio, i migliori giocatori di basket africani emigrano e talvolta cambiano addirittura cittadinanza. Anche i paesi dell'America Latina sono affetti da questa fuga di atleti, infatti, dal 1989 al 1997, ben 2.084 calciatori brasiliani sono emigrati in club professionistici stranieri, in particolare Italia, Spagna, Portogallo, Francia e Giappone. La fuga di muscoli spiega in parte le scarse prestazioni dei paesi poco sviluppati o in via di sviluppo negli eventi sportivi internazionali. Priva, infatti, il paese di origine anche di quei pochi e talentuosi atleti che hanno avuto l'opportunità di beneficiare dei rari allenatori

domestici e degli ancor più rari impianti sportivi. Nei casi in cui gli atleti cambino anche cittadinanza, il paese di origine perde la capacità di utilizzare i suoi atleti di maggior talento nelle competizioni internazionali. La "fuga di muscoli" si trasforma poi in una "fuga di cervelli" quando interessa anche gli allenatori ed i tecnici. Wladimir Andreff però propone anche alcune soluzioni per cercare di arginare questi problemi. Le poche soluzioni possibili richiedono tutte la volontà politica ed economica del governo di promuovere e sostenere lo sport di base invece di dare la priorità alla speranza di trionfare in competizioni internazionali o di migliorare la reputazione del paese con l'organizzazione di un evento sportivo mondiale. Una prima soluzione è il recupero di quelli sport considerati tradizionali della nazione, la cui pratica non necessita di sofisticati impianti sportivi o di equipaggiamenti sportivi costosi, che spesso però non fanno parte degli sport olimpici. Una seconda soluzione è la costituzione di un circuito di gare per i giovani atleti. Infine, Andreff propone l'utilizzo di una tassa, detta Coubertobin tax, con un'aliquota dell'1%, su tutte le spese di trasferimento e sul salario iniziale di ogni contratto firmato da atleti e giocatori di paesi in via di sviluppo con partner stranieri, come club professionistici, sponsor o agenti. La tassa di Coubertobin ha come obbiettivo principale quello di rallentare "la fuga di muscoli", tuttavia un vantaggio inaspettato potrebbe essere quello di ridurre il differenziale del costo del lavoro tra il paese di cui è originario l'atleta e quello sviluppato meta del suo trasferimento, disincentivando atleti e giocatori a lasciare il proprio paese. I soldi raccolti attraverso questa tassa andranno in un fondo per lo sviluppo sportivo utilizzato per finanziare la costruzione e la manutenzione di impianti sportivi e i corsi di educazione fisica nelle scuole. Andreff conclude dicendo che non è possibile trovare una soluzione a lungo termine per migliorare lo sport nei paesi in via di sviluppo senza che vi sia anche una reale crescita economica. Mentre lo sport può contribuire solo in minima parte allo sviluppo economico, quest'ultimo è ovviamente un fattore determinante per lo sviluppo dello sport.

L'anno dopo, nel 2002, esce la versione definitiva di un articolo scientifico che era già stato pubblicato nel 2000 e che sarebbe stato destinato a diventare una pietra miliare nello studio dei determinanti del successo Olimpico, rappresentando il lavoro più autorevole in questo campo e sul quale tutti i ricercatori venuti in seguito si sono basati e continuano a basarsi. Andrew B. Bernard e Meghan R. Busse nel loro studio dal titolo "Who Wins the Olympic Games: Economic Resources and Medal Totals" esaminano a livello nazionale i determinanti del successo ai Giochi Olimpici di un paese. I due ricercatori americani si chiedono quindi quante medaglie un paese dovrebbe vincere basandosi sulle variabili che determinano il successo olimpico. Essi affermano che nonostante la popolazione sia una variabile significativa del successo non è l'unica, infatti paesi come Cina, India, Indonesia e Bangladesh pur rappresentando il 43% della popolazione mondiale, nel 1996 hanno vinto poco più del 6% delle medaglie disponibili. Quindi Bernard e Busse aggiungono alla popolazione anche il PIL pro capite, portando così un sostanziale miglioramento del proprio modello. Il PIL reale è il miglior previsore delle restazioni olimpiche di un paese. Bernard e Busse iniziano la loro analisi con la congettura che immaginando gli atleti di talento equamente distribuiti tra tutti i paesi, la quota di medaglie prevista per un paese dovrebbe essere pari alla sua quota della popolazione totale dei paesi partecipanti alle Olimpiadi.

$$E\left(\text{medalshare}_{it}\right) = \frac{\text{medals}_{it}}{\sum_{j} \text{medals}_{jt}} = \frac{\text{population}_{it}}{\sum_{j} \text{population}_{jt}} = \text{popshare}_{it}.$$

Tuttavia, questa equazione non funziona completamente poiché i paesi non possono inviare atleti in proporzione alla propria popolazione per qualunque evento, ad esempio negli sport di squadra è possibile inviare un solo team. In secondo luogo gli sport di squadra portano una sola medaglia nonostante vi partecipino più atleti. Inoltre il numero massimo di atleti che un paese può inviare

alle Olimpiadi viene concordato con il Comitato Internazionale Olimpico e quindi non tutti gli atleti che potenzialmente potrebbero vincere una medaglia potranno partecipare, questo discorso vale in particolare per coloro che provengono da nazioni molto popolate. Se come detto in precedenza il talento e l'abilità sono distribuiti in modo omogeneo tra i vari paesi, in realtà formare e allenare atleti in grado di vincere una medaglia richiede un ingente investimento in infrastrutture e personale, favorendo quindi i paesi sviluppati e ricchi. La funzione di produzione per generare atleti di talento si basa sulla popolazione, sul PIL e sull'abilità organizzativa del paese. La quota di medaglie vinte ai Giochi è funzione di come viene sviluppato il talento.

$$\operatorname{E}\left(rac{\operatorname{medals}_{it}}{\sum_{j}\operatorname{medals}_{jt}}
ight)=M_{it}^{st}=g\left(T_{it}
ight).\hspace{0.5cm}T_{it}=f\left(N_{it},Y_{it},A_{it}
ight)$$

I due ricercatori americani utilizzano una funzione di produzione Cobb-Douglas per determinare la popolazione ed il PIL, e una funzione logaritmica per determinare la quota di medaglie vinte. Il modello iniziale definito da Bernard e Busse è quindi

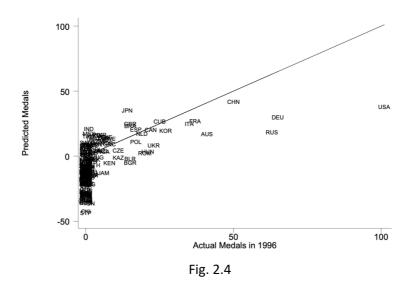
$$M_{it} = C + \alpha \ln N_{it} + \beta \ln \left(\frac{Y}{N}\right)_{it} + d_t + \nu_i + \epsilon_{it}$$

Dove dt è una variabile dummy che indica l'anno per cogliere l'effetto di una variazione del numero di atleti e nazioni partecipaneti e del numero di medaglie disponibili. L'analisi dei risultati mostra che vi è una similarità tra i coefficienti della logaritmica della popolazione e della logaritmica del PIL pro capite, questo significa che il vero determinante del successo olimpico è la logaritmica del Prodotto Interno Lordo. Tuttavia, Bernard e Busse si rendono conto che il loro modello non è completo e presenta numerose imperfezioni. In primo luogo, essi non considerano il fatto che il paese organizzatore gode del vantaggio di giocare in casa; infatti, il costo di far partecipare gli atleti è minimizzato, essi possono allenarsi nei luoghi e nelle condizioni in cui si svolgeranno le gare, saranno quindi più motivati ed anche la presenza della folla può influenzare sia gli atleti che i giudici. I due studiosi americani creano inoltre due variabili dummy: una per coprire i paesi sotto la sfera d'infleunza sovietica e l'altra per gli altri paesi comunisti. I paesi che fanno parte di queste due categorie sono noti per vincere un enorme numero di medaglie, più di quanto sarebbe prevedibile considerando i determinanti economici e sociali. Il secondo modello proposto è quindi:

$$M_{it} = C + \alpha \ln N_{it} + \beta \ln \left(\frac{Y}{N}\right)_{it} + \text{Host}_{it} + \text{Soviet}_{it} + \text{Planned}_{it} + d_t + \nu_i + \epsilon_{it}$$

Anche con questo modello la popolazione ed il GDP pro capite rimangono determinanti significativi. La variabile Soviet è positiva e significativa: infatti i paesi sotto la sfera di influenza sovietica hanno quote di medaglie superiori di oltre 6,1 punti percentuali rispetto agli altri paesi. Stesso discorso vale per la variabile Planned, gli Stati a pianificazione economica centralizzata hanno quote che sono più alte di circa 1,6 punti percentuali. L'effetto di essere il paese ospitante è anch'esso positivo e significativo, l'incremento delle medaglie vinte derivante dall'ospitare un'Olimpiade non boicottata supera i 2 punti percentuale. Durante i Giochi boicottati l'host advantage era enorme, circa del 19%, tenendo presente che entrambi gli URSS e gli USA sono stati i primi beneficiari del boicottaggio dell'avversario. Tuttavia anche questo modello risulta incompleto, osservando la Fig 2.4, che

raffigura la previsione del medagliere per i Giochi del 1996, esso tende a sottostimare il numero di medaglie per i paesi che stanno alle estremità del range e a sovrastimarlo per quelli che stanno nel mezzo.



Bernard e Busse iniziano a considerare gli atleti come assimilabili a dei beni capitali durevoli, essi infatti possono vincere medaglie in più Olimpiadi diverse. Questa considerazione suggerisce che gli investimenti per un'Olimpiade possono aumentare la probabilità di vincere medaglie nelle Olimpiadi successive. Per catturare quest'effetto aggiungiamo una variabile che indica il numero di medaglie vinte nei Giochi precedenti:

$$M_{it} = C + (1 - \delta) M_{it-1} + \alpha \ln N_{it} + \beta \ln \left(\frac{Y}{N}\right)_{it} + \delta_t + \epsilon_{it}$$

Osservando i risultati dell'analisi di regressione si nota che anche in questo caso la popolazione ed il reddito pro capite mantengono un effetto positivo e significativo. La quota di medaglie vinta nei Giochi precedenti è fortemente significativa. Infine le variabili Soviet, Planned e Host rimangono significative e hanno l'effetto positivo di incrementare il numero di medaglie vinte. Il fatto che il reddito pro capite rimanga sempre rilevante dimostra che le risorse economiche sono fondamentali nel produrre medagliati olimpici. Più sorprendente è la persistente somiglianza dei coefficienti del logaritmo della popolazione e del logaritmo del PIL pro capite. Questo risultato suggerisce che il PIL totale di un paese è determinante nella produzione degli atleti olimpici. Questo implica che due paesi con lo stesso PIL vinceranno probabilmente lo stesso numero di medaglie, anche se uno è relativamente popoloso con un reddito pro capite più basso e l'altro è più piccolo con un PIL pro capite più alto. Per valutare il loro modello i due autori hanno svolto una previsione delle medaglie vinte dalle nazioni partecipanti alle Olimpiadi estive di Sidney 2000 e per quelle di Atlanta 1996; per quest'ultime il modello funzionò correttamente per USA, Cina e Russia, sottostimando però le vittorie di Italia, Francia e Austria e sovrastimando quelle della Germania. La Tab. 2.11 contiene due previsioni per le Olimpiadi del 2000, nella prima i due autori hanno commesso degli errori nella raccolta dei dati, mentre la seconda è apparentemente esente da errori. Entrambi i set di previsioni forniscono risultati abbastanza accurati e precisi. La seconda stima indovina il valore esatto di un solo paese, tuttavia ben 9 paesi sono all'interno del range ± 1 medaglia e 23 si trovano in quello ± 3 medaglie. Questo studio è diventato da subito un classico su cui si basa gran parte della letteratura del XXI secolo sulle previsioni delle medaglie alle Olimpiadi.

	Predicted 2000	Standard error	Predicted 2000	Actual 2000
	medals (1)	estimate	medals (2)	medals
Australia	57	(5.1)	52	58
Belarus	12	(5.0)	12	17
Belgium	7	(5.1)	7	5
Brazil	18	(5.1)	17	12
Bulgaria	11	(5.1)	10	13
Canada	24	(5.1)	23	14
China	51	(6.2)	49	59
Cuba	21	(7.1)	20	29
Czech Republic	10	(5.0)	9	8
Denmark	7	(5.1)	7	6
France	39	(5.1)	38	38
Germany	66	(5.5)	63	57
Greece	8	(5.1)	8	13
Hungary	19	(5.1)	18	17
Italy	37	(5.1)	35	34
Jamaica	1	(5.1)	1	7
Japan	20	(5.2)	19	18
Kazakhstan	9	(5.0)	8	7
Kenya	5	(5.1)	5	7
Netherlands	20	(5.1)	19	23
New Zealand	5	(5.1)	5	4
Nigeria	5	(5.1)	5	3
North Korea	3	(6.2)	3	4
Norway	7	(5.1)	7	10
Poland	17	(5.0)	16	14
Romania	18	(5.1)	17	26
Russia	62	(5.4)	59	88
South Africa	6	(5.1)	6	5
South Korea	28	(5.1)	27	28
Spain	19	(5.1)	18	11
Sweden	9	(5.1)	9	12
Switzerland	8	(5.1)	8	9
Turkey	7	(5.1)	7	4
UK	18	(5.1)	18	28
Ukraine	22	(5.9)	21	23
US	102	(5.1)	97	97

Tab. 2.11: Previsioni del medagliere per le Olimpiadi del 2000

Daniel K.N. Johnson e Ayfer Ali, nel 2000 avevano pubblicato una ricerca estremamente importante sui fattori che influenzano la partecipazione ed il successo ai Giochi olimpici estivi. Nel 2002 hanno prodotto uno studio simile che prendeva in considerazione entrambi le Olimpiadi, sia quelle invernali che quelle estive. La ricerca esamina i Giochi olimpici post Seconda Guerra Mondiale in modo da definire i fattori economici e politici che determinano la partecipazione ed il successo nella manifestazione. Secondo i due autori alle Olimpiadi invernali il numero di paesi che hanno la possibilità di partecipare e di vincere è ancora più ridotto rispetto a quelle estive. Infatti, ai Giochi invernali del 1998 parteciparono solamente 2077 atleti provenienti da 72 nazioni contro i 12000 da 200 paesi delle Olimpiadi estive. Nella storia delle Olimpiadi invernali nessun atleta rappresentante di un paese senza neve si è mai classificato nei primi dieci ad una gara. Johnson e Ali iniziano elencando una serie di fattori che intuitivamente potrebbero essere legati alla partecipazione ed al successo di un paese ai Giochi. Ad esempio, un elevato PIL pro capite permette di sostenere i costi necessari ad inviare gli atleti alla manifestazione, ma permette anche di fornire un livello di allenamento e di attrezzature superiore. Una popolazione estesa permette di spalmare i costi fissi di allenamento su un maggior numero di persone e di avere un pool più ampio di potenziali atleti a cui attingere. Come già visto in precedenza il paese ospitante e quelli limitrofi hanno minori costi di trasporto, oltre il vantaggio di svolgere la preparazione nelle condizioni climatiche in cui si svolgeranno le gare, questo è un vantaggio enorme per le Olimpiadi invernali. I due studiosi suppongono che anche la struttura politica possa avere un'influenza sui risultati. Johnson e Ali scelgono di non includere come variabili né il numero di partecipanti né il numero di medaglie vinte nei Giochi precedenti. Queste variabili benché siano molto significative e siano ottimi previsori non rappresentano variabili di tipo economico o politico. Il Comitato Olimpico Internazionale, conscio del fatto che non tutti i paesi sono in grado di inviare uno stesso numero di atleti, fissa prima dei Giochi delle quote di partecipazione. Per trovare i fattori che determinano il numero di partecipanti di un paese alle Olimpiadi i due autori utilizzano una semplice funzione lineare:

$$part_{i} = \alpha + \beta_{1}GDP + \beta_{2}GDP^{2} + \beta_{3}POP + \beta_{4}POP^{2} + \beta_{5}HOME + \beta_{6}NEIGH$$
$$+ \sum_{j=7}^{10} \beta_{j}POL_{j} + \beta_{11}LFROST + \beta_{12}HFROST + \beta_{13}t + u_{i} + \varepsilon$$

Questa funzione ha come variabili il PIL pro capite e la popolazione, inclusi i loro quadrati. Tra gli altri fattori considerati da Johnson e Ali troviamo l'essere il paese ospitante o un suo limitrofo e la struttura politica. Infine, essi identificano due variabili per indicare la percentuale di terra coperta dai ghiacci e dalla neve. Osservando i risultati dell'analisi di regressione si nota come per le Olimpiadi invernali il quadrato del PIL pro capite sia fortemente significativo, sottolineando come la partecipazione aumenta con proporzionalità quadratica con il reddito. Le nazioni a basso reddito tendono a non mandare gli atleti alle competizioni, mentre le nazioni con reddito elevato inviano contingenti di atleti molto numerosi. L'impatto del PIL pro capite sulla partecipazione femminile per le Olimpiadi invernali è nullo. Mentre per i Giochi estivi i paesi con una popolazione numerosa tendono a mandare più atleti, per quelli invernali quest'effetto è totalmente assente. L'effetto paese ospitante è presente anche nelle Olimpiadi invernali anche se in maniera meno evidente, la nazione organizzatrice invia in media circa 32 atleti in più. Tuttavia, essendo le nazioni ospitanti anche quelle più sviluppate, vi è una alta correlazione tra l'host advantage ed il reddito pro capite e non è chiaro se lo stesso effetto sarebbe possibile se il paese ospitante fosse uno in via di sviluppo. Per i Giochi invernali i paesi limitrofi a quello ospitante non godono dell'aumento di atleti inviati. La struttura politica non ha nessun tipo di influenza in questo contesto, mentre al contrario, la vera scoperta fatta da Johnson e Ali è che un clima freddo sembra avere un effetto positivo sulla partecipazione sia per le Olimpiadi estive che per quelle invernali. Mentre qualsiasi clima invernale incoraggia la partecipazione ai Giochi estivi, solo i climi invernali pesanti hanno un effetto positivo sulla partecipazione ai Giochi invernali. Questo anche perché i paesi con clima più freddo o quantomeno temperato sono anche quelli più ricchi. Sulla base dei valori di queste variabili Johnson e Ali hanno proposto una previsione del numero di atleti che ogni nazione manderà alle Olimpiadi di Salt Lake City 2002, come riporta la Tab 2.12.

Nation	2002	1998	2002 predicted
	predicted	actual	dollar "cost"
	athletes [rank]	athletes [rank]	per capita of
			additional athlete
			(and % GDP)
U.S.A.	177 [1]	192 [1]	185 (0.90)
Russia*	140 [2]	126 [4]	n/a
Germany	114 [3]	131 [3]	212 (1.20)
Canada	101 [4]	81 [10]	197 (1.03)
Sweden	94 [5]	104 [7]	219 (1.29)
Italy	90 [6]	119 [5]	246 (1.62)
Switzerland	87 [7]	74 [11]	190 (0.96)
Norway	84 [8]	83 [9]	240 (1.55)
Austria	82 [9]	104 [8]	246 (1.62)
France	77 [10]	112 [6]	228 (1.39)
Czech Republic*	73 [11]	66 [12]	n/a
Japan*	73 [11]	175 [2]	n/a
Kazakhstan*	71 [13]	64 [13]	n/a
Finland	71 [13]	27 [23]	247 (1.64)
Belarus*	69 [15]	62 [14]	n/a
P.R. of China*	67 [16]	60 [15]	n/a
Ukraine*	62 [17]	56 [16]	n/a
Great Britain	61 [18]	35 [20]	249 (1.67)
Slovakia*	44 [19]	40 [18]	n/a
Poland	43 [20]	41 [17]	751 (17.37)
All nations	1485	1332	268 (2.69)
with data			
All nations	2316	2077	n/a

Tab. 2.12: Previsione del numero di atleti per nazione alle Olimpiadi di Salt Lake City 2002

I due autori stimano un aumento dell'undici percento della partecipazione totale rispetto alla precedente manifestazione olimpica; il risultato più inappropriato riscontrato è il basso numero di partecipanti degli USA, soprattutto se si pensa che è anche la nazione ospitante. Un'altra importante deduzione che ottengono è che il PIL pro capite addizionale richiesto per inviare un atleta in più alla competizione risulta minore per i paesi ricchi, sia in termini assoluti che in percentuale, confermando la proporzionalità quadratica tra la partecipazione ed il reddito, sottolineando quindi l'enorme vantaggio che i paesi sviluppati hanno ai Giochi olimpici. In sostanza, c'è un costo fisso per la partecipazione, dopodiché il costo marginale di un atleta aggiuntivo è relativamente basso. Tuttavia, fino a quando il contingente nazionale non sarà elevato, il costo fisso, che deve essere ripartito tra tutti gli atleti, è alquanto proibitivo. Per quanto riguarda, invece, l'analisi del successo olimpico Johnson e Ali utilizzano tre metodi complementari. Il primo si basa sull'analisi del numero di medaglie vinte per nazione. A livello nazionale vincere una medaglia è un processo composto da due fasi: mandare gli atleti ai Giochi e poi incoraggiarli ad eccellere. I due ricercatori utilizzano quindi la seguente funzione lineare per determinare i fattori decisivi per la conquista delle medaglie:

$$\begin{split} medals_i &= \alpha + \beta_1 GDP + \beta_2 GDP^2 + \beta_3 POP + \beta_4 POP^2 + \beta_5 HOME + \beta_6 NEIGH \\ &+ \sum_{j=7}^{10} \beta_j POL_j + \beta_{11} LFROST + \beta_{12} HFROST + \beta_{13} t + \beta_{14} MED + u_i + \varepsilon \end{split}$$

dove medals_i è il numero di medaglie vinte dalla nazione i, MED è il numero di medaglie disponibili e le altre variabili sono state già definite in precedenza. Così come sulla partecipazione, il reddito pro capite ha un impatto positivo e significativo anche sulle medaglie vinte. Le nazioni partecipanti hanno una media di una medaglia in più ogni 1000\$ di PIL pro capite aggiuntivi nei Giochi estivi, ma hanno bisogno di più del doppio per una medaglia extra nei Giochi invernali. Le nazioni più popolose vincono più medaglie, e in particolare più medaglie d'oro ai Giochi estivi ad un tasso medio di una medaglia ogni dieci milioni di abitanti ed una medaglia d'oro ogni 30 milioni di abitanti. È interessante notare che le nazioni meno popolose hanno dei risultati leggermente migliori nei Giochi invernali, anche se il vantaggio è poco significativo. Per quanto riguarda la partecipazione, la nazione ospitante ha un grande vantaggio, infatti in media porta a casa 25 medaglie estive e 3 medaglie invernali in più del previsto. Le nazioni limitrofe a quella dove si tengono i Giochi condividono solo marginalmente il vantaggio, guadagnando 2 medaglie estive in più del previsto, ma nessun vantaggio alle Olimpiadi invernali. I due autori affermano di non avere abbastanza informazioni per poter fornire una stima dell'effetto della monarchia o di un regime totalitario/militare sulle medaglie vinte. Risulta invece ovvio come i paesi che adottano sistemi a partito singolo o Comunisti outperformano le aspettative. Questi Stati in media vincono 18 medaglie estive e 10 medaglie invernali in più rispetto a nazioni che hanno le loro stesse caratteristiche economiche e geografiche. Questi paesi non hanno un contingente di atleti maggiore, semplicemente eccellono con quelli che partecipano. Le nazioni che hanno un clima più freddo hanno performances di gran lunga superiori a quelle dei loro concorrenti con climi meno rigidi, sia nei Giochi estivi che in quelli invernali. Il secondo metodo utilizzato da Johnson e Ali consiste, così come suggerito da Bernard e Busse, nello stimare il numero di medaglie vinte da un paese come percentuale del numero di medaglie disponibili; questo poiché il numero di medaglie che una nazione può vincere è limitato dal numero di medaglie disponibili. Tuttavia, l'equazione utilizzata in

questo metodo resta identica a quella vista in precedenza, con la variabile dipendente limitata tra 0 e 1. I risultati ottenuti dai due autori sono riportati nella terza colonna della tabella che segue e non differiscono di molto rispetto a quelli ottenuti utilizzando il primo metodo. L'incremento di importanza del clima e la perdita di significatività dell'effetto della struttura comunista sono le principali differenze tra i due metodi. Il terzo ed ultimo metodo si basa sull'analisi del singolo atleta, dove la sua probabilità di successo è influenzata dal supporto che lo Stato gli offre. Come nel primo metodo Johnson e Ali utilizzano una semplice funzione lineare per valutare i fattori determinanti del successo, con le stesse variabili definite in precedenza, qui basate sulla nazionalità del partecipante. Il valore della variabile dipendente per ogni atleta è il suo livello di prestazione, ovvero 3 per una medaglia d'oro, 2 per quella d'argento, 1 per quella di bronzo e 0 per la partecipazione senza successo alla medaglia. I risultati di questo metodo sono contenuti nella quarta colonna della Tab. 2.13.

Variable	OLS by			OLS by		Probit b	v partic	ipant	Ordered probit by			
	participant nation,		I	participant nation,		1	share o	•	1	pant ath		
		medals			d medal		1 '	nedals			medals	,
	Coeff	t-sta	tistic	Coeff	t-sta	tistic	Coeff	t-sta	tistic	Coeff	t-stat	istic
Constant	-5.70	2.10	**	-3.05	2.37	**	-2.07	7.58	***			
Winter only	-2.83	2.15	**	-0.92	1.64		-4.50	4.18	***	-1.06	5.16	***
GDP per capita												
(thousands)												
Winter only	0.44	2.30	**	0.10	1.28		0.39	2.35	**	0.08	4.94	***
Summer only	1.01	3.41	***	0.52	4.02	***	0.28	3.71	***			
GDP per capita												
squared Winter only	0.01	0.46		0.01	0.71		-2.7e-3	0.32		-2.3e-3	2.86	***
Summer only	-0.01	0.40		-0.01	2.27	**	-0.01	2.67	***	2.9e-4	3.77	***
Population (millions)	-0.01	0.57		-0.01	2.21		-0.01	2.07		2.96-4	3.11	
Winter only	-4.1e-4	0.06		4.5e-3	1.61		-1.7e-3	0.35		-1.5e-3	4.86	***
Summer only	0.10	6.59	***	0.03	5.56	***	0.01	6.80	***	2.5e-3	21.2	***
Population squared	0.10	0.59		0.03	3.30		0.01	0.80		2.56-3	21.2	
Winter only	-7.9e-7	0.15		-5.9e-6	2.58	***	3.5e-6	0.99		7.6e-7	2.97	***
Summer only	-6.1e-5	5.24	***	-3.9e-6 -2.1e-5	4.31	***	-1.2e-5	4.93	***	-2.0e-6	20.5	***
Home nation	-0.16-3	3.24		-2.16-3	4.31		-1.26-3	4.93		-2.06-0	20.3	
Winter only	2.82	3.53	***	1.39	3.76	***	0.61	0.73		0.13	2.46	**
Summer only	24.9	15.0	***	11.9	15.8	***	0.01	0.73		0.13	4.54	***
Neighbor nation	24.9	15.0		11.9	15.8					0.10	4.34	
Winter only	-0.06	0.12		-0.08	0.39		0.15	0.49		0.05	1 11	
Summer only		0.13 2.52	**	I			0.15			-0.02	1.11 0.88	
•	2.37	2.32		0.39	0.91		0.51	1.11		-0.02	0.88	
Political system												
Monarchy system												
Winter only												
Summer only	-4.82	1.09		-1.61	0.92							
Single party or												
Communist system Winter only	10.3	3.49	***	4.79	4.23	***	1.35	0.72		1.20	13.0	***
Summer only	17.87	4.27	***	7.71	4.62	***	0.54	0.72		0.35	10.1	***
Military system	17.87	4.27	***	7.71	4.62	***	0.54	0.52		0.33	10.1	***
Winter only												
Summer only	-6.22	1.50		-1.91	1.18		0.85	2.54	**			
Other system	-0.22	1.50		-1.91	1.16		0.83	2.34				
Winter only	-0.46	0.14		-0.25	0.21		-0.13	0.10				
Summer only	-0.46 -3.58	0.14		-0.25 -0.95	0.52		-0.13 -0.64	1.23				
Light Winter climate	-3.36-	0.70		-0.93	0.32		-0.04	1.23				
Winter only	-1.67	1.11		-0.43	0.73		3.25	3.82	***	0.94	4.91	***
Summer only	2.14	0.88		0.12	0.73		1.99	7.12	***	0.34	11.7	***
Heavy Winter climate	2.14	0.88		0.12	0.13		1.99	7.12		0.32	11.7	
Winter only	4.26	3.42	***	1.32	2.83	***	2.76	6.10	***	0.28	5.43	***
Summer only	12.46	3.92	***	4.52	3.65	***	1.37	3.06	***	0.26	12.2	***
Time trend	12.40	3.92		4.32	3.03		1.57	3.00		0.20	12.2	
Winter only	-0.71	5.30	***	-0.22	3.63	***						
Summer only	-0.71	1.83	*	-0.22	2.11	**						
Available medals	-0.50	1.03		-0.50	2.11							
Winter only	0.05	5.33	***	0.05	3.47	***	-0.15	2.64	**			
Summer only	0.03	1.34		0.03	1.66	*	-0.13	1.04				
Adjusted R ²	0.01	0.47		0.02	0.43		-0.02	n/a			0.04	
No. of panel groups		197			197			197			0.04	
No. of observations		1397			1397			1397			88386	
											20200	

Tab. 2.13: I risultati delle regressioni

La correlazione tra la previsione di medaglie, calcolata da Johnson e Ali utilizzando il primo metodo, alle Olimpiadi invernali di Nagano 1998 e le medaglie effettivamente conquistate da ciascuna nazione è di 0,94 e di 0,88 per le medaglie d'oro. Sulla base di questi risultati molto soddisfacenti i due ricercatori hanno fatto una previsione delle medaglie vinte per paese ai Giochi di Salt Lake city del 2002. In media, una medaglia aggiuntiva è associata a un aumento del PIL pro capite di circa 1750\$. Naturalmente, il costo è inferiore per le nazioni ad alto reddito, raggiungendo un minimo di 1645\$ pro capite per gli Stati Uniti. Per le medaglie d'oro, i valori sono ovviamente molto più alti, con una media di quasi 4750\$ di aumento di PIL pro capite per la maggior parte delle nazioni. Addirittura, nel caso della Corea, un'ulteriore medaglia d'oro è associata ad un sorprendente aumento del 67% del PIL pro capite. Johnson e Ali concludono quindi il loro lavoro ponendo l'accento sulle differenze che si sono delineate sia tra le Olimpiadi estive e quelle invernali che tra le diverse discipline all'interno della stessa manifestazione. Risulta una conseguenza naturale che le nazioni più popolose abbiano un vantaggio in quegli eventi che si basano sulla resistenza fisica o sulla forza, principalmente perché avranno una popolazione più vasta dalla quale selezionare gli atleti. Allo stesso modo, potrebbe esserci un vantaggio per le nazioni ad alto reddito nelle gare di sport che dipendono da attrezzature o strutture costose. Per verificare queste differenze i due ricercatori hanno suddiviso gli sport in 6 categorie, divise in tre sottocategorie in parallelo per Olimpiadi estive e invernali: sport ad alta intensità di investimento, discipline ad alta intensità di lavoro e gli sport di squadra. Gli sport ad alta intensità di lavoro nei Giochi estivi includono discipline come la boxe, la lotta, il judo, la maratona e l'atletica; nei Giochi invernali comprendono tutte le gare di sci di fondo e alpino. Gli sport ad alto investimento di capitale includono per le Olimpiadi estive sport come l'equitazione, la vela ed il pentathlon, mentre per quelle invernali le gare di slittino, bob, salto con gli sci e tutte quelle discipline che richiedono costose infrastrutture appositamente dedicate per la pratica e l'allenamento. Per quanto riguarda gli sport di squadra nei giochi estivi comprendono tutti gli sport come basket, pallavolo, calcio, pallanuoto, e altri, in quelli invernali invece sono inclusi l'hockey, il curling e la staffetta di pattinaggio di velocità. Sebbene i risultati mostrino alcune differenze statisticamente significative, la dimensione effettiva delle differenze rimane comunque piccola. Ad esempio, nei Giochi estivi il PIL pro capite ha pesato maggiormente negli sport ad alta intensità di capitale, rimanendo del tutto insignificante negli sport di squadra. La dimensione della popolazione è stata significativa e positiva in tutte le gare delle Olimpiadi estive, ma con un impatto relativo molto maggiore negli sport ad alta intensità di lavoro ed in quelli di squadra. Le differenze tra le tre categorie invernali sono state minime, forse a causa del numero più limitato di eventi e quindi di medaglie in palio.

