

POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea Magistrale



Sharing economy e differenze di genere: studio delle recensioni di Airbnb ed analisi sulla congruità di genere

Relatore

Laura Rondi

Candidato

Elisa Giacone

Anno accademico 2020/2021

Sommario

Abstract	4
1. Sharing Economy.....	6
1.1 Definizione.....	6
1.2 Driver della sharing economy.....	8
1.3 Vantaggi e svantaggi.....	9
1.4 Effetti ambientali.....	12
1.5 Valore di mercato	18
1.6 Impatto Covid-19	19
2. Piattaforme digitali	23
2.1 Ruolo delle piattaforme	23
2.2 Esternalità	26
2.3 Modelli di pricing	27
3. Regolazione	31
3.1 Inquadramento generale.....	31
3.2 Sfide regolatorie	32
3.3 Sharing Economy Act	34
3.4 I dati nella sharing economy.....	37
4. Airbnb	41
4.1 Introduzione.....	41
4.2 Impatto Airbnb su settore turistico.....	42
4.3 Airbnb vs Booking.....	45
4.4 Airbnb vs Hotel	46
4.5 Diffusione Airbnb.....	49
5. Il sistema delle recensioni per Airbnb	52
5.1 Inside the Engine Room of Digital Platforms: Reviews, Ratings, and Recommendations.....	53
5.2 The determinants of online review informativeness: evidence from field experiments on Airbnb	58
5.3 A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average.....	61

5.4 You get what you give: theory and evidence of reciprocity in the sharing economy	65
6. Uomini e donne: esiste veramente una differenza di genere?.....	70
6.1 Does gender bias exist? The impact of gender congruity on consumer's Airbnb booking intention and the mediating role of trust	72
6.2 Uomini e donne valutano le recensioni online in modo diverso?	76
7. Costruzione Database.....	79
7.1 Descrizione dati	81
8. Analisi database	84
8.1 Chi recensisce di più?	84
8.2 Aspetti delle recensioni.....	85
8.3 Confidenza	90
8.4 Host di Aibnb come lavoro.....	91
9. Analisi descrittive	95
9.1 Prezzo.....	95
9.2 Review al mese	96
9.3 Punteggi	97
9.4 Disponibilità	102
9.5 Numero reviews	103
9.6 Host response rate	104
9.7 Listing_type.....	105
9.8 Superhost	107
9.9 Numero listings	108
9.10 Caratteristiche del listing.....	109
9.11 Distanza.....	110
10. Analisi di regressione.....	112
10.1 Analisi di regressione sul prezzo	113
10.2 Entire home vs Private room	120
10.3 Relazione tra parole nelle recensioni e prezzo	124
10.4 Analisi di regressione sulla domanda	127
10.5 Tipologia di listing.....	131
10.6 Relazione tra parole nelle recensioni e domanda	132

10.7 I guest donne si rivolgono ad host donne?.....	135
11. Conclusioni.....	141
12. Bibliografia.....	144
13. Sitografia.....	145
14. Appendice.....	147
Ringraziamenti	150

Abstract

La piattaforma degli affetti brevi Airbnb è conosciuta e utilizzata in tutto il mondo. Oggi si contano oltre 260 milioni di ospiti che sono stati in oltre 65.000 città, per quasi due milioni di appartamenti presenti sul sito. Il modello di business rientra nell'ambito della sharing economy, in cui non è più importante il possesso, bensì l'accesso ad un determinato bene o servizio.

Nel presente lavoro di tesi si svolgeranno diverse analisi per individuare eventuali differenze di genere sia lato guest che lato host. Il focus principale sarà quello di verificare quali sono le parole più utilizzate nelle recensioni rilasciate sulla piattaforma per capire se queste sono differenti per guest di sesso femminile o di sesso maschile. Inoltre, si vuole indagare se esiste una tendenza delle guest donna a soggiornare presso listing di host del medesimo sesso.

La prima parte dell'elaborato si concentrerà sulla descrizione della realtà in analisi: dopo una breve introduzione sull'economia della condivisione, verranno analizzati i driver di diffusione, vantaggi e svantaggi (con focus sull'impatto ambientale), il suo valore di mercato e si analizzerà l'effetto della pandemia di Covid-19 in questa realtà dove la condivisione è il punto cardine del modello di business.

Nel secondo capitolo il focus si sposterà sulle piattaforme digitali, sui modelli di pricing e sugli effetti di rete generati, fondamentali per il successo di Airbnb e di tutte le altre piattaforme di condivisione. Infatti, il motore delle piattaforme digitali sono le persone che ne fanno parte, le quali alimentano la loro continua crescita.

Nel terzo capitolo verranno presentati i principali aspetti della regolazione di questo settore in continua evoluzione, le decisioni già in atto e le sfide da affrontare per riuscire a regolare la sharing economy.

Successivamente l'analisi si sposterà sulla piattaforma Airbnb, sul suo impatto nel settore turistico con conseguenze non solo sugli hotel ma anche su competitors come Booking.

Una risorsa fondamentale per il successo di Airbnb è il sistema di review. Diversi studi si sono concentrati sull'analisi delle recensioni e in un capitolo dedicato saranno riassunti alcuni paper su questo argomento.

A seguire, si introdurrà il tema della differenza di genere: gli uomini e le donne elaborano le informazioni ed agiscono in modo diverso. Verranno presentati due paper in cui si studia il diverso comportamento dei due sessi nei confronti di un sistema di review e se le donne guest preferiscono rivolgersi a donne host, obiettivo di questo elaborato.

Per poter svolgere l'analisi si è costruito un database scaricando i dati da Insider.Airbnb. Saranno presentate le variabili utilizzate e verrà svolta un'analisi descrittiva su di esse, sempre con focus sulla differenza tra uomini e donne. Inoltre, saranno individuate le principali parole utilizzate nelle review e si studierà se queste differiscono tra guest uomini e guest donne

Infine, saranno svolte tre differenti analisi di regressioni. Le prime due valuteranno l'influenza di alcune variabili rispettivamente sul prezzo e sul numero di review, utilizzato come proxy della domanda. Si indagherà se il sesso dell'host e le parole più utilizzate nelle review influenzano le due variabili dipendenti in analisi. La terza retta di regressione permetterà di verificare se esiste una tendenza delle donne guest a rivolgersi a donne host.

1. Sharing Economy

1.1 Definizione

“Sharing economy: an economic system that is based on people sharing possessions, product and services, either for free or for payment, usually using the internet to organize it.”

Cambridge Dictionary

“L’espressione sharing economy si riferisce ai modelli imprenditoriali in cui le attività sono facilitate da piattaforme di collaborazione che creano un mercato aperto per l’uso temporaneo di beni o servizi. L’economia collaborativa coinvolge tre categorie di soggetti: i) i prestatori di servizi che condividono beni, risorse, tempo e/o competenze, ii) gli utenti di tali servizi e iii) gli intermediari che mettono in comunicazione i prestatori e utenti”

Commissione Europea

Negli ultimi anni si è sentito sempre di più parlare di sharing economy. I modi per definire questa nuova tipologia di business sono molteplici. Una parola chiave per descriverla risulta inevitabilmente essere “condivisione”, come indica il nome stesso. La sharing economy infatti prevede l’allocazione condivisa delle risorse di spazio, tempo, beni e servizi operata tramite piattaforme digitali. Nel 2011 il New York Times individua l’economia condivisa come una delle dieci idee che cambieranno il mondo. La sharing economy ha infatti contribuito ad un profondo cambio di paradigma nelle modalità di godimento dei beni, in cui l’uso è più importante dell’averne. Si pensi infatti al mondo delle automobili: venti anni fa era impossibile pensare di riuscire a muoversi in auto senza possedere un mezzo proprio. Oggi, invece, è possibile grazie ai servizi di car sharing. L’accesso ad un determinato bene o servizio (in questo caso si parla del servizio di mobilità) è più importante del possesso.

In realtà la sharing economy non riguarda solamente la condivisione di risorse o servizi, ma prevede anche la collaborazione: più persone utilizzano la rete con l'obiettivo di realizzare una transazione da cui ognuno trarrà un beneficio anche individuale. Pertanto, le relazioni risultano essere di tipo cooperativo.

Nonostante le differenze esistenti nella letteratura di riferimento e nonostante i diversi ambiti in cui opera la sharing economy, si possono individuare degli elementi in comune tra i principali player di questo mercato, primo fra tutti la presenza di una piattaforma tecnologica, dove hanno origine le relazioni digitali. Su tali piattaforme le persone offrono beni di consumo come automobili, case, posti auto, abbigliamento, ecc. Le aziende come Airbnb in questo contesto non erogano servizi o prodotti dall'alto verso il basso ai propri clienti, ma abilitano l'accesso ai beni attraverso queste piattaforme, non solo mettendo direttamente in contatto chi cerca con chi offre, ma anche diventando garante della qualità grazie ai sistemi di recensioni, trasmettendo così fiducia agli utenti. Gli asset che generano valore per le piattaforme appartengono alle persone e non all'azienda, come avviene invece per le imprese classiche. Mentre nell'economia tradizionale il valore è creato dal prodotto o servizio, nella sharing economy il valore è ottenuto dall'interazione tra gli utenti, grazie alla creazione di un network.

Inoltre, gli stessi attori iscritti alla piattaforma possono scambiarsi i ruoli, proponendosi in alcuni casi come chi offre e in altri come chi cerca. Per esempio, su Airbnb un host che ha registrato uno o più appartamenti sulla piattaforma può tranquillamente essere un guest in un'altra situazione.

Un ulteriore aspetto in comune nel mondo della sharing economy è la condivisione da parte degli utenti di beni che non consumano sempre e forniscono loro una capacità in eccesso. Le persone hanno quindi la possibilità di concedere l'accesso temporaneo ai loro beni fisici sottoutilizzati, senza alcun investimento aggiuntivo ma anzi valorizzando questi asset per generare reddito extra.

Infine, questo modello di business è caratterizzato da rapporti orizzontali fra pari, grazie ai quali le persone entrano in relazione tra di loro. Tuttavia, negli ultimi anni

di vita del fenomeno, il fornitore del bene o servizio non è più un utente, ma un professionista. A tal proposito la Commissione Europea recentemente ha evidenziato come si possa parlare di sharing economy anche quando il fornitore del servizio sia un soggetto professionale. I prestatori di servizi possono essere pertanto sia privati che operano su base occasionale (“pari”) sia utenti con capacità professionali (prestatori di servizi professionali).

1.2 Driver della sharing economy

La sharing economy, come già sottolineato, ha subito una forte crescita: diversi cambiamenti economici e sociali hanno contribuito ad una diffusione così rapida di questo modello economico.

- Diffusione delle piattaforme digitali, le quali permettono, grazie all’utilizzo di internet, l’incontro tra utilizzatore che richiede un bene o un servizio e un altro che è in grado di fornirlo e condividerlo. Si è verificata una forte crescita del numero e della redditività delle aziende che cercano di costruire piattaforme, piuttosto che offrire direttamente prodotti o servizi.
- Diffusione dei device: la maggior parte della popolazione possiede uno smartphone, un PC o un tablet attraverso cui poter accedere alla rete internet. Secondo l’ultimo report annuale Digital 2020, in Italia sono quasi 50 milioni le persone online ogni giorno. Gli smartphone posseduti dagli utenti risultano essere pari a 80 milioni, un numero molto alto se pensiamo che la popolazione residente nel nostro paese è circa 60 milioni. Considerando invece le statistiche in tutto il mondo, sono quasi 4,54 miliardi le persone connesse a Internet. Tuttavia, nonostante nel 2019 si siano registrati circa 300 milioni di utenti che hanno avuto accesso a internet per la prima volta, ancora oggi più del 40% della popolazione mondiale non può navigare online. La penetrazione dei device ha sicuramente quindi raggiunto l’*early majority*, cioè quasi la maggioranza della popolazione, ma il potenziale di crescita è ancora alto, soprattutto per i paesi più poveri.

- Volontà di utilizzare le risorse in modo più efficiente da parte dei cittadini. Rachel Botsman nel suo libro *“What’s mine is yours, the rise of collaborative economy”*, pubblicato nel 2010, è stata la prima a teorizzare che gli oggetti non utilizzati possono essere considerati un vero e proprio asset da sfruttare. Le stanze vuote, le case non abitate, l’auto, sono tutti elementi che hanno un valore per le persone, ma in un determinato momento della vita il proprietario potrebbe non sfruttarli in modo efficiente. Grazie alla condivisione si riesce ad ottenere invece il massimo rendimento da questa capacità in eccesso o inattiva.
- Nuovi bisogni dei consumatori, non più interessati al possesso e sempre più attenti ai temi ambientali. È soprattutto la fascia d’età dei più giovani che si mostra particolarmente sensibile al tema della sostenibilità ambientale. Il sondaggio dell’Osservatorio Giovani evidenzia infatti come la grande maggioranza dei giovani si dichiara sensibile e attenta all’ambiente, mentre i disinteressati sono meno del 15%. L’interesse crescente in ambito *green* da parte della popolazione più giovane, la quale sfrutta attivamente i servizi della sharing economy, comporta un maggiore utilizzo di queste piattaforme collaborative. Quest’ultime, infatti, hanno registrato impatti positivi sulla tematica della sostenibilità ambientale e sulla riduzione degli sprechi: chi le utilizza quindi contribuisce, anche solo in minima parte, ad aiutare il nostro pianeta. Questo aspetto verrà approfondito successivamente in un paragrafo dedicato.

1.3 Vantaggi e svantaggi

La sharing economy è riuscita a diffondersi velocemente, generando vantaggi in termini di prezzi, costi e qualità. I servizi, infatti, sono offerti a prezzo basso, rendendoli attraenti agli occhi degli utenti. Questa è una diretta conseguenza di costi più bassi: le informazioni raccolte sugli utenti servono per collegare domanda e offerta, generando maggiore utilità per gli utenti e un utilizzo efficiente delle risorse. Inoltre, per molti clienti i beni condivisi sono caratterizzati da un livello di qualità superiore rispetto alle alternative tradizionali. L’accesso ai servizi è possibile

grazie a piattaforme che innalzano la qualità del servizio inserendo funzionalità innovative, semplici metodi di pagamento e con la possibilità di personalizzare la ricerca. Un vantaggio delle piattaforme è inoltre la riduzione dei costi di transazione, cioè tutti quegli elementi che rallentano, rendono più costosa o addirittura impediscono una transazione. Un esempio di costo di transazione è rappresentato dai *searching cost*, ovvero i costi sostenuti per ricercare informazioni. Le piattaforme tramite un sistema di recensioni e punteggi riducono le asimmetrie informative e limitano i costi per ottenere informazioni.

Ulteriori benefici generati dall'economia della condivisione sono riconducibili ai guadagni extra derivanti dall'offerta dei servizi e ad un orario di lavoro più flessibile, permettendo ai lavoratori di impostare le proprie ore di lavoro. Per esempio, i lavoratori come i piloti di Uber sono in grado di accettare o rifiutare una corsa in auto in base ai loro bisogni.

Altroconsumo, associazione di difesa dei consumatori, al fine di comprendere i vantaggi della sharing economy attesi dagli utenti, ha condotto nel 2015 un'indagine statistica su un campione di 8.679 consumatori europei (2.336 in Italia) per conoscere la loro esperienza. In primo luogo, si sono analizzate le percentuali di utenti che hanno utilizzato i servizi per i settori del trasporto, degli affitti e dei beni usati di seconda mano (*Figura 1.3.1*).

Sono molteplici le motivazioni per cui gli utenti usufruiscono dei servizi di sharing economy: la principale è la volontà di risparmiare e questo risultato emerge per tutti i tre settori analizzati. Seguono motivi pratici, tra cui flessibilità, comfort, aiutare persone con beni poco utilizzati, dismettere beni di cui non si ha più bisogno e la facilità d'uso. Più basse invece sono le percentuali di chi ha citato obiettivi idealistici, per esempio favorire la collaborazione tra persone e proteggere l'ambiente. Anche la percentuale di coloro che hanno deciso di accedere ai servizi di sharing economy per guadagnare è relativamente piccola.

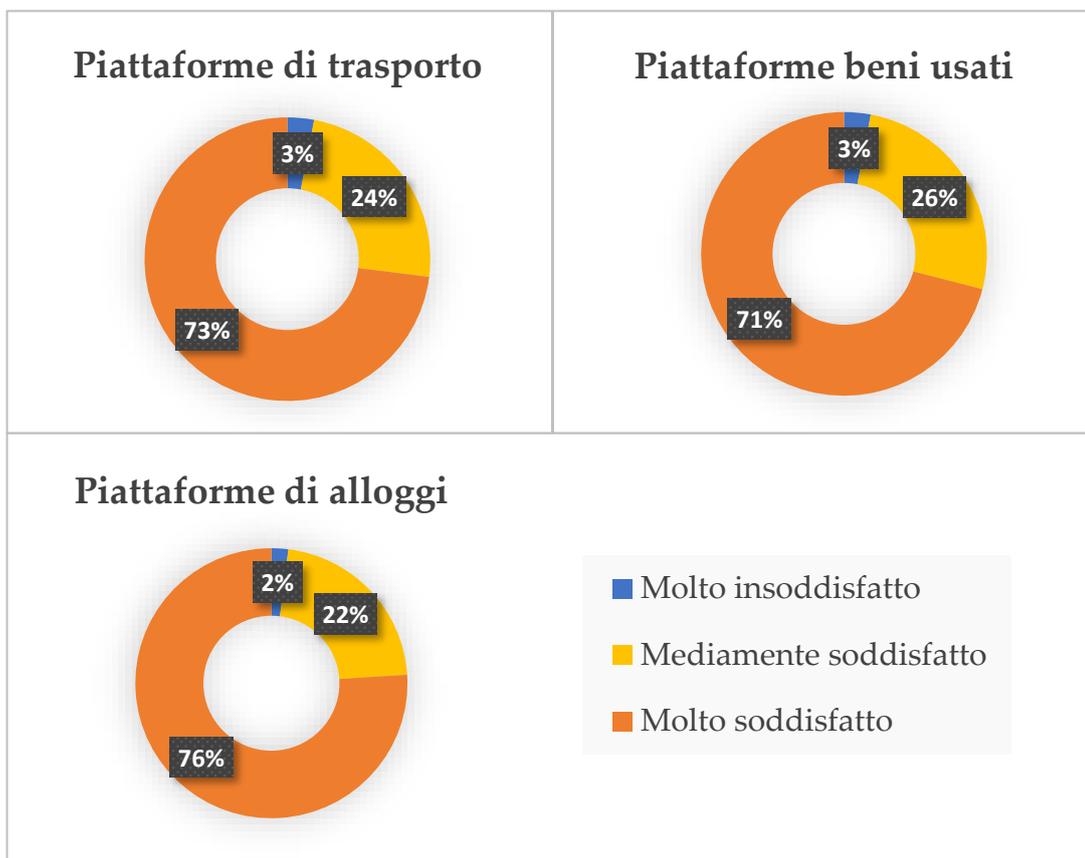


Figura 1.3.1 Soddisfazione utenti piattaforme sharing economy

Analizzando invece le barriere alla partecipazione, la principale si riconduce alla mancanza di necessità (63%) e alla mancanza di informazioni (41%). Solo un quarto di chi non ha mai partecipato ha citato come motivazione la mancanza di fiducia nelle persone.

Ogni modello di business comporta inevitabilmente anche degli svantaggi. La sharing economy, offrendo i propri servizi attraverso piattaforme e sfruttando le esternalità di rete, può rinforzare la concentrazione di alcune società favorendo il potere di mercato di pochi soggetti che possono così fissare un prezzo più alto rispetto a quello di first best, ovvero pari al costo marginale. È molto importante anche la tematica della privacy dei dati raccolti sugli utenti che partecipano all'economia della condivisione. Le piattaforme, con le informazioni acquisite sui propri consumatori, potrebbero trattare i dati per ragioni sconosciute agli utenti, venderli a terze parti o permettere ai produttori di attivare pratiche di discriminazione di prezzo, appropriandosi parte del surplus del consumatore.

1.4 Effetti ambientali

L'economia della condivisione sembra generare impatti positivi sull'ambiente grazie ad una riduzione delle risorse richieste e ad un loro utilizzo più efficiente, a minori inquinanti rilasciati, minori sprechi e minori emissioni di CO₂. A supporto di questa tesi vengono elencati di seguito due importanti risultati ottenuti in termini di sostenibilità ambientale.

La piattaforma di condivisione domestica Airbnb ha impattato positivamente sull'ambiente. Gli ospiti di Airbnb hanno consumato meno energia (dal 63% al 78% in meno), meno acqua (dal 12% al 48% in meno) e hanno generato meno emissioni di gas a effetto serra. Nel settore dei trasporti invece, il comportamento di condivisione dei veicoli può avere un impatto ambientale positivo stimolando cambiamenti di lunga durata nel comportamento dei consumatori. Inoltre, i sistemi di condivisione delle biciclette possono ridurre l'uso di auto. A Shanghai infatti, nel 2016, grazie al servizio di bike sharing, si sono ridotte le emissioni di CO₂ e ossido di azoto rispettivamente di 25.000 tonnellate e 64.000 tonnellate. La piattaforma di car pooling Blablacar ha invece dichiarato che, grazie ai loro servizi, sono state risparmiate in tutto il mondo nel solo 2015 circa 500.000 tonnellate di CO₂.

Nonostante questi esempi a favore della sostenibilità ambientale della sharing economy, vi sono ancora dibattiti e controversie sull'effettivo impatto positivo.

Il paper *"Sharing Economy: Promote Its Potential to Sustainability by Regulation"*, di Xufeng Liu and Hongmin Chen, pubblicato il 27 gennaio 2020 dal *Department of Environmental Science and Engineering, Fudan University* di Shanghai, indaga i fattori positivi e limitanti caratteristici della sharing economy per uno sviluppo sostenibile. In un secondo momento vengono esplorate nel paper le condizioni per promuovere un impatto ambientale positivo dell'economia della condivisione.

I fattori positivi principali riscontrati in questo modello economico sono i seguenti:

- Aumento dell'efficienza dell'utilizzo delle risorse: i prodotti e i servizi condivisi aiutano a ridurre la produzione e favoriscono un'allocatione delle risorse e un loro utilizzo più efficiente.
- Promozione del cambiamento sulle idee di consumo: con la sharing economy la domanda si sposta dal prodotto in sé al servizio che quel prodotto fornisce, favorendo la sostenibilità lungo tutta la catena di produzione. I consumatori aumentano la loro utilità usufruendo dei servizi derivanti dalla condivisione dei prodotti piuttosto che soddisfare i loro bisogni comprando continuamente nuovi beni. Con questo comportamento si riduce un eccessivo consumo e quindi un'eccessiva produzione.

Gli autori del paper però individuano anche dei fattori limitanti della sharing economy, che rendono ancora incerto il suo effettivo impatto positivo sull'ambiente.

- *"Rebound effect"* causato da una domanda eccessiva. La sharing economy infatti riduce i prezzi dei beni e questo spinge i consumatori ad utilizzare tali prodotti o servizi non per i loro effettivi bisogni, ma spinti dai bassi prezzi. Tutto questo si traduce in un consumo eccessivo e un impatto negativo sull'ambiente (più beni prodotti, più inquinamento). In riferimento al settore dei viaggi, a causa dei prezzi bassi molti utenti potrebbero per esempio aumentare sia la frequenza sia la durata dei viaggi, con un conseguente impatto negativo sull'ambiente.
- Utilizzo dei beni condivisi senza cura: questo comporta un ciclo di vita più breve per i prodotti. Per riuscire a soddisfare la domanda, se i beni condivisi si deteriorano velocemente, bisognerà accelerare la produzione di nuovi prodotti. L'aumento del consumo di risorse che ne deriva può determinare o addirittura superare i benefici ambientali generati dalla condivisione.
- Relazione tra l'aumento del benessere generale per la società e l'aumento del livello di consumo delle risorse quando l'economia della condivisione aumenta la domanda. Se gli utenti domandano di più, questo si traduce in un aumento del *welfare* ma, allo stesso tempo, in un aumento del consumo di risorse e in maggiori emissioni. Se il consumo di risorse marginale cresce più del beneficio

marginale che si ottiene, i vantaggi ambientali risultanti dalla sharing economy si riducono gradualmente.

Dal momento che i settori in cui opera la sharing economy sono estremamente diversi tra loro, per la varietà di prodotti e servizi offerti, è difficile stabilire una metrica generale per capire se la sharing economy genera dei benefici all'ambiente. Tuttavia, è possibile stabilire che i beni condivisi, indipendentemente dal settore, devono possedere tre caratteristiche: durabilità, compatibilità e rinnovabilità. La prima prevede che i prodotti condivisi siano di alta qualità e con un lungo ciclo di vita per durare più anni, favorendo un uso intensivo di essi. Per compatibilità invece si intende che gli item devono essere compatibili con le diverse piattaforme di condivisione. Infine, i prodotti devono essere rinnovabili, ovvero costituiti da materiali riciclabili.

Dopo aver individuato le caratteristiche principali dei beni condivisi, gli autori suggeriscono alcuni interventi regolatori per favorire uno sviluppo sostenibile della sharing economy. Mentre nell'economia tradizionale è importante regolare i produttori di beni, nell'economia della condivisione tutte le entità del sistema devono essere regolate. La *Tabella 1.4.1* riassume i principi generali per regolare i produttori, i consumatori, le piattaforme e fornisce le linee guida per i regolatori.

Tabella 1.4.1 Principi generali per regolare entità della sharing economy

Produttori	Consumatori
<ul style="list-style-type: none"> • Produrre prodotti condivisi che siano durevoli, rinnovabili e compatibili. • Applicare standard tecnici per i componenti dei prodotti. • Incorporare principi <i>environmentally friendly</i> e sostenibili nella propria strategia. 	<ul style="list-style-type: none"> • Senso di responsabilità utilizzando con cura i beni della sharing economy. • Consumo basato sulla reale domanda invece di un eccessivo consumo a causa di prezzi bassi.
Piattaforme	Regolatori governativi
<ul style="list-style-type: none"> • Aumentare il costo derivato da danneggiamenti del prodotto da parte dei consumatori. • Assumere responsabilità di manutenzione quotidiana. 	<ul style="list-style-type: none"> • Cambiare il pensiero regolatorio. • Espandere la regolazione seguendo l'espansione della sharing economy. • Cambiare la regolazione applicata nella prima fase della sharing economy.

Un ulteriore paper si pone l'obiettivo di capire se la sharing economy garantisce sostenibilità ambientale: *"Environmental impacts and potential of the sharing economy"* di John Magne Skjelvik, Anne Maren Erlandsen and Oscar Haavardsholm, pubblicato nel 2017. Gli autori inizialmente si sono concentrati sulla sharing economy e sui vantaggi da essa generati in termini di maggiore qualità e minori costi. L'analisi è stata eseguita isolando i due effetti di costi e qualità, per poi poterli studiare insieme. Partendo dall'effetto dei costi, si assume la qualità dei servizi della sharing economy pari a quella dei servizi tradizionali. In questo modello economico, offrendo il servizio a prezzo inferiore (poiché i costi sono più bassi) la curva dell'offerta si sposta verso il basso, generando un numero maggiore di transazioni (*Figura 1.4.1*). I nuovi fornitori di servizi cattureranno una parte del mercato dell'incumbent (m) e genereranno nuove attività (n).

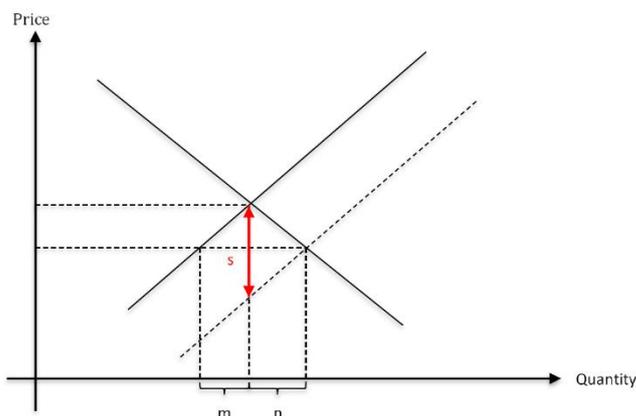


Figura 1.4.1 Effetto costi della sharing economy

Nel secondo caso il prezzo è pari a quello dell'economia tradizionale ma la qualità è più alta. La curva della domanda si sposta verso l'alto poiché gli utenti sono disposti a pagare di più per la maggior qualità offerta (*Figura 1.4.2*). Il prezzo in questo contesto cresce all'aumentare della quantità di servizi venduti. Anche per questo effetto la sharing economy aumenta il mercato (n) e cattura una parte del business dell'incumbent (m).

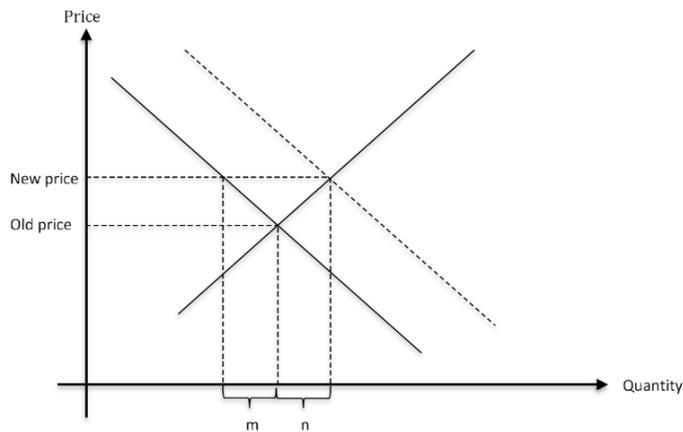


Figura 1.4.2 Effetto qualità della sharing economy

Combinando i due effetti (Figura 1.4.3) si osserva che il prezzo diminuisce leggermente, mentre la quantità invece aumenta. I guadagni della sharing economy sono uguali all'area di A+B.

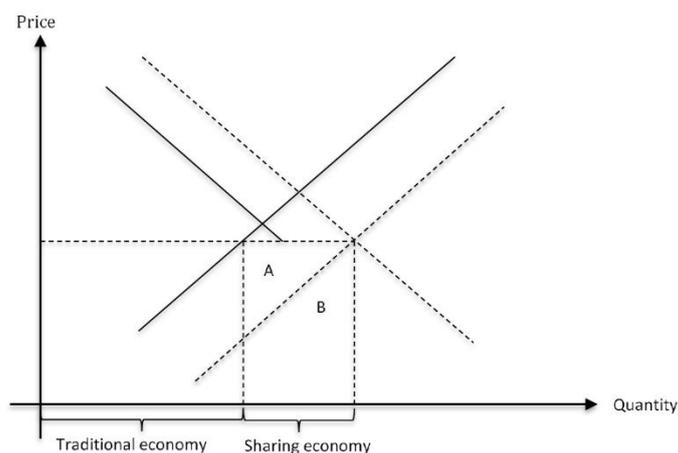


Figura 1.4.3 Effetto totale della sharing economy

Il paper successivamente si concentra sugli effetti della sharing economy analizzando quattro segmenti: trasporti, accomodation, piccoli beni (strumenti che si possono prestare ed affittare) e servizi.

Nel mercato dei Paesi del Nord sono diverse le iniziative adottate per migliorare la condizione ambientale. Le principali misure riguardano il settore dei trasporti, in cui il potenziale di riduzione delle emissioni è molto alto. Con i servizi di car sharing non diminuisce esclusivamente la CO₂ rilasciata nell'aria, ma anche la produzione di auto, il rumore, la congestione del traffico, ecc. Diversi studi hanno confermato che una macchina condivisa può sostituire da 4 fino a 13 auto private. Tutto questo

si traduce in una riduzione di CO₂ che varia da 130 kg fino a 1000 kg per famiglia per anno. Inoltre, molti servizi di car sharing sfruttano macchine ibride o elettriche, riducendo ulteriormente le emissioni. Tuttavia, come nel paper precedentemente citato, anche in questo contesto viene valutato il “*Rebound effect*”: la riduzione dei prezzi dei servizi spinge gli utenti a spendere i risparmi per utilizzare quel servizio oltre i reali bisogni o in altre attività che producono inquinamento. La diretta conseguenza di questo effetto è la riduzione dei benefici generati dal car sharing.

Nel settore degli affitti, piattaforme come Airbnb generano meno emissioni degli hotel. Gli appartamenti, infatti, a differenza degli hotel, non devono sprecare energia in servizi come bar, ristoranti, piscine, reception 24/24. Inoltre, nel lungo periodo gli affitti degli appartamenti riducono la domanda per gli hotel e, di conseguenza, le emissioni per la costruzione di nuove strutture. Tuttavia, i risparmi ottenuti possono essere utilizzati ad esempio per viaggiare in aereo, aumentando le emissioni di CO₂ (*Rebound effect*). Il settore dei piccoli beni suggerisce come la condivisione possa ridurre le emissioni. Prendendo l’esempio di un trapano, questo strumento viene utilizzato circa 18 minuti nel suo ciclo di vita. Le emissioni di CO₂ ammontano a circa 28 kg e solo il 2% sono generate dall’effettivo uso. La sharing economy quindi ridurrebbe le emissioni. Infine, i servizi condivisi come pulizia, giardinaggio, ecc. hanno meno potenziale per la riduzione delle emissioni. Il paper si conclude sottolineando l’importanza degli interventi regolatori per sfruttare al massimo il potenziale *environmentally friendly* della sharing economy.

In seguito alla lettura di questi due paper si può concludere che maggiori dati empirici sarebbero necessari per comprendere l’effetto complessivo dell’economia collaborativa sull’ambiente.

1.5 Valore di mercato

L'azienda di consulenza PwC ha stimato che, nel 2013, l'economia della condivisione valesse 15 miliardi di dollari, contro i 240 miliardi di dollari dell'economia tradizionale. È stato previsto invece che entro il 2025 il valore della sharing economy raggiungerà 335 miliardi di dollari (Figura 1.5.1). Se prima quindi la sharing economy rappresentava solo il 6% del valore di mercato, arriverà a valere circa il 50% secondo queste stime. Le opportunità di entrate previste ammontano a circa 670 milioni di dollari nel 2025.

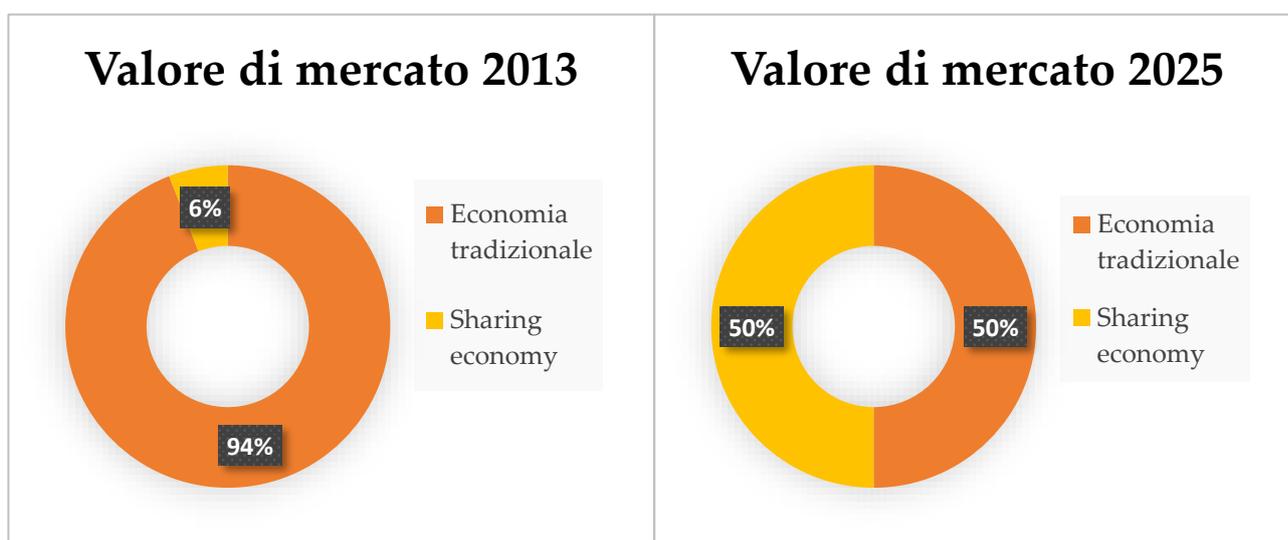


Figura 1.5.1 Confronto valore di mercato 2013 vs 2025

In Italia invece, a seconda degli scenari ipotizzati, si stima il contributo dell'economia della condivisione di un valore che oscilla tra 8,8 e 10,5 miliardi di euro entro il 2020, e tra 14,1 e 25,2 miliardi di euro entro il 2025.

