



UNIVERSITÀ
DI PAVIA

Dipartimento di Scienze Economiche

e Aziendali Corso di Laurea

Magistrale in Economia e Gestione

delle Imprese

**Capitale umano e successo delle *startup*: il ruolo
dell'istruzione, dell'esperienza e della
composizione dei *team***

Relatore:

Chiar.ma Prof. Francesca Sanguineti

Tesi di Laurea

di Mauro Zacconi

Matr. n.544656

Anno Accademico 2024-2025



UNIVERSITÀ
DI PAVIA

Ringraziamenti

Dedico questa tesi alle persone che non mi hanno abbandonato, che hanno continuato a credere in me, supportandomi nel mio percorso universitario e di vita.

Il mio pensiero va in primis alla ragazza che amo, colei che mi ha sempre incoraggiato, accolto e sostenuto nei momenti complicati, in particolare nell'ultimo anno quando sono sorti molti dubbi in me. Grazie a lei ho ritrovato lo stimolo ad impegnarmi per le cose importanti, per noi, ed è anche merito suo se questo percorso di tesi è arrivato al termine.

Ringrazio, inoltre, la mia famiglia, mamma e papà, per avermi dato l'opportunità di proseguire i miei studi, di avermi concesso la prospettiva di un futuro migliore attraverso i loro sacrifici. Spero di poter fare altrettanto per loro.

Ringrazio gli amici che mi hanno concesso dei momenti di gioia anche quando mi sentivo solo, che hanno alleggerito le mie serate semplicemente, seduti ad un tavolo a giocare e a dirci tante sciocchezze. In un mondo che rende sempre più difficile lo stare assieme, è bello trovare qualcuno con cui alleggerire il peso sulle proprie spalle.

Un pensiero speciale è dedicato a coloro che mi hanno accompagnato negli ultimi anni in università, colleghi, compagni di viaggio, ma soprattutto amici. Voglio ricordare i momenti di leggerezza, risate, complicità durante e dopo le lezioni. Voglio ricordare i momenti nei quali ci siamo dati manforte prima e dopo gli esami e quei discorsi esistenziali sul nostro futuro che facevano riflettere.

Ringrazio chi mi ha fatto sentire accolto e mi ha sfamato più volte senza obblighi, mi ha dato consigli e incoraggiamenti, senza pretendere nulla in cambio.

Ringrazio la mia relatrice, la professoressa Sanguineti per il contributo enorme a questa tesi, per la pazienza, la grande professionalità e l'interesse che ha sempre dimostrato nei miei confronti.

Infine, ringrazio Crunchbase per avermi concesso l'opportunità di accedere alla loro banca dati e aver reso possibile questa tesi.



UNIVERSITÀ
DI PAVIA

Indice

Introduzione.....	3
Capitolo 1: Stato dell'arte.....	4
1.1 Le <i>Startup</i> ed il Capitale Umano	4
1.2 I Fondatori ed il Capitale Umano	32
Capitolo 2: Metodologia della ricerca	51
2.1 Strumenti di ricerca e raccolta dati	54
2.2 Costruzione del <i>dataset</i> e descrizione delle variabili	61
2.2.1 Variabili controllo	64
2.2.2 Variabili indipendenti	67
2.2.3 Variabile dipendente	72
Capitolo 3: Risultati analisi	74
3.1 Caratteristiche del campione scelto	74
3.2 Regressione logistica	88
Capitolo 4: Discussione e implicazioni	94
4.1 Interpretazioni	94
4.1.1 Variabili di controllo	95



UNIVERSITÀ
DI PAVIA

4.1.2	Variabili indipendenti.....	102
4.2	Effetto del capitale umano dei fondatori sul successo della <i>startup</i> 110	
4.3	Effetto della composizione del <i>team</i>	117
4.4	Implicazioni.....	121
4.4.1	Implicazioni teoriche.....	122
4.4.2	Implicazioni <i>manageriali</i>	123
4.5	Limiti della ricerca e spunti per ricerche future	124
	Capitolo 5: Conclusioni.....	127
	Bibliografia.....	132
	Appendice.....	145
	Appendice A – Codici Python	145
	Appendice B – Codici R.....	149
	Appendice C – Codici Stata.....	156



UNIVERSITÀ
DI PAVIA

Indice delle figure

Figura 1.....	10
Figura 2.....	13
Figura 3: Modello di estrazione dati.....	57
Figura 4: Schema costruzione CSV Cards	58
Figura 5: Pulizia CSV Cards	59
Figura 6: Costruzione Database Finale	61
Figura 7: Matrice di correlazione	84

Indice delle tabelle

Tabella 1: Descrizione delle Variabili.....	64
Tabella 2: Statistiche descrittive variabili	74
Tabella 3: Frequenze variabili categoriali	80
Tabella 4: Confronto campione	82
Tabella 5: Multicollinearità	89
Tabella 6: Regressione logistiche <i>Odds ratio</i>	90
Tabella 7: Goodness of fit	92
Tabella 8: Goodness of fit con percentuali.....	92



UNIVERSITÀ
DI PAVIA

Introduzione

Negli ultimi anni le *startup* hanno assunto un ruolo principale quando si parla di innovazione e sono diventate attori chiave nelle iniziative imprenditoriali (Bordiga et al., 2023). La letteratura d'impresa ha spostato l'attenzione dalle risorse tangibili a quelle intangibili, riconoscendo il capitale umano dei fondatori tra gli asset più rilevanti nel ciclo di vita delle *startup*, a causa dell'incertezza che caratterizza gli ambienti di sviluppo di queste società. Alcune dimensioni sono indicate in alcuni contributi teorici ed empirici; tra esse figurano la formazione universitaria e l'esperienza lavorativa, che hanno un'influenza rilevante sulle prestazioni (Colombo & Grilli, 2005). Parallelamente, anche la dimensione collettiva del capitale umano è stata analizzata dalla ricerca che ha sottolineato l'importanza della composizione e delle dinamiche interne ai *team* imprenditoriali (Beckman & Burton, 2008; Jin et al., 2017).

I risultati disponibili in letteratura non sono omogenei, così come le diverse definizioni adottate. In particolare, una mancanza evidente riguarda il ruolo congiunto delle caratteristiche individuali nel determinare il successo delle *startup* e nella definizione di queste stesse. Il successo è definito in questa tesi come il raggiungimento di uno stadio avanzato nel ciclo di finanziamento o il conseguimento di un'*exit* positiva. Questo approccio rimane coerente con la letteratura, che identifica le *IPO* e le *M&A* come le operazioni principali di



UNIVERSITÀ
DI PAVIA

creazione e realizzazione del valore per fondatori e investitori (Certo et al., 2009; Marchetti & Manelli, 2021). Questi eventi sono esiti osservabili e misurabili come liquidazione del *business* o raggiungimento della maturità dell'impresa.

Dalle premesse indicate nasce la domanda di ricerca: *“In che misura il capitale umano dei fondatori e la composizione dei team influenzano la probabilità che una startup raggiunga un'exit positiva?”*.

L'analisi empirica si fonda su dati secondari provenienti dal database Crunchbase, ampiamente utilizzato negli studi sull'imprenditorialità e sull'ecosistema *startup* (Sanguineti et al., 2024), con l'obiettivo di adottare un approccio quantitativo. Attraverso l'analisi delle relazioni tra alcune dimensioni del capitale umano e le caratteristiche strutturali dei *team* di fondatori, si cerca di comprendere la composizione complessiva dei gruppi e il loro impatto sul successo aziendale, offrendo evidenze utili a chiarire il ruolo dei fattori considerati.

Capitolo 1: Stato dell'arte

1.1 Le *Startup* ed il Capitale Umano

“Sebbene possa sembrare che articoli e ordinamenti giuridici dei vari paesi stiano già discutendo e definendo cosa sia una startup, in particolare una startup innovativa, non è chiaro come il concetto sia definito nella letteratura



UNIVERSITÀ DI PAVIA

accademica” (Magalhães & Oliveira, 2019, p. 2). Alcune definizioni pongono l’accento sul numero di anni intercorsi dalla fondazione dell’impresa, altre si concentrano sulla dimensione aziendale, stabilendo un numero massimo di dipendenti, altre ancora privilegiano il potenziale di sviluppo e i modelli di *business* innovativi (Tomada Letizia, 2022). La varietà di criteri evidenzia la complessità intrinseca del fenomeno: non esiste un consenso unanime su come definire le *startup* (Ehsan, 2021).

Magalhães e Oliveira (2019) hanno condotto una *scoping review* per mostrare come il significato di *startup* sia variegato nella letteratura accademica. La revisione ha adottato un approccio qualitativo di *content analysis* su oltre 1.000 articoli statunitensi pubblicati tra il 2000 e il 2019, con particolare riferimento ai settori *Politics and Law* e *Management and Business*. Le teorie su cui si basano le osservazioni sono la *“creative destruction”* di Schumpeter e il ruolo dell’imprenditore come implementatore dell’innovazione, nonché l’evoluzione tecnologica e la concentrazione geografica in specifici cluster (es. Silicon Valley). I risultati da loro ottenuti confermano l’assenza di una definizione condivisa e mostrano una forte convergenza sul ruolo dell’innovazione come elemento distintivo. Le definizioni di natura istituzionale, come quella italiana, sottolineano questo aspetto, qualificando le *startup* come imprese altamente tecnologiche e evidenziando il legame tra innovazione e



UNIVERSITÀ DI PAVIA

sviluppo imprenditoriale (Bordiga et al., 2023). Il Decreto-Legge n. 179/2012 afferma: “*La startup innovativa è una società di capitali, costituita anche in forma cooperativa. Ha caratteristiche molto simili a quelle di una PMI innovativa (piccole e medie imprese)*” (Bordiga et al., 2023, p. 8). Questa prospettiva mette in rilievo la dimensione politico-istituzionale del concetto, evidenziando come dipenda sia dalle caratteristiche industriali sia dal quadro normativo.

Lo studio di Magalhães presenta alcune limitazioni metodologiche: l'unità di analisi, costituita dai singoli articoli, e le variabili considerate si limitano alla presenza o all'assenza di una definizione e ai criteri classificatori adottati (innovazione, crescita, tecnologia). La correlazione tra le relazioni e le *performance* non può essere valutata, poiché non è adottato alcun modello statistico. Nonostante tali limiti, quanto osservato rimane interessante nel dimostrare che, a fronte di una pluralità di approcci teorici, solo una quota ridotta della letteratura giuridica (circa il 40%) fornisce una definizione esplicita; inoltre, essa rimane prevalentemente ancorata al paradigma dell'innovazione tecnologica.

L'ambiguità concettuale è ribadita anche nel recente quadro normativo europeo. Secondo la Commissione Europea, una PMI è definita come un'impresa con meno di 250 dipendenti. Nell'ambito dell'AI Act dell'Unione Europea, le



UNIVERSITÀ DI PAVIA

startup non sono considerate una forma d'impresa autonoma, bensì una sottocategoria delle piccole e medie imprese (SMEs). Nonostante la legislazione sia chiara, non sempre riflette le peculiarità strutturali e le dinamiche tipiche delle *startup*; ad esempio, non è ben definito se il termine possa includere anche le imprese che superano i limiti dimensionali previsti (Tomada Letizia, 2022). Analogamente al quadro normativo italiano, il legislatore europeo non qualifica una forma giuridica distinta, ma ricomprende le *startup* tra le categorie societarie esistenti e continua a rifarsi a elementi sostanziali, quali l'innovazione tecnologica e l'alto potenziale di crescita.

L'inquadramento normativo si rivela inadeguato alla presente tesi, poiché non fornisce strumenti adeguati per analizzare il concetto di successo e non coglie le dinamiche evolutive e le traiettorie di sviluppo tipiche di queste imprese, rendendo necessaria l'adozione di una definizione operativa più coerente con il percorso che intraprendono.

Le articolazioni che caratterizzano il funzionamento di una neonata impresa favoriscono lo sviluppo di attività innovative; l'inventore non percepisce un rischio elevato di appropriazione indebita, il che incoraggia la sua attività innovativa (Gilson, 2010). Tuttavia, le *startup* operano in condizioni di elevata incertezza e necessitano di modelli di finanziamento specifici per mitigare i rischi cui sono esposte. Di solito, l'arco di sviluppo comprende un periodo



UNIVERSITÀ DI PAVIA

iniziale, uno di avvio, una fase di crescita ed espansione e, infine, un consolidamento di maturità o un'uscita e un successo (Halt et al., 2017). Inizialmente, un considerevole numero di investitori fornisce un supporto economico, sotto forma di capitale e di acquisizione di partecipazioni societarie, da privati con elevata disponibilità economica che svolgono il ruolo di *business angels* o da incubatori, ovvero società che forniscono supporto alle imprese nelle fasi iniziali, fungendo da consulenti e finanziatori (Marchetti & Manelli, 2021).

Attraverso un'analisi basata sulle fasi di sviluppo, Reisdorfer-Leite e colleghi (2020) vorrebbero proporre una definizione preliminare di *startup*, basata su un approccio concettuale che mette in relazione il ciclo di vita con il modello di *Product Lifecycle Management* (PLM). L'obiettivo dello studio è delineare il periodo temporale in cui un'impresa può essere definita come *startup*, attraverso un'analisi comparativa tra l'evoluzione dell'impresa e quella dei prodotti, con particolare attenzione alle fasi di *Beginning of Life* (BOL) e *Middle of Life* (MOL). La ricerca concettuale si fonda sulla revisione della letteratura, senza ricorrere a metodologie empiriche. Il modello teorico classifica tre stadi principali: *pre-startup* (ricavi negativi, autofinanziamento, fase BOL), *startup* (crescita rapida, investimenti di *venture capital*, fasi BOL-MOL) e *consolidated enterprise* (crescita moderata, *IPO* o *M&A*, MOL maturo). La teoria applicata integra i modelli life-cycle delle imprese, originariamente



UNIVERSITÀ DI PAVIA

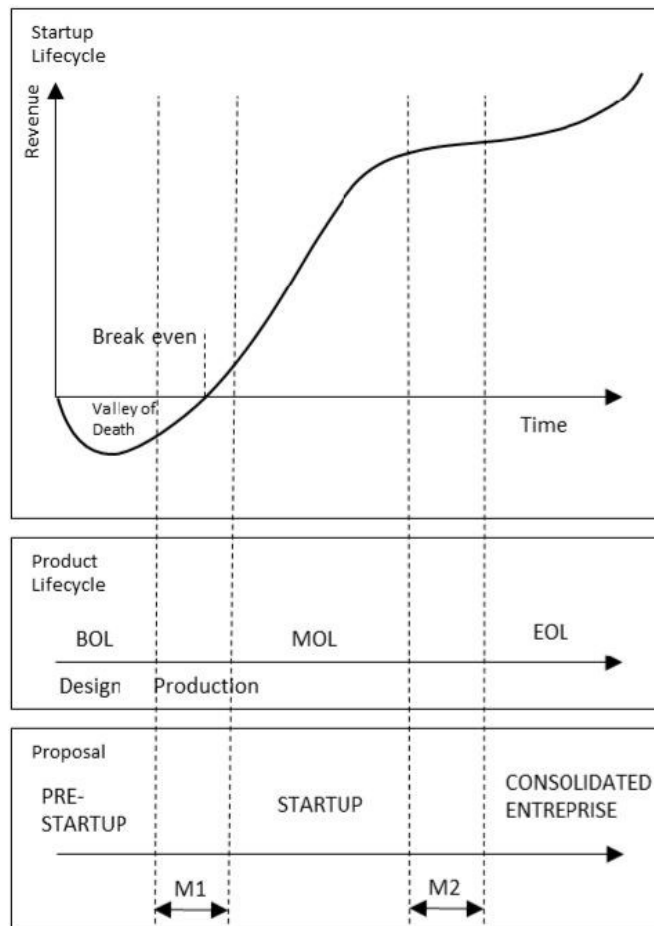
sviluppati per le PMI (Churchill & Lewis, 2000), i contributi più recenti sugli *Startup ecosystems* (Greco, 2023) e la letteratura sul PLM, inteso come la gestione integrata del prodotto lungo l'intero ciclo di vita (Ferreira et al., 2017).

Come indicato in Figura 1, gli autori individuano lo “stadio di *startup*” nel periodo compreso tra il superamento del *Valley of Death* (M1, punto di break-even) e il plateau di crescita dei ricavi (M2). Tale fase è associata alle transizioni BOL-MOL del prodotto ed è caratterizzata da una crescita accelerata dei ricavi, sostenuta da investimenti esterni e dal rafforzamento della *knowledge base* in ambito PLM.



UNIVERSITÀ
DI PAVIA

Figura 1



Fonte: tratto da Reisdorfer-Leite et al., (2020).

Come sottolineato da Bordiga e colleghi (2023), tale approccio consente di analizzare l'andamento di queste società non solo in termini di crescita dimensionale, ma anche in relazione alle fonti di capitale necessarie a sostenere lo sviluppo, evidenziando le esigenze finanziarie e le sfide gestionali associate a ciascun stadio.



UNIVERSITÀ DI PAVIA

Soffermandosi sul panorama letterario del finanziamento (Arena et al., 2018; Giraud et al., 2019), il modello illustra come l'accesso a strumenti finanziari differenti, dal *pre-seed* all'*exit*, influisca direttamente sulla capacità della *startup* di superare le barriere di crescita e di innovazione. In questo senso, è utile per identificare le fasi critiche e definire il concetto di successo in termini di transizione verso eventi di *exit* osservabili; quindi, il modello proposto da Reisdorfer-Leite e colleghi (2020) è stato adottato nella presente tesi come strumento concettuale per delimitare i momenti dell'evoluzione della *startup* e per identificare il concetto di "successo": evento osservabile nella fase di finanziamento avanzato e di transizione verso un'*exit* positiva.

L'evoluzione può essere interpretata come un processo dinamico, influenzato dall'interazione tra imprenditori e investitori, che si sviluppa attraverso i finanziamenti. Secondo la letteratura, la cooperazione reciproca tenderebbe a uniformare gli incentivi alla produzione di innovazione. L'investimento, che tipicamente è strutturato in fasi successive e caratterizzato da un elevato grado di incertezza, può avere esito negativo. Gli investitori che partecipano alle fasi iniziali, quindi, non sono necessariamente coinvolti in quelle successive; pertanto, si crea un forte impulso a migliorare le *performance* operative e innovative al fine di attrarre nuovi capitali e proseguire nel percorso di sviluppo (Sahlman, 2022); ciò porterebbe al successo.



UNIVERSITÀ DI PAVIA

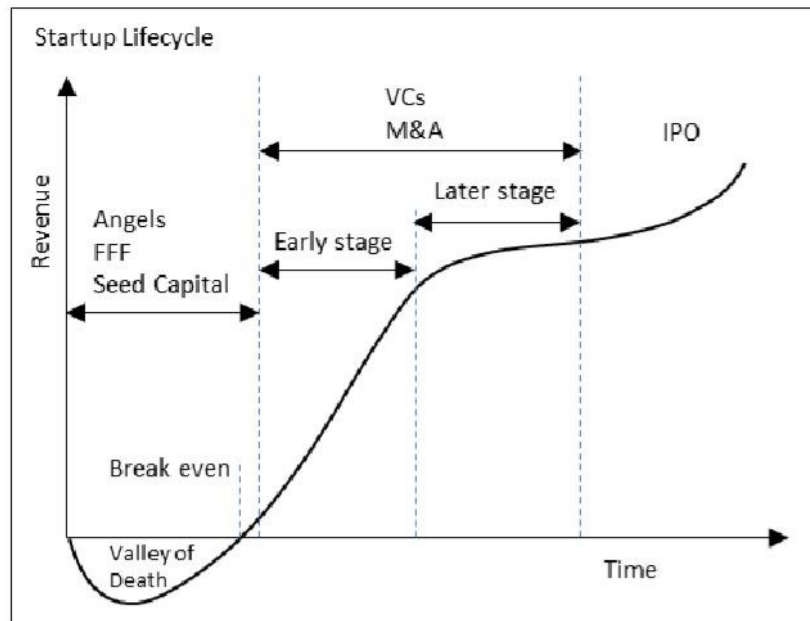
Elbahjaoui ed Elabjani (2021) forniscono prove qualitative a sostegno della tesi proposta, esaminando il processo di incubazione come strumento di innovazione mediante interviste semi-strutturate. Gli autori osservano che nelle fasi iniziali le *startup* dipendono in misura significativa da finanziamenti *pre-seed* e *seed* da parte di figure quali *friends, family and fools, business angels* e incubatori. In assenza di ricavi e con un profilo di rischio particolarmente elevato, tipico delle fasi di ideazione e prototipazione, tale dipendenza espone l'impresa a un rischio accentuato nella fase post-incubazione, soprattutto in caso di interruzione dei flussi di finanziamento o di inefficienze nella gestione.

È in questo stadio che emerge il rischio del cosiddetto *Valley of Death* (VoD), inteso come punto critico del *life-cycle* in cui l'impresa può non essere in grado di sostenere i costi operativi prima di raggiungere livelli di ricavo sufficienti. Nel modello finanziario, tale dinamica è rappresentata dall'andamento dei flussi di entrata e dalla progressiva transizione tra diverse fonti di capitale, secondo la curva a forma di S, inclinata in avanti, come illustrato in Figura 2. Il superamento di questa fase costituisce una condizione necessaria per l'accesso agli stadi successivi di crescita e, in ultima istanza, per il raggiungimento di un'*exit* positiva.



UNIVERSITÀ
DI PAVIA

Figura 2



Fonte: Adattato da web: *Comprendere il ciclo di vita di una startup aziendale*¹

Il ciclo di vita del finanziamento consente di definire la fase di *startup* come il periodo compreso tra il superamento del VoD e il raggiungimento della maturità, come illustrato in Figura 1. Il superamento della *Valley of Death* rappresenta uno snodo critico nel percorso evolutivo di queste imprese, poiché coincide con i primi anni di attività, caratterizzati da flussi di cassa negativi, elevata incertezza e rischio significativo di fallimento legato alla mancanza di

¹<https://www.zdnet.com/article/understanding-the-enterprise-Startup-lifecycle/> data visione 15/10/2025



UNIVERSITÀ DI PAVIA

sostenibilità finanziaria (Marchetti & Manelli, 2021; Zapata-Molina et al., 2025).

Un ulteriore contributo significativo è fornito da Gbadegeshin e colleghi (2022), che propongono il *Buztech Startup Model* (BSM). L'intento è fornire un framework per superare il VoD nelle *startup* ad alta intensità tecnologica. Il *Buztech Startup Model* si articola in blocchi interdipendenti e viene proposto come strumento flessibile e iterativo per accompagnare le *startup* neonate nelle fasi *pre-startup*, *startup* e *growth*. Gli autori partono dalla constatazione che la letteratura è frammentata e incentrata esclusivamente sulle dimensioni finanziarie, trascurando la natura multidimensionale delle *startup*. Non vengono formulate ipotesi statisticamente testabili; l'analisi risponde a quesiti di ricerca relativi alla definizione e alle cause del VoD, nonché alle strategie di mitigazione. Gli studi su cui si basa lo studio riguardano i principali gap dalla fase di ricerca alla commercializzazione dei prodotti, in linea con i modelli classici che evidenziano la contrazione delle risorse tra R&D e mercato (Markham, 2002): la "*Valley of Death*" sarebbe attribuibile alla presenza concomitante di vari fattori che contribuiscono in modo significativo a incrementare il rischio di insuccesso nelle *early stage*.

Il lavoro da loro proposto presenta una revisione dello stato dell'arte ed evidenze empiriche raccolte tramite interviste qualitative a *startup* tecnologiche



UNIVERSITÀ DI PAVIA

europee attive prevalentemente in contesti *high-tech*. L'analisi è ulteriormente validata da un *pilot* su tre *pre-startup* e consente agli autori di sviluppare un modello strutturato che interpreta la *Valley of Death* non come un singolo evento critico, bensì come un insieme di sfide che si manifestano lungo diverse dimensioni: vincoli finanziari, carenze nelle competenze manageriali e imprenditoriali, limitate capacità di collaborazione e di gestione del rischio.

Le evidenze emerse indicano che il superamento di questo momento è favorito dalla presenza di *team* fondatori caratterizzati da competenze diversificate e complementari, dall'accesso a fonti di finanziamento multiple, dall'avvio precoce delle attività di commercializzazione e dalla capacità di attivare *network* collaborativi con attori esterni.

In questo studio sono presenti limiti legati sia alla focalizzazione sulle sole società tecnologiche sia alla natura qualitativa dell'analisi empirica; tuttavia, quanto visto è interessante per la presente tesi. L'affermazione secondo cui il superamento della "*Valley of Death*" non sia esclusivamente determinato dalla disponibilità di capitale finanziario, ma sia significativamente influenzato dalla qualità del capitale umano e dalla composizione delle squadre di imprenditori merita un'attenta riflessione. In secondo luogo, il modello evidenzia come le competenze costituiscano un fattore chiave nel facilitare l'accesso ai



UNIVERSITÀ DI PAVIA

finanziamenti e nel favorire la transizione verso fasi avanzate, fino al potenziale raggiungimento di un'*exit* positiva.

Il superamento della *Valley of Death* è una condizione necessaria, ma non sufficiente, per il successo dell'impresa. Il raggiungimento del *break-even point* segna il termine di questa fase e l'ingresso nello "stadio di *startup*" propriamente detto, in cui "l'attenzione si sposta sull'implementazione del modello di business e sulla costruzione delle risorse organizzative necessarie a sostenere la crescita iniziale" (Pisoni & Onetti, 2018, p. 27). La prima parte di tale stadio è comunemente identificata come *Introduction Stage* o *Early Stage*. In questa fase, la proprietà si concentra sulla formazione del *team* imprenditoriale, sullo sviluppo del *minimum viable product* e sulla validazione dell'offerta attraverso l'interazione con il mercato, con l'obiettivo di attrarre ulteriori investimenti e ridurre progressivamente il rischio operativo (Salamzadeh & Kawamorita, 2015; Bordiga et al., 2023). Il rischio rimane elevato, ma tende a diminuire man mano che l'impresa si avvicina al *product-market fit*. Una volta raggiunto tale equilibrio, il focus si sposta sulla gestione della liquidità e sull'allocazione strategica degli investimenti (Reisdorfer-Leite et al., 2020).

Il successivo *Growth Stage* o *Later Stage* (Figura 2) è caratterizzato da un'espansione accelerata della base clienti, dall'aumento delle vendite e dal



UNIVERSITÀ DI PAVIA

consolidamento della posizione competitiva, spesso in mercati già presidiati da imprese *incumbent* (Aldaeej, 2019; Marchetti & Manelli, 2021). L'impresa investe in modo significativo in *marketing*, risorse umane e processi organizzativi, con l'obiettivo di scalare il modello di *business* e massimizzare il valore aziendale in vista di un'eventuale *exit* strategica (Salamzadeh & Kawamorita, 2015). In questa fase, il ruolo dei *venture capital* assume una rilevanza cruciale, poiché i finanziamenti erogati nei round di Serie A, B e C non si limitano alla mera fornitura di capitale, ma si estendono anche alla fornitura di competenze manageriali e di supporto strategico. Tali elementi si rivelano fondamentali per sostenere la crescita dell'impresa e prepararla alle fasi successive del ciclo di vita (Marchetti & Manelli, 2021). Una volta superato il punto M2 (Figura 1), la *startup* si trova di fronte a un bivio strategico: da un lato, l'uscita dal mercato attraverso una fusione o un'acquisizione (*M&A*) o un'offerta pubblica iniziale (*IPO*); dall'altro, la prosecuzione della crescita come impresa matura orientata alla stabilizzazione di lungo periodo (Bordiga et al., 2023). L'*IPO* è il processo di quotazione in borsa di una società che, a seguito di tale operazione, transita da struttura privata a struttura pubblica sul mercato. Tale processo consente ai fondatori e agli investitori iniziali di monetizzare le proprie quote e raccogliere capitali significativi per le future espansioni. Il percorso riveste un ruolo importante, in quanto consente la diversificazione delle partecipazioni azionarie per gli *shareholders* e l'uscita, anche parziale,



UNIVERSITÀ DI PAVIA

dall'azienda, favorendo ulteriori acquisizioni di risorse esterne per i nuovi investimenti. Le *M&A* si configurano come processi di acquisizione aziendale o di fusione da parte di un'entità più grande nei confronti di una più piccola, al fine di ottenere tecnologia, competenze o accesso al suo mercato di riferimento (Bordiga et al., 2023).

Solo una parte delle *startup* che accedono alla fase di crescita riesce a completare il proprio percorso attraverso un evento di *exit*, che rappresenta il principale meccanismo di creazione e di realizzazione del valore per fondatori e investitori.

IPO e *M&A* rappresentano esiti privilegiati, poiché consentono di massimizzare il ritorno per fondatori e investitori. In entrambi i casi, l'evento di *exit* consente la liquidazione delle partecipazioni e la monetizzazione del valore generato nel corso della vita dell'impresa. Nell'ambito accademico, lo studio di Pisoni e Onetti (2018) sulle strategie di uscita delle *high-tech* in Europa e negli Stati Uniti ha fornito un contributo rilevante alla comprensione di tali fenomeni come esiti positivi, non come assimilabili al fallimento. Contribuisce a superare la visione tradizionale dell'*exit* come evento negativo, rafforzando l'interpretazione secondo cui è un indicatore di successo. L'articolo rimarca che un'uscita non implica la cessazione dell'attività imprenditoriale, bensì un cambiamento di controllo e l'integrazione della *startup* all'interno di



UNIVERSITÀ DI PAVIA

organizzazioni più strutturate. Questo punto di vista si allinea a un approccio che la considera una componente essenziale e dinamica del processo imprenditoriale, nonché tra gli esiti più attesi.

L'articolo si fonda su un vasto set di dati raccolti nel periodo compreso tra il 2012 e il 2016 e, mediante un'accurata analisi empirica, propone di esaminare il ruolo delle *exit* come principale modalità di disinvestimento e di valorizzazione economica per le *startup* innovative, i loro fondatori e gli investitori. Lo studio è coerente con il modello del *Life Cycle for Entrepreneurship* estrapolato dal *World Economic Forum* del 2014, distinguendo le *exit* volontarie da quelle fallimentari. Tali attività rappresenterebbero il conseguimento di una strategia finanziaria, attraverso la quale si realizzerebbe il valore generato nelle fasi di sviluppo. Questa interpretazione sarebbe particolarmente rilevante per le imprese *venture-backed*, per le quali costituirebbe un obiettivo centrale.

I risultati empirici ottenuti sottolineano una forte concentrazione delle *exit* negli Stati Uniti rispetto all'Europa, attribuibile alla presenza di asimmetrie informative e alla *liability of foreignness*, che conferiscono alle operazioni cross-border una maggiore complessità. Un ulteriore risultato di rilievo riguarda l'età delle *startup* al momento dell'acquisizione: una quota significativa delle *exit* avviene in fasi relativamente precoci del ciclo di vita suggerendo che anche



UNIVERSITÀ DI PAVIA

l'*early exit* possa rappresentare una scelta strategica coerente con obiettivi di crescita e valorizzazione.

La fase finale, identificabile oltre il punto M2, denominata *Maturity Stage*, è caratterizzata dalla stabilizzazione dei ricavi e da una riduzione dell'intensità della crescita, con un orientamento prevalente all'ottimizzazione delle attività esistenti anziché all'espansione esponenziale (Reisdorfer-Leite et al., 2020; Orellana, 2023).

Come si evince dai precedenti studi, il risultato positivo ha natura multidimensionale; tuttavia, in questa tesi è considerata la capacità di realizzare un *exit* strategico, quale un'*IPO* o una *M&A*, o il raggiungimento di una fase di maturazione avanzata nel ciclo di vita del finanziamento (Pisoni & Onetti, 2018). Questa prospettiva consente di analizzare le dinamiche di investimento e di valutare, nelle diverse fasi, il ruolo del capitale umano all'interno di esse. Si ricorda che l'obiettivo è evidenziare come le dimensioni dello *Human Capital* e della composizione dei *team* possano influenzare gli esiti considerati positivi.