Sempre nel 2002 i tre ricercatori Robert Hoffmann, Lee Chew Ging e Bala Ramasamy pubblicano uno studio dal titolo "Public Policy and Olympic Success", nel quale cercano di trovare quali altri fattori oltre a quelli economici e politici influenzino la performance di un paese ai Giochi olimpici. Questo studio si concentra sull'analisi dei risultati delle regressioni effettuate testando una serie di potenziali fattori determinanti delle prestazioni olimpiche nazionali. La motivazione degli autori era quella di valutare la portata di politiche sportive nazionali basate sui fattori testati. La prima ipotesi che i tre ricercatori redigono è di natura demografica; la popolazione di una nazione costituisce il bacino da cui viene sviluppato ed estratto il talento. Ceteris paribus, si prevede che i paesi con popolazioni più grandi abbiano una maggiore probabilità di contenere dei potenziali campioni sportivi di livello internazionale. La seconda ipotesi è invece di tipo geografico, ovvero essi credono che le condizioni climatiche estreme come le temperature esterne diurne eccessivamente alte o basse, l'umidità eccessiva o le precipitazioni dovrebbero avere un impatto negativo sull'attività sportiva all'aperto e quindi sulla capacità di conquistare medaglie alle Olimpiadi. Gli autori decidono di stimare questo effetto attraverso l'utilizzo di una variabile che rappresenta la temperatura media

misurata in gradi Celsius della capitale del paese preso in considerazione e di quattro variabili dummy che ne indicano l'appartenenza ad uno dei quattro tipi di clima identificati dal Köppen-Geiger-Trewartha (KGT) world climate system. Per quanto riguarda l'aspetto economico è stato utilizzato il reddito pro capite per verificare la tesi che lo sviluppo delle abilità degli atleti di un paese è fortemente dipendente dalle infrastrutture sportive in termini di strutture, attrezzature, coaching e talent scouting a livello scolastico, universitario e pubblico. Hoffmann, Ging e Ramasamy ipotizzano che anche la struttura e l'orientamento politico di un paese influenzano le vittorie olimpiche, si sostiene che i paesi con un governo forte e centralizzato abbiano una maggiore capacità e propensione ad organizzare le risorse della nazione per raggiungere questo obbiettivo. Infine, l'ultima ipotesi sviluppata afferma che la cultura sportiva e la centralità dello sport nella cultura di un paese sono fortemente determinanti del successo ai Giochi olimpici. Il proxy utilizzato per verificare questo effetto è una variabile che contiene il numero totale di volte in cui un paese ha ospitato i Giochi Olimpici Estivi tra il 1948 e il 1996. Le nazioni ospitanti sono comunemente selezionate sulla base della loro capacità di organizzare con successo eventi e competizioni sportivi con i più alti standard. La Tab. 2.14 raccoglie le variabili utilizzate da Hoffmann, Ging e Ramasamy nella loro ricerca:

Variable	Description
$\overline{Y_i}$	Total Number of Sydney 2000 medals won by country i
GNP_i	GNP per capita of country i
POP_{i}	Population size of country i
$\overrightarrow{TEMP_i}$	Average annual Celsius temperature in country i's capital
D_{COM}	Dummy for current/previous socialist government (=1, 0 otherwise)
D_{AUS}	Dummy for Sydney 2000 host (=1, 0 otherwise)
D_{Host1}	Dummy for countries having hosted once (=1, 0 otherwise)
D_{Host2}	Dummy for countries having hosted twice $(=1, 0 \text{ otherwise})$
D_{CH}	Dummy for <i>Cooler Humid</i> zone (=1, 0 otherwise)
D_{WH}	Dummy for Warmer Humid zone (=1, 0 otherwise)
D_{DRY}	Dummy for Dry zone (=1, 0 otherwise)
D_{TH}	Dummy for <i>Tropical Humid</i> zone (=1, 0 otherwise)
ϵ_i	Error term

Tab. 2.14: Le variabili utilizzate da Hoffmann, Ging e Ramasamy

L'analisi di regressione è stata fatta basandosi sul medagliere delle Olimpiadi estive di Sidney 2000. Per verificare l'effetto delle politiche nazionali in modo più preciso ed accurato, i tre autori hanno condotto due test separati. Nel primo sono state eseguite le regressioni necessarie a valutare l'influenza di tutte quelle variabili che non dipendono dalle scelte politiche ed economiche del governo di un paese come la temperatura ed il clima. Sulla base dei risultati ottenuti, e riportati nella tabella sottostante, è stata stimata la seguente equazione:

$$Y_i = \beta_{10} TEMP_i + \beta_{11} TEMP_i^2 + \epsilon_i$$

I risultati mostrano quindi una relazione quadratica tra il numero di medaglie vinte nelle Olimpiadi di Sidney 2000 e la temperatura. La temperatura ottimale per il successo olimpico risulta essere di circa 15° Celsius. È stata fatta anche una regressione basata sui quattro tipi di clima definiti dal sistema KGT, da questi risultati i tre ricercatori hanno dedotto che i paesi che hanno un clima Warmer e Colder Humid vincono rispettivamente in media 14 e 20 medaglie in più. Partendo dai risultati riportati in Tab. 2.15 è stata interpolata la seguente funzione lineare:

$$Y_i = \beta_{20}D_{DRY} + \beta_{21}D_{TH} + \beta_{22}D_{WH} + \beta_{23}D_{CH} + \epsilon_i$$

Variable	Coefficient	t-value	<i>p</i> -value	Variable	Coefficient	Beta	t-value	<i>p</i> -value
$TEMP_i$	2.196	5.1439	0.0000***	D_{DRY}	2.6250	0.039	0.4145	0.6797
$TEMP_i^2$	-0.0809	-4.0956	0.0001***	D_{TH}	3.3333	0.068	0.7208	0.4734
R^2	0.317			D_{WH}	14.0000	0.471	5.0051	0.0000***
	0.517			D_{CH}	20.1667	0.367	3.9004	0.0002***

Tab. 2.15: I risultati della regressione sulle variabili climatiche

Nel secondo test invece sono stati analizzati i fattori di natura politica utilizzando la seguente funzione:

$$Y_i = \beta_{30} + \beta_{31}GNP_i^{1/2} + \beta_{32}D_{AUS} + \beta_{33}D_{Host1} + \beta_{34}D_{Host2}$$

 $+ \beta_{35}D_{COM} + \beta_{36}POP^{1/2} + \epsilon_i$

L'analisi dei risultati condotta da Hoffmann, Ging e Ramasamy mostra che i fattori considerati sono significativi e che il reddito pro capite e la popolazione hanno un effetto positivo marginalmente decrescente, come si osserva dalla Tab. 2.16. L'effetto positivo del PIL pro capite è comunque inferiore a quello della popolazione.

Variable	Coefficient	Beta	t-value	<i>p</i> -value
Constant	-9.9297	_	-2.7566	0.0075***
$GNP_i^{1/2}$	0.0866	0.247	2.8947	0.0051***
D_{AUS}	34.4701	0.213	3.0995	0.0028***
D_{Host1}	16.0267	0.306	3.993	0.0002***
D_{Host2}	72.1684	0.446	6.2466	0.0000***
D_{COM}	12.9284	0.332	4.4404	0.0000***
$D_{COM} POP_i^{1/2}$	1.1877	0.365	5.0156	0.0000***

Tab. 2.16: I risultati della regressione sulle variabili non climatiche

Le conclusioni dedotte da Hoffmann, Ging e Ramasamy differiscono, almeno parzialmente, da quelle degli studi condotti precedentemente da Condon (1999), Bernard and Busse (2000) e Johnson e Ali (2000). In primo luogo, la ricerca considera e quantifica l'effetto dei fattori climatici, ed inoltre identifica che la popolazione ed il reddito hanno ritorni marginali decrescenti. Quest'analisi rivela che esistono reali possibilità per le politiche pubbliche di migliorare le prestazioni sportive di un paese. Infatti, sebbene i fattori climatici siano significativi, il modello di Hoffmann, Ging e Ramasamy indica che i fattori politici sono ancora più influenti sulla possibilità di vincere le medaglie. Alcuni di queste variabili come il PIL pro capite o la popolazione sono modificabili solamente sul lungo periodo. A breve termine invece, per ottenere degli effetti positivi sulle vittorie olimpiche, un governo dovrebbe investire sulle infrastrutture e sull'organizzazione di importanti eventi sportivi internazionali, e seguire l'approccio alle politiche sportive tipico delle economie a pianificazione centralizzata.

Nel 2002 venne pubblicato sul Journal of the Operational Research Society una ricerca che così come quella di Bernard e Busse divenne un punto di riferimento per la misura e la previsione delle

performance olimpiche di un paese. Questo studio fu redatto congiuntamente da quattro autori, Lozano, Gabriel Villa, Flor Maria Guerrero e Pablo Cortés, presso l'Università di Siviglia. In questo studio viene utilizzato uno strumento matematico innovativo, ben noto per la valutazione dell'efficienza, ovvero il Data Envelopment Analysis (DEA), per misurare le prestazioni delle nazioni partecipanti agli ultimi cinque giochi olimpici estivi. L'approccio proposto considera due input, ovvero il PIL e la popolazione, e tre output, cioè il numero di medaglie d'oro, d'argento e di bronzo vinte. Per ogni edizione dei Giochi olimpici vengono calcolati per ogni nazione un indice di performance ed un paese benchmark. Inoltre, tenere traccia delle prestazioni di un paese specifico per le diverse edizioni delle Olimpiadi può aiutare a identificare, in modo oggettivo, sia i trend che i successi e le delusioni. La ricerca di Lozano, Villa, Guerrero e Cortés più che sviluppare delle previsioni della performance olimpica dei paesi partecipanti, ha come obbiettivo quello di utilizzare uno strumento di analisi di frontiera stocastica, il Data Envelopment Analysis, per identificare quali paesi performano in modo efficiente e quali in modo inefficiente, e per stabilire una misura oggettiva del successo di un paese alle Olimpiadi. Il Data Envelopment Analysis è uno strumento matematico utilizzato in ricerca operativa per la stima delle frontiere della funzione di produzione. In ciascuna analisi DEA, le unità produttive considerate corrispondevano alle nazioni partecipanti che hanno vinto almeno una medaglia. L'analisi si basa, come detto, su tre variabili in output, rispettivamente il numero di medaglie d'oro, il numero di medaglie d'argento e quello di medaglie di bronzo vinte alle Olimpiadi, e due sole variabili in input, ovvero il PIL e la popolazione. Questo presupposto implica che il metodo di analisi messo a punto dai quattro ricercatori spagnoli tenda a classificare come efficienti paesi piccoli e/o poveri, ad esempio Cuba, che vincono molte medaglie ed a classificare come inefficienti paesi grandi, come l'India, o ricchi, come l'Arabia Saudita, che non vincono abbastanza medaglie. Secondo gli autori vi sono altre variabili di input che potrebbero essere aggiunte, come ad esempio l'altezza media della popolazione, poiché un paese con uomini e donne alti dovrebbe essere avvantaggiato in alcuni sport. I risultati ottenuti dai quattro ricercatori iberici mostrano come degli 80 paesi partecipanti ai Giochi che hanno vinto almeno una medaglia, solo 9 siano considerati efficienti, ovvero Australia, Bahamas, Barbuda, Cuba, Estonia, Germania, Moldavia, Russia e Stati Uniti. Di questi, il paese che appare più frequentemente come uno Stato benchmark è Cuba, 58 volte, seguita dalla Russia, Bahamas e Germania e Australia. Non sorprende che, sebbene siano efficienti, gli Stati Uniti siano raramente un benchmark. Ciò è dovuto al fatto che gli Stati Uniti hanno un mix di input piuttosto atipico: un paese popolato ma ricco. Cuba rappresenta invece un caso diametralmente opposto, un piccolo paese sia per popolazione che per prodotto nazionale. Il numero di volte in cui uno Stato risulta essere un paese benchmark è una misura di quanto la performance di un paese efficiente può essere un riferimento utile per quelli non efficienti. Il punto di forza del Data Envelopment Analysis è la sua capacità di misurare la performance relativa, separando chiaramente le nazioni efficienti da quelle che non lo sono; lo fa basandosi solamente sugli input e sugli output osservati. Non è infatti in grado di identificare il motivo per il quale un paese è o non è efficiente. Tuttavia, questo metodo di analisi risulta essere molto utile per identificare correttamente i paesi efficienti ed informare quelli non efficienti di quali sono i paesi a cui dovrebbero fare riferimento, ed imitare, per migliorare le proprie prestazioni. Lozano, Villa, Guerrero e Cortés mostrano anche come si è evoluta l'efficienza per ogni paese durante le ultime cinque edizioni dei Giochi olimpici estivi. Si noti che le nazioni che sono più stabilmente e continuativamente efficienti sono Cuba, Germania, Ungheria, Russia e USA. Anche altri paesi, come la Bulgaria o la Giamaica, mostrano performance regolari con un alto livello di efficienza. Altri paesi, come l'Irlanda o il Suriname, mostrano un andamento irregolare, con un'efficienza molto alta in alcune Olimpiadi ed un'efficienza molto ridotta in altre. Infine, alcuni paesi, come l'Argentina o il Messico, sembrano essere stabilmente bloccati a bassi livelli di efficienza. L'utilità di questi risultati consiste nell'avere un metodo oggettivo di valutazione e monitoraggio della performance di un paese, ricordando però che la misura di efficienza fornita dal Data Envelopment Analysis è una misura relativa, perciò anche se un paese ha migliorato la sua performance assoluta, quindi il numero di medaglie vinte, è possibile che abbia avuto una performance relativa peggiore poiché altri paesi potrebbero essere migliorati ancora di più. Quando alcuni paesi piccoli e/o poveri, come Gibuti, Suriname o le Isole Vergini, fanno segnare delle prestazioni ottime, cioè con un'efficienza del 100%, i paesi più ricchi e/o più popolati presentano, di conseguenza, un evidente peggioramento delle proprie performance relative. Attraverso i risultati che emergono dall'analisi DEA Lozano, Villa, Guerrero e Cortés sono in grado di verificare se il paese ospitante abbia avuto o meno una performance superiore alla sua media storica.

Dopo qualche anno di immobilismo gli studi riguardanti i Giochi olimpici riprendono con forza nel 2005, anno in cui il ricercatore Bian pubblica il suo lavoro "Predicting Olympic Medal Counts: the Effects of Economic Development on Olympic Performance". Xun Bian inizia la sua ricerca chiedendosi perché alcuni paesi sono in grado di vincere un elevato numero di medaglie alle Olimpiadi, mentre altri non riescono mai a farlo. Questo studio ha l'obbiettivo di fare una previsione delle prestazioni olimpiche di una nazione indagando sulle variabili socioeconomiche che le influenzano significativamente. L'influenza della dimensione della popolazione, delle risorse economiche, della struttura politica ed economica e del vantaggio di essere la nazione ospitante viene stimata utilizzando due diversi modelli: una funzione lineare e una funzione di produzione Cobb-Douglas. Il primo metodo utilizza la seguente equazione:

$$M_t = C + \alpha_1 N_t + \alpha_2 (Y_t / N_t) + \alpha_3 P + \alpha_4 H_t + \epsilon$$

dove Mt indica il numero di medaglie di un paese in una particolare edizione delle Olimpiadi; Bian non prevede nessuna differenza tra medaglie d'oro, d'argento e di bronzo, poiché secondo lui la differenza tra i primi tre classificati di solito è così esigua che il colore della medaglia dipende più dalla fortuna che dal talento. N_t è la dimensione della popolazione del paese nell'anno t in cui si disputano i Giochi olimpici e Y_t/N_t è quindi il PIL pro capite. Infine, P e H_t sono variabili dummy che rappresentano rispettivamente la struttura politica ed economica e il fatto che un paese sia la nazione ospitante. La variabile P assume il valore 1 se il paese ha un background socialista, il che significa che il paese è o è stato un paese socialista, e assume il valore 0 in caso contrario. In modo del tutto analogo, se il paese ospita quella determinata edizione delle Olimpiadi, Ht assume il valore 1, in caso contrario assume il valore 0. I coefficienti di ciascuna variabile indipendente, che devono necessariamente essere positivi, rappresentano l'effetto marginale di quella particolare variabile sul totale di medaglie Mt. Il secondo modello utilizza come detto una funzione Cobb-Douglas, esso considera infatti la vittoria delle medaglie simile a un processo di produzione, dove i due fattori chiave sono la dimensione della popolazione N e le risorse economiche Y, che sono entrambi soggetti a un rendimento marginale decrescente. La variabile At raccoglie al suo interno l'effetto degli altri due fattori visti in precedenza.

$$M_t = A_t (N_t) \gamma (Y_t) \theta$$

Semplificandola e trasformandola opportunamente, attraverso l''uso della notazione logaritmica, Bian ottiene la seguente equazione:

$$\ln M_t = C + \alpha_1 \ln N_t + \alpha_2 \ln Y_t + \alpha_3 P + \alpha_4 H_t + e$$

I risultati dell'analisi di regressione condotta da Bian utilizzando il primo modello mostrano che la dimensione della popolazione e il PIL pro capite sono costantemente significativi nel tempo, sebbene l'entità di tali effetti differisca di anno in anno. Anche le due variabili dummy sono statisticamente significative per la maggior parte delle Olimpiadi ed il loro effetto varia drasticamente nel tempo. Il fatto di essere la nazione organizzatrice dei Giochi Olimpici del 2000 ha permesso all'Australia di vincere addirittura 41 medaglie in più. Bian utilizza invece i risultati del secondo modello per verificare che sia la popolazione che il PIL abbiano un rendimento marginale decrescente. La variabile lnYt e le due variabili dummy sono statisticamente significative, e i loro coefficienti confermano le ipotesi iniziali fatte dal ricercatore di origine asiatica. Per contro, la maggior parte dei coefficienti della variabile InNt non sono statisticamente significativi; questi risultati sono in contraddizione sia con l'ipotesi di ricerca di Bian che con i risultati ottenuti utilizzando il primo modello. L'autore si aspettava infatti che la dimensione della popolazione e le risorse economiche siano correlate positivamente con il numero di medaglie di un paese, e che l'essere un paese socialista o il paese ospitante aumenti il numero di vittorie. Tuttavia, data la forte correlazione esistente tra lnNt e lnYt, il PIL è infatti solo il prodotto del PIL pro capite per la dimensione della popolazione, è ragionevole supporre che i risultati ottenuti col secondo metodo di Bian potrebbero essere distorti dalla multicollinearità. Per risolvere questo problema l'autore utilizza quindi una funzione di Cobb-Douglas diversa:

$$M_t = A_t (N_t) 1 - \theta(Y_t) \theta$$

In linea con gli studi precedenti, anche questa ricerca rileva che le variabili socioeconomiche, tra cui

la dimensione della popolazione, le risorse economiche, il vantaggio dell'essere la nazione ospitante e la struttura politica, hanno un impatto significativo sulle prestazioni di un paese alle Olimpiadi. La dimensione della popolazione e le risorse economiche influenzano positivamente il numero di medaglie vinte. Essere la nazione organizzatrice dei Giochi ed avere in contemporanea, un background comunista aumentano il numero di medaglie conquistate. Sebbene lo studio di Bian fornisca alcune interessanti informazioni sulla correlazione tra lo sviluppo economico di un paese ed i suoi risultati ai Giochi olimpici, uno dei suoi principali difetti è quello di non riuscire a fornire delle previsioni attendibili del numero di medaglie vinte da un paese nelle future Olimpiadi. Nel dicembre del 2006 il ricercatore americano Wade Donald Pfau con il supporto del National Graduate Institute for Policy Studies del Giappone pubblica uno studio dal titolo "Predicting the Medal Wins by Country at the 2006 Winter Olympic Games: An Econometrics Approach". L'obbiettivo di tale ricerca è quello di fornire una serie di previsioni riguardo al numero di medaglie d'oro e di medaglie totali che ogni paese avrebbe dovuto vincere alle Olimpiadi invernali di Torino 2006. Lo studio si basa sulle caratteristiche economiche e demografiche, nonché sulle prestazioni passate, dei paesi partecipanti per determinare il numero di medaglie che questi ultimi vinceranno. Pfau, con questa ricerca, vuole dimostrare che un semplice modello econometrico può produrre previsioni altrettanto valide, se non migliori, rispetto ad un esperto di sport che genera delle previsioni evento per evento, semplicemente perché consente di trovare un metodo più congruo per ridurre l'incertezza. Per la redazione di questa pubblicazione scientifica Wade Donald Pfau prende ispirazione direttamente dal lavoro pubblicato dagli studiosi Bernard e Busse nel 2002, da quello di Johnson e Ali del 2004 e da quello presentato da Kuper e Sterken nel 2001. Come gli autori precedentemente citati, anche Pfau utilizza un modello di regressione per stimare il numero di medaglie che ci si può aspettare che un paese vinca in funzione di alcune caratteristiche. Il suo modello si basa sulla metodologia di Bernard e Busse, e considera la popolazione e il PIL di un paese come fattori principali per conquistare delle medaglie alle Olimpiadi. La variabile dipendente di questo studio è quindi la percentuale di medaglie MS vinte da un paese nell'Olimpiade considerata. Le variabili indipendenti utilizzate sono invece i logaritmi della popolazione POP e del PIL pro capite GDPCAP. Ci si può aspettare che un paese ospitante, variabile HOST, abbia prestazioni migliori a causa del maggiore sostegno del pubblico, della familiarità con gli impianti sportivi e delle condizioni climatiche. Inoltre, il successo alle Olimpiadi precedenti, indicato dalla variabile MSt-1, può aiutare a spiegare i successi attuali indicando che il paese ha una forte cultura sportiva. Sono state aggiunte alcune variabili dummy che rappresentano l'appartenenza geografica ad un determinato territorio o a alla sfera di influenza sovietica. La variabile SOV racchiude al suo interno Bulgaria, Germania orientale, Ungheria, Polonia, Romania, Cecoslovacchia e fino al 1988 l'Unione Sovietica. La variabile SCAND include Norvegia, Svezia e Finlandia. La variabile GERM comprende Germania, Austria, Germania ovest e Germania est. La variabile ALPINE include Svizzera, Italia e Francia, e la variabile NORTHAM annovera Stati Uniti e Canada. La funzione utilizzata dall'autore per la regressione è la seguente

$$MS_{t} = \beta_{0} + \beta_{1} \log(POP_{t}) + \beta_{2} \log(GDPCAP_{t}) + \beta_{3}HOST_{t} + \beta_{4}MS_{t-1} + \beta_{5}SOV_{t} + \beta_{6}SCAND_{t} + \beta_{7}GERM_{t} + \beta_{8}ALPINE_{t} + \beta_{9}NORTHAM_{t} + \varepsilon_{t}$$

Questa equazione è del tutto similare a quella di Bernard e Busse, solamente adattata ai Giochi invernali. Rispetto agli altri articoli che abbiamo citato, né Kuper e Sterken né Johnson e Ali includono una variabile che tiene conto del numero di medaglie vinte da un paese ai precedenti Giochi olimpici. Al contrario Johnson e Ali utilizzano il numero assoluto di medaglie anziché la percentuale di medaglie vinte, e includono una variabile per tenere conto che il numero di medaglie disponibili aumenta col passare del tempo poiché, come già detto, aumentano le discipline olimpiche. I risultati della regressione mostrano che, come previsto da Pfau, tutte le variabili hanno coefficienti positivi, sebbene la popolazione ed il reddito pro capite non siano del tutto significativi. Per quanto riguarda la popolazione, a differenza delle Olimpiadi estive, la ragione per la quale è una variabile meno importante è dovuta al fatto che i paesi che posseggono le maggiori competenze negli sport invernali sono quelli che si trovano nelle aree climatiche più fredde, in particolare quelli scandinavi, i quali tendono ad avere popolazioni meno numerose. Il motivo per cui il reddito pro capite non è significativo è spiegato dalla presenza nel modello delle variabili di appartenenza politico-geografica e da quella che riporta il numero di medaglie vinte ai precedenti Giochi. Pfau afferma che, senza queste variabili, il reddito pro capite è molto significativo. L'analisi dei risultati mostra inoltre che essere la nazione nella quale si svolgono i Giochi aumenta significativamente il numero di medaglie vinte. Anche le variabili dummy SOV, SCAND e GERM sono molto significative, i paesi che vi appartengono hanno infatti vinto più medaglie di quanto sarebbe altrimenti stato giustificato dalle loro caratteristiche economiche e demografiche. Pfau mette a confronto le sue previsioni di medaglie e medaglie d'oro per paese per le Olimpiadi invernali di Salt Lake City con quelle di Sterken e Kuper e quelle di Johnson e Ali. In termini di performance, le previsioni di Pfau del numero totale di medaglie per paese risultano leggermente migliori delle altre in termini di coefficienti di correlazione. L'approccio di questo studio è stato notevolmente più efficace di quello degli altri nel produrre previsioni accurate per le medaglie d'oro. Le previsioni di Johnson e Ali sono state le meno precise, anche se gli autori indicano che il loro obiettivo era quello di spiegare quali fattori influenzano il numero di medaglie vinte piuttosto che produrre delle previsioni il più accurate possibili. Pfau conclude la sua ricerca con la previsione del numero di medaglie totali e del numero di medaglie d'oro vinte per ogni paese ai Giochi olimpici invernali di Torino 2006, e con il confronto di queste con i risultati effettivi, come si può osservare in Tab. 2.17.

		2006 Total Medals	
Country Name	Total Medals	Prediction	Difference
Germany	29	35	-6
United States	25	31	-6
Canada	24	17	7
Austria	23	19	4
Russian Federation	22	11	11
Norway	19	24	-5
Sweden	14	10	4
Switzerland	14	11	3
China	11	7	4
Italy	11	20	-9
South Korea	11	4	7
Finland	9	10	-1
France	9	12	-3
Netherlands	9	7	2
Czech Republic	4	3	1
Croatia	3	3	0
Estonia	3	2	1
Australia	2	2	0
Poland	2	2	0
Belarus	1	1	0
Bulgaria	1	2	-1
Great Britain	1	3	-2
Japan	1	3	-2

Tab. 2.17: La previsione del medagliere per Torino 2006

Il coefficiente di correlazione complessivo per le medaglie totali è di 0,934. Come previsto, Germania e Stati Uniti sono finiti in cima alla classifica, sebbene entrambi abbiano vinto 6 medaglie in meno del previsto. Poi ci sono state delle sorprese: il Canada si è classificato terzo con 7 medaglie in più del previsto, seguito dall'Austria con 4 medaglie più di quanto era stato ipotizzato. Quinta si è piazzata la Russia, che ha portato a casa 22 medaglie, 11 medaglie in più del previsto. D'altro canto, alcuni paesi non si sono comportati come Pfau aveva previsto. In particolare, il vantaggio di essere la nazione organizzatrice non si è concretizzato per l'Italia, che ha chiuso con sole 9 medaglie. Anche la Norvegia ha avuto una prestazione deludente, Pfau si aspettava che il paese arrivasse terzo con 24 medaglie, ma alla fine la Norvegia ha ottenuto solo 19 medaglie. Infine, la previsione che questa Olimpiade fosse la più inclusiva di tutte non è si è verificata, poiché degli 11 paesi che l'autore si aspettava conquistassero almeno una medaglia, solo la Bielorussia ha raggiunto questo obiettivo.

Cartan	2006 Actual	2006 Gold Medals	D:00
Country Name	Gold Medals	Prediction	Difference
Germany	11	11	0
United States	9	9	0
Austria	9	4	5
Russian Federation	8	4	4
Canada	7	7	0
Sweden	7	2	5
South Korea	6	2	4
Italy	5	6	-1
Switzerland	5	3	2
Estonia	3	1	2
France	3	4	-1
Netherlands	3	3	0
China	2	2	0
Norway	2	11	-9
Australia	1	2	-1
Croatia	1	2	-1
Czech Republic	1	1	0
Japan	1	1	0
Finland	0	5	-5
Great Britain	0	1	-1

Tab. 2.18: La previsione delle medaglie d'oro per Torino 2006

Per quanto riguarda le medaglie d'oro, come si osserva in Tab. 2.18, il coefficiente di correlazione è 0,773; la ricerca di Pfau ha previsto con precisione 11 ori per la Germania, 9 ori per gli Stati Uniti e 7 ori per il Canada. Tuttavia, anche in cima alla classifica ci sono alcune sorprese. L'Austria ha vinto 9 medaglie d'oro e la Svezia ne ha vinte 7, 5 in più del previsto in entrambi i casi. Anche Russia e Corea del Sud hanno vinto 4 medaglie d'oro in più delle attese. Infine, mentre la Svezia è stata, tra i paesi scandinavi, quella che ha fatto meglio rispetto alle previsioni di Pfau, sia la Norvegia che la Finlandia hanno entrambe avuto una performance significativamente al di sotto delle aspettative. La Norvegia ha chiuso con 2 medaglie d'oro, 9 in meno del previsto, e la Finlandia con 0 medaglie d'oro, 5 in meno delle attese. Pfau propone infine di migliorare il suo modello provando ad includere delle variabili che tengano conto dei risultati degli eventi sportivi internazionali che hanno preceduto le Olimpiadi e del numero di atleti che ogni nazione manda ai Giochi olimpici. Questo potrebbe potenzialmente migliorare le previsioni aggiungendo informazioni su come la situazione è cambiata rispetto alle precedenti Olimpiadi, riducendo la significatività della variabile MSt-1.

Nel 2008 l'Università di Zurigo pubblica una ricerca dal titolo "Economics and the Summer Olympics: an efficiency analysis", gli autori di questo lavoro sono gli studiosi Ulrich Woitek e Alexander Rathke. In questo lavoro i due autori iniziano a pensare al successo olimpico come una funzione di produzione con il PIL e la dimensione della popolazione come input. Basandosi poi sulla letteratura degli anni precedenti, includono anche altri fattori tra quelli determinanti per la produttività di medaglie. È risaputo che i paesi socialisti e quelli che ospitano i Giochi superino sistematicamente, in termini di prestazioni, gli altri. Woitek e Rathke identificano un altro fattore, secondo loro essenziale: l'importanza data allo sport all'interno della società. Quest'ultimo influenza, infatti, la disponibilità degli enti finanziatori come il governo a fornire il denaro necessario per la formazione degli atleti e per le strutture di allenamento, come questi finanziamenti vengono utilizzati e le scelte di carriera dei potenziali atleti. L'importanza dello sport nella società è però un concetto difficile da misurare. Tuttavia, l'approccio della funzione di produzione ci consente di trattarlo in modo simile all'efficienza tecnica. Per illustrare questo punto i due ricercatori dell'istituto elvetico propongono il seguente esempio: si considerino due paesi con una diversa quota medaglie Y ai Giochi, con il paese 1 che ha performance migliori di quelle del paese 2. Questa differenza di prestazioni può essere dovuta a risorse differenti o a diverse tecnologie di allenamento, ma potrebbe anche riflettere il fatto che c'è meno sostegno pubblico per lo sport nel paese 2. La conseguenza sarebbe che il budget effettivo disponibile per la promozione dell'attività sportiva risulta inferiore ed il pool di atleti che possono essere inviati ai Giochi è più ristretto. Potrebbe anche accadere che un paese decida di concentrare le risorse su alcune discipline popolari e nelle quali si ha una tradizione vincente invece di finanziare lo sport in generale, facendo così un uso più efficiente delle scarse risorse disponibili. La frontiera di produzione rappresenta la quota massima di medaglie per un dato livello di risorse e una certa tecnologia. Nella figura a lato, questa quota massima di medaglie è rappresentata dal valore Y* in ordinate, poiché le dotazioni di risorse nei due paesi sono le stesse. Entrambi i paesi sono al di sotto di questa quota di medaglie: la distanza dalla frontiera di produzione misura la diversa importanza relativa attribuita allo sport in generale o ad alcune discipline specifiche dai due paesi. Utilizzando l'analisi di frontiera stocastica, è possibile distinguere empiricamente tra le determinanti delle differenze di prestazioni tra paesi alle Olimpiadi. L'idea di base della misura di efficienza è che i paesi potrebbero anche produrre al di sotto della frontiera, in tal caso si avrebbe uno spreco di risorse come rappresentato da Y1 e Y2 in Fig. 2.5. La frontiera di produzione definisce il massimo output ottenibile per un dato vettore di input. Per questa ricerca Woitek e Rathke utilizzano una funzione di produzione di tipo translog, che risulta essere più flessibile di una Cobb Douglas. L'output della funzione è il logaritmo della quota medaglie MSit, il PIL (PILit) e la dimensione della popolazione (POPit) sono utilizzati come input. Il modello include anche variabili dummy per indicare l'effetto di essere la nazione ospitante HOST_{it}, di essere una nazione ad economia pianificata centralmente PLAN_{it}, e di essere sotto la sfera di influenza Sovietica SOV_{it}. Per valutare se vi sono effetti a lungo termine dell'ospitare i Giochi olimpici, Woitek e Rathke includono anche due variabili dummy per indicare se la nazione ha organizzato le Olimpiadi nell'edizione precedente, PREHOSTit, o se lo farà nell'edizione futura, POSTHOST_{it}. La funzione di produzione della frontiera stocastica è:

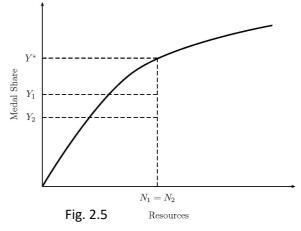
$$\ln(MS_{it}) = \beta_{0} + \beta_{1} \ln(GDP_{it}) + \beta_{2} \ln(POP_{it}) + \beta_{3} \ln(GDP_{it})^{2} + \beta_{4} \ln(POP_{it})^{2} + \beta_{5} \ln(GDP_{it}) \ln(POP_{it}) + \beta_{6}HOST_{it} + \beta_{7}SOV_{it} + \beta_{8}PLAN_{it} + \beta_{7}PREHOST_{it} + \beta_{8}POSTHOST_{it} + v_{it} - u_{it},$$

L'inefficienza tecnica è modellata come:

$$\mu_{it} = \delta_0 + \delta_1 HOST_{it} + \delta_2 SOV_{it} + \delta_3 PLAN_{it} + \delta_4 PREHOST_{it} + \delta_5 POSTHOST_{it}$$

I risultati dell'analisi mostrano che un paese sovietico che si trova sulla frontiera di efficienza può ottenere una quota di medaglie che è del 144% superiore rispetto a quella di un'economia di mercato comparabile. I due autori calcolano poi l'elasticità per il PIL e per la popolazione; l'elasticità del PIL ha segno positivo, come previsto: un aumento del PIL dell'1 % porta ad un aumento della quota medaglia dello 0,72%. L'elasticità stimata per la popolazione ha invece segno negativo: solo i paesi molto ricchi con popolazioni relativamente piccole possono trarre vantaggio da un aumento

della popolazione, il che è ovviamente ragionevole, per i paesi poveri, infatti, un aumento della popolazione riduce le risorse disponibili per raggiungere il successo alle Olimpiadi. Il segno negativo delle variabili dummy sta ad indicare che i paesi che ospitano i Giochi olimpici e quelli che non hanno una economia di mercato producono una quota di medaglie più vicina a quella di frontiera: le possibili spiegazioni sono o che lo sport ha una maggiore considerazione all'interno della società rispetto agli altri paesi, o che usano in modo più efficiente le proprie risorse. Quindi, riassumendo,



Woitek e Rathke stimano la distanza dalla frontiera utilizzando una misura dell'efficienza tecnica; quando si analizzano le determinanti del successo ai Giochi Olimpici utilizzando l'approccio della frontiera stocastica, l'efficienza tecnica cattura concetti come l'importanza dello sport nella società e gli effetti della specializzazione in determinate discipline.