Il servizio di sharing economy che si stima avere il tasso di crescita più elevato è rappresentato dalle piattaforme di condivisione dei servizi di trasporto (car sharing come Uber) che rappresenteranno nel 2025 il 40% del mercato (Figura 1.5.2). Il settore degli affitti tra pari, rappresentato da una piattaforma come Airbnb, entro il 2025 verrà superato dai servizi domestici su richiesta.

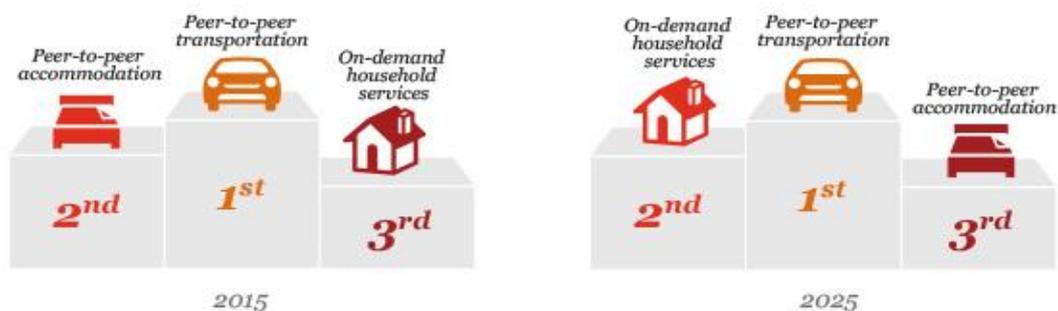


Figura 1.5.2 Servizi sharing economy

Un ruolo chiave per la crescita dei servizi di sharing economy è rappresentato dai Millennial: i giovani che appartengono ad una fascia d'età compresa tra i 18 e i 34 anni sono più propensi alla cultura della condivisione. Per questa parte della popolazione le piattaforme di condivisione riscontrano una minor resistenza nella loro diffusione. La sharing economy però, per potersi affermare come *digital disruption*, necessita di una penetrazione in tutte le fasce d'età. Il termine digital disruption indica il momento in cui una nuova tecnologia origina un cambiamento per una determinata attività, modificando il modello di business precedente. Una tecnologia disruptive ha più alto potenziale per un certo grado di sforzo e nel lungo periodo è in grado di eliminare la tecnologia incumbent.

Nel nostro Paese il coinvolgimento degli over 34 nella sharing economy porterebbe ad un mercato con valore tra lo 0,7% e l'1,3% del PIL nel 2025. È importante quindi riuscire a coinvolgere più utenti possibili per riuscire ad ottenere dei vantaggi in termini economici.

1.6 Impatto Covid-19

Nel mese di dicembre i telegiornali di tutto il mondo hanno iniziato a raccontare di un virus comparso a Wuhan, metropoli da 11 milioni di abitanti, capoluogo della provincia di Hubei, in Cina. Poche erano le informazioni possedute riguardanti questa particolare polmonite, la cui causa risultava sconosciuta: il virus non corrispondeva a nessun altro noto. L'11 gennaio è confermata la prima vittima in Cina e il 13 gennaio il primo decesso fuori confine, in Thailandia. Da quel momento la preoccupazione da parte della popolazione mondiale è aumentata, fin al 21 e il 22 febbraio, quando si registrano i primi contagi in Italia legati al virus chiamato

Covid19. In queste date l'emergenza ha raggiunto purtroppo anche il nostro Paese, con i focolai maggiori nel Lodigiano e in Veneto. Le misure attuate sono state sorprendenti: il 9 marzo l'Italia diventa zona rossa, limitando gli spostamenti solo per motivi di salute, lavoro e/o necessità. L'11 marzo l'OMS dichiara che quella di Sars-CoV-2 è una pandemia. In questi mesi di chiusura tutto si è fermato: le scuole, le aziende, gli spostamenti, i viaggi. Tutto ciò ha avuto un forte impatto sulla nostra vita e sull'economia, provocando effetti negativi anche per quelle imprese con una buona situazione finanziaria, come i grandi player della sharing economy. Il coronavirus ha avuto un impatto enorme su queste società già un mese dopo l'esplosione della pandemia. La prima conseguenza è stata il licenziamento: Uber ha ridotto il personale di 3.700 persone, il 14% della propria forza lavoro, Airbnb di 1.900, il 25%. Quest'ultima ha cominciato a perdere denaro a causa delle migliaia di cancellazioni da parte degli utenti e delle spese da sostenere per i rimborsi. Il cofondatore e amministratore delegato di Airbnb Brian Chesky si è sentito *"come il capitano di una nave colpita da un siluro"* durante questi mesi di restrizioni. Se i piani del CeO erano quelli di quotare la società in borsa verso il periodo di aprile-maggio, proprio durante quei mesi ha dovuto richiedere prestiti per due miliardi di dollari. Si osserva inoltre una riduzione degli alloggi attivi sul sito Airbnb durante i primi mesi di questa pandemia. (Figura 1.6.1).

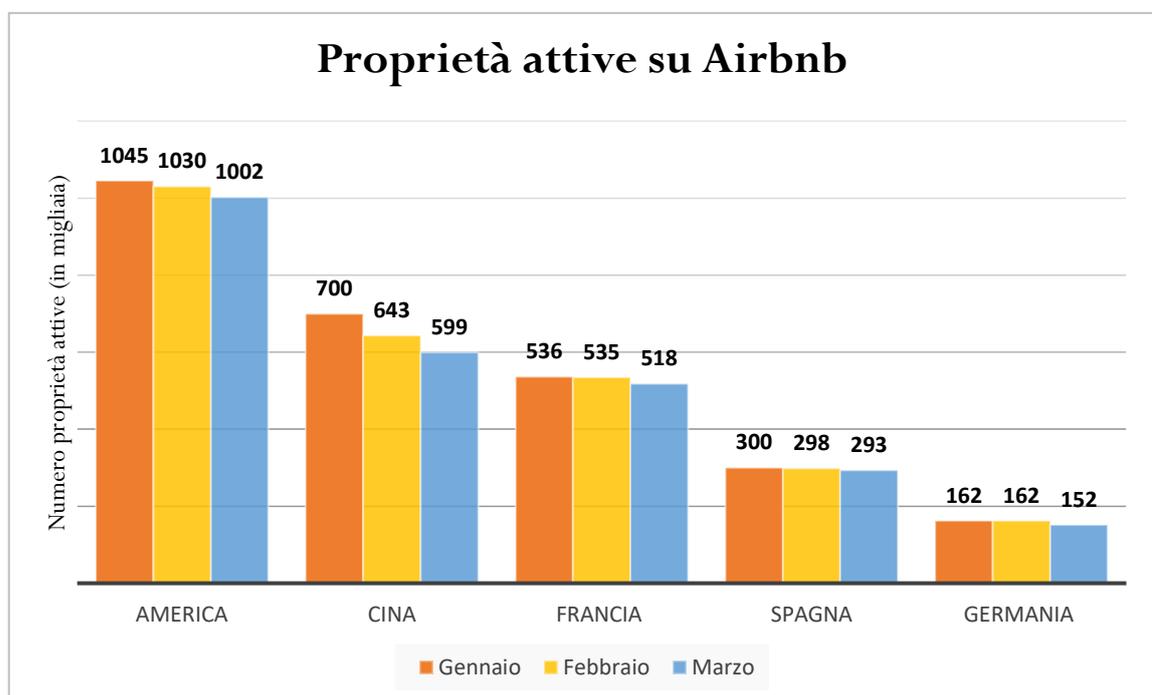


Figura 1.6.1 Alloggi attivi su Airbnb durante la pandemia

Uber invece ad aprile ha osservato una drastica riduzione del numero di corse: quasi l'80% in meno rispetto allo stesso periodo dell'anno precedente.

Le aziende si sono immediatamente attivate per riuscire a superare questo periodo di crisi, cercando soluzioni per offrire servizi di sharing economy garantendo agli utilizzatori la massima sicurezza. Airbnb ha deciso di istruire gli host su come pulire le stanze e gli alloggi affittati ed ha introdotto un periodo di attesa di 24 ore tra un soggiorno e l'altro per permettere di sanificare l'ambiente. Uber invece verifica che gli autisti indossino la mascherina attraverso una tecnologia automatica che analizza i selfie.

In questo contesto è importante inoltre considerare come sono cambiate le abitudini delle persone, con l'obiettivo di attuare la giusta strategia per riuscire a sopravvivere nel mercato. Nel settore del turismo gli utenti, anziché effettuare vacanze brevi verso grandi città, scelgono mete più vicine per periodi più lunghi. Airbnb, infatti, registra una durata media dei soggiorni più lunga e la quota di prenotazioni nazionali è più che raddoppiata, rappresentando circa l'80% del totale. Questo trend si conferma anche a livello globale: negli Stati Uniti, per esempio, molte persone sostengono di preferire le destinazioni raggiungibili in auto. I soggiorni a meno di trecento chilometri da casa, che rappresentavano circa il 33% delle prenotazioni, oggi costituiscono il 56%. Lo smart-working inoltre ha permesso alle persone di cambiare residenza per qualche giorno riuscendo comunque a rispettare gli impegni lavorativi. A seguito dello studio di questi dati il Ceo di Airbnb ha elaborato una nuova strategia chiamata "Go Near". L'obiettivo è quello di stipulare una serie di accordi con le associazioni locali di tutto il mondo per incentivare i viaggi vicino casa e la crescita economica. Grazie a queste collaborazioni sarà possibile raccogliere dati ed eseguire ricerche di viaggio per capire quali saranno le azioni più efficaci da intraprendere per sostenere il business; agli enti e alle autorità locali verrà inoltre fornita l'opportunità di utilizzare la piattaforma di Airbnb per raggiungere milioni di viaggiatori.

D'altro canto, Uber prevede invece una migrazione dai trasporti pubblici alle automobili: condividere un'auto per molte persone risulta una soluzione migliore rispetto ad autobus o treni affollati senza il rispetto del distanziamento sociale. Inoltre, la società nonostante abbia rinunciato al settore delle biciclette elettroniche, ha investito in Lime, principale rivale di Bird negli scooter elettrici, e vuole acquisire Grubhub per riuscire a rafforzare la sua presenza nel settore della consegna a domicilio di cibo, ambito in cui l'azienda ad oggi risulta in perdita.

In conclusione, la grave recessione provocata dalla crisi finanziaria del 2007-2009 aveva contribuito all'utilizzo della tecnologia per permettere di condividere oggetti e asset sottoutilizzati, generando una forte spinta per la nascita della sharing economy. Bisogna capire se la crescita dell'economia condivisa è destinata a continuare dopo questo periodo di crisi o se subirà un vero e proprio arresto. Sicuramente l'incertezza è ormai diventata l'elemento caratterizzante di ogni situazione. L'unica certezza ad oggi sono i dati della pandemia: le persone contagiate di Covid19 sono oltre 18 milioni, i morti nel mondo sono oltre 689 mila.

2. Piattaforme digitali

2.1 Ruolo delle piattaforme

La sharing economy si basa sulla condivisione di risorse. Questa pratica era comune anche prima dell'avvento di Internet: tra persone fidate, parenti, amici spesso i beni venivano prestati gratuitamente o in cambio di un minimo corrispettivo in denaro. Risultava invece insolito permettere ad estranei di sfruttare le proprie risorse, soprattutto per la mancanza di informazioni sull'affidabilità della persona. In questo contesto le piattaforme digitali hanno assunto un ruolo chiave per la diffusione della pratica di condivisione dei beni.

La Commissione Europea definisce le piattaforme online come:

“Un’impresa che opera in un mercato two-sided (o multi-sided), che utilizza internet per rendere possibili le transazioni tra due o più gruppi di user distinti e interdipendenti, per generare valore per almeno uno dei due gruppi”

Le piattaforme sono al centro della sharing economy: grazie ai sistemi di recensione permettono di stabilire fiducia tra le parti e tra estranei, assicurando gli utenti e incentivandoli a condividere asset sottoutilizzati. Le informazioni raccolte sono impiegate dal sistema che, attraverso un particolare algoritmo, sfrutta i dati sulle preferenze o sulla posizione geografica per abbinare gli utenti, aumentando così la produttività e l'efficienza.

L'infrastruttura tecnologica diventa il mezzo per facilitare la condivisione: nessuna parte deve effettuare nuovi investimenti fisici, ma con una semplice commissione pagata alla piattaforma si genera surplus attraverso l'interazione tra gli user.

In generale, le piattaforme sono diventate velocemente uno strumento necessario per la creazione di valore di molte imprese e in alcuni contesti sono l'elemento chiave su cui costruire una nuova strategia. Accenture, nota società di consulenza, ha indagato l'importanza delle piattaforme per il business di diversi settori

dell'economia e la maggior parte degli intervistati afferma che questo strumento è una risorsa *core* per la loro organizzazione. (Figura 2.1.1)

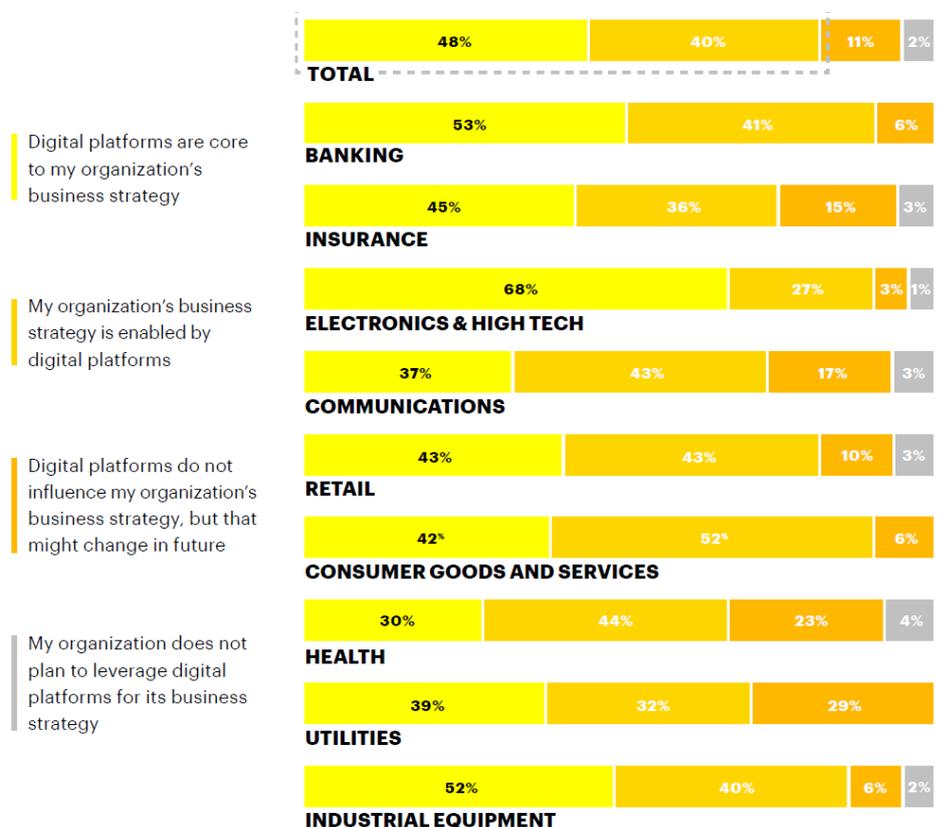


Figura 2.1.1 Ruolo piattaforme nell'organizzazione

Il valore creato dalla piattaforma cresce nel tempo all'aumentare del numero di dati raccolti e del numero di user che la utilizzano. Le imprese devono prestare attenzione ad adattare la loro strategia alla piattaforma digitale: si parla di *resources fit*, ovvero adeguare il business alla risorsa per ottenere un vantaggio competitivo. L'implementazione di questa tecnologia richiede inevitabilmente degli investimenti in denaro: sempre nel sondaggio condotto da Accenture, il 65% delle imprese in analisi dichiara di dedicare tra l'11 e il 20% del loro budget IT per le piattaforme digitali.

La società di consulenza indaga inoltre i benefici attesi dalle società: il 53% degli executives intervistati afferma che la piattaforma favorisce un nuovo modo per interagire con i clienti. General Electric è un buon esempio: l'impresa ha sviluppato

una piattaforma chiamata Predix, la quale raccoglie dati attraverso appositi sensori posizionati sui prodotti venduti per poter programmare azioni di manutenzione, per migliorare i processi, per ridurre i costi e rischi. Il 50% degli intervistati invece considera lo sviluppo di una piattaforma *as-a-business model*, come Uber e Airbnb, le quali basano la loro attività su questa infrastruttura tecnologica. Per il 44% le piattaforme rappresentano un nuovo modo per aumentare la *customer experience*: Amazon, per esempio, ha fornito nuove esperienze di shopping altamente convenienti, con la possibilità di personalizzare i prodotti mostrati al singolo utente in base alle informazioni raccolte sulle sue preferenze di consumo. Infine, il 44% considera le piattaforme uno strumento per lanciare un nuovo prodotto o servizio sul mercato.

Le società titolari delle piattaforme devono prendere la decisione di “*employ or enable*”: nel primo caso l’impresa è una delle due parti della relazione contrattuale, nel secondo caso l’azienda smette di essere parte dello scambio e si limita a consentirlo. Esistono comunque modelli ibridi: Amazon, per esempio, ha avviato il proprio business come fornitore di libri (*employ*) e successivamente ha ampliato la propria offerta offrendo ai retailers la possibilità di vendere i loro prodotti su Amazon (*enable-employ*). Un’altra scelta fondamentale è quella riguardante il livello di centralizzazione del potere decisionale. Se il potere decisionale è forte, le informazioni raccolte si riescono a sfruttare meglio, ma sono alti i costi da sostenere per assicurarsi che i venditori presenti sulla piattaforma applichino i prezzi indicati. Con la decentralizzazione si verifica la situazione opposta. Due grandi player della sharing economy come Uber e Airbnb hanno preso decisioni differenti: la piattaforma di car sharing si riserva il diritto di scegliere il prezzo che gli autisti devono applicare (centralizzazione); al contrario, Airbnb permette agli host di stabilire la tariffa per il soggiorno (decentralizzazione).

2.2 Esternalità

Le piattaforme sono caratterizzate da esternalità di rete. In generale, si possono individuare in un *two-sided market* quattro tipi di esternalità: due *same-side*, quando un utente aggiuntivo in un gruppo comporta dei benefici per i membri del gruppo stesso, due *cross-side*, quando l'aumento della numerosità di un gruppo ha delle conseguenze sull'utilità dell'altro gruppo. Le esternalità possono essere positive (all'aumentare della numerosità aumenta l'utilità) oppure negative (all'aumentare della numerosità diminuisce l'utilità). Prendendo come esempio i due player della sharing economy, si può affermare che Uber ha esternalità cross-side positive: un alto numero di consumatori sulla piattaforma genera valore per gli autisti perché aumenta la possibilità di ricevere chiamate, e viceversa un alto numero di driver aumenta l'utilità dei consumatori perché avranno più possibilità di trovare un autista disponibile al momento del bisogno. Per Airbnb il discorso è il medesimo: il valore per i proprietari di casa aumenta se è abbastanza grande il numero di potenziali affittuari; allo stesso modo il valore per gli utenti aumenta con il crescere del numero di proprietari di alloggi presenti sul sito.

La presenza di questi effetti di rete prevede una *business strategy* in cui si cerca di massimizzare il numero di agenti in entrambi i gruppi per generare il massimo surplus nel mercato. In un network è necessario quindi attrarre il maggior numero di utenti, soprattutto nelle prime fasi per riuscire a raggiungere la massa critica, ovvero il minimo livello di copertura che una tecnologia dovrebbe raggiungere per rimanere sul mercato. Airbnb, per aumentare la numerosità del suo network, ha sfruttato molto il social network Facebook per pubblicizzarsi, ha attirato l'attenzione da parte degli utenti grazie alla pubblicazione di foto degli appartamenti di alta qualità ed ha permesso di effettuare ricerche personalizzate inserendo filtri.

È molto importante considerare gli effetti di network per riuscire ad applicare la giusta politica di pricing: le società non fissano i prezzi esclusivamente basandosi

sull'elasticità della domanda e sui costi generati da ciascun gruppo, bensì internalizzando e tenendo conto degli effetti delle esternalità. Si considerino due gruppi (A e B) i quali interagiscono tra di loro attraverso una piattaforma. Se le esternalità *cross-side* generate dal gruppo A al gruppo B sono maggiori rispetto a quelle generate dal gruppo B al gruppo A, alla società converrà fissare un prezzo più basso (o addirittura nullo) agli utenti del gruppo A per riuscire ad attrarre più utenti possibili in questo gruppo e questo si rifletterà con un aumento della numerosità nel gruppo B, in cui la piattaforma andrà a monetizzare. Se, al contrario, si fissasse il prezzo di B più basso, sicuramente si attrarrebbero nuovi utenti nel gruppo A grazie agli effetti *cross-side*, ma la conseguenza sarebbe minore rispetto al caso precedente con il prezzo A più basso (questo perché esternalità $A \rightarrow B >$ esternalità $B \rightarrow A$). Facebook per la sua piattaforma utilizza questa strategia: per gli utenti l'accesso è nullo e la società monetizza con gli advertiser: gli effetti *cross-side* sono molto più alti dagli user agli advertiser rispetto al viceversa. Anzi, se si iniziasse ad applicare una tariffa per utilizzare il social network il numero di utenti diminuirebbe con conseguenze negative anche per il mercato degli advertiser. Basandosi su queste considerazioni è possibile individuare modelli di business comuni tra le piattaforme e verranno analizzati nel paragrafo successivo.

2.3 Modelli di pricing

Nella sharing economy il valore è generato usufruendo direttamente dei servizi offerti dalla piattaforma e dagli effetti di network precedentemente descritti. Le piattaforme possono catturare valore in diversi modi, talvolta anche combinando differenti modelli di business. Il primo metodo, il più diffuso, consiste nell'applicare una percentuale sulla transazione: la piattaforma guadagna attraverso una fee fissa o variabile ogniqualvolta un'interazione tra cliente e fornitore vada a buon fine. La piattaforma può monetizzare inoltre attraverso il pagamento di abbonamenti (settimanali, mensili, annuali) da parte degli utenti. Spesso le entrate provengono anche dalla presenza di inserzioni pubblicitarie sulla piattaforma. Infine, molte

società offrono i propri servizi base gratuitamente ai clienti finali mentre guadagnano dal pagamento di alcuni servizi avanzati.

Nell'articolo *"Navigating Peer-to-Peer Pricing in the Sharing Economy"* di Gemma Newlands, Christoph Lutz e Christian Fieseler, pubblicato a gennaio 2018 su SSRN Electronic Journal, sono analizzati quattro meccanismi di prezzo utilizzati nella sharing economy: *Provider-to-platform*, *Consumer-to-platform*, *Peer-exchange price* e *Platform commission*. In particolare, per descrivere al meglio il pricing, gli autori si sono concentrati principalmente sui settori del trasporto e dell'affitto.

Le transazioni peer-to-peer introducono diversi meccanismi indipendenti tra loro che possono avvenire contemporaneamente o in momenti diversi. Come si può osservare dalla *Figura 2.3.1*, ognuno dei quattro modelli di pricing rappresenta un flusso di denaro distinto.

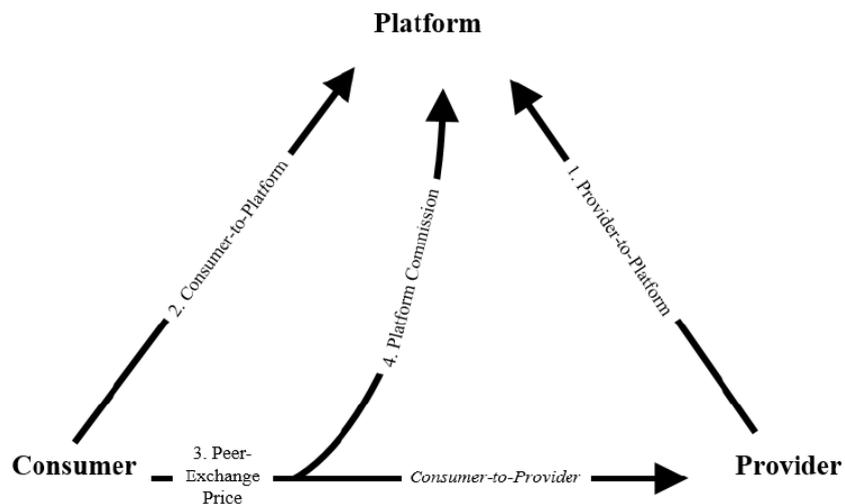


Figura 2.3.1 Flussi monetari piattaforma

Il primo meccanismo *"Provider-to-platform"* consiste in un costo diretto pagato dal fornitore per l'utilizzo della piattaforma. È prevalentemente usato nell'ambito home-sharing dove è necessario sottoscrivere una fee per avere accesso alla piattaforma. Tuttavia, la maggior parte delle piattaforme di condivisione non include questo meccanismo di prezzo, probabilmente per ridurre l'onere finanziario

per i provider e per abbassare le barriere all'ingresso. Tenzialmente questo meccanismo è utilizzato insieme agli altri (Figura 2.3.2).

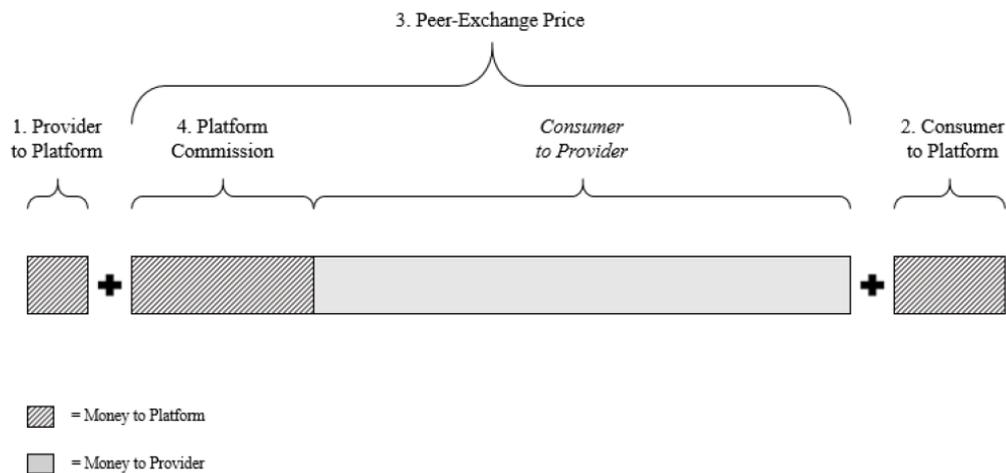


Figura 2.3.2 Quattro meccanismi di pricing

Il secondo meccanismo *“Consumer-to-platform”* rappresenta un costo diretto sostenuto dal consumatore per poter usufruire della piattaforma. Viene pagato direttamente dall’utente e nessuna percentuale di questa somma è trasferita al fornitore; la quota può essere variabile o fissa. Nel settore dei trasporti questo meccanismo è applicato da Uber: per ogni tratta l’utente deve pagare una *“Booking Fee”*, variabile a seconda della tratta. La società definisce la tassa nel seguente modo: *“The booking fee is a separate flat fee added to every trip that helps support safety initiatives for riders and drivers as well as other operational costs”* (Uber, 2017). Anche Airbnb tassa gli utenti con la *“Guest Service Fee”*, così spiegata *“Guest Fee charged for all reservations to help Airbnb run smoothly and offer 24/7 customer support. It is used to help cover the costs of running Airbnb”*.

Il terzo meccanismo *“Peer-exchange price”* è, in generale, un costo variabile determinato dalla durata dell’affitto e dalla qualità dell’item. Il prezzo può essere definito dal fornitore, dalla piattaforma o può essere una decisione condivisa.

Infine, la *“Platform commission”* è una proporzione del *“Peer-exchange price”* catturata dalla piattaforma. La somma è espressa tendenzialmente come una percentuale variabile. Uber, per esempio, applica una commissione del 25% con il nome di *“Uber*

Fee". In generale, la percentuale di commissione del settore dei trasporti è più elevata rispetto a quella di altri settori della sharing economy. Su Airbnb, viene dedotta una "Host service fee" pari al 3-5%.

L'articolo si focalizza anche sulla discriminazione di prezzo attuata dalle piattaforme. Lato consumatore, applicare una "*Consumer-to-platform*" diversa per gli individui è una forma di discriminazione. Nel paper viene analizzata la pratica del pricing dinamico: l'impresa stabilisce prezzi differenti per uno stesso item sulla base dello stato della domanda in quel preciso momento. È una strategia spesso utilizzata nel settore turistico (con prezzi diversificati per alta e bassa stagione) ma caratterizza molte transazioni che avvengono sulle piattaforme della sharing economy. Un esempio è costituito dal sistema di tariffe utilizzato da Uber per cui la stessa corsa può costare prezzi molto differenti a seconda della data o dell'orario.

Le piattaforme possono applicare una discriminazione di prezzo anche lato fornitore. Spesso infatti vengono applicate delle promozioni, rendendo altamente variabili le commissioni pagate dai provider. Il criterio maggiormente utilizzato per poter ottenere una promozione è un punteggio elevato del fornitore sulla piattaforma. Utilizzare però le recensioni e i rating come metrica di performance non è la scelta migliore perché spesso possono essere distorti o poco accurati.

In questo contesto i numerosi dati raccolti sugli utenti rendono molto più semplice l'applicazione di strategie di discriminazione. La tecnologia permette di conoscere in modo sempre più dettagliato gli user che utilizzano le piattaforme e le informazioni a disposizione aiutano a migliorare il meccanismo di pricing, permettendo alle società di catturare una percentuale sempre più alta del valore generato dalle transazioni.

3. Regolazione

3.1 Inquadramento generale

La sharing economy è un settore ancora in rapida evoluzione ed è destinata a ricoprire un ruolo sempre più rilevante nell'economia globale. Nonostante le incertezze riguardanti le dimensioni della stessa e la mancanza di una definizione chiara e precisa di sharing economy, questo nuovo modello di business necessita un intervento da parte dei regolatori. L'economia della condivisione si basa sulla fiducia tra le persone e si sviluppa in diversi ambiti: al fine di regolamentarla, invece di creare una nuova legge per ogni modello di condivisione, i regolatori dovrebbero creare leggi complete, considerando le pratiche di condivisione nel loro insieme. Se così non fosse, si creerebbero norme troppo specifiche che andrebbero continuamente modificate ogniqualvolta emerga una nuova attività di sharing economy o si modifichi una già esistente. La sfida per i regolatori risulta quindi particolarmente complessa: bisogna valutare se le leggi applicate all'economia tradizionale possano essere utilizzate anche per regolamentare la sharing economy, o se è necessario aggiornare, rivedere o creare nuove norme.

Si possono configurare diversi scenari durante una transazione in cui è necessario una regolamentazione per risolvere problematiche o comportamenti non corretti. Si pensi alla piattaforma Airbnb: si affitta un appartamento solo per scoprire che in realtà è stato subaffittato violando il contratto di locazione, come bisogna comportarsi? Oppure un host affitta un appartamento per essere poi informato che in realtà veniva utilizzato dal guest per feste. Le leggi come intervengono in queste situazioni? Considerando invece la piattaforma di car sharing Uber, chi è responsabile per la cattiva condotta di un conducente?

Questi sono solo alcuni esempi che permettono di evidenziare la varietà di casistiche che si possono verificare, dimostrando che scrivere una legge per ogni singola situazione per ogni singolo ambito è estremamente difficile, se non impossibile. I responsabili politici si trovano di fronte ad una scelta: essere molto

specifici o essere più generali, con il rischio però di non affrontare adeguatamente le singole problematiche e di non fornire soluzioni sufficientemente chiare ed esaustive.

Nei successivi paragrafi saranno analizzate le principali problematiche per regolare la sharing economy e gli interventi già attuati per questo settore.

3.2 Sfide regolatorie

Il modello di business della sharing economy, come già anticipato, ha generato diverse sfide per i regolatori. Una delle principali questioni da risolvere è capire se le regole del diritto della concorrenza nel mercato tradizionale possano applicarsi a questo settore e agli attori che entrano in gioco.

È fondamentale, prima di tutto, domandarsi se le piattaforme della sharing economy possano ritenersi imprese e se i fornitori di beni e servizi come l'host di Airbnb o l'autista di Uber siano considerati dipendenti della piattaforma o possano essere qualificati come imprese indipendenti.

Le piattaforme offrono beni e/o servizi in un determinato mercato: possono essere quindi considerate come delle imprese perché svolgono un'attività economica. Tuttavia, vi sono alcuni elementi da analizzare più dettagliatamente per capire se è corretto definirle come imprese.

Le piattaforme svolgono un ruolo da intermediario: connettono gli individui che offrono beni con le persone che domandano tali prodotti o servizi. Si potrebbe pertanto considerare questa pratica come una semplice un'attività di intermediazione su base professionale. Tuttavia, le piattaforme forniscono anche dei servizi, non si limitano a facilitare le transazioni tra utenti. La Commissione Europea a giugno 2016 decide di definire le caratteristiche essenziali che una piattaforma di sharing economy deve possedere per essere riconosciuta come fornitore di un bene o servizio, quindi come un'attività economica. L'aspetto più importante è il livello di controllo della piattaforma sui fornitori di servizi: se stabilisce i prezzi delle transazioni, se possiede le risorse e se definisce le condizioni

del rapporto contrattuale, può essere considerata come fornitore di servizi, rispondendo perfettamente alla definizione di impresa e può pertanto essere soggetta alle norme del diritto della concorrenza. In tal caso i fornitori sottostanti, come ad esempio i conducenti di Uber, devono essere considerati dipendenti della piattaforma, riservando loro determinati diritti sulle condizioni di lavoro, come le ferie retribuite o la remunerazione delle spese.

Un altro aspetto da trattare di particolare rilevanza riguarda la concorrenza diretta stabilita tra i soggetti fornitori di beni e servizi sulle piattaforme con redditi regolari e molto consistenti e i professionisti operanti nei diversi settori. L'ingresso di Uber nel mercato ha, per esempio, generato dei problemi ai taxisti per i quali si è ridotto il numero di corse (come descritto nel paper *"The Competitive Effects of the Sharing Economy: How is Uber Changing Taxis?"* di S. Wallsten, pubblicato nel 2015). La Commissione Europa ha stabilito che in alcuni casi i fornitori di beni delle piattaforme sono da considerarsi dipendenti, in altri casi sono individui autonomi; bisognerebbe comunque valutare caso per caso per capire se esiste un rapporto di lavoro tra i fornitori e la piattaforma. Se così fosse, le persone fisiche non possono essere considerate società a pieno titolo perché non svolgono un'attività economica in modo autonomo, ma operano per conto della piattaforma: le leggi sul diritto della concorrenza tra imprese non possono quindi essere applicate a questi individui. In questa situazione, è la piattaforma che offre beni o servizi a diventare competitor delle aziende tradizionali operanti nello stesso settore.

Nell'ambito della sharing economy la natura del rapporto contrattuale è un ulteriore elemento che necessita un intervento regolatorio, in particolare per indicare quali sono gli obblighi e le responsabilità per la piattaforma e per i fornitori di beni. Affinché le transazioni tra gli utenti vadano a buon fine, molte piattaforme hanno un ruolo attivo nel creare fiducia tra le parti e gestiscono uno schema di punteggi e recensioni. Tuttavia, questo sistema ha dei difetti in quanto spesso le persone possono essere manipolate o disoneste quando esprimono il proprio giudizio e la mancanza di trasparenza nel modo in cui i punteggi sono assegnati

può comportare la raccolta di informazioni distorte per gli altri utenti che utilizzano la piattaforma. In generale, molte società della sharing economy intervengono attivamente per limitare al minimo eventuali problematiche durante una transazione tra utenti: Uber, per esempio, offre una guida completa ai propri conducenti sulla pratica di car sharing, al fine di istruirli sui corretti comportamenti da seguire. Per quanto riguarda le responsabilità legali spesso queste vengono trasferite a chi fornisce i servizi: sempre in riferimento alla piattaforma di car sharing Uber, i conducenti devono essere muniti di patente di guida e di una propria assicurazione. Airbnb invece si assume in parte la responsabilità in caso di danni all'appartamento causati da guest e garantisce il risarcimento all'host fino ad un certo importo.