Le *IPO* sono spesso considerate un traguardo di grande prestigio; indicano che la startup ha raggiunto una scalabilità e una governance aziendale sufficientemente robuste da operare in un mercato regolamentato. Questi processi, tuttavia, sono complessi e costosi e richiedono un'attenta pianificazione per la conformità a stringenti requisiti normativi e di trasparenza (Bordiga et al.,



UNIVERSITÀ DI PAVIA

2023; Pisoni & Onetti, 2018). Dal punto di vista dei fondatori, le *IPO* offrono l'opportunità di massimizzare il valore delle loro quote e di ottenere liquidità, sebbene spesso comportino la diluizione della loro proprietà e un maggiore scrutinio pubblico, come emerge dall'articolo di Montanaro e colleghi (2022). Il saggio esamina i determinanti del valore di uscita nelle *startup* tecnologiche europee supportate da *venture capital*. Gli autori focalizzano l'attenzione sul valore economico generato dall'evento, interpretandolo come misura concreta del successo imprenditoriale.

L'analisi è condotta su un campione di 107 *high-tech VC-backed* che hanno realizzato un *exit* tramite *IPO* o *M&A* tra il 2010 e il 2017 e integra tre dimensioni centrali del processo: la strategia adottata, la struttura degli investitori e il capitale umano dei fondatori. Grazie a questa impostazione è possibile collegare sistematicamente le decisioni imprenditoriali e di governance al risultato finale in termini di creazione e realizzazione del valore. I risultati evidenziano che le *IPO* sono associate a valori economici significativamente più elevati rispetto alle acquisizioni e appaiono come una forma più orientata alla massimizzazione del profitto, in linea con gli obiettivi degli *stakeholder* che perseguono strategie di crescita e di visibilità a lungo termine. In relazione agli obiettivi di questa tesi, si evidenzia l'apporto significativo dello studio al dibattito sul capitale umano dei fondatori. I risultati evidenziano una correlazione positiva



UNIVERSITÀ DI PAVIA

e statisticamente significativa tra il livello medio di istruzione dei fondatori e il valore di uscita, confermando le previsioni della *Human Capital Theory*. L'istruzione formale sembra promuovere lo sviluppo di strategie di crescita efficaci, l'attrazione di investitori e la gestione di processi di *exit* complessi.

Anche questo contributo favorisce la visione del successo di una *startup* secondo la definizione della presente tesi. Questa evidenza supporta l'idea che il capitale umano dei fondatori non influisce solo sulla sopravvivenza o sulla crescita dell'impresa, ma anche sulla probabilità e sulle condizioni economiche di un'*exit*, con implicazioni dirette per le decisioni di investimento e di governance.

Dall'altro lato, le *M&A* rappresentano un'alternativa frequente; infatti, consentono ai fondatori di cedere il controllo dell'azienda a un'entità più grande in cambio di benefici economici e strategici (Suzuki et al., 2024). Queste operazioni sono le strategie più diffuse, specialmente quando si tratta di acquisizione (Bordiga et al., 2023). La *review* sistematica di Suzuki e colleghi (2024) compara i fattori che influenzano la scelta tra offerta pubblica iniziale (*IPO*) e fusione o acquisizione (*M&A*). Questo studio si inserisce nella letteratura che interpreta le *exit strategy* come esiti positivi per le *startup* tecnologiche. È messa in luce la scelta tra *IPO* e *M&A* e come essa rifletta



UNIVERSITÀ DI PAVIA

l'equilibrio tra la valorizzazione del capitale umano e la capitalizzazione d'impresa.

Lo studio si distingue per l'approccio integrato e la prospettiva internazionale e offre una revisione organica della letteratura frammentata. Gli autori organizzano i determinanti delle decisioni di *exit* in due macroaree: i fattori di contesto del mercato e le caratteristiche incentivanti dei principali *stakeholders* coinvolti. La scelta del tipo di uscita non è considerata un evento isolato, bensì l'esito di un processo decisionale complesso, influenzato da condizioni esterne e risorse interne all'impresa. Con riferimento a quest'ultime, viene attribuito un ruolo centrale al capitale umano, considerando l'intensità delle competenze incorporate nel *founder team*. Un livello elevato di questa è associato a una maggiore attrattività per i potenziali acquirenti: consentirebbe di trasferire conoscenze critiche e capacità di innovazione all'interno di essi. Le operazioni di *M&A* rappresentano un meccanismo di trasferimento di asset tecnologici e strategici ed uno strumento di integrazione di capitale umano altamente qualificato. La *review* evidenzia che le conseguenze per i fondatori spesso si limitano alla riduzione dei diritti di controllo e alla perdita di specifici *private benefits*. Per questo motivo, l'*IPO* sarebbe relativamente più attrattiva in presenza di asimmetrie informative contenute e mercati finanziari favorevoli. Tuttavia, le acquisizioni possono costituire una strategia per



UNIVERSITÀ DI PAVIA

massimizzare il valore del capitale umano imprenditoriale, nei contesti caratterizzati da elevata competizione e rapidi cicli tecnologici, favorendo l'accesso a risorse difficilmente raggiungibili in autonomia. *Le operazioni di M&A potrebbero agevolare la mobilità dei fondatori verso nuove iniziative imprenditoriali.*

Dopo aver osservato la *startup* nella sua identità come soggetto in evoluzione attraverso il modello del *life-cycle* e aver definito il concetto di “successo”, prima di procedere a un'analisi della letteratura più mirata nei confronti dei fondatori e delle loro caratteristiche, è necessario approfondire il concetto di capitale umano (HC) per comprendere al meglio la natura degli elementi osservati e come essi possano essere delineati per rispondere alla domanda di ricerca.

HC si è evoluto e modificato nel tempo all'interno del pensiero economico, partendo dal confronto con il capitale fisico fino a definizioni più complesse che lo identificano come una variabile composita non osservabile legata all'efficienza produttiva e al reddito. L'essenza del capitale umano è compresa in modo ambiguo nelle fonti: talvolta è intesa come la persona stessa, intesa come insieme di caratteristiche, o talvolta solo come la sua capacità di lavorare (Khaykin et al., 2020). Nel loro articolo *The Evolution of Human Capital Theory*, Mayilyan e Yedigaryan (2022) analizzano i passaggi storici



UNIVERSITÀ DI PAVIA

della *Human Capital Theory* adottando un approccio descrittivo e induttivo sviluppando un'analisi storica e comparativa della letteratura economica attraverso le principali trasformazioni teoriche in relazione ai mutamenti dei sistemi economici e sociali.

Gli autori propongono una classificazione articolata in cinque fasi cronologiche, partendo dai contributi degli economisti classici, tra cui William Petty, Adam Smith e David Ricardo, che attribuirono un ruolo centrale al lavoro, alle competenze e alle abilità umane nella determinazione della ricchezza delle nazioni. In particolare, Adam Smith fu il primo a introdurre esplicitamente il concetto di capitale umano nel XVIII secolo, nella sua opera *La ricchezza delle nazioni*. Smith riconobbe l'investimento nell'istruzione e nello sviluppo delle competenze come una forma di capitale in grado di accrescere la produttività individuale e, di conseguenza, il benessere economico complessivo di una nazione.

Le teorie del capitale umano considerate successivamente sono quelle formalizzate dalla scuola economica americana, in particolare i contributi di Theodore Schultz e di Gary Becker. Questi autori hanno sistematizzato il concetto dimostrando come gli investimenti in istruzione, formazione e salute contribuiscano ad aumentare la produttività e i guadagni individuali, esercitando un impatto positivo sulla crescita economica aggregata. Nell'opera *Capital*



UNIVERSITÀ DI PAVIA

Formation by Education, Schultz (1960) sottolinea l'importanza di HC per il progresso economico, soprattutto nei paesi in via di sviluppo. A tal proposito, afferma: *“Il concetto di capitale deriva da qualcosa che esiste realmente e che ha la capacità economica di fornire servizi di certo valore in futuro. È possibile distinguere il capitale umano dal non umano”*. Gary Becker, premio Nobel per l'economia, estese l'analisi applicando i principi dell'economia all'investimento in HC e alle sue implicazioni per il comportamento individuale e per le dinamiche del mercato, includendo ambiti quali la discriminazione, il matrimonio e l'educazione. Come osservato da Mahssouni e colleghi (2022), a pagina 3: *“tale approccio ha permesso di quantificare il rendimento degli investimenti in capitale umano, enfatizzando l'impatto dell'istruzione formale sulla capacità di guadagno degli individui”*. In particolare, nelle opere *Investing in Human Capital* (1962) e *Human Capital: Theoretical and Empirical Analysis* (1964), Becker definisce il capitale umano come l'insieme delle abilità e delle conoscenze acquisite dagli individui attraverso investimenti in istruzione, formazione sul posto di lavoro e altre esperienze, evidenziando come tali investimenti generino ritorni economici tangibili (Unger et al., 2009). Egli considera il capitale umano inseparabile dal suo possessore, cioè dal lavoratore stesso.



UNIVERSITÀ DI PAVIA

Riprendendo l'articolo sull'evoluzione del capitale umano, Mayilyan e Yedigaryan (2022) evidenziano gli sviluppi teorici emersi con il passaggio alle economie post-industriali e della conoscenza: Daniel Bell e Manuel Castells analizzarono il ruolo del capitale umano in sistemi economici caratterizzati dalla crescente importanza dei servizi, delle reti informative e dell'innovazione tecnologica. Analogamente si colloca anche il contributo di M. M. Kritsky (1991). Egli sostenne che l'investimento nel capitale umano rappresenta un fattore chiave per lo sviluppo economico, non solo a livello individuale, ma anche a quello macroeconomico, incidendo sulla crescita e sulla competitività delle nazioni. Kritsky definisce il capitale umano come *“una forma universale e concreta di attività umana che assimila le forme precedenti (consumo e produzione) adeguate alle epoche dell'economia di appropriazione e di produzione, e la realizza come risultato del movimento storico della società umana fino al suo stato attuale”* (Khaykin et al., 2020). Più di recente, autori come James Heckman hanno ampliato la nozione di capitale umano, includendo aspetti quali le competenze non cognitive, gli investimenti precoci e il ruolo delle istituzioni. In tal senso, *“questa evoluzione ha consolidato la comprensione che le competenze, le conoscenze e le abilità individuali non sono solo attributi personali, ma risorse strategiche che influenzano direttamente la produttività aziendale e la crescita economica degli stati”* (Blair, 2011).



UNIVERSITÀ DI PAVIA

Le teorie moderne sul capitale umano, pertanto, non si limitano più a considerare istruzione, formazione ed esperienza come unici fattori determinanti, ma integrano anche la salute, il benessere psicofisico e le relazioni sociali come componenti essenziali in grado di accrescerne il valore intrinseco, l'impatto economico e la capacità di adattamento, nonché la creatività e l'intelligenza emotiva tra le sue componenti fondamentali. Tali elementi sono evidenti nei contesti imprenditoriali innovativi, come le *startup*, già richiamati da Becker (1975) nella sua pubblicazione intitolata *"Human Capital: A Theoretical and Empirical Analysis"*.

Le fasi evolutive della teoria del capitale umano non si escludono a vicenda, come osservano Mayilyan e Yedigaryan (2022), ma tendono piuttosto a sovrapporsi. Questo rifletterebbe i cambiamenti strutturali dei sistemi economici. Il capitale umano si configurerebbe sia come fattore determinante della crescita macroeconomica sia di quella aziendale.

"L'investimento nelle persone assume un ruolo ancora più marcato, fungendo da propulsore fondamentale per l'innovazione e la competitività" (Šebestová & Popescu, 2022). Il successo di un'impresa, infatti, *"dipende in larga misura dalla qualità e dalla preparazione del suo personale per affrontare le sfide del mercato e capitalizzare le opportunità emergenti"* (Flores Laguna et al., 2020).



UNIVERSITÀ DI PAVIA

In base a quanto osservato nella letteratura, è possibile individuare i principali elementi del capitale umano pertinenti a questa tesi. Il primo è rappresentato dalle conoscenze e competenze, intese come il patrimonio intellettuale e pratico che gli individui acquisiscono attraverso l'istruzione formale, la formazione accademica e l'esperienza lavorativa, principali canali attraverso i quali le conoscenze e le competenze vengono sviluppate e progressivamente affinate. Tali componenti risultano cruciali per la produttività e la capacità innovativa delle organizzazioni, come evidenziato da Ayanponle e colleghi (2024) nella loro *review* sulle strategie innovative di gestione delle risorse umane. Inoltre, secondo Lovaglio e Vittadini (2004), questi fattori costituirebbero la base dell'accumulazione del capitale umano nel tempo e, di conseguenza, svolgerebbero un ruolo essenziale e imprescindibile. Un ulteriore elemento è costituito dalle abilità e attitudini individuali, sia innate sia acquisite. Bordiga e colleghi (2023), nel documento "*Startup: caratteristiche, finanziamento e performance*", si riferiscono ad esse come fondamenti per l'esecuzione efficiente di compiti specifici e per la risoluzione di problemi complessi. Accanto a queste componenti tradizionali, la letteratura più recente individua ulteriori dimensioni. In primo luogo, la salute e il benessere psicofisico, definiti come "l'integrità fisica e mentale dei lavoratori". Essi sarebbero essenziali per garantire la continuità operativa e massimizzare la produttività individuale e collettiva (Lovaglio & Vittadini, 2004). In secondo



UNIVERSITÀ DI PAVIA

luogo, anche i tratti individuali e la motivazione intrinseca sarebbero determinanti per l'impegno, la resilienza e la proattività nel contesto lavorativo, influenzando direttamente le *performance* a livello individuale e organizzativo, un concetto sviluppato da Česnyrienė e Stankevičienė (2011) nel loro studio "*The Role of Human Capital in Value Creation: Theoretical Insights*". Infine, le relazioni sia come reti sociali e professionali sia come insieme delle capacità comunicative agirebbero come "*catalizzatori per la diffusione delle conoscenze e per l'innovazione collaborativa*" (Kuzminov et al., 2019).

Dunque, generalmente, le caratteristiche del capitale umano evidenziate sono legate profondamente all'individuo. Come indicato dalla letteratura, quando sono analizzate specificatamente in una singola persona, possono essere definite "*Skills*", tradizionalmente distinte tra hard e soft. Esse comprendono non solo le conoscenze e abilità tecniche, ma anche gli attributi personali come l'intelligenza, la creatività e la capacità di adattamento (Baigireyeva et al., 2020). Le prime fanno riferimento a competenze cognitive e tecniche misurabili, come ad esempio le conoscenze disciplinari, le capacità numeriche e le competenze informatiche. Le seconde, invece, includono competenze di natura non cognitiva e relazionale, quali l'intelligenza emotiva, le capacità comunicative e la propensione alla collaborazione, particolarmente rilevanti nei contesti organizzativi (Lyu e Liu, 2021; Fantozzi et al., 2024; Farao et al., 2023). Accanto



UNIVERSITÀ DI PAVIA

ad esse, alcuni contributi individuano una terza categoria di competenze riconducibili alla sfera *manageriale* e *imprenditoriale*, che comprende le capacità decisionali, di coordinamento, di leadership e di visione strategica. Per trasformare il capitale umano individuale in risultati organizzativi concreti, soprattutto nelle fasi iniziali del *life-cycle*, le *skills* imprenditoriali dei fondatori sarebbero determinanti (Ceresia, 2023). Nonostante l'ampia rilevanza teorica dedicata ad esse, la loro analisi empirica risulta spesso complessa a causa delle difficoltà di misurazione e della limitata osservabilità nei *database* secondari. In particolare, le capacità non cognitive e i tratti personali difficilmente possono essere catturati in modo diretto e comparabile su ampi campioni di *startup* (Anghel e Balart, 2017; Acosta et al., 2018). Si riconosce il ruolo teorico delle *skills* come componenti del capitale umano, ma si concentra l'analisi empirica della presente tesi su dimensioni osservabili e consolidate nella letteratura, quali il livello di istruzione, la capacità imprenditoriale e la composizione delle squadre fondatrici. Tali variabili rappresentano *proxy* già utilizzati per catturare il capitale umano imprenditoriale e consentono di garantire la coerenza metodologica tra il quadro teorico e l'analisi quantitativa sviluppata nei capitoli successivi.

Il successo della *startup*, inteso come capacità di ottenere un late-stage funding e di concludere con un'*exit* positiva, non è riconducibile esclusivamente



UNIVERSITÀ DI PAVIA

a fattori di contesto o alla disponibilità di capitale finanziario, ma dipende anche dalle decisioni strategiche assunte nelle fasi iniziali del ciclo di vita. Esse sono prevalentemente attribuibili ai fondatori, che rappresentano i principali detentori e allocatori delle risorse cognitive, relazionali e *manageriali* dell'impresa. Ne consegue che l'analisi del capitale umano dei fondatori costituisce un passaggio centrale per comprendere i meccanismi con cui è possibile superare le fasi critiche del proprio sviluppo e raggiungere esiti positivi.

1.2 I Fondatori ed il Capitale Umano

Alla luce della definizione di successo adottata, è necessario approfondire il ruolo dei fondatori come decisori strategici durante lo sviluppo della *startup*. Il loro ruolo è fondamentale nelle fasi iniziali di ideazione, pianificazione e avvio; sono coloro che danno vita a una nuova attività imprenditoriale. *“Il capitale umano fa parte dell'insieme iniziale di risorse di tutte le nuove imprese, dove ci si aspetta che la combinazione tra la formazione e l'esperienza dell'imprenditore si sia trasformata in conoscenze e competenze che definiscono le caratteristiche specifiche del capitale umano della nuova impresa”* (Giones et al., 2018, p. 48).

Nelle *early stages*, i *founders* definiscono la visione e il modello di *business* e concentrano in sé una quota rilevante del capitale umano dell'impresa, influenzando direttamente il superamento delle fasi critiche di sviluppo. La letteratura sulle *exit* in questa tesi non è oggetto di analisi autonoma, bensì



UNIVERSITÀ DI PAVIA

richiama l'ambito teorico entro cui collocare il ruolo del capitale umano nella determinazione delle traiettorie di sviluppo delle *startup*.

Un contributo importante è offerto dall'articolo "*IPOs, trade sales and liquidations: Modelling venture capital exits using survival analysis*" di Giot e Schwiendbacher (2007). In questo studio, gli autori esaminano le modalità e il timing di uscita delle imprese supportate da venture capital, distinguendo tra *IPO*, acquisizioni e liquidazioni (*M&A*), e si concentrano prevalentemente sul ruolo dei *venture capitalist*. Sono analizzate le imprese statunitensi mediante modelli di sopravvivenza con rischi concorrenziali. I risultati dimostrano che le uscite sono i prodotti alternativi di un processo dinamico, influenzato dalle caratteristiche dell'impresa, nonché dal contesto istituzionale e dagli attori coinvolti nei finanziamenti.

Il raggiungimento di un'uscita positiva rifletterebbe indirettamente la qualità delle decisioni strategiche assunte nelle fasi iniziali che dipendono in larga misura dai fondatori. La letteratura sulle *IPO*, infatti, sottolinea che tale esito non dipende dalle caratteristiche degli attori che guidano l'impresa.

D'altro canto, la *review* "*IPO Research in Management and Entrepreneurship: Moving the Agenda Forward*" di Certo e colleghi (2009) sostiene la rilevanza opposta delle *IPO* come misura di *performance*. Sebbene non siano proposte analisi empiriche originali, questo articolo risulta rilevante



UNIVERSITÀ DI PAVIA

perché colloca il capitale umano dei fondatori e del *management* tra i principali determinanti della *performance*. Lo studio analizza la letteratura manageriale e imprenditoriale di alcuni articoli empirici, evidenziando il ruolo della *leadership* e della composizione dei *team* di vertice sulla quotazione in borsa. Alcune variabili utilizzate sono riconducibili all'*upper echelons theory*, che interpreta il ruolo della leadership e di altri fattori, tra cui esperienza, istruzione e prestigio, come elementi che influenzano la percezione della qualità dell'impresa da parte degli investitori, incidendo sull'*underpricing* e sulla valutazione.

Dunque, una parte della letteratura sulle *IPO* conferma che le competenze e l'esperienza influiscono anche sulla probabilità di ottenere esiti avanzati di sviluppo sul mercato finanziario.

L'attenzione si sposta dalle *exit*, intese come eventi finali di una nuova impresa, alle conseguenze delle caratteristiche intrinseche dei fondatori. Di conseguenza, è necessario analizzare i processi con cui si costituisce la squadra fondatrice per comprendere come tali caratteristiche emergano nelle fasi early stage e, più specificamente, quali siano i meccanismi di selezione e di aggregazione dei cofondatori.

Clough e Vissa (2022) propongono un modello teorico che interpreta la cofondazione come un processo sequenziale di matching, articolato in fasi di attivazione, valutazione, approccio e reciprocazione. Secondo quanto descritto



UNIVERSITÀ DI PAVIA

dagli autori, la formazione del gruppo non sarebbe il risultato di una scelta istantanea o puramente razionale, bensì emergerebbe da un processo di convergenza di valutazioni reciprocamente soddisfacenti tra i potenziali membri del gruppo. Le decisioni di *team building* sarebbero guidate da criteri di soddisfazione basati su valutazioni della competenza e della compatibilità interpersonale, anziché su criteri di ottimizzazione. L'assenza di cofondatori è interpretata come l'esito del fallimento nel processo di *matching*.

Il modello offre spunti utili per comprendere come la composizione della squadra influenzi la dotazione iniziale di capitale umano. In particolare, suggerisce che le caratteristiche osservabili dei fondatori e la struttura del *team* non siano esiti casuali, bensì il risultato di processi decisionali.

Un altro contributo centrale è la meta-analisi di Jin e colleghi (2017) intitolata "*Entrepreneurial Team Composition Characteristics and New Venture Performance*". Partendo dall'*Upper Echelons Theory* (Bonelli, 2023), lo studio estende le evidenze empiriche sul legame tra le caratteristiche dei gruppi fondatori e le *performance* delle nuove imprese. È evidenziato che il ruolo dei fondatori è stato ampliato a causa dell'elevata incertezza ambientale in cui è immersa la *startup*. Inoltre, la discrezionalità decisionale e la centralità delle scelte iniziali aumentano la rilevanza delle azioni di questi soggetti.



UNIVERSITÀ DI PAVIA

La meta-analisi distingue tre chiavi di lettura della composizione dei *team*: (i) caratteristiche aggregate dei membri, (ii) eterogeneità delle caratteristiche individuali e (iii) dimensione. Leggendo le caratteristiche aggregate secondo la *Human Capital Theory*, si evidenziano come centrali nel gruppo fondatore: il livello complessivo delle competenze, delle conoscenze e delle esperienze disponibili. L'eterogeneità, invece, è interpretata attraverso due prospettive teoriche contrapposte: da un lato, la *information processing perspective*, che attribuisce ai gruppi eterogenei una maggiore capacità di *problem solving* e di elaborazione delle informazioni; dall'altro, i potenziali costi legati a conflitti e problemi di coordinamento. Infine, la dimensione è considerata come *proxy* dell'accesso alle risorse economiche, relazionali e cognitive.

Gli autori sottolineano che la composizione rappresenta un determinante della *performance* imprenditoriale. I risultati mostrano che tutte e tre le chiavi di lettura presentano una relazione positiva e statisticamente significativa con la *performance* delle nuove imprese in termini di crescita e redditività, con un effetto particolarmente robusto associato alle caratteristiche aggregate del capitale umano. L'eterogeneità e la dimensione del *team* avrebbero effetti positivi più contenuti e potenzialmente non lineari. Tuttavia, si evince la necessità di ulteriori analisi quantitative basate su variabili osservabili.



UNIVERSITÀ DI PAVIA

Un'altra meta-analisi inerente a quanto visto è stata svolta da Unger e colleghi (2009) denominata "*Human Capital and entrepreneurial success: A meta-analytical review*". Sono stati analizzati 70 campioni indipendenti, fornendo una delle sintesi più complete disponibili sulla relazione tra il capitale umano ed il successo imprenditoriale, integrando i risultati di oltre trent'anni di ricerca empirica.

Il quadro teorico di riferimento è fondato principalmente sulla teoria del capitale umano di Becker (1975), che lo interpreta come l'insieme di conoscenze e competenze acquisite attraverso investimenti in istruzione, formazione ed esperienza lavorativa. Questa è una visione orientata all'apprendimento, secondo cui il successo dipende dalla capacità degli individui di trasferire e applicare le conoscenze rilevanti acquisite in precedenza all'attività di imprenditore. In questa teoria il capitale umano rappresenta una risorsa dinamica ed il suo valore non dipende solo dalla quantità di investimenti effettuati, ma anche dalla loro pertinenza rispetto al passato.

I risultati della meta-analisi evidenziano una relazione positiva e statisticamente significativa tra capitale umano e successo, ma essa è soggetta alle modalità con cui vengono concettualizzati e misurati. In particolare, gli effetti dell'analisi risultano più robusti quando al capitale umano è attribuito il significato di *outcome* di un investimento in conoscenze e competenze



UNIVERSITÀ DI PAVIA

effettivamente acquisite, anziché semplice indicatore dell'effettuazione di investimenti in istruzione ed esperienza. Inoltre, il capitale umano specifico imprenditoriale mostrerebbe un'associazione più forte con il successo rispetto alle forme di capitale umano generico. Inoltre, secondo gli autori, il capitale umano è più rilevante nelle fasi iniziali del *life-cycle*, coerentemente con l'idea di una maggiore vulnerabilità e con la cosiddetta *liability of newness*. Complessivamente, i risultati indicano che, pur rappresentando un determinante significativo del successo, il capitale umano esercita un effetto che dipende dalle modalità di misurazione adottate.

Questo contributo fornisce il fondamento teorico ed empirico di questa tesi per l'analisi delle componenti osservabili del capitale umano dei fondatori, quali l'istruzione e l'esperienza lavorativa e imprenditoriale. Al contempo, l'evidenza di effetti contenuti rafforza la necessità di ulteriori analisi quantitative basate su *dataset* ampi e su *proxy* coerenti con le dimensioni teoriche osservabili.

L'attenzione dell'analisi del presente studio è posta sulle *proxy* quantitative del capitale umano indicate in precedenza, coerenti con l'impostazione empirica tradizionalmente operazionalizzata attraverso indicatori misurabili quali il livello di istruzione, l'esperienza lavorativa pregressa e la partecipazione ad attività imprenditoriali, risultando particolarmente adatte alle valutazioni basate su dati secondari. Pur riconoscendo il ruolo di alcuni tratti di



UNIVERSITÀ DI PAVIA

personalità e *soft skills* (McCarthy et al., 2023) coerenti con il concetto di capitale umano, tali dimensioni non vengono considerate nella presente tesi poiché non direttamente osservabili nei dati disponibili.

Un filone consolidato di studi sull'esperienza lavorativa evidenzia come la specificità settoriale, in particolare quella tecnica e tecnologica, abbia un effetto positivo diretto sulla crescita delle imprese (Colombo & Grilli, 2005). Invece, per quanto riguarda quella generica, sebbene sia meno correlata alla crescita rispetto a quella specifica, anch'essa contribuisce allo sviluppo di competenze manageriali trasferibili e a una maggiore adattabilità strategica, elementi potenzialmente rilevanti nella valutazione degli investitori, come riportato da Colombo e Grilli (2009).

Per comprendere il ruolo dell'eterogeneità delle esperienze dei fondatori, è particolarmente interessante la pubblicazione di Beckman e Burton (2008). Gli autori contrappongono una prospettiva di tipo *life-cycle*, secondo cui la crescita e la professionalizzazione tenderebbero a rendere marginali le caratteristiche personali iniziali dei fondatori, a una prospettiva *path-dependent*, che enfatizza i meccanismi di imprinting organizzativo-culturale e di similitudine tra strutture, organi e procedure aziendali.

Lo studio osserva in che misura le condizioni iniziali di fondazione influenzino i board e il raggiungimento di milestone, come il finanziamento da



UNIVERSITÀ DI PAVIA

parte dei venture capitalist o la quotazione in borsa, ed è interessante osservare la distinzione tra l'esperienza funzionale e la struttura funzionale iniziale. La prima è intesa come capitale umano individuale derivante dalle carriere pregresse in specifiche funzioni (ad esempio ingegneria, *business*, finanza); la seconda, invece, è definita come la presenza formale di ruoli esecutivi differenziati all'interno dell'organizzazione.

Contrariamente a quanto spesso assunto nella letteratura sui *top executive team*, le due dimensioni non risulterebbero tra loro sostituibili e seguirebbero traiettorie evolutive distinte.

Per osservare questo fenomeno, gli autori hanno effettuato un'analisi longitudinale su un campione di imprese *high-tech* della *Silicon Valley* e i risultati mostrerebbero che una maggiore ampiezza delle esperienze funzionali dei fondatori sarebbe associata a una più rapida attrazione di *manager* esperti e a un accesso anticipato al finanziamento di *venture capital*. Al contrario, la presenza di una struttura funzionale iniziale più articolata risulterebbe maggiormente correlata al raggiungimento dell'*IPO*. In entrambi i casi, le imprese osservate con un'elevata varietà di esperienze funzionali combinata ad una struttura relativamente complessa avrebbero le traiettorie di crescita più rapide in seguito di un effetto di vantaggio cumulativo.



UNIVERSITÀ DI PAVIA

Le evidenze raccolte confermerebbero l'ipotesi che le caratteristiche del capitale umano dei fondatori esercitino un'influenza duratura sull'evoluzione organizzativa e sugli esiti di successo, anche oltre le fasi iniziali di vita dell'impresa. In particolare, l'eterogeneità delle esperienze risulterebbe determinante per la capacità di attrarre risorse esterne e di avanzare nello sviluppo. Tali riscontri conferiscono un solido fondamento teorico ed empirico all'analisi sviluppata nella presente tesi.

Spostando il focus all'iniziativa imprenditoriale, *“la partecipazione a iniziative di costituzione d'impresa consente ai fondatori di accumulare competenze, routine decisionali e conoscenze del mercato che facilitano l'identificazione delle opportunità e la gestione dei rischi”* (Aryadita et al., 2023, p. 5).

In merito a quanto citato, l'articolo *“Performance persistence in entrepreneurship”* di Gompers e colleghi (2010) è un contributo rilevante. Esso esamina l'esistenza e le determinanti della continuità di *performance* in ambito aziendale. Utilizzando un ampio *dataset* di imprese statunitensi finanziate da *venture capitalist* tra il 1986 e il 2000, gli autori osservano se il successo passato aumenti la probabilità di successo futuro.

I risultati mostrano che i fondatori con un precedente esito positivo – definito come *IPO* o iniziativa per *IPO* – presentano una probabilità più elevata



UNIVERSITÀ DI PAVIA

di ottenere un ulteriore esito positivo nel *venture* successivo rispetto sia ai *first-time entrepreneurs* sia a coloro che hanno sperimentato insuccessi in precedenza. In particolare, l'*odds* calcolata si attesta intorno al 30% per coloro i quali hanno un *track record* positivo, contro valori prossimi al 21–22% per le altre categorie.