Nel 2008 Wladimir Andreff decise di tornare a studiare i fattori che determinano le vittorie ai Giochi olimpici, tuttavia questa volta decide di farlo con la sorella Madeleine Andreff e la ricercatrice Sandrine Poupaux. L'articolo dal titolo "Les déterminants économiques de la performance olympiques: Prévision des médailles qui seront gagnées aux Jeux de Pékin" è uno dei più importanti nell'ambito dell'Economia dello Sport, e senza dubbio il più importante insieme a quello di Bernard e Busse, se consideriamo solamente la letteratura dei Giochi Olimpici. I tre autori propongono di non fermarsi a considerare solamente le risorse economiche e la dimensione della popolazione come fattori determinanti del successo olimpico ma di tenere in considerazione anche le variabili che rappresentano i risultati olimpici individuali per atleta, la cultura sportiva nelle diverse regioni del mondo, il doping e le naturalizzazioni di atleti stranieri. Andreff e Poupaux affermano che

l'aumento delle medaglie vinte da un paese da un'Olimpiade all'altra è connesso alla diminuzione della quota medaglie di altri paesi. Quindi, per spiegare le prestazioni olimpiche di un paese, è necessario tenere conto di tutti gli altri paesi partecipanti ai Giochi e dell'interdipendenza tra le loro prestazioni sotto il vincolo del numero totale di medaglie assegnate durante i Giochi. I tre ricercatori francesi utilizzano come variabile dipendente del loro studio il numero di medaglie vinte per paese e come principali variabili indipendenti il PIL pro capite Y/N_{i, t-4} e la dimensione della popolazione N_{i, t} - 4. Queste due variabili riportano i valori di quattro anni prima dell'Olimpiade considerata, questo perché la programmazione e la preparazione per i Giochi inizia subito dopo la fine di quelli precedenti. Ci vuole, infatti, tempo per costruire una squadra olimpica competitiva, per riunire i migliori atleti, per allenarli intensamente, per prepararli psicologicamente e per motivarli al fine di renderli competitivi ed efficienti in occasione dei prossimi Giochi. Un periodo di quattro anni è realistico. I due Andreff e Poupaux includono anche una variabile dummy Hôte che indica il paese che ospita i Giochi olimpici. Così come Bernard e Busse anche i tre autori transalpini utilizzano delle variabili dummy per indicare la struttura politica del paese considerato, migliorando tuttavia il precedente modello. La loro classificazione tiene maggiormente conto della differenziazione tra i paesi in transizione. Un primo gruppo, denominato CEEC, riunisce i paesi dell'Europa centrale e orientale che hanno vissuto un regime sovietico e un'economia pianificata centralmente fino al 1989 o 1990 e, che da allora, sono stati trasformati in regimi democratici ed economie di mercato. Un secondo gruppo, denominato TRANS, riunisce i paesi che avevano un regime comunista e un'economia pianificata di tipo sovietico, e che hanno poi intrapreso un processo di transizione verso la democrazia e l'economia di mercato nel 1992, percorso in cui però hanno fatto molti meno progressi dei paesi appartenenti al CEEC. Questi paesi non erano membri dell'UE nel 2008 e non avevano intenzione di aderirvi nel prossimo futuro. Andreff ha raggruppato in un terzo gruppo, NSCOM, i paesi che, a differenza dei due gruppi precedenti, non erano regimi sovietici pur avendo un regime di tipo comunista e un'economia pianificata centralmente, che sono entrati in una transizione verso la democrazia e un'economia di mercato intorno al 1990. Infine, vi sono due paesi che non possono essere aggregati in nessuno dei tre gruppi precedenti: Corea del Nord e Cuba, che sono ancora oggi dei regimi comunisti con un'economia pianificata (COM). Un'ultima classe di paesi, CAPME, include le economie di mercato capitaliste. Inoltre, gli autori vogliono testare l'influenza della cultura di un paese sulle medaglie vinte. Si presume che la differenza culturale sia legata alle tradizioni riguardanti l'educazione fisica, la pratica sportiva, l'abitudine ad assistere a spettacoli sportivi o a guardarli in televisione. Queste tradizioni e atteggiamenti nei confronti dello sport nella società sono profondamente radicati nella cultura locale e nelle abitudini di comportamento e di pensiero, spesso abbastanza simili tra i paesi della stessa regione del mondo, e hanno portato Andreff ed i suoi coautori a servirsi di variabili dummy regionali per catturare quest'influenza culturale. Queste variabili possono assumere nove valori diversi: AFS per i paesi situati nell'Africa subsahariana, AFN per i paesi del Nord Africa, NAM per quelli del Nord America, LSA per l'America Latina ed il Sud America, EAST per i paesi dell'Europa orientale, WEO per i paesi dell'Europa occidentale, OCE per l'Oceania, MNE per il Medio e Vicino Oriente e ASI per gli altri paesi asiatici. Infine, è interessante vedere se la specializzazione sportiva delle nazioni influisca sulle loro prestazioni olimpiche. Le discipline olimpiche sono state quindi raggruppate in classi al fine di stimare se le possibilità di ottenere una medaglia siano significativamente diverse da una classe all'altra per gli atleti partecipanti ai Giochi. La scelta degli autori è ricaduta sull'unica classificazione economica degli sport ad oggi esistente, ovvero la Nomenclatura delle Attività Fisiche e Sportive (NAPS). Trentaquattro diverse famiglie di attività sportive sono raggruppate in 9 classi utilizzando un metodo di classificazione gerarchica basato su 12 criteri, alcuni dei quali economici, in particolare: il numero di ore di copertura televisiva, l'importanza data nelle colonne del quotidiano L'Equipe, il budget della federazione sportiva destinato alla pratica professionale, il numero di praticanti, il numero di atleti di alto livello ed il tipo di attrezzatura sportiva richiesta per la pratica. Il primo modello matematico proposto da Wladimir e Madeleine Andreff e Sandrine Poupaux è quindi il seguente:

$$\begin{aligned} \boldsymbol{M}_{i,t}^* &= c + \alpha \ln N_{i,t-4} + \beta \ln \left(\frac{\boldsymbol{Y}}{N}\right)_{i,t-4} + \gamma \, \boldsymbol{H} \\ \hat{\boldsymbol{\delta}} t \boldsymbol{e}_{i,t} &+ \sum_{p} \delta_{p} \, \boldsymbol{R} \\ \boldsymbol{\epsilon} \boldsymbol{g} i \boldsymbol{m} \boldsymbol{e} \, \boldsymbol{P} \boldsymbol{o} \boldsymbol{l} t \boldsymbol{i} \boldsymbol{q} \boldsymbol{u} \boldsymbol{e}_{p,i} + \\ & \sum_{q} \kappa_{q} \, \boldsymbol{a} \boldsymbol{n} \boldsymbol{n} \\ \boldsymbol{\epsilon}_{q,i} + \boldsymbol{\epsilon}_{i,t} \end{aligned}$$

, dove la variabile année è stata introdotta per verificare la significatività dell'anno. Questa significatività non è confermata dai risultati e quindi i tre autori decidono di introdurre una seconda versione del modello, sostanzialmente uguale al precedente, dove la variabile che indica l'anno viene sostituita dalla variabile dummy che indica il cluster geografico di appartenenza. Infine, Andreff presenta un terzo modello, più fedele alle regole del CIO, nel quale presume che le Olimpiadi siano una competizione tra singoli atleti. Il modello matematico è composto da due equazioni: la prima analizza i fattori che determinano la probabilità che un atleta partecipi ad una finale olimpica, la variabile dipendente può assumere quindi solo due valori, zero se l'atleta non partecipa ad una finale e 1 se si qualifica per una finale.

$$\begin{split} \mathit{Fin}_{j,i,t} &= c + \alpha \ln N_{i,t-4} + \beta \ln \left(\frac{Y}{N}\right)_{i,t-4} + \gamma \, \mathit{Hôte}_{i,t} + \sum_{p} \delta_{p} \, \mathit{R\'{e}gime} \, \mathit{Politique}_{p,i} \\ &+ \sum_{r} \rho_{r} \, \mathit{R\'{e}gions}_{r,i} + \sum_{s} \lambda_{s} \, \mathit{Sport}_{s,j} + u_{j,i,t} \end{split}$$

La seconda equazione, invece, analizza quei fattori che determinano la probabilità che un atleta vinca una medaglia. La variabile dipendente assume quindi il valore zero se l'atleta non vince nessuna medaglia e 1 se lo fa. La regressione applicata al primo modello evidenzia che tutti i coefficienti sono positivi e significativi, eccezion fatta per la variabile année. In linea con i risultati di Bernard e Busse, il numero di medaglie vinte è determinato dal PIL pro capite della nazione partecipante, dalla sua popolazione e dalla variabile Hôte: il Paese che ospita i Giochi beneficia di un surplus di medaglie vinte. Il regime politico è una variabile con un forte potere esplicativo nel caso dei paesi comunisti e nella transizione post-comunista. Gli stessi risultati si ottengono anche con il secondo modello, seppur questo porti ad un piccolo miglioramento offrendo un'ulteriore differenziazione culturale e regionale tra i paesi partecipanti ai Giochi. La variabile che rappresenta le differenze culturali regionali è significativa, eccetto che per l'America Latina. Diversi paesi latinoamericani sono infatti permeabili all'influenza della forte cultura sportiva nordamericana. Poiché lo studio ha preso come riferimento l'Europa occidentale, i coefficienti di segno positivo vengono interpretati come l'indice di una migliore prestazione olimpica, rispetto alla regione di riferimento, in termini di medaglie vinte, come nel caso dell'Africa subsahariana, del Nord America e dell'Oceania. Se per l'America del Nord questo risultato non risulta sorprendente per gli altri due cluster geografici lo è. I paesi sub-sahariani sono infatti tra i meno sviluppati al mondo, ed avrebbero dovuto avere un numero di medaglie molto basso. Tuttavia, a causa della loro elevata specializzazione in pochi sport, alcuni paesi di questo gruppo sono stati in grado di vincere un numero significativo di medaglie, ad esempio Etiopia e Kenya nell'atletica. Anche l'Oceania ha registrato una buona performance dovuta alla presenza in questo gruppo di due paesi sviluppati come Australia e Nuova Zelanda, vincitori di numerose medaglie, che fanno parte di un continente composto da una moltitudine di piccoli paesi poco sviluppati. Analogamente, i coefficienti negativi

associati a Nord Africa, Asia, Europa orientale e Medio e Vicino Oriente indicano una performance olimpica inferiore rispetto a quella dell'Europa occidentale. Questo non risulta sorprendente per il Medio e Vicino Oriente perché l'atteggiamento culturale nei confronti di alcune discipline sportive ed il limitato accesso delle donne alla pratica sportiva spiegano probabilmente una performance inferiore a quella europea. La stessa spiegazione vale in una certa misura per l'Asia, a cui appartengono alcuni paesi che hanno ottenuto buoni risultati in termini di medaglie come la Cina, ma anche grandi paesi, come Bangladesh e India, e paesi ricchi come ad esempio Singapore, che vincono relativamente poche medaglie. Il risultato più sorprendente è quello osservato per l'Europa dell'Est, dove la maggior parte dei paesi vince più medaglie rispetto a quelli dell'Europa occidentale, ed il segno negativo del coefficiente deriva probabilmente dal fatto che la performance superiore dell'Europa orientale è già pienamente colta dalla variabile che definisce la struttura politica. Anche il terzo modello di Andreff, quello applicato ai singoli atleti, mostra lo stesso risultato di base: il numero di medaglie vinte è spiegato dal PIL pro capite e dalla dimensione della popolazione dei paesi partecipanti. Inoltre, gli atleti provenienti da tutte le altre regioni del mondo hanno una probabilità significativamente inferiore di raggiungere una finale olimpica rispetto a quelli dell'Europa occidentale, eccezion fatta per coloro che provengono dell'Oceania. In aggiunta, il fatto che un atleta appartenga alla delegazione olimpica del paese che ospita l'Olimpiade non influisce in modo significativo sulla sua probabilità di vincere una medaglia, ed anzi la sua probabilità di raggiungere una finale olimpica si riduce lievemente. Prendendo come riferimento i paesi ad economia di mercato (CAPME), risulta evidente dai risultati che gli atleti provenienti dai paesi COM abbiano maggiori probabilità di vincere una medaglia; questo ragionamento risulta sostanzialmente vero, anche se in modo meno netto, per gli atleti originari dei paesi dell'Europa centrale e orientale (CEEC). Al contrario, l'impatto della struttura politica sulla probabilità di un atleta di vincere una medaglia non è significativamente più alto nei paesi in transizione (TRANS) rispetto a quelli del gruppo di confronto (CAPME). Gli atleti dei paesi che non adottano un'economia di mercato sono caratterizzati da una maggiore probabilità di raggiungere le finali olimpiche. Infine, ogni classe di sport ha un impatto significativo sulla probabilità di partecipare alle finali e sulla probabilità di vincere una medaglia ai Giochi Olimpici. Per la maggior parte delle classi sportive la probabilità di arrivare in finale e vincere una medaglia è inferiore rispetto alla classe di riferimento IDM, quella che rappresenta gli sport individuali di massa e con il maggior numero di eventi nel programma olimpico: atletica, ciclismo, ginnastica, nuoto. Per le due classi LEA, che include il sollevamento pesi, e SOL, alla quale appartengono il ping-pong e la marcia, la probabilità di andare in finale è superiore a quella della classe di riferimento. Per quanto riguarda la probabilità di vincere una medaglia, la classe LEA si comporta meglio della classe di riferimento mentre la classe SOL si comporta peggio. Le differenze tra le diverse classi possono essere, in parte, spiegate dal fatto che ci sono più o meno atleti che partecipano a diversi sport olimpici. Nonostante i tre modelli fin qui definiti si siano dimostrati piuttosto coerenti e robusti, Madeleine e Wladimir Andreff insieme a Sandrine Poupaux, sviluppano un quarto modello che tiene conto anche di una variabile denominata culto dell'Olimpismo, che indica il numero di medaglie ottenute da ciascun paese nella precedente edizione delle Olimpiadi.

$$\begin{split} \boldsymbol{M}_{i,t}^* &= c + \alpha \ln N_{i,t-4} + \beta \ln \left(\frac{Y}{N}\right)_{i,t-4} + \gamma \, H \hat{o} t e_{i,t} + \\ &+ \sum_{p} \delta_{p} \, R \acute{e} gime \, Politique_{p,i} + \theta \, M_{i,t-4} + \varepsilon_{i,t} \end{split}$$

I tre autori producono quindi due stime delle medaglie vinte ai Giochi di Pechino 2008, visibili in Tab. 2.19, la prima utilizzando il modello appena visto e la seconda calcolando la probabilità media

che gli atleti di un paese partecipino ad una finale e la probabilità media che gli atleti dello stesso paese che hanno partecipato a una finale vincano poi una medaglia; il prodotto di queste probabilità viene moltiplicato per il numero di atleti in gara per ogni paese.

		N	lodèle Tob	it	Modèle Logit à 2 étapes		
	Mé- dailles gagnées en 2004	Gains de mé- dailles prédits en 2008	borne infé- rieure	borne supé- rieure	Gains de mé- dailles prédits en 2008	borne infé- rieure	borne supé- rieure
CEEC :							
Bulgarie	12	12	10	13	13	11	14
Hongrie	17	19	17	21	29	25	31
Pologne	10	14	12	16	33	27	34
République tchèque	8	10	8	12	22	20	24
Roumanie	19	21	19	23	26	23	28
TRANS :							
Biélorussie	15	17	14	20	21	18	24
Kazakhstan	8	11	8	14	15	13	17
Russie	92	96	93	100	117	105	127
Ukraine	23	27	24	29	48	43	53

		N.	1odèle Tob	oit	Modèl	e Logit à 2	étapes
	Mé- dailles gagnées en 2004	Gains de mé- dailles prédits en 2008	borne infé- rieure	borne supé- rieure	Gains de mé- dailles prédits en 2008	borne infé- rieure	borne supé- rieure
NSCOM							
Chine	63	80	73	86	74	59	85
Cuba	27	29	25	33	21	16	25
CAPME :	-						
Allemagne	49	52	50	54	51	44	53
Australie	49	51	47	54	42	32	46
Canada	12	15	13	18	25	20	25
États-Unis	102	106	103	110	102	88	107
France	33	36	35	38	40	34	41
Italie	32	35	34	36	32	27	33
moins développés							
Brésil	10	12	10	14	14	11	17
Corée du Sud	30	30	27	32	30	24	33
Kenya	7	2	1	4	2	1	2
Turquie	10	9	7	11	6	4	8

Tab. 2.19: Previsioni del medagliere dei Giochi di Pechino 2008

La seconda previsione non include l'effetto di inerzia portato dal considerare il culto dell'Olimpismo. Questo perché non è possibile introdurre il numero di medaglie vinte quattro anni prima in una stima basata sul singolo e non sul paese. Questa assenza produce previsioni che agli occhi degli autori risultano meno realistiche. Tra le due previsioni non vi è una differenza sui nomi dei principali paesi vincitori di medaglie, quello che cambia è la classifica. Andreff afferma quindi che la previsione effettuata con il modello basato sul paese sia più verosimile, soprattutto perché più vicina alle previsioni espresse dagli intervistati: consulenti sportivi specializzati, non economisti.

L'anno successivo, nel 2009, sul British Journal of Sports Medicine viene pubblicato un breve articolo dal titolo "Why Great Britain's success in Beijing could have been anticipated and why it should continue beyond 2012". Gli autori di questo lavoro sono Alan Nevill, Nigel Balmer ed Edward Winter. I tre autori iniziano la loro ricerca affermando come mentre sia ben dimostrato che il paese che ospita le Olimpiadi abbia un notevole incremento del numero di medaglie vinte, non è altrettanto noto che i paesi abbiano un miglioramento delle proprie performance anche nell'edizione precedente e successiva di quella che organizzano. Infatti, sebbene alcuni studi abbiano riportato un maggior numero di medaglie vinte nei giochi immediatamente prima e dopo di quelli ospitati, agli autori non sembra che sia stato fatto alcun serio tentativo per quantificare quanto questi miglioramenti siano significativi. Il numero di medaglie vinte da ciascun paese è stato analizzato attraverso una binomiale utilizzando un modello di regressione logit. Il modello logit stima la probabilità o le probabilità di un evento, in questo caso la vittoria di una medaglia, come segue:

$$p = \exp(y)/[1 + \exp(y)]$$

dove la variabile y è espressa come

Nevill, Balmer e Winter prendono il caso delle medaglie vinte dalla Gran Bretagna a Pechino come esempio della loro ricerca, il modello logit diventa y = 23,42 + 0,277 [pre-host] = 23,15. Pertanto, la probabilità che la Gran Bretagna vincesse una medaglia è data da p = exp (23.15) / [1 + exp (23.15)] = 0.0410. Dato che a Pechino furono assegnate un totale di 958 medaglie, il modello prevede che la Gran Bretagna ne avrebbe vinte 958 * 0,0410, quindi 39 medaglie. In realtà la Gran Bretagna ha fatto meglio di quanto previsto dai tre ricercatori, conquistando, in totale, 47 medaglie, questo a causa di una serie di fattori contributivi, il primo è stato l'aumento esponenziale dei finanziamenti per lo sport in Gran Bretagna negli ultimi due decenni. Infatti, nel novembre 2008, il governo del Regno Unito ha annunciato un investimento di 550 milioni di sterline per sostenere 30 sport fino ai Giochi di Londra 2012. Circa il 15% di questi fondi è stato utilizzato per fornire supporto scientifico e medico. Un secondo fattore che ha contribuito al notevole successo della Gran Bretagna a Pechino è l'aver ospitato i Giochi del Commonwealth a Manchester nel 2002, un vantaggio identificato come simile all'effetto di post-hosting dei Giochi olimpici. I Giochi del Commonwealth hanno fornito nuove strutture per l'atletica, il tennis e lo squash, ma probabilmente lo sport che ha beneficiato di più delle infrastrutture create è stato il ciclismo, con la costruzione del velodromo ciclistico nazionale. Ciò ha contribuito alla vittoria della Gran Bretagna di otto medaglie d'oro nel ciclismo alle Olimpiadi di Pechino, nonché al notevole successo ottenuto dal Regno Unito nei campionati mondiali di ciclismo su pista all'inizio del 2008. Tuttavia, questo modello ha giustamente colto l'aumento del numero di medaglie vinte a Pechino, notevolmente superiore alle 28 e 30 medaglie vinte dalla Gran Bretagna rispettivamente a Sydney e Atene. Nevill, Balmer e Winter utilizzano questo modello anche per calcolare il numero di medaglie che la Gran Bretagna vincerà alle Olimpiadi di casa di Londra 2012. Sulla base del modello logit sopra si ottiene y = 23,42 + 0,718 [hosting] = 22,702, e quindi la probabilità di vincere una medaglia olimpica a Londra 2012 è dato da $p = \exp(22.702) / [1 + \exp(22.702)] = 0.0671 / [1 + 0.0671] = 0.063$. Sapendo che sarebbero state messe in palio circa 1000 medaglie nel 2012, gli autori hanno previsto 63 medaglie vinte dalla Gran Bretagna. I tre ricercatori, in modo del tutto analogo, hanno sviluppato una previsione del numero di medaglie che la Cina avrebbe potuto vincere a Londra 2012, dopo aver ospitato le Olimpiadi di Pechino. La probabilità che la Cina vinca una medaglia a Londra diventa p = exp (22,47) / [1 + exp (22,47)] = 0,0847 / [1 + 0,0847] = 0,078. Sotto l'ipotesi che a Londra 2012 sarebbero state messe in palio all'incirca 1000 medaglie, la Cina ne avrebbe potuto vincere ben 78. Secondo Neville ed i suoi colleghi la Gran Bretagna avrebbe dovuto prevedere di avere ai Giochi olimpici del 2016 un livello di successo non dissimile da quello ottenuto alle Olimpiadi di Pechino, portandosi a casa dal Brasile ben 46 medaglie. Naturalmente, l'analisi sviluppata dai tre ricercatori britannici è molto semplice ed elementare rispetto ad altri studi che hanno preso in considerazione una più vasta gamma di fattori. Tuttavia, Nevill, Balmer e Winter dimostrano la presenza di un netto vantaggio per il paese che ospita i Giochi olimpici, che si estende anche alle Olimpiadi che precedono e seguono immediatamente l'edizione di casa.

2.2.3 Dal 2010 fino ai giorni nostri

Nel febbraio del 2011 sulla rivista Sport Management Review viene pubblicato un brevissimo studio dal titolo "Olympic medals and demo-economic factors: Novel predictors, the ex-host effect, the exact role of team size, and the population-GDP model revisited", scritto dagli studiosi George Vagenas e Eleni Vlachokyriakou. Essi hanno riesaminato il problema del produrre una stima attendibile del successo olimpico mediante alcuni indicatori socioeconomici critici. Questi indicatori sono l'estensione territoriale del paese in km², la popolazione in milioni di abitanti, Il prodotto interno lordo pro capite in dollari, la percentuale della forza lavoro, l'inflazione, il tasso di crescita economica annuo, la spesa sanitaria in percentuale del PIL, la percentuale di popolazione che vive

nei centri urbani e la percentuale di disoccupazione. Dall'analisi di regressione svolta dai due ricercatori greci si evince che la dimensione della squadra olimpica sembri essere il miglior previsore del successo olimpico. Inoltre, il tasso di crescita, la disoccupazione e la spesa sanitaria risultano essere fattori fortemente significativi. Per ultimo le tradizionali variabili popolazione e PIL pro capite si sono rivelate meno significative rispetto agli altri indicatori utilizzati in questa ricerca e, quindi, insufficienti per spiegare il successo olimpico sulla base del numero di medaglie vinte dai paesi partecipanti.

Nel 2011 i già citati fratelli francesi Wladimir e Madeleine Andreff adattano la loro ricerca, riguardante i fattori che determinano il numero di medaglie alle Olimpiadi estive, ai Giochi olimpici invernali. Furono i primi a sviluppare un modello di questo tipo per le Olimpiadi invernali, utilizzando la stessa variabile dipendente del modello per i Giochi estivi, ovvero il numero di medaglie vinte da ciascuna nazione partecipante. Alcune variabili indipendenti possono essere mantenute per i Giochi invernali mentre altre nuove variabili devono essere introdotte per catturare la specificità delle discipline olimpiche invernali. Poiché la dimensione della popolazione, il PIL pro capite e la variabile che indica l'effetto dell'essere il paese ospitante sono risultati essere dei fattori determinanti delle medaglie vinte alle Olimpiadi estive, i due ricercatori francesi li hanno mantenuti anche nel modello riguardante le Olimpiadi invernali. Wladimir e Madeleine Andreff dividono i paesi, in base al reddito, in 5 gruppi: i paesi sviluppati con un'economia di mercato (DME), nuove economie emergenti (NEC), paesi in via di sviluppo (IIC) e paesi poco sviluppati (LDC). Sebbene, a prima vista, il regime politico sembri essere una variabile meno significativa per i Giochi invernali, lo hanno mantenuto nel modello con qualche lieve modifica rispetto al modello delle Olimpiadi estive. Il gruppo di paesi di riferimento rimane quello delle economie di mercato capitaliste (CAPME); CEEC sono, invece, quelle economie post-comuniste che hanno aderito all'UE nel 2004 o nel 2007; sono quindi state riunite tutte le rimanenti economie post-comuniste in un gruppo di paesi EXCOM, anche se sarebbe stato più ragionevole considerare Cuba e la Corea del Nord come regimi ancora comunisti. Per quanto riguarda la variabile dummy Regions, gli autori non si aspettano che sia significativa per le Olimpiadi invernali tanto quanto lo è stata per quelle estive. Il motivo è molto semplice: tutti quei paesi che partecipano ai Giochi invernali hanno in comune una cultura sportiva orientata alla pratica degli sport invernali indipendentemente dalla loro posizione geografica. Questo è confermato dal fatto che, contrariamente alle Olimpiadi estive, molti paesi nel mondo non partecipano alle Olimpiadi invernali. Pertanto, i fratelli Andreff la escludono dal modello per i Giochi invernali. I due autori ipotizzano che, se un paese volesse formare, allenare e selezionare atleti altamente performanti e competitivi negli sport invernali, non potrebbe farlo senza le condizioni meteorologiche adeguate, in particolare senza una sufficiente copertura nevosa annuale, e senza una minima presenza di località e strutture per la pratica degli sport invernali. Questa ipotesi li porta ad inserire nel modello due nuove variabili. La prima, Snow, è una variabile dummy che differenzia i paesi in base al livello medio di innevamento annuale, per il quale le nazioni partecipanti sono state suddivise in POL, HIGH, MIDDLE e LOW. La seconda variabile introdotta è una variabile dummy chiamata Resort, che ripartisce in tre gruppi i paesi prendendo in considerazione le stazioni sciistiche e gli impianti per sport invernali; quando una nazione ne ha a disposizione oltre 60 rientra nel gruppo MANY. Un paese con un numero di stazioni sciistiche compreso tra 5 e 60 è considerato come parte del gruppo BETWEEN. Un paese con un numero di stazioni sciistiche e strutture per sport invernali inferiore a 5 viene valutato e classificato come avente poche opportunità di vincere medaglie a causa della sua scarsa disponibilità di strutture e quindi parte del gruppo FEW/NO. Pertanto, il modello adattato dai fratelli Andreff per determinare quali siano i fattori chiave per la vittoria delle medaglie alle Olimpiadi invernali è il seguente:

$$\begin{split} \boldsymbol{M}_{i,t}^* &= c + \alpha \ln N_{i,t-4} + \beta \ln \left(\frac{\boldsymbol{Y}}{N} \right)_{i,t-4} + \gamma \operatorname{Host}_{i,t} + \sum_{p} \delta_{p} \operatorname{Political Regime}_{p,i} \\ &+ \sum_{r} \rho_{r} \operatorname{Snow}_{r} + \sum_{l} \lambda_{l} \operatorname{Resort}_{l} + \varepsilon_{i,t} \end{split}$$

Dove M*_{i, t} rappresenta il numero di medaglie per paese per una determinata edizione dei Giochi invernali, N è invece la popolazione della nazione i e Y/N è il suo PIL pro capite, entrambe le variabili assumono il valore di quattro anni prima per lo stesso motivo delle Olimpiadi estive. Wladimir e Madeleine Andreff conducono tre regressioni; la prima prende in considerazione cinque variabili, ovvero popolazione, PIL pro capite, e le tre variabili dummy Host, Snow e Resort. Il gruppo MIDDLE che contiene il maggior numero di paesi viene preso come riferimento per la variabile dummy Snow. Con la stessa logica, il gruppo FEW/NO è preso come riferimento per la variabile Resort. La seconda regressione include in aggiunta la variabile dummy del regime politico. Infine, la terza regressione è quella utilizzata dai due ricercatori per fare le previsioni del numero di medaglie vinte ai Giochi del 2014 e comprende la variabile M_{i, t-4}, che rappresenta il numero di medaglie conquistate dalla nazione i nelle precedenti Olimpiadi invernali.

Tutte e tre le analisi di regressione mostrano che il PIL pro capite e la dimensione della popolazione sono variabili altamente significative e con segno positivo. Le medaglie si distribuiscono principalmente nelle economie sviluppate con una popolazione piuttosto numerosa. Anche la presenza di località e strutture per la pratica degli sport invernali si dimostra essere un fattore molto significativo per la vittoria di medaglie. In modo del tutto inaspettato, la copertura nevosa non appare come un fattore determinante per la vittoria delle medaglie alle Olimpiadi invernali, questo vale per tutte e tre le regressioni. In effetti, secondo Wladimir e Madeleine Andreff alcuni paesi con un elevato manto nevoso non si comportano così bene ai Giochi olimpici invernali, come ad esempio il Tagikistan e il Kirghizistan. Non è sufficiente che un paese abbia molta neve, se non ha abbastanza stazioni sciistiche ed impianti sportivi invernali per allenare e formare i potenziali vincitori di medaglie ai Giochi invernali. Nella seconda analisi di regressione l'essere la nazione ospitante diventa un fattore significativo, così come essere una nazione classificata come appartenente ai gruppi CEEC o EXCOM aumenta la probabilità di vincere medaglie alle Olimpiadi invernali. La terza regressione si dimostra essere di gran lunga la migliore: tutte le variabili sono significative, tranne la quantità di neve. Anche la variabile che rappresenta il numero di medaglie conquistate alle precedenti Olimpiadi invernali è significativa, allo stesso modo lo è anche la variabile dummy Host. Secondo i due autori questo modello si adatta bene per sviluppare la previsione del medagliere dei Giochi invernali di Sochi, riportata in Tab. 2.20.

Countries	Medals won in 2010	Medal wins predicted in 2014	Lower bound	Upper bound
USA	37	36	33	38
Germany	30	28	26	30
Canada	26	27	25	28
Russia	15	24	21	27
Norway	23	24	22	25
Austria	16	15	14	16
Sweden	11	13	12	14
France	11	12	11	13
China	11	11	9	13
South Korea	14	11	10	13
Switzerland	9	9	8	10
Japan	5	7	6	9
Italy	5	7	6	8
Netherlands	8	6	5	7
Poland	6	6	4	8
Czech Republic	6	6	4	7
Finland	5	5	3	6
Australia	3	3	1	4
Slovenia	3	2	1	4
Croatia	3	2	0	4
Slovakia	3	2	0	3
Belarus	3	1	0	3

Tab. 2.20: La previsione del medagliere dei Giochi invernali di Sochi

I due fratelli Andreff affermano che queste previsioni, così come del resto tutte le altre, non considerano le sorprese e i risultati inattesi, come ad esempio il caso in cui è l'atleta o il team sfavorito a trionfare. I due autori concludono quindi affermando che quantificare questi risultati sportivi sorprendenti potrebbe essere la strada da percorrere per migliorare il loro modello. Sempre nel 2011 Adams Ceballos, Ramon Flores, I.G. McHale e Juan de Dios Tena Horrillo pubblicarono uno studio dal titolo "Explaining and forecasting national team medals totals at the Summer Olympic Games". I quattro autori partendo dal modello teorizzato da Bernard e Busse nel 2000, lo espandono e lo modificano cercando di renderlo più accurato e preciso. Essi infatti mantengono il PIL pro capite e la dimensione della popolazione, trasformando queste due variabili da valori assoluti a percentuali, la variabile dummy che indica il paese ospitante e le variabili dummy che indicano la struttura economico-politica della nazione considerata. Inoltre, Ceballos ed i suoi colleghi hanno introdotto una variabile che indica se un paese sarà la nazione ospitante della prossima edizione dei Giochi olimpici, e volevano introdurre anche una misura della spesa pubblica per lo sport. Se una tale misura avesse potuto essere valutata, avrebbe facilitato la stima del costo marginale della vittoria di una medaglia; tuttavia non vi erano dati a sufficienza. D'altro canto, si è deciso di aggiungere una variabile proxy, recspendshareit, che indica la percentuale della spesa pubblica di un paese in "affari ricreativi, culturali e religiosi", categoria che include anche lo sport. Si tratta comunque di una proxy perché questa categoria include anche molti ambiti poco legati allo sport. Il modello stimato da Ceballos e colleghi è quindi il seguente:

> $M_{it} = f(gdpshare_{it}, populationshare_{it}, host_{it}, futurehost_{it}, sovietbloc_{it},$ otherplannedeconomy_{it}, recspendshare_{it})

Dove M_{it} indica la percentuale di medaglie vinte dal paese i in una determinata edizione dei Giochi olimpici. Dai risultati della regressione effettuata sul modello dai quattro autori non sorprende che gli effetti positivi del PIL pro capite, dello status di paese ospitante e di (ex) appartenenza al blocco sovietico siano stati tutti confermati. Il coefficiente della variabile futurehost risulta fortemente significativo e simile a quello della stessa variabile host. Quest'ultima è risultata anch'essa

significativa. Questo modello è stato poi utilizzato dai quattro ricercatori per fare delle previsioni del medagliere delle Olimpiadi di Pechino 2008. Così come altre previsioni anche questa presenta degli errori; ad esempio, il modello aveva previsto 35 medaglie per il Giappone nel 2008, ma il paese nipponico ha raccolto solamente 25 successi, una decrescita inaspettata e difficilmente prevedibile rispetto alle 37 medaglie ottenute nei Giochi del 2004. Tuttavia, anche le altre previsioni non sono riuscite a segnalare questo drastico calo di performance del Giappone, per il quale, ad esempio, lo studio di Johnson e Ali del 2008 aveva previsto 37 medaglie. D'altra parte, il modello di Ceballos, Flores, Mchale e Horrillo ha individuato i tre maggiori cambiamenti nella distribuzione delle medaglie dal 2004 al 2008: l'exploit di Cina e Gran Bretagna ed il crollo della Russia. La Cina, ovvero la nazione ospitante dei Giochi, è arrivata seconda nel medagliere, piazzandosi alle spalle degli Stati Uniti, passando da 63 podi nel 2004 a 100 nel 2008. Gli studi del 2008 di Johnson e Ali e dei due fratelli Andreff avevano predetto rispettivamente 90 e 80 medaglie. La Gran Bretagna ha aumentato il numero di medaglie conquistate da 30 a 47; il modello sovrastante ne aveva previste 44, contro le 28 di Johnson e Ali, le 34 dei fratelli Andreff e le 38 del modello di Maennig e Wellbrock. Risulta quindi evidente che la previsione di Ceballos e colleghi ha beneficiato in modo significativo dell'introduzione della variabile futurehost.

Nel 2012 due autorevoli autori, Kuper e Sterken, di cui ho già ampiamente parlato in precedenza, pubblicano, sempre congiuntamente, uno studio dal titolo "Participation and Performance at the London 2012 Olympics", il cui obbiettivo è quello di costruire una previsione del medagliere attraverso l'analisi per ciascun paese della partecipazione e del successo nelle tre edizioni più recenti dei Giochi Olimpici Estivi. I due autori ipotizzano che i fattori che determinano la partecipazione siano di natura demografica, come la dimensione della popolazione, economica come ad esempio il reddito, e geografica, come la distanza misurata dal paese nel quale si svolgono i Giochi. Anche l'essere la nazione ospitante può rappresentare, secondo Kuper e Sterken, un vantaggio per la partecipazione. La distanza dai Giochi si può tradurre anche in spese di viaggio, che potrebbero essere considerate anche come una componente economica. Sempre secondo i due ricercatori i principali fattori che determinano il numero delle medaglie vinte alle Olimpiadi sono i risultati ottenuti ai campionati del mondo e la partecipazione numerosa degli atleti; quindi il PIL pro capite, la popolazione e la distanza hanno un effetto indiretto sul successo attraverso la partecipazione. Così come per la partecipazione, includiamo una variabile che considera il vantaggio dell'essere la nazione ospitante anche per il successo; per entrambi i modelli viene inoltre considerato il fatto che il paese che ha organizzato i Giochi possa beneficiare di questo vantaggio anche nell'edizione successiva. Kuper e Sterken svolgono tre diverse regressioni per l'equazione della partecipazione: la prima considerando solo gli atleti maschili, la seconda solo quelli femminili e la terza considerando le squadre. Analizzando i risultati si nota che il PIL pro capite ha un effetto positivo sulla partecipazione delle atlete. Per le squadre l'effetto è più debole e meno significativo. La dimensione della popolazione ha un effetto positivo significativo sulla partecipazione maschile e femminile. L'effetto è particolarmente forte per gli uomini. La distanza ha un effetto negativo solamente per le donne, mentre risulta insignificante per le altre categorie. L'effetto di essere il paese ospitante sulla partecipazione è significativo ed è circa il doppio per le squadre rispetto ai singoli atleti. L'essere stata la nazione ospitante nella precedente edizione ha un effetto significativo solo per le squadre. Gli autori svolgono poi nove regressioni sull'equazione che indica il successo olimpico, considerando sia il colore delle medaglie che i piazzamenti nelle categorie maschili, femminili e a squadre. I risultati portano Kuper e Sterken a concludere che i risultati del campionato del mondo e la partecipazione sono effettivamente fattori determinanti per il successo, soprattutto per la vittoria delle medaglie individuali maschili e femminili. L'essere la nazione organizzatrice non ha alcun effetto per le squadre, al contrario ha un effetto significativo e positivo per la vittoria di medaglie d'oro e d'argento per gli atleti di genere maschile. Tramite il modello da loro sintetizzato, Kuper e Sterken fanno una previsione del medagliere di Londra 2012, osservabile in Tab. 2.21. Questa previsione si compone di due step, per prima cosa i due autori hanno stimato la partecipazione ai Giochi, quindi hanno calcolato il successo condizionato alle previsioni di partecipazione. Ciò è stato fatto per gli atleti maschili e femminili. Per le squadre la partecipazione è nota con largo anticipo rispetto ai Giochi, quindi in questo caso è stata utilizzata la partecipazione effettiva.