Infine, un'altra questione della condivisione riguarda la tassazione. Il reddito derivante dai servizi della sharing economy dovrebbe essere dichiarato e tassato, ma spesso potrebbe sfuggire all'attenzione delle autorità fiscali perché non segnalato. I regolatori devono quindi individuare delle soluzioni per tracciare questi flussi di reddito, impedendo che il problema assuma dimensioni importanti tali da rendere sempre più difficile la sua gestione.

3.3 Sharing Economy Act

La rapida crescita ed evoluzione della sharing economy, congiuntamente alle differenze emerse rispetto all'economia tradizionale, hanno reso sempre più evidente la necessità di una nuova legge in grado di gestire le nuove piattaforme digitali integrandole con il mercato attuale. Nel maggio del 2016 viene presentata in Italia una proposta di legge per promuovere e regolamentare l'economia della condivisione. Si tratta della proposta di legge 3564 - *"Disciplina delle piattaforme digitali per la condivisione di beni e servizi e disposizioni per la promozione dell'economia della condivisione"*, rinominata da molti *"Sharing Economy Act"*. L'obiettivo principale è quello di garantire equità e trasparenza per coloro che operano nell'ambito dell'economia della condivisione, oltre che tutelare i consumatori per le

questioni legate alla privacy, alla sicurezza dei loro dati personali e sulla chiarezza delle condizioni alla base del servizio. Si vuole inoltre valorizzare questo nuovo settore e cercare di promuoverne lo sviluppo.

In primo luogo, dopo aver esplicitato nell'articolo 1 le finalità della proposta di legge, è stata definita la sharing economy e i suoi limiti nell'articolo 2, oltre a identificare come destinatari della proposta di legge gli "abilitatori", ovvero i gestori delle piattaforme.

L'articolo 3 e l'articolo 4 individuano l'Autorità Garante della Concorrenza e del Mercato (AGCM) come competente per vigilare sulle pratiche di sharing economy. Inoltre, viene introdotto l'obbligo per gli abilitatori di dotarsi di un "documento di politica aziendale", soggetto al parere e all'approvazione dell'AGCM, contenente le condizioni contrattuali tra la piattaforma e gli utenti, insieme agli obblighi su eventuali polizze assicurative. Sarà inoltre istituito un "Registro elettronico nazionale delle piattaforme digitali dell'economia della condivisione", in cui è obbligatorio iscriversi. Se una piattaforma svolge la sua attività senza iscrizione, verrà immediatamente diffidata e dovrà sospendere il servizio. L'abilitatore che non rispetterà tale indicazione sarà punito con una sanzione che può valere fino al 25% del suo fatturato. Per coloro che invece non rispettano le disposizioni del documento di politica aziendale sono previste sanzioni pari all'1%-10% del fatturato.

Nella proposta di legge si affronta il tema della fiscalità, al fine di "*affermare i principi di trasparenza ed equità*". Per i redditi fino a diecimila euro ottenuti mediante le piattaforme verrà applicata un'imposta pari al 10%; d'altro canto, i redditi di valore superiore a tale soglia saranno cumulati con i redditi da lavoro autonomo o da lavoro dipendente e a questi si applicherà l'aliquota fiscale corrispondente. La legge, quindi, sottolinea la differenza tra il reddito occasionale (inferiore a diecimila euro) e il reddito superiore alla soglia indicata, senza però chiarire se i fornitori dei servizi sono considerati lavoratori autonomi o subordinati alla piattaforma.

L'articolo 6 disciplina le misure da adottare per promuovere la diffusione della sharing economy, *“al fine di rimuovere le barriere normative, di natura regolamentare o amministrativa”*. I regolatori politici sono quindi interessati alla crescita della sharing economy e hanno voluto inserire delle linee guida per valorizzare le pratiche di condivisione e per sfruttare al massimo il suo potenziale.

Un aspetto che non può assolutamente essere trascurato quando si parla di piattaforme digitali riguarda la privacy e la sicurezza dei dati degli utenti. Questi ultimi, nel momento in cui utilizzano una piattaforma di sharing economy, condividono informazioni personali che devono essere tutelate e disciplinate: in tema di trattamento dei dati, il gestore che intende cedere a terzi queste informazioni, deve comunicarlo ai soggetti interessati che hanno il diritto di eliminare i propri dati. Inoltre, viene disciplinata la possibilità di cancellare con una singola operazione tutte le informazioni memorizzate all'interno del profilo dell'utente in qualsiasi momento.

Gli ultimi cinque articoli si concentrano sulle linee guida per promuovere la sharing economy anche nella pubblica amministrazione, sulle regole in materia di controlli e sanzioni e infine sono fornite indicazioni su disposizioni finanziarie. Per quanto riguarda invece il monitoraggio e l'analisi statistica, i gestori delle piattaforme iscritti nel Registro citato nell'articolo 3 dovranno comunicare i dati relativi alle proprie performance (numero di utenti registrati, attività svolte e relativi importi, ecc.) all'Istituto Nazionale di Statistica (ISTAT). In tal modo si studierà lo sviluppo della sharing economy e l'efficacia delle azioni normative: l'obiettivo è quello di fornire ai regolatori gli strumenti per capire se sarà necessario modificare le leggi attuali o inserirne di nuove.

In conclusione, lo Sharing Economy Act ha gettato le basi per regolare un settore in evoluzione come quello dell'economia condivisa, ma proprio per la sua rapida crescita sarà necessario rivedere in futuro la proposta di legge. Non si è ancora raggiunto un livello di maturità in questo ambito tale da consentire un'adeguata normativa, tuttavia i regolatori hanno compreso il suo potenziale e la necessità di

una legge per disciplinare tali servizi: saranno necessarie modifiche e azioni correttive, ma è pur sempre un inizio.

3.4 I dati nella sharing economy

La diffusione del digitale e di Internet ha generato come prima diretta conseguenza un grande aumento del numero di dati creati e condivisi. Un termine ricorrente, citato in molte situazioni, è Big Data, ovvero la quantità di informazioni che vengono raccolte e utilizzate dalle aziende per gestire le proprie attività e prendere le giuste decisioni in ambito strategico. Molte società basano il proprio business sui dati raccolti, grazie ai quali riescono a individuare idee innovative per il futuro e ad aumentare il loro valore. La sharing economy è guidata dai Big Data: senza la mole di informazioni raccolta dalle piattaforme, sarebbe molto più complesso garantire le stesse prestazioni e lo stesso funzionamento. Tuttavia, la chiave vincente non è esclusivamente la raccolta dei dati, bensì la loro analisi, per riuscire ad estrarre da essi informazioni utili per creare sempre più valore.

Il paper *“How data analytics drive sharing economy business models?”*, di Soraya Sedkaoui e Rafika Benaichouba, affronta la tematica dell’analisi dei dati per le piattaforme di sharing economy. Inizialmente viene affrontato in generale il tema dei big data, oggi sempre più diffusi a causa di un costo di immagazzinamento basso o pari a zero, per la creazione di valore da essi generata e per la loro possibilità di essere combinati in modo da estrarre un numero elevato di informazioni. Raccogliere i dati ed analizzarli per renderli fruttuosi è l’abilità base delle società della sharing economy: non solo si riesce a migliorare le performance, ma anche a prevedere risultati economici come l’inflazione, la disoccupazione, i prezzi delle case, ecc. Nella letteratura, per avere un inquadramento generale dei Big data si fa riferimento al modello 3Vs, il quale indica le tre caratteristiche essenziali che i dati devono avere per essere di valore: volume, varietà e velocità. Per volume si intende la quantità di dati processati; la varietà è legata al tipo di dato, all’eterogeneità

dell'acquisizione e della rappresentazione mentre la velocità si riferisce al tempo impiegato per produrre, analizzare e immagazzinare i dati.

Le società della sharing economy hanno intuito il potenziale ricoperto dai dati e hanno basato la loro cultura e strategia direttamente su di essi (*data-driven*). Essere un'azienda *data-driven* implica un continuo dialogo tra i dati raccolti e la strategia della società: le decisioni prese devono basarsi sull'analisi delle informazioni. Le piattaforme di sharing economy raccolgono tre differenti tipologie di dati: quelli forniti direttamente dagli user (come informazioni del profilo, foto, ecc), dati relativi al comportamento delle persone e quelli generati dall'analisi delle due tipologie appena citate.

Per quanto riguarda l'analisi invece sono diversi gli algoritmi a disposizione, anche se i più utilizzati sono la regressione, la classificazione e l'analisi cluster. I primi due algoritmi cercano di associare gli elementi in input ad una determinata classe che può essere di tipo testuale (classificazione) o di tipo numerico (regressione). Nell'ambito della sharing economy, la classificazione può essere utilizzata per cercare parole in un testo o quando si prova a identificare le persone in una fotografia; la regressione viene utilizzata per prevedere quanto le persone sono disposte a spendere per una specifica risorsa, quanti clienti possono essere interessati a quell'item o per aiutare i conducenti Uber a capire quali parti dell'auto hanno maggiori probabilità di rompersi. Spesso questi algoritmi vengono anche combinati: la regressione e la classificazione possono essere utilizzati insieme per identificare potenziali clienti per gli appartamenti di Airbnb o per cercare destinazioni turistiche vicino ad una specifica località. L'analisi cluster invece può essere utilizzata ad esempio per studiare le attitudini degli utenti. Con questo algoritmo si parte dagli elementi in input e studiandone le caratteristiche vengono raggruppati in un certo numero di cluster, secondo le varie logiche come la similarità o la distanza.

Un caso studio reale in cui gli algoritmi vengono applicati è dato dalla piattaforma Airbnb. Attraverso l'analisi dei dati è possibile:

- Personalizzare: in base ad ogni ricerca dell'utente si riescono a trovare in tempo reale offerte che combacino perfettamente con i desideri e le richieste della persona;
- Previsione del prezzo: per determinare il valore di un appartamento o di una casa, Airbnb propone agli host un algoritmo chiamato Aerosolve: considerando diverse variabili come la città, il mese, il tipo di proprietà, i trasporti e analizzando le foto e le immagini di un listing, riesce a definire il prezzo;
- Valutare le opinioni degli utenti sulla base dei commenti: stabilire se il riscontro del cliente è positivo o negativo in base alle parole delle sue recensioni.

È chiaro quindi il ruolo essenziale ricoperto dai dati per le piattaforme di sharing economy, ma per la tutela degli utenti è necessario regolamentare il loro trattamento. Come già anticipato, nello Sharing Economy Act è stato dedicato un articolo alla riservatezza dei dati e alla privacy degli user, ma le piattaforme sono soggette anche alla disciplina del GDPR, il Regolamento Generale sulla Protezione dei Dati, principale strumento normativo in materia di *data protection*. Per poter accedere ai servizi di sharing economy gli utenti rilasciano informazioni personali come l'età, il nome, una foto, ma vengono raccolti dati anche in maniera passiva, senza che l'individuo ne sia consapevole, come l'indirizzo IP da cui avviene il collegamento o i cookies attivati. Il GDPR si pone l'obiettivo di tutelare i dati personali, così definiti: *"qualsiasi informazione relativa a una specifica persona, identificata o identificabile, facendo riferimento a caratteristiche fisiche, psicologiche, mentali, economiche dell'individuo in questione"*. Il problema è legato alla tutela dei dati non personali, non rientranti in questa definizione. Infatti, spesso le piattaforme raccolgono informazioni non riconducibili ad un singolo individuo, ma hanno un grosso potenziale perché, se combinati con altri dati non personali, potrebbero identificare un soggetto. La batteria del cellulare scarica è un'informazione generale, non personale, eppure può essere molto utile nella sharing economy: infatti, questa persona sarà probabilmente disposta a pagare un prezzo più alto rispetto alla media per un passaggio in auto offerto da piattaforme di car sharing.

Il GDPR, inoltre, in tema di trattamento dei dati, richiede da parte dell'utente il suo consenso esplicito, il quale deve essere *“una manifestazione di volontà libera, specifica, informata e inequivocabile”*. Spesso però informazioni come la località geografica o il tempo di connessione vengono raccolte senza avvisare l'individuo. Rendere l'utente pienamente consapevole di quali dati vengono raccolti e delle finalità per cui sono trattati è particolarmente complesso: tendenzialmente, infatti, le piattaforme forniscono dei moduli standard in cui sono definiti i termini e le condizioni per il trattamento dei dati, rimanendo sempre sul generale e non sottolineando le differenze di ogni specifica situazione.

La disciplina in termini di protezione dei dati presenta quindi ancora delle mancanze e dovrebbe essere revisionata per tener conto di questi aspetti, al fine di tutelare al massimo gli utenti delle piattaforme di sharing economy.

4. Airbnb

4.1 Introduzione

Airbnb è la piattaforma di sharing economy ormai diventata un punto di riferimento per la locazione turistica, in particolare per gli affitti di breve termine. Se qualche anno fa era presa in considerazione solo dai giovani in cerca di vacanze economiche, oggi è utilizzata da gran parte delle persone per trovare un appartamento o una casa in cui alloggiare durante i propri viaggi e il suo successo si osserva dai dati: 5 milioni di annunci, sistemazioni in 190 paesi nel mondo. Milioni di persone decidono oggi di creare un account su Airbnb per pubblicare annunci dei loro spazi da affittare o per prenotare una sistemazione. Le piattaforme della sharing economy fungono da intermediari per creare un punto di contatto tra domanda e offerta: infatti Airbnb mette in comunicazione coloro che offrono sistemazioni, gli host, con coloro che sono in viaggio e ricercano uno spazio in cui alloggiare, i guest. I benefici sono per tutti: gli host riescono a guadagnare da appartamenti poco sfruttati e sono molti coloro che gestiscono sulla piattaforma più di un annuncio, i guest hanno una valida alternativa agli hotel per scegliere la sistemazione più adatta alle loro esigenze e Airbnb aumenta il suo fatturato con il passare degli anni. Nel 2016, secondo L'Espresso, gli host italiani sono riusciti a guadagnare circa 621 milioni di euro grazie ad almeno 5,6 milioni di viaggiatori che hanno preferito un alloggio sulla piattaforma piuttosto che una sistemazione in hotel. Ovviamente questi dati generano alcune riflessioni sull'impatto di Airbnb nel settore turistico. Sicuramente la piattaforma ha cambiato il modo di viaggiare degli utenti ed è destinata ad avere sempre più successo nel tempo.

4.2 Impatto Airbnb su settore turistico

Airbnb ha avuto effetti sui consumatori, sugli hotel e sugli host che registrano i propri appartamenti sulla piattaforma.

Il paper *“The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb”* di Chiara Farronato e Andrey Fradkin, pubblicato nell’ottobre del 2017, valuta l’impatto di Airbnb sul settore turistico, in particolare sul surplus generato e sulle conseguenze per gli hotel.

Gli utenti ricercano sistemazioni presso il quale alloggiare durante le loro esperienze di viaggio e l’offerta a loro disposizione può essere flessibile o dedicata. Con la prima si intendono i proprietari di alloggi o case che decidono di ospitare viaggiatori solo in alcuni periodi (peer host), con la seconda invece si parla degli hotel, i quali dedicano le proprie stanze per ospitare i turisti tutto l’anno. La principale differenza tra le due tipologie di offerte è rappresentata dai diversi costi da sostenere: gli hotel hanno maggiori costi di investimento mentre i fornitori flessibili tipicamente hanno costi marginali più alti (devono preparare le camere per l’arrivo di nuovi ospiti, devono interagire con i viaggiatori durante il loro pernottamento, ecc.) e tendenzialmente non hanno bisogno di alti investimenti iniziali.

I viaggiatori beneficiano da Airbnb per due ragioni: gli host offrono prodotti differenziati rispetto agli hotel ed espandendo il numero di stanze disponibili, ampliando quindi l’offerta. Questo effetto è particolarmente importante nei periodi di alta domanda dove gli hotel hanno una capacità limitata. Questi ultimi non sono in grado di assorbire i picchi di domanda perché il numero di stanze a loro disposizione è fisso. Per costruire nuove strutture in media ci vogliono da 3 ai 5 anni. I fornitori flessibili invece sono in grado di fornire un’offerta aggiuntiva durante i picchi di domanda, quando le loro sistemazioni hanno grande valore per i turisti.

In generale, la piattaforma contribuisce a generare benefici per i guest e per gli host. Il surplus dei consumatori per ogni notte prenotata su Airbnb risulta essere pari a 70\$. Il surplus totale guadagnato invece da Airbnb è di 432 milioni di dollari mentre in media gli host ricevono per notte un surplus pari a 28\$.

Il dataset utilizzato per riportare tali risultati comprende 50 città degli USA in cui sono state analizzate le prenotazioni di hotel e gli affitti su Airbnb nel periodo tra il 2011 e il 2014. Risulta esserci eterogeneità in termini di aree geografiche (in alcune città le stanze disponibili su Airbnb sono aumentate più velocemente) e in termini temporali (molta disponibilità di stanze nei periodi estivi e di festività). Questo suggerisce che i fornitori flessibili di Airbnb sono più propensi a registrare le loro case sulla piattaforma nelle città e nei periodi in cui i ritorni per gli affitti sono più alti.

È stata costruita una retta di regressione per valutare l'elasticità degli hotel e degli host rispetto ai prezzi.

$$\text{Log}(Q_{mt}) = x\log(K_{mt}) + w\log(p_{mt}) + \mu_{mt} + \varepsilon_{mt}$$

Q_{mt} → Numero di prenotazioni in una città m al giorno t ;

K_{mt} → Capacità nella città m nel giorno t ;

P_{mt} → Prezzo della transazione;

w → Elasticità rispetto al prezzo.

Le altre due variabili racchiudono effetti fissi sulla città, sulla stagionalità, ecc.

L'analisi è svolta separatamente per gli hotel e per gli host.

Risultati: l'elasticità degli hotel è pari a 2.1 mentre risulta essere il 92% maggiore per i fornitori flessibili. Airbnb ha quindi elasticità più alta: vuol dire che ci sono più host disposti ad affittare le loro stanze quando i prezzi sono alti ma preferiscono non farlo quando i prezzi sono bassi.

Gli autori evidenziano un problema di endogeneità: il prezzo della transazione dipende dalla domanda. Sono utilizzate allora due variabili strumentali per rendere

i prezzi esogeni rispetto alla domanda: il numero di voli atterrati in una determinata città per quel giorno e i Google Trends, dove la query di interesse è “Hotel c” con c=nome città.

Si vuole inoltre indagare l’effetto di Airbnb non solo sui prezzi delle transazioni degli hotel, ma anche sui loro ricavi e sul tasso di occupazione delle camere. La retta di regressione è la seguente:

$$Y_{mt} = \alpha \log(\text{airbnb}_{mt}) + \beta \log(\text{gtrend}_{mt}) + \gamma \log(\text{travelers}_{mt}) + \theta_{mt} + v_{mt}$$

Y_{mt} → Outcome degli hotel (ricavi, prezzi, tasso occupazione camere);

airbnb_{mt} → Sistemazioni disponibili su Airbnb;

gtrend_{mt} → Google Trend;

travelers_{mt} → Numero voli atterrati.

L’effetto di interesse è α e risulta essere statisticamente significativo: utilizzando come outcome i ricavi, si osserva un valore di α pari a -0.36%. Questo significa che in media, un 10% in più di appartamenti su Airbnb riduce i ricavi per gli hotel dello 0.36%. I minori guadagni degli hotel sono causati da una riduzione dei prezzi dei pernottamenti piuttosto che da una riduzione del tasso di occupazione degli stessi. Se l’outcome sono i prezzi degli hotel infatti si osserva che l’effetto è sempre significativo e questi diminuiscono. Questo effetto è maggiore nelle città dove gli hotel hanno capacità limitata, dove un 10% in più di appartamenti su Airbnb riduce i prezzi degli hotel del 0,52%.

Infine, si è svolta anche un’analisi per capire quali sarebbero gli effetti se Airbnb non fosse presente sul mercato. Il surplus dei consumatori si ridurrebbe: se infatti le camere degli hotel fossero tutte occupate e non ci fossero i listing di Airbnb, gli utenti non avrebbero alternative tra cui scegliere. In realtà risulta che il 70% delle prenotazioni effettuate sulla piattaforma degli affitti brevi non sarebbero sostituibili con pernottamenti in hotel. In un contesto senza Airbnb gli hotel interverrebbero alzando i prezzi per notte, aumentando così i loro ricavi.

4.3 Airbnb vs Booking

Nel mercato delle prenotazioni a breve termine uno dei principali concorrenti di Airbnb è Booking. La piattaforma offre agli utenti la possibilità di fermare una sistemazione per le proprie vacanze scegliendo non solo tra i vari appartamenti registrati dagli host: a differenza di Airbnb infatti sono mostrate anche camere di hotel.

Analizzando più nel dettaglio le due società, si presentano differenze anche in termini di acquisizioni, investimenti e finanziamenti.

Booking ha acquisito esclusivamente Evanture Technologies Ltd nel 2017, società incentrata nello sviluppo di una tecnologia di ricerca testuale per l'industria dei viaggi, basata sull'intelligenza artificiale.

Airbnb invece solo nel 2019 ha acquisito tre società: Urbandoor, marketplace online che permette di cercare e prenotare appartamenti arredati, HotelTonight che sviluppa un servizio di prenotazione alberghiera e Gaest che consente agli utenti di prenotare luoghi per riunioni e servizi fotografici.

In termini di finanziamenti, Booking ha raccolto 0\$ di finanziamenti al contrario di Airbnb, la quale ha partecipato a 14 round di finanziamenti, raccogliendo in totale 6,6 miliardi di dollari. L'ultimo round di finanziamento è stato ad aprile 2020 per una cifra pari a 1 miliardo di dollari dal fondo di investimenti Silver Lake.

Gli investimenti invece sono stati di più per Airbnb ma di cifre inferiori rispetto alla sua rivale. Infatti, Booking ha investito 180 milioni di dollari in Yanolja nel 2019, piattaforma online che consente di prenotare hotel e circa 225 mila dollari in Authenticook, piattaforma che collega i viaggiatori agli host locali per cenare a casa loro. Airbnb solo nel 2019 ha investito una cifra pari a 55 milioni di dollari in Zeus, piattaforma che consente di affittare case arredate, 60 milioni di dollari in Tigets, applicazione online che permette di acquistare biglietti per musei e spettacoli e 19,1 milioni in Atlas Obscura, rivista di viaggi.

La *Tabella 4.3.1* riassume le principali differenze tra Airbnb e Booking (fonte: Owler).

	Anno fondazione	Entrate all'anno	Finanziamenti	Sede	N° dipendenti	Settore
Airbnb	2008	\$ 4,8 miliardi	\$ 6,6 miliardi	San Francisco, USA	5600	Servizi Immobiliari
Booking	1996	\$ 4 miliardi	\$ 0	Amsterdam	25000	Software applicativo

Tabella 4.3.1 Airbnb vs Booking

Le due piattaforme di affitti brevi mostrano un'ulteriore differenza nei prezzi proposti per le sistemazioni. Booking registra prezzi più dinamici, tendenzialmente più alti rispetto a quelli della sua competitor ma molto più variabili. Airbnb invece ha prezzi in media più bassi e soprattutto fissi.

Concentrandosi infine sulla popolarità delle due società, Booking è primo in termini di visibilità e numero di utenti grazie all'inserimento tra le sue proposte di camere di hotel e B&B, mentre Airbnb ha il primato sul numero di alloggi inseriti da privati. Nel 2018 Booking ha registrato sul sito visite mensili pari a 428 milioni, mentre Airbnb circa 80 milioni.

4.4 Airbnb vs Hotel

Airbnb, come dimostrato nel paragrafo 4.2, ha impattato anche sul settore degli hotel. Per molti viaggiatori in realtà le due sistemazioni non sono confrontabili e non possono essere considerate sostitute. Alcuni utenti apprezzano la libertà offerta dagli appartamenti, senza alcun vincolo di orario, con la possibilità di avere più spazio e una cucina ben fornita. Gli amanti degli hotel d'altro canto non rinuncerebbero mai a tutti i servizi offerti, come la pulizia della camera ogni giorno, i miniclub per i bambini, l'animazione, la presenza di una reception 24/24 per ogni problema.

Esistono invece molte persone che prediligono una proposta piuttosto che un'altra in base al prezzo di queste. In media gli hotel presentano prezzi più alti rispetto agli appartamenti proprio per i numerosi servizi offerti e forse non tentano di competere

con Airbnb fissando prezzi più bassi, ma sull' user-experience generata. Sicuramente le due alternative seguono strategie completamente diverse per fissare i loro prezzi. Due variabili che influenzano entrambe le soluzioni sono il posizionamento (vicino/lontano dal centro) e la stagionalità.

Il prezzo degli hotel è influenzato da aspetti come le dotazioni della struttura, il valore dei servizi offerti, la tipologia di camera, la concorrenza con gli altri hotel, il brand (si pensi agli Hilton), ecc. In realtà, anche aspetti come l'aumento delle ricerche di volo per una determinata destinazione o le condizioni atmosferiche possono incidere sul prezzo della camera. Raramente gli hotel stabiliscono delle tariffe fisse per ogni stanza: la teoria dello Yield management suggerisce che è necessario proporre il prezzo che quel determinato cliente è disposto a pagare. Molte strutture, infatti, applicano tariffe differenti in base al segmento di mercato servito, cioè clienti accumulati da caratteristiche simili. In base ai canali di distribuzione le camere sono inserite con prezzi differenti. Capita spesso poi che quando la domanda sale, le offerte con i prezzi più bassi per quella camera vengono eliminate, in modo da riuscire ad avere il maggior ricavo possibile.

Diverse invece sono le variabili considerate per decidere quale tariffa applicare per un appartamento su Airbnb. Il paper *"Price Determinants on Airbnb: How Reputation Pays Off in the Sharing Economy"* di Tim Teubner, David Dann e Florian Hawlitschek pubblicato nel 2017, indaga le variabili significative per la determinazione del prezzo di un listing su Airbnb. I principali aspetti considerati dagli autori come influenti per il prezzo sono rappresentati dalla *Figura 4.4.1*.

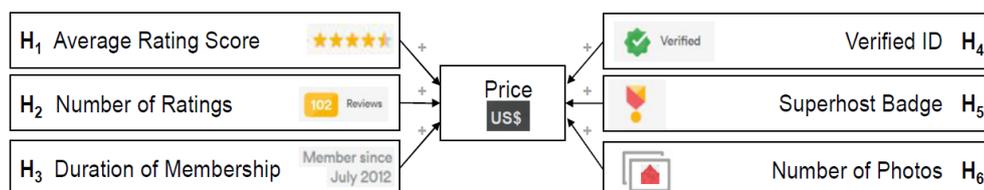


Figura 4.4.1 Variabili influenti nella scelta del prezzo

Le ipotesi studiate sono le seguenti:

H1: *Higher average rating scores are associated with higher listing prices.*

Gli autori hanno analizzato i punteggi espressi con le stelle nelle diverse categorie: location, check-in, comunicazione, precisione, rapporto qualità-prezzo, pulizia. Grazie a diversi modelli di regressione è emerso che ad alti punteggi sono associati alti prezzi. Uno studio di Wang e Nicolau condotto nel 2017 conferma la medesima ipotesi: una stella in più nel punteggio corrisponde ad un mark-up sul prezzo dello 0,87%.

H2: *Higher numbers of reviews are associated with higher prices.*

Un alto numero di recensioni rende i punteggi più affidabili, mentre se compaiono poche valutazioni aumentano i dubbi sulla loro veridicità. Per esempio, un basso numero di recensioni potrebbe essere raccolto esclusivamente da amici familiari ecc. fornendo un'informazione distorta all'utente.

H3: *A longer duration of membership is associated with higher prices.*

La permanenza su Airbnb può beneficiare sulla reputazione dell'host e quindi impattare sui prezzi. I profili caricati richiedono tempo per la loro gestione e in media hanno meno probabilità di creare nei clienti false aspettative. Con l'esperienza si riesce inoltre a fissare il giusto prezzo.

H4: *Verified IDs are associated with higher prices*

Airbnb permette ai propri utenti di verificare la loro identità scansionando un documento di identità. Una volta approvato, il sistema registra l'autenticità dell'utente e questo serve per aumentare la fiducia.

H5: *Superhost badges are associated with higher prices*

Lo status di Superhost è sicuramente un segnale di qualità e può essere utilizzato per costruire la propria reputazione e quindi aumentare il prezzo. Da evidenze

empiriche di Wang e Nicolau del 2017 si registra un mark-up dell'8,37% con questo status.

H6: *A higher number of apartment photos is associated with higher prices.*

Le foto di un appartamento permettono agli ospiti di farsi un'idea sulla sistemazione e forniscono informazioni sulla qualità della casa, permettendo di aumentare i prezzi perché si riducono i rischi per i guest.

L'analisi è stata svolta su un dataset contenente 13.884 listing di 86 città tedesche. È stata costruita una retta di regressione per valutare l'impatto di queste variabili sul prezzo di una permanenza di due notti in una sistemazione per due persone.

Risultati: gli effetti significativi risultano quelli di H1 (punteggi delle valutazioni), H3 (membership) e H6 (numero di foto). H4 (verifica ID) e H5 (status di super host) non sono significative. In contrasto con le ipotesi, c'è un effetto negativo, significativo e consistente tra H2 (numero di valutazioni) e il prezzo.

La conclusione a cui giungono gli autori è l'importanza della reputazione in una piattaforma come Airbnb, in cui non si conoscono gli utenti e le transazioni si basano sulla fiducia generata tramite il sistema di recensioni.

4.5 Diffusione Airbnb

Dai paragrafi precedenti il successo di Airbnb è evidente: dal suo anno di fondazione la piattaforma è cresciuta sempre di più, aumentando i guadagni per sé stessa, per gli host che decidono di caricare i propri appartamenti e creando sempre più benefici per i viaggiatori.

Per valutare quanto effettivamente il servizio offerto da Airbnb si sia diffuso sul mercato, si può fare riferimento al modello "Diffusione dell'innovazione" di M. Rogers. Secondo tale teoria, sono quattro i principali elementi che influenzano la diffusione di una nuova idea: l'innovazione in sé, i canali di comunicazione, il tempo e le caratteristiche della società.

Il modello di Rogers individua diverse categorie di persone che adottano una determinata innovazione (Figura 4.5.1):

- Innovator: coloro che sono più interessati alla componente tecnologica e alle performance del prodotto piuttosto che ai servizi da esso offerti. Sono meno price-sensitive. Rappresentano i primi 2,5% di individui che adottano l'innovazione;
- Early adopters: sono disponibili a pagare un prezzo relativamente alto ma iniziano ad essere interessati ai servizi connessi all'innovazione. Sono i successivi 13,5% che adottano il prodotto/servizio;
- Early majority: sono coloro che ricercano una user experience migliore possibile, per loro è importante il tema dell'usabilità e i servizi offerti sono il motivo principale per cui decidono di utilizzare l'innovazione. Rappresentano il 34,5%;
- Late majority: non adottano l'innovazione fin quando non risentono una pressione sociale causata dagli altri utenti che hanno già adottato il prodotto/servizio. Sono il 34,5%;
- Laggards: sono scettici riguardo l'innovazione, decidono di utilizzarla esclusivamente se non c'è alternativa e dopo aver sentito i pareri positivi degli altri utenti. Il 16% di utenti rimanenti nell'adottare la tecnologia è da loro rappresentato.

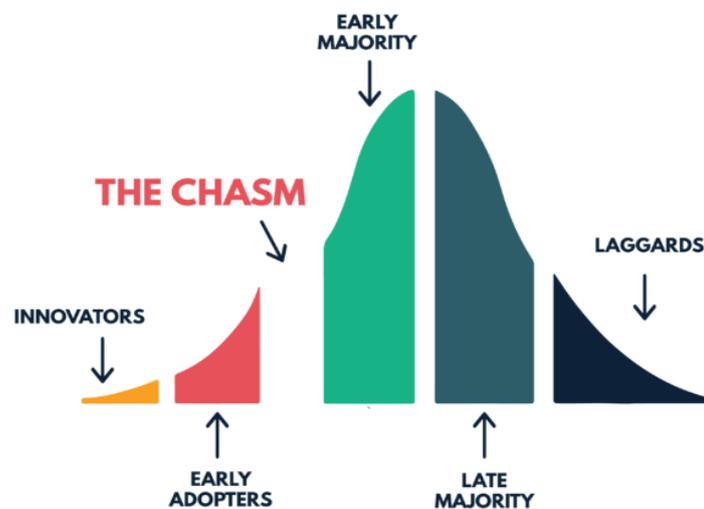


Figura 4.5.1 Modello di Rogers

In base ai dati raccolti sull'utilizzo di Airbnb è possibile definire a quale fase di diffusione è arrivata la piattaforma. Per far ciò, si possono utilizzare il numero di arrivi di guest negli appartamenti o nelle case. Nel 2009, la piattaforma registrava appena 21.000 arrivi di ospiti, aumentati poi fino a 80 milioni nel 2016 e nel 2019 si sono registrati circa mezzo miliardo di ospiti (Fonte: *"The diffusion of Airbnb: a comparative look at earlier adopters, later adopters, and non-adopters"* di Daniel Guttentang). È necessario confrontare questi numeri con quanti utenti hanno deciso di mettersi in viaggio in un determinato anno. La World Tourism Organization, l'agenzia delle nazioni unite che si occupa di monitorare l'andamento del settore turistico, ha registrato un numero di turisti in tutto il mondo nel 2019 pari a circa 1,4 miliardi. Questo significa che circa il 36% dei viaggiatori hanno deciso di trovare una sistemazione per la propria vacanza su Airbnb. Si può pertanto affermare, in base a questi calcoli, che il mercato abbia raggiunto l'early majority e si sia superato il Chasm, ovvero il "baratro" tra gli Early Adopter e la maggioranza iniziale, così chiamato a causa della grande differenza tra le due categorie nei criteri di scelta per l'adozione di una tecnologia.

È abbastanza alto quindi il numero di viaggiatori che utilizzano la piattaforma e con il tempo è destinato ad aumentare, Covid19 permettendo.

5. Il sistema delle recensioni per Airbnb

Airbnb è un network basato sulla fiducia e per tale ragione si è dotato di uno strumento, all'interno della piattaforma stessa, per permettere ai propri utenti di effettuare prenotazioni in modo sicuro: le recensioni. Quando un guest visita il sito dell'azienda per prenotare un soggiorno, la sua prima attività consiste proprio nella lettura delle recensioni degli altri utenti che hanno precedentemente alloggiato nella medesima sistemazione. Grazie a questo strumento, è possibile beneficiare di un feedback autentico che, auspicabilmente, potrebbe abbattere le asimmetrie informative e contemporaneamente aumentare la customer experience dell'utente, dal momento che gli viene fornita un'idea ex ante di come sarà il suo viaggio.

La letteratura in merito ai sistemi di recensione delle piattaforme è particolarmente vasta e, concentrandosi in modo più dettagliato sulla società in analisi, sono stati selezionati quattro studi che verranno riportati di seguito.

Il primo articolo, intitolato *"Inside the Engine Room of Digital Platforms: Reviews, Ratings, and Recommendations"*, affronta le principali tematiche associate alle recensioni, evidenziandone anche eventuali punti di debolezza, tra cui la possibilità di incorrere nella lettura di false informazioni. In riferimento a tale problematica, il paper *"The determinants of online review informativeness: evidence from field experiments on Airbnb"* descrive le principali fonti di bias delle recensioni, cercando di individuare la causa principale della distorsione delle informazioni. Il terzo studio, *"A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average"* esegue un confronto tra i modelli di review di Airbnb e del competitor TripAdvisor. Infine, il quarto elaborato denominato *"You get what you give: theory and evidence of reciprocity in the sharing economy"* studia l'effetto della reciprocità (comportarsi in base all'aspettativa che si ha sul comportamento della controparte), coerentemente con il sistema *reciprocal review* di Airbnb, sui punteggi degli host e sui prezzi da loro applicati.

5.1 Inside the Engine Room of Digital Platforms: Reviews, Ratings, and Recommendations

Il paper, scritto da Paul Belleflamme e Martin Peitz, pubblicato nel febbraio 2018, affronta il tema delle recensioni, dalla fiducia che creano tra le parti a come guidano gli utenti nelle loro transazioni.

Le piattaforme si possono definire come imprese il cui core business è quello di creare valore per gli utenti attraverso la loro interazione. L'attrattività di queste tecnologie dipende dal numero di utenti presenti, ma in quale modo i sistemi di valutazione e di raccomandazione sono utili per la creazione del *network*? Per rispondere a questa domanda si pensi al caso di Amazon, più utenti acquistano i prodotti e li recensiscono, più i nuovi acquirenti avranno delle informazioni ex-ante che permetteranno loro di prendere una decisione migliore.