Gli autori distinguono dal punto di vista teorico due componenti del successo: il primo sarebbe riconducibile alla capacità di *market timing*, ossia l'abilità di avviare nuove imprese in settori e periodi caratterizzati da condizioni favorevoli; il secondo riguarderebbe la capacità manageriale, intesa come *overperformance* rispetto ai concorrenti nello stesso settore. Entrambi gli elementi sarebbero persistenti nel tempo e contribuirebbero al successo futuro, sottolineando che esso rifletterebbe abilità sistematiche e non sarebbe il risultato di eventi casuali. Nell'articolo è identificato un ulteriore meccanismo chiamato "*success breeds success*" secondo cui il successo passato migliorerebbe la capacità dell'imprenditore di attrarre risorse di elevata qualità in termini di capitale economico, umano e strategico, riducendo i problemi di avvio d'impresa. In questo senso, il *track record* positivo fungerebbe da segnale credibile per gli *stakeholder* e amplificherebbe le differenze iniziali tra imprenditori. Tali evidenze indicherebbero che non è la mera serialità



UNIVERSITÀ DI PAVIA

imprenditoriale a generare persistenza di *performance*, bensì la qualità dell'esperienza imprenditoriale accumulata.

Nel complesso, dunque, si fornisce un supporto empirico all'idea che l'esperienza imprenditoriale non sia omogenea, ma differisca in funzione degli esiti precedenti. Tali risultati giustificano l'introduzione nella presente tesi di misure che osservano la serialità imprenditoriale dei fondatori, interpretabile come il numero di imprese fondate e la presenza o meno di altre esperienze come fondatore; inoltre, supportano l'analisi dell'impatto di queste sulla probabilità di *exit* delle *startup*.

Anche l'articolo di Hsu (2007) "*Experienced entrepreneurial founders, organizational capital, and venture capital funding*" esplora il ruolo dell'esperienza imprenditoriale a fianco del capitale organizzativo nella capacità delle *startup* di ottenere finanziamenti. Hsu evidenzia che la presenza di fondatori con esperienze pregresse positive rappresenterebbe un segnale positivo per gli investitori poiché ridurrebbe le asimmetrie informative ed il rischio percepito.

L'autore distingue tra capitale umano specifico (inteso come competenze tecniche e *manageriali*) e capitale organizzativo (rappresentato da reputazione, relazioni e reti sviluppate in precedenza). In particolare, quest'ultimo faciliterebbe l'accesso a risorse finanziarie e strategiche. Hsu



UNIVERSITÀ DI PAVIA

mostra empiricamente che le società guidate da imprenditori seriali o da aggregati con esperienze complementari hanno maggiori probabilità di ottenere capitale di rischio, sottolineando il “*certification effect*”: i fondatori esperti agirebbero come certificatori del potenziale della *startup* riducendo il rischio percepito dagli investitori. Quanto detto è coerente anche con i contributi di Colombo e Grilli (2005), secondo i quali le competenze dei fondatori influenzano direttamente la *performance* iniziale e la crescita delle imprese ad alto contenuto tecnologico. La partecipazione ad attività imprenditoriali dei fondatori sarebbe un asset strategico per la raccolta di capitale, confermando l’importanza del capitale umano nel determinare le scelte e il successo.

Sebbene questa letteratura enfatizzi il ruolo dell’esperienza imprenditoriale strettamente precedente, altri studi empirici basati su dati secondari adottano *proxy* più ampie, che catturano l’intensità dell’attività imprenditoriale complessiva, indipendentemente dalla sequenza temporale delle iniziative come lo studio di Sanguineti e colleghi (2024). La presente tesi segue questo filone ed utilizza come *proxy* del *background* imprenditoriale l’intensità complessiva di esperienza accumulata dai membri del *team* fondatore, ossia il numero totale di imprese fondate dai singoli riconducibili ad ogni gruppo osservato, in considerazione dell’impossibilità di ricostruire con precisione la sequenza temporale delle iniziative imprenditoriali.



UNIVERSITÀ DI PAVIA

Logicamente la presenza di più fondatori consente di combinare competenze e prospettive differenti, generando un capitale umano collettivo più robusto. Infatti, una parte della letteratura suggerisce che i gruppi moderatamente eterogenei favorirebbero innovazione e creatività, a fronte di un aumento dei costi di coordinamento per i quali, tuttavia, sarebbero necessari meccanismi di governance efficaci (Bordiga et al., 2023; Colombo & Grilli, 2005). *“Si osserva che le imprese con più fondatori hanno maggiori probabilità di successo, ..., le imprese con tre o più fondatori hanno probabilità di successo più che doppie rispetto alle startup fondate da una sola persona.”* (McCarthy et al., 2023, p. 6).

Considerare gli elementi di capitale a livello di *team* permette di derivare misure di eterogeneità e complementarità coerenti con i dati secondari disponibili. Tale impostazione collega direttamente la composizione del gruppo fondatore alla probabilità di successo, misurata in termini di avanzamento nel *life-cycle* del finanziamento ed *exit*, obiettivo centrale del lavoro in oggetto.

Un titolo più elevato in ambito economico-gestionale risulterebbe associato a migliori capacità decisionali e strategiche (Colombo & Grilli, 2005, 2009). La già citata istruzione può essere interpretata come formazione accademica; infatti, numerosi studi evidenziano una correlazione positiva tra il livello di istruzione dei fondatori e le *performance* delle *startup*. La presenza di



UNIVERSITÀ DI PAVIA

titoli avanzati, come il dottorato, può fungere da segnale di capacità di ricerca e sviluppo, aumentando l'attrattiva per gli investitori in capitale di rischio. In linea con l'obiettivo previsto, quanto detto è interpretabile come segnale osservabile empiricamente nella valutazione delle probabilità di accesso a stadi avanzati di finanziamento o di *exit*.

L'articolo "*The impact of digital startup founders' higher education on reaching equity investment milestones*" di Ratzinger e colleghi (2018) analizza l'impatto dell'istruzione universitaria dei fondatori sulla probabilità che le *startup* digitali raggiungano due traguardi chiave di *equity*: l'ottenimento di finanziamenti esterni (da *angel* o *venture capital*) e l'*exit* tramite acquisizione o *IPO*, il tutto all'interno del *framework* della *Human Capital Theory*. Utilizzando un *dataset* tratto da Crunchbase, gli autori stimano dei modelli di *ordered probit* per valutare l'effetto di differenti tipologie di istruzione (tecnica, economico-manageriale e generale) e dei diversi livelli di titolo (laurea, *master*, dottorato) sugli esiti finanziari delle imprese del loro campione. Sul piano empirico, l'articolo si inserisce nel filone che analizza i segnali osservabili *ex ante* utilizzati dagli investitori per valutare le *startup*; l'istruzione universitaria sarebbe percepita sia come fonte di competenze sia come segnale di qualità.

I loro risultati descrivono che il capitale umano universitario dei fondatori è positivamente associato alla probabilità di raggiungere traguardi di



UNIVERSITÀ DI PAVIA

equity, ma in modo eterogeneo e non lineare. In particolare, la presenza di almeno un fondatore con formazione tecnica aumenta significativamente la probabilità di ottenere finanziamenti ed *exit*, sebbene l'effetto marginale decresca ai livelli di titolo più elevati. La formazione economico-manageriale risulta invece rilevante soprattutto a livello di dottorato, mentre le lauree e i *master* in ambito *business* riportano effetti più deboli o non robusti. Inaspettatamente, l'istruzione generale *undergraduate* (arti, umanistiche, scienze sociali) è positivamente associata ai traguardi di *equity*, mentre tale relazione non emerge per livelli di istruzione superiori. I risultati suggeriscono che non sarebbe l'accumulazione lineare di anni formativi a determinare le *performance* imprenditoriali, bensì le specifiche combinazioni di aree disciplinari e livelli di formazione, nonché la loro integrazione all'interno del gruppo fondatore. In questo senso, l'evidenza di rendimenti decrescenti o non monotoni mette in discussione una lettura puramente lineare della teoria del capitale umano e si allinea ai contributi che ipotizzano i fenomeni di *over-investment* educativo.

Il mercato del capitale di rischio valorizzerebbe maggiormente le forme di capitale umano direttamente allineate alle esigenze tecnologiche e strategiche del settore digitale, ma si riconosce un ruolo non trascurabile alle competenze generaliste nello sviluppo iniziale dell'impresa.



UNIVERSITÀ DI PAVIA

La presente tesi integra l'analisi della persistenza di *performance* basata sull'esperienza imprenditoriale con il capitale umano formale, osservando come esso contribuisca a strutturare le probabilità di successo nei mercati. Quanto detto è coerente con le presunzioni fatte perché rafforza l'argomento per il quale le *performance* e l'accesso al finanziamento dipendono da caratteristiche individuali e di *team*, ma in modo eterogeneo e mediato da meccanismi di selezione e percezione.

L'assenza di una definizione accademica e normativa univoca rende necessario adottare un inquadramento concettuale che consenta di cogliere le specificità strutturali ed evolutive di queste imprese. Per questo motivo, l'approccio basato sul ciclo di vita dell'impresa innovativa si è rivelato particolarmente utile per interpretare le fasi di sviluppo e per individuare i momenti critici, quali il superamento della *Valley of Death* e la transizione verso gli stadi avanzati della crescita.

Il concetto di successo della *startup* assume una natura multidimensionale. Coerentemente con la letteratura sull'imprenditorialità innovativa e con le dinamiche dei mercati *venture-backed*, nel presente lavoro il successo viene operazionalizzato attraverso il raggiungimento di un'*exit positiva*, quale un'operazione di fusione o acquisizione (*M&A*) o un'offerta pubblica iniziale (*IPO*). Tali eventi, infatti, rappresentano esiti osservabili e



UNIVERSITÀ DI PAVIA

costituiscono il principale meccanismo di realizzazione del valore per fondatori e investitori e segnalano una validazione del modello di *business* e il raggiungimento di uno stadio avanzato di maturazione dell'impresa.

La letteratura converge nel riconoscere il capitale umano come una risorsa centrale nei contesti imprenditoriali. Le caratteristiche umane come l'istruzione e le esperienze lavorative ed imprenditoriali incidono significativamente sulla capacità delle *startup* di attrarre risorse, superare le fasi di criticità e avanzare verso esiti positivi, come è stato osservato in numerosi studi. Accanto alle dimensioni individuali, dalla letteratura osservata emerge inoltre l'importanza della composizione dei *team*, sia in termini di aggregazione sia di eterogeneità del capitale umano. Essi sarebbero fattori in grado di influenzare le *performance* e le traiettorie delle nuove imprese.

I risultati empirici accademici disponibili appaiono eterogenei e spesso condizionati dalle modalità di misurazione del capitale umano e del successo. Molte analisi si concentrano su indicatori di sopravvivenza o crescita, trascurando il ruolo delle *exit* come esito finale. Le misurazioni empiriche delle componenti non osservabili del capitale umano, quali ad esempio le già menzionate *soft skills* e i tratti di personalità, risultano complesse e difficilmente applicabili su ampi campioni.



UNIVERSITÀ DI PAVIA

In conclusione, la presente tesi si colloca nel filone di studi che analizzano il ruolo del capitale umano dei fondatori e delle caratteristiche dei *team* imprenditoriali nella determinazione del successo, adottando un approccio quantitativo basato su dati secondari. In particolare, l'attenzione è rivolta a dimensioni osservabili quali il livello di istruzione, l'esperienza lavorativa e imprenditoriale dei fondatori, nonché alle caratteristiche strutturali dei *team* fondatori, al fine di valutarne l'impatto sulla probabilità di raggiungere un'*exit* positiva.

Le ipotesi che questa tesi si propone di sviluppare sono le seguenti:

H1: Un livello più elevato di istruzione media del *team* dei fondatori è positivamente associato alla probabilità che una *startup* raggiunga il successo.

Questa supposizione si basa sui contributi che evidenziano come il capitale umano formale favorisca l'acquisizione di competenze cognitive, manageriali e analitiche rilevanti nei contesti imprenditoriali ad alta intensità di conoscenza.

H2: Una maggiore quantità di esperienza accumulata a livello di *team* nella fondazione di imprese è positivamente associata alla probabilità che una *startup* raggiunga il successo.



UNIVERSITÀ DI PAVIA

L'ipotesi si fonda sulla letteratura che attribuisce all'esperienza di fondazione un ruolo chiave nella riduzione dell'incertezza e nell'aumento delle capacità decisionali e di gestione delle fasi del ciclo di vita.

H3: Una maggiore eterogeneità del *team* fondatore in termini di esperienze settoriali è positivamente associata alla probabilità che una *startup* raggiunga il successo.

Questo riflette l'idea diffusa in letteratura che i *team* più diversificati per esperienza settoriale precedente siano in grado di integrare prospettive complementari, favorendo lo sviluppo della *startup*.

Il capitolo successivo presenta il disegno metodologico adottato per testare le ipotesi formulate, illustra le fonti dei dati utilizzate, le modalità di costruzione del *dataset* e delle variabili, nonché il modello statistico impiegato nell'analisi.

Capitolo 2: Metodologia della ricerca

Il successo imprenditoriale è definito in questa tesi come il raggiungimento o il superamento della fase di finanziamento avanzato (*late stage funding*), poiché esso rappresenta l'avvicinamento a un evento di *exit*. Quest'ultimo può concretizzarsi attraverso un'operazione di fusione o acquisizione (*M&A*) oppure mediante un'offerta pubblica iniziale (*IPO*). Tali eventi sono considerati esiti



UNIVERSITÀ DI PAVIA

tangibili e verificabili del ciclo di vita di una *startup*, costituiscono i principali meccanismi attraverso cui viene generato valore e permettono di ridurre le ambiguità interpretative associate a misure soggettive o multidimensionali delle *performance* aziendali. Questa definizione è in linea con la letteratura accademica riguardante il ciclo di vita delle *startup*, la quale interpreta l'*exit* come l'esito osservabile di un processo di crescita e selezione del mercato (Giot & Schwienbacher, 2007; Pisoni & Onetti, 2018).

Per supportare il quadro teorico, l'analisi empirica integra evidenze di tipo statistico-probabilistico e si focalizza su tre dimensioni osservabili: il livello di istruzione, l'esperienza lavorativa settoriale e la serialità imprenditoriale. La scelta di queste dimensioni è direttamente derivata dalla *Human Capital Theory* (Becker, 1975) che le individua come elementi centrali e altamente rilevanti per le *performance* delle nuove imprese (Colombo & Grilli, 2005; Unger et al., 2011). L'aggregazione è svolta a livello di organizzazione seguendo l'approccio basato sui gruppi di fondatori adottato dalla letteratura sulle *entrepreneurial teams* secondo cui i risultati delle *startup* dipendono dalla dotazione complessiva di capitale umano del gruppo e non esclusivamente dalle caratteristiche dei singoli individui (Beckman & Burton, 2008; Jin et al., 2017).

Sulla base delle ipotesi di ricerca precedentemente presentate, la domanda che guida questa tesi è la seguente: “*In che misura il livello di*



UNIVERSITÀ DI PAVIA

istruzione, l'esperienza imprenditoriale complessivamente accumulata e l'eterogeneità settoriale del team fondatore influenzano la probabilità che una startup raggiunga un'exit positiva?"

L'obiettivo è stimare la probabilità statistica che una *startup* raggiunga la fase di *late stage funding* e degli eventi di *IPO* o *M&A*, in funzione delle caratteristiche del capitale umano e della composizione del *team* imprenditoriale. Nel capitolo si presenta la metodologia seguita per ottenere i dati e costruire il *database* necessario per l'analisi.

Poiché la variabile dipendente ricercata è dicotomica, si è scelto di utilizzare modelli di regressione logistica. Dal punto di vista statistico, il modello *logit* consente di sviluppare una relazione non lineare tra regressori e probabilità di successo, garantendo che le probabilità stimate siano comprese tra zero e uno, a differenza dei modelli lineari applicati a variabili binarie (Long & Freese, 2001; Greene, 2018). Si mira a chiarire come le caratteristiche delle variabili indipendenti e le loro interazioni possano tradursi in differenti esiti.

Per questo studio è stato utilizzato un modello *cross-sectional* basato sull'analisi di dati secondari relativi al 2025, non essendo disponibili dati longitudinali completi per tutti i percorsi temporali. Questo approccio è coerente con la natura dei dati disponibili e con l'obiettivo principale dello studio: stimare la variabile dipendente rappresentata dalla probabilità di raggiungere l'evento di



UNIVERSITÀ DI PAVIA

exit attraverso l'analisi della relazione tra il capitale umano e la percentuale di successo (Wooldridge, 2010; Unger et al., 2011; Jin et al., 2017). Il processo di raccolta e analisi dei dati è stato progettato considerando la natura quantitativa dell'intero disegno metodologico. Il *dataset* utilizzato è costituito a partire da Crunchbase, un *database* ampiamente impiegato nella letteratura accademica per lo studio delle *startup* e delle dinamiche del finanziamento. L'unità di analisi è rappresentata dalla singola impresa, mentre le informazioni relative ai fondatori vengono elaborate per riflettere la composizione complessiva del *team* imprenditoriale al momento della fondazione.

2.1 Strumenti di ricerca e raccolta dati

I dati sono stati ottenuti attraverso Crunchbase², un *database online* sviluppato e lanciato nel 2007 da TechCrunch³. È stato già utilizzato nella ricerca imprenditoriale ed è considerato una fonte affidabile e completa di informazioni sia sulle caratteristiche dei fondatori che sulle loro aziende. (Sanguineti et al., 2024). Crunchbase aggrega dati storici e conserva anche dati ed eventi rilevanti antecedenti alla sua creazione. I dati forniti hanno portata globale e coprono cinque aree principali: persone, organizzazioni, attività di investimento, uscite ed eventi pubblici. Unendo diversi set di dati, è possibile identificare il numero

² www.crunchbase.com

³ www.techcrunch.com



UNIVERSITÀ DI PAVIA

di aziende fondate, le eventuali *exit* di queste imprese e le loro posizioni lavorative (Sanguineti et al., 2024).

Per ottenere i dati è stato seguito un processo complesso suddiviso in sottoprocessi. L'estrazione è avvenuta tra ottobre e dicembre 2025, a causa della mole di dati ampia, aggiornati al 2025; pertanto le informazioni inerenti sono osservate fino a tale momento. A partire dai dati grezzi estratti da Crunchbase, l'estrazione è stata effettuata mediante l'utilizzo di programmi sviluppati in linguaggio Python, impiegati per interrogare Crunchbase attraverso chiamate *API* (*Application Programming Interface*). Gli *script* utilizzati per le interrogazioni e per l'estrazione dei dati sono riportati in Appendice A, al fine di garantire la trasparenza e la replicabilità del processo.

Questo sottoprocesso si è concentrato sull'identificazione delle imprese e dei fondatori tramite i rispettivi identificativi univoci (UUID). Inizialmente, è stata effettuata una chiamata *API* di tipo *Search* finalizzata all'estrazione dei gruppi di categorie utilizzati da Crunchbase per la classificazione delle imprese, Codice 1 (vedi Appendice A). Lo *script* riportato segue la struttura generale della *pipeline* di acquisizione dati tramite *API* ed è una ricostruzione degli *script* originali. Mostra la logica di interrogazione del *database*, del salvataggio strutturato in formato JSON e della gestione incrementale del processo, senza includere le credenziali o i percorsi locali specifici. Per ciascun gruppo sono stati



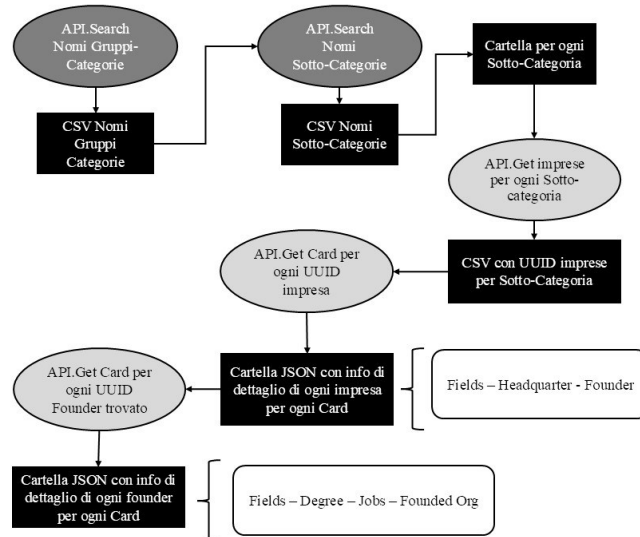
UNIVERSITÀ DI PAVIA

raccolti l'UUID e il nome identificativo, successivamente salvati in un file in formato CSV. Tale passaggio ha consentito di mappare il livello gerarchico più elevato della classificazione settoriale adottata da Crunchbase. Successivamente, è stata effettuata l'*API* per ottenere l'elenco completo delle sottocategorie associate a ciascun gruppo categoriale. Per ognuna di esse sono stati estratti i nomi, anch'essi archiviati in file CSV separati. Questo elenco è stato utilizzato con la stessa modalità al fine di identificare le imprese associate a ciascuna sottocategoria. È stato generato un file CSV contenente gli UUID delle imprese corrispondenti. A partire dall'insieme delle imprese così ottenute, sono state effettuate ulteriori chiamate *API* per ciascun identificativo per recuperare le informazioni di dettaglio. Contestualmente, sono stati scaricati anche gli UUID dei fondatori presenti nelle informazioni di dettaglio delle imprese. Di seguito viene riportato il modello di estrazione dei dati.



UNIVERSITÀ
DI PAVIA

Figura 3: Modello di estrazione dati



Fonte: Elaborazione dell'autore

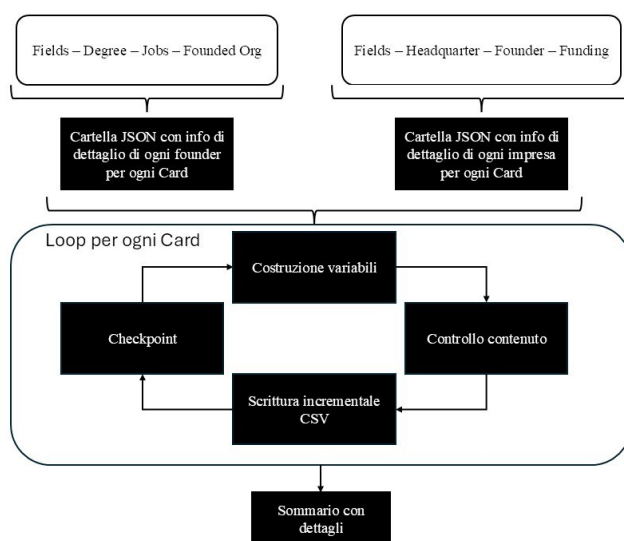
Attraverso lo *script* Codice 2 (vedi Appendice A), è stato possibile ottenere informazioni molto specifiche su imprese e fondatori contenute nelle cosiddette Card: microcategorie di dati relative a un singolo identificativo, correlate da parole chiave associabili a tipologie di informazioni. Ad esempio, la Card “Degree” contiene dati relativi al percorso di istruzione di un singolo soggetto. Lo *script* mostra la logica di interrogazione delle Card, il salvataggio e la gestione incrementale. I dati sono stati estratti in formato JSON e suddivisi per ciascuna Card. Per ognuna di esse è stato costruito un file CSV organizzato in base agli UUID di imprese, o fondatori, in base all’origine delle informazioni. Sono stati sviluppati programmi in linguaggio R, i cui *script* sono riportati in Appendice B. Tramite la lettura di tutti i file JSON omogenei di ogni cartella



UNIVERSITÀ DI PAVIA

Card, organizzati per argomento, i dati vengono convertiti in stringhe o valori numerici in un unico file CSV per ciascuna cartella Card. Gli *script* “appiattiscono” i dati annidati filtrandoli in base alle necessità, includendo solo i contenuti non vuoti e gestiscono la scrittura in blocchi con checkpoint e cache efficienti. Per l’intero processo di estrazione sono stati implementati meccanismi di controllo, come salvataggi periodici e procedure di gestione degli errori di rete, per garantire la continuità e l’affidabilità dell’elaborazione dei dati. In Figura 4 è riportato il modello logico seguito per questa fase.

Figura 4: Schema costruzione CSV Cards



Fonte: Elaborazione dell'autore

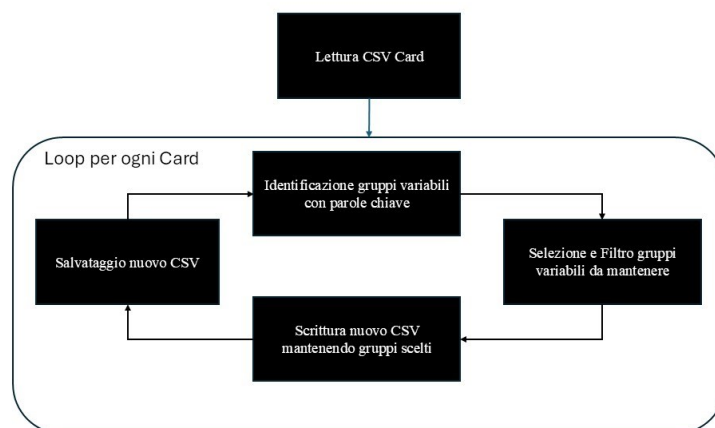
Al termine del processo, è stato prodotto un CSV completo per ciascuna Card, organizzato per UUID delle organizzazioni o dei fondatori, a seconda della



UNIVERSITÀ
DI PAVIA

tipologia di dati. Tuttavia, tali CSV risultavano troppo grandi e complessi per essere importati direttamente nello strumento di analisi utilizzato, Stata. Per questo motivo, mediante *script* R dedicati, sono state selezionate solo le informazioni rilevanti in base agli obiettivi previsti e alla letteratura di riferimento, come riportato in Figura 5.

Figura 5: Pulizia CSV Cards



Fonte: Elaborazione dell'autore

Una volta ottenuti i CSV Cards puliti dai dati superflui, è stato necessario unirli in due *database* intermedi: uno organizzato per UUID delle imprese e uno per UUID dei fondatori. La separazione è stata necessaria per procedere alla normalizzazione dei dati. Gli *script* R sviluppati per questa fase leggevano più CSV Cards di organizzazioni o fondatori (*fields*, *degrees*,



UNIVERSITÀ DI PAVIA

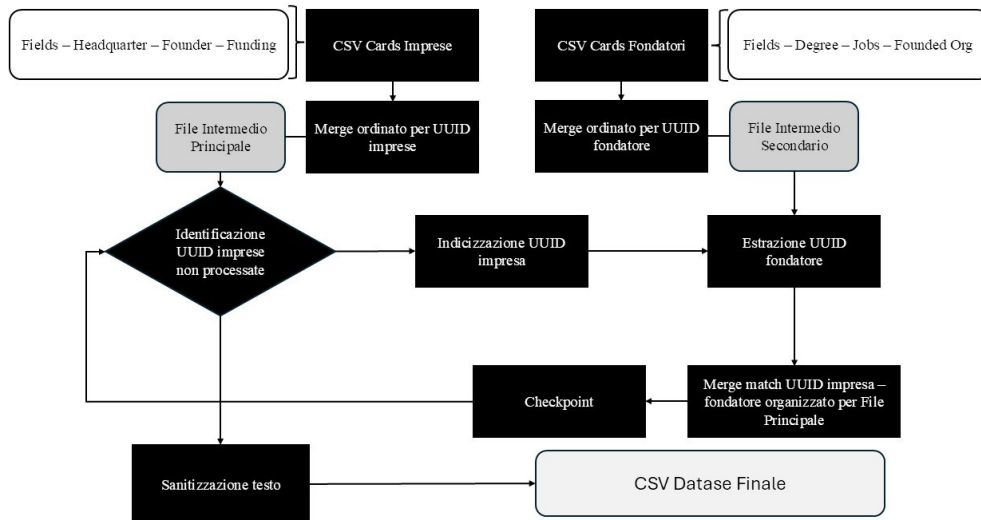
headquarter, ecc.) gestendo caratteri problematici quali virgole, parentesi, ritorni a capo e spazi. Gli *script* aggregavano i dati per UUID, combinando le variabili ripetute e, in caso di necessità, creando variabili aggiuntive separate, producendo un unico CSV finale per ciascuna entità (Organizzazioni o Fondatori), con record uniti, pronti per ulteriori operazioni di merge.

Infine, alcune variabili dei *database* intermedi contenevano valori testuali non standardizzati, derivanti dai JSON iniziali, poiché non esistono vincoli formali su quel formato. La stessa informazione poteva essere rappresentata in modi diversi (ad esempio con abbreviazioni, maiuscole/minuscole o punteggiatura). Per garantire la coerenza e la comparabilità dei dati, è stato sviluppato uno *script* R per uniformare le diverse rappresentazioni della stessa variabile, previa verifica manuale. Il modello logico è riportato di seguito.



UNIVERSITÀ
DI PAVIA

Figura 6: Costruzione Database Finale



Fonte: Elaborazione dell'autore

Il campione così ottenuto comprendeva 132.661 osservazioni per quanto riguarda il *database* Intermedio dei Fondatori, per i quali sono state raccolte informazioni demografiche, lavorative e imprenditoriali a livello singolo. Per quanto riguarda, invece, il *database* Intermedio per le Organizzazioni, il campione iniziale comprendeva 112.851 osservazioni, delle quali sono state raccolte informazioni demografiche e finanziarie al momento dell'osservazione, nonché gli identificativi dei fondatori di ciascuna società.

2.2 Costruzione del *dataset* e descrizione delle variabili

I *database* intermedi importati in Stata sono stati oggetto di ulteriori operazioni di pulizia e ristrutturazione. In primo luogo, è stato costituito un *dataset*



UNIVERSITÀ DI PAVIA

comprensivo di tutti gli UUID, sia dei fondatori che delle organizzazioni, organizzato sulla base degli UUID di queste ultime, ed è stato filtrato mantenendo esclusivamente le osservazioni che:

- avevano natura *for-profit*;
- risultavano attive al momento della rilevazione;
- sono state fondate a partire dal 2010 fino al 2020;
- Avevano dati sul *funding stage*.

I criteri consentono di delimitare il campione a imprese comparabili, mirati a identificare imprese potenzialmente esposte al rischio di *exit* che hanno operato in un contesto economico relativamente omogeneo, riducendo al contempo le distorsioni dovute a imprese inattive o appartenenti a modelli organizzativi non profit. Il periodo osservato è relativo alla scelta di escludere le imprese che, ad oggi, potrebbero non aver ancora raggiunto fasi avanzate del ciclo di vita a causa della crescita non uguale per tutte. Il *dataset* risultante, ridotto a 47.808 osservazioni, costituisce la base di riferimento per tutte le successive elaborazioni.

Successivamente, è stato costruito il *dataset* di collegamento tra fondatori e organizzazioni. Lo scopo era creare una struttura relazionale *many-to-many* tra fondatori e imprese, indispensabile per le analisi a livello di *team*



UNIVERSITÀ DI PAVIA

imprenditoriale. Il *dataset*, comprensivo di organizzazioni e fondatori, è stato ridotto alle sole variabili identificative: l'UUID dell'organizzazione e gli UUID dei fondatori, ottenendo una riga per ciascuna coppia impresa–fondatore. Le osservazioni prive di identificativo sono state eliminate, garantendo la validità dei collegamenti.

Per quanto riguarda le variabili, i dati originali contenuti nei *database* intermedi che si presentavano in formati non standardizzati sono stati codificati e trasformati in variabili numeriche o *dummy* per una rappresentazione coerente e comparabile di ciascuna categoria di interesse. Questo processo ha incluso la selezione delle informazioni rilevanti, l'aggregazione di valori multipli in capo allo stesso soggetto e la creazione di indicatori binari o categoriali, producendo *dataset* puliti e pronti per l'analisi statistica. Il *dataset* risultante è stato utilizzato in tutte le fasi successive di aggregazione delle caratteristiche individuali dei fondatori a livello di impresa.

La Tabella 1 di seguito riassume le variabili impiegate nel modello. La costruzione operativa, la logica di misurazione adottata e la coerenza con il quadro teorico delineato sono descritte in seguito, facendo riferimento ai concetti di capitale umano e di composizione del *team* discussi nel Capitolo 1.