Rank	Country	Gold	Silver	Bronze
1	China	44	11	22
2	United States of America	33	36	34
3	Russia	27	28	34
4	Great Britain	21	19	19
5	Australia	13	19	13
6	Japan	12	9	10
7	Germany	10	13	20
8	Italy	10	8	13
9	France	9	12	15
10	South Korea	8	11	12
11	Netherlands	8	8	7
12	Romania	8	4	7
13	Ukraine	7	7	13
14	Cuba	6	8	11
15	Hungary	6	5	3
16	Belarus	4	6	9
17	Greece	4	6	O
18	Brazil	4	4	6
19	Norway	4	3	1
20	Spain	3	7	6
21	Canada	3	6	7
22	Kenya	3	5	3
23	Poland	3	4	3
24	Jamaica	3	3	3
25	Ethiopia	3	2	3 3 3
26	New Zealand	3	1	3
27	Iran	3	1	1
28	Kazakhstan	2	4	3
29	Turkey	2 2 2	3	3
30	Czech Republic	2	3	2

Tab. 2.21: Previsione del medagliere di Londra 2012

Nel 2014 gli autori Jennifer L. Berdahl, Eric Luis Uhlmann e Feng Bai pubblicano sulla prestigiosa rivista "Journal of Experimental Social Psychology" una ricerca intitolata "Win-win: female and male athletes from more gender equal nations perform better in international sports competitions". I tre autori iniziano il loro lavoro affermando che, come dimostrato da precedenti studi, una maggiore parità di genere aumenta la redditività delle aziende, l'impatto scientifico delle collaborazioni intellettuali e la crescita economica di intere nazioni. Le cause di questi cambiamenti sono da ricercarsi nei maggiori guadagni ottenuti impiegando il capitale umano femminile, precedentemente sottoutilizzato, e nei cambiamenti all'interno delle dinamiche di squadra causate dalla diversità di genere. Berdahl, Uhlmann e Bai calcolano, come prima variabile del loro modello, un indice di gender gap per ogni paese, si tratta di una misura composita dell'uguaglianza economica, politica, sanitaria ed educativa tra i sessi. Valori di gender gap più elevati riflettono una maggiore uguaglianza sociale tra uomini e donne. I tre ricercatori tengono conto anche di altre variabili tra cui il PIL, la dimensione della popolazione, la latitudine geografica e la disuguaglianza di reddito, rappresentata dal coefficiente di Gini; valori elevati di tale coefficiente riflettono una maggiore disuguaglianza economica all'interno di un paese. La ricerca è stata condotta su 121 paesi e sono state analizzate le medaglie vinte alle Olimpiadi estive di Londra 2012 e alle Olimpiadi invernali di Sochi 2014. Dai risultati delle regressioni effettuate Berdahl e colleghi deducono che l'effetto del gender gap sia notevolmente significativo, e che questo indice sia un fattore determinante per la conquista di medaglie alle Olimpiadi. Anche il prodotto interno lordo e la dimensione della popolazione hanno un effetto positivo e significativo sulla vittoria delle medaglie, infatti osservando i dati sono i paesi più ricchi e quelli più grandi ad aver vinto più medaglie. La latitudine non si è dimostrata essere un fattore significativo.

I tre autori affermano che anche la disparità di reddito influenza attivamente il numero di medaglie ottenute. Maggiori le disuguaglianze economiche all'interno di un paese, inferiore il numero di medaglie conquistate. La suddivisione, per ogni paese, del gender gap in ciascuna delle sue quattro componenti, uguaglianza economica, politica, sanitaria ed educativa tra i sessi, ha rivelato che è proprio l'uguaglianza educativa ad avere l'effetto più rilevante, dal punto di vista dimensionale, e

ad essere il fattore migliore per predire il numero di medaglie conquistate ai Giochi olimpici. Una maggiore parità di genere, soprattutto a livello educativo ed accademico, permette di incrementare le vittorie olimpiche sia degli atleti di sesso maschile che di quelli di sesso femminile. Questi risultati, a dir poco sorprendenti, contraddicono con fermezza la convinzione comune secondo la quale un aumento delle opportunità per le donne si traduca inevitabilmente in uno svantaggio per gli uomini. Berdahl, Uhlmann e Bai concludono affermando che, piuttosto, è probabile che la disuguaglianza di genere danneggi sia le donne che gli uomini a causa degli stereotipi che limitano la loro capacità di raggiungere il loro pieno potenziale come individui. La diminuzione della discriminazione di genere ed una maggiore parità tra i sessi rappresenta quindi una situazione win-win, che consente ai membri di entrambi i generi di esprimere il loro vero potenziale.

Nel 2016 i ricercatori Julia Bredtmann, Carsten J. Crede and Sebastian Otten pubblicano sulla prestigiosa rivista "The Royal Statistical Society" una breve ricerca dal titolo "Olympic medals: Does the past predict the future?", nella quale i tre autori investigano su quali siano i fattori che determinano il successo di una nazione alle Olimpiadi. La ricerca dei tre autori si basa su due modelli; il primo, quello più semplice, è

$$Medals_{it} = \beta_1 + \beta_2 Lag Medals_{it} + \beta_3 Year_t + \epsilon_{it}$$

dove Medals_{it} indica il numero totale di medaglie vinte da un paese alle Olimpiadi corrispondenti, la variabile Lag Medals_{it}, invece, rappresenta le medaglie ottenute dal medesimo paese nell'edizione precedente dei Giochi e Year_t indica l'anno nel quale si svolge la manifestazione. Quest'ultima variabile è utilizzata per tenere conto del costante aumento nel tempo del numero totale di medaglie messe in palio ai Giochi Olimpici. Il secondo modello, più sofisticato, include delle variabili aggiuntive che tengono conto di alcune caratteristiche specifiche di un paese:

$$\begin{aligned} \textit{Medals}_{it} &= \gamma_{1} + \gamma_{2} \textit{Lag Medals}_{it} + \gamma_{3} \ln \textit{GDP}_{it} + \gamma_{4} \ln \textit{Pop}_{it} + \gamma_{5} \textit{Host}_{it} + \gamma_{6} \textit{Next Host}_{it} \\ &+ \gamma_{7} \textit{Planned}_{it} + \gamma_{8} \textit{Muslim}_{it} + \gamma_{9} \textit{Year}_{t} + \zeta_{it} \end{aligned}$$

dove In GDP_{it} indica il logaritmo naturale del PIL pro capite e In Pop_{it} è il logaritmo naturale della dimensione della popolazione. Entrambe le variabili sono espresse come logaritmi per tenere in considerazione il fatto che gli effetti positivi del PIL pro capite e delle dimensioni della popolazione sulle medaglie olimpiche diminuiscono all'aumentare dei valori di queste variabili. Host_{it} e Next Hostit indicano rispettivamente il paese ospitante dell'attuale edizione della manifestazione ed il paese ospitante nell'edizione successiva. Plannedit è una variabile che indica se un paese ha o ha avuto in passato un'economia centralmente pianificata, come ad esempio gli ex membri dell'Unione Sovietica, la Cina o Cuba. Infine, Muslimit è un indicatore per i paesi con una popolazione a maggioranza musulmana, che tendono a inviare meno atlete alle Olimpiadi e che di conseguenza hanno una quota inferiore della popolazione femminile attiva negli sport professionistici. Utilizzando i due modelli visti i tre ricercatori hanno sviluppato due diverse previsioni del medagliere delle Olimpiadi di Londra 2012, osservabili in Tab. 2.22, e le hanno poi confrontate con i risultati effettivi per comprendere quale metodo fosse più preciso. Per un miglior raffronto Bredtmann, Crede ed Otten hanno calcolato l'errore medio assoluto (MAE) e l'errore medio di previsione (MFE) sia per il modello più semplice che per quello più sofisticato. Se si considera il campione esteso comprendente 181 paesi, l'errore medio assoluto è di 1,43 per il primo modello e di 1,41 per il secondo modello. In altre parole, in media, nel modello più semplice le previsioni si sono discostate

di 1,43 medaglie dal numero reale di medaglie vinte da ciascun paese e di 1,41 nel modello più sofisticato. Tuttavia, considerando solo quei paesi classificatisi nelle prime 15 posizioni del medagliere, la differenza dell'errore medio assoluto tra i due modelli risulta più marcata, passando rispettivamente a 6,6 per il primo modello e a 5,8 per il secondo.

Allo stesso tempo, le previsioni ottenute utilizzando il modello raffinato sono soggette a una minore incertezza, infatti in questo caso l'errore medio di previsione si attesta a 3,2 rispetto a 3,4 del modello elementare. La previsione effettuata usando la prima equazione si discosta di 106 medaglie dalle 654 medaglie effettivamente vinte dai primi quindici paesi. L'outlier più significativo è il Regno Unito, il paese che ha ospitato l'edizione del 2012, che ha vinto 18 medaglie in più rispetto a quanto previsto dal modello, che ignorava totalmente l'host effect. D'altra parte, osservando le previsioni fornite dagli autori utilizzando il secondo modello, si nota che queste sono molto più vicine ai risultati effettivi rispetto a quelle ottenute con il primo modello. Questo risultato è in linea con quanto visto in precedenza con l'analisi dei valori degli errori medi. La differenza assoluta tra le medaglie previste e quelle effettive si riduce da 106 a 93 medaglie. L'aumento della precisione delle previsioni deriva principalmente da minori differenze per il Regno Unito e per gli Stati Uniti. Per il Regno Unito, lo scostamento passa da 18 a 5 medaglie, poiché in questo caso si è preso in considerazione, nel modello, il fatto che gli UK fossero il paese ospitante. Secondo Bredtmann, Crede ed Otten, questi risultati, riassunti nella tabella a destra, suggeriscono che l'aggiunta delle variabili socioeconomiche al modello aumenta leggermente la precisione delle previsioni. Siccome il modello più raffinato si è dimostrato sufficientemente preciso ed accurato per Giochi olimpici di Londra 2012, i tre autori decidono di utilizzarlo per fornire una previsione del medagliere olimpico di Rio 2016, anch'essa osservabile in Tab. 2.22, come si può osservare nella tabella sottostante. L'unico modo per valutare l'affidabilità di tale previsione è di calcolarne l'errore medio di previsione, che scende a 3,1 per le Olimpiadi di Rio 2016 da 3,2 per i Giochi di Londra 2012. I tre ricercatori si attendono che Brasile e Giappone registreranno il maggior incremento di medaglie. Infatti, come paese ospitante, il Brasile porterà alla manifestazione un maggior numero di atleti rispetto ai Giochi precedenti, e godrà dell'incremento degli investimenti in ambito sportivo; in più gli atleti brasiliani potranno godere anche del supporto del pubblico di casa. Allo stesso modo, il Giappone, che ospiterà le Olimpiadi del 2021, dovrebbe già raccogliere i frutti degli investimenti fatti per preparare i suoi atleti e le sue strutture per l'edizione casalinga dei Giochi. Bredtmann, Crede ed Otten concludono affermando che sebbene i risultati delle precedenti edizioni e le condizioni socioeconomiche siano dei fattori determinanti, rimane un certo livello di imprevedibilità tipico di qualunque competizione sportiva.

	Actua	l results	Naive mo	del	Sophist	icated	model
Country	2012 rank	Medals	Predicted medals for 2012	Diff.	Predicted medals for 2012	Diff.	Predicted 2012 rank
USA	1	103	110	7	105	2	1
China	2	88	100	12	96	8	2
Russia	3	81	73	8	70	11	3
UK	4	65	47	18	60	5	4
Germany	5	44	41	3	39	5	6
Japan	6	38	25	13	24	14	12
Australia	7	35	46	11	44	9	5
France	8	34	41	7	39	5	6
Italy	9	28	27	1	26	2	10
South Korea	9	28	31	3	30	2	8
Nether- lands	11	20	16	4	16	4	17
Ukraine	11	20	27	7	27	7	9
Canada	13	18	19	1	19	1	14
Hungary	13	18	10	8	10	8	21
Brazil	15	17	15	2	26	9	10
Spain	15	17	18	1	18	1	15
Sum		654	646	106	649	93	
MAE				6.6		5.8	
MFE			3.4		3.2		

Country	Predicted rank for 2016	Predicted medals for 2016	Rank in 2012	Medals in 2012	Diff. in ranks	Diff. in medals to 2012
USA	1	98	1	103	0	-5
China	2	84	2	88	0	-4
Russia	3	77	3	81	0	-4
UK	4	62	4	65	0	-3
Japan	5	46	6	38	1	8
Germany	6	42	5	44	-1	-2
Australia	7	33	7	35	0	-2
Brazil	7	33	15	17	8	16
France	7	33	8	34	1	-1
Italy	10	27	9	28	-1	-1
South Korea	10	27	9	28	-1	-1
Ukraine	12	20	11	20	-1	0
Netherlands	13	19	11	20	-2	-1
Canada	14	18	13	18	-1	0
Hungary	14	18	13	18	-1	0
MFE		3.1				

Tab. 2.22: Previsioni del medagliere di Londra 2012 e Rio 2016

Sempre nel 2016 i ricercatori David Forrest, Ian G. McHale, Ismael Sanz e J. D. Tena pubblicano un articolo dal titolo "An analysis of country medal shares in individual sports at the Olympics" nel quale analizzano come alcuni sport siano più o meno correlati con i livelli di reddito e l'host effect sia più pronunciato su alcuni sport rispetto ad altri. L'innovazione portata dai quattro autori in questo lavoro è quella di presentare una versione leggermente modificata del modello di Bernard e Busse e utilizzarla a livello di singolo sport. Il modello prende in considerazione 15 sport nel periodo considerato dai Giochi del 1992 a quelli del 2012, per i quali sono state messe in palio almeno 100 medaglie. Forrest, McHale, Sanz e Tena hanno calcolato per ciascuno dei quindici sport, per ogni edizione dei Giochi nel periodo preso in considerazione, l'indice di Herfindahl Hirschman (HHI), questo è un indicatore statistico, che va da 0 ad 1, concepito originariamente per sintetizzare il grado di concentrazione di mercato in un settore: maggiore è l'indice, maggiore è il grado di concentrazione del mercato. In questo caso, invece, viene utilizzato dai ricercatori per misurare il grado di concentrazione della vittoria di medaglie in uno sport. Ad esempio, osservando Tab. 2.23 sottostante, l'indice HHI per i tuffi è sempre maggiore di quello calcolato per l'atletica, quindi, il successo nelle competizioni acquatiche è limitato a meno paesi, mentre nell'atletica le nazioni che riescono a trionfare sono di più.

	1992	1996	2000	2004	2008	2012
Athletics	0.10	0.06	0.04	0.07	0.07	0.08
Boxing	0.07	0.08	0.08	0.08	0.08	0.07
Canoe/kayak	0.10	0.09	0.09	0.08	0.08	0.12
Cycling	0.11	0.11	0.11	0.09	0.13	0.08
Diving	0.31	0.26	0.24	0.24	0.28	0.24
Equestrian	0.22	0.17	0.15	0.18	0.17	0.19
Fencing	0.15	0.18	0.13	0.14	0.13	0.13
Gymnastics	0.22	0.13	0.20	0.12	0.16	0.12
Judo	0.09	0.09	0.07	0.08	0.06	0.07
Rowing	0.12	0.09	0.08	0.06	0.07	0.10
Sailing	0.16	0.05	0.09	0.07	0.08	0.09
Shooting	0.10	0.08	0.07	0.09	0.09	0.07
Swimming	0.14	0.13	0.17	0.13	0.15	0.14
Weightlifting	0.16	0.10	0.07	0.10	0.10	0.08
Wrestling	0.12	0.07	0.09	0.08	0.06	0.07
All sports modelled	0.07	0.05	0.04	0.05	0.05	0.04
All sports in Games	0.06	0.04	0.04	0.04	0.05	0.04

Tab. 2.23: L'indice HHI per gli sport olimpici

Il modello messo in atto da Forrest e colleghi ha l'obbiettivo di rivelare se tali differenze sono causate dal modo in cui gli sport sono influenzati dal PIL pro-capite e da altre variabili socioeconomiche simili a quelle utilizzate da Bernard e Busse. La funzione generata è la seguente:

 $medal\ share_{it} = f(\log\ population_{it},\ \log\ gdp_{it},\ host_{it},\ ex-Soviet\ Union_{it},\ ex-Soviet\ bloc_{it},\ planned_{it},\ t_1\ \ldots\ t_6),$

dove, oltre al logaritmo della popolazione e del PIL pro capite, troviamo tre variabile dummy di natura politica ed economica, ed una che tiene conto dell'host effect. La variabile ex-Soviet Union identifica i paesi che prima del 1990 facevano parte dell'Unione Sovietica. Ex-Soviet bloc comprende invece quei paesi i cui territori erano sotto la sfera di influenza sovietica. Infine, planned descrive un'economia pianificata centralmente, come Cina e Cuba. Su questo modello i quattro ricercatori conducono poi un'analisi di regressione i cui i risultati sono riportati nella Tab. 2.24.

	Log	Log per capita				Ex-Soviet	
	population	GDP	Host	Planned	Ex-USSR	bloc	ρ
Athletics (815)	0.012***	0.014***	0.008	0.008	0.020**	0.010	0.823***
Boxing (283)	0.026***	0.029***	0.058***	0.066	0.104***	0.042**	0.597***
Canoe/kayak (288)	0.030***	0.058***	0.038	0.052	0.091***	0.119***	0.642***
Cycling (259)	0.044***	0.100***	0.005	0.055**	0.127***	0.047	0.486***
Diving (120)	0.112***	0.104***	0.100*	0.265	0.193	-0.616	0.702***
Equestrian (222)	0.087***	0.219***	0.113	-9.332	9.937	-8.895***	0.703***
Fencing (174)	0.082***	0.11***	-0.088	0.183**	0.147**	0.189***	0.684***
Gymnastics (594)	0.067***	0.076***	0.071**	-0.089*	0.203***	0.179***	0.472***
Judo (336)	0.025***	0.033***	0.032**	0.070**	0.078***	0.044**	0.700***
Rowing (252)	0.032***	0.062***	0.052***	0.016	0.092***	0.105***	0.719***
Sailing (189)	0.037***	0.091***	0.058**	0.038	0.058*	0.037	0.576***
Shooting (274)	0.027***	0.044***	0.023	0.086***	0.095***	0.102***	0.499***
Swimming (589)	0.029***	0.046***	0.015	0.041	0.032	0.057**	0.725***
Weightlifting (233)	0.043***	0.040***	-0.008	0.094***	0.152***	0.106***	0.587***
Wrestling (740)	0.026***	0.027***	0.020	0.035	0.119***	0.050***	0.722***
All sports modelled (4576)	0.008***	0.009***	0.017***	0.012	0.019***	0.009**	0.854***
All sports in Games (5401)	0.008***	0.008***	0.020***	0.014*	0.016***	0.009**	0.861***

Tab. 2.24: Analisi di regressione

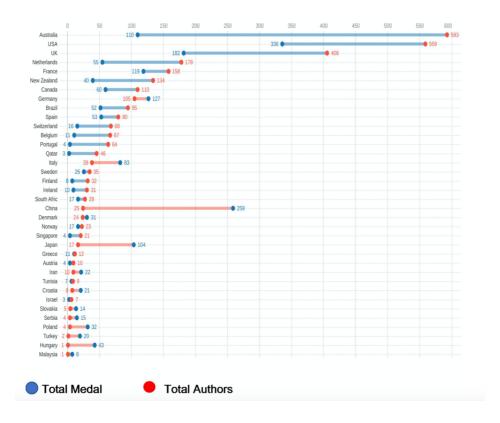
Per tutti e quindici gli sport considerati i coefficienti dei logaritmi del reddito e della popolazione sono entrambi significativi e positivi; l'effetto dell'host effect è più eterogeneo: la variabile host è statisticamente significativa, sempre positiva, solo in sei dei 15 sport esaminati individualmente.

Dal confronto dei coefficienti del logaritmo della popolazione e del logaritmo del PIL pro capite di questo modello con quelli di Bernard e Busse si nota che i primi sono dimensionalmente inferiori rispetto ai secondi, questo risultato è in linea con il suggerimento di De Bosscher, che nel 2006 affermava che l'importanza del reddito e della popolazione, sebbene ancora molto elevata, fosse diminuita nel tempo. Una delle possibili spiegazioni che forniscono i quattro autori è che con l'aumento della globalizzazione gli atleti di paesi relativamente poveri possono allenarsi nelle nazioni più ricche grazie ai flussi migratori e alle borse di studio. La relazione tra medaglie e PIL pro capite come già visto in precedenza si basa sul fatto che i cittadini di paesi ricchi hanno più tempo da dedicare allo sport e maggiore margine di spesa; di conseguenza i paesi più ricchi trionferanno negli sport più costosi da praticare, come l'equitazione, la scherma ed i tuffi. I paesi ospitanti invece tenderanno a dominare in quegli sport come tuffi, ginnastica e pugilato, che si basano quasi interamente sul verdetto dei giudici, o negli sport come la vela ed il canottaggio, in cui la familiarità con le condizioni ambientali risulta essere un vantaggio non trascurabile. Forrest, McHale, Sanz e Tena concludono il lavoro affermando che i paesi più poveri dovrebbero concentrarsi e specializzarsi in quegli sport sui quali l'effetto del PIL pro capite è meno significativo.

Nel 2020 i due fratelli francesi Wladimir e Madeleine Andreff riprendono la loro ricerca del 2008, dal titolo "Les déterminants économiques de la performance olympiques: prévision des médailles qui seront gagnées aux Jeux de Pékin", ed insieme ad altri tre ricercatori, Nicolas Scelles, Liliane Bonnal e Pascal Favard, la modificano, migliorandola e rinnovandola. Al modello originale per i Giochi olimpici estivi dei fratelli Andreff vengono aggiunte tre variabili; la prima variabile indica se il paese considerato ha ospitato l'edizione dei Giochi precedente, la seconda indica se il paese ospiterà l'edizione successiva ed infine la terza variabile indica il numero di atleti che la nazione presa in considerazione porta alla manifestazione olimpica. Partendo da questo rinnovato modello gli autori hanno fornito una previsione delle medaglie vinte da ciascun paese per le Olimpiadi di Tokyo 2020.

Infine, nell'aprile del 2020 due scienziati turchi, Süleyman Şahin e Fatih Şenduran, hanno pubblicato sull'African Educational Research Journal l'importante articolo "Effect of sports health and exercise research on Olympic game success: an analytical and correlational survey", il cui scopo è quello di indagare quale sia l'influenza del livello di sviluppo della medicina sportiva di un paese sui suoi successi ai Giochi Olimpici. I due autori iniziano l'articolo affermando che oltre ai fattori

socioeconomici e politici, ad influenzare i risultati delle Olimpiadi sono anche e soprattutto le abilità fisiche e tattiche degli atleti. Oggi, grazie ad allenamenti mirati e alla ricerca medica con il supporto della scienza e dello sport, gli atleti raggiungono limiti fisici e tattici sempre più elevati: i paesi utilizzano molto le scoperte provenienti dal mondo accademico ed universitario per avere successo alle Olimpiadi. Şahin e Şenduran pongono quindi l'attenzione sul nesso tra l'efficacia degli studi scientifici sullo sport e della ricerca medica in ambito sportivo e la vittoria di medaglie ai Giochi olimpici. I due autori raccolgono, analizzano e catalogano gli articoli scientifici pubblicati sulla rivista Sport Medicine nel periodo che va dal 2010 al 2018, dividendoli in base alla nazionalità dell'autore o degli autori; hanno catalogato circa 1027 articoli scritti da 3011 autori. Al fine di verificare la relazione tra il numero totale di autori di un paese ed il numero di medaglie vinte da quel paese nelle Olimpiadi estive del 2016, del 2012 e del 2008, è stato utilizzato un test di correlazione di Pearson, con un livello di significatività inferiore al 5%.



Tab. 2.25: Il totale delle medaglie e degli autori per paese

Osservando la Tab. 2.25 elaborata dai due studiosi si nota che, considerando il numero totale di autori, delle nazioni ed il numero totale di medaglie olimpiche conquistate nelle tre edizioni dei Giochi olimpici considerate, Australia, USA, Inghilterra, Olanda e Francia si sono classificati nelle prime cinque posizioni. Per il numero totale di medaglie conquistate nelle tre Olimpiadi, il coefficiente esplicativo R² si è attestato al 55% per il numero totale degli autori. In altre parole, il numero di medaglie alle Olimpiadi può essere spiegato per il 55% dal numero totale di autori, ed il restante 45% può essere previsto tramite altri fattori come il PIL, la dimensione della popolazione, il livello d'istruzione e gli investimenti in impianti sportivi e strutture per l'allenamento. Lo studio condotto dai due ricercatori turchi mostra che esiste una correlazione tra il numero di pubblicazioni accademiche in un paese e il numero di medaglie vinte da quel paese, e che questa correlazione è significativa e assume un valore pari a 0,74. Tuttavia, secondo i due autori turchi questa ricerca non è completa; infatti, sulla rivista Sport Medicine sono pubblicati solamente gli articoli scientifici scritti

in lingua inglese. In questo studio vengono poco considerati paesi come Cina, Giappone, Grecia e Russia, dove si tende a scrivere e pubblicare le ricerche nella lingua locale, che risultano erroneamente svantaggiati. Şahin e Şenduran concludono affermando che si deduce che un aumento sia a livello qualitativo che a livello quantitativo della ricerca nell'ambito delle Scienze dello Sport aumenta statisticamente la possibilità di vincere una medaglia olimpica.

2.3 Conclusione

La letteratura socio-economica ha cercato di spiegare cosa determina il successo sportivo di una nazione alle Olimpiadi. L'approccio e la metodologia utilizzata negli studi esaminati si sono decisamente evoluti nel tempo. Le ricerche pubblicate negli anni '70 hanno un approccio più politico e mettono in luce il vantaggio dato dall'adottare una struttura politico-economica di tipo socialista. Infatti, sia le ricerche di Ball e di Nevill che quelle di Grimes, Kelly e Rubin, tutte risalenti ai primi anni '70, prendono in considerazione solamente le variabili politiche che identificano i paesi socialisti da quelli che adottano un'economia di mercato e le variabili che riportano il vantaggio dato dall'essere la nazione ospitante. Analogamente anche gli studi scientifici degli anni '90 mantengono il profilo politico cercando però di spiegare, attraverso i cambiamenti avvenuti nel panorama politico mondiale, le variazioni del medagliere olimpico. Nel 2000 poi, grazie a Clarke, si ripartì prendendo in considerazione la variabile dicotomica che coglie l'host effect, che non sarebbe più stata abbandonata negli studi successivi. Il lavoro di Clarke non si preoccupava minimamente di fornire una spiegazione del successo olimpico nel suo insieme, ma, come altri lavori pubblicati in seguito, voleva solamente dimostrare la presenza di un incremento di medaglie per la nazione ospitante. Nello stesso anno Johnson e Ali riuniscono insieme quanto era emerso dagli studi precedenti sviluppando, per la prima volta, un modello per la partecipazione ed un modello per il successo, che oltre a basarsi sulle variabili economiche, demografiche, politiche e geografiche introducevano anche un effetto dovuto al colonialismo, che pur dimostrandosi significativo non venne più ripreso nelle ricerche successive. Tuttavia, questi due modelli non si dimostrano in grado di fornire previsioni accurate del successo olimpico. La ricerca di Johnson e Ali venne ripresa un anno dopo da Sterken e Kuper, la cui innovazione evidente risulta nel fatto di modellizzare la quota di medaglie di un paese come funzione della sua quota di partecipazione, e affermano i due letterati che la partecipazione è determinata da fattori economici e geografici, mentre il successo è fortemente legato all'home advantage ed al PIL pro capite. Tuttavia, nemmeno questo modello è in grado di fornire delle previsioni accurate. Al contrario, l'articolo di Bernard e Busse identifica lo sviluppo economico come principale determinante del successo olimpico, e fornisce per la prima volta delle previsioni che si dimostrano accurate e precise, dimostrandosi il modello migliore tra quelli teorizzati in quegli anni. I lavori, che seguono, continuano a fondarsi sulla ricchezza, sulla dimensione della popolazione e sulla struttura politica, ma aggiungono alle variabili geografiche anche quelle climatiche, che si dimostrano significative. Negli anni a venire vennero introdotti altri metodi come il Data Envelopment Analysis la cui funzione più che sviluppare delle previsioni della performance è quella di misurare l'efficienza di un paese alle Olimpiadi. Pfau fu il primo autore ad inserire, nel 2005, delle variabili che colgono l'appartenenza a determinate regioni geografiche e che, per quanto riguarda i Giochi olimpici invernali, gli permisero di fornire le previsioni migliori in termini di coefficiente di correlazione. Nel 2008 i fratelli Andreff pubblicano il lavoro, a mio avviso, migliore nella letteratura dei Giochi olimpici, dal quale ho preso spunto per la stesura di questa tesi. Essi pongono l'accento sulla struttura politica e sulla localizzazione geografica, le previsioni che ne derivano sono infatti più vicine alle previsioni espresse dai consulenti sportivi specializzati. Col passare degli anni alle variabili più classiche viste fin ora sono state inserite anche delle variabili di tipo educativo, sanitario o scientifico, il cui obbiettivo era rappresentare una proxy di un

determinato effetto sulle medaglie, e sul numero di medaglie vinte nelle Olimpiadi precedenti. L'host effect è stato, anche dai fratelli Andreff, esteso sia all'edizione precedente che a quella successiva all'Olimpiade organizzata in casa; grazie all'introduzione di queste due variabili dicotomiche le previsioni del medagliere dei Giochi olimpici estivi si avvicinano molto di più ai dati reali, rispetto ai modelli passati. Nella seconda metà degli anni Dieci del terzo millennio si è cercato di individuare altri fattori che possano determinare il numero di medaglie vinte da un paese alle Olimpiadi, e sono state quindi introdotte due variabili di tipo sociale come l'indice di parità dei sessi e la religione. Nel corso degli anni sono state utilizzate varie tipologie di variabili negli studi che trattano questo argomento e queste variabili sono state associate in modo diverso tra loro, ma alla fine, la maggioranza degli autori, sono giunti alla conclusione che lo sviluppo economico è il principale fattore determinante per il successo sportivo. Che si tratti di Olimpiadi estive o invernali, la ricchezza, l'host effect e la dimensione della popolazione sono le tre variabili che hanno maggiore influenza nel determinare il successo olimpico di uno stato e quindi ad ottenere il maggior numero di medaglie. Proprio per questo motivo ho deciso di basare la mia ricerca su queste tre variabili, per valutare se mantengano ancora il loro potere esplicativo e se confermino ancora una volta quanto emerge dalla letteratura anche per le ultime edizioni olimpiche. Estrema attenzione è stata rivolta anche alla struttura politica ed al boost in termini di medaglie che ricevevano i paesi comunisti e quelli che erano sotto la sfera d'influenza sovietica. Risulta opportuno chiedersi se ancora oggi, a distanza di molti anni dallo scioglimento dell'Unione Sovietica e dalla caduta del muro di Berlino, questo vantaggio continui a mantenersi. Anche le variabili di appartenenza geografica, che sono state incluse nei più importanti lavori, necessitano di essere ulteriormente testate per comprendere se il loro effetto sul successo alle Olimpiadi si sia acuito o si sia ridotto. Vi sono poi le variabili che non rappresentano direttamente un fenomeno ma che ne sono solamente una proxy, come quelle sociali, educative, sanitarie e scientifiche, che ci costringono ad ulteriori verifiche, per poter cogliere come si relazionano con le medaglie vinte e con le altre variabili più tradizionali, con cui spesso risultano essere fortemente correlate. Nel corso degli anni si sono susseguite numerose ricerche che hanno utilizzato le variabili economiche e demografiche ritardate di 4 anni e ricerche che hanno utilizzato i valori corrispondenti all'anno delle Olimpiadi considerate. Sarà quindi interessante verificare se esiste una vera e propria differenza in termini di significatività nell'utilizzare le variabili ritardate nel tempo; infatti, i parametri economici e demografici cambiano molto lentamente nel tempo e lo scostamento nei quattro anni risulta quasi sempre impercettibile. Questa letteratura presenta comunque dei limiti poiché non considera alcuni fattori, come ad esempio il doping, di cui sono soliti fare uso alcuni paesi, come la Russia. In conclusione, quantificare questi fattori potrebbe essere una delle strade da percorrere per la ricerca in questo ambito.

3 Metodologia

L'obbiettivo della presente dissertazione consiste nello studiare le determinanti della performance sportiva alle Olimpiadi.

In questo capitolo vengono individuati i dati da raccogliere, per poi soffermarsi sulle variabili indipendenti, cioè i fattori che influenzano e determinano il numero di medaglie vinto da un Paese alle Olimpiadi, che rappresenta la variabile dipendente. Tali covariate indipendenti si possono classificare in sei macro categorie: economiche, geografiche, politiche, sociali, scientifiche, e demografiche.

3.1 Dati e fonti

Per questa ricerca sono state prese in considerazione cinque edizioni dei Giochi Olimpici estivi: Sydney 2000, Atene 2004, Pechino 2008, Londra 2012 e Rio 2016. Questa scelta è motivata dal fatto che si tratta delle kermesse più recenti e più facilmente comparabili.

La forte presenza di dati missing ha suggerito di escludere dall'analisi i Giochi olimpici estivi di Atlanta 1996 e di Seul 1988.

Le Olimpiadi di Barcellona 1992 non sono oggetto della ricerca in quanto si sono svolte immediatamente dopo la caduta del muro di Berlino e lo scioglimento dell'Unione Sovietica; nei Giochi catalani tutti i paesi appartenenti al blocco sovietico si presentarono sotto il nome di Unified Team, generando problemi di confrontabilità con le edizioni successive.

Inoltre, i risultati delle Olimpiadi di Los Angeles 1984 e di Mosca 1980 furono influenzati da importanti boicottaggi, pertanto si è preferito non includere tali edizioni, per evitare di generare distorsioni nelle stime.

Infine, i Giochi precedenti a quelli dei boicottaggi risultano troppo differenti da quelli odierni per il contesto politico, economico, sociale e sportivo nel quale si si sono svolti, per cui si è optato per la loro esclusione.

Le fonti utilizzate per questa ricerca sono principalmente tre: il database del fondo monetario internazionale, la base dati della World Bank e le pagine di Wikipedia relative alle diverse edizioni olimpiche.