Inoltre, è possibile personalizzare il sistema di rating mostrando, ad esempio, valutazioni e recensioni solo di acquirenti con determinati profili. Una tale funzionalità può rappresentare una guida migliore per l'acquisto, poiché ciò che è preferibile per un gruppo di acquirenti non è necessariamente buono per gli altri: si pensi ad esempio ad un viaggiatore d'affari che valuterà degli aspetti differenti rispetto ad un gruppo di amici in vacanza.

Le recensioni costituiscono uno strumento utile per la riduzione delle asimmetrie informative soprattutto nell'ambito dei beni di esperienza, in quanto gli acquirenti hanno meno informazioni rispetto ai venditori in merito alla qualità del prodotto. Ulteriori strumenti utilizzati tradizionalmente per risolvere tali problematiche ex-ante sono le certificazioni che garantiscono il livello di qualità minima offerta e si possono classificare come obbligatorie e volontarie: Uber controlla i registri dei suoi autisti per assicurarsi che siano idonei alla guida, tale certificazione è obbligatoria. Airbnb invece offre agli host la possibilità di certificare l'autenticità delle foto della rispettiva proprietà, riducendo così il rischio di spiacevoli sorprese per l'acquirente. I problemi di asimmetria informativa possono essere affrontati anche ex post

attraverso le assicurazioni, le quali possono, in linea di principio, essere fornite dai venditori stessi, ma le piattaforme sono spesso in una posizione migliore per fornirle, poiché interagiscono più frequentemente e direttamente con gli acquirenti. Ad esempio, Airbnb assicura i venditori contro il vandalismo da parte degli acquirenti.

I sistemi di valutazione e di revisione completano questi strumenti classici e tendono a diventare più efficienti tanto maggiore è il numero di transazioni che le piattaforme facilitano. Infatti, la loro capacità di fronteggiare i problemi di informazione affrontati dagli acquirenti aumenta con il volume, la varietà e la velocità dei dati che le piattaforme possono raccogliere sui loro utenti e sulle transazioni che conducono.

Chevalier e Mayzlin nel 2006 hanno analizzato l'effetto delle recensioni dei libri sui modelli di vendita delle due principali librerie online negli Stati Uniti: Amazon e Barnes & Noble. La questione centrale dello studio consiste nel valutare se un giudizio negativo su Amazon porti a un calo delle vendite maggiore rispetto a Barnes & Noble. Per rispondere a questa domanda è stato utilizzato l'approccio '*differences-in-differences*', regredendo il logaritmo naturale delle vendite su una serie di variabili, tra cui alcuni *fixed effects* come i prezzi e le quote di recensioni positive (5 stelle) e negative (1 stella). Si ottiene che un'ulteriore recensione positiva per un libro porta ad un aumento delle vendite. Vi sono anche alcune prove che una recensione negativa è più potente nel diminuire le vendite di un libro rispetto a un'ulteriore recensione positiva nell'aumentarle.

Nel 2018 due studiosi Vana e Lambrecht hanno effettuato un approfondimento sull'importanza dell'ordine delle recensioni ed hanno mostrato che le valutazioni delle prime recensioni visualizzate ha un forte impatto sulla futura probabilità di acquisto. In particolare, se queste recensioni hanno un rating elevato (quattro o cinque stelle su cinque), la probabilità di acquisto stimata aumenta in modo significativo.

Sorge, quindi, spontaneo domandarsi quanto ci si dovrebbe fidare delle recensioni. Ovviamente più recensioni sono presenti più la valutazione media che si ottiene rispecchierà maggiormente l'effettiva qualità del prodotto. Per tale motivazione, il sistema di valutazione potrebbe essere progettato per incoraggiare gli acquirenti a lasciare una recensione, fornendo magari incentivi non monetari e monetari per lasciare le recensioni.

Un certo numero di opere empiriche hanno inoltre dimostrato che i venditori più affidabili hanno maggiore successo, come dimostrato dallo studioso Cai, che nel 2014 ha riscontrato una correlazione positiva tra il tasso di successo cumulativo dei venditori e la frazione dei compratori abituali, dimostrando che se la transazione è avvenuta con successo è più probabile che le due parti vogliano interagire nuovamente tra di loro.

Le recensioni sono utili solo se contengono informazioni rilevanti. Si possono identificare tre cause che potrebbero condurre a distorsioni: (i) valutazioni rumorose, (ii) recensioni strategicamente distorte e (iii) asymmetric herding behavior.

Per la prima causa di distorsione, di seguito verranno riportate quattro cause per cui gli acquirenti potrebbero rilasciare tale tipo di recensione:

- Mancata comprensione: gli utenti non comprendono quali aspetti effettivamente stanno valutando, ad esempio nel caso di un prodotto acquistato su Amazon potrebbero recensire i tempi di attesa invece che le caratteristiche proprie;
- Idiosyncratic tastes: i buyers potrebbero commentare in modo eterogeneo in base ad alcune caratteristiche soggettive. Per esempio, recensore potrebbe dare una valutazione negativa del prodotto perché non ne gradisce il colore, mentre altri potrebbero apprezzarlo;
- Uncontrollable shock: eventi che non sono sotto il controllo del venditore, ad esempio se l'azienda di trasporto non ha consegnato in tempo il prodotto;
- Variazioni di prezzo = le recensioni sono spesso basate su quanto l'acquirente è soddisfatto rispetto a quanto ha pagato.

Per quanto riguarda le recensioni strategicamente distorte, entrambe le parti potrebbero intraprendere azioni che generano valutazioni distorte.

Chiaramente, poiché i venditori beneficiano di una reputazione positiva, possono pagare altri affinché lascino recensioni positive oppure negative sui competitors. Tuttavia, occorre costatare che generare tali revisioni false è costoso, come sostengono Ott, Cardie e Hancock (2012) che si sono occupati del caso degli hotel, in cui le recensioni sono fondamentali per incentivare l'acquisto della camera. Le piattaforme di prenotazione alberghiera Booking ed Expedia richiedono un acquisto effettivo prima di lasciare una review, mentre Tripadvisor lo permette a chiunque dichiari di aver effettuato una prenotazione senza verificarne la veridicità. Quindi, le recensioni false sono più costose nel primo caso, mentre nel secondo chiunque potrebbe scriverle senza conseguenze.

Un ulteriore studio sempre nel settore alberghiero effettuato da Mayzlin, Dover e Chevalier ha dimostrato che gli hotel indipendenti sono più propensi a sponsorizzare recensioni false, in quanto in caso vengano scoperti rischiano di perdere meno reputazione rispetto a quelli appartenenti a famose catene.

Nel 2017 Fradkin, Grewal e Holtz hanno condotto alcuni esperimenti sul campo e hanno scoperto che coloro che non forniscono recensioni tendono ad avere esperienze peggiori di quelli che lo fanno. Nel caso di Airbnb hanno affermato che dal momento in cui acquirente e venditore possono interagire, potrebbero essere meno inclini a lasciare feedback negativi. Inoltre, le recensioni non sono anonime, per cui chi affitta un appartamento può controllare il *track record* di qualcuno che vorrebbe affittarlo e se questa persona tende a lasciare recensioni negative, un futuro proprietario potrebbe essere meno incline a confermare la richiesta.

In conclusione, se gli utenti della piattaforma effettuano queste azioni il network effect positivo diminuirà.

Per quanto riguarda i sistemi di raccomandazione dei prodotti, sono utilizzati dalle piattaforme per suggerire ai propri utenti quei beni che maggiormente potrebbero rispondere ai loro gusti. I consumatori possono essere classificati come "amatoriali"

o "esperti". I primi basano la loro decisione sulla popolarità, mentre i secondi acquisiscono informazioni sulle caratteristiche intrinseche del prodotto ed in base a queste effettuano l'acquisto. La piattaforma permette di "raccomandare" quei beni che potrebbero essere maggiormente apprezzati dagli utenti, basandosi non solo sulle osservazioni del singolo fruitore, bensì anche sul comportamento degli altri. Dunque, una maggiore partecipazione da parte di un gruppo di utenti aumenta la possibilità per la piattaforma di proporre match più attraenti per i buyers futuri.

Per progettare un corretto sistema di raccomandazione occorre valutare il fenomeno denominato "*long tail*", che rappresenta la distribuzione delle vendite che si osserva per molti prodotti digitali, costituita da pochi articoli che coprono la maggior parte della domanda e moltissimi articoli caratterizzati da poche vendite. Alcuni acquirenti, ad esempio, sono interessati ai prodotti di nicchia, mentre altri potrebbero rendersi conto di preferire prodotti standard che soddisfano il gusto del mercato di massa. I sistemi di raccomandazione possono contenere informazioni sulla popolarità, ovvero mostrare in termini relativi la frequenza di acquisto del bene, così che un buyer caratterizzato da un gusto comune alla massa acquisterà prodotti che in passato hanno venduto molto. Amazon si contraddistingue per la presenza di un mix di vari sistemi di raccomandazione, tra i quali il più degno di nota è che quando si acquista un bene ne consiglia altri che sono stati acquistati dai buyers precedenti dell'oggetto in questione. Ovviamente tale sistema non si limita a segnalare la popolarità dei prodotti, bensì permette ai consumatori di scoprire quali prodotti servono ad altri acquirenti con gusti simili ai loro, permettendogli di trovarli senza ulteriori ricerche. Questa metodologia però sembra coinvolgere solo prodotti che sono già piuttosto popolari tra l'intera popolazione.

In antitesi a tale argomentazione si sono susseguiti diversi studi empirici che hanno dimostrato come un sistema di raccomandazione possa portare, dopo la sua introduzione, ad una vendita maggiore anche dei prodotti meno popolari e che quindi potrebbe essere uno strumento ideale proprio nei mercati di nicchia. Ad esempio, gli studiosi Oestreicher-Singer e Sundararajan hanno effettuato delle

ricerche sul settore dei libri venduti su Amazon e sono giunti ad affermare che proprio i prodotti di nicchia hanno ottenuto risultati relativamente migliori in termini di vendite.

5.2 The determinants of online review informativeness: evidence from field experiments on Airbnb

Lo studio, condotto da Andrey Fradkin, Elena Grewal e David Holtz e pubblicato nell'aprile 2018, indaga come il sistema di reputazione influisce sulla misura in cui le recensioni forniscono informazioni sulla piattaforma Airbnb. In generale, le recensioni rappresentano un fattore critico di successo per i mercati digitali e riducono le asimmetrie informative. Per funzionare e creare vantaggi agli utenti però devono essere veritiere: la probabilità che gli user utilizzino una piattaforma caratterizzata da un alto numero di recensioni distorte diminuisce nel tempo.

Gli autori hanno studiato 59981 viaggi avvenuti tra maggio e giugno 2014. Dall'analisi dei dati emerge che gli host hanno un tasso di recensione più alto e recensiscono prima principalmente per due ragioni: entrano nella piattaforma più frequentemente dei guest (ad esempio per verificare le prenotazioni) e monetizzano da Airbnb, quindi ottengono un maggior guadagno nel lasciare una recensione al guest.

Spesso molte informazioni vengono perse nelle review perché gli utenti non recensiscono o perché non descrivono la vera esperienza vissuta in modo oggettivo. Per capire il grado di veridicità delle review su Airbnb, gli autori descrivono tre esperimenti per spiegare la causa della presenza di informazioni distorte nelle recensioni.

Esperimento 1

Prima fonte di bias nelle review: individui con esperienze diverse recensiscono con un tasso diverso; alcuni utenti recensiscono quindi più di altri.

Gli host sono stati divisi in due gruppi (trattamento e controllo). Ai guest che hanno alloggiato da host del gruppo di trattamento viene dato un coupon se lasciano una

recensione, mentre per i guest che soggiornano dal gruppo di controllo è inviata una semplice e-mail di reminder. Il tasso di recensione nel gruppo di controllo risulta inferiore rispetto al gruppo di trattamento. Inoltre, più passa il tempo per lasciare la recensione, più i punteggi dei guest sono bassi.

Esperimento 2

Seconda fonte di bias: la decisione di recensire e il tipo di testo utilizzato sono influenzate dalla risposta prevista dalla controparte.

L'esperimento consiste nel cambiare il tempo in cui le recensioni diventano pubbliche. Per il gruppo di trattamento, 1/3 degli host non vedono le recensioni fin quando sia loro che i guest non le hanno scritte. Per il gruppo di controllo invece 1/3 degli host ha 14 giorni per lasciare una recensione e sono mostrate appena vengono scritte. Infine, un terzo gruppo ha a disposizione 30 giorni per lasciare una review.

Dall'esperimento risulta che il tasso di recensione aumenta sia per gli host che per i guest. Quest'ultimi, tuttavia, rilasciano meno punteggi con 5 stelle e aumentano le recensioni con testo negativo.

Esperimento 3

Terza fonte di bias: natura sociale delle transazioni, cioè la volontà del guest di scrivere recensioni è in funzione del numero di contatti sociali con l'host.

Prenotare su Airbnb prevede spesso un'interazione sociale con l'host e questo può influenzare la decisione di recensire. La probabilità di interagire dipende molto da alcune caratteristiche dell'host: se si affitta una stanza privata rispetto ad una intera proprietà, aumentano probabilmente le interazioni sociali perché si condividono gli spazi. Inoltre, se l'host ha registrato più di tre appartamenti su Airbnb (host non occasionale), tendenzialmente avrà meno contatti sociali con il guest.

Dall'esperimento risulta che i guest che alloggiano in appartamenti di host occasionali (meno di 3 listings) tendono a lasciare un punteggio di 5 stelle rispetto ai soggiorni in sistemazioni di host non occasionali.

Gli esperimenti dimostrano come le recensioni rilasciate non rappresentano la qualità reale dell'esperienza vissuta. Gli autori si pongono l'obiettivo di calcolare il grado di informazione persa.

Primo bias: differenza tra il punteggio medio dell'esperienza reale e la media dei punteggi delle recensioni rilasciate.

$$B_{avg} = (1 - g) \frac{rlp}{Pr(r)} - g \left(1 - \frac{rp}{Pr(r)} \right)$$

Il primo termine rappresenta gli utenti che hanno vissuto un'esperienza negativa ma scrivono una recensione positiva, mentre il secondo sono gli utenti che vivono un'esperienza positiva ma scrivono una recensione negativa.

g → probabilità di avere viaggio con esperienza positiva;

rlp → probabilità di lasciare una recensione positiva dopo esperienza negativa;

$1-g$ → probabilità di avere viaggio con esperienza negativa;

rp → probabilità di lasciare recensione positiva dopo esperienza positiva;

ru → probabilità di lasciare recensione negativa con esperienza negativa;

$Pr(r)$ → tasso totale di recensione → $Pr(r) = g*rp + (1-g)(rlp+ru)$.

Secondo bias: quota delle esperienze negative non recensite.

$$B_{neg} = 1 - \frac{Nn/n}{N_{all}(1 - g)}$$

Nn/n → numero di report negativi dati da user con esperienza negativa;

N_{all} → numero totale di viaggi.

Si calcolano i bias delle recensioni dei guest valutando 5 scenari differenti.

Scenario 1: operano tutti e 3 le fonti di bias precedentemente descritte nei tre esperimenti;

Scenario 2: si considera il gruppo di trattamento del secondo esperimento (non si considerano quindi gli aspetti strategici);

Scenario 3: si considerano tutte le fonti di bias esclusa la natura sociale delle transazioni (terzo esperimento);

Scenario 4: non si considera la prima fonte di bias (tasso di recensione diverso tra gli utenti)

Scenario 5: calcolo del bias nel caso in cui tutti rilasciano reviews.

Risultati:

Scenario 1: le recensioni positive sono di 1,32 punti percentuali in più rispetto alle esperienze positive.

Scenario 2 e 3: il bias medio si riduce di 0.88 punti percentuali e il numero di recensioni negative non riportate è di 4.6 punti percentuali.

Scenario 4: il bias medio scende di 1.1 punto percentuale. Le differenti esperienze degli utenti sono quindi la più importante fonte di bias.

Scenario 5: il bias medio non cambia perché non cambia il numero di recensioni sbagliate.

In conclusione, esistono bias dovuti ad aspetti strategici e alla natura sociale delle transazioni, ma gli effetti sono più piccoli rispetto a quelli causati da un tasso di recensione diverso tra gli utenti.

5.3 A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average

Lo studio in analisi, pubblicato nell'aprile 2015 e condotto da Georgios Zervas e Davide Proserpio, si pone l'obiettivo di valutare l'andamento delle recensioni sulla piattaforma di Airbnb. Basandosi su una collezione di 600000 proprietà presenti nel 2015 sulla piattaforma, si è riscontrato che circa il 95% presentava una media di valutazione pari a 4,5 su 5 stelle e pochissime ne avevano meno di 3,5. Gli studiosi che si sono occupati del suddetto lavoro hanno effettuato un confronto con un'altra piattaforma del settore turistico ovvero TripAdvisor, la cui media delle valutazioni si aggira intorno a 3,8 stelle. Inizialmente sono state valutate le proprietà in base alla tipologia di alloggio ed alla posizione, da cui si è riscontrata una notevole variabilità nelle valutazioni, osservando che le proprietà in affitto per le vacanze su TripAdvisor avevano valutazioni più simili a quelle delle proprietà Airbnb. Infine,

si sono considerate diverse migliaia di proprietà elencate su entrambe le piattaforme, da cui si è riscontrato che sebbene le valutazioni medie sulle due piattaforme siano simili, proporzionalmente più abitazioni ricevono valutazioni più alte su Airbnb rispetto che TripAdvisor.

Le recensioni online rappresentano un tool fondamentale per comprendere il comportamento dei futuri consumatori. Occorre però costatare che spesso non rappresentino l'effettivo livello di qualità offerta, infatti alcuni studi precedenti mostrano come le recensioni siano generalmente positive per quattro motivazioni salienti:

- *Herding behavior*: per cui le valutazioni precedenti influenzano quelle successive (Salganik et al. 2006, Muchnik et al. 2013);
- sotto-rapporto delle recensioni negative: per cui i revisori temono ritorsioni su piattaforme che permettono e incoraggiano la revisione bilaterale (Dellarocas e Wood 2008, Bolton et al. 2013, Fradkin et al. 2014);
- l'autoselezione: i consumatori che a priori hanno più probabilità di essere soddisfatti di un prodotto sono anche più propensi ad acquistarlo e recensirlo (Li e Hitt 2008);
- la manipolazione strategica delle recensioni: tipicamente intrapresa da imprese che cercano artificialmente di influenzare la loro reputazione online (Mayzlin et al. 2014, Luca e Zervas 2013).

Occorre costatare che, nonostante gli aspetti sopraelencati, oltre il 70% dei consumatori afferma di avere fiducia nelle recensioni online.

Gli autori hanno deciso di concentrarsi su Airbnb poiché si contraddistingue per il carattere personale delle recensioni, dal momento che gli utenti valutano l'esperienza nella casa o nell'appartamento di un altro individuo. Per lo studio è stato valutato il database completo disponibile sul sito di Airbnb e per il confronto con TripAdvisor quello della suddetta piattaforma. Entrambe le aziende operano su scala mondiale, sebbene la seconda non si occupi solo di case vacanze, bensì

anche di B&B e hotels; su entrambi i siti i guests posso esprimere la loro valutazione mediante una scala di 5 stelle.

Nell'ultima fase dello studio è stato effettuato un confronto sulle abitazioni presenti su entrambe le piattaforme, così da controllare l'effetto dell'eterogeneità delle tipologie di strutture. Anche in questo caso si è riscontrato che il 14% delle proprietà ha ricevuto una valutazione più elevata su Airbnb. Per spiegare questo fenomeno, la prima possibile motivazione individuata è il modello di gestione bilaterale delle recensioni adottato, che incentiva gli host a dare feedback positivi ai guest per riceverne a loro volta. Per cercare di limitare questo fenomeno, dal luglio del 2014, è stata introdotta una nuova metodologia che, sebbene consista sempre nella reciproca recensione, permette ad entrambe le parti di visionare la recensione ottenuta dalla controparte solo nel momento in cui entrambe le parti si sono espresse, fino ad un massimo di 14 giorni dopo il soggiorno.

Il database analizzato, come precedentemente accennato è stato scaricato dal sito di Airbnb, contenente 600000 proprietà, a cui è seguita una fase di pulizia, eliminando gli immobili con un numero inferiore alle 3 recensioni, ottenendo un campione di 226.549. Lo stesso procedimento è stato applicato per TripAdvisor per essere coerenti con il metodo utilizzato, ottenendo così 412.223 hotels e 54.008 case vacanze.

Dal momento in cui le proprietà su Airbnb non sono dotate di una posizione precisa è stata utilizzata una tecnica euristica per valutarne approssimativamente la latitudine e la longitudine. Una volta individuate, si è proceduto ad abbinare ogni struttura di Tripadvisor ad una potenziale di Airbnb in base alla distanza geografica e si sono ottenuti 2234 matches, tra 1959 proprietà di TripAdvisor (di cui 827 hotels o B&B e 1132 case vacanze) e 2234 di Airbnb. Il suddetto metodo potrebbe sembrare eccessivamente approssimativo e per tale ragione è stata effettuata un'ulteriore analisi campionaria per valutare la qualità delle coppie individuate e non si sono riscontrati errori sistematici.

È risultato che le valutazioni su Tripadvisor sono meno elevate di quelle di Airbnb, infatti solo il 4% degli hotels presentano una media di 5 stelle e circa il 26% maggiore di 4,5. Tale differenza potrebbe essere spiegata dall'eterogeneità delle tipologie di strutture, infatti i clienti di un hotel potrebbe essere più severi nel giudizio piuttosto che nei confronti di un proprietario di una casa vacanze. Per eliminare tale *confounding factor* si sono considerati solo gli immobili presenti su TripAdvisor e si è ottenuto un valore medio pari a 4.6 stelle. La regressione effettuata ha confermato che la differenza tra Airbnb e Tripadvisor è significativa.

Infine, per analizzare ulteriormente la significatività sono state valutate le proprietà presenti su entrambe le piattaforme e come precedentemente anticipato è risultata significativa, in particolare il 14% in più di strutture sono state valutate maggiormente su Airbnb.

Successivamente gli autori del paper si sono chiesti in quale misura le valutazioni su Airbnb prevedono quelle di TripAdvisor, ovvero se le recensioni della seconda piattaforma, sebbene siano quasi sempre inferiori, potrebbero prevedere quelle della prima e viceversa. Dalla retta di regressione analizzata si può affermare che esista una correlazione positiva tra le valutazioni delle due piattaforme, sebbene R^2 del modello sia relativamente basso (0,17). Tale valore suggerisce che le valutazioni su una piattaforma spiegano solo una piccola parte della variabilità delle recensioni sull'altra. Si è riscontrato, ad esempio, che una parte della variabilità potrebbe dipendere dal fatto che gli utenti su TripAdvisor abbiano una *willingness to pay* maggiore rispetto a quelli di Airbnb, infatti sul primo sito è possibile valutare di alloggiare anche presso costosi hotels. Inoltre, un'altra determinante potrebbe essere la posizione geografica. Per tali due ragioni sono state inserite nel modello due variabili dummy che hanno portato R^2 ad aumentare a 0.22. Per esaminare ulteriormente la correlazione è stato utilizzato il *Kendall rank correlation* (τ) che ha confermato i risultati ottenuti con la regressione precedente: esiste infatti una correlazione positiva sebbene sia debole.

In conclusione, le recensioni su Airbnb sono superiori rispetto alla piattaforma del competitor per la metodologia del “*reciprocal reviews*”.

5.4 You get what you give: theory and evidence of reciprocity in the sharing economy

Nel presente articolo, pubblicato nel giugno 2018 e scritto da Davide Proserpio, Wendy Xu e Georgios Zervas, viene valutata la reciprocità, cioè la tendenza ad aumentare (o diminuire) lo sforzo in risposta all’aumento (o diminuzione) dello sforzo della controparte nella transazione. Questo effetto è molto più forte nei sistemi di recensione bilaterali, dove entrambe le parti possono lasciare una review.

Dopo una breve introduzione del mercato peer to peer (come Airbnb), gli autori si concentrano sulla definizione di reciprocità. In particolare, questa viene descritta come una norma sociale sotto la quale le persone rispondono alle azioni di altri con un’azione equivalente. Nella letteratura sono fornite due diverse definizioni:

- Letteratura delle recensioni online: la reciprocità si riferisce al comportamento strategico di recensire quando un feedback positivo (negativo) presenta un’alta probabilità di essere ricambiato con un feedback positivo (o negativo) dalla controparte;
- Letteratura dei behavioral economics: la reciprocità è la tendenza dei partecipanti del mercato di rispondere con un buon (cattivo) comportamento ad un buon (cattivo) comportamento. Nel paper si farà riferimento a questa definizione.

Per studiare l’effetto della reciprocità il quadro di riferimento è il seguente: si assume un unico host, il quale fissa i prezzi per le proprie sistemazioni e diversi guest. Il modello si articola in tre stadi. Nel primo periodo l’host sceglie il prezzo P_1 e ogni ospite decide se entrare o meno nel mercato. Nel secondo periodo entrambi gli attori decidono il loro livello di sforzo. Ogni guest inoltre pubblica un punteggio e viceversa l’host assegna un rating al guest i . Nel terzo periodo altri consumatori entrano nel mercato: l’host e i guest osservano le domande del primo periodo e i

punteggi assegnati nel secondo. L'host definisce allora un prezzo P3 sulla base dei punteggi ottenuti. Si risolve con backward induction, dal periodo due al periodo uno. Le utilità dell'host e dei guest nel secondo periodo sono le seguenti:

$$U_i(e_i/e_h, r_i) = v_h + \alpha u(e_i; e_h) - C_i(e_i) + \beta r_i$$

$$U_h(e_h/e_i; r_{h,i}) = v_i + \alpha u(e_i; e_h) - C_h(e_h) + \beta r_{h,i}$$

Con il pedice h si fa riferimento all'host, mentre con il pedice i si denota l'i-esimo guest e la sua utilità. V_h e V_i sono due fattori esogeni che influenzano la qualità della permanenza (per esempio V_h potrebbe essere la posizione dell'appartamento, mentre V_i se il guest ha un animale con sé durante il soggiorno). Il secondo termine $u(e_i, e_h)$ cattura l'utilità derivante dall'interazione tra i due attori, definita "*shared experienced utility*" ed è funzione dello sforzo dell'host e del guest (per esempio e_h potrebbe aumentare se sono fornite informazioni su attrazioni locali all'ospite, e_i invece aumenta se per esempio vengono rispettate dal guest le regole della casa). " α " è invece il peso della reciprocità. La derivata prima rispetto allo sforzo è positiva (all'aumentare dello sforzo dell'host o del guest aumenta l'utilità); anche la derivata seconda è positiva (i rendimenti sono quindi crescenti). $C(e)$ rappresenta il costo di sostenere un determinato livello di sforzo. " β " è il peso della reputazione, cioè quanto l'agente si preoccupa del punteggio.

Nel primo periodo invece l'utilità dell'host è composta da due parti: i ritorni monetari ($P \cdot Q$) e l'utilità attesa del secondo periodo. L'utilità del guest è data da quella attesa nel secondo periodo meno il prezzo pagato per la permanenza. Entrambe le utilità dipendono dai livelli di sforzo ottimali esercitati da guest e host, il cui obiettivo è quello di ottenere la massima utilità.

I dati utilizzati per risolvere il modello sono valutati su 17 mesi tra luglio 2014 e novembre 2015. Il panel contiene informazioni su 198.743 appartamenti 137687 host. Gli autori vogliono verificare due preposizioni:

Preposizione 1: *Il punteggio medio dell'host dipende dal peso della reciprocità dell'host (α).*

Un host con un peso della reciprocità più alto ha una volontà maggiore di migliorare l'esperienza del guest esercitando più sforzo. A causa della reciprocità anche i guest esercitano più sforzo. L'aumento del livello di sforzo di entrambi aumenta l'utilità dell'esperienza condivisa, che si riflette in punteggi più alti lasciati all'host.

Preposizione 2: *Il prezzo cresce su Airbnb dopo uno shock positivo sui punteggi, mentre decresce dopo uno shock negativo”.*

Un host con un punteggio medio alto sarà più attraente per guest futuri e avrà una domanda attesa più alta rispetto ad un host con lo stesso prezzo ma con un punteggio più basso.

Per testare la prima preposizione bisogna conoscere il peso della reciprocità dell'host α . La prima proxy utilizzata per indicare il peso della reciprocità è se l'host scrive recensioni lunghe sul guest. Un host che si preoccupa sull'esperienza vissuta su Airbnb spenderà più tempo a scrivere una recensione. La seconda proxy è se l'host utilizza l'instant book, cioè se la prenotazione da parte di un ospite non richiede la sua approvazione esplicita. Se non si utilizza questo strumento, l'host assegna alla reciprocità un peso più alto perché si interessa maggiormente dei suoi ospiti.

Si regredisce il punteggio ottenuto dall'host per queste due variabili.

Risultati: i coefficienti sono entrambi positivi e significativi, suggerendo che gli host che scrivono lunghe review e che non usano l'instant book (e quindi danno un peso alla reciprocità più alto) hanno in media punteggi più alti. Si ottiene lo stesso risultato considerando un set di variabili di controllo che possano influenzare il punteggio dell'host.

$$\text{Star-rating} = \beta_1 \log(\text{lunghezza review dell'host}) + \beta_2 (\text{no instant book}) + X_{ij} + e$$

Per trovare evidenze sulla seconda preposizione, è difficile stimare l'impatto dei punteggi sui prezzi perché ci sono cambiamenti non riconducibili ad una minore o

maggior qualità dei listing ma che possono simultaneamente creare effetti sia sui prezzi che sui rating. Per ridurre al minimo i problemi di endogeneità dovuti a tali cambiamenti non osservabili, ci concentriamo su listing i cui rating (che sono un proxy della qualità) sono costanti. Dal momento che i punteggi non sono resi pubblici fin quando non si raccolgono 3 recensioni, si può dedurre che se un listing ottiene 5 stelle come punteggio medio, ha ricevuto 3 recensioni con 5 stelle. Si considerano pertanto solo i listing in cui i punteggi resi pubblici sono in media di 5 stelle dopo 3 reviews.

$$\text{Log (prezzo)} = \beta \text{Dit} + \text{Xit} + \alpha + \tau + e$$

La variabile dipendente è il log del prezzo del listing i al periodo t . D è una dummy che vale 1 se il punteggio medio del listing i è stato reso pubblico nel tempo t . X è un set di variabili di controllo per cambiamenti di qualità nel tempo. Le altre due variabili sono effetti fissi per il tempo e per i listing.

Risultati: il coefficiente di interesse è positivo e significativo. Un punteggio di 5 stelle reso pubblico aumenta il prezzo di 1,6%. Se poi si includono le variabili di controllo e gli effetti fissi il coefficiente rimane positivo e significativo.

Si verifica inoltre che gli host professionisti hanno punteggi più bassi di quelli casuali. Quelli professionisti hanno una partecipazione di mercato alta, quelli non professionisti bassa. Si regredisce il punteggio in stelle sulla partecipazione di mercato. Aumentando la partecipazione di mercato diminuisce il punteggio.

Infine, si valuta se le proxy per la prima preposizione siano correlate negativamente con la partecipazione di mercato.

$$\text{Market participation} = \beta_1 \log(\text{lunghezza review host}) + \beta_2 (\text{no instant book}) + X_{ij} + e$$

Senza includere nessun controllo, i coefficienti sono correlati negativamente alla partecipazione di mercato (queste variabili infatti avevano effetto positivo sul punteggio dell'host, ma dato che il punteggio è correlato negativamente alla

partecipazione di mercato, qua l'effetto è negativo). Anche includendo variabili di controllo il risultato non cambia.

In conclusione, la reciprocità ha effetti sui punteggi e favorisce l'esercizio di un maggiore sforzo lato guest e lato host. La diretta conseguenza è una maggiore utilità per entrambe le parti. Un buon meccanismo di matching della piattaforma induce una reciprocità positiva, aumentando il benessere sia per i guest che per gli host.

6. Uomini e donne: esiste veramente una differenza di genere?

Uomini e donne guardano le stesse cose? Nelle più comuni conversazioni viene posta spesso questa domanda, ma ancora oggi non è possibile rispondere a questo quesito.

Uomini e donne appartengono alla medesima specie, ma è differente il modo in cui elaborano ed agiscono, poiché il procedimento di raccolta di informazioni che raggiungono la corteccia celebrale è differente, in accordo con quanto sostenuto dalle ricerche scientifiche. Numerosi studi, infatti, hanno dimostrato che esistono differenze comportamentali, strutturali e funzionali tra i due sessi, causate da piccole ma decisive difformità dello sviluppo cerebrale. Nonostante le neuroscienze affermino che più del 99% del codice genetico di uomini e donne sia il medesimo, quell'1% di variazione tra i due sessi influisce su qualsiasi cellula del corpo.

Si pensi ad esempio ai processi emotivi, valutati attraverso TAC e radiografie, vissuti in modo dissimile da uomini e donne: queste ultime sanno esprimere meglio le proprie emozioni, come evidenziato da diversi esami eseguiti. Gli uomini tendono ad assumere più rischi come si può evidenziare da diverse ricerche effettuate nel mondo del lavoro, le donne invece risultano essere più timorose delle conseguenze e quindi tendono ad essere più prudenti. Questi sono solo due dei numerosi esempi che caratterizzano l'eterogeneità degli atteggiamenti dei due sessi in ogni situazione ed aspetto della vita quotidiana e professionale.

Il paper *"Gender Differences in Preferences"* di Rachel Croson e Uri Gneezy studia proprio le differenze di genere, focalizzandosi principalmente su tre aspetti: avversione al rischio, preferenze sociali e reazione alla competizione. Per tale analisi sono stati valutati diversi articoli pubblicati in precedenza da altri autori e dallo studio di alcuni di questi emerge la prima considerazione: gli uomini assumono più rischio rispetto alle donne. Una possibile spiegazione di questa differenza risiede nell'emotività: le donne vivono le emozioni più intensamente, sono più sensibili al

nervosismo ed alla paura. Gli uomini, invece, tendono a essere più sicuri di sé stessi e questo li spinge a rischiare di più. Inoltre, è diversa l'interpretazione della situazione rischiosa: per le donne è una minaccia che si può evitare, per gli uomini rappresenta una sfida. Queste diverse attitudini non sono confermate se vengono prese in analisi alcune figure professionali, come i manager: in questa circostanza infatti non si evidenziano differenze tra uomini e donne, forse perché chi ricopre tali ruoli tende a essere in generale più disposto a rischiare.

L'analisi delle preferenze sociali riporta risultati contrastanti: in alcuni studi le donne tendono a fidarsi di più rispetto agli uomini, sono più altruiste e cooperano di più, ma questo non emerge in tutti i paper analizzati dagli autori. Si è però intuito che le donne sono molto più sensibili all'ambiente esterno nel determinare il loro comportamento verso gli altri.

L'elaborato si conclude con la seguente riflessione: gli uomini sembrano essere più stimolati dalla competizione nell'eseguire le loro performance rispetto alle donne. Questo è dovuto sia alla naturale differenza negli atteggiamenti tra i due sessi, sia dal fatto che le donne spesso tendono a non competere, perché sanno già di avere meno possibilità rispetto agli uomini. Si pensi ad una determinata posizione lavorativa: le candidature femminili sono molte volte penalizzate rispetto a quelle maschili e proprio questo spinge le donne a non proporsi per tali mansioni.

Cogliendo le suddette diversità tra uomini e donne, le aziende spesso differenziano le strategie di marketing in base al genere a cui sono rivolte e tale approccio è denominato "*gender marketing*". Le imprese cercano di comprendere i bisogni e le diverse esigenze dei due sessi per comprendere le loro future scelte di acquisto, ad oggi influenzate non solo più dal prodotto in sé, bensì dal valore e dal messaggio ad esso associato. Questo vale per tutti ma soprattutto per le donne che, nell'atto del comprare, spesso sono più ricettive ai contenuti emotivi. Inoltre, richiedono prodotti di qualità senza tralasciare la variabile del prezzo, confrontando offerte e promozioni, contrariamente all'uomo che acquista subito anche se l'importo è più elevato. Sono state rilevate ulteriori differenze persino nella ricerca delle

informazioni: le donne sono più sensibili ai dettagli e tendono ad elaborare le notizie in modo completo ed esaustivo; al contrario, gli uomini solitamente valutano solo quelle informazioni che considerano utili.