UNIVERSITÀ DI PAVIA

Tabella 1: Descrizione delle Variabili

Nome variabile	Variabile nel dataset	TIPO variabile	Caratteristiche	Formato
Presenza femminile	female_presence	controllo	1=almeno una presenza femminile nel <i>team</i>	int
Residenza fondatore	presence_north_american_founder	controllo	1=almeno un fondatore residente in Nord America	int
Età impresa	org_age	controllo	Differenza tra anno 2026 e anno fondazione	int
Dimensioni del <i>team</i>	<i>team_size</i>	controllo	Numero componenti <i>team</i>	int
Titolo di studio medio	mean_study_title	indipendente	Media dei titoli di studio dei fondatori per <i>team</i>	float
Eterogeneità delle esperienze lavorative	blau_team_job_exp	indipendente	Indice di Blau calcolato tra il numero di settori eterogenei dei membri di ogni <i>team</i>	float
Quantità esperienza accumulata	sum_prev_org_founded_team	indipendente	Somma complessiva delle organizzazioni fondate dai membri di uno stesso <i>team</i> esclusa l'unità di analisi	int
Successo	success	dipendente	1 = raggiungimento di <i>IPO</i> o <i>M&A</i> o late funding stage	int

Fonte: Elaborazione dell'autore

2.2.1 Variabili controllo

Per quanto riguarda le variabili di controllo, sono state incluse per isolare l'effetto delle variabili indipendenti sulla probabilità di *exit* in quanto attenuanti di distorsioni causate da fattori contestuali. A livello di fondatore, sono



UNIVERSITÀ DI PAVIA

considerati (i) il genere, (ii) la residenza geografica e (iii) il numero totale di fondatori per *team*. Queste dimensioni sono state aggregate per organizzazione.

In particolare, per quanto riguarda il genere, la dummy sulla presenza di una fondatrice donna è positiva se nel *team* c'è almeno una fondatrice donna. La scelta è dettata dalla necessità di valutare l'eterogeneità del *team* nelle sue caratteristiche e il *gender* rappresenta un importante fattore che influisce sulle decisioni di *networking*, rischio e *performance*. Come evidenziato in letteratura (Gottschalk & Niefert, 2013), infatti, esistono differenze marcate tra imprese fondate da *team* di sole donne o soli uomini. Le imprese femminili presentano risultati leggermente inferiori in termini di capitalizzazione e dimensioni del *team*, mentre presentano livelli medi più elevati di istruzione formale.

La variabile relativa alla residenza dei fondatori segnala la presenza di almeno un fondatore residente negli USA o in Canada nel *team* (Nord America). La residenza geografica, in particolare negli Stati Uniti, può influenzare l'accesso alle reti di finanziamento e alle reti culturali, oltre che l'eterogeneità del *team* (Díaz-Santamaría & Bulchand-Gidumal, 2021). La letteratura evidenzia che le differenze geografiche nei fondatori sono impattanti, il vantaggio dell'ecosistema imprenditoriale statunitense rispetto ad altre economie innovative deriva da diverse fonti che producono guadagni considerevoli per le startup (Conti & Guzman, 2023).



UNIVERSITÀ DI PAVIA

L'età dell'impresa è stata calcolata come la differenza tra l'anno corrente, 2026, e l'anno di fondazione. *Le startup* più mature possono avere maggiori possibilità di *exit*, semplicemente perché è passato più tempo. In particolare, secondo quanto emerge dallo studio di (Díaz-Santamaría & Bulchand-Gidumal, 2021), l'età della *startup* emerge come variabile centrale: una maggiore anzianità aumenta la probabilità di raggiungere fatturati significativi e di ottenere finanziamenti. Controllare questo aspetto consente di eliminare l'effetto spurio.

Successivamente, è stata calcolata la dimensione del *team* imprenditoriale come numero totale dei fondatori associati a ciascuna organizzazione. La dimensione incide sulla capacità decisionale, sul capitale sociale e sulle reti e specializzazioni dei membri. Secondo quanto emerso dalla letteratura (Shrivastava & Tamvada, 2011), la dimensione del *team* ha una relazione a campana con il numero medio empiricamente osservato nel loro test, che si attesta a 3. La relazione con la *performance* aziendale delle imprese da loro osservate suggerisce che i *team* di dimensioni superiori alla media hanno *performance* inferiori.

Queste variabili di controllo permettono di isolare l'effetto delle caratteristiche del *team* fondatore tenendo conto di fattori demografici e strutturali potenzialmente rilevanti. Le variabili indipendenti sono tre indicatori



UNIVERSITÀ DI PAVIA

che esprimono le caratteristiche del capitale umano e la composizione del *team* fondatore.

2.2.2 Variabili indipendenti

Come indicato dalla teoria di Becker, la formazione accademica è un elemento centrale del capitale umano, correlata a capacità cognitive, tecniche e organizzative, e deve essere considerata come un investimento produttivo del capitale umano. Basandomi su questi criteri, il titolo di studio accademico di ciascun fondatore è stato codificato come misura della distanza ordinata nel capitale umano formale, in linea con quanto comunemente adottato nella letteratura empirica (Unger et al., 2011). I titoli di studio sono stati codificati come una variabile ordinale numerica secondo una scala crescente ispirata alla Classificazione Internazionale Standard dell'Istruzione (ISCED). In particolare, le lauree triennali o equivalenti (*bachelor's degree* - BA) sono state assegnate il valore 1, le lauree magistrali o *master's degree* (MA) hanno ottenuto il valore 2, mentre i dottorati e titoli superiori, come incarichi di ricerca o insegnamento (PhD), hanno ricevuto il valore 3. I titoli equiparabili sono stati raggruppati sotto uno stesso valore, poiché nel contesto analizzato non si riteneva necessario distinguere tra differenze di disciplina. Analogamente, i titoli superiori al dottorato, essendo ottenibili solo previa acquisizione di un PhD, sono stati inclusi nella stessa categoria. Questi dati sono stati successivamente aggregati a



UNIVERSITÀ DI PAVIA

livello di *team* imprenditoriale attraverso un indicatore sintetico calcolato come media aritmetica dei valori assegnati, che rappresenta il livello medio di istruzione del *team* dei fondatori. Tale indicatore, denominato titolo di studio medio, è una variabile continua numerica che risponde all'Ipotesi H1 secondo la quale a livelli di istruzione più elevati, aggregati a livello di *team*, si possa aumentare la probabilità di successo.

Le meta-analisi presentate nel capitolo precedente relative ai *team* imprenditoriali mostrano che l'eterogeneità delle esperienze ha effetti positivi sulla *performance*: sembrerebbe generare benefici in termini di problem solving e capacità di elaborazione delle informazioni, pur potendo comportare costi di coordinamento ed eventuale conflitto con effetti non lineari. La letteratura dunque conferma che il capitale umano legato alle esperienze lavorative è più fortemente associato al successo rispetto a misure puramente di investimento come gli anni di istruzione (Unger et al., 2011). Gli studi di Beckman e Burton (2008) sul *Path Dependence* evidenziano che l'esperienza settoriale specifica e la varietà dei percorsi di carriera dei fondatori favoriscono la crescita in termini di attrazione di *manager* esperti e accesso anticipato a *venture capital*. Tenuto conto delle evidenze precedenti, si può affermare che l'eterogeneità dei percorsi genera vantaggi lungo il ciclo di vita d'impresa. Nel complesso, la quantità di esperienze diverse può rappresentare un determinante chiave da osservare



UNIVERSITÀ DI PAVIA

empiricamente (Jin et al., 2017). Per fare ciò, è stata costruita la variabile dell'eterogeneità delle esperienze lavorative, un indice di diversità (Blau index) calcolato sulla base del numero di esperienze lavorative passate dei fondatori, escludendo i fondatori senza dati disponibili e aggregando a livello di *team*. In una prima fase, si è proceduto a uniformare le variabili settoriali disponibili mediante un ulteriore processo di pulizia testuale, al fine di evitare duplicazioni artificiali di categorie. Successivamente, sono stati individuati tutti i settori distinti presenti nel *dataset* e, per ciascuno di essi, è stata creata una variabile *dummy* che segnala la presenza di esperienza per ogni singolo fondatore. A partire da tali *dummy*, si è calcolato:

- il numero di settori distinti per singolo fondatore (*n_settori*): k ;
- l'indice di Blau, definito come:

$$Blau = 1 - \sum_i^k p_i^2$$

Dove p_i rappresenta la quota di esperienza attribuibile al settore i .

L'indice così calcolato misura il grado di eterogeneità settoriale dell'esperienza professionale del singolo fondatore, assumendo valori crescenti al crescere della varietà. Attraverso il *dataset* di collegamento fondatori–imprese si è aggregato il tutto a livello di organizzazione tramite una funzione di



UNIVERSITÀ DI PAVIA

massimo, identificando la presenza di ciascun settore all'interno del *team*. A partire da queste informazioni sono stati ricalcolati:

- il numero di settori distinti coperti dal *team* ($n_settori_team$): k ;
- l'indice di Blau medio a livello di *team* ($blau_team$):

$$Blau_team = 1 - \sum_{j=1}^k p_j^2$$

dove $p_j = \frac{n_j}{N}$

con:

- n_j che rappresenta il numero di fondatori con esperienza nel settore generico j ;
- $IniziaN$ che rappresenta il numero totale di membri del *team*.

Il risultato è una misura sintetica dell'eterogeneità settoriale complessiva del *team* imprenditoriale. L'indice di Blau non misura la quantità di esperienza, ma la sua distribuzione tra ambiti differenti, catturando la complementarità potenziale delle competenze all'interno del *team* (Beckman & Burton, 2008). Il valore della variabile oscilla tra lo 0, che corrisponde ad omogeneità totale nelle esperienze di un *team*, ed un valore teorico $1 - \frac{k}{1}$ che corrisponde ad eterogeneità totale tra i settori coperti da un *team*.



UNIVERSITÀ DI PAVIA

La variabile sull'esperienza fondativa rappresenta la somma complessiva delle organizzazioni fondate dai membri di uno stesso *team*, al netto dell'impresa oggetto di analisi. Come discusso nel Capitolo 1, la letteratura sull'imprenditorialità seriale evidenzia come l'esperienza fondativa pregressa e il track record dei fondatori siano associati a una maggiore probabilità di accesso a finanziamenti avanzati e di raggiungimento di *exit* positive (Gompers et al., 2010; Hsu, 2007). Nel presente studio non è possibile distinguere con certezza la sequenza temporale delle imprese fondate, in assenza di informazioni longitudinali complete. Di conseguenza, la variabile non misura strettamente la serialità imprenditoriale in senso cronologico, bensì l'intensità complessiva dell'esperienza fondativa accumulata dal *team*. Tale misura può essere interpretata come indicatore di capitale umano imprenditoriale cumulato (Unger et al., 2011) o di imprenditorialità abituale (Ucbasaran et al., 2008), riflettendo la propensione dei membri del *team* a intraprendere iniziative imprenditoriali e la loro esposizione a processi di creazione d'impresa.

Operativamente, dopo aver selezionato e convertito correttamente le variabili numeriche relative al numero di organizzazioni fondate da ciascun individuo, dall'indicatore individuale è stata sottratta l'unità di analisi, al fine di evitare sovrastime meccaniche dell'esperienza. La misura individuale così ottenuta è stata successivamente aggregata a livello di *team* mediante la somma,



UNIVERSITÀ DI PAVIA

utilizzando il dataset di collegamento tra fondatori e imprese. Il risultato è un indicatore continuo che cattura l'intensità accumulata dell'esperienza imprenditoriale del *team* fondatore, coerentemente con l'ipotesi teorica secondo cui una maggiore dotazione di capitale umano esperienziale può aumentare la probabilità di successo dell'impresa.

Tutti i *database* costruiti per ogni singola variabile sono stati uniti in un unico *dataset* finale, utilizzato per le analisi statistiche e di regressione. Attraverso una serie di fusioni sequenziali è stato possibile combinare le informazioni relative al capitale umano e le variabili di controllo, armonizzandole a livello organizzativo. Il *dataset* finale è stato sottoposto a controlli di coerenza e di pulizia delle variabili, con verifica della presenza di dati mancanti, e rinominato per migliorarne la leggibilità e l'interpretazione economica. Il risultato è un *dataset* strutturato e coerente di 15.035 osservazioni, che costituisce la base empirica dell'analisi quantitativa presentata nei capitoli successivi della tesi.

2.2.3 Variabile dipendente

In linea con la letteratura sul ciclo di vita delle *startup VC-backed* (Giot & Schwienbacher, 2007; Reisdorfer-Leite et al., 2020), è stata operazionalizzata la variabile del successo come il raggiungimento di un'*exit* positiva (*IPO* o *M&A*), in quanto evento osservabile (Pisoni & Onetti, 2018), o lo stadio avanzato del



UNIVERSITÀ DI PAVIA

finanziamento che vedrà un'uscita in futuro. La variabile dipendente, quindi, assume valore 1 nel caso in cui l'impresa abbia raggiunto uno stadio di *late stage funding*, *M&A* o *IPO* o superiore, e valore 0 negli altri casi. Dal punto di vista statistico, la variabile è dicotomica; ciò riduce il bias soggettivo; infatti, l'osservazione di un evento specifico permette la replicabilità ed evita che le misure non siano comparabili. La probabilità condizionata è funzione delle caratteristiche osservabili dell'impresa:

$$Y_i \in \{0,1\}, P(Y_i = 1 | X_i)$$

Il campione che è stato utilizzato per la costruzione del *dataset* strutturato non è filtrato *ex post* sulla base dell'esito del *funding stage*; pertanto, non sono state selezionate solo le imprese che rispettano l'esito positivo dell'evento. Sono state incluse nel campione finale anche *startup* che non hanno ancora raggiunto il *late stage funding*. Il fine è evitare distorsioni nella selezione e consentire una stima corretta della probabilità di successo (Wooldridge, 2010).

La specificazione e la costruzione delle variabili indipendenti e di controllo qui descritte garantiscono che il modello empirico sia coerente con i meccanismi teorici delineati nel Capitolo 1, consentendo una stima robusta mediante regressione logistica.



UNIVERSITÀ
DI PAVIA

Capitolo 3: Risultati analisi

3.1 Caratteristiche del campione scelto

Dopo ulteriori pulizie del *dataset*, quello finale comprende 9.637 osservazioni su imprese che hanno effettuato almeno un *funding stage* tra *pre-seed* e *late stage* (M&A o IPO), ottenute tramite il processo descritto nel Capitolo 2. Il campione è adeguato rispetto al numero di regressori inclusi nei modelli successivi, garantendo un rapporto tra osservazioni e variabili ampiamente superiore a quanto suggerito dalla letteratura (Greene, 2018; Wooldridge, 2010). In Tabella 2 sono riportate le variabili non categoriali e le relative statistiche descrittive approssimate a due cifre decimali, descritte in dettaglio in seguito:

Tabella 2: Statistiche descrittive variabili

Variabile	Media	Dev.Std	Min	Max	Mediana	Skewness	Kurtosis
Età impresa	10,15	2,82	6	16	10	0,3	2,14
Dimensione del <i>team</i>	2,41	1,17	1	11	2	1,26	6,42
Eterogeneità esperienze lavorative	0,76	0,17	0	0,95	0,86	-2,64	11,66
Quantità esperienza accumulata	1,63	2,99	0	69	1	7,84	120,01
Titolo di studio medio	1,61	0,55	1	3	1,5	0,68	2,94
Presenza femminile	0,24	0,43	0	1	0	1,21	2,48
Residenza fondatore	0,64	0,48	0	1	1	-0,62	1,38

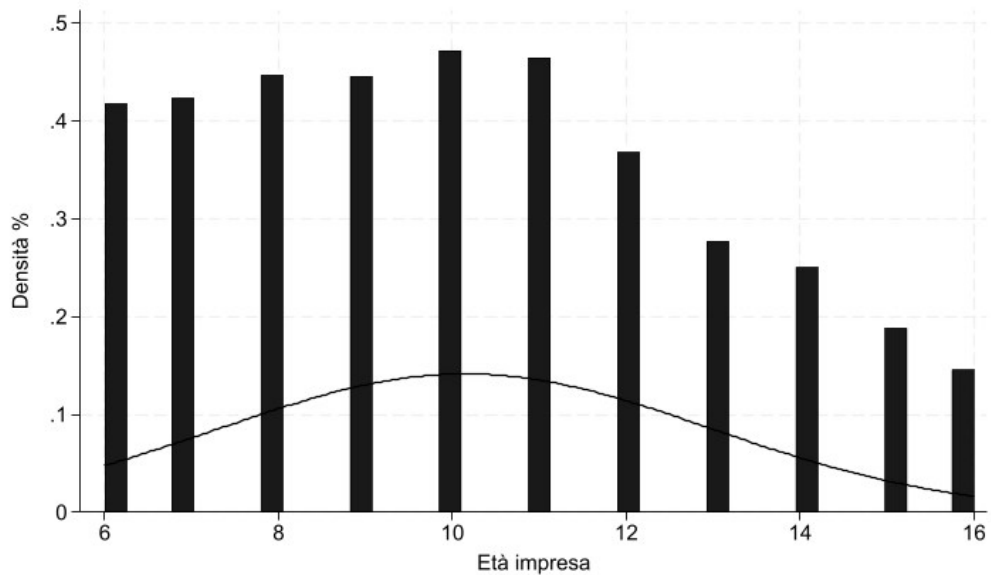
Fonte: Elaborazione dell'autore



UNIVERSITÀ DI PAVIA

L'età media delle imprese osservate è pari a 10,15 anni, con un intervallo compreso tra 6 e 16 anni. La deviazione standard pari a 2,82 colloca la maggioranza delle imprese tra i 7 ed i 13 anni circa, con fondazione collocabile tra il 2013 ed il 2019, come visibile nel Grafico 1:

Grafico 1: Distribuzione organizzazione age



Fonte: Elaborazione dell'autore

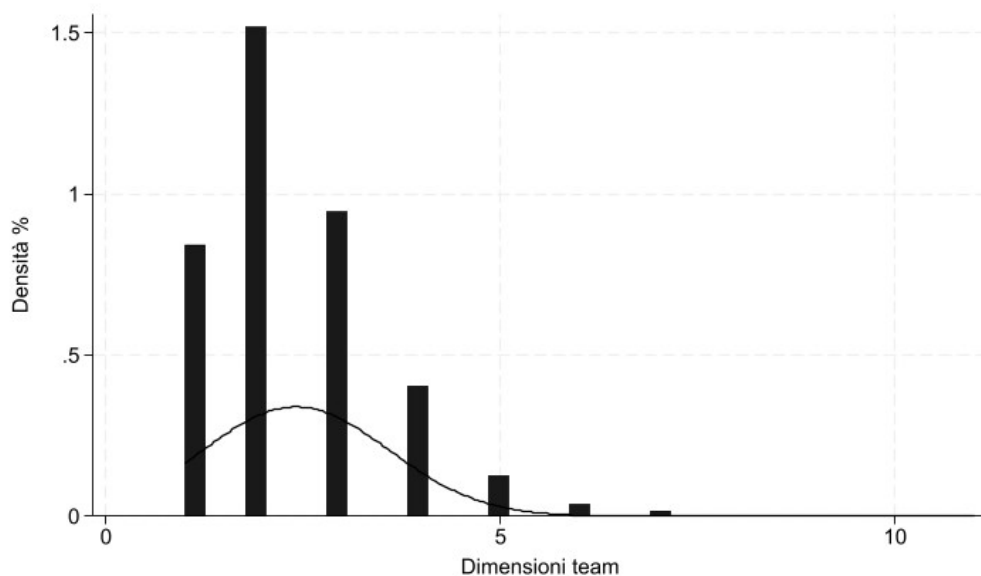
La distribuzione presenta asimmetria contenuta ($\text{skewness} = 0,3$) e il grado di curtosi, inferiore a 3 (2,82), evidenzia una coda meno pronunciata rispetto alla normale. Dal punto di vista teorico, la mancanza di *outlier* estremi è positiva; essi potrebbero infatti influenzare le stime dei modelli di regressione (Wooldridge, 2010).



UNIVERSITÀ DI PAVIA

La dimensione media dei *team* è di 2,41 membri (mediana = 2) e la loro distribuzione è positivamente asimmetrica (skewness = 1,26). I *team* risultano quindi generalmente di piccole dimensioni; la maggior parte si colloca tra 2 e 3 membri. Solo 4 *team* rientrano nel numero massimo di 11 membri. Come visibile nel Grafico 2, il grado di curtosi è molto alto (6,42); questo suggerisce la presenza di outlier marcati, ma si tratta di fenomeni rari e marginali sul totale delle imprese osservate (Wooldridge, 2010).

Grafico 2: Distribuzione dimensioni dei team



Fonte: Elaborazione dell'autore

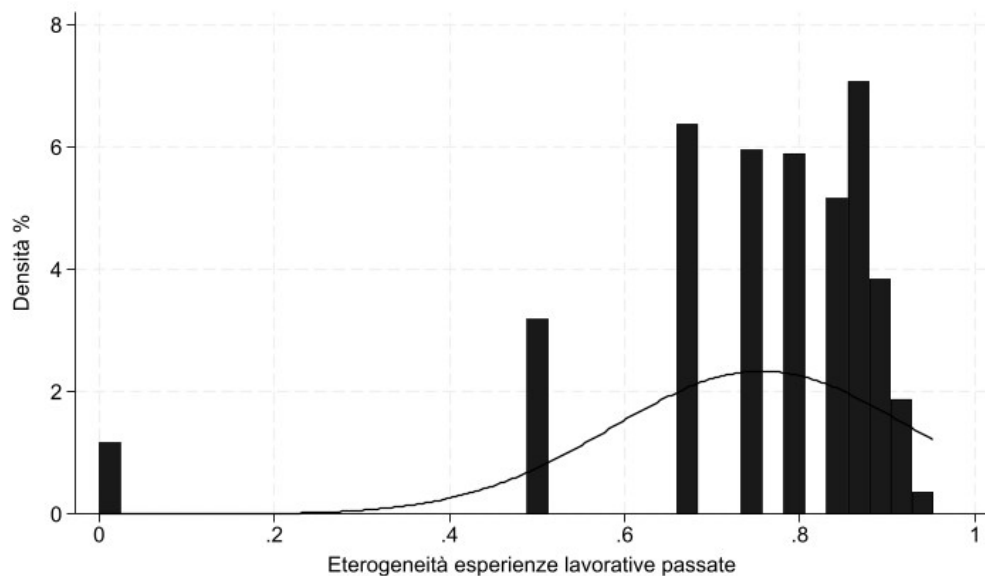
Il grado di eterogeneità settoriale complessiva del *team* ha un valore medio di 0,76, compreso tra 0 e 0,95. Solo 277 imprese sono distribuite al valore



UNIVERSITÀ
DI PAVIA

minimo, che indica la totale omogeneità. La distribuzione è negativamente asimmetrica ($\text{skewness} = -2,64$), come visibile nel Grafico 3.

Grafico 3: Distribuzione eterogeneità esperienze lavorative team



Fonte: Elaborazione dell'autore

Il grado elevato di curtosi pari a 11,66 conferma che la concentrazione è attorno ai valori elevati, ma, trattandosi di un indice compreso nell'intervallo $[0,1]$, le caratteristiche sono coerenti con la natura della variabile (Blau, 1977).

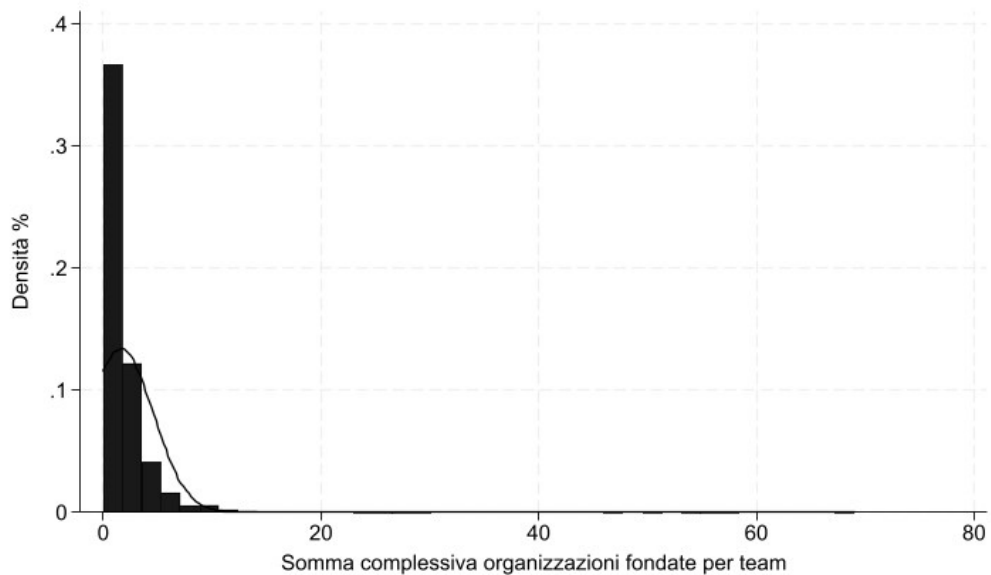
Anche la distribuzione del numero di imprese fondate per *team* è fortemente asimmetrica, con una lunga coda verso destra, come visibile nel Grafico 4. Il numero massimo di 69 imprese è associato a un'unica impresa. La media è pari a 1,63 imprese per *team* e la deviazione standard è pari a 2,99.



UNIVERSITÀ DI PAVIA

Pochissimi *team* presentano un'elevata intensità di esperienza accumulata. Complessivamente, il numero di osservazioni superiori alla media è pari a 3.388 (pari al 35%); tuttavia, di queste, il valore medio è pari a 4,04.

Grafico 4: Distribuzione intensità accumulata di esperienza imprenditoriale



Fonte: Elaborazione dell'autore

Data l'elevata asimmetria, con un grado di curtosi pari a 120, la maggioranza dei *team* è alla prima esperienza imprenditoriale; l'impatto dei valori estremi verrà monitorato successivamente nei modelli econometrici.

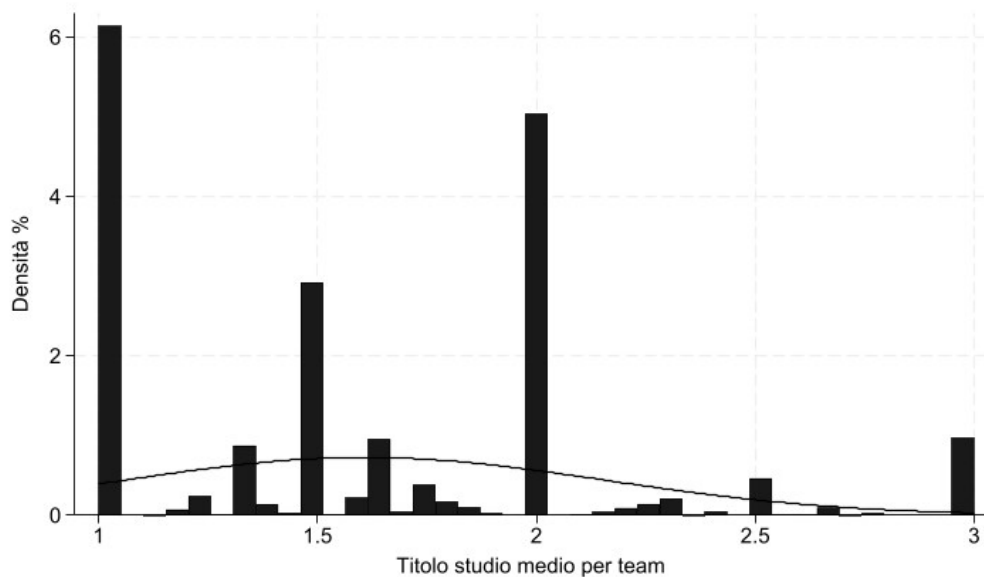
Infine, la variabile relativa al titolo di studio medio è ordinale (1 = *Bachelor*, 2 = *Master*, 3 = *PhD*) aggregata a livello di *team*; per questo motivo può essere trattata come variabile continua, assumendo il valore equidistante tra



UNIVERSITÀ DI PAVIA

i valori delle categorie (Wooldridge, 2010). La distribuzione nel Grafico 5 mostra che i valori si collocano mediamente tra la laurea triennale e quella magistrale; non c'è un'elevata incidenza di titoli di studio elevati come PhD; infatti, la media è pari a 1,61 (dev. std. = 0,55), e la mediana è esattamente pari a 1,5.

Grafico 5: Distribuzione titoli di studio medi per team



Fonte: Elaborazione dell'autore

Si evidenzia che almeno il 25% dei *team* ha un titolo medio di valore pari a 2, mentre soltanto 481 *team* (circa il 5%) presentano tutti i membri con valore 4. La distribuzione è positivamente asimmetrica, verso i valori minori, ma non sono presenti outlier eccessivamente marcati (curtosi = 2,94).



UNIVERSITÀ DI PAVIA

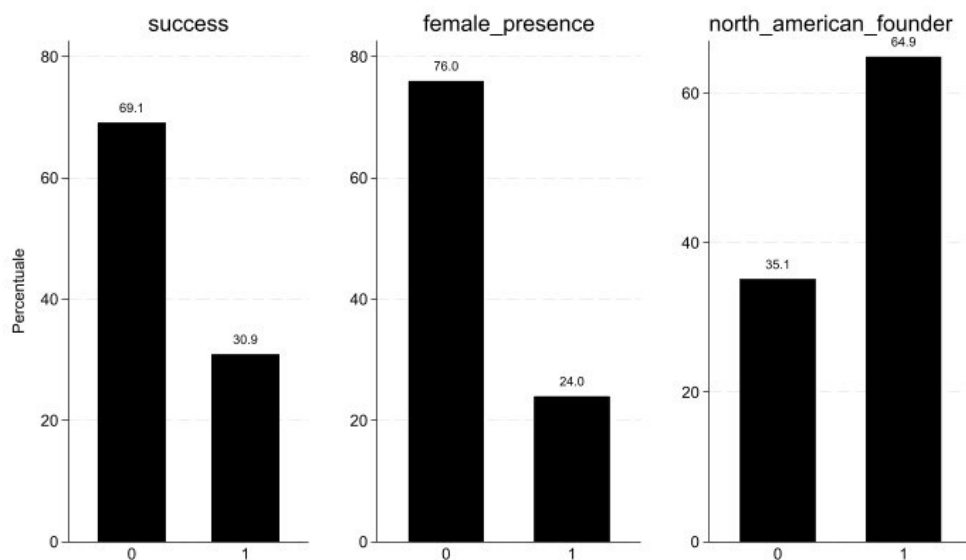
Per quanto riguarda le variabili categoriali, nella Tabella 3 sono riportate le frequenze osservate nel campione disponibile per la positività dell'evento. Nel grafico 6 sono rappresentate come distribuzione percentuale. Si approfondiscono questi risultati nel capitolo 4, paragrafo 4.1.

Tabella 3: Frequenze variabili categoriali

Variabile	Frequenza	Percentuale
Successo =1	2983	30,9%
Presenza femminile =1	2313	24%
Fondatore residente in Nord America =1	6264	64,89%

Fonte: Elaborazione dell'autore

Grafico 6: Distribuzione variabili categoriali



Fonte: Elaborazione dell'autore



UNIVERSITÀ DI PAVIA

Il successo dell'impresa (success) è un valore positivo per il 30,9% delle *startup* osservate (2.983 su 9.637), mentre il 69,1% ha esito negativo. Il raggiungimento del *late stage funding*, quindi, non è molto frequente; tuttavia, per la regressione logistica è accettabile, poiché non si incorre nel bias dell'evento eccezionale (Wooldridge, 2010). Il bilanciamento dell'evento è di circa 1:2,2. La percentuale sembra superiore alla media degli articoli analizzati, che solitamente riportano percentuali attorno all'11%. (Potanin et al., 2023).