3.2 La partecipazione e le performances

In Tab. 3.1 si osserva che nell'arco di sedici anni, dal 2000 al 2016, i Paesi partecipanti alle Olimpiadi sono aumentati del 3%. Più altalenante è il dato che riguarda il numero di atleti partecipanti che, nel periodo considerato, è cresciuto del 4,33%, pur subendo una battuta d'arresto nelle Olimpiadi di Londra 2012. Questi dati sono stati reperiti consultando la base dati presente sul sito del Comitato Olimpico Internazionale.

Si è deciso di includere nell'analisi soltanto le osservazioni relative alle nazioni che hanno partecipato a tutt'e cinque le edizioni olimpiche considerate.¹ Inoltre, sono state espunte le unità statistiche affette da una fortissima presenza di dati missing.²

Così facendo si ottiene un dataset composto da 945 osservazioni.

¹ A questo proposito, risultano esclusi dall'analisi i seguenti paesi: la Serbia, il Montenegro, la Repubblica del Congo, gli Atleti Olimpici Rifugiati, gli Atleti Olimpici Indipendenti, il Kosovo, il Gibuti, l'Afghanistan, le Isole Marshall, la Repubblica Democratica del Congo, la Sierra Leone, il Brunei, il Sudan del Sud, la Repubblica Democratica di Timor Est, il Tuvalu, il Kuwait, la Repubblica delle Kiribati, le Antille Olandesi e la Jugoslavia.

² Si tratta di Taiwan e delle Isole Vergini Britanniche.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di paesi partecipanti	200	201	204	205	206
Numero di atleti partecipanti	10771	10931	11176	10768	11237
Numero medio di atleti per paese	53,86	54,38	54,78	52,53	54,55

Tab. 3.1: I numeri della partecipazione ai Giochi olimpici estivi.

3.3 Le variabili dipendenti

3.3.1 Il numero di medaglie

Coerentemente con quanto asserito in letteratura, il numero di medaglie conquistate da un paese risulta essere la variabile dipendente, la cui variabilità viene determinata da fattori sociali, economici, politici, geografici e scientifici; l'obbiettivo di questa ricerca consiste nell'identificare compiutamente quali siano questi fattori e quantificare la loro influenza.

Le variabili dipendenti considerate sono quattro: il numero di medaglie d'oro, d'argento e di bronzo ed il numero di medaglie totali. I dati sono stati tratti dalle pagine Wikipedia relative alle diverse edizioni delle Olimpiadi.

In linea con la letteratura, il conteggio delle medaglie conquistate può rappresentare una proxy della performance olimpica; quindi, il numero di medaglie vinte da un paese in una determinata edizione dei Giochi olimpici estivi rappresenta la principale variabile dipendente utilizzata nei modelli presentati in questo studio.

In Tab. 3.2 si può osservare come il numero di medaglie messe a disposizione dal Comitato Olimpico Internazionale sia aumentato nel corso del tempo, evidenziando una crescita del 4,96% tra il 2000 ed il 2016. Questo accade perché è aumentato nel tempo il numero di discipline olimpiche e di conseguenza il numero di eventi e di gare. Anche il numero di paesi che hanno conquistato almeno una medaglia è aumentato del 7,5% nel periodo compreso tra i Giochi di Sydney e quelli di Rio, nonostante abbia mostrato un andamento altalenante. Per contro, il tasso dei paesi che conquistano almeno un successo alle Olimpiadi sul totale delle nazioni partecipanti rimane pressoché costante.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di medaglie distribuite	927	927	958	957	973
Numero di paesi partecipanti	200	201	204	205	206
Numero di paesi che hanno vinto almeno una medaglia	80	74	87	86	86
Numero di paesi che hanno vinto almeno una medaglia %	0,40	0,37	0,43	0,42	0,42
Media di medaglie per paese	4,64	4,61	4,70	4,67	4,72
Media di medaglie per partecipante	0,086	0,085	0,086	0,089	0,087

Tab. 3.2: Le performance olimpiche

3.3.2 Il numero di atleti

Questa variabile raccoglie il numero di atleti che ciascun paese invia ad una determinata edizione dei Giochi olimpici estivi. Anche in questo caso per reperire i dati sono state utilizzate le pagine Wikipedia relative alle diverse edizioni delle Olimpiadi.

Il numero di atleti inviati da ciascun paese può essere considerato sia una variabile indipendente sia una variabile dipendente. Infatti, come suggerito dai lavori di Johnson e Ali (2000) e di Kuper e Sterken (2001), la partecipazione è strettamente influenzata dal benessere economico e dalla

dimensione della popolazione; come si può osservare in Tab. 3.3 ed in Tab. 3.4, i paesi più ricchi e popolosi inviano più atleti alle Olimpiadi. Inoltre, anche essere la nazione ospitante, o quantomeno un paese geograficamente limitrofo a quello dove si svolgeranno i Giochi, porta ad un aumento del numero di atleti che prendono parte alla manifestazione. Tuttavia, per quanto riguarda il numero di medaglie vinte, il numero di atleti inviati risulta essere una variabile indipendente, che come affermato da Sterken e Kuper (2001), si dimostra uno dei fattori che determinano maggiormente il successo olimpico.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Percentuale di paesi economicamente sviluppati	17%	23%	30%	30%	29%
Percentuale di atleti provenienti da paesi economicamente sviluppati	42%	51%	55%	57%	52%
Percentuale di paesi in via di sviluppo	14%	15%	20%	26%	24%
Percentuale di atleti provenienti da paesi in via di sviluppo	16%	11%	24%	29%	31%
Percentuale di paesi economicamente non sviluppati	63%	59%	46%	40%	43%
Percentuale di atleti provenienti da paesi economicamente non sviluppati	41%	37%	20%	13%	16%

Tab. 3.3: Percentuale di atleti per livello di benessere economico

Come si evince dalla Tab. 3.3, i paesi economicamente sviluppati, ovvero quelli che presentano un PIL pro capite superiore ai 15000 USD all'anno, rappresentano meno di un terzo del totale dei paesi partecipanti, ma inviano più della metà del totale di atleti. Al contrario, i paesi non sviluppati dal punto di vista economico, ovvero quelli che presentano un reddito pro capite inferiore ai 5000 USD annuali, pur costituendo un po' meno della metà dei paesi partecipanti portano ai Giochi una percentuale di atleti inferiore al 20% del totale. Le percentuali risultano influenzate dall'aumento collettivo dei valori del PIL pro capite nel periodo considerato.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di paesi partecipanti	200	201	204	205	206
Numero di atleti partecipanti	10771	10931	11176	10768	11237
Percentuale dei 25 paesi più popolosi sul totale dei paesi	12,5%	12,4%	12,3%	12,2%	12,1%
Percentuale di atleti provenienti dai 25 paesi più popolosi	40,1%	38,9%	40,1%	39,6%	39,9%

Tab. 3.4: Percentuale di atleti provenienti dai venticinque paesi più popolosi

Come si può osservare dalla Tab. 3.4, i venticinque paesi più popolosi costituiscono circa il 12% del totale dei paesi che prendono parte ai Giochi olimpici estivi, ma inviano dei contingenti molto numerosi, che in aggregato rappresentano il 40% del totale degli atleti partecipanti. Queste percentuali rimangono stabili col passare degli anni, poiché vi sono comunque dei limiti, fissati dal CIO, sul numero di atleti che ciascun paese può iscrivere alle diverse discipline olimpiche.

3.4 Le variabili indipendenti

Come già affermato in precedenza, le variabili indipendenti che sono state prese in considerazione per questa ricerca possono essere raggruppate in sei macro categorie principali: variabili economiche, geografiche, politiche, sociali, scientifiche, e demografiche. Ciascuna categoria ha l'obbiettivo di riuscire a spiegare come l'omonimo effetto influenzi il numero di medaglie conquistate da un paese ai Giochi olimpici estivi.

3.4.1 Le variabili economiche

Le variabili economiche risultano estremamente rilevanti nel determinare il successo olimpico di un paese, in quanto avere a disposizione maggiori risorse economiche significa poter attuare degli investimenti finalizzati a migliorare le performance olimpiche. Maggiori risorse economiche portano ad un aumento dei fondi investiti per ammodernare e potenziare strutture di allenamento, stadi e palestre, ma anche ad un incremento degli investimenti profusi per incentivare l'educazione fisica in ambito scolastico e per arruolare allenatori di livello internazionale. Uno Stato più ricco può permettersi una maggiore spesa per lo sport e può riuscire più facilmente ad identificare, a formare e ad allenare atleti di alto livello.

La prima variabile selezionata è quella riportante il Prodotto Interno Lordo, denominata GDP_{i,t}, i cui valori sono stati raccolti consultando il database della Banca Mondiale. I due ricercatori Bernard e Busse (2000) avevano identificano il PIL come il vero e principale determinante del successo olimpico e, analogamente, l'economista Xun Bian (2005) afferma come il PIL sia un fattore chiave per spiegare il numero di medaglie vinte da un paese. Tuttavia, come spiega Bian, dovrebbe esserci una forte correlazione tra l'effetto del PIL e quello della popolazione, in quanto il PIL rappresenta il prodotto del PIL pro capite per la dimensione della popolazione. Proprio per questo motivo, il Prodotto Interno Lordo si configura solo in parte come una misura dell'effetto della ricchezza sul totale delle medaglie, in quanto annovera al suo interno anche gran parte dell'effetto derivante dalla dimensione della popolazione. Alla luce di quanto suggerito dalla letteratura, è lecito aspettarsi che i paesi che presentano un elevato Prodotto Interno Lordo siano quelli che ottengono il maggior numero di successi ai Giochi olimpici estivi, come testimoniato in Tab. 3.5.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di medaglie distribuite	927	927	958	957	973
Numero di paesi partecipanti	200	201	204	205	206
Percentuale dei 10 paesi con il PIL più elevato	5,0%	5,0%	4,9%	4,9%	4,9%
Numero di medaglie conquistate dai primi 10 paesi per PIL	357	407	495	508	454
Percentuale di medaglie conquistate dai primi 10 paesi per PIL	38,5%	43,9%	51,7%	53,1%	46,7%

Tab. 3.5: Numero di medaglie conquistate dai dieci paesi con il PIL più elevato

Osservando la Tab. 3.5, si nota come i dieci paesi con il PIL più alto, che rappresentano una percentuale del 5% sul totale dei paesi partecipanti, sono soliti conquistare in media circa metà delle medaglie messe in palio alle Olimpiadi estive.

Parimenti, sono state incluse le variabili GDP_{i,t-2} e GDP_{i,t-4}, che riportano i valori del Prodotto Interno Lordo riferiti rispettivamente a due ed a quattro anni prima dell'Olimpiade considerata, poiché, come affermato dai ricercatori Wladimir e Madeleine Andreff (2008), un paese inizia ad investire capitali nella preparazione dell'evento olimpico già due o quattro anni prima del suo svolgimento. Anche il PIL pro capite fa parte delle variabili economiche considerate in questo ricerca, rappresentato dalla covariata GDP per capita_{i,t}, i cui valori sono stati ottenuti, anche in questo caso, dal database della World Bank. Dalla letteratura si evince che il PIL pro capite è un fattore estremamente significativo nel determinare il numero di medaglie conquistate da un paese alle Olimpiadi estive. Infatti, come affermato da Johnson e Ali (2000), un alto PIL pro capite è associato ad una migliore qualità di allenamento e ad una maggiore possibilità di sostenere le spese per mandare gli atleti ai Giochi. Anche il lavoro di Kuper e Sterken (2001) asserisce che il PIL pro capite influenza le condizioni di allenamento e di salute dei cittadini. Alla luce di quanto emerge dalla letteratura, è lecito aspettarsi che le nazioni con un elevato PIL pro capite tendano a conquistare più medaglie ai Giochi olimpici estivi. Quest'ipotesi trova un'ulteriore conferma nella Tab. 3.6, in cui si nota come i venti paesi con il reddito pro capite più elevato, che rappresentano il 10% del totale

dei paesi partecipanti, tendono a vincere in media il 25% delle medaglie messe in palio dal Comitato Olimpico Internazionale. Si nota, inoltre, che questa percentuale mostra un andamento altalenante e che il PIL pro capite risulta essere un fattore meno determinante rispetto al Prodotto Interno Lordo.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di medaglie distribuite	927	927	958	957	973
Numero di paesi partecipanti	200	201	204	205	206
Percentuale dei primi 20 paesi per PIL pro capite	10,0%	10,0%	9,8%	9,8%	9,7%
Numero di medaglie conquistate dai primi 20 paesi per PIL pro capite	210	226	215	253	234
Percentuale di medaglie conquistate dai primi 20 paesi per PIL pro capite	22,7%	24,4%	22,4%	26,4%	24,0%

Tab. 3.6: Numero di medaglie conquistate dai venti paesi con il PIL pro capite più elevato

Sono state prese in considerazione in questa ricerca anche le variabili GDP per capite_{i,t-2} e GDP per capite_{i,t-4}, per motivazioni analoghe a quelle addotte in precedenza in merito alle covariate ritardate riportanti il Prodotto Interno Lordo; tali variabili quantificano il reddito pro capite relativo, rispettivamente, a due e quattro anni prima dei Giochi olimpici considerati.

Si considera, inoltre, il debito pubblico calcolato come percentuale del Prodotto Interno Lordo, riportato dalla variabile Debt (%GDP)_{i,t}, una covariata non ancora presa in considerazione dalla letteratura di settore. I dati sono stati ottenuti attraverso il database del sito web del Fondo Monetario Internazionale. Dalla teoria economica emerge che un debito pubblico elevato obbliga lo Stato a tagliare con decisione la sua spesa pubblica e che uno dei settori maggiormente colpiti da queste politiche restrittive è sicuramente quello dello sport.

Per descrivere la tendenza secondo cui sono necessari dai due ai quattro anni per preparare gli atleti alle gare olimpiche, è opportuno considerare anche le variabili Debt (%GDP)_{i,t-2} e Debt (%GDP)_{i,t-4}, che riportano i valori del debito pubblico in funzione del PIL riferiti a due e quattro anni prima dei Giochi olimpici considerati.

Viene inoltre presa in considerazione la spesa di uno Stato nell'istruzione, espressa sia come percentuale del PIL sia come percentuale del totale della spesa pubblica, introducendo le variabili indipendenti denominate rispettivamente Education expenditure (%GDP)_{i,t} e Education expenditure (%GOV)_{i,t}, due covariate ancora non contemplate dagli studi scientifici di riferimento. I dati relativi a quanto ciascun paese spenda per istruire le future generazioni sono stati desunti dal database della World Bank. Queste variabili permettono di quantificare la quota parte della spesa pubblica dedicata al sistema scolastico, considerando l'intero percorso formativo dalla scuola primaria all'Università. Un paese che investe molto nel sistema scolastico riuscirà ad avere un'istruzione efficiente e ad abbassare i tassi di analfabetismo e di abbandono scolastico. Non erogare alle nuove generazioni una buona istruzione non solo compromette la loro esistenza, ma anche quella delle generazioni future. Inoltre, l'istruzione è un importante indicatore del miglioramento delle condizioni di salute ed economiche dei cittadini. Istruzione e sport si influenzano vicendevolmente, perché, come già asserito da Andreff (2001), il primo contatto con l'attività fisica avviene normalmente nelle palestre degli istituti scolastici. La spesa nell'istruzione appare una misura più significativa se espressa come percentuale del Prodotto Interno Lordo, piuttosto che se venisse espressa come percentuale del totale della spesa pubblica. Infatti, benché un paese spenda una notevole percentuale della propria spesa pubblica nell'istruzione, una spesa pubblica esigua limiterebbe comunque i miglioramenti apportati al sistema scolastico. È da osservare che il paese che investe percentualmente di più nel sistema scolastico è Cuba, che rappresenta un'eccezione per i Giochi olimpici, in quanto vince regolarmente molte più medaglie di quanto ci si potrebbe aspettare basandosi solamente sul suo PIL pro capite o sulla sua popolazione.

Si considera poi la crescita economica espressa come variazione del Prodotto Interno Lordo di un paese, calcolata tra i quattro anni ed i due anni prima dell'inizio delle Olimpiadi. Tale variabile è in grado di riflettere un'eventuale crescita economica oppure una contrazione del PIL nel periodo considerato. I dati raccolti per questa covariata provengono dal database della Banca Mondiale. Un incremento del Prodotto Interno Lordo di un paese porta ad un aumento degli investimenti, che riguarda anche il mondo dello sport; la crescita economica porta ad un incremento non soltanto della spesa per la costruzione ed il mantenimento degli impianti sportivi e delle strutture di allenamento, ma anche dei fondi utilizzati per individuare e formare atleti, allenatori e dirigenti sportivi di alto profilo.

Si osserva che l'inclusione della variabile GDP growth (%)_{i,t} rappresenta una novità per quanto riguarda la ricerca nell'ambito dei Giochi olimpici estivi.

Sono state poi prese in considerazione la spesa sanitaria pro capite, rappresentata dalla variabile indipendente Health expenditure per capitai,t, e la spesa sanitaria espressa come percentuale del Prodotto Interno Lordo, descritta dalla covariata Health expenditure (%GDP)_{i,t}. I dati per entrambe le variabili sono stati raccolti attraverso il database della Banca Mondiale. La decisione di includere queste due variabili trova riscontro nella letteratura: gli studiosi George Vagenas e Eleni Vlachokyriakou (2011) affermano, infatti, che la spesa sanitaria risulta essere un fattore estremamente significativo nel determinare il numero di medaglie vinte da un paese alle Olimpiadi. Analogamente, anche Wladimir Andreff (2001) asserisce come nei paesi in via di sviluppo lo sport risulti influenzato dall'assenza di un sistema sanitario adeguato. Parimenti, l'economista Amartya Sen (1999) dimostra come un'elevata spesa sanitaria non soltanto migliori l'aspettativa di vita e la salute dei cittadini, ma sia anche responsabile di una crescita economica e di un appianamento delle disuguaglianze. La spesa sanitaria pro capite non soltanto descrive il livello di sviluppo del sistema sanitario, ma rappresenta anche una misura del livello di benessere e di ricchezza di un paese decisamente più attendibile del PIL pro capite, soprattutto nel caso dei paesi la cui ricchezza è concentrata nelle mani di poche persone. Inoltre, un sistema sanitario sviluppato può meglio dedicarsi agli atleti, alla loro preparazione ed alla ripresa dai loro infortuni. Dall'osservazione della Tab. 3.7 si nota come i primi venti paesi per spesa sanitaria pro capite tendano a vincere circa il 40% delle medaglie messe in palio alle Olimpiadi; tale percentuale risulta notevolmente più bassa se si individuano i venti paesi di riferimento sulla base della spesa sanitaria espressa come funzione del PIL, poiché, come detto, la spesa sanitaria pro capite riesce a descrivere in modo ottimale il welfare di una nazione.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di medaglie distribuite	927	927	958	957	973
Numero di medaglie conquistate dai primi 20 paesi per spesa sanitaria pro capite	381	372	397	344	448
Percentuale di medaglie conquistate dai primi 20 paesi per spesa sanitaria pro capite	41,1%	40,1%	41,4%	35,9%	46,0%
Numero di medaglie conquistate dai primi 20 paesi per spesa sanitaria % GDP	220	235	284	303	346
Percentuale di medaglie conquistate dai primi 20 paesi per spesa sanitaria % GDP	23,7%	25,4%	29,6%	31,7%	35,6%

Tab. 3.7: Numero di medaglie conquistate dai venti paesi con la spesa sanitaria più elevata

In seguito è stata considerata l'inflazione, rappresentata dalla variabile indipendente Inflation_{i,t}, i cui dati sono stati estratti dal database della Banca Mondiale. L'idea di includere il fenomeno dell'inflazione in questa ricerca trova origine nell'articolo scientifico pubblicato dagli studiosi George Vagenas e Eleni Vlachokyriakou (2011), nel quale essi affermano come l'inflazione sia un fattore molto significativo nel determinare il numero di medaglie vinte da un paese alle Olimpiadi estive. L'inflazione si riflette negativamente sugli investimenti e, di conseguenza, sullo sport, che al pari di tutti gli altri settori economici si trova a fare i conti con quest'effetto, che si traduce in una

diminuzione degli investimenti dedicati agli impianti sportivi ed in una contrazione dei fondi destinati ad individuare e formare atleti e allenatori.

In seguito, sono state considerate anche la spesa militare espressa come percentuale del PIL e la spesa militare espressa come percentuale della spesa pubblica totale di un paese, i cui valori sono riportati rispettivamente dalle variabili Military expenditure (%GDP)_{i,t} e Military expenditure (%GOV)_{i,t}. I dati riguardanti queste due covariate sono stati estratti dal database della Banca Mondiale. Si tratta di due variabili mai testate in letteratura prima d'ora, che sono state incluse nella presente ricerca sulla base della considerazione per cui grandi investimenti in ambito militare possono portare a scoperte tecnologiche e scientifiche che, se applicate in ambito agonistico, possono portare ad un miglioramento delle performance sportive.

Si è deciso di includere all'interno di questo studio anche la spesa in ricerca e sviluppo di una nazione espressa come percentuale del Prodotto Interno Lordo. Essa è rappresentata dalla covariata Research expenditure (%GDP)_{i,t}, i cui dati sono stati ottenuti consultando il database della World Bank. Nonostante questo fenomeno non sia stato mai studiato in letteratura, si è scelto di considerarlo in quanto la spesa in attività di R&S costituisce un volano per la crescita economica del paese e porta ad innovazioni scientifiche e tecnologiche che possono essere applicate anche in ambito sportivo, migliorando le prestazioni olimpiche.

La successiva variabile che si è deciso di considerare in questo lavoro, Total expenditure (%GDP)_{i,t}, rappresenta la spesa pubblica espressa in funzione del Prodotto Interno Lordo. Si è scelto di includere questa variabile, per quanto nessun ricercatore lo abbia fatto prima d'ora, dato che la teoria economica suggerisce che un aumento della spesa pubblica determina un'accelerazione della crescita economica. Inoltre, per incrementare la spesa pubblica sono necessari numerosi investimenti, anche in ambito sportivo, i quali possono portare ad un boost nelle performance olimpiche. Per le stesse ragioni appena elencate è stata aggiunta anche la covariata Total investment (%GDP)_{i,t}, che indica il totale degli investimenti di un paese espresso come percentuale del PIL. I dati per entrambe le variabili sono stati ricavati dal database del Fondo Monetario Internazionale.

Infine, l'ultima variabile economica presa in considerazione è il tasso di disoccupazione, rappresentato dalla covariata Unemployment_{i,t}, i cui valori sono stati estratti dal database dell'Istituto Internazionale del Lavoro. L'idea di includere il tasso di disoccupazione trova origine nel paper pubblicato dagli studiosi George Vagenas e Eleni Vlachokyriakou (2011), nel quale affermano che la percentuale di persone non occupate risulta essere un fattore fortemente significativo nel determinare il numero di medaglie conquistate da un paese alle Olimpiadi estive. La teoria economica suggerisce, inoltre, come un'elevata disoccupazione porti ad una contrazione del Prodotto Interno Lordo, con relative conseguenze sugli investimenti in ambito sportivo.

3.4.2 Le variabili geografiche

Le variabili geografiche hanno l'obbiettivo di cogliere l'effetto sul totale delle medaglie fornito da quei fattori che dipendono dalla posizione geografica del paese.

La prima variabile presa in considerazione, Hosti,t, è una variabile dicotomica, che riporta l'effetto dato dall'essere la nazione che ospita i Giochi olimpici estivi. Fin dai primi anni '70, con le ricerche pubblicate da Levine (1972) e Ball (1972), la letteratura afferma come le performance dei paesi che organizzano le Olimpiadi surclassino quelle degli altri paesi. Anche il ricercatore australiano Clarke (2000) afferma che sui diciassette paesi ospitanti considerati nel suo studio ben quattordici hanno vinto la maggior percentuale di medaglie disponibili proprio nell'edizione in cui erano il paese organizzatore. I motivi che spiegano la presenza dell'host effect sono la familiarità con gli impianti sportivi, di cui possono godere gli atleti di casa, l'abitudine al clima e alle condizioni ambientali ed il fatto che gli sportivi della nazione che ospita i Giochi non devono sottoporsi a lunghi e stancanti

viaggi per raggiungere il luogo dove si svolgeranno le Olimpiadi, un fattore che tenderebbe a peggiorare sensibilmente le performance. Un aspetto determinante del giocare in casa è quello di poter giovare del tifo del pubblico, una componente che, soprattutto in alcuni sport, risulta essere fondamentale. Il supporto del pubblico appare ancora più significativo per quegli sport per i quali il verdetto di un arbitro o dei giudici è determinante. L'host effect oltre ad essere un determinante diretto del successo, lo è anche indirettamente; infatti, come affermato da Johnson e Ali (2000), i paesi che ospitano i Giochi olimpici partecipano alle gare con un contingente di atleti decisamente più nutrito rispetto a quello di cui possono disporre nelle altre edizioni, aumentando sensibilmente le proprie probabilità di vincere una medaglia.

Sono state, altresì, prese in considerazione due variabili dummy, denominate Previous Host_{i,t} e Next Host_{i,t}, che raappresentano rispettivamente l'effetto dovuto all'essere stato il paese ospitante della precedente edizione delle Olimpiadi e l'effetto dovuto ad essere il paese che ospiterà i prossimi Giochi olimpici. Le motivazioni per includere queste variabili sono da ricercare in letteratura; infatti, sia Nevill, Balmer e Winter (2009) sia Andreff e Andreff (2020) affermano che l'host effect perduri nel tempo. Questo accade perché un paese ha a disposizione circa 7 anni dall'attribuzione del ruolo di paese ospitante per prepararsi e per investire nell'ammodernamento degli impianti sportivi e nella formazione degli atleti; poiché questi investimenti hanno come target l'edizione casalinga dei Giochi, appare ovvio che i primi risultati inizino a vedersi già nelle Olimpiadi precedenti a quelle di casa. Come detto, l'effetto di aver ospitato i Giochi olimpici non svanisce subito, ma perdura negli anni, arrivando ad influenzare anche l'edizione successiva, poiché l'effetto degli investimenti sulle infrastrutture e sulla formazione degli atleti dura per almeno quattro anni, secondo quanto asserito dalla letteratura. Dalla Tab. 3.8 si desume che i paesi che ospitano i Giochi olimpici tendono a vincere più medaglie, e che questo vantaggio perdura nel tempo.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Australia	58	50	46	35	29
Grecia	13	16	3	2	6
Cina	58	63	100	91	70
Gran Bretagna	28	30	51	65	67
Brasile	12	10	17	17	19

Tab. 3.8: Numero di medaglie conquistate dai paesi che hanno ospitato le cinque Olimpiadi considerate

Inoltre, sono state prese in considerazioni delle variabili dummy che indicano l'appartenenza di uno Stato ad una determinata regione o area geografica, raccolte sotto il nome di Regions. Si è deciso di considerare queste variabili seguendo quanto fatto da Wladimir e Madeleine Andreff (2008), i quali hanno appunto introdotto delle variabili dicotomiche regionali che risultano essere significative nel determinare il numero di medaglie vinte alle Olimpiadi. Lo scopo di queste covariate è di testare l'influenza culturale dei paesi sulle medaglie vinte. Si presume che la differenza culturale sia legata alle tradizioni riguardanti l'educazione fisica, la pratica sportiva e l'abitudine di assistere a spettacoli sportivi o di guardarli in televisione. Queste tradizioni e questi atteggiamenti nei confronti dello sport sono profondamente radicati nella cultura locale e nelle abitudini di comportamento e di pensiero, che risultano essere tendenzialmente simili nei paesi appartenenti alla stessa regione del mondo. I paesi che fanno parte di questo studio son stati suddivisi in sette macro gruppi, che identificano una specifica regione geografica di appartenenza, rappresentati dalle dummy "Europe and Central Asia", "East Asia and Pacific", "Latin America and The Caribbean", "Middle East and North Africa", "North America", "South Asia" e "Sub-Saharan Africa". Dalla Tab. 3.9 si evince come i paesi europei e dell'Asia Centrale e quelli nordamericani tendano a vincere più medaglie e

rappresentino le regioni con la più alta percentuale di paesi che vincono almeno una medaglia ai Giochi olimpici estivi.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Europe and Central Asia - numero di paesi	50	50	50	50	50
Europe and Central Asia - numero di medaglie conquistate	515	493	449	456	480
Europe and Central Asia - numero medio di medaglie per paese	10,3	9,9	9,0	9,1	9,6
Europe and Central Asia - percentuale di paesi che ha vinto almeno una medaglia	80%	68%	80%	74%	72%
East Asia and Pacific - numero di paesi	28	28	28	28	28
East Asia and Pacific - numero di medaglie conquistate	180	204	238	230	207
East Asia and Pacific - numero medio di medaglie per paese	6,4	7,3	8,5	8,2	7,4
East Asia and Pacific - percentuale di paesi che ha vinto almeno una medaglia	32%	36%	46%	43%	50%
Latin America and The Caribbean - numero di paesi	37	37	37	37	37
Latin America and The Caribbean - numero di medaglie conquistate	71	64	81	76	67
Latin America and The Caribbean - numero medio di medaglie per paese	1,9	1,7	2,2	2,1	1,8
Latin America and The Caribbean - percentuale di paesi che ha vinto almeno una medaglia	32%	32%	35%	35%	32%
Middle East and North Africa - numero di paesi	18	18	18	18	18
Middle East and North Africa - numero di medaglie conquistate	18	18	10	26	24
Middle East and North Africa - numero medio di medaglie per paese	1,0	1,0	0,6	1,4	1,3
Middle East and North Africa - percentuale di paesi che ha vinto almeno una medaglia	33%	33%	33%	44%	56%
North America - numero di paesi	3	3	3	3	3
North America - numero di medaglie conquistate	107	113	132	122	143
North America - numero medio di medaglie per paese	35,7	37,7	44,0	40,7	47,7
North America - percentuale di paesi che ha vinto almeno una medaglia	67%	67%	67%	67%	67%
South Asia - numero di paesi	7	7	7	7	7
South Asia - numero di medaglie conquistate	1	1	3	6	2
South Asia - numero medio di medaglie per paese	0,1	0,1	0,4	0,9	0,3
South Asia - percentuale di paesi che ha vinto almeno una medaglia	14%	14%	14%	14%	14%
Sub-Saharan Africa - numero di paesi	43	43	43	43	43
Sub-Saharan Africa - numero di medaglie conquistate	28	26	38	31	36
Sub-Saharan Africa - numero medio di medaglie per paese	0,7	0,6	0,9	0,7	0,8
Sub-Saharan Africa - percentuale di paesi che ha vinto almeno una medaglia	16%	16%	21%	16%	16%

Tab. 3.9: La ripartizione delle medaglie in base alle regioni di appartenenza

La successiva variabile che si è deciso di includere all'interno di questa ricerca è Temperature_{i,t}, che indica la temperatura media di un paese, i cui dati sono stati raccolti su Wikipedia. Si è scelto di considerare questa variabile sulla base della letteratura. Gli studiosi Johnson e Ali (2002) affermano che le nazioni che hanno un clima più freddo hanno performance superiori a quelle dei loro concorrenti, sia nei Giochi olimpici estivi che in quelli invernali. Parimenti, i ricercatori Robert Hoffmann, Lee Chew Ging e Bala Ramasamy (2002) asseriscono come la temperatura ottimale per il successo olimpico risulti essere di circa 15° Celsius.

3.4.3 Le variabili politiche

Le variabili politiche hanno l'obbiettivo di cogliere quale sia l'effetto della struttura economica e politica del paese considerato sul numero di medaglie conquistate alle Olimpiadi estive.

Per ottenere questo risultato, si è deciso di creare, prendendo spunto da quanto fatto da Wladimir e Madeleine Andreff (2008), cinque variabili dummy che suddividono i paesi presi in considerazione in questa ricerca in cinque gruppi, sulla base della struttura politica ed economica. Un primo gruppo, denominato CEEC, riunisce i paesi dell'Europa centrale e orientale che hanno vissuto un regime sovietico e un'economia pianificata centralmente fino al 1989 o 1990 e che da allora sono stati trasformati in regimi democratici, con economie di mercato. Un secondo gruppo, denominato TRANS, riunisce i paesi contraddistinti da un regime comunista e da un'impostazione economica di tipo sovietico, che hanno poi intrapreso un processo di transizione verso la democrazia e l'economia di mercato nel 1992, secondo un percorso che ha comportato molti meno progressi rispetto a quelli ottenuti dai paesi appartenenti al CEEC. È da osservare come questi paesi non fossero membri dell'UE nel 2008 e non abbiano intenzione di aderirvi nel prossimo futuro. Il terzo gruppo, NSCOM, comprende i paesi che, a differenza dei due gruppi precedenti, non erano contraddistinti da regimi sovietici, pur avendo un'impostazione di tipo comunista ed un'economia pianificata centralmente,

e che sono entrati in una transizione verso la democrazia e un'economia di mercato intorno al 1990. Vi sono due paesi che non possono essere aggregati in uno dei tre gruppi precedenti: Corea del Nord e Cuba, che sono ancora oggi dei regimi comunisti con un'economia pianificata; questi due paesi vengono, quindi, inclusi un quarto gruppo, denominato COM. Tutti gli altri paesi fanno parte di un'ultima classe di paesi, CAPME, che include le economie di mercato capitaliste.

Osservando la Tab. 3.10 si nota che i paesi tutt'oggi comunisti e quelli che hanno risentito dell'influenza sovietica sono soliti conseguire un numero di medaglie per paese più elevato e sono caratterizzati da una percentuale di paesi che conquistano almeno una medaglia superiore rispetto ai paesi che adottano un'economica di libero mercato.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
CAPME - numero di paesi	157	157	157	157	157
CAPME - numero di medaglie conquistate	634	644	708	719	734
CAPME - numero medio di medaglie per paese	4,0	4,1	4,5	4,6	4,7
CAPME - percentuale di paesi che ha vinto almeno una medaglia	33%	32%	36%	36%	38%
CEEC - numero di paesi	10	10	10	10	10
CEEC - numero di medaglie conquistate	96	87	63	67	56
CEEC - numero medio di medaglie per paese	9,6	8,7	6,3	6,7	5,6
CEEC - percentuale di paesi che ha vinto almeno una medaglia	100%	100%	100%	100%	90%
COM - numero di paesi	2	2	2	2	2
COM - numero di medaglie conquistate	33	32	36	21	18
COM - numero medio di medaglie per paese	16,5	16,0	18,0	10,5	9,0
COM - percentuale di paesi che ha vinto almeno una medaglia	100%	100%	100%	100%	100%
NSCOM - numero di paesi	6	6	6	6	6
NSCOM - numero di medaglie conquistate	4	8	6	6	10
NSCOM - numero medio di medaglie per paese	0,7	1,3	1,0	1,0	1,7
NSCOM - percentuale di paesi che ha vinto almeno una medaglia	50%	33%	33%	17%	17%
TRANS - numero di paesi	14	14	14	14	14
TRANS - numero di medaglie conquistate	154	148	138	134	141
TRANS - numero medio di medaglie per paese	11,0	10,6	9,9	9,6	10,1
TRANS - percentuale di paesi che ha vinto almeno una medaglia	79%	57%	93%	71%	79%

Tab. 3.10: La ripartizione delle medaglie in base alla struttura politica ed economica

3.4.4 Le variabili sociali

La prima variabile sociale che è stata presa in considerazione è il numero di anni di scuola dell'obbligo, rappresentato dalla covariata Compulsory years of education_{i,t}, i cui dati sono stati ottenuti dal database della World Bank. Far leva sul numero di anni della scuola dell'obbligo è una politica cruciale per un paese, finalizzata a migliorare la qualità dell'istruzione, a minimizzare gli studenti che smettono di frequentare la scuola per motivi economici o familiari ed a bilanciare le differenze economiche tra i cittadini. Nonostante questa variabile non sia mai stata testata in letteratura, il ricercatore Wladimir Andreff (2001) asserisce che i bambini entrano per la prima volta in contatto con lo sport a scuola attraverso le lezioni di educazione fisica e che se non frequentano la scuola ben difficilmente hanno l'occasione per iniziare a praticare sport; in quest'ottica, tale aspetto influisce sul numero di atleti di alto livello che viene prodotto da una nazione.