Tuttavia, è molto importante non cadere negli stereotipi e quindi nella discriminazione di genere, come è accaduto per la pubblicità dell'azienda Huggies, di cui è stata vietata la trasmissione nel 2015. La società, infatti, aveva pensato a due differenti tipologie di pannolini per bambini e bambine e aveva condotto una campagna pubblicitaria sfruttando i classici stereotipi di genere: "lei penserà a farsi bella, lui a fare goal. Lei cercherà tenerezza, lui avventure. Lei si farà correre dietro, lui invece la cercherà. Così piccoli e già così diversi."

Tali differenze di atteggiamenti si riflettono anche nelle piattaforme di Sharing Economy, come si cercherà di dimostrare nel seguente capitolo, attraverso lo studio di un ulteriore paper *"Does gender bias exist? The impact of gender congruity on consumer's Airbnb booking intention and the mediating role of trust"* ed un articolo *"Uomini e donne valutano le recensioni online in modo diverso?"*, redatto da Jamie Pitman.

6.1 Does gender bias exist? The impact of gender congruity on consumer's Airbnb booking intention and the mediating role of trust

Il seguente studio, condotto da Na Su e Anna Mattila, esamina l'impatto della congruità di genere sull'intenzione di prenotare un appartamento su Airbnb da parte dei guest.

I viaggiatori scelgono la loro sistemazione sulla piattaforma non solo valutando la qualità della proprietà, ma anche in base alle caratteristiche dell'host. Per creare fiducia tra le parti infatti, guest e host sono incoraggiati a fornire più informazioni personali per segnalare la loro credibilità e affidabilità, come il caricamento di foto personali o l'inserimento dei loro canali social. Tuttavia, la divulgazione di informazioni personali potrebbe sollevare una questione di discriminazione digitale basata su razza, sesso, età ed altre caratteristiche demografiche. Ci si può

domandare se effettivamente il caricamento di una foto personale è veramente necessaria al fine di prenotare un appartamento e creare fiducia tra gli utenti, o se invece aumenta la possibilità di discriminare determinati individui sulla piattaforma.

Sicuramente i consumatori devono interagire con l'host sia online che offline per completare una transazione e pertanto il proprietario rappresenta un fattore critico, quando si sceglie di soggiornare in una sistemazione di Airbnb. Nel contesto della piattaforma in esame, la fiducia nell'host è più critica di quella nella piattaforma. Proprio per l'importanza della figura dell'host nella decisione dei guest, gli autori si domandano se i consumatori mostrano un'intenzione di prenotare più alta per un proprietario di casa del medesimo sesso.

Studi precedenti sulla congruità di genere si basano principalmente sulla teoria della demografia relazionale e sul paradigma della somiglianza-attrazione. La prima sostiene che è nella natura umana utilizzare caratteristiche demografiche come razza, sesso, livello di istruzione o stato socio-economico per valutare la somiglianza con gli altri. La teoria dell'attrazione per somiglianza indica invece che gli individui che si percepiscono simili in termini di dati demografici sono più propensi a comunicare tra loro.

Gli autori nel presente lavoro propongono la seguente ipotesi:

H1: *I consumatori mostrano livelli più elevati di intenzione di prenotare quando il sesso dell'ospite è congruente con il proprio. Inoltre, l'effetto è particolarmente rilevante per i consumatori di sesso femminile rispetto a quelli di sesso maschile.*

Nell'ambito della piattaforma Airbnb, la fiducia ricopre indubbiamente un ruolo fondamentale per portare a termine una transazione. La letteratura sulle relazioni interpersonali suggerisce due forme fondamentali di fiducia: cognitiva ed affettiva. La prima si riferisce alla propria convinzione sulla capacità del fornitore di servizi di svolgere i propri compiti in modo soddisfacente. La fiducia cognitiva si basa, infatti, su un giudizio razionale basato sulle proprie esperienze passate ed

altre conoscenze disponibili. La fiducia affettiva, d'altra parte, si riferisce alle proprie percezioni su un fornitore di servizi. A differenza della fiducia cognitiva, la fiducia affettiva si basa sui sentimenti. Ricerche precedenti hanno mostrato che le due tipologie sono correlate positivamente e che un aumento di una tipologia di fiducia può portare ad un aumento dell'altro tipo. Da queste considerazioni gli autori propongono una seconda ipotesi:

H2: *La fiducia affettiva e cognitiva mediano l'effetto della congruità di genere sull'intenzione di prenotazione dei consumatori. Inoltre, l'effetto di mediazione della fiducia è più importante per le donne rispetto agli uomini.*

Lo studio è stato condotto attraverso un questionario, coinvolgendo 200 consumatori che sono stati assegnati ad host di sesso maschile o femminile. In particolare, 41 utenti di sesso femminile erano associati ad un host donna e 56 uomini sempre ad una proprietaria, mentre 45 donne e 56 uomini sono stati studiati con un host uomo. I partecipanti sono stati esposti a un annuncio di proprietà di Airbnb che riportava una foto personale dell'host, per capirne il sesso. In un secondo momento è stato chiesto agli individui di valutare la loro intenzione di prenotare attraverso 4 enunciati e 7 classi di valutazione (ad esempio, soggiornare in questa struttura Airbnb sarebbe un'esperienza piacevole: 1 = fortemente in disaccordo, 7 = assolutamente d'accordo). In seguito, i partecipanti hanno valutato la loro fiducia affettiva percepita e la fiducia cognitiva nei confronti dell'host. La fiducia affettiva è stata misurata con quattro enunciati su una scala di valutazione da 1 a 7 (ad esempio, l'ospite mostrerà un atteggiamento caldo e premuroso nei miei confronti: 1 = fortemente in disaccordo, 7 = fortemente d'accordo). La fiducia cognitiva è stata misurata con tre enunciati e con la medesima scala (ad esempio, posso fidarmi fiduciosamente di questo host: 1 = fortemente in disaccordo, 7 = fortemente d'accordo).

Nel modello per testare H1, sono stati inseriti come variabili indipendenti: il sesso dell'host, il sesso dei partecipanti al sondaggio e la loro interazione, mentre

l'intenzione di prenotare è stata inserita come variabile dipendente. Si è dimostrato che le partecipanti di sesso femminile hanno livelli più elevati di intenzione di prenotare presso un host dello stesso sesso rispetto ad uno di sesso opposto; la differenza è statisticamente significativa. Allo stesso modo, i partecipanti di sesso maschile hanno dimostrato un livello più alto di intenzione di prenotare presso un host dello stesso genere. Tuttavia, questa differenza non ha raggiunto livelli di significatività.

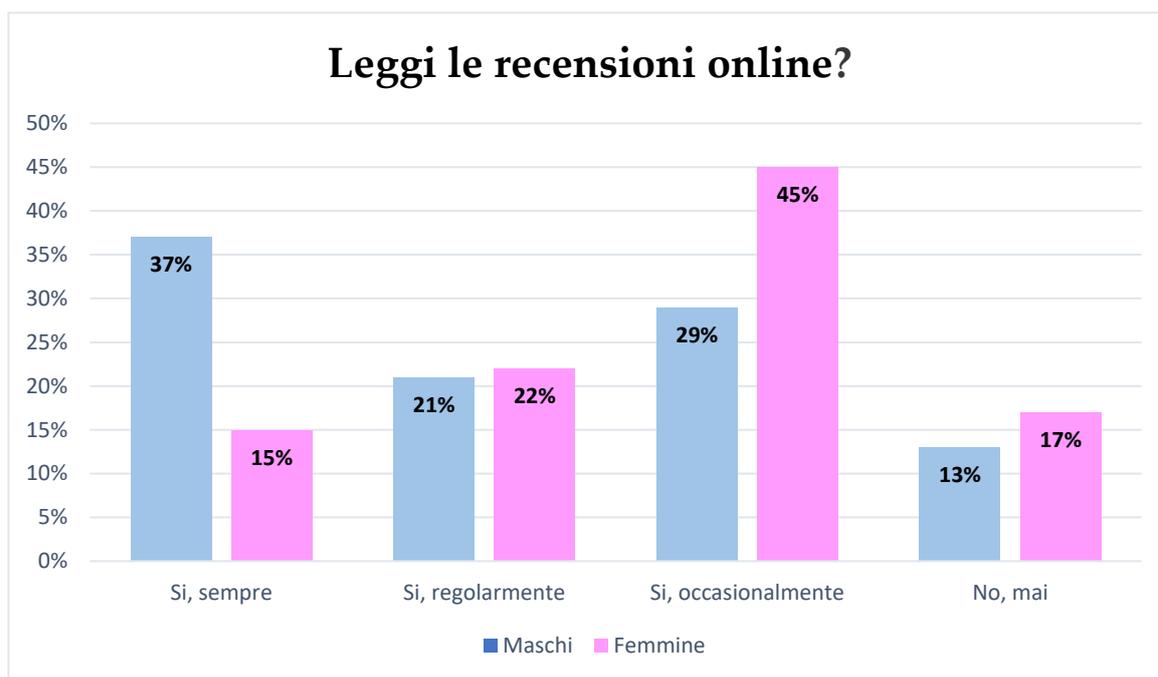
Per testare H2, l'intenzione di prenotare è sempre stata utilizzata come variabile dipendente, mentre il genere dell'ospite come variabile indipendente, la fiducia affettiva e cognitiva come mediatori. I risultati mostrano che, tra le donne, un host dello stesso sesso può indurre fiducia affettiva che influenzerà positivamente l'intenzione di prenotare. Tuttavia, un tale risultato non è stato osservato tra i consumatori di sesso maschile. La congruità di genere sembra influenzare principalmente la fiducia affettiva dei consumatori rispetto a quella cognitiva, supportando solo parzialmente H2.

In conclusione, gli autori si focalizzano sulle conseguenze in ambito manageriale derivanti da queste considerazioni. In primo luogo, dal momento che i guest di sesso femminile hanno maggiori probabilità di prenotare una proprietà su Airbnb quando l'host è dello stesso sesso, potrebbe essere una buona strategia di marketing mettere in risalto le proprietà di host donne raccomandate da viaggiatrici. In secondo luogo, lo studio suggerisce che la mancanza di fiducia è il motivo per cui i consumatori di sesso femminile mostrano un'intenzione di prenotare inferiore con le proprietà di host uomini. Pertanto, è importante che gli host di sesso maschile dimostrino la loro affidabilità mettendo in evidenza i commenti e le recensioni positive rilasciate da guest di sesso femminile. Inoltre, i padroni di casa di sesso maschile potrebbero prendere in considerazione l'idea di affittare l'intero appartamento, per aumentare la privacy dei propri ospiti e farli sentire più tranquilli.

6.2 Uomini e donne valutano le recensioni online in modo diverso?

L'articolo pubblicato da Jamie Pitman il 14 dicembre 2018 si focalizza sui diversi comportamenti degli uomini e delle donne di fronte alle recensioni di una determinata piattaforma. Queste considerazioni possono essere considerate valide anche per la piattaforma sotto esame, ovvero Airbnb.

La prima analisi si focalizza sul tasso di lettura delle recensioni: mentre una proporzione simile di uomini e donne afferma di leggere "regolarmente" le recensioni per le transazioni online, c'è un grande divario tra i due sessi in merito a coloro che leggono "sempre" le review lasciate. Infatti, il 37% degli uomini ha affermato di leggere sempre le recensioni online, ma una percentuale relativamente piccola di donne, il 15%, fa lo stesso, preferendo invece leggere "occasionalmente" le recensioni (Figura 6.2.1).



Figur 6.2.1 Tasso lettura recensioni

Le strategie portate avanti da una piattaforma possono dipendere dal sesso dei propri clienti: se la maggior parte degli utenti sono uomini, è molto importante investire tempo e sforzi per ottenere recensioni con punteggi elevati. Tuttavia, il punteggio in stelle non è l'unico elemento su cui concentrarsi, poiché il sondaggio

riporta che il 40% dei consumatori intervistati non presta attenzione alle recensioni rilasciate oltre due settimane prima.

In generale, la tendenza a recensire da parte di uomini e donne risulta differente (Figura 6.2.2). Solo ad un quarto degli uomini non è mai stato chiesto di recensire, mentre per le donne, la percentuale a cui non è stato richiesto di lasciare una recensione sale al 44%. Se questo dipende dal fatto che gli uomini sono più disponibili a recensire rispetto alle donne, i numeri suggeriscono comunque che qualcosa stia ostacolando la crescita delle recensioni lasciate dalle donne. Un altro aspetto da sottolineare è la differenza tra il numero di uomini e donne che scelgono di lasciare una recensione quando richiesto. Più della metà di tutti i consumatori di sesso maschile intervistati ha lasciato una recensione quando richiesto, dimostrando la volontà di fornire un feedback, mentre una percentuale minore (sebbene non meno significativa) di donne risponde favorevolmente a una richiesta di recensire.

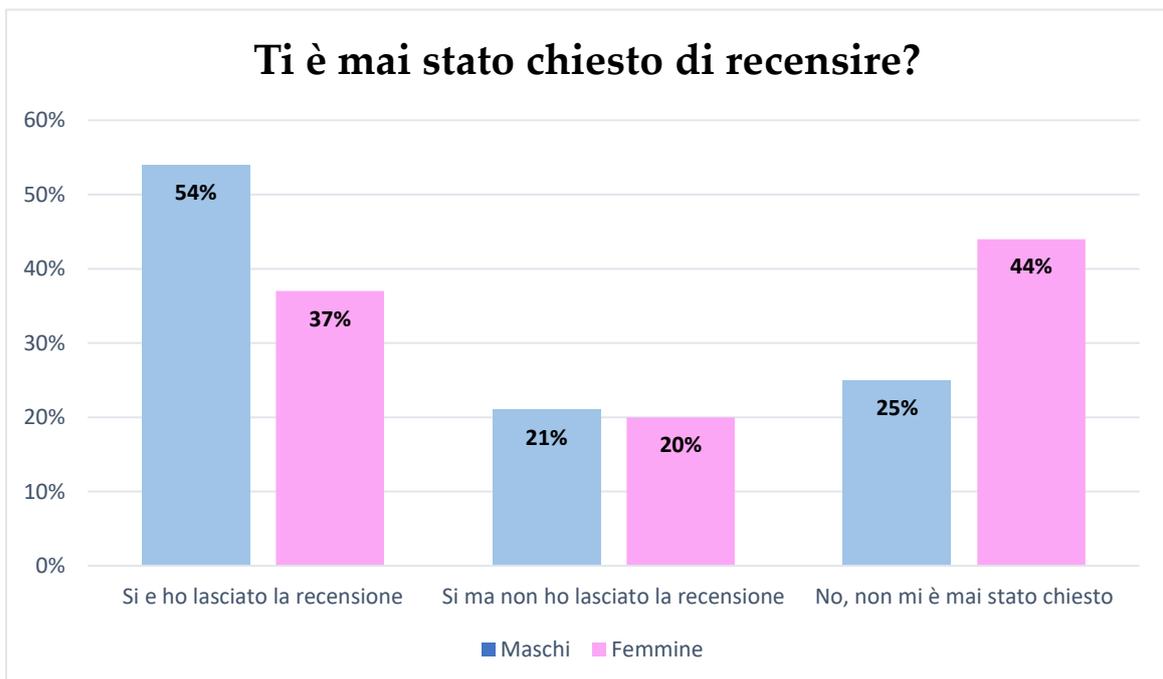


Figura 6.2.2 Richiesta di recensire

Nell'articolo si sottolinea un'ulteriore differenza riguardo le aspettative di una risposta da parte della piattaforma ad una determinata recensione. Il 30% circa sia di uomini che di donne afferma che le recensioni false dovrebbero ricevere risposte. Si osserva invece come gli uomini (in particolare il 58%) ritengono che le

risposte alle recensioni positive siano quelle più importanti, al contrario delle donne, il cui 63% ritiene che siano fondamentali le risposte a quelle negative (Figura 6.2.3).

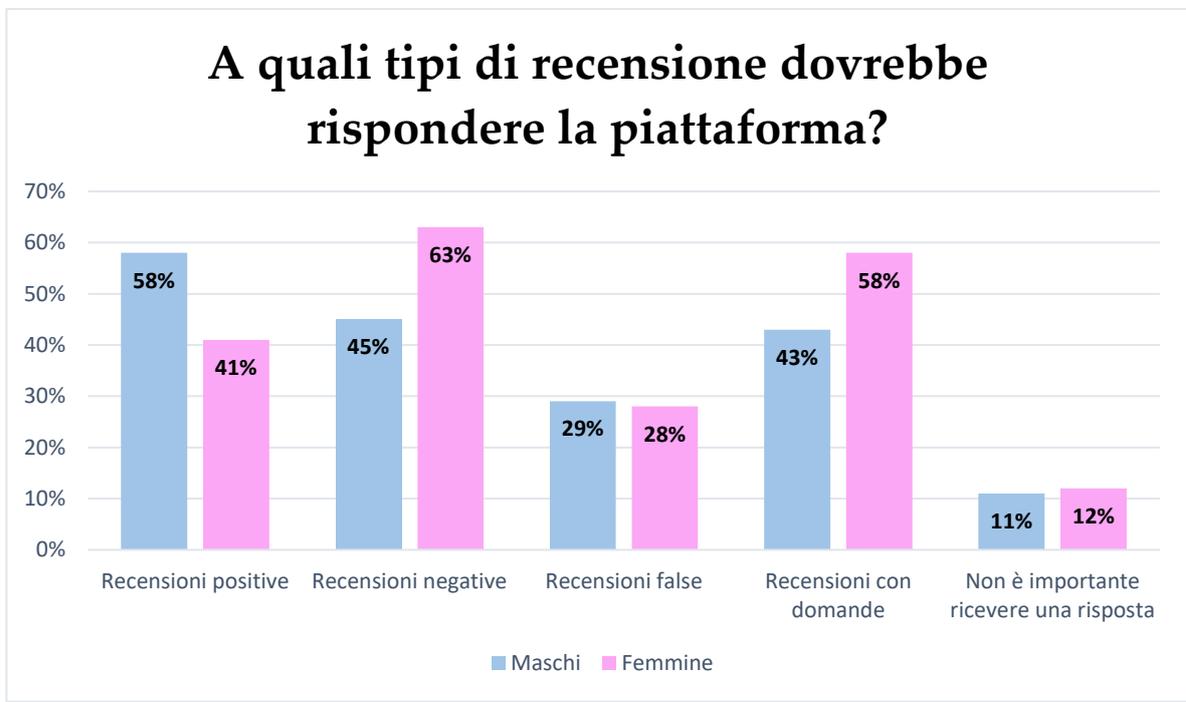


Figura 6.2.3 Risposte alle recensioni

In conclusione, si può affermare che le attitudini di fronte alle recensioni si differenziano in base al sesso dell'utente. Le piattaforme operanti con un sistema di recensione devono tenere in considerazione questo aspetto se vogliono attuare delle strategie per aumentare il numero di transazioni effettuate, per incrementare la fiducia tra le parti ed il tasso di soddisfazione degli utenti.

7. Costruzione Database

Nel presente capitolo saranno descritti i passaggi eseguiti per la costruzione del database necessario per le successive analisi. I dati sono stati scaricati dal sito di sorveglianza "Insider Airbnb", lanciato da Murray Cox nel 2016. In particolare, i file utilizzati per la definizione delle variabili sono stati "listings.csv.gz" e "reviews.csv", relativi alla città di Bologna.

Inizialmente è stato costruito il database per gli host partendo dal file "listings.csv.gz", trasformato in formato xls. Dal momento che il file risultava difficilmente leggibile, attraverso l'ausilio di un programma scritto in linguaggio C++, è stato possibile ordinare le singole informazioni. Infatti, molti dati erano riportati nella medesima cella ed era necessario suddividerli tra le varie colonne per poterli studiare. Il processo utilizzato per la riorganizzazione del database è riassumibile nelle seguenti fasi dell'algoritmo:

1. si processa una riga per volta;
2. la stringa di testo ottenuta viene suddivisa in tanti token quanti sono i caratteri separatore di campo del file csv ',' così da ottenere una lista di token;
3. si processa la lista di token in modo da individuare quando inizia un nuovo record. Per individuare un record è sufficiente controllare che il primo token della lista sia un numero, in caso contrario significa che il record è iniziato in una riga precedente. Per identificare i campi che compongono un record si sono individuati alcuni tag che hanno un carattere di fine/inizio campo. Se ad esempio un token inizia con 'http' significa che quel token rappresenta un campo intero; se il token termina con doppio apice "" o con "]" significa che il campo che si sta costruendo è terminato. Inoltre, sono stati inseriti alcuni controlli per monitorare la costruzione del record in modo da individuare quando l'allineamento è sbagliato.

Sono anche state eliminate tutte le righe bianche presenti nel file.

Dopo queste prime operazioni, dalle 8226 righe iniziali si sono ottenute 3520 righe.

Allo stesso modo, il file "reviews.csv", utilizzato per il database sui guest e sulle review, è stato prima salvato in formato xls e poi ordinato sempre utilizzando il programma in linguaggio C++. Sono state eliminate le righe bianche e le recensioni in cui al posto delle lettere comparivano simboli. I dati considerati sono inerenti alla fascia temporale 2015-2020, quindi sono state cancellate tutte le recensioni risalenti agli anni precedenti. Inoltre, si è deciso di mantenere i dati del 2020 fino al 31/01/2020, questo perché nel mese di febbraio sono stati individuati i primi pazienti affetti da Covid-19. In questo periodo sono iniziate le varie restrizioni che non hanno permesso di spostarsi e quindi di usufruire della piattaforma Airbnb. I dati relativi a soggiorni successivi alla data prescelta rischiavano di compromettere l'analisi e di generare risultati distorti.

Il database iniziale di 216.681 righe è diventato di 129.844 righe.

Una delle informazioni principali necessarie per questo lavoro di tesi è stata la definizione del sesso degli host e dei guest. Per quanto concerne i primi, il sesso è stato definito manualmente poiché il numero di nomi da analizzare non era particolarmente eccessivo (2160 host). Il sesso dei guest, d'altro canto, è stato identificato mediante un programma scritto in Python. Il codice restituiva cinque differenti alternative: *male*, *female*, *most_male*, *most_female*, *unknown*. Nei casi di *most_male* e *most_female* sono stati sostituiti rispettivamente con *male* e *female*. In un secondo momento invece sono stati analizzati i casi riconosciuti dal programma come *unknown* e manualmente è stato definito il sesso del guest, ove possibile.

Infine, le recensioni sono state tutte tradotte in un'unica lingua: la scelta è ricaduta sulla lingua inglese perché la maggior parte delle reviews erano già scritte in questo modo. Anche per questa operazione è stato utilizzato un programma scritto in Python, riportato in appendice.

7.1 Descrizione dati

Per una miglior chiarezza dei dati si è scelto di costruire due database separati per le recensioni e per i listings.

Il database delle recensioni comprende i seguenti dati:

- codice identificativo del listing presso il quale il guest ha soggiornato;
- codice identificativo del guest;
- nome del guest;
- sesso;
- data della recensione;
- codice identificativo della review;
- numero_parole: mediante l'utilizzo della formula "LUNGHEZZA" di Excel è stato possibile contare il numero di parole all'interno della recensione;
- nome_host_in_recensione: attraverso alcune formule di Excel si è valutato se il guest chiama l'host per nome nella review. Questa informazione permette di valutare il livello di confidenza raggiunto tra guest e host durante il soggiorno. La variabile assume valore 1 se compare il nome dell'host, 0 altrimenti.

Dal momento le recensioni risultavano particolarmente pesanti, è stato deciso di tenerle separate. Si mantiene traccia in un ulteriore file dei codici identificativi del guest e del listings, del sesso del guest e della recensione rilasciata.

Per quanto riguarda il database dei listings, le variabili considerate sono state le seguenti:

- Codice_Listing: codice identificativo del listings;
- Codice_Host: codice identificativo dell'host;
- Nome: nome dell'host;
- Sesso: sesso dell'host, identificato manualmente come spiegato in precedenza;

- Listing_type: i listing si dividono in quattro tipologie: entire home/apt, private room, hotel room, shared room. Questa informazione è particolarmente rilevante per le analisi successive: nel prevedere se c'è una tendenza dei guest a soggiornare in case di host del medesimo sesso, si potrebbe pensare che questo effetto sia più evidente nelle situazioni in cui bisogna condividere la casa e quindi quando c'è più interazione con l'host, rispetto ai soggiorni in cui si ha l'intero appartamento a disposizione e gli incontri con l'host si limitano la maggior parte delle volte esclusivamente al check-in e al check-out.
- Prezzo: prezzo per una notte. Tutte le sistemazioni che registravano un prezzo a notte superiore a 800 euro sono state rimosse;
- Numero_reviews: numero di review rilasciate per il listing;
- Notti_minime: numero di notti minime necessarie per soggiornare in quel listing;
- Disponibilità: giorni all'anno in cui è disponibile il listing sulla piattaforma. Il database è stato ripulito eliminando le righe in cui la disponibilità risultava pari a zero;
- Review_al_mese: numero di review rilasciate in media in un mese;
- Distanza_avg: distanza in km dai principali punti di attrazione. Nel file scaricato da Insider Airbnb per ogni listing è presente il quartiere di Bologna in cui si trova. Su Google Maps è stata calcolata la distanza del quartiere del listing dalle tre principali attrazioni della città: la torre degli Asinelli, Piazza Maggiore e Piazza del Nettuno. È stata fatta la media delle tre distanze e il risultato è stato inserito nel database;
- Host_response_rate: tasso di risposta dell'host;
- Host_is_superhost: booleana, 1 se l'host possiede la qualifica di superhost, 0 altrimenti;
- Host_has_profile_pic: booleana, 1 se l'host possiede una foto profilo, 0 altrimenti

- Host_identity_verified: booleana, 1 se l'identità dell'host è verificata, 0 altrimenti
- Accommodates: massimo numero di persone che l'host può ospitare;
- bedrooms: numero di camere da letto disponibili nel listing;
- letti: numero di letti presenti nel listing;
- bagni: numero di bagni presenti nel listing
- review_scores_rating: media dei punteggi delle recensioni, da 1 a 100;
- review_scores_accuracy: punteggio su precisione dell'host, da 1 a 10;
- review_scores_cleanliness: punteggio su pulizia, da 1 a 10;
- review_scores_checkin: punteggio sul momento del check-in, da 1 a 10;
- review_scores_communication: punteggio sulla comunicazione tra l'ospite e l'host, da 1 a 10;
- review_scores_location: punteggio su posizione listings, da 1 a 10;
- review_scores_value: punteggio su rapporto qualità/prezzo, da 1 a 10.

8. Analisi database

In primo luogo, dal database costruito si può osservare la percentuale di uomini e donne, lato guest e lato host. Il numero di nomi per i quali non si è riusciti a definire il sesso risulta maggiore per i guest rispetto agli host (14% vs 9%), dal momento che chi possiede un appartamento a Bologna da affittare ha tendenzialmente un nome italiano più o meno facilmente identificabile, al contrario dei guest che provengono invece da ogni parte del mondo. Una seconda osservazione riguarda il sesso prevalente: risulta essere quello femminile sia per gli host che per i guest (*Figura 8.1*).

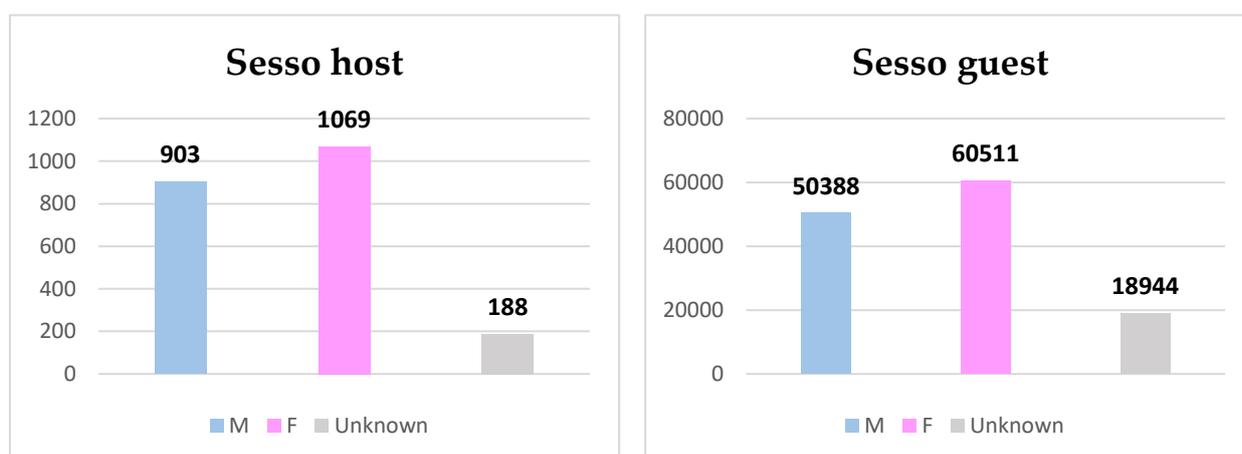


Figura 8.1 Sesso guest e host

8.1 Chi recensisce di più?

Si potrebbe pensare che i guest di sesso femminile siano più propensi a rilasciare reviews successive al loro soggiorno, dal momento che nel database il numero di guest donne è superiore rispetto a quello dei guest uomini. Per capire la solidità di questa affermazione è stato eseguito un test di ipotesi. Per ogni listing, attraverso la costruzione di una tabella pivot, è stato contato il numero di recensioni rilasciate dai guest donna e dai guest uomini. Si vuole verificare se la media delle differenze tra il numero di review scritte dagli uomini e quelle scritte dalle donne possa ritenersi uguale a zero.

Descrizione variabili:

Recensioni donne: numero di recensioni rilasciate da guest donne per il listing;

Recensioni uomini: numero di recensioni rilasciate da guest uomini per il listing;

Diff: differenza per ogni listing del n° di review dei guest donna e dei guest uomo.

$$\left\{ \begin{array}{l} H_0: \mu(\text{Diff}) = 0 \\ H_a: \mu(\text{Diff}) \neq 0 \end{array} \right.$$

La funzione utilizzata su STATA è stata:

Ttest recensioni donne=recensioni uomini

Risultati:

μ Diff	t calc	p-value	Intervallo di confidenza (95%)
3,46	17,05	0,0000	[3,0663 ; 3,8627]

Dal momento che l'intervallo di confidenza non include lo zero e che la statistica calcolata non ricade nella zona di accettazione, si rifiuta l'ipotesi nulla che la media delle differenze tra le recensioni rilasciate dai guest donna e dai guest uomini sia uguale a zero. Esiste quindi una differenza significativa: le donne sembrano recensire di più.

8.2 Aspetti delle recensioni

Come descritto nei capitoli precedenti, uomini e donne sono molto diversi tra loro nel modo di pensare, di vivere le emozioni, di valutare le esperienze vissute. Quando si viaggia e si affitta un alloggio, probabilmente saranno diverse le impressioni che avranno i guest uomo e i guest donna, come diversi saranno gli aspetti che vorranno sottolineare nell'istante in cui rilasciano una recensione. Sarà probabilmente diverso anche il modo di recensire. Si è voluto valutare se, in media, le donne rilasciano review più lunghe rispetto agli uomini. Per far ciò, si è eseguito un test di ipotesi sul numero di parole nelle recensioni.

Descrizione variabili:

Numero_parole: n° di parole di una review;

Sesso: sesso dei guest, F → femmina, G → maschio. In media le review delle donne contengono 279 parole, quelle degli uomini 239. Si vuole testare se c'è una differenza significativa tra questi valori.

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest numero_parole, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
279,20	238,76	29,93	0,0000	[38,7277 ; 44,1561]

Con un p-value inferiore al 5% si rifiuta l'ipotesi nulla H_0 e si giunge alla prima considerazione: le recensioni delle donne sono più lunghe.

In secondo luogo, è stata svolta un'analisi per capire quali sono gli aspetti che emergono di più nelle recensioni dei guest e capire se ci sono differenze tra uomini e donne. Le donne citano la pulizia più degli uomini? Per gli uomini l'elemento più importante è la posizione? Per rispondere a queste domande è stata creata una funzione macro su Excel per separare le parole delle recensioni in celle singole. Questo procedimento è stato eseguito prima per le recensioni delle donne e poi filtrando per le review rilasciate dagli uomini. La funzione macro utilizzata è riportata in appendice. Successivamente, attraverso un programma scritto in linguaggio Python (riportato in appendice), è stato possibile calcolare per ogni parola il numero di occorrenze nelle recensioni. Il programma restituiva le parole in ordine decrescente, dalla più utilizzata a quella meno utilizzata.

La parola più utilizzata da guest donne e guest uomini risulta essere "and". Eliminando tutte le preposizioni, i pronomi e le congiunzioni, per entrambi i sessi

sono state selezionate le prime 28 parole più utilizzate e i risultati sono riportati in

Tabella 8.2.1:

Parole	Occorrenze guest uomini	Parole	Occorrenze guest donne
Apartment	17337	Apartment	20582
Bologna	15895	Bologna	19488
Place	12020	Place	14786
Great	11388	Great	14051
Location	9757	Clean	11763
Host	9377	Host	11365
Clean	9022	Location	11080
Nice	7008	Nice	8498
City	6869	City	8367
Recommend	6043	Recommend	7390
Good	5045	Good	6295
Perfect	4940	Walk	5985
Walk	4877	Perfect	5982
Comfortable	4617	Comfortable	5628
Helpful	4145	Centre	5037
Centre	4036	Helpful	4973
Close	3813	Close	4790
Great	3659	Great	4336
Easy	3392	Easy	4186
Lovely	3206	Lovely	4126
Super	3055	Restaurants	3895
Restaurants	3032	Super	3802
Quiet	2900	Quiet	3569
Station	2964	Station	3476
Beautiful	2569	Beautiful	3337
Bus	2491	Friendly	2788
Friendly	2263	Train	2786
Spacious	2147	Spacious	2513

Tabella 8.2.1 Parole più utilizzate nelle review

Per effettuare un confronto più puntuale, si è calcolata la percentuale di utilizzo di ogni singola parola, cioè si è diviso il numero di occorrenze del termine in analisi per il numero totale di parole delle recensioni. Per le donne le parole totali presenti nelle recensioni sono pari a 1.960.243,00, per gli uomini invece 1.669.056,00.