Osservando le differenze tra le caratteristiche delle *startup* di successo (success 1) e di non successo (success 0) nella Tabella 4, è possibile notare alcuni elementi i cui dati variano marcatamente: l'età media delle imprese di successo è maggiore di 1,94 anni rispetto alla controparte. Un altro dato interessante è la *skewness*, che diventa negativa; per un valore positivo si trasforma in un valore negativo pari a -0,15. L'asimmetria della distribuzione è bassa, ma opposta a quella della controparte.



UNIVERSITÀ
DI PAVIA

Tabella 4: Confronto campione

	Età impresa	Dim. team	Titolo studio medio	Eterogeneità esperienze	Esperienza imp. cumulata	Presenza Femminile	Presenza residenza in Nord America
success 0							
Mean	9.55*	2.33	1.60	0.75	1.42	0.26	0.62
Dev. Std.	2.66	1.12	0.55	0.17	2.60	0.44	0.49
Min	6.00	1.00	1.00	0.00	0.00	0	0
Max	16.00	11.0 0	3.00	0.95	54.00	1	1
Skewness	0.51*	1.20	0.68	-2.61	7.91	1.07	-0.48
Kurtosis	2.45	6.22	2.95	11.30	126.12	2.14	1.23
success 1							
Mean	11.49*	2.57	1.62	0.76	2.11	0.19	0.72
Dev. Std.	2.72	1.27	0.55	0.16	3.65	0.39	0.45
Min	6.00	1.00	1.00	0.00	0.00	0	0
Max	16.00	11.0 0	3.00	0.95	69.00	1	1
Skewness	-0.15*	1.28	0.67	-2.68	7.24	1.22	-0.98
Kurtosis	2.14	6.36	2.93	12.49	97.88	2.48	1.96

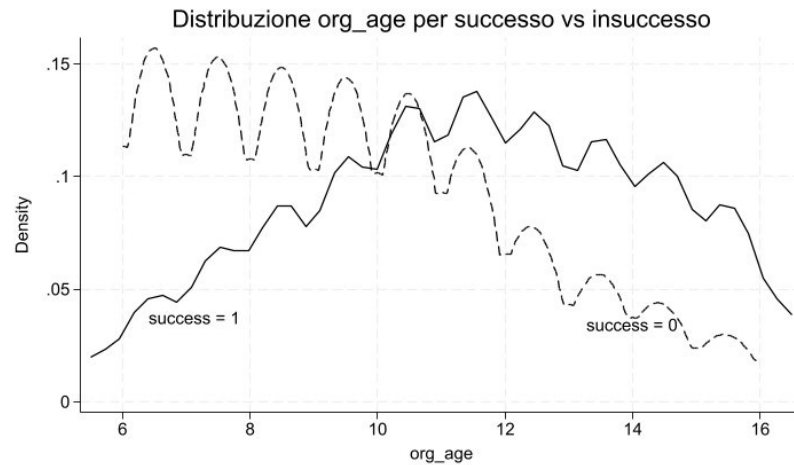
Fonte: Elaborazione dell'autore

Nel Grafico 7 si mostra il confronto tra la distribuzione delle *startup* di successo e quella delle *startup* non ancora di successo o insuccesso.



UNIVERSITÀ
DI PAVIA

Grafico 7: Distribuzioni età imprese tra successo e insuccesso



Fonte: Elaborazione dell'autore

Le differenze per le altre variabili sono contenute: la dimensione media dei team varia leggermente e aumenta in caso di successo; allo stesso modo, la concentrazione di esperienza fondativa varia. L'indice di Blau resta pressoché invariato, con entrambi i gruppi ad alta asimmetria e a concentrazioni ai poli estremi.



UNIVERSITÀ DI PAVIA

Per effettuare una comparazione dei dati efficace è necessario svolgere l'analisi tramite matrice di correlazione tra le variabili:

Figura 7: Matrice di correlazione

	Età impresa	Dim. team	Presenza femminile	Presenza residenza Nord America	Titolo studio medio	Eterogeneità esperienze	Esperienza imp. cumulata
Età impresa	1,00						
Dim. team	0,03	1,00					
Presenza femminile	-0,04	0,10	1,00				
Presenza residenza Nord America	0,06	-0,10*	0,09	1,00			
Titolo studio medio	-0,01	0,02	0,02	-0,10	1,00		
Eterogeneità esperienze	0,04	0,25*	0,01	0,05	-0,01	1,00	
Esperienza imp. cumulata	0,03	0,27*	-0,02	0,07	-0,02	0,12*	1,00

Fonte: Elaborazione dell'autore

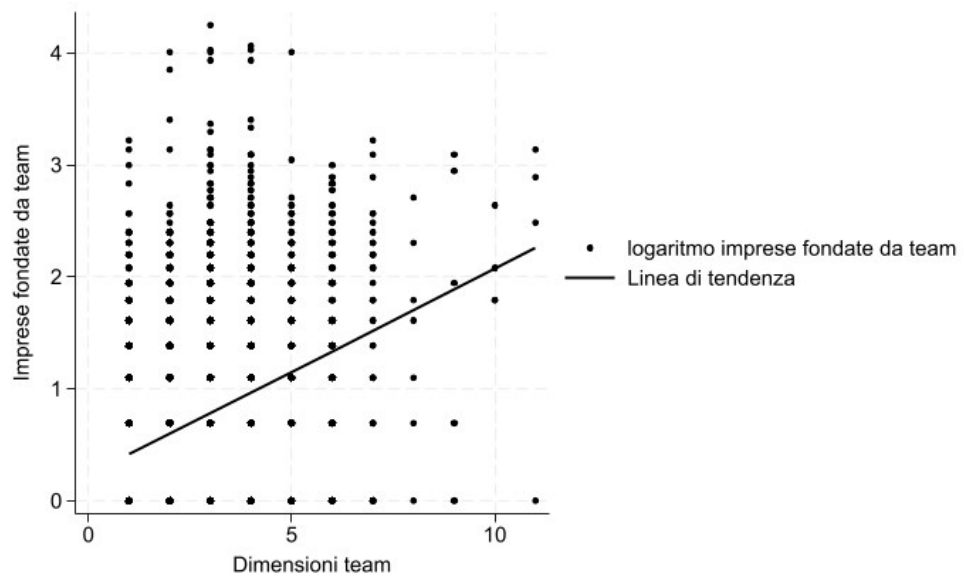
È possibile notare che le correlazioni tra le altre variabili sono molto basse e, in generale, dalla matrice non emergono particolari criticità; nessuna correlazione supera 0,30. Tuttavia, a causa dell'elevata asimmetria evidenziata in precedenza in alcune variabili, è necessario agire con cautela nell'interpretazione delle correlazioni lineari, che potrebbero essere influenzate dalla presenza di *outlier*. Per questo motivo sono state effettuate delle trasformazioni logaritmiche sulle variabili con maggiore asimmetria $\log(1 + X)$ per diminuirne gli effetti degli estremi:



UNIVERSITÀ
DI PAVIA

- dimensioni del *team* – esperienza imprenditoriale accumulata = 0,27. Si è preso in considerazione la variabile con asimmetria maggiore per effettuare $\log(1 + \text{sum_prev_org_founded_team})$. La dispersione ottenuta è accettabile e la tendenza è lineare come visibile nel Grafico 8 qui sotto:

Grafico 8: Dispersione con tendenza *team_size* - $\log(1 + \text{sum_prev_org_founded_team})$



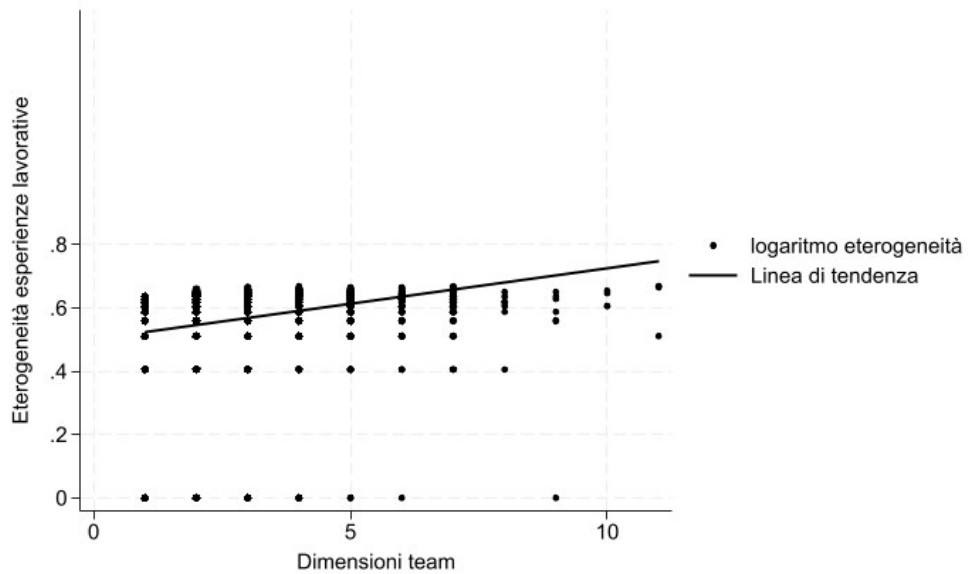
Fonte: Elaborazione dell'autore

- dimensioni del *team* – eterogeneità delle esperienze lavorative = 0,25. Analogamente a quanto effettuato in precedenza, si procede ad osservare la dispersione in seguito alla trasformazione logaritmica della variabile con distribuzione asimmetrica più accentuata $\log(1 + \text{blau_team_job_exp})$. Anche per questa coppia di variabili, la dispersione è accettabile e la tendenza è lineare:



UNIVERSITÀ
DI PAVIA

Grafico 9: Distribuzione con tendenza $\text{team_size} - \log \text{blau_team_job_exp}$



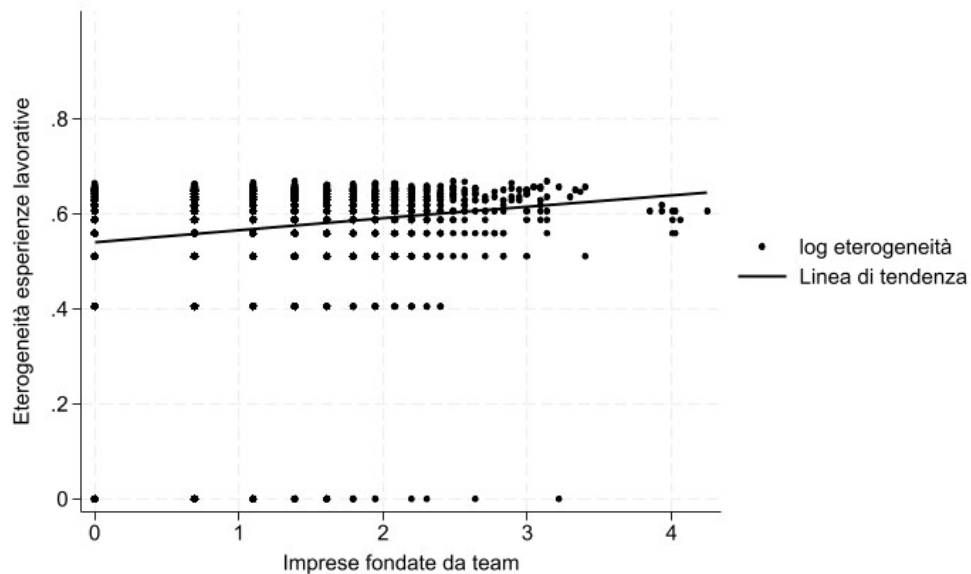
Fonte: Elaborazione dell'autore

- eterogeneità esperienze lavorative - esperienza imprenditoriale accumulata = 0,12; questa debole correlazione positiva è il risultato di una coppia di variabili altamente asimmetriche che necessita di una trasformazione logaritmica per entrambe per valutare la veridicità di questa correlazione. La dispersione ha tendenza lineare ed è accettabile come si può osservare nel grafico 10 sottostante:



UNIVERSITÀ
DI PAVIA

Grafico 10: dispersione con linea di tendenza $\log \text{blau_team_job_exp} - \log \text{sum_prev_org_founded_team}$



Fonte: Elaborazione dell'autore

- dimensioni del team – fondatore residente in Nord America = -0,1; non richiede particolari attenzioni.

In conclusione, dalla matrice di correlazione è evidente che le correlazioni moderate sono legate alle strutture aggregative delle variabili. Inoltre, le stesse non sono tra loro sovrapposte concettualmente, così come i regressori. La variabile dell'esperienza imprenditoriale accumulata ha superato la prova di robustezza mediante trasformazione logaritmica e non evidenzia problemi particolari di correlazione con altre variabili.



UNIVERSITÀ
DI PAVIA

3.2 Regressione logistica

Come discusso nel paragrafo precedente, sebbene siano state necessarie alcune trasformazioni logaritmiche ai fini dell'analisi descrittiva, non si evidenziano particolari problemi di distribuzione tali da richiedere trasformazioni anche nei modelli di regressione (Wooldridge, 2010). Di conseguenza, le variabili sono state incluse nella loro forma originale.

Prima della stima dei modelli *logit*, nonostante i risultati della matrice di correlazione lineare, è stata verificata l'assenza di multicollinearità tra le variabili per un ulteriore controllo. A tal fine, è stata stimata una regressione lineare ausiliaria comprendente tutte le variabili per calcolare i *Variance Inflation Factors* (VIF) (Hair et al., 2019):

$$VIF_j = \frac{1}{1 - R_j^2}$$

dove: R_j^2 è l' R^2 della regressione della variabile X_j su tutte le altre variabili indipendenti.

Come riportato nella Tabella 5, tutti i valori di VIF risultano ampiamente inferiori alla soglia massima consigliata di 10, "la cui tolleranza corrisponderebbe a errori standard gonfiati di oltre tre volte" (Hair et al., 2019, p. 317). Infatti, il fattore medio di varianza dell'inflazione è pari a 1,06. È un



UNIVERSITÀ
DI PAVIA

risultato positivo: indica l'assenza di multicollinearità rilevante e consente l'inclusione simultanea di tutte le variabili nei modelli.

Tabella 5: Multicollinearità

Variabile	VIF
Dimensione <i>team</i>	1,15
Esperienza imp. Accumulata	1,09
Eterogeneità esperienze	1,07
Presenza residente Nord America	1,03
Presenza femminile	1,02
Titolo di studio medio	1,01
Età impresa	1,01
Mean VIF	1,06*

Fonte: Elaborazione dell'autore

L'analisi empirica si articola in 5 modelli di regressione logistica binaria stimati in forma annidata secondo l'approccio standard consigliato da (Wooldridge, 2010). In primo luogo, è stato adottato un modello *baseline* M0 che include esclusivamente le variabili di controllo, al fine di isolarne l'effetto. Successivamente, M1, M2, M3 estendono il modello; ciascuno aggiunge a M0 una specifica dimensione del capitale umano dei *team* per testare le Ipotesi M1–M3 presentate nel Paragrafo 1.2. Infine, attraverso il modello completo M4, che include tutte le variabili di interesse, si osservano le differenze tra i modelli (Wooldridge, 2010).



UNIVERSITÀ
DI PAVIA

La funzione generale del modello logistico è la seguente:

$$\log \left(\frac{P(\text{success}_i = 1)}{1 - P(\text{success}_i = 1)} \right) = \sum_{k=1}^K \beta_k X_{ki}$$

success_i È la variabile binaria positiva in caso di successo e X_{ki} È il vettore delle variabili e β_k rappresenta i coefficienti stimati sulla scala dei log-odds; il modello rappresenta il logaritmo del rapporto delle probabilità.

Tabella 6: Regressione logistiche Odds ratio

Variable	M0	M1	M2	M3	M4
Età impresa	1.286*** (0.0110)	1.286*** (0.0111)	1.287*** (0.0111)	1.286*** (0.0111)	1.288*** (0.0112)
Dim. Team	1.207*** (0.0247)	1.206*** (0.0247)	1.225*** (0.0259)	1.163*** (0.0249)	1.180*** (0.0252)
Presenza Femminile	0.590*** (0.0347)	0.588*** (0.0346)	0.587*** (0.0347)	0.600*** (0.0354)	0.594*** (0.0350)
Residenza NordAmerica	1.573*** (0.0796)	1.599*** (0.0818)	1.586*** (0.0804)	1.536*** (0.0780)	1.574*** (0.0812)
Titolo studio medio		1.134*** (0.0487)			1.138*** (0.0491)
Eterogeneità esperienze			0.656*** (0.0945)		0.619*** (0.0884)
Esperienza Imp. Accumulata				1.056*** (0.0108)	1.059*** (0.00922)
N Observations	9637	9637	9637	9637	9637
Wald chi ²	1061.9	1071.9	1068.8	1077.3	1278.3
Log pseudolikelihood	-5355.6	-5351.4	-5351.3	-5333.6	-5323.5
Pseudo R ²	0.1018	0.103	0.103	0.106	0.107
Note: Odds ratios riportati, Std error tra parentesi, Significance levels *** p<0.01, ** p<0.05, * p<0.10					

Fonte: Elaborazione dell'autore



UNIVERSITÀ DI PAVIA

La Tabella 6 riporta i risultati delle regressioni logistiche stimate. Per interpretare gli effetti marginali sul successo, i risultati della regressione sono presentati in termini di *odds ratio*, per un'immediata interpretazione dell'effetto delle variabili (Long & Freese, 2001). I coefficienti stimati sulla scala dei log-*odds* non sono riportati per brevità.

La funzione di massima verosimiglianza converge regolarmente in tutte le stime è stabile a un valore massimo *Log pseudolikelihood* pari a -5355,64; inoltre, il test di Wald indica che i modelli sono tutti significativi poiché ipotizzando il χ^2 varia leggermente da 1061,9 a 1278,3.

Il test dei minimi quadrati non è possibile in quanto non si tratta di una regressione lineare di conseguenza, per misurare la bontà di adattamento del modello, è stato utilizzato lo *Pseudo R²* di McFadden, che confronta la verosimiglianza logaritmica del modello in questione rapportandola al modello *Null* che include solo la costante (McFadden, 1972).

$$Pseudo R^2 = 1 - \frac{\log H_0}{\log Null}$$

Il confronto con i valori di *log-pseudolikelihood* consente di valutare l'*improvement* della bontà di adattamento dei modelli rispetto alla specificazione *baseline*. L'indicatore varia da 0,1018 a 0,107.



UNIVERSITÀ
DI PAVIA

Il test di bontà dell'adattamento di Pearson, necessario per verificare la capacità dei modelli di descrivere adeguatamente il fenomeno osservato, non consente di rifiutare l'ipotesi nulla di corretta specificazione del modello ($p\text{-value} = 0,427$), suggerendo che il modello logistico fornisce una rappresentazione adeguata dei dati osservati ($p > 0,05$), come visibile nelle Tabelle 7 e 8:

Tabella 7: Goodness of fit

	Successo (D)	Insuccesso (~D)	Totale
Predetto positivamente (+)	897 (TP)	628 (FP)	1525
Predetto negativamente (-)	2086 (FN)	6026 (TN)	8112
Totale	2983	6654	9637

Fonte: Elaborazione dell'autore

Dove: TP sono i successi correttamente previsti, FP i successi falsi positivi, FN i falsi negativi e TN gli insuccessi correttamente previsti.

Tabella 8: Goodness of fit con percentuali

Sensitività	$\Pr (+ D) = \frac{TP}{TP+FN}$	30,07%
Specificità	$\Pr (- \sim D) = \frac{TN}{TN+FP}$	90,56%
Predizioni positive	$\Pr (D +) = \frac{TP}{TP+FP}$	58,82%
Predizioni negative	$\Pr (\sim D -) = \frac{TN}{TN+FN}$	74,29%
Grado falso positivi	$\Pr (+ \sim D) = 1 - \frac{FP}{FP+FN}$	9,44%
Grado falso negativi	$\Pr (- D) = 1 - \frac{FN}{TP+FN}$	69,39%



UNIVERSITÀ
DI PAVIA

Quota previsioni positive errate	$\Pr(\sim D +) = \frac{FP}{TP+FP}$	41,18%
Quota previsioni negative errate	$\Pr(D -) = \frac{FN}{TN+FN}$	25,71%

Fonte: Elaborazione dell'autore

Per quanto riguarda la capacità di classificazione, il modello riconosce correttamente circa il 30% delle *startup* che hanno effettivamente raggiunto il successo, evidenziando una presenza significativa di falsi negativi (69,39%). Al contrario, la specificità è elevata, pari a circa il 91%, indicando una buona capacità di identificare le *startup* che non hanno avuto successo. Le previsioni positive risultano accurate in circa il 59% dei casi, mentre quelle negative lo sono nel 74% circa. La probabilità di classificare erroneamente come successo una *startup* che non lo ha raggiunto è contenuta, attestandosi intorno al 9%, mentre la probabilità di non riconoscere una *startup* di successo si aggira attorno al 70%. Inoltre, circa il 41% delle previsioni positive e il 26% delle previsioni negative risultano errati.

L'analisi della matrice di classificazione mostra una buona capacità complessiva di classificazione, trainata da un'elevata specificità, mentre la sensibilità è molto più contenuta, in linea con la natura sbilanciata dell'evento di successo nel campione.

Per ulteriori controlli, è stato effettuato il confronto tra i modelli. È importante notare che il miglioramento degli indicatori di bontà di adattamento



UNIVERSITÀ
DI PAVIA

è molto contenuto passando dal modello baseline (71,68% correttamente classificato) alle specificazioni estese in Tabella 7 ed 8 (71,84%). Questi risultati suggeriscono che le variabili, pur risultando statisticamente significative, incidono principalmente sulle *odds* di successo piuttosto che sulla bontà classificatoria complessiva, in linea con la natura esplicativa dei modelli.

Il Capitolo 4 sarà dedicato all'analisi e alla discussione dei risultati ottenuti, con particolare attenzione alla verifica delle ipotesi di ricerca e alle implicazioni teoriche ed empiriche.

Capitolo 4: Discussione e implicazioni

In questo capitolo vengono sviluppati i risultati ottenuti dai modelli spiegati nel precedente capitolo 3. L'attenzione è rivolta all'effetto delle caratteristiche del capitale umano e della composizione del *team* fondatore sulla probabilità relativa di successo della *startup*.

I risultati saranno utilizzati per verificare le ipotesi di ricerca del Capitolo 1 (H1, H2, H3) rispetto alle evidenze empiriche esistenti in letteratura.

4.1 Interpretazioni

I modelli suggeriscono che molte variabili hanno un effetto statisticamente significativo sulle *odds* di successo, sul raggiungimento del *late stage funding* (*IPO, M&A*).



UNIVERSITÀ DI PAVIA

4.1.1 Variabili di controllo

Partendo dalle variabili di controllo, l'età dell'impresa mostra un *odds ratio* compreso tra 1,28 e 1,29 nei diversi modelli stimati. L'analisi descrittiva del campione indica che le *startup* di successo presentano un'età media superiore di circa 2 anni (1,94) rispetto alle controparti non riuscite. Questo risultato suggerisce che le imprese più "mature" avrebbero una probabilità maggiore di raggiungere fasi avanzate di finanziamento. Ciò è coerente con la definizione di successo adottata nello studio: il raggiungimento di un *late stage funding* richiede generalmente un periodo più lungo per consolidare il modello di *business* e ottenere l'accesso a finanziamenti significativi. Gli *odds ratio* stimati mediante la regressione logistica indicano che ogni anno aggiuntivo di attività aumenta le probabilità di raggiungere l'*exit* del 28–29%, a parità di tutte le altre variabili. L'effetto è stabile e significativo in tutti i modelli considerati. La relazione tra la maturità dell'impresa e la probabilità di successo nel processo di finanziamento trova conferma nella letteratura sul ciclo di vita delle *startup* e sul ruolo della maturità aziendale nella raccolta di capitale. In particolare, Potanin e colleghi (2023) evidenziano che le *startup* più mature hanno maggiori probabilità di attrarre investimenti *late stage* poiché dispongono di un *track record* consolidato e di *performance* storiche che riducono il rischio percepito dagli investitori. Analogamente, la teoria del *venture capital stage model* (Gompers & Lerner, 2001) sottolinea come i finanziatori privilegino le imprese



UNIVERSITÀ DI PAVIA

con un adeguato livello di consolidamento operativo e di *performance* verificabile. L'analisi di regressione logistica condotta da Colombo e Grilli (2010) presenta risultati analoghi: ogni anno aggiuntivo di attività incrementa le *odds* di raggiungere fasi avanzate di finanziamento di circa il 25–30%, confermando la coerenza con le evidenze empiriche esistenti.

La presenza femminile è limitata nel *dataset* analizzato: solo il 24% delle *startup* include almeno una donna nel *team*. La variabile in questione mostra un effetto negativo e statisticamente significativo sulle probabilità di successo in tutte le specificazioni del modello, come riportato nella Tabella 6. Gli *odds ratio* variano tra 0,587 e 0,600 e, nei team con presenza femminile, sono inferiori rispetto ai *team* esclusivamente maschili di circa il 40%, a parità di condizioni. La stabilità di questi valori persiste anche dopo l'inclusione delle variabili relative al capitale umano nei modelli M1–M4. Il confronto tra *startup* di successo e insuccesso, tuttavia, mostra differenze rilevanti: nelle imprese che non raggiungono l'obiettivo la presenza femminile media è pari a 0,26, mentre nelle *startup* di successo scende a 0,19. Anche la distribuzione mantiene caratteristiche simili nei due gruppi.

L'interpretazione di questi risultati richiede cautela e va letta alla luce delle dinamiche strutturali dell'ecosistema delle *startup*. Montanaro e colleghi (2022), nel loro “*Determinants of the exit value in European venture capital-*



UNIVERSITÀ DI PAVIA

backed technology startup”, evidenziano, attraverso modelli di regressione lineare, che la presenza femminile nel loro campione ha un coefficiente positivo sulle probabilità di exit, ma non è statisticamente significativo. Ciò è in contrasto con quanto verificabile nell’analisi condotta per questa tesi.

Alcuni studi evidenziano l’esistenza di bias sistemici nei processi di finanziamento delle *startup*, che tendono a penalizzare i *team* con una presenza femminile. Ad esempio, Brush e colleghi (2004) analizzano le barriere all’accesso per *startup* fondate da donne nell’ambito delle *venture capital*. Da un loro studio emerge un significativo divario: solo il 4-9% dei finanziamenti è destinato alle imprese fondate da donne. Gli autori attribuiscono tale valore alla struttura delle reti di *venture capital*, storicamente caratterizzate da relazioni professionali prevalentemente maschili. Un minor finanziamento forzerebbe le imprese con componenti femminili a ricorrere a fonti alternative più costose o a rallentare la crescita. Ne deriverebbero maggiori difficoltà nelle fasi iniziali, minore visibilità e minore opportunità di ricchezza da *exit* (es. *IPO*). “*Non avendo contatti con gli investitori di capitale di rischio, le imprenditrici hanno meno possibilità di sedersi al tavolo delle trattative. Le donne ritengono che le loro reti di contatti comprendano sia uomini che donne, mentre gli uomini dichiarano di avere nella propria rete di contatti solo uomini (e non donne).*” (Brush et al., 2004, p. 4). Contrariamente, uno studio più recente di Gompers e



UNIVERSITÀ DI PAVIA

Kovvali (2018) integra un'analisi di *team* con composizione eterogenea di genere; gli autori osservano che, nelle fasi post-finanziamento, i *team* con presenza femminile nel loro campione presentano un tasso di aumento delle *exit* positive relativo di ~10,7%. Le cause da loro attribuite sarebbero la maggiore creatività e strategia.

Alla luce dei tre studi osservati, il risultato ottenuto nei modelli della presente analisi non dovrebbe essere interpretato come un errore o come una minore capacità imprenditoriale dei *team* con presenza femminile, bensì come un possibile riflesso delle barriere strutturali e dei bias presenti nei mercati del capitale di rischio. In altre parole, l'effetto negativo osservato potrebbe riflettere principalmente differenze nelle opportunità di accesso alle risorse finanziarie, per cui le probabilità di successo, come definite, sarebbero inferiori rispetto a quelle di *team* omogenei maschili.

Passando invece alla variabile dimensione del *team*, essa esercita un effetto positivo e statisticamente significativo sulle probabilità di successo delle *startup* in tutte le specificazioni dei modelli stimati. Ogni incremento unitario nella dimensione del *team* aumenta gli *odds* di successo compresi tra il 16% e il 22,5%, a seconda della specificazione del modello (*odds ratio* minimo pari a 1,163 nel modello M3 e massimo pari a 1,225 nel modello M2), come riportato in Tabella 6. L'analisi comparativa tra *startup* di successo e insuccesso evidenzia



UNIVERSITÀ DI PAVIA

che le imprese di successo presentano, in media, 0,24 membri in più rispetto a quelle non riuscite, confermando l'associazione positiva tra il numero di membri del team e la probabilità di raggiungere fasi avanzate di finanziamento.

L'effetto positivo della dimensione del *team* potrebbe essere interpretato alla luce dei vantaggi derivanti dalla combinazione di competenze complementari. Un precedente studio (Jin et al.,2017) approfondisce questo aspetto, evidenziando che la relazione tra dimensione del *team* e *performance* non è lineare: i *team* più piccoli (≤ 3 membri) eccellerebbero in termini di integrazione comportamentale e comunicazione informale, mentre i *team* più grandi (≥ 6 membri) avrebbero migliori *performance* grazie alla maggiore capacità di gestione della complessità. Secondo lo studio, l'*odds ratio* medio per ciascun membro aggiuntivo è pari a circa 1,107, corrispondente a un incremento marginale della probabilità di successo del 10,7%, che può variare da un massimo del 23,4% per *team* ad alta *performance* a un minimo del 2% per *team* a bassa *performance*, con valori medi del 16,2% per *team* piccoli (≤ 3 membri). Gli *odds ratio* stimati risultano coerenti con la letteratura del 2017, a supporto dell'idea che una maggiore dimensione del *team* rappresenti un fattore rilevante, sebbene l'entità dell'effetto possa dipendere dal contesto.

L'ultima variabile di controllo considerata è la presenza, nel team imprenditoriale, di almeno un fondatore residente in Nord America. Dall'analisi



UNIVERSITÀ DI PAVIA

descrittiva del campione emerge che il 64,89% delle *startup* analizzate include almeno un fondatore residente negli Stati Uniti. Gli *odds ratio* riportati nella Tabella 6, compresi tra 1,536 e 1,599, indicano un effetto positivo e statisticamente significativo in tutti i modelli stimati. La presenza di un fondatore localizzato negli Stati Uniti suggerisce un incremento delle probabilità di successo della *startup*, compreso tra il 53,6% e il 59,9%, mantenendo costanti le altre variabili incluse nel modello.

Il confronto tra imprese che hanno raggiunto un esito positivo e quelle che non lo hanno conseguito mostra una differenza significativa di circa 10 punti percentuali: il 62% delle *startup* appartenenti al gruppo *success = 0* presenta almeno un fondatore residente negli Stati Uniti, mentre tale quota sale al 72% nel gruppo *success = 1*.

Sebbene la letteratura abbia raramente esaminato direttamente la relazione tra la residenza geografica dei fondatori e la probabilità di exit, e risultino limitate le evidenze empiriche specifiche su questo aspetto, il risultato può essere interpretato in riferimento agli studi sugli ecosistemi imprenditoriali. Diversi di essi sottolineano come l'ecosistema statunitense si distingua per una maggiore concentrazione di capitale di rischio e infrastrutture di supporto alle *startup* in alcune aree geografiche, per esempio, uno studio (Florida & Mellander, 2016) evidenzia come “*capital investment remains highly*



UNIVERSITÀ DI PAVIA

concentrated in a small number of global startup hubs”, tra i quali figura in modo prominente la Silicon Valley che accumula fino al 40% di tutte le risorse disponibili per *startup* innovative degli Stati Uniti. La presenza di fondatori residenti negli USA, in aree geografiche di alto interesse, potrebbe agevolare l’accesso ai finanziamenti.