Per queste stesse ragioni si è deciso di prendere in considerazione le due variabili % children out of school_{i,t} e Enrollment rate at primary_{i,t}, le quali indicano, rispettivamente, la percentuale di bambini in età scolare che non risultano iscritti né ad una scuola primaria né ad una scuola secondaria e il tasso netto d'iscrizione a scuola, calcolato come il numero di alunni che risultano iscritti ad una scuola primaria sul totale dei bambini che dovrebbero per età avere accesso all'istruzione di primo grado. Per entrambe le variabili i dati sono stati estratti dal database della Banca Mondiale.

Analogamente, sono stati inclusi in questo studio anche il tasso di completamento della scuola primaria ed il tasso di avanzamento alla scuola secondaria, rappresentati rispettivamente dalle covariate Primary completion rate_{i,t} e Progress to secondary_{i,t}. I dati in questione sono stati raccolti attraverso il database della Banca Mondiale. Queste due variabili vanno considerate insieme ed

appaiono piuttosto rilevanti per determinare la qualità del sistema scolastico di un paese. Come già detto, avere un sistema scolastico in grado di erogare un'istruzione di qualità può influenzare il numero di medaglie vinte ai Giochi olimpici.

Le successive variabili sociali introdotte riportano la principale religione professata all'interno di un paese. A questo proposito, si è deciso di suddividere le correnti religiose in sette gruppi principali e di assegnare oggi paese oggetto di questa ricerca ad uno dei gruppi, in base alla religione prevalentemente praticata. Le informazioni circa le religioni ed i paesi in cui vengono praticate sono state ottenute consultando il sito web ourworldindata.org. Le sette dummy create sono "Buddhism", "Christianity", "Folk Religions", "Hinduism", "Islam", "Judaism" e "Unaffiliated Religions", dove quest'ultima include atei e agnostici. L'intuizione di utilizzare la corrente religiosa per predire il numero di medaglie conquistato da un paese ai Giochi olimpici estivi è da ricercarsi nella letteratura di settore; lo studio dei ricercatori Julia Bredtmann, Carsten J. Crede e Sebastian Otten (2016) rivela come la variabile Muslim_{it}, una dummy che rappresenta i paesi con una popolazione a maggioranza musulmana, risulti essere estremamente significativa nel descrivere il successo olimpico, in quanto i paesi musulmani tendono a inviare meno atlete alle Olimpiadi e di conseguenza a conquistare meno medaglie. Dall'osservazione della Tab. 3.11 si nota come i paesi cristiani e quelli atei siano soliti avere un alto numero di medaglie per paese.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Buddhism - numero di paesi	8	8	8	8	8
Buddhism - numero di medaglie conquistate	4	9	11	11	9
Buddhism - numero medio di medaglie per paese	0,5	1,1	1,4	1,4	1,1
Buddhism - percentuale di paesi che ha vinto almeno una medaglia	25%	25%	38%	38%	38%
Christianity - numero di paesi	127	127	127	127	127
Christianity - numero di medaglie conquistate	752	710	716	693	706
Christianity - numero medio di medaglie per paese	5,9	5,6	5,6	5,5	5,6
Christianity - percentuale di paesi che ha vinto almeno una medaglia	44%	40%	45%	43%	43%
Folk Religions - numero di paesi	1	1	1	1	1
Folk Religions - numero di medaglie conquistate	1	0	1	0	2
Folk Religions - numero medio di medaglie per paese	1,0	0,0	1,0	0,0	2,0
Folk Religions - percentuale di paesi che ha vinto almeno una medaglia	100%	0%	100%	0%	100%
Hinduism - numero di paesi	3	3	3	3	3
Hinduism - numero di medaglie conquistate	1	1	4	6	2
Hinduism - numero medio di medaglie per paese	0,3	0,3	1,3	2,0	0,7
Hinduism - percentuale di paesi che ha vinto almeno una medaglia	33%	33%	67%	33%	33%
Islam - numero di paesi	40	40	40	40	40
Islam - numero di medaglie conquistate	43	49	46	58	88
Islam - numero medio di medaglie per paese	1,1	1,2	1,2	1,5	2,2
Islam - percentuale di paesi che ha vinto almeno una medaglia	28%	25%	35%	38%	40%
Judaism - numero di paesi	1	1	1	1	1
Judaism - numero di medaglie conquistate	1	2	1	0	2
Judaism - numero medio di medaglie per paese	1,0	2,0	1,0	0,0	2,0
Judaism - percentuale di paesi che ha vinto almeno una medaglia	100%	100%	100%	0%	100%
Unaffiliated Religions - numero di paesi	7	7	7	7	7
Unaffiliated Religions - numero di medaglie conquistate	119	148	172	179	150
Unaffiliated Religions - numero medio di medaglie per paese	17,0	21,1	24,6	25,6	21,4
Unaffiliated Religions - percentuale di paesi che ha vinto almeno una medaglia	86%	100%	86%	100%	86%

Tab. 3.11: La ripartizione delle medaglie in base alla religione professata

Si è poi deciso di prendere in considerazione la percentuale di cittadini che usa Internet, rappresentata dalla variabile Use of internet_{i,t}. Questa variabile riporta la percentuale del numero di utenti sul totale della popolazione che hanno utilizzato Internet, da qualsiasi luogo, almeno una volta negli ultimi 3 mesi. I dati per questa variabile sono stati ottenuti consultando il database della Banca Mondiale. L'utilizzo massiccio delle tecnologie dell'informazione e della comunicazione, denominate ICT, offre un'opportunità di crescita economica e comporta un miglioramento della salute e progressi sociali e culturali.

È stato quindi introdotto il Gender Parity Index: si tratta di un indicatore creato appositamente, che indica quanto i paesi possano essere percepiti vicini o lontani dal configurarsi come una società in cui vige la parità di genere. La scelta di includere questa variabile è da ricercarsi nella letteratura di settore; i ricercatori Jennifer L. Berdahl, Eric Luis Uhlmann e Feng Bai (2014) affermano che una maggiore parità di genere aumenta la redditività delle aziende, l'impatto scientifico delle collaborazioni intellettuali, la crescita economica di intere nazioni e anche il numero di successi nelle gare olimpiche, non soltanto per gli atleti di sesso femminile, come è facilmente intuibile, ma persino per gli atleti di sesso maschile. Anche in questo caso, le informazioni per costruire questo indicatore sono state ottenute consultando il database della World Bank.

Infine, si è deciso di includere in questa ricerca anche l'aspettativa di vita, rappresentata dalla variabile Life expectancy_{i,t}, i cui dai sono stati raccolti attraverso il database della Banca Mondiale. Tra i numerosi fattori che influenzano l'aspettativa di vita, quelli più rilevanti risultano i fattori economici, come affermano anche i ricercatori Sebastián Lozano, Gabriel Villa, Flor Maria Guerrero e Pablo Cortés (2002). Essi asseriscono, inoltre, che l'aspettativa di vita risulta essere estremamente significativa nel determinare il numero di medaglie conquistate da un paese ai Giochi olimpici estivi.

3.4.5 Le variabili scientifiche

La prima variabile scientifica che si è deciso di prendere in considerazione in questa ricerca è il numero di medici ogni 1000 abitanti, rappresentato dalla covariata Physicians_{i,t}, i cui dati sono stati estratti dal database della World Bank. Questa variabile rappresenta un indicatore di quanto sia efficiente il sistema sanitario di un paese. Maggiore è il numero di medici per abitante, migliore sarà l'assistenza medica garantita ai cittadini. Un paese che ha un elevato numero di medici per 1000 abitanti sarà in grado di dedicarsi non solo alla cura dei cittadini, ma anche al supporto medico, scientifico e psicologico degli atleti e degli allenatori che dovranno poi competere ai Giochi olimpici. Si è scelto di includere in questo studio la variabile Researchers for population_{i,t}, che indica il numero di ricercatori impegnati in attività di ricerca e sviluppo, espressi per milione di abitanti. Anche per questa variabile i valori sono stati ottenuti consultando il database della World Bank. Nonostante questo fenomeno non sia stato mai studiato in letteratura, si è scelto di considerarlo in quanto le attività di R&S costituiscono un adeguato indicatore per la crescita economica del paese, in quanto portano ad innovazioni scientifiche e tecnologiche, che trovano applicazione anche in ambito sportivo, comportando un miglioramento per le prestazioni olimpiche.

Infine, ho considerato la variabile Scientific articles_{i,t}, che riporta il numero di articoli scientifici e ingegneristici pubblicati in un paese in un anno; tali dati sono stati raccolti consultando il database della World Bank. Questa variabile è già stata presa in considerazione in letteratura: i due scienziati turchi Süleyman Şahin e Fatih Şenduran (2020) affermano l'esistenza di una relazione tra il livello di sviluppo della medicina sportiva di un paese ed i suoi successi ai Giochi Olimpici. Infatti, le scoperte effettuate dalla ricerca scientifica spostano i limiti fisici degli atleti sempre più in alto, grazie ad allenamenti mirati ed alla ricerca medica e, per l'appunto, i paesi utilizzano le scoperte provenienti dal mondo accademico ed universitario per cercare di conquistare un maggior numero di successi alle Olimpiadi. Un paese che pubblica un elevato numero di articoli scientifici è un paese che investe molto nelle attività di ricerca e sviluppo, con un conseguente beneficio per le performance ai Giochi olimpici. Questa variabile risulta influenzata, inoltre, sia dalla ricchezza sia dalla dimensione della popolazione di un paese, di cui dovrebbe riflettere l'effetto sul totale delle medaglie.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di medaglie distribuite	927	927	958	957	973
Numero di medaglie conquistate dai primi 20 paesi per numero di medici	256	266	212	231	257
Percentuale di medaglie conquistate dai primi 20 paesi per numero di medici	27,6%	28,7%	22,1%	24,1%	26,4%
Numero di medaglie conquistate dai primi 20 paesi per numero di ricercatori	488	492	416	398	421
Percentuale di medaglie conquistate dai primi 20 paesi per numero di ricercatori	52,6%	53,1%	43,4%	41,6%	43,3%
Numero di medaglie conquistate dai primi 20 paesi per numero di articoli scientifici	604	616	642	656	641
Percentuale di medaglie conquistate dai primi 20 paesi per numero di articoli scientifici	65,2%	66,5%	67,0%	68,5%	65,9%

Tab. 3.12: La ripartizione delle medaglie in base alle variabili scientifiche

Dalla Tab. 3.12 si deduce che i primi venti paesi per numero di articoli scientifici pubblicati conquistano in media circa il 65% delle medaglie disponibili, dimostrando che questa variabile, delle tre scientifiche, è quella maggiormente in grado di descrivere il successo olimpico.

3.4.6 Le variabili demografiche

La prima variabile demografica che si è deciso di considerare è la dimensione della popolazione, rappresentata dalla covariata Population_{i,t}, i cui valori sono stati ottenuti dal database della Banca Mondiale. Secondo la letteratura, la popolazione, insieme con il reddito, è il fattore che maggiormente determina la vittoria delle medaglie alle Olimpiadi. Johnson e Ali (2000) affermano che in una popolazione numerosa i costi di allenamento, i costi di mantenimento degli impianti sportivi e gli investimenti nello sport sono redistribuiti su un maggior numero di persone; quando i governi decidono di affrontare queste spese e di investire, sono i cittadini, in qualità di contribuenti, che copriranno l'investimento. Inoltre, immaginando che il talento sia ugualmente distribuito, avere una popolazione più ampia significa avere un bacino maggiore dal quale estrarre possibili campioni e, di conseguenza, una maggior probabilità di vincere delle medaglie. Hoffmann, Ging e Ramasamy (2002) affermano, addirittura, che l'effetto della dimensione della popolazione sul totale delle medaglie è superiore a quello del reddito.

È stata anche inserita la variabile Population_{i,t-4}, che riporta i valori della dimensione della popolazione riferiti a quattro anni prima dell'Olimpiade considerata, poiché, come detto, per un paese sono necessari dai due ai quattro anni per preparare adeguatamente la propria partecipazione alla manifestazione.

La successiva variabile che è stata introdotta in questa ricerca è il numero di neonati morti su base annua, rappresentata dalla covariata Neonathal deaths_{i,t}. A questo proposito, si definisce morte neonatale quando un bambino muore durante i primi 28 giorni di vita. I dati per questa variabile sono stati ottenuti attraverso il database della Banca Mondiale. Il numero di morti neonatali riflette la dimensione della popolazione, con cui risulta strettamente correlata, come conferma il coefficiente di determinazione pari a 0,73. I paesi più popolati e con un sistema sanitario meno sviluppato sono quindi quelli che presentano un maggior numero di neonati morti.

Olimpiade	Sydney 2000	Atene 2004	Pechino 2008	Londra 2012	Rio 2016
Numero di medaglie distribuite	927	927	958	957	973
Numero di medaglie conquistate dai primi 20 paesi per dimensione della popolazione	366	400	401	411	397
Percentuale di medaglie conquistate dai primi 20 paesi per dimensione della popolazione	39,5%	43,1%	41,9%	42,9%	40,8%
Numero di medaglie conquistate dai primi 20 paesi per numero di morti neonatali	105	101	162	143	123
Percentuale di medaglie conquistate dai primi 20 paesi per numero di morti neonatali	11,3%	10,9%	16,9%	14,9%	12,6%

Tab. 3.13: La ripartizione delle medaglie in base alle variabili demografiche

Dalla Tab. 3.13 si osserva che i primi venti paesi per dimensione della popolazione vincono circa il 40% delle medaglie messe in palio dal Comitato Olimpico Internazionale, mentre i primi venti paesi per numero di morti neonatali vincono in media il 10% delle medaglie a disposizione.

Si è deciso di includere altre due variabili demografiche, ovvero % Popolazione 15-34 anni_{i,t} e % Popolazione 15-34 anni_{i,t-4}, che riportano rispettivamente la percentuale della popolazione di età compresa tra i 15 ed i 34 anni e la stessa percentuale riferita a quattro anni prima dell'Olimpiade considerata. Poiché nell'evento sportivo considerato, ovvero i Giochi olimpici estivi, competono principalmente atleti che sono compresi tra i 15 ed i 34 anni di età, per avere dei potenziali campioni che possano trionfare alle Olimpiadi è necessaria un'elevata percentuale di popolazione compresa in questa fascia di età; pertanto, si è scelto di includere queste due covariate, i cui dati sono stati ottenuti dal database della World Bank.

La successiva variabile che si è deciso di utilizzare nella presente ricerca è la densità di popolazione, rappresentata dalla covariata Population density_{i,t}, i cui valori sono stati raccolti attraverso la consultazione del database della Banca Mondiale. La densità di popolazione rappresenta una proxy del livello di urbanizzazione di un paese. La scelta di analizzare questo fenomeno è da ricondursi agli studiosi George Vagenas e Eleni Vlachokyriakou (2011), che hanno mostrato come il livello di urbanizzazione di un paese influisca sul successo olimpico.

L'ultima variabile demografica presa in considerazione è la crescita della popolazione avvenuta nel periodo compreso tra i venti ed i quindici anni prima dell'Olimpiade considerata, rappresentata dalla covariata Population growth_{i,t}. Si è scelto di includere questa variabile perché gli atleti che si presentano per la prima volta ad un'Olimpiade sono solitamente ragazzi nati tra i 15 ed i 20 anni prima dell'evento considerato; pertanto, un boom demografico dovrebbe portare 20 anni dopo ad un boom di successi olimpici. I dati per questa variabile sono stati anch'essi ottenuti dal database della Banca Mondiale.

4 I fattori che determinano le medaglie alle Olimpiadi dal 2000 al 2016: quattro stime

Dopo aver presentato le variabili dipendenti ed indipendenti su cui si basa la mia ricerca, in questo capitolo vengono sviluppati quattro modelli di regressione per ottenere altrettante funzioni lineari che permettano di comprendere quali siano i fenomeni che influiscono sul numero di medaglie vinte da un paese ai Giochi olimpici estivi, a fini descrittivi e predittivi.

4.1 L'importanza della struttura politico-economica

4.1.1 Il modello statistico

In linea con la letteratura, la variabile dipendente della quale si vogliono stimare le determinanti è la variabile $Med_{i,t}$, cioè il numero di medaglie vinte dal paese i (i = 1,...,189), ai Giochi olimpici t, (t = 2000, 2004, 2008, 2012 e 2016). La funzione di regressione costruita è la seguente:

$$\begin{split} \textit{Med}_{i,t} &= c + \alpha \textit{GDP}_{i,t} + \beta \textit{Population}_{i,t} + \gamma \textit{Host}_{i,t} + \delta \textit{Next Host}_{i,t} + \varphi \textit{Previous Host}_{i,t} \\ &+ \sum_{p} \theta_{p} \textit{Politics}_{p,i} + \lambda \textit{Health expenditure per capita}_{i,t} \\ &+ \omega \textit{Population growth}_{i,t} + \psi \textit{Neonathal deaths}_{i,t} + \varepsilon_{i,t} \end{split}$$

Si tratta di un modello panel OLS, che prende spunto da quello di Bernard e Busse (2000) e da quello di Andreff e Andreff (2008), esaminati in precedenza; è da osservare, però, che rispetto al primo il numero di medaglie non viene calcolato in percentuale, bensì in valore assoluto. Inoltre, rispetto a quanto presente in letteratura, dove utilizzano il PIL pro capite, in questa funzione si è deciso di includere il Prodotto Interno Lordo.

Ho costruito questo modello partendo dai regressori più tradizionali contemplati in letteratura, ovvero GDP_{i,t}, Population_{i,t} e Host_{i,t}. A questi ho aggiunto le due dummy Next Host_{i,t} e Previous Host_{i,t} per verificare se il vantaggio dato dall'essere la nazione ospitante dei Giochi olimpici estivi si traduca in un effetto temporale capace di influenzare le prestazioni anche nelle edizioni immediatamente precedente ed immediatamente successiva a quella di volta in volta considerata. Ho inserito le dummy che indicano la struttura politico-economica assunta dal paese per verificare se ancora oggi esista un vantaggio per i paesi che un tempo risentirono dell'influenza sovietica o che sono tutt'ora comunisti.

La covariata Health expenditure per capita_{i,t} è stata inserita per più motivi: in primo luogo, permette di cogliere adeguatamente l'effetto del livello di ricchezza diffuso sulle medaglie. In secondo luogo, ho preferito questo regressore al PIL pro capite, poiché quest'ultimo risulta non significativo, in contrasto con quanto asserito in gran parte della letteratura. Queste due variabili presentano una forte correlazione³, come testimonia un coefficiente di correlazione lineare pari a 0,83, pertanto la spesa sanitaria riesce a catturare adeguatamente l'effetto che veniva prima colto dal reddito. Inoltre, la spesa sanitaria mi permette di apprezzare l'effetto che il livello di sviluppo del sistema medico-sanitario esercita sul totale delle medaglie.

Infine, sono state incluse nel modello le covariate Population growth_{i,t} e Neonathal deaths_{i,t}, che permettono di definire in modo ancora più preciso l'impatto della ricchezza e della dimensione della

-

³ Per questo e per i successivi valori dei coefficienti di correlazione si veda l'Appendice I.

popolazione sul totale delle medaglie vinte, che secondo la letteratura rappresentano due degli effetti dominanti.

I risultati della regressione effettuata su questa funzione lineare sono visibili in Tab 4.1.

Dall'output della stima di regressione emerge come tutte le variabili esplicative considerate risultino essere significative ad un livello di significatività dello 0,001, ad eccezione del regressore Population growth_{i,t} e della dummy NSCOM, in quanto la prima si dimostra significativa soltanto ad un livello dello 0,01, mentre la seconda risulta essere completamente non significativa, persino ad un livello dello 0,1.

La variabile GDP_{i,t} presenta un coefficiente positivo ed un effetto marginale tale che è necessario un incremento di circa 158 miliardi di USD per conquistare una medaglia in più ai Giochi olimpici estivi. Questo spiegherebbe perché nazioni che presentano un Prodotto Interno Lordo estremamente basso, o comunque inferiore a questa cifra, raramente riescono ad aggiudicarsi delle medaglie. Le nazioni povere o comunque molto piccole non dispongono di risorse sufficienti a supportare adeguatamente le performance dei propri atleti.

La variabile che riporta la dimensione della popolazione risulta anch'essa significativa, confermando quindi quanto affermato dalla letteratura, che ritiene il numero di abitanti uno dei fattori più determinanti, se non il più determinante. Tuttavia, secondo questo modello è necessario un considerevole aumento della dimensione della popolazione per ottenere un effetto tangibile sul numero di medaglie conquistate alle Olimpiadi. Infatti, serve una crescita della popolazione pari a circa 45 milioni di abitanti per incrementare il bottino olimpico di una medaglia. Si tratta di una cifra estremamente grande, ed in contrasto con quanto emerge dagli studi precedenti. Tuttavia, quest'effetto marginale così estremo potrebbe essere causato dal fatto che in ben 549 osservazioni, sulle 945 totali contenute nel dataset, il numero di medaglie conquistate risulta essere pari a zero. Vi è infatti un'enorme moltitudine di paesi che terminano i Giochi senza alcun successo e l'inclusione di quest'ultimi nella regressione altera, come si vede, l'effetto marginale. A tal proposito, è da osservare che in questa ricerca si utilizza un modello panel OLS, mentre nella maggior parte dei lavori scientifici pubblicati, per evitare questo problema, si tende ad impiegare modelli Tobit.

La regressione conferma che essere la nazione che ospita le Olimpiadi è ancora un fattore determinante, infatti consente al paese che organizza i Giochi di vincere in media circa 30,52 medaglie aggiuntive. Ospitare le Olimpiadi permette alla nazione padrona di casa di inviare un contingente di atleti ben più nutrito, di contare sul supporto del pubblico e su valutazione benevole da parte dei giudici di gara negli sport in cui l'attribuzione delle medaglie passa attraverso il voto di una giuria. Questo risultato appare in contrasto con alcuni lavori scientifici che mostrano, invece, che il vantaggio determinato dall'host effect sia sempre meno impattante col trascorrere delle edizioni.

Le due dummy che catturano l'estensione temporale del vantaggio dato dall'essere la nazione ospitante si dimostrano estremamente significative, a conferma di quanto veniva asserito sia nell'ultimo lavoro dei fratelli Andreff (2020) sia in quello di Nevill, Balmer e Winter (2012). Appare quindi sensato includere nella funzione le variabili Next Host_{i,t} e Previous Host_{i,t}, come asserito nelle più recenti ricerche in materia. Aver ospitato l'edizione precedente dei Giochi olimpici estivi garantisce alla nazione in questione di aggiudicarsi un extra di circa 21,76 medaglie nell'edizione che segue. Parimenti essere il paese che ospiterà le prossime Olimpiadi permette di vincere in media 12,99 medaglie aggiuntive nell'edizione corrente. Risulta evidente che l'impatto sul numero di medaglie del regressore Previous Host_{i,t} è nettamente superiore a quello della dummy Next Host_{i,t}; questa tendenza è dovuta al fatto che quando un paese inizia a prepararsi per ospitare i Giochi olimpici, circa quattro anni prima dell'evento, gli effetti degli investimenti effettuati si apprezzano sin da subito, per poi raggiungere il loro apice nell'edizione casalinga ed infine protrarsi, seppur con una flessione, fino all'edizione ancora successiva.

Dall'output della stima di regressione emerge, inoltre, che le variabili che riportano la struttura politica ed economica dei paesi oggetto dello studio si rivelano, nel complesso, significative. Per valutare l'effetto sulle medaglie di queste covariate sono state inserite nel modello di regressione solamente le dummy CEEC, COM, NSCOM e TRANS, mentre è stata esclusa la variabile CAPME, che rappresenta i paesi che si sono sempre basati su un sistema politico-economico di libero mercato, la quale è stata utilizzata come variabile di riferimento; in altre parole, i coefficienti delle covariate socialiste presenti nel modello sono calcolati proprio rispetto alla dummy CAPME.

I paesi che appartengono al gruppo CEEC, ovvero quelle nazioni che erano sotto la sfera d'influenza sovietica, sono soliti vincere alle Olimpiadi estive, in media, 3,82 medaglie in più rispetto ai paesi che si basano su un'economia di mercato. Quest'effetto risulta ancora più marcato per la dummy COM, i cui paesi che ne fanno parte, ovvero Corea del Nord e Cuba, vincono molte più medaglie, esattamente 19,77, di quanto non farebbero se adottassero al posto di un sistema comunista un modello di tipo capitalistico, ceteris paribus. Anche la variabile TRANS risulta essere significativa, ed i paesi che ne fanno parte sono soliti conquistare un extra di circa 7,50 medaglie ai Giochi olimpici estivi rispetto ai concorrenti che utilizzano un sistema economico capitalistico. Per contro, la dummy NSCOM, che rappresenta i paesi che pur non essendo sotto la sfera di influenza sovietica adottavano comunque un regime economico e politico di stampo comunista, risulta, come detto, non significativa, e presenta un coefficiente pari a -0,75.

Questi risultati mostrano che, nonostante siano passati molti anni dalla caduta del muro di Berlino e dallo scioglimento dell'Unione Sovietica, l'effetto sulle medaglie dato dall'avere una struttura politico-economica di tipo socialista è ancora presente, significativo e molto marcato. Tutto ciò ad eccezione della dummy NSCOM, che raccoglie al suo interno solo nazioni che sono ormai da tempo a stampo capitalistico, motivo per cui tale variabile non mostra alcuna differenza significativa con la variabile di riferimento; infatti, i paesi che fanno parte delle due dummy hanno praticamente la stessa struttura politico-economica. Ancora oggi, quindi, i paesi che hanno subito negli anni l'influenza sovietica continuano a manifestare un vantaggio in termini di medaglie rispetto a paesi che presentano analoghe caratteristiche sociali e demografiche, ma che si basano su un'economia di libero mercato. Allo stesso modo, i paesi che sono tutt'oggi comunisti vincono molte più medaglie di quanto si potrebbe prevedere prendendo in considerazione i determinanti economici e sociali. Questi risultati confermano, quindi, quanto asserito dalla letteratura: essi sono, infatti, in linea con lo studio di Johnson e Ali del 2000, secondo cui i paesi comunisti e socialisti erano soliti vincere in media circa 18 medaglie aggiuntive ai Giochi olimpici estivi.

La variabile che riporta la spesa sanitaria pro capite risulta essere significativa, confermando quando asserito dagli studiosi George Vagenas e Eleni Vlachokyriakou (2011). Inoltre, analizzando l'output statistico della regressione si osserva che è necessario un incremento della spesa sanitaria pro capite pari a 1226,39 USD per permettere al paese in questione di vincere una medaglia in più alle Olimpiadi estive. Questo risultato evidenzia come la covariata Health expenditure per capitai,t rappresenti due effetti appartenenti a due sfere diverse, ma fortemente legate tra loro: ovvero l'effetto derivante della ricchezza economica e quello dato dal livello di sviluppo del sistema sanitario. La spesa sanitaria pro capite rappresenta, infatti, una misura del livello di benessere e di ricchezza di un paese decisamente più attendibile del PIL pro capite, soprattutto nel caso dei paesi la cui ricchezza è concentrata nelle mani di poche persone. La spesa sanitaria, invece, riesce a rappresentare in modo appropriato il welfare di un paese. Inoltre, come detto, essa indica anche il livello di sviluppo del sistema sanitario e rappresenta una proxy dell'ammontare degli investimenti da parte dello Stato nella sanità. Di conseguenza, il benessere economico dei cittadini ed il livello di sviluppo del sistema sanitario si dimostrano essere due fattori fortemente determinanti nello spiegare il numero di medaglie vinte da un paese alle Olimpiadi. Come già affermato, utilizzare Health expenditure per capitai, al posto del reddito permette di cogliere ancor meglio l'effetto del livello di sviluppo economico sulle medaglie; questi risultati confermano quindi anche quanto asserito nella letteratura, ovvero che la ricchezza pro capite risulta essere fortemente significativa. Infine, poter contare su una sanità di qualità e all'avanguardia permette di incrementare il numero di successi olimpici, in quanto il sistema sanitario può dedicarsi agli atleti, alla loro preparazione ed alla ripresa dai loro infortuni.

Le ultime due variabili della nostra funzione, ovvero Population growth_{i,t} e Neonathal deaths_{i,t}, si dimostrano significative e presentano entrambe un coefficiente negativo. La regressione mostra come una decrescita della popolazione pari a 1,39 punti percentuali, riferita al periodo compreso tra i 20 ed i 15 anni precedenti ai Giochi olimpici estivi considerati, comporti un incremento nel bottino olimpico di una medaglia extra. Questa variabile rappresenta una proxy dell'effetto della ricchezza sul totale delle medaglie; infatti, essa risulta negativamente correlata sia con la covariata che riporta il Prodotto Interno Lordo sia con quella che descrive il PIL pro capite. I paesi più ricchi sono soliti registrare un calo demografico, mentre quelli in via di sviluppo sono caratterizzati da veri e propri boom di nascite. La variazione nella dimensione della popolazione gioca un ruolo importante nel quantificare l'effetto ricchezza, come confermato dal fatto che la sua esclusione dal modello aumenterebbe il coefficiente della variabile Health expenditure per capita_{i,t}, in quanto così facendo quest'ultima catturerebbe da sola tutto l'effetto del benessere economico sul totale delle medaglie. Il regressore Population growth_{i,t} ci permette, insieme alla spesa sanitaria pro capite, di identificare l'effetto reddito come un fattore estremamente determinante per ottenere dei successi olimpici.

Come previsto, l'esclusione dal modello delle variabili riportanti la spesa sanitaria e la variazione percentuale della popolazione aumenta l'effetto marginale della covariata GDP_{i,t}. Tuttavia, questa variazione risulta lieve, poiché, come già detto, il Prodotto Interno Lordo riesce a spiegare solamente in parte il benessere economico di un paese, in quanto è una misura anche e soprattutto della dimensione della nazione stessa. Infatti, si può osservare come un modello composto dalle sole variabili GDP_{i,t} e Population_{i,t} riesca a spiegare già in modo adeguato la variabilità del totale delle medaglie, sottolineando come il Prodotto Interno Lordo e la dimensione della popolazione siano i due principali fattori che determinano il successo alle Olimpiadi estive, a conferma di quanto asserito nella letteratura.

Infine, l'analisi dell'output di regressione mostra come sia necessaria una diminuzione nel numero di bambini morti entro le prime quattro settimane di vita pari a circa quarantamila unità per poter conquistare una medaglia aggiuntiva ai Giochi olimpici estivi. Il numero di morti neonatali è estremamente legata alla dimensione della popolazione, come si evince dal coefficiente di correlazione pari a 0,73, e come conferma il fatto che la sua esclusione dal modello di regressione porti il coefficiente della covariata Population_{i,t} a diminuire. Il regressore Neonathal deaths_{i,t} riesce a rafforzare l'impatto della dimensione della popolazione sul totale delle medaglie; esso, infatti, è una buona proxy dell'effetto dato dal numero di abitanti e contribuisce, seppur in minima parte, a descrivere l'effetto dato dal livello di sviluppo del sistema sanitario e quello derivante dal benessere economico.

Nel suo complesso, il modello presenta un'ottima adattabilità ai dati, come confermato dal coefficiente di determinazione lineare corretto, pari a 0,743.

4.1.2 Verifica degli assunti

Una volta costruito questo primo modello risulta opportuno verificare se siano rispettati gli assunti alla base della teoria della regressione lineare multipla.

In prima istanza, si verifica che i residui risultino omoschedastici, sfruttando il test di Breusch-Pagan:

un p-value decisamente basso ci obbliga a rifiutare l'ipotesi di omoschedasticità a qualsiasi livello di significatività. Quindi, i residui presentano eteroschedasticità, ovvero la loro varianza non è costante, ma varia in base alla distribuzione dei dati. D'altra parte, come affermato da James H. Stock e Mark W. Watson nel loro manuale "Introduction to Econometrics" del 2020, spesso e volentieri nella pratica econometrica è prudente mettere in conto che l'assunto di omoschedasticità non venga rispettato.

In secondo luogo, per valutare la normalità dei residui si è fatto ricorso al test di Shapiro-Wilk: dall'osservazione dell'output del test si deduce che i residui non si distribuiscono secondo una variabile aleatoria normale, in quanto un valore estremamente basso assunto dal p-value implica il rifiuto dell'ipotesi nulla del test.

L'ultima assunzione prevede di verificare che la distribuzione dei residui presenti valore atteso pari a zero. Il p-value del test assume un valore estremamente elevato, addirittura approssimabile alla certezza probabilistica, per cui non vi è sufficiente evidenza statistica per rigettare l'ipotesi che la media sia pari a zero.

Dopo aver verificato gli assunti, è opportuno procedere con un'analisi grafica esplorativa dei residui del modello costruito.

Per prima cosa si valuta la linearità: Il grafico in Fig. 4.1 mostra una distribuzione dei punti non particolarmente simmetrica rispetto all'asse delle ascisse, suggerendo la presenza di una qual certa non linearità. Questa tendenza si accentua al crescere dei valori previsti.

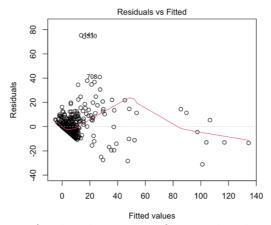
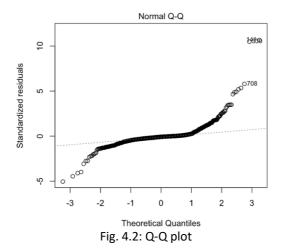


Fig. 4.1: Grafico dei valori residui in funzione dei valori stimati

Successivamente viene verificata la normalità dei residui: ovvero si valuta se i residui assumono una distribuzione campanulare simmetrica.



Come si osserva, l'andamento del grafico in Fig. 4.2 non ricalca perfettamente quello di una retta inclinata a 45 gradi, sottolineando una non normalità dei residui. Si può denotare, inoltre, un caratteristico andamento a S, che è indice di una forte asimmetria.

In terza istanza, occorre valutare l'eventuale presenza di eteroschedasticità: il grafico in Fig. 4.3 mostra un andamento medio che si discosta da quello di una retta parallela all'asse delle ascisse, confermando la presenza di eteroschedasticità. D'altro canto, non emergono pattern riconoscibili, pertanto i residui si distribuiscono in modo casuale attorno alla tendenza predominante.

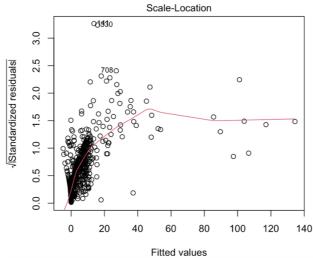


Fig. 4.3: Grafico Scale-Location

Infine, in Fig. 4.4 si prende in esame l'andamento dei residui standardizzati in funzione della leverage, ovvero della sensibilità dei valori stimati a "reagire" a variazioni dei valori osservati. Emerge come la quasi totalità dei punti rientri all'interno della banda individuata dai valori che non si discostano dalla linea di tendenza per un valore della distanza di Cook superiore a 0,5.