Le percentuali ottenute per guest di sesso maschile e guest di sesso femminile sono rappresentate rispettivamente in *Figura 8.2.1* e in *Figura 8.2.2*

Recensioni guest uomini

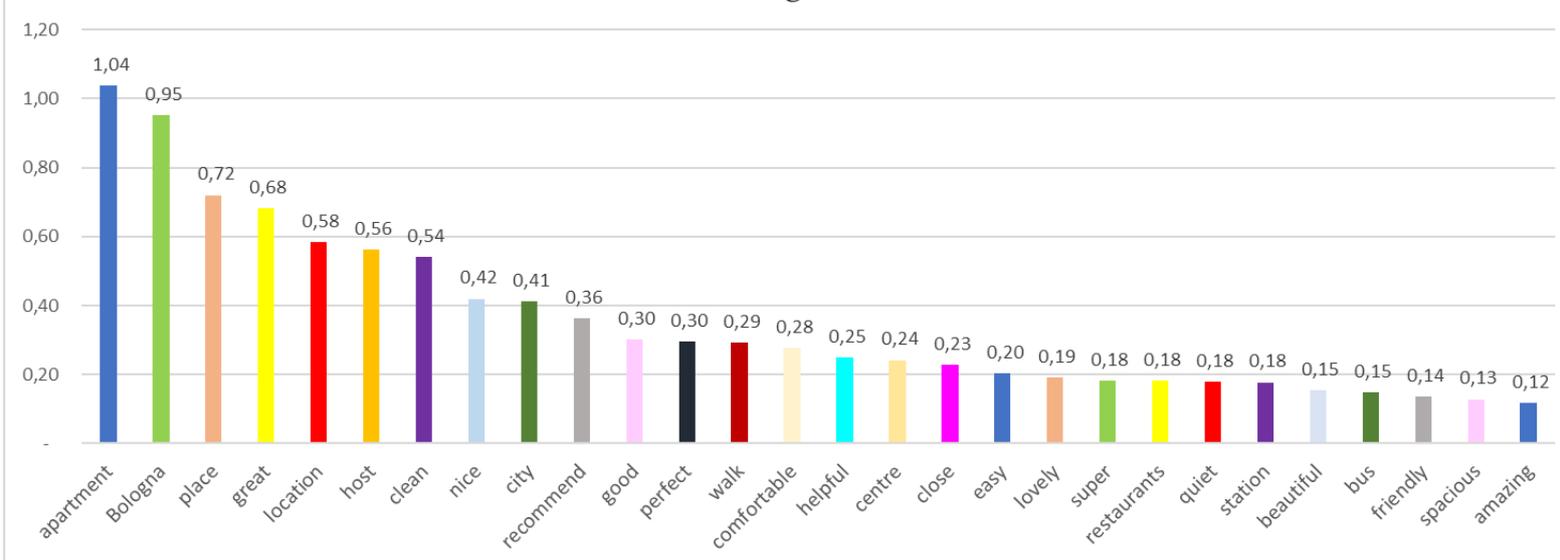


Figura 8.2.1 Parole guest uomini

Recensioni guest donne

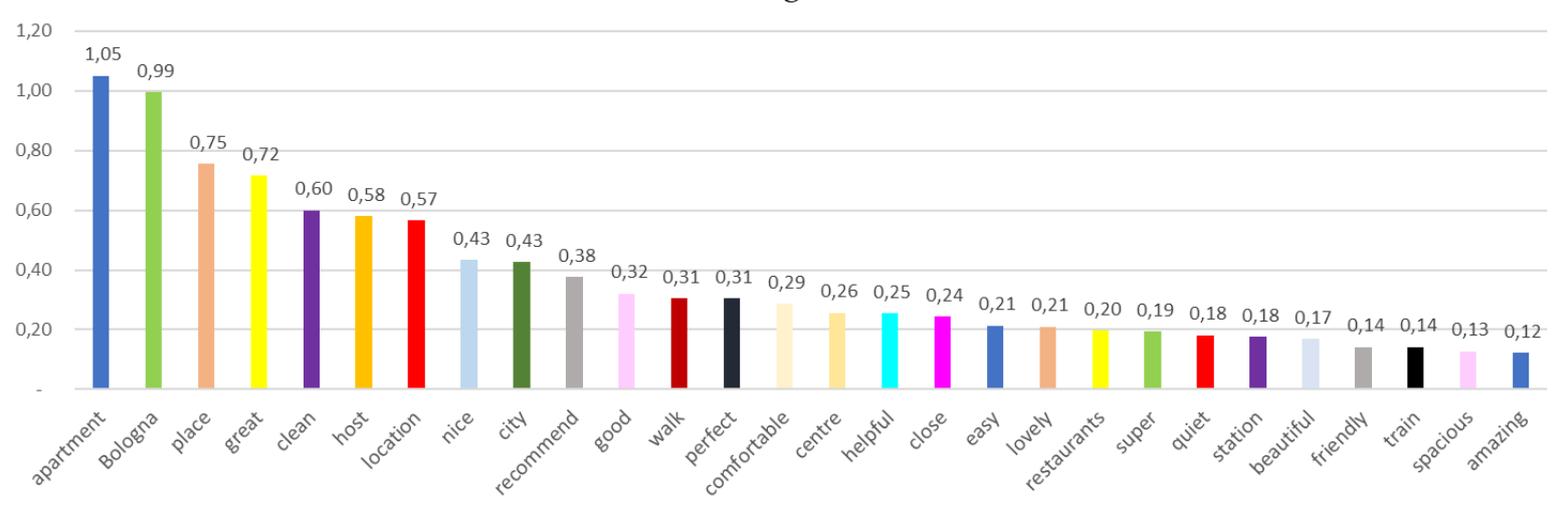


Figura 8.2.2 Parole guest donne

Le prime quattro parole più utilizzate sono le medesime: si cita l'appartamento, quindi ci saranno commenti positivi e/o negativi relativi al listing, viene nominata la città in cui si soggiorna e, più in generale, il luogo; come aggettivo più utilizzato compare il termine "great". A questo punto si può evidenziare la prima differenza: per le donne la quinta parola più utilizzata è la pulizia, per gli uomini la location. Quindi i due sessi probabilmente assegnano un'importanza diversa a questi due aspetti. Procedendo con l'analisi si osserva che sia uomini che donne lasciano

commenti sull'host, sulla città, sulla raccomandazione del listing nella recensione. Anche gli aggettivi utilizzati per descrivere la propria esperienza sono i medesimi: nice, good, perfect, comfortable, helpful, close, centre, easy, lovely, super, quite. Alcuni di questi si riferiscono probabilmente all'host, altri al listing (come il fatto che sia confortevole e vicino al centro), altri ancora al soggiorno in generale. Le donne usano la parola "walk" di più rispetto agli uomini: probabilmente sono più interessate al fatto che una città possa essere vissuta a piedi.

Sia uomini che donne, dopo aver descritto l'appartamento e l'host, citano elementi come ristoranti e stazioni, quindi è importante per un guest che il listing sia vicino a questi luoghi per migliorare il soggiorno. Una curiosità riguarda il mezzo di trasporto citato: mentre gli uomini parlano di bus, le donne invece scrivono di più la parola train nelle loro recensioni. Questo spinge a riflettere sul fatto che è diverso il modo in cui ci si preferisce spostare per uomini e donne.

Infine, probabilmente sempre nell'ottica di descrivere il soggiorno viene utilizzato il termine amazing, mentre una caratteristica importante del listing che viene citata spesso nelle recensioni da parte di entrambi i sessi è la sua spaziosità e quindi quanto è vivibile.

Questa analisi è interessante perché evidenzia alcune differenze sugli aspetti che i due sessi preferiscono commentare nelle loro recensioni, ma mostra anche che gli aggettivi utilizzati sono più o meno gli stessi e che le review si concentrano su tre macro-aree: listing, esperienza in generale e host. Inoltre, non compaiono termini con accezione negativa e tutti gli aggettivi più utilizzati sono positivi, quindi sembrerebbe che un numero importante di persone che hanno soggiornato in un listing a Bologna si siano trovate bene e abbiano vissuto un'esperienza piacevole.

8.3 Confidenza

Nelle review rilasciate dopo un soggiorno si commenta spesso l'host, la sua disponibilità, la sua gentilezza e la sua attenzione per i guest. Per capire il grado di confidenza che si crea tra host e guest è stata costruita una variabile che indicasse se il guest nella recensione si rivolge all'host chiamandolo per nome. Per ogni listing è stata calcolata la percentuale delle recensioni in cui l'host viene citato per nome rispetto alle review totali del listing. Le frequenze sono rappresentate in *Figura 8.3.1*.

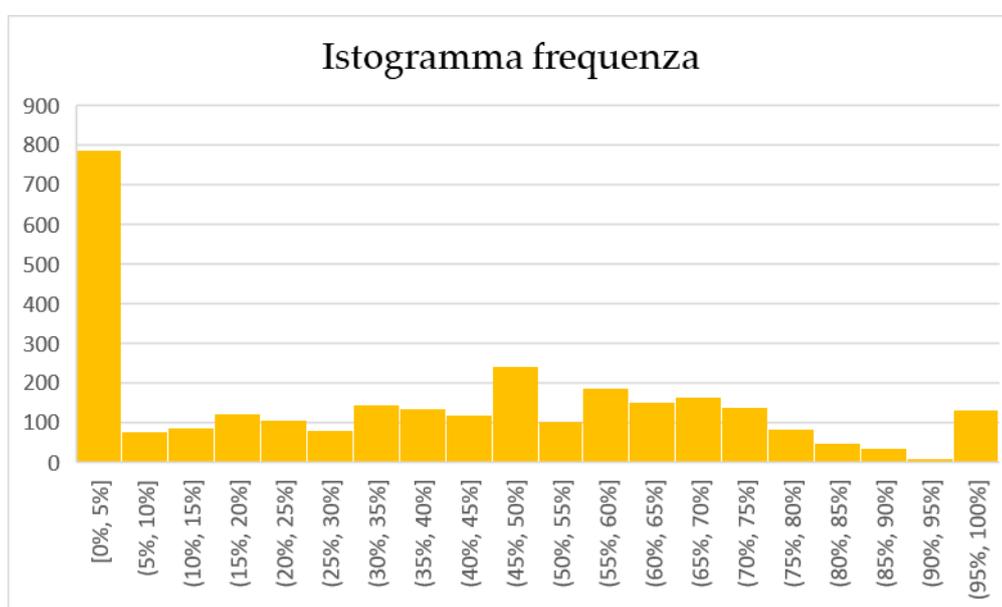


Figura 8.3.1 Frequenza listings

La maggior parte dei listing ricade nella classe dove la percentuale delle recensioni in cui compare il nome dell'host è compresa tra lo 0 e il 5%. Sono molti quindi le sistemazioni che raccolgono poche recensioni in cui si è creata una certa confidenza tra il proprietario del listing e l'ospite.

Si vuole indagare se esiste una differenza significativa tra la percentuale delle recensioni in cui si cita l'host per nome nel caso in cui quest'ultimo sia donna o sia uomo.

Descrizione variabili:

Recensioni_con_nome: rapporto tra review in cui ci si rivolge all'host con il suo nome sulle review totali.

Sesso: sesso dell'host, F → femmina, G → maschio. Il test di ipotesi è il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest recensioni_con_nome, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
0,496	0,439	2,837	0,0046	[0,01760 ; 0,09637]

Con un p-value inferiore al 5% si rifiuta l'ipotesi nulla. Esiste quindi una differenza significativa tra i due sessi: gli host donna ricevono in media più recensioni in cui sono chiamate per nome, quindi creano probabilmente un legame più forte con le persone che ospitano nelle loro sistemazioni.

8.4 Host di Aibnb come lavoro

“The gender employment gap has stagnated over the past few years, and women are still less likely to be employed than men”

Eurostat

L'uguaglianza di genere è uno dei principali obiettivi che i Paesi dell'Unione Europea cercano di perseguire ormai da anni. Purtroppo, questa condizione risulta essere ancora lontana da raggiungere, in particolar modo nel mondo del lavoro. Il divario tra occupazione maschile e occupazione femminile è alto, come emerge dall'ultimo report pubblicato da Eurostat a maggio 2020 (*Figura 8.4.1.*). Analizzando i dati raccolti si osserva come la situazione delle donne è ancora fortemente svantaggiata anche nel nostro Paese, il cui divario risulta essere maggiore rispetto

alla media europea. Una ragione alla base di questa differenza risiede nella difficoltà di conciliare gli impegni di lavoro con quelli familiari. Il divario del tasso di occupazione aumenta infatti con il numero di figli e sono molte le donne che lavorano part time rispetto agli uomini.

Nonostante l'impegno delle Nazioni Unite per ridurre al minimo il divario di occupazione, la strada è ancora lunga.

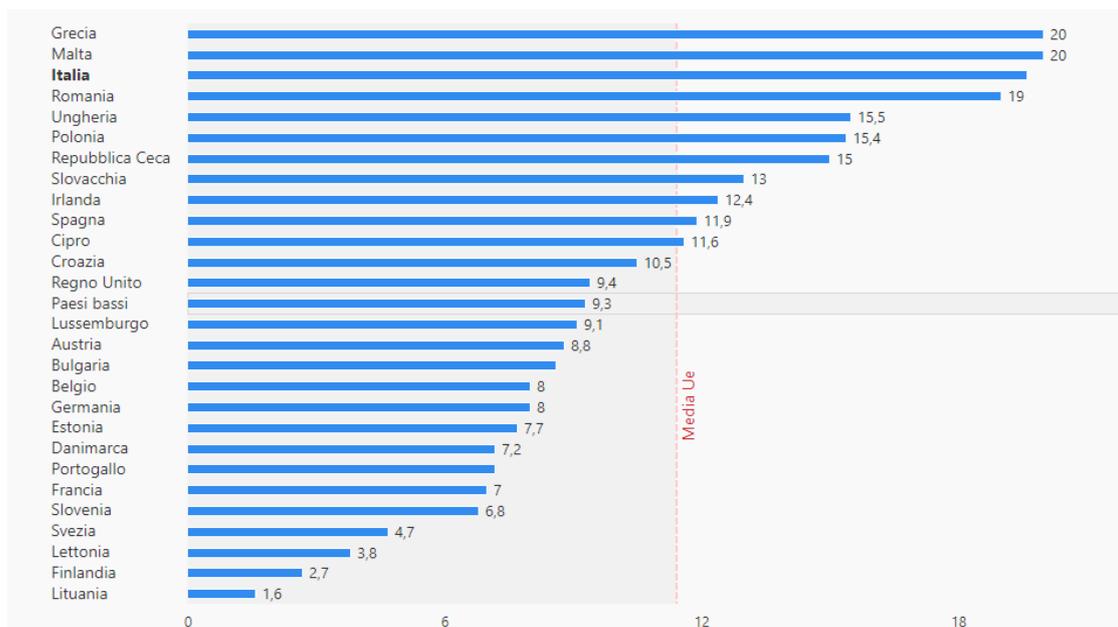


Figura 8.4.1 Divario tra tasso di occupazione maschile e femminile (Fonte Eurostat)

La situazione sopra descritta offre diversi spunti di riflessione. Le donne devono organizzare le loro giornate per conciliare tutti i loro vari impegni. Ma allora, possedere uno o più appartamenti su Airbnb potrebbe essere considerato un lavoro solido? Questa soluzione permetterebbe loro di guadagnare ma anche di avere del tempo a disposizione per occuparsi della famiglia.

In questo paragrafo l'obiettivo è quello di indagare se per le donne essere un host su Airbnb può essere un lavoro, quindi se una donna può essere imprenditrice di sé stessa attraverso questa piattaforma di sharing economy. Si è valutato il guadagno minimo e il guadagno massimo che una donna potrebbe ottenere affittando un appartamento su Airbnb. Nel peggiore dei casi, un appartamento

inserito nella piattaforma è affittato sicuramente tante volte quante sono le recensioni rilasciate dal guest: il numero di review rilasciate al mese viene considerato come tasso di domanda. Se un listing in media al mese riceve 3 recensioni, sicuramente ci saranno stati almeno tre soggiorni, più ovviamente tutti quelli non recensiti, di cui però non si conosce il numero. Questi 3 soggiorni devono perdurare almeno per le notti minime indicate dall'host. Tornando all'esempio, se un listing ha in media 3 recensioni e le notti minime richieste sono 2, si può dedurre che è rimasto affittato per un minimo di sei giorni quel mese. Ragionando in questo modo, sui listing appartenenti alle sole host donne si è visto come il tasso di recensione medio al mese per i soggiorni nella città di Bologna sia di 1,3 recensioni/mese, valore molto basso. Il numero medio di notti minime necessarie per soggiornare in un determinato appartamento è in media 2,3 notti. Considerando che, in media, le host donne posseggono circa 1,4 appartamenti registrati sulla piattaforma e che il prezzo medio per notte è 85,82€, il guadagno al mese per una donna risulta essere pari a $2,3 * 1,3 * 1,4 * 85,82 = 410,6$ €/mese, ovvero 4927 €/anno. Il risultato ottenuto sicuramente non è veritiero sui guadagni realmente ottenuti dalle host di sesso femminile dal momento che è stata utilizzata come proxy del tasso di occupazione di un appartamento il numero di recensioni al mese ed è noto che non tutti i guest decidono di rilasciare una review al termine di un soggiorno. Inoltre, il risultato ottenuto non rappresenta il vero guadagno dell'host perché dovrà essere depurato dalle spese sostenute dall'host già comprese nel prezzo di affitto del listing, come la pulizia al termine del soggiorno.

Lo scenario descritto può essere considerato come *worst case*. Concentrandosi invece sullo studio del *best case*, l'analisi si sposta sulla disponibilità dell'appartamento durante l'intero anno. Valutando la disponibilità media di un listings a Bologna, pari a 291 giorni all'anno, se questo fosse effettivamente affittato per tutti i giorni in cui è disponibile si otterrebbe un guadagno pari a $291 * 85,52 = 24973$ €/anno per ogni listing. Se in media le host donne possiedono 1,4 listings, il guadagno sale a circa 40 mila €/anno. Alcune considerazioni vengono spontanee anche in questo scenario: è

molto raro che un appartamento sia affittato per tutta la sua effettiva disponibilità, soprattutto nel periodo attuale che stiamo vivendo, con una pandemia in atto, in cui tra un soggiorno e l'altro Airbnb impone all'host di sanificare l'appartamento e quindi lasciare che intercorrano almeno 24 ore prima di accogliere un guest diverso. Si può dedurre da questi ragionamenti che il risultato sia nel mezzo tra il best case e il worst case. In base ai risultati ottenuti possedere un appartamento a Bologna da affittare su Airbnb potrebbe non essere la soluzione ideale per sostituire effettivamente un lavoro. Sicuramente potrebbe essere un'integrazione rispetto allo stipendio ottenuto. Tutto dipende se la soluzione all'analisi si avvicina di più al guadagno minimo calcolato, quindi se le recensioni possono effettivamente essere considerate una buona proxy del tasso di occupazione e in questo scenario i ricavi ottenuti non sarebbero sufficienti per sopravvivere; al contrario, se gli appartamenti fossero affittati per un numero di giorni che si avvicina alla disponibilità massima, l'importo sarebbe in questo caso sufficiente per poter definire questo un vero e proprio lavoro.

9. Analisi descrittive

In questo capitolo verranno mostrate le statistiche descrittive delle principali variabili di interesse, utilizzate successivamente nel capitolo seguente per le analisi di regressione. Inoltre, verrà eseguita un'analisi su queste variabili per capire le differenze presenti tra uomini e donne, sia lato guest che lato host. Per far ciò si eseguiranno una serie di test di ipotesi su STATA per verificare se esiste una differenza significativa tra i valori assunti dalle variabili per i due sessi.

9.1 Prezzo

La variabile prezzo presenta le seguenti statistiche descrittive:

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
85,82	73,52	9	700	65

Si è voluta studiare se vi è una differenza significativa per il prezzo fissato per un listing su Airbnb da host donne e host uomini. In media, i prezzi delle sistemazioni di host donne possono essere considerati pari a quelli fissati dagli host uomini? Per rispondere a questa domanda è stato eseguito un test di ipotesi sulla differenza tra il prezzo medio proposto da proprietari di sesso maschile e femminile.

Descrizione variabili:

Prezzo: prezzo per il listing presente sulla piattaforma;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest prezzo, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
82,18	85,26	-1,18	0,2369	[-8,1785 ; 2,0237]

La statistica calcolata cade nella zona di accettazione e il p-value è superiore al 5%: non si hanno sufficienti ragioni per rifiutare l'ipotesi nulla per cui la differenza dei prezzi medi può ritenersi pari a zero. Sotto questo punto di vista non c'è alcuna differenza di genere, le donne hanno le stesse opportunità di guadagno dei loro colleghi host uomini.

9.2 Review al mese

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
1,30	1,67	0	14	1

Il numero di review al mese registrate per un determinato listing può essere utilizzato come proxy della domanda di una determinata sistemazione. Infatti, non si hanno dati su quanto effettivamente una sistemazione sia affittata, ma sicuramente lo è tante volte quante sono le review rilasciate in quel mese. Bisogna però considerare che non tutti i guest decidono di recensire il loro soggiorno quindi le review che si hanno a disposizione non rappresentano tutti gli affitti e quindi l'effettiva domanda.

Per capire se esiste una differenza significativa tra la domanda di listings appartenenti a host donne e host uomini si esegue un test di ipotesi.

Descrizione variabili:

Review_al_mese: numero di review al mese rilasciate per un determinato listing;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest review_al_mese, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
1,329	1,403	-1,1248	0,2608	[-0,20215 ; 0,054776]

La statistica calcolata ricade nella zona di accettazione, pertanto non vi sono sufficienti ragioni per rigettare l'ipotesi nulla. I listings di host donne e host uomini presentano la medesima domanda.

9.3 Punteggi

Per i punteggi, le statistiche descrittive sono le seguenti:

	Media	Deviazione std	Valore minimo	Valore massimo	Mediana
Review_Score_Rating	94	9,32	10	100	96
Review_Scores_Accuracy	9,68	0,76	0	10	10
Review_Scores_Cleanliness	9,58	0,85	0	10	10
Review_Scores_Checkin	9,81	0,63	0	10	10
Review_Scores_Communication	9,79	0,69	0	10	10
Review_Scores_Location	9,58	0,74	0	10	10
Review_Scores_Value	9,36	0,93	0	10	10

Si vuole indagare se esiste una differenza significativa tra i punteggi delle review di host uomini e host donne. Sono stati eseguiti dei test di ipotesi sul punteggio globale (review score rating) e sui punteggi assegnati alle singole voci come check-in o posizione listing. Il comando utilizzato su STATA per i singoli test di ipotesi è stato

ttest review_score, by (sesso)*

Descrizione variabili:

Review score rating: media dei punteggi delle recensioni;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
94,45	94,08	-2,095	0,0362	[0,04764 ; 1,4381]

L'intervallo di confidenza non include lo zero e la statistica calcolata non ricade nella zona di accettazione. Si rifiuta l'ipotesi nulla per cui la media dei punteggi tra host uomini e host donne sia uguale. In media le donne presentano punteggi più alti.

Descrizione variabili:

Review score accuracy: punteggio su precisione dell'host;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
9,75	9,66	3,035	0,0024	[0,03389 ; 0,15763]

Il p-value è inferiore al 5%: si rifiuta l'ipotesi nulla per cui il punteggio sulla precisione dell'host sia uguale per uomini e donne. Le host donne registrano punteggi più alti per questa voce, risultando quindi più precise.

Descrizione variabili:

Review score cleanliness: punteggio su pulizia del listing;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
9,66	9,54	3,468	0,0005	[0,05366; 0,19329]

Con un p-value di 0,0005 si rifiuta l'ipotesi nulla: anche per questa voce le donne presentano punteggi più alti e la differenza è significativa.

Descrizione variabili:

Review score check-in: punteggio sul check in;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
9,86	9,79	2,540	0,0111	[0,01514 ; 0,11764]

L'intervallo di confidenza non include lo zero, pertanto si rifiuta l'ipotesi nulla. I punteggi più alti per il check-in sono assegnati alle host di sesso femminile.

Descrizione variabili:

Review score communication: punteggio sull'efficacia della comunicazione con l'host;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
9,85	9,78	2,564	0,0104	[0,01638 ; 0,12301]

Come nei casi precedenti, si rifiuta l'ipotesi nulla dal momento che il p-value è inferiore al 5%. La comunicazione con host donne è più efficace rispetto a quella con host uomini.

Descrizione variabili:

Review score location: punteggio su posizione listing;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
9,56	9,60	-1,239	0,2154	[-0,10265 ; 0,02316]

La statistica calcolata ricade nella zona di accettazione, pertanto non vi sono sufficienti ragioni per rifiutare l'ipotesi nulla: non c'è una differenza significativa tra i punteggi sulla location del listing registrati da host donne e host uomini.

Descrizione variabili:

Review score value: punteggio su rapporto qualità/prezzo;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
9,42	9,36	1,6643	0,0962	[-0,01155 ; 0,14114]

L'intervallo di confidenza contiene lo zero e la statistica calcolata ricade nella zona di accettazione. Non vi sono sufficienti ragioni per rifiutare l'ipotesi nulla. Per questa voce quindi non esiste una differenza significativa tra i punteggi delle host donne e degli host uomini.

In sintesi, i punteggi medi registrati nelle singole categorie da host uomini e host donne sono rappresentati in *Figura 9.3.1*.

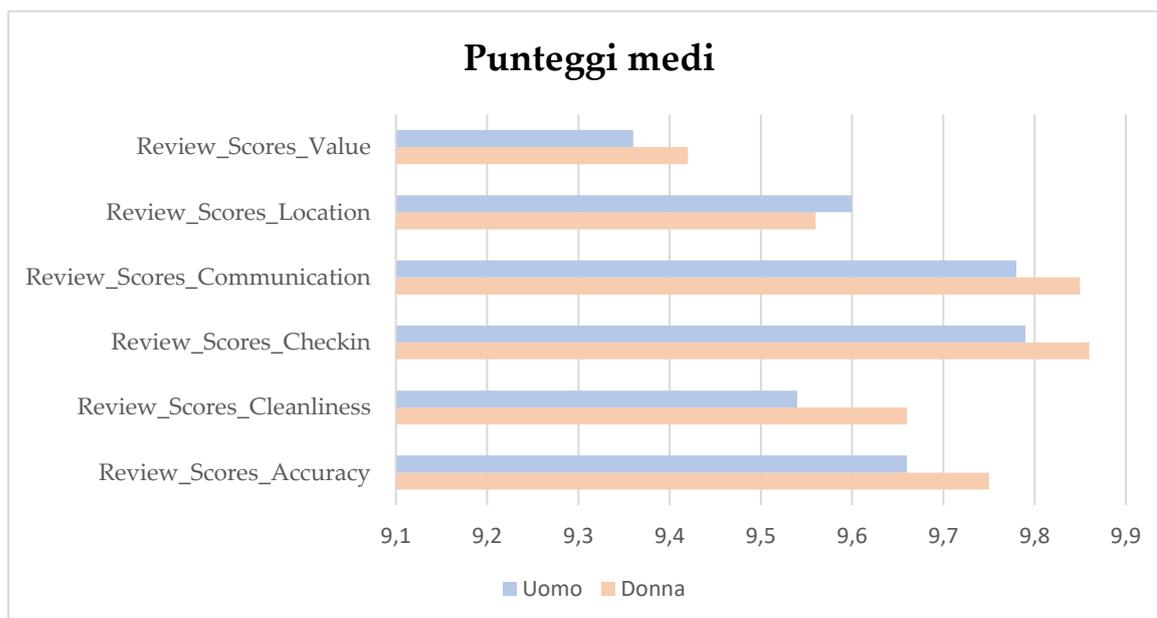


Figura 9.3.1 Punteggi medi host uomini e host donne

9.4 Disponibilità

La statistica descrittiva della variabile disponibilità è la seguente:

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
200,96	122,80	1	365	179

Esiste una differenza significativa tra i giorni in cui un listing è disponibile per gli host uomini e gli host donne? Si esegue un test di ipotesi per rispondere a questa domanda.

Descrizione variabili:

Disponibilità: giorni in cui il listing è disponibile sulla piattaforma per poter essere affittato;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest disponibilità, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
201,59	202,75	-0,2554	0,7985	[-10,0268; 7,7161]

L'intervallo di confidenza contiene lo zero, pertanto non vi sono sufficienti ragioni per rifiutare H_0 . In media la disponibilità dei listing degli host donna e degli host uomini è la medesima

9.5 Numero reviews

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
38,10	67,20	0	640	11

Per ogni listing è indicato il numero di review che può essere approssimato al tasso di domanda. Si vuole indagare se esiste una differenza significativa tra la media delle review rilasciate alle host donne e quelle agli host uomini.

Descrizione variabili:

Numero_review: numero di review per ogni listing;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest numero_review, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
41,51	39,67	0,7163	0,4738	[-3,20764 ; 6,90052]

L'intervallo di confidenza non include lo zero, pertanto non vi sono sufficienti ragioni per rifiutare l'ipotesi nulla. In media host uomini e host donne registrano lo stesso numero di recensioni per il loro listing. Questo risultato conferma quello precedentemente ottenuto per la variabile review_al_mese: la domanda è la medesima per uomini e donne.

9.6 Host response rate

Il tasso di risposta dell'host è una variabile importante per comprendere il comportamento da parte dei due sessi, il grado di attenzione per le persone che ospitano nei loro listing.

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
0,97	0,18	0	1	1

Esiste una differenza significativa tra il tasso di risposta di host uomini e host donne?

Descrizione variabili:

Host response rate: tasso di risposta dell'host;

Sesso: sesso dell'host, M→maschio, F→ femmina. Il test di ipotesi eseguito è stato il seguente:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato su STATA è stato il seguente:

ttest host_response_rate, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
0,965	0,959	0,7095	0,4781	[-0,01165 ; 0,02486]

In primo luogo, si osserva che sia gli host donna che gli host uomini presentano un tasso di risposta molto alto, vicino al 100%. La differenza tra i due sessi inoltre non risulta significativa, in quanto la statistica calcolata ricade nella zona di accettazione.

9.7 Listing_type

Nel database sono state individuate quattro diverse tipologie di listing:

- Entire home/apartment;
- Private room;
- Hotel room;
- Shared room.

Per quanto concerne la città di Bologna, su 3519 listing appartenenti a host di sesso maschile, femminile e unknown, ne risultano 2431 del tipo Entire home/apartment, 997 Private room, 61 Hotel room e 30 Shared room. (Figura 8.7.1)

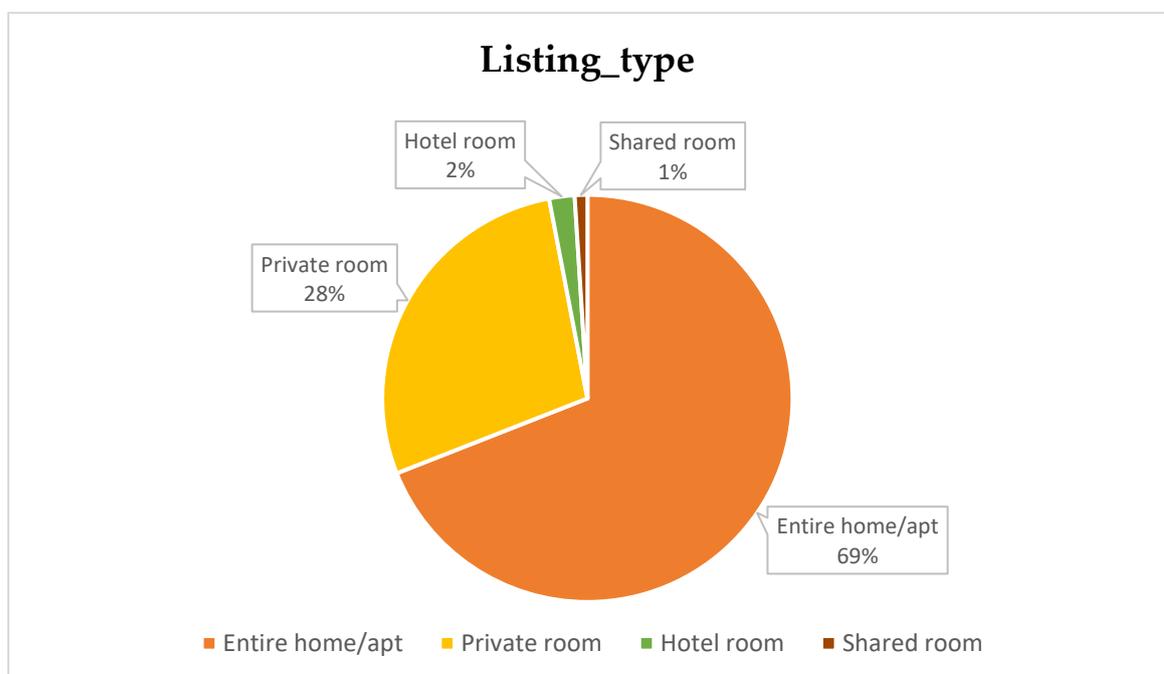


Figura 8.7.1 Suddivisione tipologie listings

Entrando più nel dettaglio nell'analisi, per gli host di sesso femminile la suddivisione dei listing è la seguente (Figura 8.7.2):

Entire home/apt	1010
Private room	444
Hotel room	22
Shared room	8

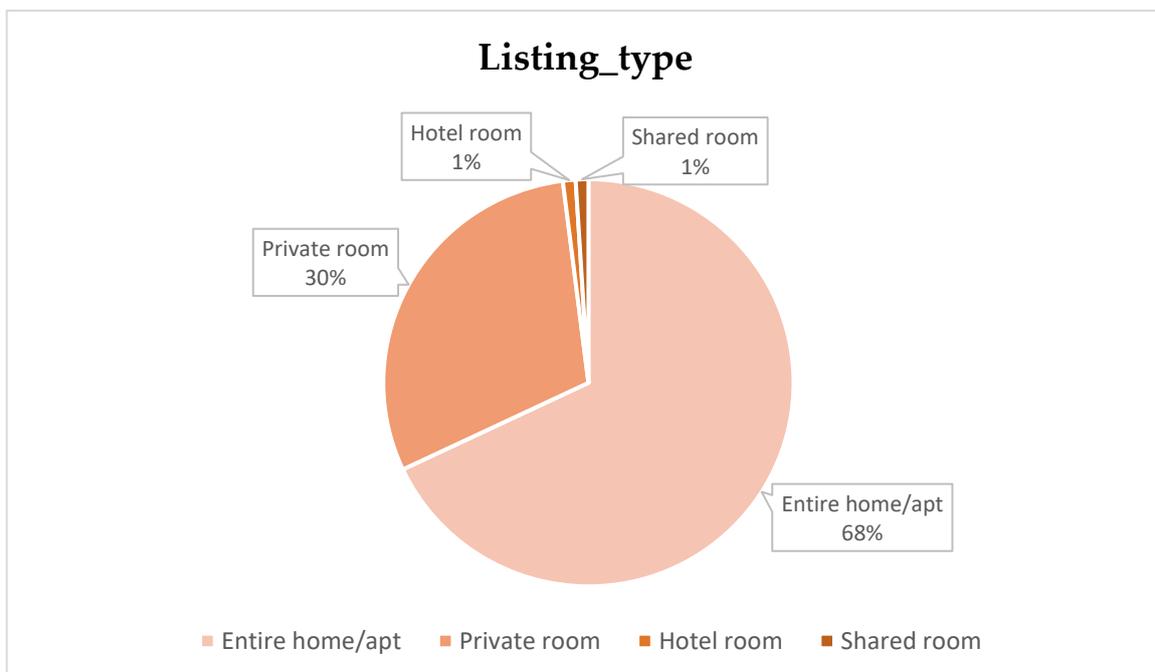


Figura 8.7.2 Suddivisione tipologia listing per host donne

I listing appartenenti ad host di sesso maschile invece sono così distribuiti (Figura 8.7.3):

Entire home/apt	965
Private room	443
Hotel room	32
Shared room	18

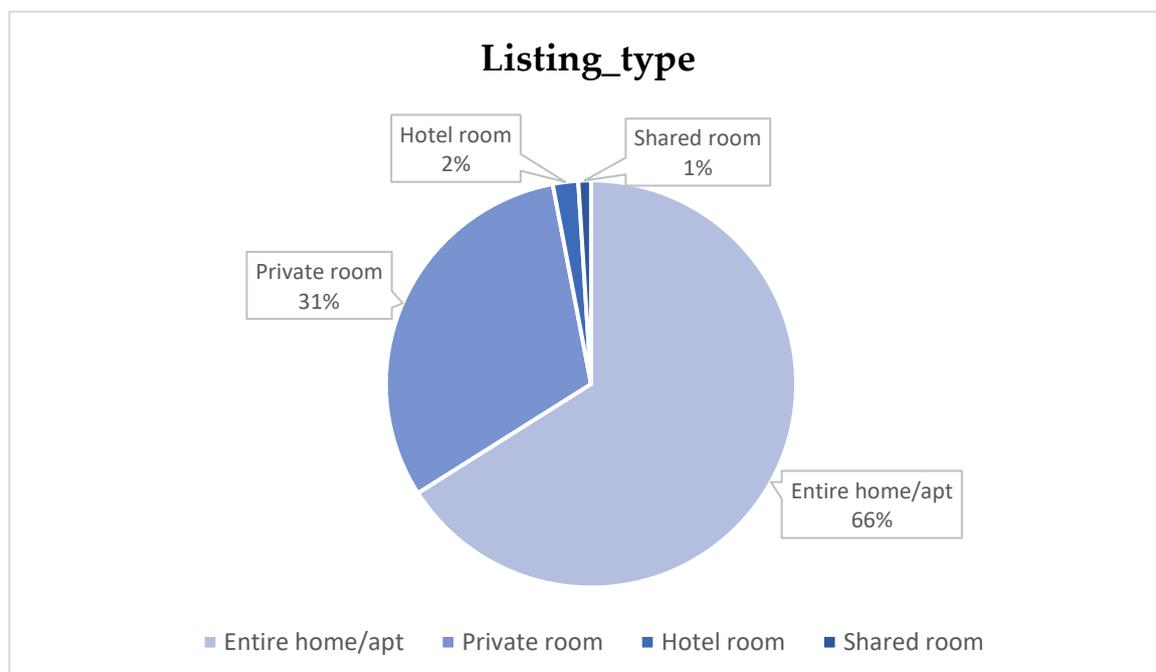


Figura 8.7.3 Suddivisione tipologia listing per host uomini

La tipologia prevalente per i listings di host donne e di host uomini è l'intero appartamento, sicuramente soluzione preferibile per aver maggior privacy. Questa sarà una variabile fondamentale nell'analisi econometrica, perché la propensione a scegliere un host del medesimo sesso potrebbe aumentare nel caso il listing appartenga alla tipologia Private room o Shared room.