Anche nella letteratura sul *venture capital*, il ruolo della prossimità geografica tra imprenditori e investitori è ritenuto rilevante nei processi di selezione e di monitoraggio. Essa tenderebbe a favorire la costruzione di relazioni fiduciarie e il supporto operativo, aumentando così la probabilità di accesso ai round di finanziamento successivi. La presenza di un fondatore residente negli Stati Uniti può essere interpretata come un fattore che facilita l’integrazione della *startup* nelle reti imprenditoriali e finanziarie particolarmente sviluppate. Sorenson e Stuart (2001) osservano, infatti, che una maggiore distanza riduce le *probabilità di follow-on funding, specie per le startup* giovani. Lo studio è antecedente ai più moderni mezzi di comunicazione e oggi potrebbe non essere così marcato l’effetto della vicinanza fisica a fonti finanziarie sull’accesso ai finanziamenti; tuttavia, anche alcune evidenze empiriche più recenti supportano questa interpretazione. La prova più evidente proviene da uno studio (Guzman, 2018) che ha rilevato che i fondatori emigrati nella Silicon Valley da altri stati hanno registrato un aumento del 5,8% nella



UNIVERSITÀ DI PAVIA

probabilità di uscita (dall'1,4% al 7,2% di riferimento). Ciò contribuirebbe a dimostrare l'effetto sostanziale della posizione geografica per i fondatori con sede negli Stati Uniti. *“I fondatori immigrati negli USA guidano gran parte del successo: il 55% delle startup miliardarie americane ha fondatori stranieri, che spesso risiedono lì per scalare”* (tradotto da Anderson, 2018, p.1⁴).

Le prove suggeriscono che la residenza negli Stati Uniti è un fattore sostanziale, ma le fonti non forniscono confronti quantitativi direttamente correlati alla probabilità di uscita. I risultati del campione analizzato sono coerenti con questo quadro teorico ed empirico. La presenza di un fondatore residente negli Stati Uniti potrebbe rappresentare un vantaggio strategico effettivo grazie alla vicinanza ad ecosistemi imprenditoriali più sviluppati e reti professionali più estese.

4.1.2 Variabili indipendenti

L'analisi dei risultati della regressione logistica evidenzia che il titolo di studio medio del *team* ha un effetto positivo e statisticamente significativo sulle *odds* di successo. In particolare, esso risulta pari a 1,134 nel modello M1 e 1,138 nel modello M4, con significatività al livello dell'1%. Questi valori indicano che un

⁴ <https://www.fosterglobal.com/blog/55-of-americas-billion-dollar-startups-have-an-immigrant-founder/>



UNIVERSITÀ DI PAVIA

incremento unitario del titolo di studio medio nel *team* è associato a un aumento della probabilità di successo compreso tra il 13,4% e il 13,8%, a parità delle altre variabili incluse nel modello.

Il confronto tra *startup* di successo e di non successo ha differenze limitate ma coerenti con i risultati della regressione. In particolare, le *startup* che hanno realizzato un'*exit* positiva (success = 1) presentano un valore medio della variabile pari a 1.62, leggermente superiore a quello osservato per le *startup* senza *exit* (success = 0, media = 1.60). Nel complesso, sebbene le differenze nelle statistiche descrittive tra i due gruppi risultino relativamente contenute, il livello medio dei titoli di studio dei fondatori contribuisce, seppur con un impatto moderato, ad aumentare le probabilità che una *startup* raggiunga esiti positivi nel proprio ciclo di vita.

Le *startup* che raggiungono un'*exit* positiva presentano, in media, una maggiore presenza di fondatori con istruzione universitaria. In particolare, circa il 7,9% dei *team* con fondatori laureati raggiunge tale risultato, rispetto al 4,1% dei *team* senza laurea. Inoltre, circa il 10% delle *startup* con *exit* presenta almeno un fondatore con dottorato, una quota superiore rispetto alla distribuzione osservata nel campione complessivo.

Un confronto utile per comprendere meglio l'impatto sui modelli creati è lo studio di Ratzinger e colleghi (2018), che analizza l'impatto dell'educazione



UNIVERSITÀ DI PAVIA

dei fondatori sulle probabilità di finanziamento e di *exit* nelle *startup* digitali. Gli autori stimano la relazione tra il *background* educativo del *founder team* e i possibili esiti ordinali del ciclo di vita della *startup*, tra cui *il funding e l'exit*. I modelli da loro creati sono *ordered probit* e includono una serie di variabili rilevanti, simili a quelle considerate nel presente elaborato, tra cui l'età della *startup*, il numero di cofondatori, il genere e l'esperienza imprenditoriale.

I risultati da loro ottenuti evidenziano che la presenza di almeno un fondatore con titolo universitario aumenterebbe le probabilità di progressione lungo le diverse fasi di sviluppo del finanziamento della *startup* di circa 4–9% e la probabilità di ottenere *equity funding* di circa 3–7% rispetto ai *team* privi di laureati. Tutti i livelli di istruzione risultano associati a una maggiore probabilità di *funding* ed *exit*, con effetti particolarmente rilevanti per i livelli di *master* e PhD.

Poiché i modelli *ordered probit* non riportano direttamente gli *odds ratios*, non è possibile effettuare un confronto diretto; tuttavia, le statistiche e l'analisi del campione della presente analisi seguono i risultati dello studio di Ratzinger; entrambe le analisi convergono nell'indicare il capitale umano formale dei fondatori come uno dei fattori rilevanti nel determinare le *performance* e le prospettive di *exit* delle *startup* innovative.



UNIVERSITÀ DI PAVIA

Un risultato opposto è ottenuto dalla variabile relativa all'eterogeneità delle esperienze lavorative all'interno dei *team* fondatori. Essa ha un effetto negativo e statisticamente significativo sulla probabilità di successo della *startup* in entrambi i modelli stimati. In particolare, gli *odds ratio* risultano pari a 0.656 nel modello M2 e 0.619 nel modello M4. L'incremento unitario dell'indice è associato a una riduzione delle *odds* di successo compresa tra il 34,4% e il 38,1%, a parità delle altre variabili.

Anche il confronto tra *startup* di successo e insuccesso mostra differenze molto contenute nelle statistiche descrittive: le *startup* senza exit (success = 0) presentano una media pari a 0,75, mentre quelle che hanno raggiunto un'*exit* positiva (success = 1) mostrano una media pari a 0,76. La maggior parte delle *startup* del campione ha livelli elevati di eterogeneità nelle esperienze professionali dei fondatori. Nel complesso, l'analisi evidenzia che livelli più elevati di eterogeneità sono associati a una riduzione significativa delle probabilità di *exit*.

Per spiegare questo fenomeno, è bene osservare altri studi in merito. La meta-analisi di Jin e colleghi (2017) fornisce un utile termine di confronto. La relazione tra l'eterogeneità delle esperienze lavorative dei membri dei *team* imprenditoriali e le *performance* delle *startup* è analizzata attraverso l'aggregazione dei risultati di numerosi studi empirici. Questa varietà di



UNIVERSITÀ DI PAVIA

background lavorativo è stata misurata attraverso alcuni indicatori di dispersione come l'indice di Blau, lo stesso indicatore utilizzato nella presente tesi. I loro risultati mostrano che l'eterogeneità sarebbe positivamente correlata alla *performance*, a differenza di quanto osservato nel campione della tesi in oggetto. Sebbene l'effetto ottenuto da Jin e colleghi sia contenuto, è statisticamente significativo. Per gli autori, i *team* caratterizzati da *background* professionali diversi dispongono di una maggiore capacità di elaborazione delle informazioni e di *problem solving*. Lo studio, però, non considera eventi di *exit*, come *IPO* o *M&A*, ma valuta le *performance* attraverso indicatori di crescita e redditività. I risultati, quindi, non permettono di stimare direttamente l'impatto dell'eterogeneità del *team* sulle probabilità di raggiungere *exit* o finanziamenti e un ulteriore limite riguarda l'assenza di modelli di regressione logistica o di stime espresse in termini di *odds ratio*.

Per ottenere un termine di confronto migliore, è stato preso in considerazione lo studio di Cristofaro e colleghi (2026), che esamina l'impatto della diversità delle esperienze professionali sulla capacità delle startup di raggiungere lo status di Unicorn, utilizzando una versione normalizzata dell'indice di Blau, applicata a diverse dimensioni, attraverso modelli di regressione di Cox, una metodologia più simile a quella sviluppata in questa tesi. Contrariamente alle aspettative teoriche della *Human Capital Theory*, i risultati



UNIVERSITÀ DI PAVIA

ottenuti mostrano che l'eterogeneità non ha un effetto statisticamente significativo sulla probabilità di raggiungere l'obiettivo dello studio; la diversità delle esperienze professionali non accelererebbe sistematicamente il processo di crescita. Gli autori attribuiscono questo risultato alla limitata variabilità nelle esperienze all'interno del campione analizzato.

La divergenza dei risultati può essere interpretata alla luce delle differenze metodologiche tra i due studi e tra la presente tesi. In primo luogo, la metanalisi considera principalmente indicatori di *performance* operativa, mentre il secondo studio e la presente tesi utilizzano come variabile dipendente un indicatore più restrittivo, il "successo imprenditoriale", rappresentato dallo status di Unicorn o dal raggiungimento di un *exit* positiva o di fasi avanzate del *funding stage*. In secondo luogo, la metanalisi aggrega risultati provenienti da studi diversi, mentre l'analisi sulle imprese Unicorn analizza un campione europeo e i dati a disposizione per la presente tesi derivano da un *dataset* specifico di *startup*.

Come spiegare l'effetto negativo sulle *odds ratio* del presente studio, rispetto alla meta-analisi e alla significatività che esso presenta?

Lo studio di Kaiser e Müller (2015) indica che le difficoltà associate all'eterogeneità, come i conflitti affettivi o i costi di coordinamento, possono, in alcuni casi, superare i benefici derivanti dalla maggiore varietà di competenze.



UNIVERSITÀ DI PAVIA

In particolare, per quanto riguarda i *background* professionali, possono sorgere conflitti affettivi tra i membri del *team*, derivanti da divergenze nei valori, nelle esperienze e nei modelli decisionali, che tendono a ridurre la fiducia reciproca e a rallentare i processi. Un ulteriore problema riguarderebbe l'aumento dei costi di coordinamento. Fondatori con esperienze lavorative molto differenti potrebbero avere dei modelli mentali o linguaggi tecnici e delle priorità strategiche diversi. Queste renderebbero più complesse l'integrazione delle competenze e la definizione di una strategia condivisa. Nel loro studio, questo aspetto risulta particolarmente rilevante nel contesto delle *startup*, caratterizzato da elevata incertezza e necessità di prendere decisioni rapide. Anche piccoli disallineamenti potrebbero ridurre la capacità dell'impresa di adattarsi alle opportunità di mercato o alle richieste degli investitori.

Mentre parte della letteratura indica che l'eterogeneità dei *team* possa favorire la *performance* delle nuove imprese attraverso una maggiore varietà di competenze e prospettive, i risultati empirici ottenuti nel contesto specifico analizzato, mostrano che una maggiore diversità nelle esperienze lavorative dei fondatori potrebbe invece essere associata a una minore probabilità di raggiungere una *exit* positiva, probabilmente a causa dei maggiori costi di coordinamento e delle difficoltà decisionali che tali differenze possono generare nelle fasi critiche dello sviluppo della *startup*.



UNIVERSITÀ DI PAVIA

Infine, l'osservazione della variabile relativa all'intensità di esperienze imprenditoriali accumulate dai membri dei *team* fondatori mostra un effetto positivo e statisticamente significativo sulla probabilità di successo dell'impresa in entrambi i modelli stimati. In particolare, gli *odds ratio* risultano pari a 1,056 nel modello M3 e 1.059 nel modello M4. L'incremento unitario della variabile è associato a un aumento delle *odds* di successo compreso tra il 5,6% e il 5,9%, a parità delle altre variabili incluse nei modelli.

È importante ricordare che la maggior parte dei *team* fondatori possiede un numero molto limitato di esperienze pregresse, mentre una piccola quota presenta numeri molto elevati.

Il confronto tra *startup* di successo e insuccesso conferma il ruolo positivo dell'esperienza imprenditoriale: le *startup* per cui la variabile dipendente è positiva presentano valori medi più elevati di intensità dell'esperienza rispetto a quelle che non hanno ottenuto un *exit* o un *funding stage avanzato*. Nel gruppo delle *startup* non di successo (*success* = 0), la media è pari a 0,62, con una deviazione standard di 0,49, mentre tra le *startup* di successo (*success* = 1) la media sale a 0,72, con una deviazione standard di 0,45.

Sebbene la differenza non sia particolarmente marcata, questi risultati sono coerenti con la letteratura sulla *Human Capital Theory*, secondo cui l'esperienza accumulata dai fondatori costituisce una componente fondamentale



UNIVERSITÀ DI PAVIA

del capitale umano. Com'è evidente nella letteratura economica e *manageriale*, inoltre, contribuisce alla *performance* delle imprese e alla creazione di valore (Schultz, 1960; Becker, 1975; Unger et al., 2011).

Nel contesto delle *startup*, l'esperienza imprenditoriale accumulata può influenzare positivamente le probabilità di successo attraverso diversi meccanismi. Le esperienze precedenti consentono ai fondatori di sviluppare conoscenze e capacità decisionali che migliorano la gestione delle fasi iniziali dell'impresa e contribuiscono a ridurre gli errori strategici grazie a processi di *learning-by-doing*, consentendo di identificare più efficacemente le opportunità di mercato. Inoltre, la presenza attiva nel mercato, come imprenditore fondatore, permetterebbe di accedere a reti di contatti e a una maggiore credibilità nei confronti di investitori e partner, facilitando l'accesso a risorse finanziarie e a reti professionali (Colombo & Grilli, 2005; Hsu, 2007; Gompers et al., 2010). A tal proposito, il risultato positivo osservato suggerisce che l'intensità è un fattore rilevante nel favorire esiti positivi.

4.2 Effetto del capitale umano dei fondatori sul successo della *startup*

Emerge con chiarezza il ruolo che il capitale umano ha in un effetto significativo sulla probabilità di successo delle *startup*. Come discusso nel Capitolo 1, la letteratura ha ampiamente evidenziato le caratteristiche individuali dei fondatori, tra le quali l'istruzione, l'esperienza imprenditoriale e il *background*



UNIVERSITÀ DI PAVIA

professionale. Esse sono fattori chiave per la capacità delle nuove imprese di svilupparsi e raggiungere risultati positivi nel proprio ciclo di vita (Becker, 1975; Schultz, 1960; Unger et al., 2011).

L'interpretazione dei risultati, nel contesto del dataset utilizzato, va intesa tenendo in considerazione la variabile dipendente: il raggiungimento di eventi di successo, definiti come exit positive (IPO o M&A), o l'accesso a stadi avanzati di finanziamento (*late stage funding*).

In primo luogo, l'effetto positivo e statisticamente significativo del livello di istruzione medio dei *team* associa ad un incremento unitario del titolo di studio un aumento delle *odds* di successo tra il 13,4% e il 13,8%. Questo risultato appare coerente con le previsioni formulate e con l'ipotesi H1: un livello più elevato di istruzione dei fondatori è positivamente associato al successo della *startup*. Dal punto di vista teorico, nella *Human Capital Theory* l'istruzione rappresenta un investimento che accresce le capacità produttive degli individui attraverso l'acquisizione di conoscenze, competenze analitiche e capacità decisionali (Becker, 1975). In contesti come quelli delle *startup*, tali competenze possono contribuire alla qualità delle decisioni strategiche e alla capacità di interpretare le opportunità. I risultati empirici sono coerenti con le evidenze della letteratura, come discusso in precedenza. Complessivamente, il capitale umano formale svolge un ruolo rilevante nel migliorare la capacità delle



UNIVERSITÀ DI PAVIA

startup di svilupparsi lungo il proprio ciclo di vita, confermando la validità empirica dell'Ipotesi H1.

La seconda dimensione riguarda l'intensità dell'esperienza imprenditoriale accumulata all'interno del *team* fondatore. La variabile che misura il numero complessivo di esperienze di fondazione maturate dai membri del *team*. L'incremento unitario è associato a un aumento delle *odds* di successo compreso tra il 5,6% e il 5,9%. Sebbene l'entità dell'effetto sia più contenuta rispetto a quella osservata per il livello di istruzione, il risultato è comunque marcato come componente specifica del capitale umano dei fondatori. Questa dimensione può essere interpretata alla luce del concetto di *entrepreneurial Human Capital*, discusso nel Capitolo 1, che identifica l'insieme delle competenze sviluppate attraverso l'esperienza diretta nella creazione e gestione di nuove imprese (Unger et al., 2011). Le esperienze precedenti consentono ai fondatori di sviluppare competenze e conoscenze che riducono gli errori nelle fasi iniziali. Questo processo è spesso descritto in letteratura come un meccanismo di *learning-by-doing*. Inoltre, l'esperienza imprenditoriale favorisce lo sviluppo di reti relazionali e la reputazione, elementi che possono facilitare l'accesso a risorse finanziarie e partnership (Colombo & Grilli, 2005; Gompers et al., 2010; Hsu, 2007). Anche questi risultati empirici sono coerenti con la letteratura esistente e forniscono supporto empirico all'Ipotesi H2: una



UNIVERSITÀ DI PAVIA

maggior intensità dell'esperienza imprenditoriale nel *team* fondatore aumenta la probabilità di successo della *startup*.

Diversamente, per quanto riguarda l'eterogeneità delle esperienze lavorative dei fondatori, sono stati ottenuti dei risultati inaspettati rispetto alle aspettative teoriche. La variabile misura il grado di diversità nei *background* professionali del *team* attraverso l'indice di Blau (Blau, 1977). L'effetto è negativo e statisticamente significativo. Un incremento unitario dell'eterogeneità è associato a una riduzione delle *odds* di successo compresa tra il 34,4% e il 38,1%. Questo risultato appare in forte contrasto con parte della letteratura esaminata (Jin et al., 2017; Unger et al., 2011). Tuttavia, come evidenziato anche da alcuni contributi del paragrafo precedente, l'eterogeneità può generare costi organizzativi e di coordinamento particolarmente rilevanti nelle fasi iniziali dello sviluppo delle *startup* (Kaiser & Müller, 2015). In contesti caratterizzati da elevata incertezza, da risorse limitate e dalla necessità di prendere decisioni rapidamente, le differenze marcate tra i fondatori possono generare conflitti decisionali e difficoltà di coordinamento. Questi fattori rallentano e riducono la capacità del *team* di reagire tempestivamente alle opportunità di mercato. Le differenze tra i risultati ottenuti e le evidenze più diffuse nella letteratura possono essere interpretate alla luce delle misure di *performance* utilizzate. Alcuni studi utilizzano indicatori di *performance*



UNIVERSITÀ DI PAVIA

operativa come la crescita del fatturato o la redditività. Nel presente studio, invece, il successo è definito tramite eventi selettivi che rappresentano traguardi da raggiungere nel ciclo di vita. In questo contesto, è possibile che i benefici derivanti dalla diversità delle competenze siano compensati o superati dai costi di coordinamento interni. Alla luce di tali evidenze, l'Ipotesi H3, che prevedeva un effetto positivo dell'eterogeneità delle esperienze lavorative sulla probabilità di successo della *startup*, non trova supporto empirico nel campione analizzato. Livelli più elevati di eterogeneità professionale nel *background* dei fondatori sono associati a una minore probabilità di raggiungere eventi di successo.

Considerando congiuntamente i coefficienti stimati dai modelli logistici, emerge che le diverse dimensioni contribuiscano in modo differenziato al possibile successo della *startup*. L'effetto del livello di istruzione è più marcato rispetto a quello dell'esperienza imprenditoriale accumulata dai fondatori, e l'eterogeneità dei *background* professionali presenta l'effetto più incisivo, ma negativo. Le dimensioni osservate contemporaneamente possono rafforzarsi a vicenda. In termini interpretativi, un *team* fondatore caratterizzato sia da un livello di istruzione più elevato della media sia da una maggiore esperienza imprenditoriale aggregata presenta *odds* di successo complessivamente più elevati rispetto a un *team* che possiede soltanto una di queste caratteristiche. Dai valori stimati nei modelli, l'effetto combinato di un



UNIVERSITÀ DI PAVIA

incremento unitario di entrambe le variabili corrisponde a un aumento complessivo delle *odds* di successo di circa il 19–20%, risultante dalla moltiplicazione degli *odds ratio* associati alle due variabili. Da quanto emerge, il capitale umano non opera esclusivamente attraverso singole dimensioni, bensì attraverso una combinazione di competenze formali e imprenditoriali. La presenza di fondatori con elevato livello di istruzione può favorire le capacità analitiche e le competenze tecniche, mentre l'esperienza imprenditoriale contribuisce allo sviluppo di competenze pratiche legate alla gestione e alle relazioni con gli *stakeholder* (Colombo & Grilli, 2005; Unger et al., 2011). Contemporaneamente, l'effetto negativo associato all'eterogeneità delle esperienze lavorative dei fondatori potrebbe attenuare i benefici derivanti dalle altre dimensioni del capitale umano. La qualità del capitale umano formale e dell'esperienza imprenditoriale potrebbe essere particolarmente rilevante se inserita in *team* relativamente omogenei dal punto di vista professionale, in combinazione con una minore frammentazione dei *background*. L'efficacia del *team* potrebbe dipendere anche dalla sua configurazione oltre che dall'equilibrio tra competenze complementari e capacità di coordinamento.

Per comprendere meglio l'impatto della composizione del *team*, è possibile confrontare le diverse configurazioni delle variabili del modello completo M4. Poiché la regressione logistica stima gli effetti in termini di *odds*



UNIVERSITÀ
DI PAVIA

ratio, l'impatto combinato è interpretato moltiplicando gli effetti associati ai singoli coefficienti.

Poiché gli effetti nel modello *logit* sono moltiplicativi, l'impatto complessivo della configurazione delle variabili osservate può essere calcolato moltiplicando le tre *odds ratio*:

$$OR_{totale} = 1,138 \times 0,619 \times 1,059$$

Il risultato è pari a:

$$OR_{totale} \approx 0,746$$

Considerando simultaneamente tutte le caratteristiche descritte e l'aumento di un'unità di tutte le variabili contemporaneamente, le *odds* di successo risultano pari a circa 0,75 rispetto allo scenario di riferimento. Gli effetti della composizione del *team*, quindi, derivano dall'interazione tra i fattori che esercitano influenze di segno opposto.

Gli effetti della composizione del *team*, dunque, non possono essere letti come il semplice risultato della presenza isolata di singole caratteristiche, ma vanno interpretati come l'esito della loro combinazione complessiva.

I risultati della presente analisi mostrano infatti che alcune dimensioni del capitale umano esercitano un effetto positivo sulla probabilità di successo



UNIVERSITÀ DI PAVIA

della *startup*, mentre altre, come l'eterogeneità delle esperienze, risultano associate a un effetto negativo. La performance della *startup* non dipende esclusivamente dal livello medio di competenze o di esperienza presente nel *team*, bensì anche dal modo in cui tali risorse risultano distribuite, integrate e rese complementari all'interno del gruppo. In questa prospettiva, la composizione rappresenta una dimensione rilevante, poiché il valore del capitale umano non emerge soltanto dalla qualità delle risorse individuali, ma anche dalla coerenza e dall'equilibrio della loro configurazione complessiva.

4.3 Effetto della composizione del *team*

Come visto, la configurazione complessiva del *team* imprenditoriale è centrale, poiché determina come le competenze e le esperienze siano combinate all'interno della *startup*. Nei modelli stimati nel Capitolo 3 e discussi nei paragrafi precedenti emergono quattro caratteristiche principali intrinseche alla composizione dei *team* fondatori, considerabili significativamente associate alla probabilità di successo della startup: la dimensione del team, il grado di eterogeneità delle esperienze lavorative, la presenza di almeno una donna nel *team* e la presenza di almeno un fondatore residente in Nord America.

L'interpretazione congiunta di questi risultati permette di contestualizzare empiricamente anche le ipotesi di ricerca H1–H3. La



UNIVERSITÀ DI PAVIA

grandezza dei *team* è una delle variabili di controllo più rilevanti tra quelle incluse nei modelli. In tutte le specificazioni della regressione logistica presenta un effetto positivo e significativo; in particolare, nel modello completo (M4) è associata ad un aumento delle probabilità di evento positivo di circa il 18%. Questo risultato è coerente con quanto raccontato nello Stato dell'arte: *team* più ampi possono disporre di una maggiore dotazione complessiva di capitale umano grazie alla combinazione dei singoli. Ogni fondatore contribuirebbe con la propria rete professionale e personale, favorendo le opportunità di accesso a risorse esterne. Le statistiche descrittive del campione mostrano, tuttavia, che i *team* analizzati sono generalmente di dimensioni ridotte: la media è pari a 2,41 membri, con una mediana di 2 fondatori. È possibile osservare che l'effetto positivo non deriva da *team* particolarmente numerosi, bensì dal passaggio da configurazioni di singolo a *team* composti da due o tre membri.

Come ribadito nei paragrafi precedenti, l'eterogeneità delle esperienze professionali ha un effetto negativo e statisticamente significativo sulla probabilità di successo. Tra il modello M2 ed il modello completo M4 osserviamo una diminuzione leggera degli *odds*, a cui l'aumento unitario dell'indice è associata una riduzione delle probabilità compresa tra il 34,4% e il 38,1%. Un'elevata diversità ridurrebbe la probabilità di raggiungere eventi



UNIVERSITÀ DI PAVIA

di *exit* o fasi avanzate di finanziamento. Dal punto di vista teorico, questo risultato si inserisce nel dibattito discusso in precedenza riguardo agli effetti ambivalenti dell'eterogeneità nei *team* come evidenziato anche da (Kaiser & Müller, 2015). I livelli elevati di eterogeneità possono comportare costi di coordinamento e conflitti interpersonali, riducendo l'efficienza decisionale. In questo contesto, *team* con *background* professionali molto distanti potrebbero incontrare maggiori difficoltà; pertanto, i risultati suggeriscono che l'efficacia del capitale umano dipende anche dalla capacità del *team* di integrarlo. Nel campione della presente tesi, i livelli più moderati di eterogeneità risultano più funzionali rispetto a configurazioni estremamente frammentate.

Tra gli effetti statisticamente negativi osservati figura anche la presenza di almeno una fondatrice nel *team*. In particolare, nel modello completo M4 l'*odds ratio* è pari a 0,594. L'evidenza empirica non implica che le imprenditrici abbiano *performance* inferiori in termini di capacità o qualità. Piuttosto, come evidenziato nella letteratura citata, i valori ottenuti possono riflettere fattori strutturali del contesto imprenditoriale. La disparità nell'accesso al capitale finanziario, alle reti professionali e alle opportunità di crescita. Lo studio di Gottschalk e Niefert (2013), già osservato in precedenza, evidenzia come le imprese fondate da sole donne tendano a presentare dimensioni inferiori e livelli di capitalizzazione più contenuti, pur mostrando



UNIVERSITÀ DI PAVIA

livelli medi più elevati di istruzione formale. Esistono barriere sistemiche nei mercati e nelle reti di conoscenza, oltre a differenze intrinseche nelle capacità. Nel campione analizzato, la presenza femminile è relativamente limitata al 24% delle *startup* osservate. I risultati possono essere influenzati dalla distribuzione non equilibrata dei *team* nel *dataset* e dalla loro rappresentanza limitata. L'effetto osservato può essere interpretato come un indicatore delle condizioni strutturali dell'ecosistema, anziché come misura del capitale umano.

Infine, la presenza in un *team* di almeno un fondatore residente in Nord America sembra essere un fattore positivamente associato al successo. Nel modello completo M4, l'incremento unitario della stessa è associato ad un aumento delle *odds* di successo di circa il 57%. Com'è stato evidenziato in precedenza, gli ecosistemi imprenditoriali nordamericani svolgono nel panorama globale un ruolo di primo piano grazie a una combinazione favorevole di fattori istituzionali, finanziari e culturali. La letteratura evidenzia che una quota significativa delle *startup* più successe negli Stati Uniti è guidata da fondatori internazionali che operano all'interno di tali ecosistemi. Nella presente tesi, la presenza positiva di un fondatore residente in Nord America può essere interpretata come un indicatore di accesso privilegiato a sistemi



UNIVERSITÀ DI PAVIA

altamente sviluppati, con reti di investitori più ampie, maggiore disponibilità di capitale di rischio e una cultura imprenditoriale consolidata.

Team leggermente più ampi, inseriti in ecosistemi imprenditoriali più favorevoli, mostrano probabilità di successo significativamente più elevate. Al contrario, livelli molto elevati di eterogeneità professionale possono generare costi di coordinamento che riducono l'efficacia del capitale umano disponibile.

Queste evidenze contribuiscono a integrare le ipotesi formulate nel Capitolo 1 mostrando come gli effetti del capitale umano (H1 e H2) siano mediati dalla struttura del *team* imprenditoriale e dal contesto istituzionale in cui opera. Sotto questo punto di vista, il capitale umano non rappresenta una risorsa puramente individuale, bensì una risorsa collettiva la cui efficacia dipende dall'organizzazione e dall'integrazione delle competenze all'interno dei *team*.

4.4 Implicazioni

L'analisi empirica ha evidenziato che le diverse dimensioni del capitale umano e la composizione del *team* fondatore sono legate al successo delle *startup* del campione. È necessario discutere le principali implicazioni della ricerca. In questa sezione sono analizzate le implicazioni teoriche, che riguardano il contributo dello studio alla letteratura sul capitale umano, e le implicazioni manageriali, che riguardano le possibili applicazioni pratiche dei risultati per



UNIVERSITÀ DI PAVIA

imprenditori, investitori e organizzazioni di supporto all'ecosistema delle *startup*.

4.4.1 Implicazioni teoriche

Le evidenze empiriche supportano quanto previsto dalla *Human Capital Theory*: gli investimenti in istruzione e competenze contribuiscono ad aumentare la produttività e la capacità decisionale degli individui (Becker, 1975; Schultz, 1960). Il capitale umano è una risorsa strategica nelle fasi iniziali del ciclo di vita imprenditoriale; un livello di istruzione più elevato all'interno dei *founder teams* aumenta le probabilità che la startup raggiunga una fase avanzata di finanziamento o realizzi un'*exit* positiva. Le competenze cognitive e manageriali incidono direttamente sulla qualità delle decisioni strategiche, sulla capacità di attrarre risorse finanziarie e sulla gestione delle fasi di sviluppo.

In secondo luogo, i risultati contribuiscono alla letteratura sugli *entrepreneurial teams*, evidenziando come il capitale umano debba essere analizzato come dotazione aggregata del gruppo fondatore, non solo come elemento caratterizzante del singolo. L'aggregazione di competenze e reti relazionali aumenta la probabilità di accesso alle risorse e facilita la divisione del lavoro. Questa evidenza contribuisce agli studi che interpretano il capitale umano come una risorsa collettiva (Beckman & Burton, 2008; Jin et al., 2017).



UNIVERSITÀ DI PAVIA

Infine, l'incremento dell'eterogeneità delle esperienze lavorative all'interno del *team* fondatore è associato a una riduzione delle probabilità di successo della *startup*, contrariamente alle ipotesi formulate sulla base della letteratura. Sebbene la diversità possa favorire l'innovazione e la creatività, può anche generare costi di coordinamento e conflitti decisionali che incidono negativamente sulle *performance*. In questo senso, i risultati contribuiscono al dibattito teorico evidenziando come *background* lavorativi diversi possano avere effetti non lineari e dipendere dal contesto organizzativo e dalle capacità di coordinamento del gruppo.