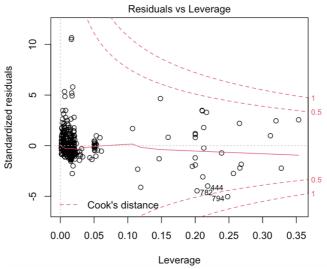


Fig 4.4: Grafico dei residui standardizzati in funzione della leverage

Si nota la presenza di osservazioni particolarmente influenti, che impattano sensibilmente sul modello stimato: nello specifico, si tratta di Cina 2016, Brasile 2016 e Grecia 2008. A questo proposito, è suggestivo notificare come si tratti di osservazioni relative ai paesi ospitanti. Per meglio

apprezzare l'influenza di tali osservazioni, potrebbe risultare utile escluderle dal dataset e stimare nuovamente il modello di regressione.

4.1.3 L'analisi della robustezza

Data la presenza di eteroschedasticità, è opportuno procedere con un test di robustezza, che verifichi se una volta eliminate alcune osservazioni sia possibile ricondurre il modello ad una condizione di omoschedasticità. A tal fine, sono state individuate le unità statistiche candidate ad essere espunte, definite outlier, tramite l'applicazione del test di Bonferroni. Vengono identificati come valori outlier i seguenti paesi: la Cina, la Germania e la Russia nel 2000, la Russia nel 2004, l'Australia, la Grecia e la Russia nel 2008, la Russia nel 2012 ed infine il Brasile e la Russia nel 2016. Una volta identificati i valori anomali, è stata condotta una stima di regressione su questo primo modello, sfruttando il dataset di partenza al netto degli outlier: dall'ispezione dell'output si nota come il coefficiente di determinazione cresca notevolmente passando da 0,743 a 0,855; il modello stimato senza la presenza di valori anomali presenta quindi una maggiore capacità di spiegare la variabilità della variabile dipendente, adattandosi meglio ai dati empirici.

L'analisi dei residui su questa nuova versione del modello ha però evidenziato un acutizzarsi dei problemi di eteroschedasticità e di non normalità dei residui; non si apprezzano, invece, variazioni in relazione all'assunto concernente la nullità della media dei residui.

Inoltre, è stata sviluppata anche un'analisi della robustezza mossa da una valutazione qualitativa: si è cercato di escludere l'effetto dato dalle cinque nazioni che hanno raccolto più medaglie nelle singole edizioni olimpiche considerate in questa ricerca. Quindi sia per le Olimpiadi di Sydney 2000 sia per quelle di Atene 2004 sono stati tolti gli USA, la Russia, l'Australia, la Cina e la Germania. Invece, per i Giochi olimpici di Pechino 2008 sono stati rimossi gli USA, la Cina, l'Australia, la Russia e la Gran Bretagna, e per Londra 2012 è stata esclusa la Germania al posto del Regno Unito. Infine, le nazioni rimosse per l'edizione di Rio 2016 sono le stesse dell'Olimpiade precedente; in realtà la Francia e la Germania presentano lo stesso numero di medaglie conquistate, ma si è scelto di escludere i Teutonici, in quanto hanno racimolato un maggior numero di ori.

Viene condotta una stima di regressione, utilizzando questa nuova versione del dataset: si nota come questa stima generi un coefficiente di determinazione lineare notevolmente inferiore a quello calcolato utilizzando il dataset completo, in quanto passa dal valore di 0,743 all'essere pari a 0,664. La riduzione dell'indice R² Adjusted si traduce in un peggioramento dell'adattabilità del modello ai dati empirici.

L'esclusione delle osservazioni riguardanti i paesi che vincono più medaglie ai Giochi olimpici estivi peggiora indiscutibilmente la bontà del modello di regressione. Con l'eliminazione di queste unità statistiche si sperava se non di risolvere, almeno di migliorare i problemi di eteroschedasticità di questo primo modello. Tuttavia, si nota come l'eteroschedasticità non soltanto persista, ma risulti persino accentuata. Allo stesso modo non sembra esserci alcun miglioramento nemmeno per quel che riguarda la non normalità. Infine, non si registrano cambiamenti in relazione all'assunzione di nullità della media dei residui.

4.1.4 Conclusione

Da questo primo modello di regressione risulta che gli effetti dominanti nell'influenzare il numero di medaglie conquistato da un paese ai Giochi olimpici estivi sono quelli dati dalla dimensione della popolazione e dalla ricchezza economica. Ciò conferma i risultati ottenuti dalla letteratura in questi anni. Inoltre, traspare come i paesi ex-sovietici e quelli tutt'ora comunisti tendano a vincere più medaglie di quanto ci si aspetterebbe basandosi solamente sulle loro caratteristiche economiche e

demografiche. La struttura politica ricopre ancora oggi un ruolo determinante nello spiegare il successo di una nazione alle Olimpiadi estive. Allo stesso modo, i risultati derivanti dalla funzione di regressione analizzata mostrano come l'essere la nazione ospitate continui a rimanere un fattore significativo e si è scoperto che l'host effect ha un'influenza che si protrae nel tempo. Infine, si è dimostrato che anche lo sviluppo del sistema sanitario ricopre un ruolo importante nel determinare il numero di medaglie conquistato da un paese ai Giochi olimpici estivi.

Le variabili che non sono state inserite nel modello risultano assenti perché non si dimostrano significative o perché, seppur significative, fanno perdere di significatività gli atri regressori inclusi nella funzione. Questo accade perché l'effetto che le variabili escluse, come ad esempio quelle che riportano la qualità del sistema scolastico od il livello di avanzamento scientifico, vogliono cogliere risulta essere già ampiamente spiegato dalle covariate presenti nel modello.

In conclusione, il primo modello realizzato, considerato nella sua versione completa di tutte le osservazioni, risulta accettabile in quanto mostra un buon adattamento ai dati nell'ambito della descrizione della variabilità della variabile risposta. Pur presentando problemi di eteroschedasticità e di non normalità dei residui, appare un apprezzabile strumento per la previsione del numero di medaglie vinte da un paese ai Giochi olimpici estivi.

4.2 Una specifica che tiene conto delle differenze culturali e regionali

4.2.1 Il modello statistico

In questo secondo modello si è voluta testare l'influenza della cultura di un paese sulle medaglie vinte. Si presume che la differenza culturale sia legata alle tradizioni riguardanti l'educazione fisica, la pratica sportiva, l'abitudine ad assistere a spettacoli sportivi o a guardarli in televisione. Queste tradizioni e questi atteggiamenti nei confronti dello sport sono profondamente radicati nella cultura locale e nelle abitudini di comportamento e di pensiero, e risultano spesso abbastanza simili tra i paesi appartenenti alla stessa macro-zona. È risultato necessario servirsi di variabili dummy regionali per catturare quest'influenza culturale.

Si è quindi partiti dalla funzione di regressione teorizzata precedentemente, a cui sono state aggiunte le variabili dicotomiche che indicano il cluster geografico di appartenenza, ottenendo la seguente equazione:

$$\begin{split} \textit{Med}_{i,t} &= c + \alpha \textit{GDP}_{i,t} + \beta \textit{Population}_{i,t} + \gamma \textit{Host}_{i,t} + \delta \textit{Next Host}_{i,t} + \varphi \textit{Previous Host}_{i,t} \\ &+ \sum_{p} \theta_{p} \textit{Politics}_{p,i} + \sum_{r} \eta_{r} \textit{Regions}_{r,i} + \lambda \textit{Health expenditure per capita}_{i,t} \\ &+ \omega \textit{Population growth}_{i,t} + \psi \textit{Neonathal deaths}_{i,t} + \varepsilon_{i,t} \end{split}$$

Si è scelto di aggiungere queste variabili seguendo quanto fatto dai due fratelli Andreff nel 2008, i quali avevano appunto introdotto delle variabili dummy regionali, che risultarono fornire un apporto significativo nella determinazione del numero di medaglie vinte alle Olimpiadi. Si tratta di sette dummy geografiche che hanno l'obbiettivo di cogliere l'influenza culturale dei paesi sulle medaglie vinte. Sono state inserite nel modello sei di queste sette dummy, lasciando come variabile di riferimento quella che indica i paesi dell'Europa e dell'Asia Centrale. Le sei dummy aggiunte al modello sono quindi "East Asia and Pacific", "Latin America and The Caribbean", "North America", "South Asia", "Sub-Saharan Africa" e "Middle East and North Africa". L'effetto marginale delle

covariate regionali inserite nel modello viene calcolato rispetto alla variabile di riferimento esclusa dalla regressione.

L'output della procedura di regressione viene riportato in Tab 4.1.

Nonostante l'inserimento nella funzione lineare delle variabili dicotomiche regionali, le covariate che riportano il Prodotto Interno Lordo e la dimensione della popolazione rimangono fortemente significative ad un livello di significatività dello 0,001, e mantengono completamente inalterati i propri effetti marginali. Questo risultato fornisce un'ulteriore riprova di quanto l'effetto della dimensione della nazione sia importante nel determinare il numero di medaglie conquistato da un paese ai Giochi olimpici estivi, confermandosi come l'effetto predominante.

Anche le tre variabili dummy Host_{i,t}, Next Host_{i,t} e Previous Host_{i,t} si mantengono significative ad un livello di significatività dello 0,001 e non presentano variazioni nel valori dei coefficienti stimati, ad eccezione di quello della variabile che indica il paese che ospiterà la futura edizione delle Olimpiadi estive. Quest'ultima, infatti, registra un lieve incremento del proprio effetto marginale, tale che il paese indicato come prossima nazione ospitante è solito portarsi a casa, in media, circa 13,32 medaglie in più. Tutto ciò evidenzia come l'host effect ed il perdurare di tale effetto nel tempo risultino determinanti fondamentali anche in questo nuovo modello, ad ulteriore riprova del loro impatto estremamente significativo.

La prima variazione rilevante rispetto al modello precedente si può osservare analizzando le dummy che indicano la struttura politica ed economica dei paesi considerati nella regressione. Infatti, mentre le variabili TRANS e COM si mantengono estremamente significative, ad un livello di significatività dello 0,001, nonostante l'aggiunta all'equazione delle variabili dicotomiche regionali, la dummy CEEC perde completamente di significatività, arrivando a non risultare pregnante neppure ad un livello dello 0,1. Questo accade poiché i paesi che vengono descritti dalla covariata CEEC appartengono tutti ad uno stesso cluster geografico, in particolare alla variabile binaria "Europe and Central Asia", che, tra l'altro, rappresenta la variabile di riferimento. Si mantiene, invece, inalterata la dummy NSCOM, che risulta essere non significativa a qualunque livello di significatività. La presenza delle variabili regionali modifica anche i coefficienti delle dummy politiche: infatti, l'effetto marginale delle covariate CEEC e TRANS diminuisce significativamente. Per contro, i paesi tutt'ora comunisti tendono ad aggiudicarsi un extra di 20,98 medaglie aggiuntive rispetto a quanto farebbe un paese con analoghe caratteristiche socio-demografiche che si basa però su un sistema economico di tipo capitalistico. Anche il coefficiente della dummy NSCOM aumenta in modo considerevole, garantendo ai paesi che appartengono a questo gruppo di vincere in media circa 2,26 medaglie in meno di quanto ci si aspetterebbe. Si può quindi affermare che, seppur con una lieve riduzione, continui ad esistere un chiaro vantaggio in termini di medaglie per i paesi comunisti o che furono parte dell'Unione Sovietica.

Dall'output della stima della regressione emerge, inoltre, che le variabili che riportano l'appartenenza di un paese ad una regione si rivelano, nel complesso, significative. Si ricorda che i coefficienti ottenuti per questi parametri sono da interpretare rispetto alla variabile di riferimento. La covariata "East Asia and Pacific" si dimostra significativa ad un livello di significatività dello 0,01, ed i paesi che ne fanno parte tendono a vincere in media circa 3,43 medaglie in meno rispetto alle nazioni dell'Europa e dell'Asia Centrale. La dummy che identifica i paesi situati in America Latina e nei Caraibi risulta essere ancora più significativa, dimostrandosi tale ad un livello dello 0,001 e presenta un effetto marginale ancora più marcato: queste nazioni sono solite aggiudicarsi circa 4,65 medaglie in meno di quanto non faccia un paese con analoghe caratteristiche ubicato però in Europa o in Asia Centrale. Analogamente, anche i paesi che si trovano in Medio Oriente ed in Africa del Nord hanno l'abitudine di conquistare in media 3,84 medaglie in meno di quelli che appartengono al cluster geografico individuato dalla variabile di riferimento; la dummy in questione risulta comunque significativa ad un livello dello 0,01. Per contro, la variabile dicotomica "North America"

si dimostra completamente non significativa anche ad un livello di significatività dello 0,1 e i paesi che si trovano in questa regione sono soliti aggiudicarsi 1,89 medaglie in meno di quanto ci si aspetterebbe. Si tratta della regione che presenta il differenziale minore con la variabile di riferimento, e questo accade perché al suo interno si trovano Canada e Stati Uniti, paesi in cui vi è una grandissima cultura sportiva ed in cui l'attività fisica assume un ruolo centrale nella società. Contrariamente, la dummy che presenta la differenza più marcata in termini di medaglie con la variabile di riferimento è "South Asia". Quest'ultima risulta essere fortemente significativa, rimanendo tale anche ad un livello dello 0,001, ed i paesi che si trovano in questa regione vincono in media addirittura 7,93 medaglie in meno di quanto facciano le nazioni europee e dell'Asia Centrale. Infine, la variabile dicotomica che riporta i paesi ubicati nell'Africa Sub-Sahariana si dimostra debolmente significativa ad un livello di significatività dello 0,05, e tali paesi tendono ad aggiudicarsi alle Olimpiadi estive 3.17 medaglie in meno rispetto a paesi con eguali caratteristiche economiche e demografiche situati però in Europa o in Asia Centrale. Questi risultati mostrano chiaramente come l'effetto dell'ubicazione geografica di un paese sia un fattore fortemente determinante nell'influenzare il numero di medaglie conquistato alle Olimpiadi estive. Infatti, queste dummy geografiche riescono a cogliere numerosi effetti: primo fra tutti quello legato alle differenze riguardanti la cultura sportiva e non sportiva tra i diversi paesi. Inoltre, queste variabili riescono a spiegare anche gli effetti sul totale delle medaglie dati dal clima e dalla temperatura, ed in alcuni casi sono capaci di convogliare al loro interno anche gli effetti dovuti alla struttura politica ed al benessere economico di un paese. Si evince, quindi, che le variabili dicotomiche regionali rappresentano una buona proxy per dipingere al meglio il quadro dei fenomeni economici, geografici e politici. Pertanto, questi risultati confermano quanto asserito dalla letteratura: essi sono, infatti, in linea con lo studio dei fratelli Andreff (2008), i quali affermavano che l'ubicazione geografica era uno dei fattori più rilevanti nello spiegare il numero di medaglie conquistate ai Giochi olimpici estivi. Tutt'oggi, a 12 anni di distanza, viene confermato come il suo effetto risulti estremamente significativo.

Tuttavia, con l'inserimento delle dummy regionali la spesa sanitaria pro capite diventa completamente non significativa anche ad un livello di significatività dello 0,1 ed il suo effetto marginale subisce una cospicua riduzione. Questo accade perché i due effetti che tale variabile si proponeva di spiegare, ovvero il livello di benessere economico ed il livello di sviluppo del sistema sanitario di un paese, vengono completamente catturati dalle variabili dicotomiche regionali. Esse, infatti, suddividendo il mondo in cluster geografici, riescono a suddividere i paesi, in modo più o meno efficiente, anche in base alla ricchezza economica e, conseguentemente, in base alla qualità offerta dalla sanità nazionale.

Infine, i due regressori Population growth_{i,t} e Neonathal deaths_{i,t} si mantengono significativi per gli stessi livelli di significatività del modello senza le dummy regionali, ovvero rispettivamente ad un livello dello 0,01 e dello 0,001. L'effetto marginale della variazione percentuale della popolazione diminuisce, anche se solo lievemente, mentre quello della variabile che riporta il numero di morti neonatali rimane pressoché invariato.

L'aggiunta al modello delle variabili dicotomiche che suddividono i paesi in base alla propria posizione geografica migliora anche la capacità del modello di spiegare la variabilità del totale delle medaglie e di adattarsi ai dati; infatti, il coefficiente di determinazione lineare corretto aumenta lievemente e si attesta sul valore di 0,747.

4.2.2 Verifica degli assunti

Una volta costruito ed analizzato questo secondo modello, risulta opportuno verificare se vengono rispettati gli assunti alla base della teoria della regressione lineare multipla.

In primo luogo, si è verificata l'omoschedasticità dei residui attraverso il test di Breusch-Pagan, che genera un p-value estremamente piccolo, obbligandoci a rifiutare l'ipotesi di omoschedasticità a qualsiasi livello di significatività. Quindi, anche in questo secondo modello, i residui risultano essere eteroschedastici.

In seconda istanza, si è valutata la normalità dei residui facendo ricorso al test di Shapiro-Wilk, da cui si deduce che i residui non si distribuiscono secondo una variabile aleatoria normale.

Infine, l'ultima assunzione prevede di verificare che i residui presentino media nulla. Il p-value del test assume un valore estremamente elevato, approssimabile all'unità, che ci obbliga a non poter rifiutare l'ipotesi secondo cui la distribuzione dei residui presenta valore atteso pari a zero.

4.2.3 L'analisi della robustezza

Poiché il modello presenta dei problemi di eteroschedasticità, è stato effettuato un test di robustezza per cercare di risolverli e ricondurre il modello ad una condizione di omoschedasticità, attraverso l'eliminazione di alcune osservazioni.

Vengono identificati come valori outlier, quindi da eliminare dal dataset, i seguenti paesi: la Cina, la Germania e la Russia nel 2000, la Russia nel 2004, l'Australia e la Russia nel 2008, la Russia nel 2012 ed infine il Brasile, la Cina e la Russia nel 2016.

È stata quindi condotta una stima di regressione sfruttando la funzione comprendente le dummy regionali ed utilizzando il dataset di partenza, dal quale sono state escluse le osservazioni menzionate sopra. Dall'analisi dei risultati della regressione si osserva nitidamente come l'indice R² Adjusted del modello aumenti significativamente, passando da 0,747 a 0,851. Quindi, anche in questo caso il modello stimato in assenza degli outlier presenta una maggiore capacità di spiegare la varianza campionaria del totale delle medaglie, adattandosi meglio ai dati empirici.

Tuttavia, nonostante le modifiche apportate al dataset, permangono sia i problemi di eteroschedasticità sia quelli di non normalità; analogamente non si apprezzano variazioni per quanto riguarda la nullità della media dei residui.

Come per il modello precedente, è stata sviluppata anche una seconda analisi della robustezza, mossa da una valutazione qualitativa in cui si è cercato di escludere l'effetto dato dalle cinque nazioni che hanno raccolto più medaglie nelle singole edizioni olimpiche considerate in questa dissertazione.

Si è quindi effettuata una regressione sulla funzione lineare considerata in questo secondo modello, utilizzando però un dataset dal quale sono stati rimossi per ogni Olimpiade i cinque paesi che hanno conquistato più medaglie; dall'ispezione dell'output statistico si osserva una riduzione del coefficiente di determinazione lineare corretto in questa nuova versione del modello rispetto al valore riportato in quella originaria. L'indice R² Adjusted passa infatti da 0,747 al valore di 0,684, mostrando quindi una minore capacità del modello di spiegare la variabilità della variabile dipendente ed una peggior capacità di adattamento ai dati.

Inoltre, nemmeno per mezzo di questa seconda analisi di robustezza si riescono a risolvere i problemi di non normalità ed eteroschedasticità dei residui, anzi, quest'ultima appare accentuarsi se si utilizza il dataset al netto dei paesi che dominano il medagliere. Non si osserva, invece, alcuna variazione in merito alla non normalità ed alla nullità della media.

4.2.4 Conclusione

Riassumendo, è possibile affermare senza tema di smentita che l'ubicazione geografica dei paesi risulta essere, a conferma di quanto asserito dalla letteratura, un fattore estremamente significativo nel determinare il numero di medaglie vinto da un paese ai Giochi olimpici estivi. Questo accade

poiché le variabili regionali utilizzate possono essere considerate delle ottime proxy di diversi effetti rilevanti. In primo luogo, queste dummy sono in grado di spiegare le differenze culturali tra i vari paesi e la diversa importanza che viene data allo sport all'interno della società. Le variabili dicotomiche regionali sono anche in grado di cogliere l'effetto dettato dalla temperatura e dal clima sul totale delle medaglie. Infine, le covariate che riportano la collocazione geografica riescono a descrivere in modo piuttosto preciso anche il livello di benessere economico di una nazione, rendendo, infatti, completamente non significativa la variabile Health expenditure per capita_{i,t}.

Si può quindi concludere che i paesi che non si trovano in Europa od in Asia Centrale tendono a vincere meno medaglie delle nazioni con eguali caratteristiche economiche e demografiche che appartengono alla variabile di riferimento. Se ne deduce che l'Europa e l'Asia Centrale siano favorite da una forte cultura sportiva, dal ruolo centrale giocato dallo sport nella società e da una condizione climatica estremamente favorevole.

Le variabili regionali non sono, invece, in grado di cogliere l'effetto dato dalla dimensione della popolazione sul totale delle medaglie, e, per l'appunto, le variabili che riportano quest'effetto si dimostrano ancora significative.

Si può quindi dedurre che anche secondo questo secondo modello di regressione lineare l'effetto della dimensione della popolazione rimane il più rilevante ai fini di determinare il successo di un paese alle Olimpiadi. Parimenti viene confermata la rilevanza sia della struttura politica sia del vantaggio dato dall'essere la nazione ospitante. Si dimostra, ancora una volta, che quest'ultimo effetto è in grado di protrarsi nel tempo, influenzando sia l'edizione precedente sia quella successiva a quella ospitata.

4.3 L'importanza della religione

4.3.1 Il modello statistico

L'obbiettivo di questo terzo modello consiste nell'indagare l'effetto che la religione professata in un paese esercita sul numero di medaglie conquistate ai Giochi olimpici estivi. Si è quindi partiti dalla funzione lineare costruita per testare l'importanza della struttura politica a cui sono state aggiunte le variabili dummy che rappresentano la religione professata. Si ottiene quindi la seguente equazione:

$$\begin{split} \textit{Med}_{i,t} &= c + \alpha \textit{GDP}_{i,t} + \beta \textit{Population}_{i,t} + \gamma \textit{Host}_{i,t} + \delta \textit{Next Host}_{i,t} + \varphi \textit{Previous Host}_{i,t} \\ &+ \sum_{p} \theta_{p} \textit{Politics}_{p,i} + \sum_{r} \eta_{r} \textit{Religions}_{r,i} \\ &+ \lambda \textit{Health expenditure per capita}_{i,t} + \omega \textit{Population growth}_{i,t} \\ &+ \psi \textit{Neonathal deaths}_{i,t} + \varepsilon_{i,t} \end{split}$$

L'idea di introdurre queste variabili dicotomiche trova origine nello studio dei ricercatori Julia Bredtmann, Carsten J. Crede e Sebastian Otten (2016). Essi utilizzavano la variabile Muslim_{it}, che indicava i paesi con una popolazione a maggioranza musulmana, e hanno dimostrato che queste nazioni sono solite avere una quota inferiore della popolazione femminile attiva negli sport professionistici, tendono ad inviare meno atlete alle Olimpiadi e, di conseguenza, vincono meno medaglie. Risulta però importante scoprire quali altri effetti è in grado di spiegare la religione e se anche le altre religioni hanno un effetto sul totale delle medaglie.

Pertanto, sono state introdotte nella funzione di regressione lineare, al posto delle dummy regionali, sei variabili inerenti alla religione praticata e si è scelto di utilizzare come variabile di riferimento quella che indica i paesi che si basano sulla religione cristiana. Le sei dummy aggiunte al modello sono quindi "Buddhism", "Folk Religions", "Hinduism", "Islam", "Judaism" e "Unaffiliated Religions". Gli effetti marginali delle covariate, che indicano la religione professata, sono quindi stati calcolati rispetto alla variabile di riferimento, che è stata esclusa dalla funzione di regressione.

In Tab 4.1 è possibile osservare i risultati della procedura di regressione.

Alla stregua di quanto affermato in merito alle variabili regionali, l'aggiunta delle variabili dicotomiche di tipo religioso conferma come i regressori GDP_{i,t} e Population_{i,t} siano estremamente significativi, anche ad un livello di significatività dello 0,001. I coefficienti di tali parametri rimangono sostanzialmente invariati, anche se l'effetto marginale della dimensione della popolazione subisce un incremento. Infatti, il modello attesta come sia necessario un aumento di 38 milioni di abitanti per ottenere una medaglia in più alle Olimpiadi estive mentre il primo modello, che includeva solamente le dummy politiche, mostrava come fosse richiesto un aumento della dimensione della popolazione pari a 45 milioni di persone per ottenere lo stesso risultato. Pertanto, si rafforza maggiormente quanto detto in precedenza, ovvero che il numero di abitanti di un paese rappresenta il fattore che maggiormente influenza il numero di medaglie conquistate ai Giochi olimpici estivi.

Analogamente, le dummy Host_{i,t}, Next Host_{i,t} e Previous Host_{i,t}, che rappresentano il vantaggio in termini di medaglie dato dall'ospitare le Olimpiadi estive, si confermano estremamente significative, anch'esse ad un livello di significatività dello 0,001. Parimenti, gli effetti marginali sul totale delle medaglie rimangono quasi del tutto invariati, ad eccezione di quello della variabile che indica il paese che ospiterà la prossima edizione dei Giochi, che fa registrare un lieve incremento. Infatti, secondo questo modello, essere il Next Host garantisce ben 13,51 medaglie in più, contro le 12,99 predette dal primo modello. Questi risultati confermano ulteriormente quanto affermato sia dai due precedenti modelli sia in letteratura; vale a dire che essere la nazione che ospita i Giochi olimpici risulta essere un fattore estremamente rilevante nel determinare il numero di medaglie, poiché la nazione organizzatrice può contare su un contingente di atleti più numeroso e su ingenti investimenti in infrastrutture per le gare e per gli allenamenti. La regressione mostra che quest'effetto, come sottolineavano anche i due modelli precedenti, si protrae nel tempo influenzando il medagliere sia dell'edizione che precede le Olimpiadi estive ospitate sia di quella che le segue.

Dall'output della stima della regressione emerge, inoltre, che le variabili dicotomiche che indicano la struttura politico-economica si confermano nel complesso significative, anche secondo questo modello. Esse mantengono infatti lo stesso livello di significatività che presentavano nel modello che includeva solo le dummy politiche: infatti, le covariate CEEC, TRANS e COM risultano significative ad un livello dello 0,001, mentre la variabile riportante i paesi che un tempo erano comunisti pur non essendo sotto l'influenza dell'Unione Sovietica si conferma non significativa anche ad un livello di significatività dello 0,1. Tuttavia, si possono apprezzare delle variazioni per quanto riguarda i coefficienti di questi parametri rispetto al primo modello costruito. Il coefficiente della variabile dicotomica CEEC mostra un lieve incremento; infatti, i paesi che appartengono a tale gruppo di ricerca sono soliti vincere circa 4,66 medaglie in più rispetto all'andamento medio. Parimenti, anche l'effetto marginale della dummy TRANS mostra un incremento piuttosto rilevante, secondo cui i paesi che erano parte dell'Unione Sovietica tendono a vincere circa 8,86 medaglie aggiuntive rispetto a paesi con analoghe caratteristiche economiche e demografiche, ma che adottano un sistema di tipo capitalistico. Per contro, il coefficiente della variabile binaria COM non mostra alcuna variazione rispetto a quello ottenuto con il primo modello, così come quello della dummy NSCOM. La regressione conferma, anche in questo caso, che i paesi che adottano un sistema comunista o che sono stati sotto l'influenza dell'URSS vincono più medaglie di quanto sarebbe lecito aspettarsi basandosi solamente sui parametri economici, demografici e sociali. L'inserimento delle covariate religiose sembra addirittura rafforzare l'effetto sul totale delle medaglie dettato dall'assumere una struttura politica socialista.

Dall'osservazione dell'analisi dei risultati della regressione si nota come, nel complesso, le variabili dicotomiche religiose aggiunte siano piuttosto significative nello spiegare il numero di medaglie conquistate da un paese ai Giochi olimpici estivi. Si ricorda che i coefficienti ottenuti per questi parametri sono da interpretare rispetto alla variabile di riferimento. La variabile "Buddhism" si dimostra significativa solamente ad un livello di significatività dello 0,1. Si dimostra quindi un parametro poco rilevante e l'effetto marginale che ne consegue è tale che i paesi che professano questa religione sono soliti vincere circa 2,22 medaglie in meno di quanto non facciano nazioni con analoghe caratteristiche socio-demografiche, in cui però la religione ufficiale risulta essere quella cristiana. Per contro, la seconda dummy religiosa aggiunta, denominata "Folk Religions", risulta essere estremamente significativa, anche ad un livello dello 0,001. Le nazioni in cui si professano questi tipi di religione, come il Vietnam, sono solite aggiudicarsi addirittura 11,86 medaglie in meno di quanto ci si aspetterebbe basandosi solamente sugli altri parametri di questa funzione. I paesi individuati da questa dummy presentano il maggior differenziale negativo in termini di medaglie vinte con le nazioni della variabile di riferimento "Christianity". La variabile binaria "Hinduism" si dimostra, invece, significativa ad un livello di significatività dello 0,01; anch'essa presenta un coefficiente negativo, il cui effetto marginale indica come i paesi nei quali la religione principale risulta essere quella induista mostrano una tendenza a conquistare in media circa 5,99 medaglie in meno rispetto ad analoghe nazioni in cui si professa il Cristianesimo. La dummy "Islam", che rappresenta la variabile per la quale si è deciso di testare l'influenza delle religioni sulle medaglie conquistate alle Olimpiadi estive, risulta anch'essa significativa ad un livello di significatività dello 0,01. Si dimostra quindi un parametro rilevante nel determinare il successo olimpico, e, come prevedibile, presenta un effetto marginale negativo, tale per cui i paesi islamici sono soliti vincere, in media, 1,96 medaglie in meno di quanto sarebbe altrimenti prevedibile. Si tratta della variabile che presenta il più basso differenziale in termini di successi olimpici con la variabile di riferimento. Contrariamente a quanto visto finora, si può osservare che la covariata "Judaism" risulta essere completamente non significativa, dimostrandosi tale persino ad un livello di significatività dello 0,1. L'unico paese in cui la religione ufficiale è il giudaismo risulta essere Israele, un paese che, nonostante una popolazione ridotta, è in grado di conquistare almeno un paio di medaglie alle Olimpiadi estive, proprio grazie ad un elevato livello di benessere e di sviluppo del sistema sanitario. Il coefficiente di questa dummy sembra annoverare al suo interno quest'evidenza empirica, generando un effetto marginale secondo cui i paesi che professano la religione ebraica sono soliti vincere circa 2,98 medaglie in meno rispetto ai concorrenti con analoghe caratteristiche economiche e demografiche, che si basano invece sulla dottrina cristiana. Infine, la variabile dicotomica denominata "Unaffiliated Religions" risulta essere significativa solamente ad un livello di significatività dello 0,05 e mostra anch'essa un coefficiente negativo. I paesi che si dichiarano atei tendono ad aggiudicarsi circa 4,92 medaglie in meno rispetto ai paesi che appartengono alla variabile di riferimento. I risultati della regressione evidenziano come la religione sia un fattore che influenza il numero di medaglie conquistate alle Olimpiadi estive, in quanto è in grado di fornire un'ottima proxy per molti parametri sociali, quali ad esempio il livello di uguaglianza di genere. Inoltre, le variabili religiose riescono a cogliere, anche se solo parzialmente, l'effetto della ricchezza e del livello di benessere di un paese sul totale delle medaglie; infatti, i paesi in cui la religione ufficiale è quella cristiana risultano essere correlati positivamente sia con il PIL pro capite sia, in misura ancor più consistente, con la spesa sanitaria pro capite.

La variabile Health expenditure per capita_{i,t} si conferma essere estremamente significativa, addirittura ad un livello di significatività dello 0,001. Il coefficiente di tale parametro fa registrare una riduzione, generando così un effetto marginale per cui risulta necessario un aumento di circa 1300 USD nella spesa sanitaria pro capite per permettere al paese in questione di conquistare un ulteriore medaglia olimpica. Questa diminuzione dell'effetto marginale può essere attribuita al fatto che parte dell'effetto che nel primo modello veniva spiegato solo dalla spesa sanitaria pro capite ora viene in parte colto dalle dummy religiose: si tratta dell'effetto sulle medaglie dettato dal livello di benessere economico di cui godono i cittadini.

Il regressore che riporta la variazione demografica percentuale perde di significatività rispetto ai due modelli precedenti, rimanendo però significativo ad un livello di significatività dello 0,05. Il coefficiente della covariata Population growth_{i,t} rimane negativo ed evidenzia una diminuzione dell'effetto marginale; la regressione mostra, infatti, come una decrescita demografica di 2,11 punti percentuali sia condizione necessaria per un paese per aggiudicarsi una medaglia extra alle Olimpiadi estive. Anche in questo caso, la diminuzione dell'effetto marginale è da attribuirsi all'introduzione nel modello delle dummy religiose; infatti, il regressore Population growth_{i,t} contribuisce alla spiegazione dell'effetto sulle medaglie dato dal livello di ricchezza dei cittadini, che in questo modello viene in parte descritto anche dalle variabili dicotomiche religiose.

Inoltre, la covariata Neonathal deaths_{i,t} mantiene esattamente lo stesso livello di significatività mostrato nei due modelli precedenti, rimanendo quindi fortemente significativa ad un livello dello 0,001. Analogamente, il suo coefficiente rimane quasi del tutto inalterato.

In conclusione, l'aggiunta al modello delle variabili dicotomiche che raggruppano i paesi di questa ricerca in base al tipo di religione professata migliora l'adattabilità del modello ai dati empirici, come testimonia l'incremento subito dall'R² Adjusted, che si attesta ora su un valore di 0,751.

4.3.2 Verifica degli assunti

Una volta costruito questo terzo modello risulta opportuno verificare se siano rispettati gli assunti alla base della teoria della regressione lineare multipla.

In prima istanza, si verifica che i residui risultino omoschedastici, sfruttando il test di Breusch-Pagan, che genera, in questo caso, un p-value estremamente basso, che porta a rifiutare l'ipotesi di omoschedasticità a qualsiasi livello di significatività. Quindi, anche in questo terzo modello, i residui risultano essere eteroschedastici.

In secondo luogo, si è valutata la normalità dei residui attraverso il test di Shapiro-Wilk, da cui si deduce che i residui non si distribuiscono secondo una variabile aleatoria normale, in quanto il valore estremamente basso assunto dal p-value implica il rifiuto dell'ipotesi nulla del test.

Infine, si è verificato che la distribuzione dei residui presenti valore atteso pari a zero. Il p-value del test assume un valore estremamente elevato, addirittura approssimabile alla certezza probabilistica, per cui non vi è sufficiente evidenza statistica per rigettare l'ipotesi che la media sia pari a zero.