9.8 Superhost

La qualifica di superhost, secondo il sito di Airbnb, viene assegnata automaticamente se vengono rispettati i seguenti requisiti:

- aver completato almeno 10 viaggi OPPURE 3 prenotazioni per una durata complessiva di almeno 100 notti;
- aver mantenuto un tasso di risposta pari almeno al 90%;
- aver mantenuto un tasso di cancellazione non superiore all'1%;
- aver mantenuto una valutazione complessiva di un punteggio pari ad almeno il 4,8.

Possedere la qualifica può essere una variabile che viene valutata nel momento della scelta di un determinato listing perché fornisce informazioni sulle prenotazioni passate dell'host e sulla qualità di esse. Potrebbe quindi avere un effetto significativo sul tasso di domanda.

Nel database in analisi, il 39% dei listing viene certificato come superhost. Di questi, il 47,8% sono listing di host donne mentre il 42,2% sono di host uomini (Figura 8.8.1).

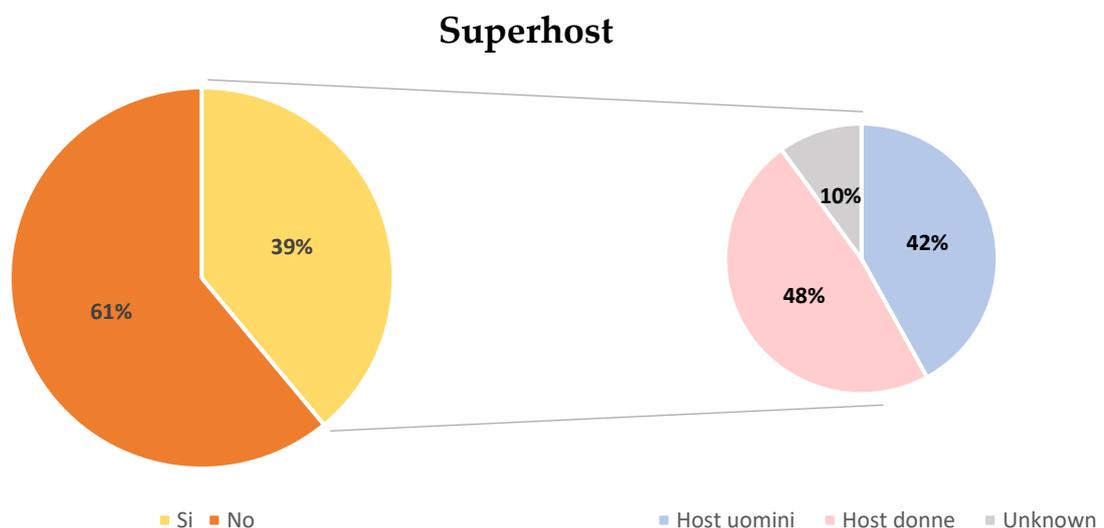


Figura 8.8.1 Percentuale superhost

Osservando invece i dati da un altro punto di vista, sul totale dei listings appartenenti ad host donne, il 42% di questi possiede la qualifica di superhost, mentre considerando tutti i listings di host uomini, solo il 39% è un superhost (Figura 8.8.2)

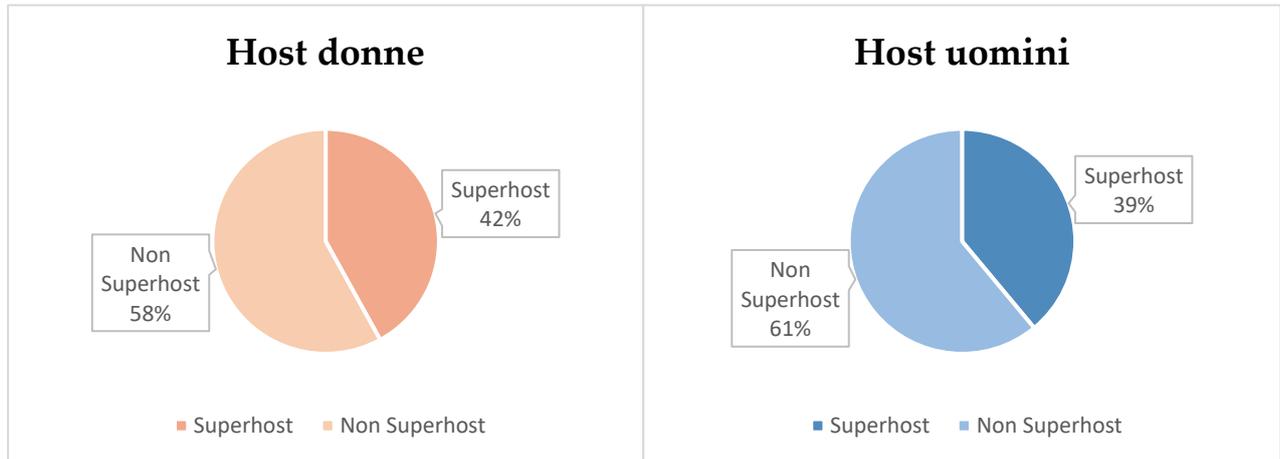


Figura 8.8.2 Percentuale superhost host donne e host uomini

9.9 Numero listings

Molti host posseggono più di un listing registrato sulla piattaforma. Questa variabile è importante perché permette di capire se l'host può essere considerato occasionale oppure un professionista.

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
1,64	2,65	1	62	1

Si vuole verificare se il numero medio di listings posseduti da host uomini e host donne possa ritenersi uguale.

Descrizione variabili:

Numero_Listing: numero di proprietà registrate sulla piattaforma per ogni host;

Sesso: sesso dell'host, F → femmina, M → maschio. Si testa la seguente ipotesi:

$$\left\{ \begin{array}{l} H_0: \mu_F - \mu_M = 0 \\ H_a: \mu_F - \mu_M \neq 0 \end{array} \right.$$

Il comando utilizzato è stato:

ttest numero_listing, by (sesso)

Risultati:

μ Donne	μ Uomini	t differenza	p-value	Intervallo di confidenza (95%)
1,4	1,6	-2,89	0,0039	[-0,3586 ; -0,0687]

L'intervallo di confidenza non include lo zero, la statistica calcolata non ricade nella zona di accettazione, pertanto si rifiuta l'ipotesi nulla per cui la media delle sistemazioni possedute dalle donne sia uguale a quella posseduta dagli uomini. Le donne possiedono meno listings registrati.

9.10 Caratteristiche del listing

Ogni listing è caratterizzato da un determinato numero di bagni, di camere da letto e di letti. Queste caratteristiche sono importanti per comprendere la dimensione del listing e potrebbero avere un effetto significativo sul prezzo del soggiorno.

Per ogni listing, inoltre, è indicato il numero massimo di persone che può ospitare, dato che dovrebbe essere coerente con il numero di letti presenti, e un numero minimo di notti necessarie per soggiornarci.

Numero di bagni

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
1,5	1,3	1	12	1

Per alcuni listings del database originale scaricato da Insider Airbnb veniva indicata anche la tipologia del bagno, se privato o condiviso. È stato deciso di non tenere conto di questa informazione poiché erano molti i listings in cui non era presente nessuna indicazione a riguardo.

Numero di camere da letto

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
1,29	0,71	0	6	1

Numero di letti

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
1,97	1,41	1	11	2

Accomodate

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
3	1,78	1	16	3

Si può osservare che in media il numero massimo di persone che si possono ospitare è superiore rispetto al numero di letti presenti in un listing. Bisognerebbe capire se per numero di letto si intendono tutte le postazioni per dormire come anche divani letti o magari letti a due piazze che possono accogliere due persone o se effettivamente fa riferimenti solo al numero di letti.

Notti minime

Media	Deviazione std	Valore minimo	Valore massimo	Mediana
2,27	4,01	1	90	2

9.11 Distanza

La metodologia utilizzata per calcolare la distanza individua sei classi in cui i listing si distribuiscono, corrispondenti alla distanza dei sei quartieri individuati rispetto ai principali punti turistici di Bologna.

La *Figura 9.11.1* rappresenta la frequenza delle distanze.

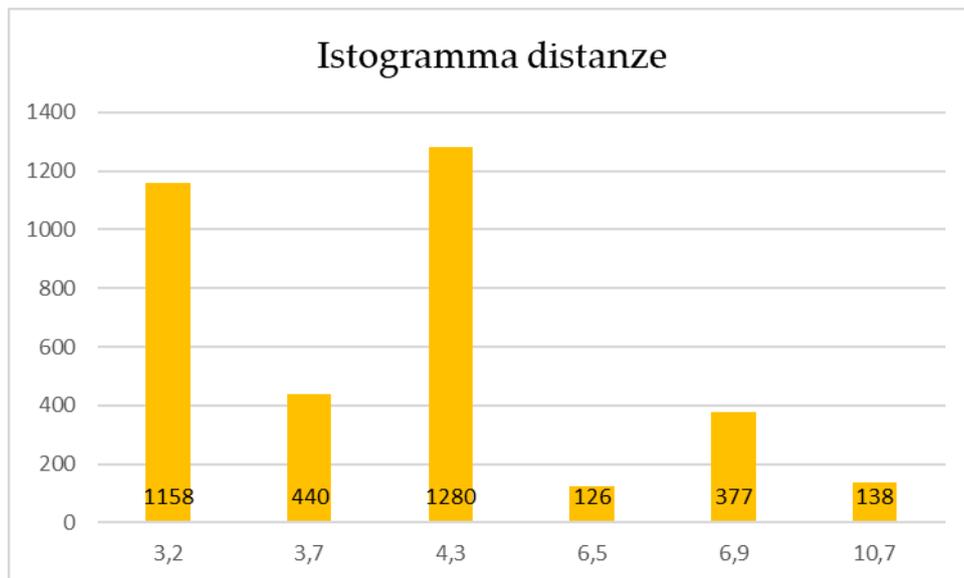


Figura 9.11.1 Istogramma distanze

In base alle frequenze si può affermare che la maggior parte dei listings si trova a poca distanza dalle principali attrazioni. Questa variabile influenzerà il prezzo, perché ovviamente più una sistemazione è vicina al centro e ai principali punti della città, maggiore sarà il prezzo per notte. Gli appartamenti più lontani invece non possono puntare sulla distanza per attrarre turisti ma dovranno valutare altri aspetti per essere competitivi e ospitare le persone, come magari un prezzo più basso.

10. Analisi di regressione

Nel presente capitolo verranno eseguite tre differenti analisi di regressioni e saranno mostrati i risultati ottenuti. Come nei capitoli precedenti si utilizzerà il software STATA per effettuare queste analisi.

Per quanto riguarda la prima retta di regressione si andranno ad indagare cosa influenza la variabile dipendente prezzo, con particolare attenzione alla variabile del sesso dell'host, per capire se quest'ultima ha un'influenza significativa sulla determinazione del prezzo.

Nella seconda analisi di regressione si studierà la domanda come variabile dipendente e si porrà sempre attenzione all'influenza del sesso dell'host.

Infine, la terza ed ultima analisi econometrica sarà eseguita con la funzione probit, con l'obiettivo di verificare se esiste una tendenza dei guest ad andare da host del medesimo sesso.

Nei modelli presentati si terrà conto in un primo momento di variabili più di tipo quantitativo, mentre saranno aggiunte in un secondo momento variabili qualitative, legate agli aspetti principali citati nelle recensioni, descritti nel capitolo 8, paragrafo 8.2.

Le analisi saranno eseguite utilizzando la retta di regressione lineare, con l'obiettivo di stimare l'effetto causale sulla variabile dipendente di una variazione unitaria della variabile indipendente. Il modello utilizzato sarà il seguente:

$$Y = \beta_0 + \beta_1 X_i + \mu_i$$

dove l'errore di regressione è rappresentato da μ_i , costituito da fattori omessi.

I coefficienti β_0 e β_1 verranno calcolati mediante lo stimatore OLS (Ordinary Least Squares), il quale minimizza la somma dei quadrati delle differenze tra i valori reali osservati della variabile dipendente ed i valori predetti in base alla retta di regressione stimata.

Occorre constatare che nella retta di regressione lineare molte variabili sarebbero omesse e questo potrebbe comportare una distorsione dello stimatore OLS. Pertanto, nelle analisi effettuate verrà utilizzato il modello di regressione multipla. Per valutare se i modelli utilizzati si adattino ai dati a disposizione, si valuterà come misura della bontà di adattamento l' R^2 che misura la frazione della varianza in Y spiegata da X . Per la regressione multipla si utilizza l' R^2 corretto, il quale permette di correggere il valore dell' R^2 che aumenta necessariamente quando si aggiunge un nuovo regressore. Inoltre, risulta fondamentale inserire il comando "robust" nella parte finale della formula affinché si eviti che gli errori standard e di conseguenza la statistica-t siano errati. Utilizzando tale comando STATA calcola gli errori standard robusti all'eteroschedasticità, dove con tale termine si intende che la varianza di μ condizionata a X dipende da X .

Negli output mostrati nelle parentesi saranno riportati gli errori standard robusti all'eteroschedasticità e la significatività delle variabili sarà così indicata:

*** $p < 0.1$, ** $p < 0.05$, * $p < 0.1$.

10.1 Analisi di regressione sul prezzo

Si vuole indagare se esiste un'influenza significativa del sesso dell'host sul prezzo del listing. Per far ciò è necessario inserire nel modello ulteriori variabili esplicative, perché vi sono altri aspetti che potrebbero influenzare il prezzo di una determinata sistemazione presente sulla piattaforma di Airbnb. Alcune di queste variabili erano già inserite nel database, altre invece sono state costruite in un secondo momento per essere utilizzate nel modello.

Inizialmente sono state inserite variabili che fornissero qualche indicazione sull'host. Infatti, ci sono alcune caratteristiche dei proprietari dei listings che potrebbero avere un'influenza sul prezzo. Di seguito sono riportate le variabili che si è deciso di inserire nel modello:

- Dumsesso: variabile binaria che assume il valore 1 se l'host è di sesso femminile, 0 altrimenti.

- Superhost: variabile binaria che assume valore 1 se l'host possiede lo status di superhost, 0 altrimenti. Chi ha ottenuto questa qualifica da Airbnb potrebbe applicare un sovrapprezzo per il proprio listing poiché la sua reputazione è maggiore rispetto a quella di altri host senza lo status di superhost.
- Host_verified: variabile binaria che assume valore 1 se l'host ha l'identità verificata dalla piattaforma, 0 altrimenti. Airbnb garantisce in questo modo l'identità dell'host, fornendo maggiore fiducia al guest.
- Host_multiprop: variabile binaria che assume valore 1 se l'host possiede più di un listing registrato su Airbnb, 0 altrimenti. Questa variabile è stata costruita per essere inserita nella retta di regressione perché un host che possiede più sistemazioni dovrebbe essere più esperto e potrebbe quindi fissare un prezzo più alto.
- Anzianità: numero di anni in cui l'host è presente sulla piattaforma. Anche in questo caso la variabile è sinonimo di esperienza: più anni sei sulla piattaforma, più hai avuto modo di ospitare persone e di migliorare. Il calcolo degli anni di anzianità è stato effettuato osservando per ogni listing a che anno risale la prima recensione rilasciata.

Un'altra variabile inserita nel modello è la domanda. Infatti, se la domanda aumenta per la relazione domanda-offerta ci si aspetterebbe una diminuzione del prezzo. Nel database si ha a disposizione sia il numero di reviews per ogni listing, sia il numero di reviews al mese. Entrambe le variabili possono essere utilizzate come proxy della domanda. Per questa analisi si è deciso di inserire la variabile `reviews_al_mese`.

Successivamente si è ragionato su un altro gruppo di variabili che potrebbero influenzare il prezzo, ovvero quelle che descrivono il listing. Nel database a disposizione le variabili in questa categoria erano le seguenti:

- Accomodates: numero massimo di ospiti che si possono accogliere nel listing;
- Bagni: numero di bagni;
- Bed: numero di letti;

- Bedrooms: numero di camere da letto,
- Distanza: distanza media del listing dai principali punti di attrazione di Bologna,
- Listing_Type: variabile binaria che assume valore 1 se l'appartamento appartiene alla categoria Entire home/apartment, 0 altrimenti. Se si affitta un intero appartamento ci si aspetta un soprapprezzo.

Queste variabili (esclusa la distanza) forniscono un'idea sulla grandezza del listing ma sono correlate tra di loro: più ospiti si possono accogliere, più letti ci saranno e conseguentemente più camere da letto. Si è deciso di inserire comunque sia la variabile sul numero di ospiti sia quella sul numero di camere per avere un'idea sulla grandezza del listing, oltre alle variabili Distanza e Listing Type. Nel modello di regressione più variabili si inseriscono, più è probabile che esse siano correlate tra loro. In questo caso, si parla di collinearità imperfetta, ovvero quando due o più regressori sono correlati. La conseguenza è la stima imprecisa di uno o più coefficienti di regressione. Pertanto, è importante riflettere su quali variabili inserire nel modello per non generare grandi errori standard per i coefficienti OLS.

Infine, si voleva inserire un variabile che fornisse qualche indicazione sulla qualità del listing e delle esperienze passate. Nel database erano a disposizione le seguenti variabili:

- review_score_rating;
- review_score_accuracy;
- review_score_cleanliness;
- review_score_checkin;
- review_score_communication;
- review_score_location;
- review_score_value.

Dal momento che il review_score_rating è una media dei punteggi ottenuti nelle singole categorie e risulta quindi correlata alle altre variabili, si è deciso di inserire nel modello solo questa variabile. In *Tabella 10.1.1* sono riportati i tassi di

correlazione della variabile `review_score_rating` con gli altri punteggi delle singole voci e si può osservare come i valori siano molto alti, mai inferiori allo 0,5, escluso il `review_score_location` che riporta un valore di correlazione pari a 0,4 circa.

Tabella 10.1.1 Correlazione variabili con punteggi

	Review score accuracy	Review score check-in	Review score cleanliness	Review score communication	Review score location	Review score value
Review score rating	0,7293	0,5753	0,6960	0,6402	0,4130	0,758

Le variabili descritte sono state inserite nel modello in quattro step successivi per osservare se e come si modificassero i singoli effetti. I risultati ottenuti sono mostrati in Tabella 10.1.2.

Tabella 10.1.2 Analisi regressione con variabile dipendente prezzo

Prezzo	Caratteristiche host (1)	Domanda (2)	Caratteristiche listing (3)	Punteggio (4)
Dumsesso	-2,3950 (2,6044)	-3,5858 (6,6921)	-2,6789 (2,5060)	-2,5132 (2,7727)
Superhost	7.5059*** (2,5368)	6.6921*** (2,6758)	2,2926 (2,6007)	-0,0364 (2,8614)
Host_multiprop	0.5514 (2,5976)	1.0006 (2,6674)	1,3884 (2,5143)	1,6464 (2,7529)
Host_verified	8.0633*** (2,8935)	4.9791*** (3,0367)	-3,6314 (3,2616)	-6,2362 (3,5184)
Anzianità	-2.5839*** (0,7764)	0.6879 (0,8835)	1,6449** (0,8406)	1,3379 (0,9107)
Reviews_al_mese		-6,1567*** (0,6931)	-6,6628*** (0,6823)	-6,9570*** (0,7667)

Accomodate			9,3541*** (2,6998)	8,8515*** (2,7915)
Bagni			27,6852*** (5,9390)	27,8370*** (5,9891)
Bedrooms			13,9674** (5,6076)	13,8419*** (5,7418)
Distanza			-1,9347*** (0,6674)	-2,2400*** (0,7271)
Listing_type			24,7883*** (2,7142)	25,8155*** (2,9864)
Review_score_rating				-0,0657 (0,2295)
Cons	84,4816*** (3,1236)	81,9141*** (3,7913)	16,3283* (9,0721)	29,2549 (22,7421)
R2	0,0097	0,0248	0,1021	0,1306
Numero obs	2942	2471	2471	2471
F-all	4,97***	14,30***	26,80***	22,24***

Nel modello di regressione lineare utilizzato la variabile di interesse dumsesso non è mai significativa, in nessuno dei quattro casi. I guest non sono disposti a pagare di più se l'host è uomo o se è donna, quindi automaticamente chi possiede listing non può applicare un sovrapprezzo.

Spostando l'analisi sulle altre variabili, ci concentriamo prima su quelle legate alle caratteristiche dell'host. La variabile superhost è significativa nei modelli (1) e (2) ed ha un'influenza positiva sul prezzo, come si era ipotizzato: se l'host possiede questa qualifica si registra un aumento di prezzo di 7,5 nel primo modello e 6,7 nel secondo. Stesso discorso vale per la variabile host_verified, significativa nei primi due casi: si osserva un aumento del prezzo di 8,06 e successivamente di 4,98. La

variabile `host_multiprop` non è mai significativa, mentre l'anzianità ha un'influenza significativa nei modelli (1) e (3): mentre nel primo scenario ha segno opposto rispetto a quanto previsto, nel terzo l'effetto è positivo.

Per quanto concerne la domanda, si osserva un'influenza significativa e negativo sul prezzo, come previsto. Un aumento della domanda comporta una riduzione di 6 e 7 sul prezzo.

L'analisi si sposta sulle caratteristiche del listing: tutte le variabili sono significative. Le variabili `Bagni`, `Accomodate` e `Bedrooms` hanno un'influenza positiva molto alta sul prezzo, rispettivamente di 27,9 e 13, la variabile `Distanza` ha un'influenza negativa sul prezzo di 1,9 nel modello (3) e 2,24 nel modello (4) (maggiore è la distanza dai principali punti di attrazione della città, minore sarà il prezzo) mentre la variabile `Listing_type` ha un'influenza positiva di circa 25 nel modello (3) e circa 27 nel modello (4), perché ovviamente un appartamento intero rispetto ad una camera prevede un sopraprezzo.

Infine, la variabile `review_score_rating` non ha un effetto significativo sul prezzo.

Dal momento che la variabile di interesse non risulta mai significativa con la forma lineare, si è provato il modello di regressione utilizzando come variabile dipendente il log del prezzo. Infatti, la relazione tra il prezzo e il sesso dell'host potrebbe essere non lineare. In *Tabella 10.1.3* sono riportati i risultati.

Tabella 10.1.3 Analisi regressione con variabile dipendente logprezzo

Logprezzo	
Dumsesso	0,0164*** (0,0057)
Superhost	-0,0130 (0,0078)
Host_multiprop	0,0009 (0,0056)
Host_verified	0,0038

	(0,0057)
Anzianità	0,0121***
	(0,0023)
Log review_al_mese	-0,0070***
	(0,0020)
Accomodate	0,0012
	(0,0027)
Bagni	0,0062
	(0,0060)
Bedrooms	0,1567
	(0,0126)
Distanza	-0,0023
	(0,0021)
Listing_type	0,0052
	(0,0083)
Review_score_rating	-0,0003
	(0,0004)
Cons	0,9713***
	(0,0156)
R2	0,1553
Numero obs	2471
F-all	5,35***

In questo caso la variabile dumsesso risulta significativa: se l'host è di sesso femminile, si osserva un aumento del prezzo del listing di circa il 2%. Questo potrebbe accadere perché le host donne sono più attente e più disponibili, quindi si possono permettere di fissare un prezzo leggermente più alto rispetto ai propri colleghi uomini. Questo risultato è interessante e forse è diverso rispetto a quanto

ci si potesse aspettare, in un periodo in cui tendenzialmente sono gli uomini a guadagnare di più. Le altre variabili che risultano significative in tale modello sono l'anzianità con un'influenza positiva sul prezzo dell'1% e il log delle reviews al mese con un effetto negativo. Pertanto, un host presente sul sito da più tempo e quindi con più esperienza può permettersi di applicare un prezzo più alto, mentre se la domanda aumenta il prezzo diminuisce.

10.2 Entire home vs Private room

L'analisi effettuata si basa sul dataset contenente tutti i listing, i quali risultano però molto diversi tra loro. La variabile Listing_type tiene conto delle differenze tra le varie sistemazioni, tuttavia si può ipotizzare che i risultati in realtà possano cambiare se il listing è di tipo Entire home/apartment o un'altra tipologia come Private Room o Shared room. In merito a tale considerazione si è deciso di dividere il database in due, separando i listings affittati interamente e quelli in cui si prenota una camera in un appartamento condiviso. L'analisi è stata eseguita su entrambi per capire se le determinanti del prezzo variano ma soprattutto per verificare se il sesso dell'host influenza il prezzo in un caso o nell'altro, in entrambi o in nessuno dei due database. La variabile dipendente è il prezzo e il modello è lineare. La *Tabella 10.2.1* mostra i risultati ottenuti.

Tabella 10.2.1 Analisi regressione con variabile dipendente prezzo con database separati

ENTIRE HOME		PRIVATE/SHARED ROOM	
Prezzo		Prezzo	
Dumsesso	-0,1926 (3,6238)	Dumsesso	-5,1290 (4,7939)
Superhost	-4,1032 (3,5948)	Superhost	8,6613* (5,1530)

Host_multiprop	2,5487 (3,5495)	Host_multiprop	-0,3004 (4,6029)
Host_verified	-2,7359 (4,5663)	Host_verified	-11,9782** (5,4578)
Anzianità	1,0570 (1,2424)	Anzianità	-0,4296 (1,2407)
Reviews_al_mese	-7,9739*** (1,0363)	Reviews_al_mese	-6,5531*** (1,1268)
Accomodate	5,8102** (2,2210)	Accomodate	19,8723*** (5,3310)
Bagni	29,2763*** (4,5020)	Bagni	15,9563 (10,8197)
Bedrooms	14,7616** (6,2054)	Bedrooms	23,1313* (12,6351)
Distanza	-2,0690** (0,9986)	Distanza	-3,3991*** (-1,0137)
Review_score_rating	-0,1477 (0,3352)	Review_score_rating	0,3885** (0,2055)
Cons	65,6932** (34,3548)	Cons	-10,2186 (19,6766)
R2	0,0940	R2	0,2018
Numero obs	1525	Numero obs	726
F-all	11,87***	F-all	10,51***

Da una prima analisi dei dati nei due output si osserva che la variabile di interesse dumsesso non è significativa. Il dataset contenente i listing del tipo Entire home/apartment presenta gli stessi risultati ottenuti nel modello generale, cioè sono significative le variabili Reviews_al_mese, Accomodate, Bedrooms, Bagni e

Distanza. La situazione invece cambia considerando le altre tipologie di listing, quindi Private room o Shared room. Le Review_al_mese, la distanza e il numero di ospiti espressi dalla variabile Accomodates continuano ad essere significative, mentre invece la variabile Bagni perde di significatività. Infatti, quando si prenota una camera in un appartamento condiviso lo si fa per vivere un'esperienza diversa o anche per risparmiare, pertanto le persone potrebbero non essere disposte a pagare per avere dei bagni in più. Contrariamente al caso generale e all'Entire home, se si affitta una singola camera il prezzo di questa aumenta all'aumentare del punteggio medio delle review: un guest che deve condividere un appartamento probabilmente tiene più conto delle recensioni del listing ed è disposto a pagare un sovrapprezzo quando il punteggio aumenta. Tuttavia, il prezzo diminuisce significativamente se l'host ha l'identità verificata, effetto contrario a quanto ci si aspettava.

La medesima analisi si è eseguita utilizzando come variabile dipendente il log del prezzo. [Tabella 10.2.2]

Tabella 10.2.2 Analisi regressione con variabile dipendente prezzo con database separati

ENTIRE HOME		PRIVATE/SHARED ROOM	
Log prezzo		Log prezzo	
Dumsesso	0,0201** (0,0084)	Dumsesso	0,0222 (0,0143)
Superhost	0,0171** (0,0087)	Superhost	-0,0045 (0,0144)
Host_multiprop	0,0028 (0,0086)	Host_multiprop	-0,0156 (0,0144)
Host_verified	0,011 (0,0084)	Host_verified	-0,0126 (0,0160)
Anzianità	0,0248***	Anzianità	0,0085**

	(0,0039)		(0,0040)
Log review_al_mese	-0,0146**	Log review_al_mese	-0,0210**
	(0,0070)		(0,0091)
Accomodate	0,0020	Accomodate	-0,0088
	(0,0033)		(0,0085)
Bagni	0,0006	Bagni	0,0153
	(0,0064)		(0,0149)
Bedrooms	0,0077	Bedrooms	0,0372
	(0,0157)		(0,0387)
Distanza	-0,0022	Distanza	-0,0026
	(0,0023)		(0,0041)
Review_score_rating	0,0004	Review_score_rating	0,0005
	(0,0006)		(0,0005)
Cons	0,9800***	Cons	0,9148***
	(0,0632)		(0,0548)
R2	0,1697	R2	0,1847
Numero obs	1551	Numero obs	726
F-all	4,51***	F-all	2,19**

La variabile di interesse si comporta in modo differente nei due database: per gli appartamenti interi è significativa, per le camere in case condivise non lo è. Una possibile spiegazione dei risultati ottenuti potrebbe essere dovuta alle caratteristiche dei guest che decidono di condividere la camera, tendenzialmente giovani che probabilmente non sono disposti a pagare di più in base al sesso dell'host, mentre chi decide di pernottare in un appartamento intero potrebbe pagare di più pur di avere una garanzia di un buon soggiorno e le host donne potrebbero essere più rassicuranti e attente su questo aspetto rispetto agli host uomini. Al contrario del caso generale con il log prezzo, oltre alle variabili anzianità

e `Log reviews_al_mese`, anche la variabile `superhost` è significativa e prevede un aumento di prezzo di circa il 2% nel dataset `Entire home`, mentre non è significativa se si sceglie una camera in un appartamento condiviso, sempre perché probabilmente chi alloggia in questa tipologia di sistemazione vuole risparmiare e non è disposto a pagare di più per la qualifica di `superhost`.

10.3 Relazione tra parole nelle recensioni e prezzo

Una delle analisi principali svolte in questo lavoro di tesi è stata quella di individuare quali fossero le parole più utilizzate nelle recensioni rilasciate dai guest uomini e dai guest donne, con l'obiettivo di capire se gli aspetti valutati dai due sessi fossero i medesimi, o se questi invece differissero tra loro. Nel complesso non sono risultati grossi scostamenti tra le parole più utilizzate dagli uomini e dalle donne. In generale, tutti i guest descrivono nelle loro review la posizione del listing, la pulizia, l'host e l'esperienza vissuta. Si è deciso allora di verificare se questi temi possano influenzare la determinazione del prezzo. Per ogni listing è stato calcolato il numero di review in cui compaiono le parole scelte, rapportandolo alle recensioni totali registrate dal listing per ottenere la percentuale di reviews per ogni listing in cui compaiono le parole `clean`, `helpful`, `location` e `great`. Si sono costruite in tal modo quattro variabili:

- `Clean`
- `Location`
- `Helpful`
- `Great`

Queste parole compaiono tra quelle più utilizzate e si collegano ai punti principali che tendenzialmente vengono analizzati e descritti nelle recensioni. Dal momento che nelle analisi precedenti è emerso che la variabile di interesse `dum sesso` è significativa con la variabile dipendente `log prezzo`, si è deciso di valutare l'effetto delle variabili nel modello `log-lineare` e non in quello `lineare`. L'analisi è stata svolta

sul database contenente tutti i litings e sui due database separati per tipologia, ovvero Entire home e Private/Shared room.

In *Tabella 10.3.1* sono riportati i risultati per le tre casistiche.

Tabella 10.3.1 Analisi regressione con variabile dipendente prezzo con parole più utilizzate

	DATABASE COMPLETO	ENTIRE HOME	PRIVATE/SHARED ROOM
Log prezzo			
Dumsesso	0,0242** (0,0086)	0,0192** (0,0083)	0,0230 (0,0143)
Superhost	-0,0151 (0,0094)	-0,0177 (0,0087)	-0,0059 (0,0140)
Host_multiprop	0,0035 (0,0084)	-0,0033 (0,0086)	-0,0174 (0,0141)
Host_verified	0,0027 (0,0073)	0,0127 (0,0083)	-0,0176 (0,0156)
Anzianità	0,0212*** (0,0030)	0,0228*** (0,0038)	0,0075** (0,0041)
Log review_al_mese	-0,0153** (0,0056)	-0,0149** (0,0071)	-0,0205*** (0,0084)
Accomodate	-0,0007 (0,0025)	-0,0015 (0,0032)	-0,0075 (0,0087)
Bagni	0,0061 (0,0059)	0,0000 (0,0064)	0,0154 (0,0154)
Bedrooms	0,0151 (0,0146)	0,0059 (0,0151)	0,0367 (0,0390)
Distanza	-0,0022 (0,0022)	-0,0019 (0,0023)	-0,0020 (0,0041)
Listing type	0,0010 (0,0082)	/	/
Review_score_rating	-0,0003 (0,0004)	-0,006 (0,0005)	0,0002 (0,0005)
Clean	-0,0344 (0,0247)	-0,0339 (0,0294)	-0,0367 (0,0462)
Location	-0,0423	-0,0504	-0,0149

	(0,0238)	(0,0281)	(0,0462)
Helpful	0,07727	0,1073**	-0,0314
	(0,0472)	(0,0544)	(0,0964)
Great	0,0955**	0,0843**	0,1395
	(0,0360)	(0,0351)	(0,0881)
Cons	0,9715***	0,9935***	0,9348***
	(0,0391)	(0,0625)	(0,0579)
R2	0,17352	0,1692	0,1843
Numero obs	2471	1551	726
F-all	5,00***	3,41***	1,69**

La variabile *dum sesso* rimane significativa per il database completo e per i listings di tipologia *Entire home*, mentre continua a non essere significativa per le camere in appartamenti condivisi. Anche le variabili *Anzianità* e *Log review_al_mese* continuano ad essere significative in tutti e tre gli scenari.

Invece, spostando l'analisi sulle nuove variabili inserite, si può osservare che *Location* ha un livello di significatività al 10% per il database completo e per gli *Entire home*, mentre *Clean* non risulta mai significativa, quindi le persone probabilmente non sono disposte a pagare un prezzo più alto se sono tante le recensioni contenenti questa parola. Al contrario, nel caso del database completo e per le *Entire home* risulta significativa la variabile *great*: se ci sono molte review che descrivono le esperienze vissute con questo aggettivo positivo, i guest accettano un sovrapprezzo del 9% nel database completo e del 10% se optano per interi alloggi. Inoltre, per gli *Entire home* anche *Helpful* è significativa, quindi chi sceglie questa tipologia desidera che l'host sia disponibile durante l'esperienza nel listing e pagherebbe il 10% in più.

Nel database con le camere in appartamenti condivisi nessuna delle variabili inserite risulta significativa e la spiegazione potrebbe essere la medesima delle analisi precedenti: che sceglie questa esperienza probabilmente possiede una *willingness to pay* più bassa e non è disposto a pagare di più per questi aspetti citati nelle recensioni.

10.4 Analisi di regressione sulla domanda

L'analisi si sposta sullo studio dei fattori che influenzano un'altra variabile dipendente importante, ovvero la domanda. Come più volte specificato, si utilizzano le recensioni come proxy della domanda.

Nel database a disposizione per ogni listing si ha informazione sia del numero di reviews al mese, sia del numero di reviews totali raccolte per la sistemazione ed entrambe possono essere considerate come tasso di domanda. La variabile di interesse risulta essere sempre il sesso dell'host e si è inizialmente valutata l'influenza di quest'ultima sul numero di reviews utilizzando una retta di regressione lineare semplice. La variabile dumsesso risulta significativa e si è deciso allora di effettuare l'analisi su questa variabile [Tabella 10.4.1].

Tabella 10.4.1 Analisi regressione semplice con variabile dipendente numero reviews

Numero reviews	
Dumsesso	5,6978** (2,5916)
Cons	39,6694 (1,8088)
R2	0,0016

La variabile dumsesso risulta significativa e ha un'influenza positiva: se l'host è donna, la domanda è più alta di 5,7 punti circa. Questo risultato è interessante perché significherebbe che i guest non solo valutano il sesso dell'host nel momento in cui decidono di soggiornare presso un listings su Airbnb, ma prediligono anche rivolgersi ad host donne. Tuttavia, per comprendere la vera influenza del sesso dell'host sulla domanda, è necessario inserire ulteriori variabili di controllo.