4.4.2 Implicazioni *manageriali*

I risultati della tesi presentano alcune implicazioni potenzialmente interessanti anche per i soggetti coinvolti nei processi di creazione e sviluppo delle *startup*, quali fondatori, investitori e organizzazioni di supporto all'imprenditorialità.

Il titolo di studio dei fondatori rappresenta un fattore importante insieme alla formazione accademica e all'acquisizione di competenze analitiche e manageriali. Questi elementi potrebbero favorire la capacità degli imprenditori di interpretare le opportunità di mercato o di sviluppare modelli di *business* e di attrarre investitori. Di conseguenza, gli ambienti universitari e accademici svolgono un ruolo significativo nel capitale umano dei futuri imprenditori e nelle prospettive di crescita economica.



UNIVERSITÀ DI PAVIA

Inoltre, la struttura e composizione dei *team* in ottica di divisione del lavoro, o di integrazione delle competenze sembra rimanere un elemento centrale, così come l'ampliamento delle reti relazionali è utile per l'accesso a risorse finanziarie e opportunità di mercato. Un *team* con competenze diversificate ma complementari rappresenta una scelta strategica proprio per quanto indicato. È necessario gestire con attenzione il grado di eterogeneità all'interno del *team*, in particolare per quanto riguarda la storia e le esperienze dei membri. Sebbene la diversità possa favorire l'innovazione, come visto in precedenza, i livelli eccessivi generano difficoltà di coordinamento e conflitti. Pertanto, a fianco della diversità, sarebbero necessari meccanismi organizzativi e di governance in grado di facilitare la collaborazione e l'allineamento degli obiettivi.

Fondamentalmente, il successo delle *startup* non dipende esclusivamente dalla disponibilità di risorse finanziarie o dalle condizioni di mercato, ma anche dalla qualità e dalla configurazione del capitale umano, che rappresentano tra i principali driver delle traiettorie di sviluppo.

4.5 Limiti della ricerca e spunti per ricerche future

Ogni studio empirico è inevitabilmente soggetto a limitazioni da considerare.

Anche la presente tesi ha alcuni vincoli metodologici ed empirici.



UNIVERSITÀ DI PAVIA

Il primo limite riguarda la natura cross-sectional del *dataset* utilizzato. L'analisi è stata condotta sulla base di dati relativi a un singolo periodo temporale. In assenza di informazioni longitudinali sui percorsi evolutivi delle imprese del campione, non è possibile osservare dinamicamente come evolvano le relazioni nel tempo. Questo approccio consente di identificare associazioni statistiche tra le caratteristiche del campione dei fondatori e delle imprese e la probabilità di successo delle *startup*, ma non consente di analizzare l'evoluzione dei *team*, i cambiamenti nelle competenze o l'impatto delle decisioni adottate nelle diverse fasi del ciclo di vita.

Il secondo limite è relativo alla misurazione del successo che è stato definito come il raggiungimento di eventi osservabili quali operazioni di *IPO*, acquisizioni (*M&A*) o l'accesso a fasi avanzate di finanziamento. Sebbene questa definizione sia coerente con una parte rilevante della letteratura e consenta di utilizzare indicatori oggettivi e verificabili, rappresenta solo una delle possibili dimensioni delle *performance* analizzabili. Altre dimensioni, come ad esempio la crescita dei ricavi, la redditività, la sopravvivenza nel lungo periodo o l'impatto innovativo, non sono state considerate e potrebbero portare a valutazioni differenti del successo delle *startup*. Inoltre, la misurazione del capitale umano dei fondatori si concentra su poche dimensioni osservabili: il livello medio di istruzione del *team*, l'esperienza imprenditoriale accumulata e



UNIVERSITÀ DI PAVIA

l'eterogeneità delle esperienze lavorative. Il concetto di capitale umano è multidimensionale e comprende anche competenze non osservabili come le *soft skills*, le capacità di leadership, le reti relazionali e le caratteristiche caratteriali. Il modello empirico cattura solo una parte delle componenti che possono influenzare il successo delle *startup*.

Un ulteriore elemento di cautela riguarda il *dataset* utilizzato per l'analisi. È stato costruito a partire da dati secondari di Crunchbase. Nonostante la sua ampia copertura e diffusione, questo tipo di *database* può presentare alcune limitazioni legate alla completezza delle informazioni disponibili, poiché mancavano dati per molte *startup*, probabilmente quelle che non hanno ottenuto visibilità nel mercato. Ciò potrebbe aver introdotto distorsioni nella composizione del campione.

Infine, il modello si basa su un numero limitato di variabili esplicative e su una scelta metodologica specifica, ossia l'uso della regressione logistica binaria. Sebbene questo approccio sia coerente con la natura dicotomica della variabile dipendente, altri fattori rilevanti per il successo, come il settore di mercato, il contesto istituzionale, le caratteristiche degli investitori o la qualità delle reti imprenditoriali, non sono stati inclusi nell'analisi e potrebbero influenzare le relazioni osservate.



UNIVERSITÀ DI PAVIA

Questi limiti non annullano i risultati ottenuti, ma offrono opportunità per sviluppi futuri. Un primo studio potrebbe utilizzare dei *dataset* longitudinali per osservare l'evoluzione delle *startup* nel tempo e comprendere meglio la dinamica temporale del capitale umano lungo le diverse fasi del ciclo di vita. In secondo luogo, si potrebbe ampliare la misurazione includendo ulteriori variabili relative alle competenze non cognitive, alle caratteristiche di personalità e alla qualità delle reti relazionali imprenditoriali. Anche alcune variabili contestuali, come il settore di attività, l'ecosistema o la presenza di investitori istituzionali, consentirebbero di comprendere come il capitale umano interagisce con il contesto in cui operano le *startup*. Infine, sarebbe utile affiancare alla regressione logistica delle metodologie differenti, come i modelli di sopravvivenza o gli approcci di inferenza causale, per approfondire i meccanismi attraverso cui le caratteristiche dei fondatori influenzano le traiettorie di crescita e di *exit* delle *startup*.

Capitolo 5: Conclusioni

Questa tesi si è occupata di analizzare il capitale umano dei fondatori e la composizione dei *team* imprenditoriali come determinanti del successo delle *startup*. In un contesto caratterizzato da forte incertezza, capire quali elementi incidono maggiormente sulle strade che favoriscono la crescita è un tema rilevante sia per la ricerca sia per gli attori dell'ecosistema imprenditoriale. La



UNIVERSITÀ DI PAVIA

domanda di ricerca che ha guidato la tesi è: “*In che misura il livello di istruzione, l’esperienza imprenditoriale complessivamente accumulata e l’eterogeneità settoriale del team fondatore influenzano la probabilità che una startup raggiunga un’exit positiva?*”. I risultati dell’analisi mostrano che alcune dimensioni osservabili del capitale umano risultano positivamente associate alla probabilità di successo, in particolare il livello di istruzione e l’esperienza imprenditoriale, mentre l’eterogeneità delle esperienze professionali presenta un effetto negativo nel campione analizzato.

Nel dettaglio, l’evidenza empirica correlata al livello accademico è coerente con i fondamenti della *Human Capital Theory*, secondo cui gli investimenti in istruzione accrescono le competenze e le capacità decisionali degli individui, generando ritorni economici e migliorando la qualità delle scelte in contesti complessi (Becker, 1975; Schultz, 1960). Analogamente, l’esperienza imprenditoriale accumulata dai *team* si conferma un fattore rilevante, in linea con le evidenze secondo cui i fondatori seriali sono in grado di gestire più efficacemente l’incertezza grazie a competenze acquisite e reti relazionali consolidate (Aryadita et al., 2023).

L’eterogeneità delle esperienze professionali passate, invece, contrariamente alle aspettative teoriche iniziali, ha una relazione negativa con la diversità dei *background* e probabilità di successo. Questo suggerisce che,



UNIVERSITÀ DI PAVIA

almeno nelle fasi iniziali del ciclo di vita delle *startup*, i costi di coordinamento, le difficoltà decisionali e i possibili conflitti interni possano prevalere sui benefici derivanti dalla complementarità delle competenze (Kaiser & Müller, 2015).

Globalmente, i risultati ottenuti dimostrano che il capitale umano non agisce come somma di caratteristiche individuali, bensì come una risorsa collettiva, la cui efficacia è legata alla configurazione e all'integrazione delle competenze all'interno dei *team*. Il successo delle *startup* è il risultato di un equilibrio tra le qualità individuali e le dinamiche di gruppo.

Dal punto di vista teorico, la tesi in essere estende la prospettiva della *Human Capital Theory* applicandola al contesto delle *startup*, evidenziando come le competenze e le esperienze dei fondatori influenzino non solo le *performance* individuali, ma anche gli esiti organizzativi. Inoltre, contribuisce alla letteratura sugli *entrepreneurial teams*, fornendo evidenze su come la composizione dei gruppi possa modificare l'efficacia del capitale umano e generare effetti non lineari.

Dal punto di vista manageriale, per gli imprenditori emerge l'importanza di costruire *team* dotati di un adeguato livello di competenze ed esperienze pregresse significative, evitando al contempo un'eccessiva eterogeneità nelle fasi iniziali. Per gli investitori, le caratteristiche dei fondatori



UNIVERSITÀ DI PAVIA

possono essere interpretate come segnali informativi utili alla valutazione del potenziale e per ridurre l'asimmetria informativa. In questo senso, le variabili del livello di istruzione e dell'esperienza imprenditoriale cumulata possono rappresentare indicatori sintetici della qualità del *team*.

Nonostante i contributi forniti, questa tesi presenta alcune limitazioni. In primo luogo, l'utilizzo di dati secondari di natura cross-sectional non consente di osservare l'evoluzione temporale né di cogliere le traiettorie di sviluppo lungo il ciclo di vita. In secondo luogo, la misurazione del capitale umano si basa su *proxy* evidenti. Queste non permettono di catturare le dimensioni più complesse, come le soft *skills*, i tratti di personalità o le reti relazionali, la cui rilevanza è riconosciuta in letteratura, ma è difficilmente misurabile. Infine, la definizione di successo adottata, pur coerente con la letteratura sulle *startup venture-backed*, rappresenta solo una delle possibili modalità di valutazione.

Alla luce di tali limiti, la ricerca futura potrebbe più direzioni. L'utilizzo di *dataset* longitudinali consentirebbe di tenere in considerazione le dinamiche temporali e i meccanismi causali che collegano il capitale umano agli esiti delle *startup*. Secondariamente, si potrebbero integrare misure più articolate per il capitale umano, includendo le dimensioni non cognitive e relazionali citate in precedenza, al fine di cogliere più completamente la complessità. Infine, ulteriori studi potrebbero esplorare il ruolo dei fattori contestuali: il settore di attività o le



UNIVERSITÀ
DI PAVIA

caratteristiche degli ecosistemi imprenditoriali. Il capitale umano interagisce con l'ambiente esterno per determinare il successo.

In conclusione, la tesi sottolinea che il capitale umano dei fondatori è una risorsa fondamentale per lo sviluppo delle *startup*. Esso influenza in modo significativo la loro capacità di affrontare contesti incerti e di ottenere esiti positivi. Comprendere queste dinamiche contribuisce non solo ad arricchire la letteratura accademica, ma anche a fornire strumenti utili per migliorare la comprensione di tali temi.



UNIVERSITÀ
DI PAVIA

Bibliografia

- Acosta, P. M. & others. (2018). The role of cognitive and socio-emotional *skills* in labor markets. *IZA World of Labor*.
- Aldaej, A. (2019). Towards effective technical debt *decision-making* in software *Startups*. *ACM SIGSOFT Software Engineering Notes*, 44(3), 22–22.
- Anderson, S. (2018). *55% of America's billion-dollar startups have an immigrant founder*. Forbes. [https://www.forbes.com/sites/stuartanderson/2018/10/25/55-of](https://www.forbes.com/sites/stuartanderson/2018/10/25/55-of-...)
- Anghel, B., & Balart, P. (2017). Non-cognitive *skills* and individual earnings: New evidence from PIAAC. *SERIEs*, 8(4), 417–473.
- Arena, M., Bengo, I., Calderini, M., & Chiodo, V. (2018). Unlocking finance for social tech start-ups: Is there a new opportunity space? *Technological Forecasting and Social Change*, 127, 154–165.
- Aryadita, H., Sukoco, B. M., & Lyver, M. (2023). Founders and the success of start-ups: An integrative review. *Cogent Business & Management*, 10(3). <https://doi.org/10.1080/23311975.2023.2284451>
- Ayanponle, L. O., Awonuga, K. F., Asuzu, O. F., Daraojimba, R. E., Elufioye, O. A., & Daraojimba, O. D. (2024). A review of innovative HR strategies



UNIVERSITÀ
DI PAVIA

in enhancing workforce efficiency in the US. *International Journal of Science and Research Archive*, 11(1), 817–827.

Baigireyeva, Z. Z., Niyazbekova, S. U., Borisova, E., & Ivanova, O. (2020). The role of *Human Capital* in improving the competitiveness of enterprises. *First International Volga Region Conference on Economics, Humanities and Sports (FICEHS 2019)*, 100–102.

Becker, G. S. (1975). *Human Capital: A theoretical and empirical analysis*. National Bureau of Economic Research.

Beckman, C. M., & Burton, M. D. (2008). Founding the Future: Path Dependence in the Evolution of Top Management Teams from Founding to IPO. *Organization Science*, 19(1), 3–24.
<https://doi.org/10.1287/orsc.1070.0311>

Blair, M. M. (2011). An economic perspective on the notion of *Human Capital*. *The Oxford handbook of Human Capital*, 49–70.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure* (Vol. 7). Free press New York.

Bonelli, M. I. (2023). A Review of the Upper Echelon Theory and Subsequent Refinements. *SSRN Electronic Journal*.
<https://doi.org/10.2139/ssrn.4517218>

Bordiga, A., Bruno, C., & Abrate, G. (2023, gennaio). *Start-up: Caratteristiche, finanziamento e performance*.



UNIVERSITÀ
DI PAVIA

- Brush, C. G., Carter, N. M., Gatewood, E. J., Greene, P. G., & Hart, M. (2004). Gatekeepers of venture growth: A Diana Project report on the role and participation of women in the venture capital industry. *Available at SSRN 1260385*.
- Ceresia, F. (2023). A system dynamics-based interactive learning environment for online formative (self-) assessment of Wanna-Be entrepreneurs' performance management capabilities. *Administrative Sciences, 14*(1), 3.
- Certo, S. T., Holcomb, T. R., & Holmes Jr, R. M. (2009). IPO research in management and entrepreneurship: Moving the agenda forward. *Journal of Management, 35*(6), 1340–1378.
- Česynienė, R., & Stankevičienė, A. (2011). THE ROLE OF HUMAN CAPITAL IN VALUE CREATION: THEORETICAL INSIGHTS. *Ekonomika, 90*(4), 49–62. <https://doi.org/10.15388/ekon.2011.0.925>
- Churchill, N. C., & Lewis, V. L. (2000). The five stages of small business growth. *Small business: critical perspectives on business and management, 291*.
- Clough, D. R., & Vissa, B. (2022). How do founding teams form? Toward a behavioral theory of founding team formation. In *Entrepreneurialism and Society: Consequences and Meanings* (pp. 115–147). Emerald Publishing Limited.



UNIVERSITÀ
DI PAVIA

- Colombo, M. G., & Grilli, L. (2005). Founders' *Human Capital* and the growth of new technology-based firms: A competence-based view. *Research Policy*, 34(6), 795–816. <https://doi.org/10.1016/j.respol.2005.03.010>
- Colombo, M. G., & Grilli, L. (2009). On growth drivers of *high-tech* start-ups: Exploring the role of founders' *Human Capital* and venture capital. *Journal of Business Venturing*, 25(6), 610–626. <https://doi.org/10.1016/j.jbusvent.2009.01.005>
- Colombo, M. G., & Grilli, L. (2010). On growth drivers of *high-tech* start-ups: Exploring the role of founders' *Human Capital* and venture capital. *Journal of Business Venturing*, 25(6), 610–626.
- Conti, A., & Guzman, J. A. (2023). What is the US comparative advantage in entrepreneurship? Evidence from Israeli migration to the United States. *Review of Economics and Statistics*, 105(3), 528–544.
- Cristofaro, M., Zannoni, A., Mari, M., & Poggesi, S. (2026). The pace of scaling: Founding *team* heterogeneity and the timing of Unicorn status. *Journal of Small Business and Enterprise Development*, 33(8), 1–31. <https://doi.org/10.1108/JSBED-12-2024-0650>
- Díaz-Santamaría, C., & Bulchand-Gidumal, J. (2021). Econometric estimation of the factors that influence *Startup* success. *Sustainability*, 13(4), 2242.
- Ehsan, Z.-A. (2021). Defining a *Startup*—A Critical Analysis. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3823361>



UNIVERSITÀ
DI PAVIA

- Elbahjaoui, M., & Elabjani, A. (2021). Incubation process: A key innovation lever for successful start-up *businesses*—*International Journal of Research & Review* (www.ijrrjournal.com).
- Fantozzi, I. C., Martuscelli, L., Di Luozzo, S., & Schiraldi, M. M. (2024). Soft *skills*, attitudes, and personality traits: How does the human factor matter? A systematic review and taxonomy proposal through the ProKnow-C methodology. *Businesses*, 4(2), 156–176.
- Farao, C., Bernuzzi, C., & Ronchetti, C. (2023). The crucial role of green soft *skills* and leadership for sustainability: A case study of an Italian small and medium enterprise operating in the food sector. *Sustainability*, 15(22), 15841.
- Ferreira, F., Faria, J., Azevedo, A., & Marques, A. L. (2017). Product lifecycle management in knowledge-intensive collaborative environments: An application to the automotive industry. *International Journal of Information Management*, 37(1), 1474–1487.
- Flores Laguna, O. A., Basurto Gutiérrez, K. S., & Sánchez Valdez, J. A. (2020). Validez y confiabilidad de una escala de medición del capital intelectual en PyMEs. *Contaduría y Administración*, 65(4).
- Florida, R., & Mellander, C. (2016). Rise of the *Startup* city: The changing geography of the venture capital-financed innovation. *California Management Review*, 59(1), 14–38.



UNIVERSITÀ
DI PAVIA

- Gbadegeshin, S. A., Al Natsheh, A., Ghafel, K., Mohammed, O., Koskela, A., Rimpiläinen, A., Tikkanen, J., & Kuoppala, A. (2022). Overcoming the *Valley of Death: A new model for high technology Startups. Sustainable Futures, 4*, 100077.
- Gilson, R. J. (2010). Locating innovation: The endogeneity of technology, organizational structure, and financial contracting. *Colum. L. Rev.*, *110*, 885.
- Giones, F., Gozun, B., Miralles, F., & others. (2018). Unbundling the Influence of *Human Capital* on the New Venture's Performance. *DLSU Business & Economics Review. 2019, Vol. 28, 3*.
- Giot, P., & Schwienbacher, A. (2007). IPOs, trade sales, and liquidations: Modeling venture capital exits using survival analysis. *Journal of Banking & Finance, 31*(3), 679–702.
- Giraud, E., Giudici, G., & Grilli, L. (2019). Entrepreneurship policy and the financing of young innovative companies: Evidence from the Italian *Startup Act. Research Policy, 48*(9), 103801.
- Gompers, P., Kovner, A., Lerner, J., & Scharfstein, D. (2010). Performance persistence in entrepreneurship. *Journal of Financial Economics, 96*(1), 18–32. <https://doi.org/10.1016/j.jfineco.2009.11.001>
- Gompers, P., & Kovvali, S. (2018). Diversity dividend. *Harvard Business Review, 72*, 77.



UNIVERSITÀ
DI PAVIA

- Gompers, P., & Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives*, 15(2), 145–168.
- Gottschalk, S., & Niefert, M. (2013). Gender differences in the *business* success of German start-up firms. *International Journal of Entrepreneurship and Small Business*, 18(1), 15–46.
- Greco, F. (2023). *Startup ecosystems. Studies on Entrepreneurship, Structural Change and Industrial Dynamics; Springer: Berlin/Heidelberg, Germany.*
- Greene, W. H. (2018). *Econometric analysis/Limdep users manual. Retrieved on July, 15, 2022.*
- Guzman, J. (2018). Go west, young firm: The value of entrepreneurial migration for *Startups* and founders. *Columbia Business School, working paper*, 18–49.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (Eighth edition). Cengage.
- Halt, G. B., Donch, J. C., Stiles, A. R., & Fesnak, R. (2017). *Intellectual property and financing strategies for technology startups*. Springer.
- Hsu, D. H. (2007). Experienced entrepreneurial founders, organizational capital, and venture capital funding. *Research Policy*, 36(5), 722–741.
<https://doi.org/10.1016/j.respol.2007.02.022>



UNIVERSITÀ
DI PAVIA

- Jin, L., Madison, K., Kraiczy, N. D., Kellermanns, F. W., Crook, T. R., & Xi, J. (2017). Entrepreneurial *Team* Composition Characteristics and New Venture Performance: A Meta-Analysis. *Entrepreneurship Theory and Practice*, 41(5), 743–771. <https://doi.org/10.1111/etap.12232>
- Kaiser, U., & Müller, B. (2015). *Skill* heterogeneity in *Startups* and its development over time. *Small Business Economics*, 45(4), 787–804.
- Khaykin, M. M., Lapinskas, A., & Кочергина, О. А. (2020, gennaio). The Development of the Theory of *Human Capital* in the Historical Dimension. *Proceedings of the International Conference on Economics, Management and Technologies 2020 (ICEMT 2020)*. <https://doi.org/10.2991/aebmr.k.200509.090>
- Kuzminov, Y., Сорокин, П., & Froumin, I. (2019). Generic and Specific *Skills* as Components of *Human Capital*: New Challenges for Education Theory and Practice. *Foresight-Russia*, 13(2), 19–41. <https://doi.org/10.17323/2500-2597.2019.2.19.41>
- Long, J. S., & Freese, J. (2001). *Regression models for categorical dependent variables using Stata*. Stata Press.
- Lovaglio, P., & Vittadini, G. (2004). *Il concetto di capitale umano e la sua stima*.
- Lyu, W., & Liu, J. (2021). Soft *skills*, hard *skills*: What matters most? Evidence from job postings. *Applied Energy*, 300, 117307. <https://doi.org/10.1016/j.apenergy.2021.117307>



UNIVERSITÀ
DI PAVIA

- Magalhães, R. P. C., & Oliveira, E. T. V. D. (2019). *What is a Startup? A Scoping Review on How the Literature Defines Startup*. Católica Porto Business School.
- Mahssouni, R., Touijer, M. N., & Makhroute, M. (2022). Employee compensation, training, and financial performance during the COVID-19 pandemic. *Journal of Risk and Financial Management*, 15(12), 559.
- Marchetti, R., & Manelli, A. (2021, gennaio). *I FONDI VENTURE CAPITAL E IL FINANZIAMENTO DELLE START-UP INNOVATIVE*.
- Markham, S. K. (2002). Moving technologies from lab to market. *Research-technology management*, 45(6), 31–42.
- Mayilyan, F. N., & Yedigaryan, K. (2022). THE EVOLUTION OF HUMAN CAPITAL THEORY. *ALTERNATIVE*, 68–79.
<https://doi.org/10.55528/18292828-2022.4-68>
- McCarthy, P. X., Gong, X., Braesemann, F., Stephany, F., Rizoïu, M.-A., & Kern, M. L. (2023). The impact of founder personalities on *Startup* success. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-41980-y>
- McFadden, D. (1972). *Conditional logit analysis of qualitative choice behavior*.
- Montanaro, B., Cavallo, A., Giudici, G., & Ghezzi, A. (2022a). Determinants of the *exit* value in European venture capital-backed technology *Startups*.



UNIVERSITÀ
DI PAVIA

Competitiveness Review: An International Business Journal, 32(7), 62–84.

Montanaro, B., Cavallo, A., Giudici, G., & Ghezzi, A. (2022b). Determinants of the *exit* value in European venture capital-backed technology *Startups*. *Competitiveness Review: An International Business Journal*, 32(7), 62–84. <https://doi.org/10.1108/CR-03-2021-0032>

Pisoni, A., & Onetti, A. (s.d.). *When startups exit: Comparing strategies in Europe and the USA*.

Pisoni, A., & Onetti, A. (2018). *When Startups exit: Comparing strategies in Europe and the USA*. *Journal of Business Strategy*, 39(3), 26–33. <https://doi.org/10.1108/jbs-02-2017-0022>

Potantin, M., Chertok, A., Zorin, K., & Shtabtsovsky, C. (2023). *Startup success prediction and VC portfolio simulation using CrunchBase data*. *arXiv preprint arXiv:2309.15552*.

Ratzinger, D., Amess, K., Greenman, A., & Mosey, S. (2018). The impact of digital start-up founders' higher education on reaching equity investment milestones. *The Journal of Technology Transfer*, 43(3), 760–778. <https://doi.org/10.1007/s10961-017-9627-3>

Reisdorfer-Leite, B., Oliveira, M. M. de, Rudek, M., Szejka, A. L., & Canciglieri, O. (2020). *Startup Definition Proposal Using Product Lifecycle Management*. In *IFIP advances in information and*



UNIVERSITÀ
DI PAVIA

communication technology (pp. 426–435). Springer Science+Business Media. https://doi.org/10.1007/978-3-030-62807-9_34

Sahlman, W. A. (2022). The structure and governance of venture-capital organizations. In *Venture capital* (pp. 3–51). Routledge.

Salamzadeh, A., & Kawamorita, H. (2015). *Startup Companies: Life Cycle and Challenges*. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2628861>

Sanguineti, F., Majocchi, A., & Lechner, C. (2024). From one to many: Education, diversity, and international experience in habitual vs. One-time founders. *Small Business Economics*. <https://doi.org/10.1007/s11187-024-00970-8>

Schultz, T. W. (1960). Capital formation by education. *Journal of Political Economy*, 68(6), 571–583.

Šebestová, J. D., & Popescu, C. R. G. (2022). Factors influencing investments into human resources to support company performance. *Journal of Risk and Financial Management*, 15(1), 19.

Shrivastava, M., & Tamvada, J. P. (2011). *Entrepreneurial teams, optimal team size, and founder exits*. Working Paper, University of Jena.

Smith, A. (2013). *La ricchezza delle nazioni*. Utet Libri.



UNIVERSITÀ
DI PAVIA

Sorenson, O., & Stuart, T. E. (2001). Syndication Networks and the Spatial Distribution of Venture Capital Investments. *American Journal of Sociology*, 106(6), 1546–1588. <https://doi.org/10.1086/321301>

Suzuki, H., Miyashita, S., & Sengoku, S. (2024, agosto). A Comprehensive Review of Factors Influencing Mergers and Acquisitions in Technology Startups. In 2022, *Portland International Conference on Management of Engineering and Technology (PICMET)* (Vol. 101, pp. 1–11). <https://doi.org/10.23919/picmet64035.2024.10653150>

Tomada Letizia. (2022). Start-ups and the proposed EU AI Act: Bridges or Barriers in the path from Invention to Innovation? *Journal of Intellectual Property, Information Technology and E-Commerce Law*. <https://local.forskningsportal.dk/local/dki-cgi/ws/cris-link?src=ku&id=ku-e74b43f3-a896-48d8-aa50-0bf1fe755cd9&ti=Start-ups%20and%20the%20proposed%20EU%20AI%20Act%20%3A%20Bridges%20or%20Barriers%20in%20the%20path%20from%20Invention%20to%20Innovation%3F>

Ucbasaran, D., Alsos, G. A., Westhead, P., Wright, M., & others. (2008). Habitual entrepreneurs. *Foundations and Trends® in Entrepreneurship*, 4(4), 309–450.

Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2011). Human Capital and entrepreneurial success: A meta-analytical review. *Journal of*



UNIVERSITÀ
DI PAVIA

Business Venturing, 26(3), 341–358.
<https://doi.org/10.1016/j.jbusvent.2009.09.004>

Unger, J., Rauch, A., Fresé, M., & Rosenbusch, N. (2009, novembre). *Human Capital and entrepreneurial success: A meta-analytical review*. In *Journal of Business Venturing* (Vol. 26, Fascicolo 3, pp. 341–358). Elsevier BV. <https://doi.org/10.1016/j.jbusvent.2009.09.004>

V., L. (2019). *Le forme dell'innovazione nell'ideologia californiana: le retoriche, i modelli e le trasformazioni nell'economia delle startup*. EGEA Spa.

Wooldridge, J. M. (2010). *Econometric analysis of cross-section and panel data*. MIT Press.

Zapata-Molina, C., Bedoya-Villa, M., Castro-Gómez, J., Gutiérrez-Broncano, S., Román, E., & Rave-Gómez, E. (2025). Factors Affecting the Financial Sustainability of *Startups* During the *Valley of Death*: An Empirical Study in an Innovative Ecosystem. *International Journal of Financial Studies*, 13(2), 73.