4.3.3 L'analisi della robustezza

Data la presenza di eteroschedasticità, è stato effettuato un test di robustezza per cercare di ricondurre il modello ad una condizione di omoschedasticità, attraverso l'eliminazione di alcune osservazioni anomale.

Tramite l'utilizzo del test di Bonferroni, vengono identificate come outlier, e quindi da eliminare dal dataset, le seguenti unità statistiche: la Cina, la Germania e la Russia nel 2000, la Russia nel 2004, l'Australia e la Russia nel 2008, la Russia nel 2012 ed infine il Brasile, la Cina e la Russia nel 2016.

È stata quindi effettuata una stima di regressione tramite la funzione lineare che include le variabili dicotomiche religiose, utilizzando un dataset dal quale sono stati rimossi gli outlier; si nota immediatamente come al netto delle osservazioni anomale il modello si dimostri capace di spiegare una maggior porzione della variabilità della variabile dipendente, come testimonia l'aumento del coefficiente di determinazione lineare corretto, che passa da un valore di 0,751 ad un valore di 0,850.

Tuttavia, nonostante le modifiche apportate al dataset, permangono sia i problemi di eteroschedasticità, che si sono addirittura acutizzati, sia quelli di non normalità, mentre non si osservano variazioni per quanto riguarda la nullità della media dei residui.

Anche in questo caso, è stata sviluppata una seconda analisi della robustezza mossa da una valutazione qualitativa, secondo cui si è cercato di escludere l'effetto dato dalle cinque nazioni che hanno raccolto più medaglie nelle singole edizioni olimpiche considerate in questa ricerca.

Si è condotta un'analisi di regressione sul dataset completo dal quale sono stati rimossi per ogni Olimpiade i cinque paesi che hanno conquistato più medaglie. Dall'osservazione dei risultati della procedura si osserva, rispetto al modello che si basa sul dataset completo, una riduzione dell'indice R² Adjusted, che passa infatti da 0,751 a 0,687. Questa nuova versione del modello non solo si dimostra meno capace di predire la variabilità del totale delle medaglie, ma non riesce nemmeno a risolvere i problemi relativi all'eteroschedasticità ed alla non normalità dei residui. Viene, anche in questo caso, rispettato l'assunto di nullità della media dei residui.

4.3.4 Conclusione

Riassumendo, si può affermare che anche in questo caso il fattore che sembra più di tutti influenzare il numero di medaglie vinto ai Giochi olimpici estivi da un paese è la dimensione della popolazione. Parimenti, viene confermata l'importanza della struttura politica che caratterizza i paesi oggetto dello studio; anche in questo caso essere un paese comunista od essere stato sotto l'influenza dell'Unione Sovietica garantisce ancora oggi un vantaggio piuttosto evidente in termini di medaglie conquistate.

Anche in questo terzo modello si conferma il vantaggio generato dall'essere la nazione padrona di casa: chi ospita l'evento olimpico tende ad ottenere molti più successi di quanto non farebbe da semplice nazione partecipante. Inoltre, il vantaggio dato dall'ospitare le Olimpiadi è solito perdurare nel tempo, influenzando sia l'edizione che segue sia quella che precede quella ospitata. Questi risultati confermano sia quanto affermato dai due modelli precedenti sia quanto asserito dalla letteratura.

Le variabili dicotomiche che suddividono i paesi in base alla religione professata si dimostrano nel complesso rilevanti nell'influenzare il successo olimpico: si tratta di un risultato innovativo, non ancora descritto in letteratura. Infatti, attraverso la matrice religiosa, si riescono a catturare diverse caratteristiche sociali ed economiche di un paese, come ad esempio il livello di sviluppo del sistema scolastico, la parità di genere, il ruolo della donna nella società e, di conseguenza, il suo ruolo nello sport. Inoltre, suddividendo i paesi in base alla religione professata, si riesce a cogliere, seppur parzialmente, l'effetto derivante dal benessere dei cittadini sul totale delle medaglie: infatti, come detto, i paesi cristiani tendono a caratterizzarsi per una qualità di vita superiore.

Quindi, grazie alle variabili inerenti alla pratica religiosa ed alla covariata che indica la spesa sanitaria pro capite si osserva che la ricchezza della popolazione risulta essere uno dei fattori principali nel determinare il numero di medaglie che un paese conquista alle Olimpiadi estive.

Si è provato anche ad includere nello stesso modello sia le variabili dicotomiche regionali sia le dummy concernenti la religione praticata: così facendo, queste ultime risultano poco significative, in quanto gli aspetti culturali ed economici che si propongono di cogliere vengono spiegati in modo

più completo dalle covariate geografiche. Si può, infatti, osservare una correlazione tra questi due gruppi di variabili binarie, che suggerisce di inserirle separatamente in due modelli distinti.

4.4 Un modello innovativo

4.4.1 Il modello statistico

L'idea alla base di questo quarto modello poggia su una strategia completamente differente da quelle viste finora. Si è deciso di non partire da quanto asserito in letteratura, come invece è stato fatto per i tre modelli precedenti, ma di applicare uno degli algoritmi alla base della stepwise regression. Si tratta della forward selection, che prevede di iniziare con un modello senza variabili e di testare l'aggiunta di ciascun regressore a disposizione, provando ad inserire per primo quello che presenta la correlazione più intensa in valore assoluto con il totale delle medaglie, per poi decidere di aggiungere al modello la variabile la cui aggiunta fornisce il miglioramento statisticamente più significativo alla bontà di adattamento del modello; questo processo viene reiterato, portando ad includere nel modello via via altre variabili, per poi interrompere le iterazioni quando nessuno dei regressori candidati è in grado di migliorare l'adattabilità del modello in misura statisticamente significativa.

Così facendo, si è giunti alla seguente funzione lineare:

$$Med_{i,t} = c + \alpha Athletes_{i,t} + \beta Scientific \ articles_{i,t} + \lambda Health \ expenditure \ per \ capita_{i,t} + \omega Life \ expectancy_{i,t} + \psi Neonathal \ deaths_{i,t} + \xi Unemployment_{i,t} + \varepsilon_{i,t}$$

Gli effetti delle covariate ricomprese nella funzione lineare sono visibili in Tab 4.1.

Quasi tutte le variabili indipendenti che fanno parte di questa funzione di regressione lineare risultano essere molto significative ed in grado di influenzare il numero di medaglie conquistate da un paese alle Olimpiadi estive.

In particolare, la covariata che indica il numero di atleti inviati da una nazione ad una determinata edizione dei Giochi si dimostra estremamente significativa, ad un livello di significatività dello 0,001. Essa presenta un coefficiente positivo con un effetto marginale piuttosto rilevante, tale che risulta necessario un aumento di soli 11,12 atleti per vincere una medaglia in più. Questo avviene perché la covariata Athletes_{i,t} risulta essere funzione di altre variabili, come la dimensione della popolazione, il livello di sviluppo economico e l'essere la nazione ospitante, che rappresentano i principali fattori che determinano il successo olimpico di un paese. Infatti, come suggerito dai lavori di Johnson e Ali (2000) e di Kuper e Sterken (2001), la partecipazione è strettamente influenzata dal PIL pro capite e dalla popolazione, in quanto i paesi più ricchi e popolosi inviano più atleti alle Olimpiadi. Inoltre, essere la nazione ospitante, o quantomeno un paese geograficamente limitrofo a quello dove si svolgono i Giochi, porta ad un aumento del numero di atleti che prendono parte alla manifestazione. Il numero di atleti risulta essere la variabile più determinante di quelle viste finora, proprio perché raccoglie al suo interno gli effetti delle principali variabili economiche, demografiche e sociali. Ovviamente, più nutrito è il contingente di atleti che una nazione presenta ai Giochi olimpici estivi, maggiore è la probabilità di conquistare delle medaglie, sulla scorta di una deduzione puramente numerica. Quindi, presentarsi alle Olimpiadi con un elevato numero di atleti permette di conquistare un ricco bottino di medaglie, a conferma di quanto asserito dalla letteratura.

Parimenti, il regressore Scientific articles_{i,t} si dimostra fortemente significativo, dimostrandosi tale ad un livello di significatività dello 0,001. Questa covariata presenta un coefficiente positivo che genera un effetto marginale per cui risulta necessario un incremento nel numero di articoli scientifici pubblicati in un anno pari a 7148 papers. Anche questo fattore si dimostra estremamente rilevante nel determinare il numero di medaglie conquistate da un paese ai Giochi olimpici estivi, in quanto raccoglie al suo interno molti effetti, da quelli economici a quelli demografici. Innanzitutto, il numero di articoli scientifici è una variabile in grado di ritrarre in modo chiaro e trasparente lo stato di salute e di sviluppo del sistema scientifico di un paese. Infatti, la letteratura evidenzia che le scoperte effettuate dalla ricerca scientifica spostano i limiti fisici degli atleti sempre più in alto, grazie ad allenamenti mirati ed alla ricerca medica e che, per l'appunto, i paesi utilizzano le scoperte provenienti dal mondo accademico ed universitario per incrementare il proprio successo alle Olimpiadi. Questa variabile è quindi in grado di cogliere pienamente gli effetti della ricerca scientifica e dell'avanzamento tecnologico sul totale delle medaglie. Inoltre, il numero di articoli pubblicati risulta essere una buona proxy della dimensione della popolazione, in quanto le due variabili risultano correlate fra loro, come testimoniato dal coefficiente di correlazione pari a 0,54. Poiché il numero di articoli scientifici non è diviso per il numero di abitanti, risulta evidente che i paesi più grandi abbiano più persone da dedicare alla ricerca scientifica e che quindi possano pubblicare un elevato numero di papers. Inoltre, il regressore in questione riesce anche a cogliere l'effetto sul totale delle medaglie dato dal livello di sviluppo economico e dal livello di benessere dei cittadini. Infatti, il numero di articoli scientifici risulta estremamente correlato con il Prodotto Interno Lordo, come conferma il coefficiente di correlazione pari a 0,97; d'altro canto, come detto, la variabile GDP_{i,t} descrive solamente in parte l'effetto della ricchezza economica, perché coglie principalmente quello fornito dalla dimensione della popolazione. La variabile Scientific articles_{i,t} è in grado di descrivere anche l'effetto del benessere economico sul successo olimpico, come sottolineato dal fatto che risulta essere positivamente correlata sia con il PIL pro capite sia, in modo ancora più intenso, con la spesa sanitaria pro capite. Inoltre, può essere anche considerata una proxy del livello di sviluppo del sistema medico e di quello scolastico, che rappresentano fattori necessari per poter supportare lo sviluppo della ricerca e dell'innovazione. Infine, la significatività del numero di articoli scientifici conferma quanto asserito in letteratura dagli scienziati turchi Süleyman Şahin e Fatih Şenduran (2020).

La variabile Health expenditure per capita_{i,t} risulta essere estremamente significativa, anche ad un livello di significatività dello 0,001, esattamente come nel primo e nel terzo modello, ma rispetto a questi ultimi presenta un coefficiente negativo. Questo accade perché la spesa sanitaria pro capite si propone di spiegare l'effetto sulle medaglie causato dal livello di benessere economico e dal livello di sviluppo del sistema scientifico-sanitario; tuttavia, questi due effetti vengono già ampiamente catturati dalle due covariate Scientific articles_{i,t} e Athletes_{i,t}, che descrivono a fondo come l'aspetto economico e quello scientifico influenzino il successo olimpico di un paese.

Anche la variabile Life expectancy_{i,t} si dimostra estremamente significativa, addirittura ad un livello di significatività dello 0,001, e contraddistinta da un coefficiente negativo. Questa covariata, come la precedente, è un'ottima proxy del benessere economico e della qualità del servizio erogato dal sistema sanitario nazionale. L'aspettativa di vita risulta, infatti, correlata positivamente sia con il PIL pro capite sia con la stessa spesa sanitaria per persona, come testimoniato dai coefficienti di correlazione che risultano essere rispettivamente pari a 0,55 ed a 0,57. Il coefficiente della variabile Life expectancy_{i,t} appare quindi negativo semplicemente perché gli effetti economici e scientificosanitari che essa si propone di spiegare vengono completamente colti, così come per la variabile precedente, dai regressori che riportano il numero di atleti inviati alle Olimpiadi estive ed il numero di articoli scientifici pubblicati in un anno.

Una sorte simile tocca alla covariata Neonathal deaths_{i,t}, che, come negli altri modelli costruiti, risulta essere molto significativa, dimostrandosi tale ad un livello di significatività dello 0,001, e che presenta un coefficiente negativo. Come già detto per i modelli precedenti, il numero di morti neonatali risulta essere una proxy in grado di cogliere gli effetti della dimensione della popolazione sul totale delle medaglie conquistate da un paese. Tuttavia, in questo caso gli effetti del numero di abitanti sono stati già ampiamente colti dalle due covariate Scientific articles_{i,t} e Athletes_{i,t}, e per questo motivo la variabile in questione presenta un coefficiente negativo.

Infine, il regressore Unemployment_{i,t} si dimostra significativo ad un livello dello 0,01 e presenta anch'esso un coefficiente negativo, per le stesse ragioni menzionate per le precedenti variabili. La disoccupazione è una misura del livello di benessere economico di una nazione e per quanto questa variabile si proponga di coglierne l'effetto sul successo olimpico, l'impatto economico viene già colto in modo chiaro e esauriente dalle prime due covariate di questa funzione di regressione lineare.

Come si può osservare, in questo modello non vengono inserite le variabili dicotomiche che identificano il vantaggio dato dall'ospitare i Giochi olimpici estivi, poiché non risultano essere significative e non migliorano l'R² Adjusted del modello.

Discorso analogo è stato fatto per le dummy che suddividono i paesi in base alla loro posizione geografica ed in base alla religione professata, che risultano non significative in quanto gli effetti economici, demografici, sociali e geografici che si propongono di cogliere vengono già ampiamente spiegati dalle variabili presenti.

Il modello così creato risulta comunque più capace di spiegare la varianza campionaria del totale delle medaglie di quanto non lo siano i tre modelli precedenti. Dall'analisi dei risultati della regressione emerge che questo quarto modello genera un coefficiente di determinazione lineare corretto pari 0,878, dimostrando un'elevata adattabilità ai dati empirici.

4.4.2 Verifica degli assunti

Una volta costruito questo quarto modello, risulta opportuno verificare se siano rispettati gli assunti alla base della teoria della regressione lineare multipla.

In primo luogo, si è verificata l'omoschedasticità dei residui per mezzo del test di Breusch-Pagan, il cui p-value estremamente piccolo ci obbliga a rifiutare l'ipotesi di omoschedasticità a qualsiasi livello di significatività. Quindi, anche in questo quarto modello, i residui risultano essere eteroschedastici.

In seconda istanza, si è valutata la normalità dei residui facendo ricorso al test di Shapiro-Wilk, da cui si deduce che i residui non si distribuiscono secondo una variabile aleatoria normale.

In terzo luogo, si è verificato che i residui presentino media nulla. Il p-value del test assume un valore tendente al valore uno, che ci obbliga a non poter rifiutare l'ipotesi secondo cui la distribuzione dei residui presenta valore atteso pari a zero.

4.4.3 L'analisi della robustezza

Dato che il modello presenta dei problemi di eteroschedasticità, è stato effettuato un test di robustezza per cercare di ricondurre il modello ad una condizione di omoschedasticità, attraverso l'eliminazione di alcune osservazioni.

Grazie all'utilizzo del test di Bonferroni vengono identificati come valori outlier, e quindi eliminati dal dataset, i seguenti paesi: la Cina e la Russia nel 2000, la Grecia e la Russia nel 2004, la Russia nel 2012 ed infine il Brasile, la Cina, la Gran Bretagna, l'India e la Russia nel 2016.

È stata quindi condotta una stima di regressione utilizzando lo stesso modello, ma con il dataset di partenza dal quale sono state escluse le osservazioni di cui sopra. Dall'analisi dei risultati della regressione si nota come l'R² Adjusted del modello aumenti in modo considerevole, passando da 0,878 al valore di 0,925. Anche in questo caso il modello stimato in assenza degli outlier presenta, quindi, una maggiore capacità di spiegare la varianza campionaria del totale delle medaglie, adattandosi meglio ai dati empirici. Tuttavia, nonostante le modifiche apportate al dataset, permangono sia i problemi di eteroschedasticità sia quelli di non normalità; analogamente, non si apprezzano variazioni per quanto riguarda la nullità della media dei residui.

Come per il modello precedente, è stata sviluppata anche una seconda analisi della robustezza mossa da una valutazione qualitativa, secondo cui si è cercato di escludere l'effetto dato dalle cinque nazioni che hanno raccolto più medaglie nelle singole edizioni olimpiche considerate in questa ricerca.

Si è quindi effettuata la regressione sulla funzione lineare costruita, utilizzando però un dataset dal quale sono stati rimossi per ogni Olimpiade i cinque paesi che hanno conquistato più medaglie; dall'ispezione dell'output si osservano una diminuzione della significatività dei regressori ed una riduzione del coefficiente di determinazione lineare corretto. Infatti, l'R² Adjusted passa da 0,878 a 0,798, testimoniando una diminuzione della capacità del modello di spiegare la variabilità del totale delle medaglie.

Tuttavia, nemmeno attraverso questa seconda analisi di robustezza ha permesso di risolvere i problemi di non normalità e di eteroschedasticità dei residui, che si sono addirittura acutizzati. Non si osservano, invece, apprezzabili variazioni per quanto riguarda la nullità della media.

4.4.4 Conclusione

In definitiva, il modello costruito per mezzo della forward selection risulta essere, tra i quattro proposti, quello che meglio si adatta ai dati empirici. Tuttavia, si basa soltanto sul numero di atleti inviati ai Giochi olimpici estivi, sul numero di articoli scientifici pubblicati in un anno e su quattro regressori che risentono negativamente dell'effetto catturato da queste due variabili. Nonostante la genesi del modello affondi le radici su un approccio molto diverso da quella adottato in letteratura, in realtà non fa altro che confermare quanto asserito nei lavori scientifici pubblicati fino ad oggi.

Il modello afferma che il numero di medaglie conquistate da un paese può essere adeguatamente catturato da due sole variabili, ovvero Scientific articles_{i,t} e Athletes_{i,t}, che sono in grado di cogliere pienamente anche tutti gli altri effetti prevalenti, descritti in letteratura. Il numero di articoli scientifici è espressione sia della dimensione della popolazione sia del benessere economico e, di conseguenza, anche del livello di sviluppo del settore della ricerca e sviluppo, della medicina e dell'avanzamento tecnologico. Analogamente, la numerosità del contingente olimpico è funzione del livello di sviluppo economico, della dimensione della popolazione, dell'essere la nazione ospitante e della ricchezza dei cittadini.

Pertanto, questo modello conferma che l'host effect e il benessere economico sono due fattori determinanti e che la dimensione della popolazione rappresenta l'effetto che maggiormente influenza il numero di medaglie conquistate alle Olimpiadi estive, a conferma di quanto affermato in letteratura.

La "novità" portata da questo modello consiste nel fatto che anche il livello di sviluppo del settore della ricerca e delle scoperte scientifiche riesce a spiegare una piccola parte del successo olimpico.

Variabili indipendenti	Modelle	<u>o 1</u>	Mode		Model	lo 3	Modell	<u>o 4</u>
Constant	2,036	***	5,605	***	2,191	***	5,451	***
GDP (t) (milioni)	0,00000633	***	0,000006413	***	0,000006303	***		
Population (t) (milioni)	0,02211		0,02206	***	0,02629	***		
Host	30,52	***	30,00	***	29,88	***		
Next Host	12,99	***	13,32	***	13,51	***		
Previous Host	21,76	***	21,05	***	21,16	***		
Politics (variabile riferimento CAPME)								
CEEC	3,824	***	0,762		4,655	***		
СОМ	19,770	***	20,980	***	19,450	***		
NSCOM	-0,748		-2,260		-0,543			
TRANS	7,504	***	4,456	***	8,860	***		
Health expenditure per capita (t)	0,0008154	***	0,0001591		0,0007699	***	-0,0005291	***
Population growth (t)	-0,7181	**	-0,6322	**	-0,4747	*		
Neonathal deaths (t)	-0,00002513	***	-0,00002348	***	-0,000025420	***	-0,00000819	***
Regions (rif. Europe and Central Asia)			•		•		•	
East Asia and Pacific			-3,430	**				
Latin America and The Caribbean			-4,650	***				
Middle East and North Africa			-3,840	**				
North America			-1,890					
South Asia			-7,930	***				
Sub-Saharan Africa			-3,166	*				
Religions (rif. Christianity)								
Buddhism					-2,219			
Folk Religions					-11,860	***		
Hinduism					-5,990	**		
Islam					-1,963	**		
Judaism					-2,981			
Unaffiliated Religions					-4,919	*		
Athletes (t)							0,08849	***
Scientific articles (t)							0,0001452	***
Life expectancy (t)							-0,08143	***
Unemployment (t)							-0,08946	**
R2 Adjusted	0,7425		0,7471		0,7507		0,8784	
Numero di osservazioni	829		823		823		799	

Tab. 4.1: i determinanti del numero di medaglie conquistate ai Giochi olimpici estivi

5 Conclusione

La presente dissertazione ha messo in luce quali siano i fattori che maggiormente influenzano il numero di medaglie che un paese è in grado di conquistare ai Giochi olimpici estivi.

Dopo aver preso in esame la letteratura di settore, sono state passate in rassegna le determinanti che hanno permesso di individuare quattro modelli di regressione atti a descrivere il successo olimpico.

Sulla base dei primi tre modelli, la dimensione della popolazione si configura come un fattore estremamente significativo nel quantificare le medaglie vinte da un paese; il suo effetto viene colto sia attraverso il numero di abitanti sia per mezzo del Prodotto Interno Lordo, che rappresenta una misura economica ed al contempo dimensionale della nazione considerata. Questo risultato si dimostra in linea con quanto descritto in altri studi sul tema delle Olimpiadi, a sottolineare come una popolazione più ampia comporti non soltanto una miglior distribuzione tra i cittadini dei costi necessari per supportare adeguatamente il movimento sportivo, ma anche un bacino più ampio dal quale estrarre possibili campioni e, di conseguenza, una maggior probabilità di vincere medaglie olimpiche.

Dalle regressioni effettuate si evince come l'host effect garantisca un netto incremento nel numero di medaglie conquistate alle Olimpiadi estive. Tale effetto è contraddistinto da una dinamica temporale, secondo cui è possibile apprezzare un miglioramento delle performance sia nell'edizione che precede quella casalinga sia, in misura più consistente, in quella che la segue. Queste evidenze confermano quanto asserito in letteratura, dimostrando che ancora oggi i paesi che ospitano i Giochi olimpici estivi traggono un vantaggio in termini di medaglie conquistate. Le possibili spiegazioni di tale beneficio si possono ricondurre a fattori ambientali e climatici, al supporto del pubblico ed al trattamento di favore talvolta ricevuto dai giudici di gara.

Questa tesi sottolinea come la struttura politica ed economica assunta da un paese influenzi in modo considerevole il medagliere olimpico. Nonostante siano trascorsi molti anni dallo scioglimento dei regimi di stampo sovietico, le nazioni in questione continuano, in media, a vincere più medaglie di quanto ci si potrebbe aspettare. Il retaggio dell'impostazione sovietica nella pianificazione della preparazione sportiva continua ad influenzare le performance olimpiche di questi paesi. Anche le nazioni che adottano tutt'oggi un regime di tipo comunista sono solite godere di un consistente vantaggio in termini di medaglie conquistate. Tutto ciò a conferma di quanto suggerito dalla letteratura di settore.

Come si può notare dal secondo modello, l'appartenenza ad una diversa area geografica influenza significativamente il successo nelle discipline olimpiche. I paesi situati in Europa ed in Asia Centrale tendono a vincere di più di quanto si potrebbe predire osservando solamente le caratteristiche economico-demografiche che li contraddistinguono. Questo accade perché le diverse tradizioni culturali, relative alla pratica sportiva ed al ruolo dello sport nella società, giocano un ruolo fondamentale nel determinare il successo olimpico, come asserito dai precedenti papers scientifici che si sono occupati di questo tema. Si evince, inoltre, che i paesi Nordamericani presentano le maggiori similarità in termini di cultura sportiva con i paesi europei e dell'Asia Centrale, mentre le nazioni ricomprese nelle altre zone vincono sensibilmente meno medaglie.

Il terzo modello introduce le covariate che descrivono la principale religione professata: le nazioni a prevalenza cristiana tendono, in media, a vincere un maggior numero di medaglie rispetto ai paesi nei quali si pratica più diffusamente un altro tipo di religione, a parità di condizioni economico-demografiche. Queste variabili sociali influenzano in modo significativo le performance olimpiche: ad esempio, risulta rilevante constatare come i paesi musulmani tendano ad inviare meno atlete alle Olimpiadi e di conseguenza a conquistare meno medaglie.

Questo lavoro mette in evidenza come si possa proficuamente utilizzare la spesa sanitaria pro capite per cogliere l'influenza del livello di benessere della popolazione sul totale delle medaglie. Questo risultato conferma che, come asserito nei precedenti studi scientifici pubblicati, il livello di sviluppo economico e la distribuzione della ricchezza si dimostrano fattori estremamente rilevanti nel determinare il successo olimpico. La significatività della spesa sanitaria pro capite in tre dei quattro modelli costruiti mostra come non solo il livello di sviluppo economico, ma anche quello sanitario, influenzino le performance sportive dei paesi.

Sulla scorta delle regressioni effettuate, emerge come le variabili demografiche influenzino in misura apprezzabile il totale delle medaglie; infatti, oltre alla dimensione della popolazione, anche la crescita demografica ed il numero di morti neonatali risultano essere due determinanti significative del successo olimpico.

Dal quarto modello si evince, invece, che, come asserito da numerosi articoli pubblicati in precedenza, il numero di atleti inviati ai Giochi da una nazione risulta essere un fattore di grande rilevanza, in quanto un contingente composto da un maggior numero di unità aumenta la probabilità di conquistare un bottino di medaglie più ricco. Inoltre, l'alta correlazione di questa covariata con le altre variabili economiche e demografiche sottolinea come il numero di atleti riesca a catturare anche gli effetti sul totale delle medaglie dati dalle variabili economico-demografiche. Si può osservare che anche il numero di articoli scientifici pubblicati in un paese in un anno risulta essere un fattore estremamente significativo. Se ne può dedurre che il livello di sviluppo del settore

essere un fattore estremamente significativo. Se ne può dedurre che il livello di sviluppo del settore della ricerca influenza direttamente le performance sportive. Un paese che pubblica un elevato numero di articoli scientifici tende ad investire in modo consistente nelle attività di ricerca e sviluppo, con un conseguente beneficio per le performance ai Giochi olimpici. Inoltre, la significatività di questa variabile non fa altro che rimarcare la rilevanza del livello di sviluppo economico e della dimensione della popolazione, essendo correlata positivamente con entrambe.

Parimenti, si osserva come l'aspettativa di vita e la disoccupazione siano due fattori determinanti per il successo olimpico, in quanto sono in grado di cogliere l'effetto sulle medaglie dovuto a fattori sociali ed economici. Infatti, un'elevata disoccupazione porta ad una contrazione del Prodotto Interno Lordo, con conseguenti ripercussioni negative sugli investimenti in ambito sportivo.

Concludendo, si può affermare che questa tesi conferma e rafforza quanto asserito dalla letteratura di settore; risultano, infatti, come principali fattori significativi il livello di benessere, la dimensione della popolazione, l'host effect, la struttura politica e le influenze culturali e sociali.

I primi tre modelli costruiti, nonostante mostrino un coefficiente di determinazione lineare corretto più basso, appaiono migliori e più coerenti con la letteratura.

Il lavoro presentato suggerisce una duplice applicabilità: in prima istanza, le indicazioni suggerite dall'analisi forniscono un florido supporto alle decisioni nell'ambito delle politiche sportive, finalizzato a delineare su quali determinanti si possa far leva per migliorare le prestazioni ai Giochi olimpici estivi; in secondo luogo, i modelli presentati possono essere impiegati a scopi previsionali, ad esempio dalle società di scommesse sportive.

Nell'ambito della ricerca futura in questo settore, la metodologia presentata può fornire spunti per raffinare le capacità descrittive e predittive dei modelli finalizzati a determinare la variabilità del totale delle medaglie olimpiche conquistate da un paese.

<u>Bibliografia</u>

- D. W. Ball, "Olympic Games Competition: Structural Correlates of National Success", 1972, International Journal of Comparative Sociology
- N. Levine, "Why do Countries Win Olympic Medals? Some Structural Correlates of Olympic Games Success: 1972", 1972, Sociology and Social Research
- A. Grimes, W. J. Kelly, P. H. Rubin, "A Socioeconomic Model of National Olympic Performance", 1974, Social Science Quarterly
- W. F. Shughart II, R. D. Tollison, "Going for the Gold: Property Rights and Athletic Effort in Transitional Economies", 1993, Kyklos
- A. Sen, "Development as Freedom", 1999, New York: Random House
- S. R. Clarke, "Home Advantage in the Olympic Games", 2000, Proceedings of the Fifth Australian Conference on Mathematics and Computers in Sport
- D. K. N. Johnson, A. Ali, "Coming to Play or Coming to Win: Participation and Success at the Olympic Games", 2000, Wellesley College Dept. of Economics Working Paper
- G. H. Kuper, E. Sterken, "The Olympic Winter Games: Participation and Performance", 2001, SSRN Electronic Journal
- A. B. Bernard, M. R. Busse, "Who Wins the Olympic Games: Economic Resources and Medal Totals", 2002, Review of Economics and Statistics
- R. Hoffmann, L. C. Ging, B. Ramasamy, "Public policy and olympic success", 2002, Applied Economics Letters
- D. K. N. Johnson, A. Ali, "A Tale of Two Seasons: Participation and Medal Counts at the Summer and Winter Olympic Games", 2002, Wellesley College Working Paper
- S. Lozano, G. Villa, F. Guerrero, P. Cortés, "Measuring the performance of nations at the Summer Olympics using data envelopment analysis", 2002, Journal of the Operational Research Society
- X. Bian, "Predicting Olympic Medal Counts: the Effects of Economic Development on Olympic Performance", 2005, The Park Place Economist Volume XIII
- W. D. Pfau, "Predicting the Medal Wins by Country at the 2006 Winter Olympic Games: An Econometrics Approach", 2006, The Korean Economic Review. Vol. 22
- W. Andreff, "The correlation between economic underdevelopment and sport", 2007, European Sport Management Quarterly

- M. Andreff, W. Andreff, S. Poupaux, "Les déterminants économiques de la performance olympiques: Prévision des médailles qui seront gagnées aux Jeux de Pékin", 2008, Revue d'Economie Politique Dalloz
- A. Rathke, U. Woitek, "Economics and the Summer Olympics: An Efficiency Analysis", 2008, Journal of Sports Economics
- A. Nevill, N. Balmer, E. Winter, "Why Great Britain's success in Beijing could have been anticipated and why it should continue beyond 2012", 2009, British Journal of Sports Medicine
- M. Andreff, W. Andreff, "Economic Prediction of Sport Performances: From Beijing Olympics to 2010 FIFA World Cup in South Africa", 2010, Conference: 12th Conference of the International Association of Sport Economists, 85th Western Economic Association International Conference
- M. Andreff, W. Andreff, "Economic Prediction of Medal Wins at the 2014 Winter Olympics", 2011, International Association of Sports Economists
- A. Ceballos, R. Flores, I. G. McHale, J. Horrillo, "Explaining and forecasting national team medals totals at the Summer Olympic Games", 2011, International Journal of Forecasting
- G. Vagenas, E. Vlachokyriakou, "Olympic medals and demo-economic factors: Novel predictors, the ex-host effect, the exact role of team size, and the 'population-GDP' model revisited", 2011, Sport Management Review
- G. H. Kuper, E. Sterken, "Participation and Performance at the London 2012 Olympics", 2012, University of Groningen
- J. L. Berdahl, E. L. Uhlmann, F. Bai, "Win—win: Female and male athletes from more gender equal nations perform better in international sports competitions", 2014, Journal of Experimental Social Psychology
- J. Bredtmann, C. J. Crede, S. Otten, "Olympic medals: Does the past predict the future?", 2016, The Royal Statistical Society
- D. Forrest, I. G. McHale, I. Sanz, J. D. Tena, "An analysis of country medal shares in individual sports at the Olympics", 2016, European Sport Management Quarterly
- N. Scelles, W. Andreff, L. Bonnal, M. Andreff, P. Favard, "Forecasting National Medal Totals at the Summer Olympic Games Reconsidered", 2020, Social Science Quarterly
- S. Şahin, F. Şenduran, "Effect of sports health and exercise research on Olympic game success: An analytical and correlational survey", 2020, African Educational Research Journal

<u>Sitografia</u>

https://data.worldbank.org/

https://en.wikipedia.org/wiki/List of countries by average yearly temperature

https://ourworldindata.org/grapher/main-religion-of-the-country-in?tab=table

https://it.wikipedia.org/wiki/Giochi della XXVII Olimpiade

https://it.wikipedia.org/wiki/Giochi della XXVIII Olimpiade

https://it.wikipedia.org/wiki/Giochi della XXIX Olimpiade

https://it.wikipedia.org/wiki/Giochi della XXX Olimpiade

https://it.wikipedia.org/wiki/Giochi della XXXI Olimpiade

https://it.wikipedia.org/wiki/Medagliere dei Giochi della XXVII Olimpiade

https://it.wikipedia.org/wiki/Medagliere dei Giochi della XXVIII Olimpiade

https://it.wikipedia.org/wiki/Medagliere dei Giochi della XXIX Olimpiade

https://it.wikipedia.org/wiki/Medagliere dei Giochi della XXX Olimpiade

https://it.wikipedia.org/wiki/Medagliere dei Giochi della XXXI Olimpiade

https://ourworldindata.org

https://www.ilo.org/global/statistics-and-databases/lang--en/index.htm

https://www.imf.org/en/Data

I siti qui riportati sono stati verificati in data 29/03/2021.

<u>Ringraziamenti</u>

Ormai giunto alla fine del mio percorso di studi, ritengo doveroso dedicare questo spazio del mio elaborato alle persone che hanno contribuito, con il loro instancabile supporto, alla realizzazione dello stesso.

Un particolare ringraziamento va al Professor Buzzacchi, che ha seguito, con pazienza e grande disponibilità, la stesura di questa tesi. Lo ringrazio, inoltre, per avermi spronato e stimolato nel corso di questo percorso finale nel tentativo di realizzare un lavoro all'altezza della situazione.

Una dedica speciale va a tutta la mia famiglia, in particolare a mio nonno, Illy, ed ai miei genitori, che hanno sempre creduto in me aiutandomi a superare anche i momenti più difficili che ho dovuto affrontare durante questo percorso.

Vorrei anche ringraziare tutti gli amici, che da sempre rallegrano le mie giornate; a chi c'è dai tempi delle medie, a chi dalle elementari e a chi dalla nascita, agli amici con cui ho condiviso momenti indelebili sui campi di calcio, in uno stadio e su un ring.

Infine, vorrei dedicare questo piccolo traguardo a tutti quelli che hanno sempre creduto in me, persino quando non ci credevo più nemmeno io.

Per tutte le volte che vi ho deluso...

Appendice I

Di seguito, viene riportata la matrice di correlazione delle variabili dipendenti ed indipendenti considerate nello studio, ad eccezione delle variabili dicotomiche.