Come per lo studio sulla variabile dipendente prezzo, anche in questa analisi sono state inserite inizialmente variabili indicanti caratteristiche dell'host, ovvero il sesso, se possiede più proprietà, se ha l'identità verificata e se è superhost. Inoltre, si è controllato che sul sito di Airbnb, quando si visualizza una sistemazione, compare sia la descrizione dell'host sia l'anno in cui si è registrato come tale sulla piattaforma. Si inserisce quindi come ulteriore caratteristica dell'host la sua anzianità sulla piattaforma. In un secondo momento si valuta l'influenza del prezzo sulla domanda e per la relazione domanda-offerta ci si aspetta che un prezzo maggiore comporti una riduzione della domanda. Contrariamente all'analisi sul prezzo, l'unica variabile inserita sulle caratteristiche del listings è la distanza.

Infine, si valuta l'influenza del review score rating. I risultati sono riportati in *Tabella 10.4.2*

Tabella 10.4.2 Analisi regressione con variabile dipendente numero reviews

Numero_reviews	Caratteristiche	Prezzo (2)	Caratteristiche	Punteggio
	host (1)		listing (3)	(4)
Dumsesso	2,8869 (3,0922)	4,5594* (2,3614)	4,6096** (2,3632)	0,6439 (2,9691)
Superhost	21,041*** (3,1686)	15,4673*** (2,5002)	13,8257*** (2,5170)	15,7796*** (3,0563)
Host_multiprop	-0,7962 (3,0453)	-0,0659 (2,3407)	-0,0884 (2,3242)	-1,9997 (2,8829)
Host_verified	22,0245*** (3,1157)	21,3488*** (2,2324)	21,0921*** (2,2198)	23,6950*** (2,7050)
Anzianità	13,8382*** (0,9230)	12,5378*** (0,7355)	12,4178*** (0,7336)	14,6052*** (1,0844)
Prezzo		-0,0803*** (0,0165)	-0,0924*** (0,0187)	-0,0861*** (0,0187)
Distanza			-2,3548 *** (0,5906)	-2,6974*** (0,7559)
Listing_type			9,5408*** (3,1678)	12,5654*** (3,1678)
Review_score_rating				0,1563 (0,1383)
Cons	-20,4160*** (3,8966)	-7,3646*** (2,5381)	-1,1888 (3,9118)	-26,4862* (13,8838)
R2	0,1565	0,1892	0,1930	0,1720
Numero obs	2942	2942	2277	2942
F-all	72,07***	92,04***	46,31***	73,71***

La variabile di interesse *dum sesso* risulta significativa solo nel secondo e terzo scenario, in cui sono state considerate le caratteristiche dell'host, il prezzo e le caratteristiche del listing. L'influenza di questa variabile è positiva: se l'host è donna, la domanda è più alta di quasi 5 punti. I guest in generale sembrano prediligere host di sesso femminile. Analizzando le altre variabili, si osserva che le caratteristiche dell'host, eccetto la variabile *multi prop*, sono tutte significative e impattano positivamente sulla domanda: se l'host è superhost si osserva un aumento di domanda di 21 punti iniziali che poi scendono a circa 16 nel momento in cui si inseriscono altre variabili. Anche l'identità verificata e l'anzianità dell'host generano una domanda più alta rispettivamente di 22 e 14 punti circa nei quattro scenari. I guest, quindi, prestano attenzione a questi aspetti e probabilmente preferiscono rivolgersi ad host che hanno queste caratteristiche, perché sono sinonimo di affidabilità, esperienza e rappresentano una garanzia sulla qualità dell'esperienza. Proseguendo con l'analisi, anche la variabile prezzo è significativa ma un aumento di questa ha un'influenza negativa sulla domanda, come previsto. Per quanto riguarda le caratteristiche del listing, entrambe le variabili considerate sono significative. Se la distanza aumenta rispetto i punti principali della città di Bologna si osserva una domanda più bassa di circa 2 punti, mentre se il listing è un appartamento intero il numero di review è più alto di 9 punti e poi di 12 quando si inserisce la variabile sui punteggi. L'introduzione di quest'ultima variabile non risulta essere significativa sulla domanda, pertanto si è deciso di effettuare le analisi successive omettendola. Questa decisione è stata presa anche osservando l' R^2 che diminuisce nel modello (4).

10.5 Tipologia di listing

È interessante capire se il sesso dell'host influenza la domanda nel caso in cui la tipologia di listing sia differente. Si esegue allora l'analisi separando i due database in base al listing: l'intero appartamento o la camera in un listing condiviso. I risultati sono riportati in *Tabella 10.5.1*.

Tabella 10.5.1 Analisi regressione con variabile dipendente numero reviews con database separa

ENTIRE HOME		PRIVATE/SHARED ROOM	
Numero_reviews		Numero_reviews	
Dumsesso	2,8536 (2,3632)	Dumsesso	6,7956** (0,0143)
Superhost	11,1633*** (2,5170)	Superhost	16,7527*** (0,0140)
Host_multiprop	-2,9291 (2,3242)	Host_multiprop	6,7420* (0,0150)
Host_verified	19,0710*** (2,2198)	Host_verified	18,1309*** (0,0158)
Anzianità	17,0194*** (0,7336)	Anzianità	10,3755*** (0,0052)
Prezzo	-0,0985*** (0,0187)	Prezzo	-0,0709*** (0,0029)
Distanza	-2,9682 *** (0,5906)	Distanza	-0,5175 (0,0040)
Cons	2,0840 (4,3955)	Cons	-9,2222 (0,0416)
R2	0,2812	R2	0,1750
Numero obs	1943	Numero obs	999
F-all	67,04***	F-all	19,44***

Un risultato che sorprende è la variabile *dumpresso*: per il primo scenario non risulta significativa, mentre si rifiuta l'ipotesi nulla che il coefficiente sia pari a zero con un livello di significatività pari al 10% per il secondo scenario.

Per quanto concerne le altre variabili invece continuano ad essere significative *superhost*, l'identità verificata dell'host, l'anzianità e il prezzo. La domanda per la tipologia *Entire home/apartment* è più bassa di circa 3 punti se aumenta la distanza, mentre lo stesso effetto non risulta per le camere in listing condivisi, in cui contano solo le caratteristiche dell'host e il prezzo.

In generale, si può concludere che le variabili che più influiscono sul tasso di domanda sono sempre quelle legate all'host e alla sua immagine e questo effetto lo si osserva in tutti gli scenari analizzati.

10.6 Relazione tra parole nelle recensioni e domanda

I guest, quando ricercano una sistemazione per i loro viaggi e spostamenti, dopo aver inserito i vari filtri per trovare i listings che più si adattano alle loro necessità, si concentrano sull'host, come emerso nell'analisi di regressione precedente, per raccogliere più informazioni possibili su di lui e tendenzialmente, in un secondo momento, si dedicano alla lettura delle recensioni. Infatti, Airbnb aumenta la sua reputazione e quella degli host che si registrano sulla piattaforma proprio grazie alle recensioni dei diversi utenti che hanno soggiornato presso una struttura presente sulla piattaforma. I guest, leggendo le recensioni di altri viaggiatori, si potrebbero lasciare influenzare sulla scelta di un listing. Si è scoperto nei paragrafi precedenti quali sono le parole più utilizzate e quindi gli aspetti più importanti per chi soggiorna nei listings. A priori, prima di prenotare, i guest ricercano probabilmente queste tematiche nelle recensioni rilasciate. Si vuole allora indagare se le parole più utilizzate, ovvero *clean*, *location*, *great* e *helpful* influenzano le recensioni rilasciate per un listings, ovvero la domanda. Bisogna considerare il fatto che una persona che vuole pernottare presso una determinata sistemazione legge un numero di recensioni limitato e non tutte quelle rilasciate all'host. Supponendo

però che sulle prime 20 recensioni, in 10 di queste ci sia la parola helpful riguardante l'host, è una rassicurazione per il guest sull'affidabilità e sulla disponibilità della persona che lo ospiterà. Nella *Tabella 10.6.1* si osserva se effettivamente queste parole hanno un'influenza sulla domanda. L'analisi è stata svolta sia sul database completo, sia nel caso di Entire home e sia per la tipologia Private/Shared room.

Tabella 10.6.1 Analisi regressione con variabile dipendente numero reviews con parole delle recensioni

	DATABASE COMPLETO	ENTIRE HOME	PRIVATE/SHARED ROOM
Numero reviews			
Dumsesso	4,7083*** (2,3618)	3,0518 (2,8393)	6,6212* (3,7891)
Superhost	12,8890*** (2,5320)	10,7222*** (2,9635)	16,6757*** (4,3813)
Host_multiprop	-0,2092 (2,3239)	-3,2035 (2,7895)	6,5821* (3,7673)
Host_verified	20,5107*** (2,1972)	18,9224*** (2,6556)	17,9593*** (3,3294)
Anzianità	11,6401*** (0,8237)	16,6463*** (1,0582)	10,2603*** (1,1383)
Prezzo	-0,0919*** (0,0171)	-0,0983*** (0,0208)	-0,0723*** (0,0260)
Distanza	-1,9858*** (0,5803)	-2,7283** (0,7114)	-0,3750 (1,0216)
Listing type	8,1511*** (2,4726)	/	/
Clean	-2,9334 (6,1366)	-8,3558 (7,7489)	-2,6325 (8,2654)

Location	7,2408 (6,8491)	-3,9965 (7,4897)	22,1390* (12,8510)
Helpful	-1,9639 (10,1199)	-8,7050 (10,7626)	-18,5814 (19,8735)
Great	29,3358*** (8,9364)	29,3115*** (10,0522)	-0,56066 (11,3525)
Cons	-2,8938 (3,8407)	0,9029 (4,4307)	-9,8768 (6,4960)
R2	0,2007	0,2714	0,1770
Numero obs	2942	1943	999
F-all	32,27***	44,70***	12,79***

Nei tre scenari la variabile dumsesso risulta significativa solo nel database completo, dove sono considerati tutti i listing indipendentemente dalla tipologia, esattamente come nelle analisi precedentemente effettuate.

Le variabili sulle parole più utilizzate nelle recensioni non risultano mai significative in nessuno scenario, eccetto per la variabile location per le camere in appartamenti condivisi e per la variabile great. Se si considerano tutti i listing e solo gli Entire home, il numero di reviews è più alto di circa 29 se aumenta la percentuale di reviews contenente questa parola. I guest probabilmente non si fanno influenzare dalle altre parole nella scelta di un listing, ma sono più concentrati sull'esperienza in generale: se le recensioni sono globalmente positive e descrivono il soggiorno come grandioso, questo è sufficiente per spingere a pernottare in quella sistemazione. Tuttavia, per le stanze in appartamenti condivisi neanche la presenza della parola great nelle recensioni influenza il numero di review. Chi sceglie questa tipologia di listing non tiene conto degli aspetti più citati nelle recensioni. Rimangono significative le variabili sulle caratteristiche dell'host (anzianità, superhost e identità verificata), il prezzo e la distanza (quest'ultima non per la tipologia private room).

Si può concludere che i listing preferiti dai guest, i quali registrano un numero più alto di recensioni, sono quelli in cui l'host detiene un certo grado di affidabilità, dimostrabile attraverso le informazioni che Airbnb mette a disposizione su di lui, probabilmente consapevole che questo aspetto è importante per chi decide di rivolgersi alla piattaforma. Il prezzo anche è importante per la scelta di un listing insieme alla distanza: sia chi viaggia per lavoro, sia chi viaggia per piacere vorrà essere vicino ai punti centrali della città di Bologna. Infine, le review dei guest influiscono sul tasso di domanda solo nel caso in cui l'esperienza sia descritta come "great", quindi quando la review si può considerare positiva. È importante ricordare che questa analisi si fonda sull'ipotesi che il numero di review possa essere approssimato al tasso di domanda, anche se molti guest decidono di non rilasciare una recensione dopo i loro soggiorni e questi dati sono emessi nelle analisi effettuate in questo lavoro di tesi.

10.7 I guest donne si rivolgono ad host donne?

In ultima istanza viene effettuata un'analisi per valutare se i guest donna preferiscano o meno affittare listings appartenenti ad host del medesimo sesso. Per raggiungere tale scopo non si utilizza la classica regressione, bensì la funzione probit sul software STATA.

La funzione probit esprime una probabilità e per stimare i coefficienti della funzione si utilizza il metodo della stima della massima verosimiglianza: i coefficienti trovati non risolvono problemi di minimo tra la retta di regressione e l'errore (non utilizza infatti gli OLS), bensì massimizzano la funzione di verosimiglianza mediante un procedimento iterativo che converge al risultato migliore.

Nelle analisi di regressione effettuate nei paragrafi precedenti la variabile dipendente è sempre stata continua. Nelle successive invece verrà invece utilizzata una variabile dipendente dummy, ovvero sesso_guest che assume valore pari ad 1 nel caso in cui la viaggiatrice sia donna e 0 in caso contrario.

Per indagare qual è la probabilità delle donne guest ad affittare listing da donne host si potrebbe utilizzare un modello di regressione lineare, in cui il valore predetto della variabile `sexo_guest` è interpretato come la probabilità predetta che tale variabile assuma valore 1. In tal caso β_1 è la variazione di tale probabilità predetta generata da una variazione della variabile indipendente, ovvero `sexo_host` che assume valore 1 se è donna, 0 altrimenti.

Questa soluzione comporterebbe due differenti problematiche:

- la variazione della probabilità prevista per un determinato cambiamento in X è la stessa per tutti i valori di X , ma questo potrebbe non avere senso;
- le probabilità previste potrebbero essere minori di zero o maggiori di 1.

Per tali motivazioni, l'analisi sfrutterà la regressione probit, la quale modella la probabilità che la variabile dipendente sia uguale ad uno utilizzando la funzione di ripartizione normale $\Phi(z)$, valutata in $z = \beta_0 + \beta_1 X$.

Nello specifico, il modello di regressione probit è il seguente:

$$\Pr(Y = 1 | X) = \Phi(\beta_0 + \beta_1 X)$$

La funzione di ripartizione normale soddisfa le seguenti condizioni:

- $\Pr(Y = 1 | X)$ crescente in X per $\beta_1 > 0$;
- $0 \leq \Pr(Y = 1 | X) \leq 1$ per tutte le X

In merito agli output presentati di seguito si precisa che Wald chi2 rappresenta il test eseguito con la "Chi quadro" il quale testa congiuntamente che tutti i coefficienti siano pari a zero e lo pseudo R2 invece ha il medesimo significato del classico R2.

È stata eseguita inizialmente un'analisi probit considerando come unica variabile indipendente il sesso dell'host. I risultati sono mostrati in *Tabella 10.7.1*.

Tabella 10.7.1 Analisi probit con variabile dipendente sesso del guest recensioni

Sesso_guest	
Sesso_host	2,0597*** (0,0062)
Cons	-1,9307 (0,0026)
Pseudo R2	0,2684

Si può osservare che la variabile Sesso_host è significativa. Per capire qual è la probabilità che il guest sia donna, ovvero che la variabile Sesso_guest assuma valore pari a uno, si valuta il valore espresso dalla funzione $\Phi(\beta_0 + \beta_1 X)$. In questo caso, se l'host è di sesso femminile, si avrà $\Phi(-1,9307 + 1 * 2,0597)$, pari a $\Phi(0,129)$. L'area sotto la distribuzione normale standard a sinistra di $z = 0,129$ è 55,17%. Quest'ultima rappresenta la probabilità che il guest sia donna dato che l'host sia del medesimo sesso. Per comprendere meglio questo fenomeno sono state inserite nel modello ulteriori variabili. In particolare, si sono valutate in un primo momento le caratteristiche dell'host, ovvero se è un superhost, se possiede più proprietà, se ha l'identità verificata e la sua anzianità. Nel modello è inserito anche il prezzo e successivamente le caratteristiche del listing come il numero di bagni, il numero di camere da letto e la tipologia di listing. La Tabella 10.7.2 mostra gli output ottenuti.

Tabella 10.7.2 Analisi probit con variabile dipendente sesso del guest recensioni e più scenari

Sesso_guest	Caratteristiche host (1)	Prezzo (2)	Caratteristiche listing (3)
Sesso host	0,1900*** (0,0078)	0,1896*** (0,0078)	0,1869*** (0,0079)
Superhost	0,0037 (0,0081)	0,0035 (0,0080)	-0,0079 (0,0083)
Host_multiprop	0,0028 (0,0078)	0,0024 (0,0078)	0,0060 (0,0080)

Host_verified	0,0754*** (0,0095)	0,0751*** (0,0095)	0,0889*** (0,0096)
Anzianità	0,1862*** (0,0029)	0,1862*** (0,0029)	0,1852*** (0,0028)
Prezzo		0,0001 (0,0000)	0,0001 (0,0000)
Listing_type			-0,0076 (0,0094)
Bagni			-0,0402*** (0,0053)
Bedrooms			-0,0036 (0,0054)
Cons	-1,0165*** (0,0150)	-1,0212*** (0,0156)	-0,9709*** (0,0168)
Pseudo R2	0,0389	0,0389	0,0393
Numero obs	110898	110898	110895
Chi2	5977,67***	5978,98***	6650,04***

La variabile Sesso host rimane statisticamente significativa in tutti e tre i modelli. Le altre caratteristiche del guest che risultano essere significative sono l'anzianità e l'identità verificata. Contrariamente alle analisi svolte nei paragrafi precedenti, le variabili prezzo e listing_type non risultano significative in nessuno dei tre scenari. Stesso discorso vale per il numero di camere da letto, mentre il numero di bagni è una variabile significativa. Nel primo scenario in cui si considera l'influenza delle caratteristiche dell'host, la probabilità che il guest sia donna dato l'host donna è pari a $\Phi(-1,0165+1*0,1900+0,0754+0,1862*4,4)$, cioè $\Phi(0,06) = 52,39\%$. La probabilità, quindi, diminuisce se si considera oltre al sesso dell'host anche l'anzianità e se l'host possiede l'identità verificata. Anche la variabile che esprime il numero di bagni è significativa. La probabilità che sesso_guest sia uguale a uno quando sesso_host è

uguale a uno è pari a $\Phi(-0,9709+1*0,1869+0,0889+0,1852*4,4-0,0076*1,2)$, cioè $\Phi(0,11)$ = 54,38%.

Al fine di fornire una chiara interpretazione dei risultati, si è deciso di eseguire l'analisi probit separando in due database differenti i listing affittabili interamente e quelli in cui si affitta una sola camera (Tabella 10.7.3).

Tabella 10.7.3 Analisi probit con variabile dipendente sesso del guest recensioni con database separati

ENTIRE HOME		PRIVATE/SHARED ROOM	
Sesso_guest		Sesso_guest	
Sesso_host	0,0197*** (0,0092)	Sesso_host	0,1569*** (0,0160)
Superhost	-0,0088 (0,0095)	Superhost	-0,0053 (0,0165)
Host_multiprop	0,0028 (0,0092)	Host_multiprop	0,0235 (0,0163)
Host_verified	0,1013*** (0,0112)	Host_verified	0,0471** (0,0194)
Anzianità	0,1870*** (0,0031)	Anzianità	0,1767*** (0,0058)
Prezzo	0,0001*** (0,0000)	Prezzo	-0,0000 (0,0002)
Bagni	-0,0389*** (0,0054)	Bagni	-0,0837** (0,0314)
Cons	-1,0016*** (0,0178)	Cons	-0,8484 (0,0463)
Pseudo R2	0,0418	Pseudo R2	0,0316
Numero obs	83760	Numero obs	27138
Chi2	5337,64***	Chi2	1305,34***

La variabile sesso_host è significativa in entrambi i database, quindi il sesso dell'host influenza la probabilità che il guest sia donna per qualunque tipologia di listing. Come nell'analisi dei listing generale, anche separando il database si osserva che le variabili statisticamente significative sono le medesime: l'anzianità dell'host, l'identità verificata ed il numero di bagni.

Per gli entire home la probabilità che il guest sia donna dato che l'host è donna è pari a $\Phi(-1,0016+1*0,0197+0,1013+0,1870*4,4-0,0389*1,2)$, cioè $\Phi(-0,10) = 46,02\%$. La probabilità nel caso di private room diventa $\Phi(-0,8484+1*0,1569+0,0471+0,1767*4,5-0,0837*1,2)$, cioè $\Phi(0,06) = 52,39\%$. Il risultato ottenuto è molto interessante: se l'host è donna, la probabilità che il guest sia del medesimo sesso aumenta nel caso in cui si condivida l'appartamento. Esiste quindi una tendenza delle viaggiatrici a rivolgersi ad host di sesso femminile soprattutto quando è necessario condividere gli spazi, come avviene per la tipologia private/shared room. Questa conclusione offre diversi spunti di riflessione. La piattaforma Airbnb potrebbe utilizzare il sesso degli host e dei guest per riuscire ad eseguire i giusti match tra chi cerca e chi offre un listing, mettendo in risalto gli appartamenti delle donne quando chi viaggia è una persona di sesso femminile. Inoltre, gli host di sesso maschile potrebbero decidere di affittare l'intero listing per aumentare la privacy dei guest e riuscire ad ospitare quindi anche le donne.

11. Conclusioni

Il presente lavoro di tesi si poneva l'obiettivo di indagare eventuali differenze di genere presenti nella piattaforma di affitti di breve termini Airbnb. Inizialmente è stato presentato il mondo della sharing economy: il suo valore economico, i driver di diffusione, i principali aspetti regolatori e l'importante ruolo delle piattaforme digitali. Successivamente è stato analizzato il database delle recensioni di Airbnb relativo alla città di Bologna al fine di individuare eventuali differenze di comportamento tra i due sessi, sia lato guest sia lato host.

Dall'analisi descrittiva delle variabili utilizzate nel database sono emerse differenze significative tra uomini e donne, quali la tendenza delle viaggiatrici a scrivere recensioni di lunghezza maggiore rispetto agli uomini, oltre al fatto le host donna registrano più recensioni in cui sono chiamate direttamente per nome, segno di un livello di confidenza maggiore. Inoltre, è stato riscontrato che le donne guest scrivono più recensioni rispetto agli uomini e che sono più le donne host a possedere la qualifica di superhost. Per quanto concerne i punteggi è emerso che le donne registrano punteggi significativamente più alti nelle categorie review score rating, review score accuracy, review score cleanliness, review score check-in e review score communication. In merito al prezzo e alla domanda, non si evidenziano differenze significative tra i due sessi e tale risultato denota quindi una pari opportunità di guadagno.

È stata effettuata un'ulteriore analisi sulle recensioni per capire per capire quali fossero gli aspetti più rilevanti citati nelle review per i due generi. Le parole più frequenti, escluse le preposizioni e i pronomi, sono risultate per i due sessi apartment, Bologna, place, great. Tuttavia, è interessante sottolineare che le donne nelle loro recensioni parlano più della pulizia, mentre gli uomini più della location. Tali risultati sono stati utilizzati anche nelle successive analisi econometriche. In primo luogo si è studiata l'influenza sulla variabile dipendente prezzo, ponendo particolare attenzione al sesso dell'host e a sue caratteristiche (se possiede la

qualifica di superhost, se possiede più listings, ecc.), alla domanda, alle caratteristiche del listing (numero di letti, la distanza, ecc.) e ai punteggi. Inoltre, è stata valutata l'influenza di variabili di tipo qualitativo come le parole più utilizzate nelle recensioni (great, location, clean, helpful). Dal momento in cui l'analisi di regressione lineare non forniva alcun risultato interessante in merito alla variabile di interesse sesso dell'host, si è deciso di utilizzare un modello logaritmico dal quale è emersa un'influenza significativa e positiva del sesso dell'host: se l'host è donna, il prezzo è più alto. Questo risultato potrebbe essere giustificato da una maggiore attenzione da parte delle donne come si evince dai maggiori punteggi ottenuti nelle recensioni. Occorre sottolineare che, in un secondo momento, l'analisi è stata effettuata separando i listing di tipo Entire home da quelli di tipologia Private/Shared room. La variabile sesso dell'host è risultata significativa solo per la tipologia Entire home, probabilmente perché chi sceglie la tipologia Private room vuole risparmiare e non è disposto a pagare un sovrapprezzo per il sesso dell'host. Per quanto concerne le variabili rappresentanti le parole più utilizzate nelle recensioni, helpful ha un'influenza significativa e positiva per i listing Entire home, great è significativa per il dataset completo e per il dataset contenente la tipologia Entire home. I guest sono probabilmente disposti a pagare un prezzo più alto se le recensioni descrivono l'esperienza generale come positiva.

Le medesime analisi sono state effettuate considerando come variabile dipendente la domanda; è emerso che il sesso dell'host ha un'influenza significativa: se l'host è donna si registra una domanda più alta, sottolineando una preferenza dei guest per le host di sesso femminile. Inoltre, come nell'analisi di regressione sul prezzo, la variabile great ha un'influenza positiva e significativa.

Infine, si è indagato se esiste una tendenza delle viaggiatrici a rivolgersi ad host del medesimo sesso utilizzando la regressione probit. Si è osservato che la probabilità che il guest sia donna se l'host è donna è pari a circa il 54%. Inoltre, separando i listing in base alla tipologia è emerso che la probabilità è 46% per la tipologia Entire home, 52% se si condivide l'appartamento, quindi per la tipologia Private/Shared

room. Esiste quindi una tendenza delle viaggiatrici a rivolgersi ad host di sesso femminile soprattutto quando è necessario condividere gli spazi, come avviene per la tipologia Private/Shared room.

In conclusione, le differenze emerse tra i due sesso potrebbero essere utilizzate per migliorare il servizio offerto dalla piattaforma e si potrebbe attuare una strategia per migliorare il match tra guest e host, come mostrare alle viaggiatrici listings condivisi di proprietarie donne oppure porre in evidenza recensioni positive di altre donne nel caso il cui l'host sia di sesso maschile.

12. Bibliografia

- “Sharing Economy: Promote Its Potential to Sustainability by Regulation”, Xufeng Liu and Hongmin Chen, Department of Environmental Science and Engineering, Fudan University, 2020.
- “Environmental impacts and potential of the sharing economy”, John Magne Skjelvik, Anne Maren Erlandsen and Oscar Haavardsholm, 2017.
- “Navigating Peer-to-Peer Pricing in the Sharing Economy”, Gemma Newlands, Christoph Lutz e Christian Fieseler, 2018.
- “The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb”, Chiara Farronato e Andrey Fradkin, 2017.
- “Five Ways to Win with Digital Platforms”, Accenture, 2016
- “Price Determinants on Airbnb: How Reputation Pays Off in the Sharing Economy”, Tim Teubner, David Dann e Florian Hawlitschek, 2017.
- “Inside the Engine Room of Digital Platforms: Reviews, Ratings, and Recommendations”, Paul Belleflamme, Martin Peitz, 2018.
- “The Determinants of Online Review Informativeness: Evidence from Field Experiments on Airbnb”, Andrey Fradkin, Elena Grewal, David Holtz, 2018.
- “A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average”, Georgios Zervas, Davide Proserpio, John W. Byers, 2015.
- “You Get What You Give: Theory and Evidence of Reciprocity in the Sharing Economy”, Davide Proserpio, Wendy Xu, Georgios Zervas, 2018.
- “Does gender bias exist? The impact of gender congruity on consumer’s Airbnb booking intention and the mediating role of trust”, Na Su, Anna Mattila, 2020.
- “Strong Evidence for Gender Differences in Risk Taking”, Gary Charness, Uri Gneezy, 2011.
- “When guests trust hosts for their words: Host description and trust in sharing economy”, Iis P. Tussyadiah, Sangwon Park, 2017.

13. Sitografia

- <http://insideairbnb.com>
- https://www.repubblica.it/economia/rapporti/energitalia/lascossa/2019/11/15/news/sharing_mobility_e_sharing_economy_per_la_sostenibilita_-_241171429/
- <https://www.agendadigitale.eu/cultura-digitale/digital-disruption-ecco-la-formula-che-riunisce-presente-e-futuro/>
- <https://www.pwc.com/hu/en/kiadvanyok/assets/pdf/sharing-economy-en.pdf>
- www.altroconsumo.it
- <https://voltaitalia.org/wp-content/uploads/2016/03/Paper-05-sharing-economy-2.pdf>
- <https://www.letture.org/la-sharing-economy-chi-guadagna-e-chi-perde-mario-a-maggioni>
- <https://www.digitaldictionary.it/blog/report-digital-2020-scenario-digitale-mondo-e-italia>
- https://www.repubblica.it/viaggi/2020/06/11/news/riprendono_le_prenotazioni_in_italia_e_airbnb_cambia_strategia_turismo_di_prossimita_-_258953447/
- <https://www.ilsole24ore.com/art/il-boom-airbnb-spacca-mercato-affitti-locazioni-lunghe-crisi-ACXJNzj>
- <https://www.internazionale.it/notizie/2020/06/12/covid-sharing-economy>
- <https://www.investopedia.com/terms/n/network-effect.asp>
- <https://www.getsmarter.com/blog/career-advice/what-is-a-two-sided-market-and-why-does-it-matter/>
- <https://policyreview.info/articles/analysis/regulating-sharing-economy>
- <https://theconversation.com/the-three-regulatory-challenges-for-the-sharing-economy-37808>

- <https://chicagopolicyreview.org/2017/08/26/risks-and-regulations-of-the-sharing-economy/>
- <http://www.rosels.eu/2018/12/31/lapplicabilite-du-droit-europeen-de-la-concurrence-a-leconomie-collaborative/>
- <https://www.sciencedirect.com/science/article/pii/S2352146519301309>
- <https://www.intereconomics.eu/contents/year/2017/number/1/article/regulating-the-european-sharing-economy-state-of-play-and-challenges.html>
- <https://www.tandfonline.com/doi/full/10.1080/13683500.2020.1782855>
- <https://medium.com/enrique-dans/airbnb-a-case-study-in-diffusion-of-innovations-99b22444f276>
- [owler.com](https://www.owler.com)
- <https://business.trustedshops.it/blog/sistema-recensioni-chiuso-integrita-azienda>
- <https://www.lettera43.it/neuroscienze-cervello-uomini-donne-marketing/>
- <https://www.ildigitale.it/perche-il-marketing-parla-alle-donne-sono-il-motore-delleconomia-italiana/>
- <https://hbr.org/2018/05/what-most-people-get-wrong-about-men-and-women>
- <https://www.tandfonline.com/doi/abs/10.1080/10548408.2017.1284030>
- <https://www.hotelmanagement.net/operate/how-gender-and-safety-play-a-role-airbnb-s-popularity>
- <https://searchengineland.com/do-men-and-women-value-online-reviews-differently-309485>

14. Appendice

- Funzione Macro Excel per separazione parole recensioni

```
Sub CopiaInColonna()  
Copia i dati da una tabella, trasponendo tutte le righe su un'unica colonna  
Dim i As Integer  
Dim NumRows As Integer  
Application.ScreenUpdating = False  
'Disattiva l'update dello schermo  
Range("F100000").Select  
Selection.End(xlUp).Select  
NumRows = ActiveCell.Row  
'Salva il numero di righe piene in NumRows(a patto che ci siano meno di 1000000 righe)  
Range("F1").Select  
Range(Selection, Selection.End(xlToRight)).Select  
Selection.Copy  
Range("A1").Select  
Selection.PasteSpecial Paste:=xlPasteAll, Operation:=xlNone, SkipBlanks:= _  
    False, Transpose:=True  
'Copia la prima riga, per dare valori di start  
i = 1  
Do While i < NumRows  
    Range("F1").Select  
    ActiveCell.Offset(i, 0).Select  
    Range(Selection, Selection.End(xlToRight)).Select  
    Selection.Copy  
    Range("A1").Select  
    Selection.End(xlDown).Select  
    ActiveCell.Offset(1, 0).Select  
    Selection.PasteSpecial Paste:=xlPasteAll, Operation:=xlNone, SkipBlanks:= _  
        False, Transpose:=True  
    i = i + 1  
Loop  
'ciclo su tutte le altre righe  
Columns("A:A").Select  
    Application.CutCopyMode = False  
    Selection.Copy  
    Range("C1").Select  
    ActiveSheet.Paste  
    Columns("C:C").Select  
    Application.CutCopyMode = False  
    Selection.RemoveDuplicates Columns:=1, Header:= _
```

```

xlNo
'Crea un'altra colonna senza duplicati
  Application.ScreenUpdating = True
'Riattiva gli update a schermo
End Sub

```

- Traduzione review in inglese e definizione sesso guest

```

from google.colab import drive
drive.mount('/content/gdrive')
%cd /content/gdrive/My Drive/Code/
!pip install gender_guesser
!pip install translate
!pip install swifter
import gender_guesser.detector as gender
from translate import Translator
import swifter
import pandas as pd
df= pd.read_excel("Database_Bologna.xlsx")
#df.columns = ['Listing_id', 'Id', 'Guest_id', 'Name_Guest', 'Review']
d = gender.Detector()
df['Gender_Guest'] = df['Name_Guest'].apply(lambda x: d.get_gender(x))
translator= Translator(to_lang="en")
df['Review_ita'] = df['Review'].swifter.apply(lambda x: translator.translate(x))
df.drop(columns=['Unnamed: 5', 'Unnamed: 6'], inplace=True)
df.to_excel('Database_Bologna.xlsx')

```

- Parole più utilizzate nelle recensioni

```

nomeFileExcelInput= "Parole_femmine.xlsx"
nomeFileOutput= 'Parole_femmine.txt'
timeStart=datetime.now()
print(timeStart)
# Read and count
try:

```

```

print("Inizio lettura...")
pd.set_option('display.max_rows', None)
s = pd.read_excel(nomeFileExcelInput, sheet_name=None)
print("Fine lettura "+str(datetime.now()))
print("Inizio concat...")
stot=pd.concat(s) #Concatenamento fogli
print("Fine concat "+str(datetime.now()))
print("Inizio append...")
#stot2=stot['Numeri'].append(stot['Nomi']).append(stot['Prove']).append(stot['
Cose']).reset_index(drop=True)
    stot2=stot['f1'].append(stot['f2']).append(stot['f3']).append(stot['f4']).append(st
ot['f5']).append(stot['f6']).reset_index(drop=True)
    #stot2=stot['f1'].append(stot['f2']).reset_index(drop=True)
print("Fine append "+str(datetime.now()))
print("Inizio calcolo...")
res=stot2.value_counts(sort=True, ascending=False)
print("Fine calcolo "+str(datetime.now()))
print("Inizio scrittura file...")
with open(nomeFileOutput, 'w') as f:
    #print(res)
    print(res, file=f)
except Exception as e:
    print(e)
else:
    print("Calcolo finito. Risultati in file "+nomeFileOutput)
    timeEnd=datetime.now()
    print(timeEnd)

```

Ringraziamenti

Grazie alla Professoressa Laura Rondi per la sua disponibilità durante tutti questi mesi di stesura dell'elaborato, per i suoi consigli e per i numerosi spunti di analisi che ci ha offerto.

Grazie ai miei genitori. Mi avete sostenuto e supportato più di chiunque altro, credendo sempre in me anche quando io non lo facevo. Avete vissuto questi cinque anni con la stessa intensità con cui li ho vissuti io, sopportando tutti i miei musì, i miei mutismi, ma sempre in prima fila a festeggiare con me. Non avrei potuto chiedere genitori migliori.

Grazie a Roberta e Francesca. Saper di poter contare sempre su amiche come voi è una fortuna. Ringrazierò per sempre il Politecnico per averci fatte incontrare. Lo dirò sempre, una parte dei risultati che ho ottenuto è merito vostro, merito del vostro aiuto e dei vostri consigli. Grazie per essere sempre dalla mia parte e per sostenermi in tutto.

Grazie ad Arianna, Chiara, Giulia e Sabrina. Cosa sarebbero stati questi anni senza le nostre gite, i nostri viaggi, le nostre serate passate a spettegolare? Ormai siamo insieme da 10/13 anni, stiamo crescendo insieme e affronteremo le sfide che il futuro ci riserverà insieme.

Grazie ai miei compagni di classe, o meglio, compagni di viaggio. Nessuno meglio di voi sa cosa abbiamo passato per raggiungere questo traguardo. I momenti che abbiamo vissuto insieme hanno reso questi anni indimenticabili.

Grazie a Federico per avermi chiesto un fazzoletto il primo giorno di università fuori dall'aula 2. Grazie a Domenico per avermi parlato sul 58. Grazie per essere stati sempre con me in questi anni, affrontando insieme gioie e dolori del Politecnico dal 2015.