UNIVERSITÀ
DI PAVIA

Appendice

Appendice A – Codici Python

Codice 1: Estrazione Categorie, Sottocategorie e UUID Organizzazioni

```
import os
import csv
import time
import requests
API_KEY = os.getenv("CRUNCHBASE_API_KEY")
BASE_URL = "https://API.crunchbase.com/API/v4/searches"
HEADERS = {
    "accept": "application/json",
    "Content-Type": "application/json",
    "X-cb-user-key": API_KEY }
DATA_DIR = "data"
CATEGORY_DIR = os.path.join(DATA_DIR, "categories")
ORG_DIR = os.path.join(DATA_DIR, "organizations")
os.makedirs(DATA_DIR, exist_ok=True)
os.makedirs(CATEGORY_DIR, exist_ok=True)
os.makedirs(ORG_DIR, exist_ok=True)
def post_request(endpoint, payload, delay=0.5):
    response = requests.post(f"{BASE_URL}/{endpoint}", headers=HEADERS, json=payload)
    if response.status_code != 200:
        return None
    at time.sleep(delay)
    return response.json().get("entities", [])
def export_categories(endpoint, output_file):
    payload = {
        "field_ids": ["identifier", "name"],
        "limit": 1000 }
    entities = post_request(endpoint, payload)
    with open(output_file, "w", newline="", encoding="utf-8") as f:
        writer = csv.writer(f)
        writer.writerow(["uuid", "name"])
        if entities:
            for e in entities:
                props = e.get("properties", {})
                uuid = props.get("identifier", {}).get("uuid", "")
                name = props.get("name", "")
                writer.writerow([uuid, name])
def export_organizations_by_category(category_csv):
    with open(category_csv, encoding="utf-8") as f:
        reader = csv.reader(f)
        next(reader)
```



UNIVERSITÀ DI PAVIA

```
for uuid, name in reader:
    payload = {
        "field_ids": ["identifier"],
        "query": [
            {
                "type": "predicate",
                "field_id": "categories",
                "operator_id": "includes",
                "values": [uuid]
            },
            {
                "type": "predicate",
                "field_id": "facet_ids",
                "operator_id": "includes",
                "values": ["company"]
            }
        ],
        "limit": 1000
    }
    entities = post_request("organizations", payload)
    output_path = os.path.join(ORG_DIR, f"{uuid}.csv")
    with open(output_path, "w", newline="", encoding="utf-8") as out:
        writer = csv.writer(out)
        writer.writerow(["organization_name", "organization_uuid"])
    if entities:
        for org in entities:
            props = org.get("properties", {})
            writer.writerow([
                props.get("identifier", {}).get("value", ""),
                props.get("identifier", {}).get("uuid", "")])
def export_founder_data():
    output_file = os.path.join(DATA_DIR, "organization_founders.csv")
    with open(output_file, "w", newline="", encoding="utf-8") as fout:
        writer = csv.writer(fout)
        writer.writerow([
            "organization_name",
            "organization_uuid",
            "founded_year",
            "founders"
        ])
    for file in os.listdir(ORG_DIR):
        path = os.path.join(ORG_DIR, file)
        with open(path, encoding="utf-8") as f:
            reader = csv.reader(f)
            next(reader)
            for name, uuid in reader:
                payload = {
                    "field_ids": ["identifier", "founder_identifiers", "founded_on"],
                    "query": [
                        {
                            "type": "predicate",
                            "field_id": "uuid",
                            "operator_id": "eq",
                            "values": [uuid]
                        }
                    ],
                    "limit": 1
                }
                entities = post_request("organizations", payload)
                if not entities:
                    writer.writerow([name, uuid, "", ""])
                    continue
                props = entities[0].get("properties", {})
                founded = props.get("founded_on", {}).get("value", "")
                founders = props.get("founder_identifiers", [])
                founders_str = " | ".join(
                    f"{f.get('value', '')} ({f.get('uuid', '')})"
                    for f in founders
                )
                writer.writerow([name, uuid, founded, founders_str])
def main():
    export_categories(
```



UNIVERSITÀ DI PAVIA

```
endpoint="category_groups",
output_file=os.path.join(DATA_DIR, "category_groups.csv") )
export_categories(
    endpoint="categories",
    output_file=os.path.join(DATA_DIR, "categories.csv") )
export_organizations_by_category(
    os.path.join(DATA_DIR, "categories.csv") )
export_founder_data()
if __name__ == "__main__":
    main()
```

Codice 2: Estrazione Cards

```
import os
import csv
import json
import signal
import requests
from typing import Set
API_KEY = os.getenv("CRUNCHBASE_API_KEY")
BASE_DIR = "data"
INPUT_CSV = os.path.join(BASE_DIR, "founders_dataset.csv")
OUTPUT_DIR = os.path.join(BASE_DIR, "cards")
CHECKPOINT_DIR = os.path.join(BASE_DIR, "checkpoints")
CARD_ENDPOINTS = {"organization_cards":
    "entities/organizations/{uuid}?card_ids={card}",
    "people_cards":
    "entities/people/{uuid}/cards/{card}"}
CARDS = [("organization_cards", "founders"), ("organization_cards",
    "headquarters_address"), ("organization_cards", "fields"), ("people_cards",
    "degrees"), ("people_cards", "jobs"),
    ("people_cards", "primary_job"), ("people_cards",
    "primary_organization"), ("people_cards", "founded_organizations")]
os.makedirs(OUTPUT_DIR, exist_ok=True)
os.makedirs(CHECKPOINT_DIR, exist_ok=True)
stop_requested = False
def handle_interrupt(signum, frame):
    global stop_requested
    stop_requested = True
signal.signal(signal.SIGINT, handle_interrupt)
signal.signal(signal.SIGTERM, handle_interrupt)
def parse_founders(cell):
    result = []
    for item in cell.split("|"):
        if "(" in item and ")" in item:
            name = item[:item.find("(")].strip()
            uuid = item[item.find("(")+1:item.find(")")].strip()
            result.append((name, uuid))
    return result
def load_checkpoint(card_name) -> Set[str]:
    path = os.path.join(CHECKPOINT_DIR, f"{card_name}.txt")
    if os.path.exists(path):
        with open(path, "r", encoding="utf-8") as f:
            return set(line.strip() for line in f if line.strip())
```



UNIVERSITÀ DI PAVIA

```
return set()
def save_checkpoint(card_name, uuid):
    path = os.path.join(CHECKPOINT_DIR, f"{card_name}.txt")
    with open(path, "a", encoding="utf-8") as f:
        f.write(uuid + "\n")
def fetch_card(entity_type, card, uuid):
    template = CARD_ENDPOINTS[entity_type]
    url = f"https://API.crunchbase.com/API/v4/" + template.format(uuid=uuid,
card=card)
    headers = {"accept": "application/json", "X-cb-user-key": API_KEY}
    try:
        response = requests.get(url, headers=headers, timeout=15)
        if response.status_code == 200:
            return response.json()
    except requests.exceptions.RequestException:
        pass
    return None
def main():
    with open(INPUT_CSV, "r", encoding="utf-8") as file:
        reader = csv.DictReader(file)
        for row in reader:
            founders = parse_founders(row["Fondatori (Nome e UUID)"])
            for entity_type, card in CARDS:
                completed = load_checkpoint(card)
                for name, uuid in founders:
                    if stop_requested:
                        return
                    if uuid in completed:
                        continue
                    data = fetch_card(entity_type, card, uuid)
                    if not data:
                        continue
                    card_dir = os.path.join(OUTPUT_DIR, card)
                    os.makedirs(card_dir, exist_ok=True)
                    with open(os.path.join(card_dir, f"{uuid}.json"), "w",
encoding="utf-8") as f:
                        json.dump(data, f, ensure_ascii=False, indent=2)
                    save_checkpoint(card, uuid)
if __name__ == "__main__":
    main()
```



UNIVERSITÀ
DI PAVIA

Appendice B – Codici R

Codice 3: Creazione CSV Cards

```
library(jsonlite)
library(dplyr)
BASE_DIR <- "data"
INPUT_SUBDIRS <- c("fields", "headquarters", "founded_organizations",
"primary_organization", "primary_jobs", "jobs", "degrees")
OUTPUT_DIR <- file.path(BASE_DIR, "processed")
dir.create(OUTPUT_DIR, recursive = TRUE, showWarnings = FALSE)
BLOCK_SIZE <- 500
PREVENT_STANDBY <- FALSE
prevent_standby <- function() {
  if (PREVENT_STANDBY && .Platform$OS.type == "windows") {
    system("powercfg -change -standby-timeout-ac 0", show.output.on.console =
FALSE)
    system("powercfg -change -standby-timeout-dc 0", show.output.on.console =
FALSE) }}
restore_standby <- function() {
  if (PREVENT_STANDBY && .Platform$OS.type == "windows") {
    system("powercfg -change -standby-timeout-ac 15", show.output.on.console =
FALSE)
    system("powercfg -change -standby-timeout-dc 15", show.output.on.console =
FALSE) }}
flatten_json <- function(x, prefix = "") {
  result <- list()
  if (is.null(x)) return(result)
  if (is.list(x)) { nms <- names(x)
    if (is.null(nms)) nms <- seq_along(x)
    for (i in seq_along(x)) {
      name <- gsub("[^A-Za-z0-9_]", "_", as.character(nms[i]))
      new_prefix <- if (prefix == "") name else paste0(prefix, "_", name)
      result <- c(result, flatten_json(x[[i]], new_prefix)) } }
  else if (length(x) > 1) {result[[if(prefix == "") "value" else prefix]] <-
paste(as.character(x), collapse = "|")} else {
  result[[if(prefix == "") "value" else prefix]] <- as.character(x) }
  result}process_json <- function(path) {
  tryCatch({ data <- fromJSON(path, simplifyVector = FALSE)
  flatten_json(data) }, error = function(e) NULL)}
has_content <- function(flattened) { if (is.null(flattened)) return(FALSE)
  cols <- names(flattened)[grepl("^cards_", names(flattened))]
  any(sapply(flattened[cols], function(v) ! is.null(v) && v != "" && ! is.na(v)))}
```



UNIVERSITÀ DI PAVIA

```
order_columns <- function(cols) {
  props <- sort(cols[grepl("^properties_", cols)])
  cards <- sort(cols[grepl("^cards_", cols)])
  others <- sort(setdiff(cols, c(props, cards)))
  c(props, cards, others)}
sanitize <- function(x) {if (is.null(x) || is.na(x)) return("")
  x <- gsub("[\r\n\t]+", " ", as.character(x)) trimws(x)}
prevent_standby()
on.exit(restore_standby())
for (subdir in INPUT_SUBDIRS) {
  json_dir <- file.path(BASE_DIR, "cards", subdir)
  if (!dir.exists(json_dir)) next
  output_csv <- file.path(OUTPUT_DIR, paste0(subdir, ".csv"))
  checkpoint_file <- file.path(OUTPUT_DIR, paste0(subdir, "_checkpoint.txt"))
  columns_cache <- file.path(OUTPUT_DIR, paste0(subdir, "_columns.rds"))
  json_files <- list.files(json_dir, pattern = "\\\\.json$", full.names = TRUE)
  if (length(json_files) == 0) next
  processed <- if (file.exists(checkpoint_file)) readLines(checkpoint_file)
  else character(0)
  remaining <- setdiff(json_files, processed)
  if (length(remaining) == 0) next
  columns <- if (file.exists(columns_cache)) {readRDS(columns_cache)} else
{ cols <- character(0) for (f in json_files) { data <- process_json(f)
  if(!is.null(data))cols<-union(cols,names(data))}cols<- rder_columns(cols)
  saveRDS(cols, columns_cache)cols }
  first_write <- !file.exists(output_csv)
  buffer <- list()
  for (i in seq_along(remaining)) {
    data <- process_json(remaining[i])
    if (!is.null(data) && has_content(data)) {
      row <- sapply(columns, function(col) {
        if (col %in% names(data)) sanitize(data[[col]]) else ""})
      buffer[[length(buffer) + 1]] <- row }
    write(remaining[i], checkpoint_file, append = TRUE)
    if (length(buffer) >= BLOCK_SIZE || i == length(remaining)) {
      if (length(buffer) > 0) { df <- as.data.frame(do.call(rbind, buffer),
stringsAsFactors = FALSE)
        colnames(df) <- columns
      write.table(df,output_csv,append = !first_write,sep = ",",row.names = FALSE,
col.names = first_write,quote = TRUE,fileEncoding = "UTF-8")
      first_write <- FALSE
      buffer <- list()
      gc() } }
    summary_file <- file.path(OUTPUT_DIR, paste0(subdir, "_summary.txt"))
    summary_text <- paste(
      "JSON Processing Summary",
      paste("Processed on:", Sys.time()),
      paste("Total JSON files:", length(json_files)),
      paste("Columns:", length(columns)),
      sep = "\n" )
    writeLines(summary_text, summary_file)}
```



UNIVERSITÀ DI PAVIA

Codice 4: Modifica CSV Cards

```
library(readr)
library(dplyr)
library(shiny)
library(DT)
# Percorsi dei file (da definire a seconda dell'ambiente)
file_input <- "file_input.csv"
file_output <- "file_output.csv"
file_colonne <- "colonne_mantenute.txt"
chunk_size <- 1000
estrai_nome_base <- function(nome_colonna) {
  nome_base <- gsub("_?\\d+$", "", nome_colonna)
  if (nchar(nome_base) < 2) return(nome_colonna)
  return(nome_base)}
colnames_total <- names(read_csv(file_input, n_max = 1, col_types =
cols(.default = col_character()))))
df_colonne <- data.frame(
  nome_originale = colnames_total,
  nome_base = sapply(colnames_total, estrai_nome_base),
  stringsAsFactors = FALSE)
gruppi_colonne <- df_colonne %>%
  group_by(nome_base) %>%
  summarise(count = n(),colonne = list(nome_originale),esempi =
paste(head(nome_originale, 3), collapse = ", "), .groups = 'drop') %>%
  arrange(desc(count), nome_base)
ui <- fluidPage(
  titlePanel("Selezione Colonne per Gruppi e Parole Chiave"),
  tabsetPanel(
    tabPanel("Selezione per Gruppi", sidebarLayout(
      sidebarPanel(actionButton("seleziona_tutti", "Seleziona Tutti"),
actionButton("deseleziona_tutti", "Deseleziona Tutti"),
actionButton("conferma_gruppi", "Conferma Selezione")),
      mainPanel(DT::dataTableOutput("tabella_gruppi"))),
    tabPanel("Selezione per Parole Chiave", sidebarLayout(sidebarPanel(textInput("parole_chiave", "Parole
chiave (separate da virgola):", ""),actionButton("cerca_parole", "Cerca
Colonne"),actionButton("conferma_parole", "Conferma Selezione")
), mainPanel(DT::dataTableOutput("tabella_parole"))) ) ) )
server <- function(input, output, session) {
  selected_rows <- reactiveVal(integer(0))
  colonne_selezionate_parole <- reactiveVal(character(0))
  colonne_finali <- reactiveVal(character(0))
  output$tabella_gruppi <- DT::renderDataTable({
    DT::datatable(gruppi_colonne %>% select(nome_base, count, esempi),
      selection = list(mode='multiple', selected =
selected_rows()))})
```



UNIVERSITÀ
DI PAVIA

```
observeEvent(input$tabella_gruppi_rows_selected, {
selected_rows(input$tabella_gruppi_rows_selected) })
observeEvent(input$seleziona_tutti, {
selected_rows(1:nrow(gruppi_colonne)) })
observeEvent(input$deseleziona_tutti, { selected_rows(integer(0)) })
observeEvent(input$conferma_gruppi, {sel <- selected_rows()
if(length(sel) > 0) {
colonne_finali(unique(c(colonne_finali(),
unlist(gruppi_colonne$colonne[sel]))))
showNotification("Gruppi selezionati aggiunti!") } })
observeEvent(input$cerca_parole, {
parole <- trimws(unlist(strsplit(input$parole_chiave, ",")))
colonne_trovate <- unique(unlist(lapply(parole, function(p) grep(p,
colnames_total, value=TRUE, ignore.case=TRUE))))
colonne_selezionate_parole(colonne_trovate)
showNotification(paste("Trovate", length(colonne_trovate), "colonne"))
})
observeEvent(input$conferma_parole, {
colonne_finali(unique(c(colonne_finali(),
colonne_selezionate_parole())))
showNotification("Colonne trovate aggiunte!") })
observeEvent(input$conferma_gruppi, { stopApp(colonne_finali()) })}
colonne_mantenute <- shiny::runApp(list(ui=ui, server=server))
writeLines(colonne_mantenute, file_colonne)
write_lines(paste(colonne_mantenute, collapse = ","), file_output)
read_csv_chunked( file_input, callback = DataFrameCallback$new(function(df,
pos) {
df_filtered <- df %>% select(all_of(colonne_mantenute))
write_csv(df_filtered, file_output, append=TRUE, col_names=FALSE) }),
chunk_size = chunk_size,
col_types = cols(.default = col_character()))
cat("Elaborazione completata!\n")
cat("Colonne originali:", length(colnames_total), "\n")
cat("Colonne mantenute:", length(colonne_mantenute), "\n")
```

Codice 5: Creazione database Intermedi

```
required_packages <- c("data.table", "progress", "stringr")
for (pkg in required_packages) { if (!require(pkg, character.only = TRUE,
quietly = TRUE)) {
install.packages(pkg, quiet = TRUE)
library(pkg, character.only = TRUE, quietly = TRUE) }}
options(datatable.optimize = Inf)
setDTthreads(0)
prevent_standby <- function() {
if (.Platform$OS.type == "windows") {
system("powercfg -change -standby-timeout-ac 0", show.output.on.console
= FALSE)
system("powercfg -change -standby-timeout-dc 0", show.output.on.console
= FALSE) }}
log_message <- function(message, log_file = NULL, level = "INFO") {
```



UNIVERSITÀ DI PAVIA

```
ts <- format(Sys.time(), "%Y-%m-%d %H:%M:%S")
msg <- paste0("[", ts, "] [", level, "] ", message)
cat(msg, "\n")
if (!is.null(log_file)) write(msg, file = log_file, append = TRUE)}
sanitize_text <- function(x) {
  x <- as.character(x)
  x[is.na(x)] <- ""
  x <- gsub("[\r\n\t]+", " ", x)
  x <- gsub("\"", "'", x)
  x <- gsub(",", ";", x)
  trimws(gsub(" +", " ", x))}
read_csv_robust <- function(file, error_log, header_pattern = NULL) {
  tryCatch({
    if (!is.null(header_pattern)) {dt <- fread(file, encoding = "UTF-8", sep
= ",", quote = "\"", fill = TRUE, skip = header_pattern, showProgress =
FALSE)
    } else {dt <- fread(file, encoding = "UTF-8", sep = ",", quote = "\"",
fill = TRUE, showProgress = FALSE)
    }
    dt }, error = function(e) { log_message(paste("Errore lettura:",
basename(file), e$message), error_log, "ERROR") NULL })}
load_checkpoint <- function(f) if (file.exists(f)) readLines(f) else
character(0)
save_checkpoint <- function(v, f) writelines(unique(v), f)
merge_principale_secondario <- function(principale_file, secondario_file,
output_file, checkpoint_file = NULL, error_log = NULL, batch_size = 1000) {
  if (is.null(checkpoint_file)) checkpoint_file <- tempfile("checkpoint_")
  if (is.null(error_log)) error_log <- tempfile("error_log_")
  prevent_standby()
  log_message("Lettura CSV principale")
  main_dt <- read_csv_robust(principale_file, error_log)
  if (is.null(main_dt)) stop("Errore file principale")
  if (!("uuid_org" %in% names(main_dt))) stop("uuid_org mancante")
  uuid_founder_cols <- grep("^uuid_founder_", names(main_dt), value = TRUE)
  if (length(uuid_founder_cols) == 0) stop("Nessuna colonna uuid_founder
trovata")
  main_dt[, .row_id := .I]
  log_message("Lettura CSV secondario")
  sec_dt <- read_csv_robust(secondario_file, error_log)
  if (is.null(sec_dt)) stop("Errore file secondario")
  if (!("uuid_fondatore" %in% names(sec_dt))) stop("uuid_fondatore
mancante")
  setkey(sec_dt, uuid_fondatore)
  processed <- load_checkpoint(checkpoint_file)
  org_ids <- setdiff(unique(main_dt$uuid_org), processed)
  pb <- progress::progress_bar$new(total = length(org_ids), format = "Merge
[:bar] :percent | ETA: :eta")
  for (i in seq_along(org_ids)) {
    fid <- org_ids[i]
    idx <- which(main_dt$uuid_org == fid)
    rows <- main_dt[idx]
    founder_values <- unique(unlist(rows[, ..uuid_founder_cols]))
    founder_values <- founder_values[founder_values != ""]
    if (length(founder_values) == 0) next
  }
}
```



UNIVERSITÀ DI PAVIA

```
matched <- sec_dt[J(founder_values), nomatch = 0]
if (nrow(matched) == 0) next
matched <- matched[, lapply(.SD, sanitize_text)]
updates <- list()
main_dt[idx, names(updates) := updates]
processed <- c(processed, fid)
save_checkpoint(processed, checkpoint_file)
pb$tick()
if (i %% 500 == 0) gc() }
setorder(main_dt, .row_id)
main_dt[, .row_id := NULL]
fwrite(main_dt, output_file, sep = ",", quote = TRUE, na = "", encoding =
"UTF-8")
log_message(paste("Merge completato:", output_file))}
if (interactive()) {
  merge_principale_secondario(
    principale_file = "sample_principale.csv",
    secondario_file = "sample_secondario.csv",
    output_file = "sample_output.csv" )}
```

Codice 6: Pulizia Database Intermedi

```
library(data.table)
library(stringr)
library(progress)
prevent_standby <- function() {
  if (.Platform$OS.type == "windows") {
    system("powercfg -change -standby-timeout-ac 0", show.output.on.console
= FALSE)
    system("powercfg -change -standby-timeout-dc 0", show.output.on.console
= FALSE) }}
read_normalization_txt <- function(txt_path, type = "simple") {
  lines <- readlines(txt_path, warn = FALSE, encoding = "UTF-8")
  if (type == "column") { norm_list <- list() current_col <- NULL
  for (line in lines) {
    line <- str_trim(line)
    if (str_starts(line, "==== COLONNA:")) {
      current_col <- str_match(line, "====
COLONNA:\\s*(.*?)\\s*====")[1,2]
      norm_list[[current_col]] <- list()
    } else if (!is.null(current_col) & line != "") {
      m <- str_match(line, '^"?(.*?)"?\\s*\\((.*?)\\)$')
      if (!is.na(m[1,1])) {
        key <- str_to_lower(str_replace_all(str_trim(m[1,2]), "\\s+", ""))
        norm_list[[current_col]][[key]] <- str_trim(m[1,3]) } } }
  return(norm_list)} else if (type == "priority") {
  df <- data.frame(key = character(), level = character(), stringsAsFactors
= FALSE)
  for (line in lines) {
    line <- str_trim(line)
    m <- str_match(line, '^"?(.*?)"?\\s*\\((BA|MA|PhD)\\)$')
```



UNIVERSITÀ DI PAVIA

```
if (!is.na(m[1,1])) {
  df <- rbind(df, data.frame(key = str_to_lower(m[1,2]), level =
m[1,3], stringsAsFactors = FALSE)) } }
df$priority <- c(BA=1, MA=2, PhD=3)[df$level]
return(df) } else { mapping <- list() for (line in lines) {
  line <- str_trim(line)
  m <- str_match(line, '^"?(.*?)"?\\s*\\(((.*?)\\))$')
  if(!is.na(m[1,1]))mapping[[str_to_lower(m[1,2])]] <- str_trim(m[1,3])
} return(mapping) }}
normalize_column <- function(col_vec, map_df, pb = NULL) {
result <- character(length(col_vec))
col_lower <- tolower(col_vec)
for (i in seq_len(nrow(map_df))) {
  idx <- str_detect(col_lower, fixed(map_df$key[i]))
  replace_idx <- idx & (result == "" | map_df$priority[i] > c(0,
map_df$priority[i])[match(result, map_df$level, nomatch=0)])
  result[replace_idx] <- map_df$level[i]
  if (!is.null(pb)) pb$tick() } return(result)}
tryCatch({ prevent_standby()
  dt <- fread("input.csv", encoding = "UTF-8")
  norm_cols <- intersect(names(dt),
names(read_normalization_txt("mapping1.txt", type="column")))
  pb <- progress_bar$new(total = length(norm_cols), format="Normalizzazione
[:bar] :percent ETA: :eta", width=40)
  for (col in norm_cols) {
    mapping <- read_normalization_txt("mapping1.txt", type="column")[[col]]
    dt[[col]] <- sapply(dt[[col]], function(x) {
      if (is.na(x)) return("")
      key <- str_to_lower(str_replace_all(as.character(x), "\\s+", ""))
      if (key %in% names(mapping)) mapping[[key]] else "" } ) pb$tick() }
  col_single <- "funding_stage"
  mapping_single <- read_normalization_txt("mapping2.txt")
  dt[[col_single]] <- sapply(dt[[col_single]], function(x) {
    if (is.na(x) || x=="") return("")
    x_norm <- str_to_lower(as.character(x))
    for (k in names(mapping_single)) if (str_detect(x_norm, fixed(k,
ignore_case=TRUE))) return(mapping_single[[k]] " " })
  cols_multi <- grep("^titolo_studio_\\d+$", names(dt), value=TRUE)
  map_priority <- read_normalization_txt("mapping3.txt", type="priority")
  pb <- progress_bar$new(total = length(cols_multi)*nrow(map_priority),
format="Normalizzazione [:bar] :percent ETA: :eta", width=40)
  for (col in cols_multi) dt[[col]] <- normalize_column(dt[[col]],
map_priority, pb)
  cols_generic <- c("genere_fondatore", "job_type_1") # esempio generico
  for (col in cols_generic) {
    mapping_gen <- read_normalization_txt(paste0(col, ".txt"))
    keys <- names(mapping_gen)
    dt[[col]] <- sapply(dt[[col]], function(x) {
      if (is.na(x) || x=="") return("")
      x_norm <- str_to_lower(as.character(x))
      for (k in keys) if (str_detect(x_norm, fixed(k, ignore_case=TRUE)))
return(mapping_gen[[k]] " " }) }
  fwrite(dt, "output.csv", sep=",", quote=TRUE, na="", encoding="UTF-8")
```



UNIVERSITÀ
DI PAVIA

```
cat("\n☑ CSV normalizzato salvato\n")  
}, error=function(e){cat("\n✗ ERRORE:", e$message, "\n")}}
```

Appendice C – Codici Stata

Codice 7: Costruzione dataset

```
clear  
set more off  
set linesize 255  
  
*-----  
* Apertura dataset organizzazioni  
*-----  
use "dataset_organizzazioni.dta", clear  
  
gen founded_date = date(founded_on, "YMD")  
format founded_date %td  
  
keep if org_type=="for_profit" & operating_status=="active" &  
inrange(year(founded_date),2010,2020)  
save "dataset_organizzazioni_filtrato.dta", replace  
  
*-----  
* Creazione dataset collegamento fondatori  
*-----  
use "dataset_organizzazioni_filtrato.dta", clear  
keep uuid_org uuid_founder_*  
reshape long uuid_founder_, i(uuid_org) j(n_founder)  
rename uuid_founder_ uuid_fondatore  
drop if missing(uuid_fondatore)  
bysort uuid_org: gen team_size = _N  
save "dataset_link_org_fondatori.dta", replace  
  
*-----  
* Ordinale titoli studio fondatori  
*-----  
use "dataset_fondatori_normalizzato.dta", clear  
  
gen BA=0  
gen MA=0  
gen PhD=0  
forvalues i=1/5 {  
    replace BA = BA + 1 if titolo_studio_`i'=="BA"  
    replace MA = MA + 1 if titolo_studio_`i'=="MA"  
    replace PhD = PhD + 1 if titolo_studio_`i'=="PhD"  
}
```



UNIVERSITÀ DI PAVIA

```
save "tmp_titoli_fondatore.dta", replace

*-----
* Aggregazione a livello team
*-----
use "dataset_link_org_fondatori.dta", clear
merge m:1 uuid_fondatore using "tmp_titoli_fondatore.dta"
drop if _merge!=3
drop _merge

collapse (sum) BA MA PhD, by(uuid_org)

gen numero_team = BA + MA + PhD
gen titolo_max_team = cond(PhD>0,3,cond(MA>0,2,cond(BA>0,1,0)))
gen titolo_medio_team = (BA*1 + MA*2 + PhD*3)/(BA+MA+PhD)
drop if BA==0 & MA==0 & PhD==0
save "dataset_titoli_team.dta", replace

*-----
* Dummy settoriali, n_settori, Blau index
*-----
use "dataset_fondatori_normalizzato.dta", clear

foreach var of varlist sec_settore_ex_jobs_* {
  gen `var'_clean = lower(`var')
  replace `var'_clean = subinstr(`var'_clean, " ", "_", .)
  replace `var'_clean = subinstr(`var'_clean, "&", "and", .)
  replace `var'_clean = subinstr(`var'_clean, "-", "_", .)
  replace `var'_clean = subinstr(`var'_clean, "/", "_", .)
}

local all_sectors ""
foreach var of varlist sec_settore_ex_jobs_*_clean {
  levelsof `var', local(tmp)
  foreach s of local tmp local all_sectors `all_sectors' "`s'"
}
local all_sectors : list uniq all_sectors

local i=1
local short_names ""
foreach s of local all_sectors {
  local short = "s`i'"
  local short_names `short_names' `short'
  local ++i
}

local i=1
foreach s of local all_sectors {
  local short = "s`i'"
  gen byte `short' = 0
  foreach var of varlist sec_settore_ex_jobs_*_clean {
    replace `short'=1 if `var'=="`s'"
  }
}
```



UNIVERSITÀ
DI PAVIA

```
    local ++i
  }

  local dummies `short_names'
  egen n_settori=rowtotal(`dummies')
  gen blau_index=0
  foreach var of local dummies {
    replace blau_index = blau_index + (`var'/n_settori)^2 if n_settori>0
  }
  replace blau_index = 1 - blau_index if n_settori>0

  keep uuid_fondatore s* n_settori blau_index
  drop sec*
  drop stage_*
  drop if missing(n_settori)|n_settori==0
  save "dataset_blau_singolo_fondatore.dta", replace

  * Aggregazione Blau a livello team
  use "dataset_link_org_fondatori.dta", clear
  merge m:1 uuid_fondatore using "dataset_blau_singolo_fondatore.dta"
  drop if _merge!=3
  drop _merge

  collapse (max) s* (first) team_size, by(uuid_org)
  egen n_settori_team=rowtotal(s*)
  gen blau_team=0
  foreach var of varlist s* {
    replace blau_team = blau_team + (`var'/n_settori_team)^2 if
n_settori_team>0
  }
  replace blau_team = 1-blau_team if n_settori_team>0
  save "dataset_blau_team.dta", replace

  *-----
  * Variabili controllo fondatori
  *-----
  use "dataset_fondatori_normalizzato.dta", clear
  keep uuid_fondatore residenza_fondatore genere_fondatore
  save "dataset_genere_residenza.dta", replace

  use "dataset_organizzazioni_filtrato.dta", clear
  keep uuid_org founded_date
  save "dataset_data_fondazione.dta", replace

  use "dataset_link_org_fondatori.dta", clear
  merge m:1 uuid_fondatore using "dataset_genere_residenza.dta"
  drop if _merge!=3
  drop _merge

  replace genere_fondatore = strtrim(lower(genere_fondatore))
  gen female = (genere_fondatore=="female") if genere_fondatore!=""
  replace residenza_fondatore = strtrim(residenza_fondatore)
```



UNIVERSITÀ DI PAVIA

```
gen north_america_f = (residenza_fondatore=="North America") if
residenza_fondatore!="""

collapse (max) female_presence=female (max) north_america=north_america_f
(count) n_fondatori=female, by(uuid_org)
merge 1:1 uuid_org using "dataset_data_fondazione.dta"
drop if _merge!=3
drop _merge
gen org_age = 2026 - year(founded_date)
drop founded_date
drop if missing(north_america)
save "dataset_variabili_controllo.dta", replace

*-----
* Numero di organizzazioni fondate dai fondatori
*-----
use "dataset_fondatori_normalizzato.dta", clear
keep uuid_fondatore num_org_fondate
destring num_org_fondate, replace force
save "dataset_num_org_intermedio.dta", replace

use "dataset_link_org_fondatori.dta", clear
merge m:1 uuid_fondatore using "dataset_num_org_intermedio.dta"
drop if _merge!=3
drop _merge
gen exp_fondatore_net = num_org_fondate - 1
replace exp_fondatore_net = 0 if exp_fondatore_net<0
collapse (sum) num_org_fondate_prec=exp_fondatore_net (first) team_size,
by(uuid_org)
keep uuid_org team_size num_org_fondate_prec
save "dataset_num_org_fondate_prec.dta", replace

*-----
* Variabile binaria successo organizzazione
*-----
use "dataset_organizzazioni_filtrato.dta", clear
keep uuid_org funding_stage
gen success = inlist(funding_stage,"late_stage","M&A","IPO")
save "dataset_variabile_binaria.dta", replace

*-----
* Merge master dataset variabili
*-----
use "dataset_variabili_controllo.dta", clear
merge 1:1 uuid_org using "dataset_titoli_team.dta", keepusing(BA MA PhD
titolo_max_team titolo_medio_team)
drop if _merge!=3
drop _merge

merge 1:1 uuid_org using "dataset_blau_team.dta", keepusing(s* team_size
n_settori_team blau_team)
drop if _merge!=3
drop _merge
```



UNIVERSITÀ
DI PAVIA

```
merge 1:1 uuid_org using "dataset_num_org_fondate_prec.dta",
keepusing(num_org_fondate_prec team_size)
drop if _merge!=3
drop _merge

merge 1:1 uuid_org using "dataset_variabili_controllo.dta"
drop if _merge!=3
drop _merge

merge 1:1 uuid_org using "dataset_variabile_binaria.dta"
drop if _merge!=3
drop _merge

*-----
* Pulizia finale e rinomina variabili
*-----
drop n_fondatori
drop if missing(north_america)
rename north_america presence_north_american_founder
rename num_org_fondate_prec sum_prev_org_founded_team
rename blau_team blau_team_job_exp
rename org_age org_age
rename titolo_max_team top_study_title
rename titolo_medio_team mean_study_title

order uuid_org success funding_stage org_age team_size female_presence
presence_north_american_founder mean_study_title blau_team_job_exp
sum_prev_org_founded_team BA MA PhD top_study_title n_settori_team, first

save "dataset_definitivo_regressione.dta", replace
```