

# **POLITECNICO DI TORINO**

Collegio di Ingegneria Gestionale - Classe LM-31

Corso di Laurea in Ingegneria Gestionale



Tesi di Laurea Magistrale

## **IL FUNZIONAMENTO DEL MERCATO DEI DATI: DATA PROVISION E DATA PRICING**

Relatore:

Prof. Carlo Cambini

Candidato:

Alessandra Canova

Ottobre 2020

Anno Accademico 2019-2020

*“Women move mountains”*

# SOMMARIO

<b>INTRODUZIONE.....</b>	<b>6</b>
<b>CAPITOLO I – PANORAMICA GENERALE SUL DATA MARKET.....</b>	<b>8</b>
1.1. I big data.....	8
1.1.1. Il processamento dei dati .....	12
1.1.2. Tecnologie abilitanti il processamento dei dati .....	12
1.1.3. Il valore economico dei dati e la loro gestione .....	14
1.2. Le piattaforme digitali.....	16
1.2.1. Classificazione delle <i>digital platforms</i> .....	17
1.2.2. Le caratteristiche chiave .....	19
1.2.3. Gli effetti delle <i>digital platforms</i> per gli attori economici.....	20
1.3. La struttura del <i>data market</i> .....	21
<b>CAPITOLO II – RASSEGNA DELLA LETTERATURA.....</b>	<b>33</b>
2.1. Fallimenti di mercato .....	34
2.1.1. Monopolio naturale.....	34
2.1.2. Potere di mercato .....	35
2.1.3. Esternalità di rete .....	36
2.1.4. Costi di switching .....	43
2.2. Discriminazione di prezzo e inefficienze nel mercato dei dati .....	45
2.2.1. Modelli di pricing nel mercato dei dati.....	48
2.3. Abuso di posizione dominante: abusi di esclusione.....	62
2.3.1. La difesa degli innovatori .....	62
2.3.2. Il problema di Mergers and Acquisitions .....	64
2.3.3. I prezzi predatori: “Zero-Price problem” .....	65
2.4. Abuso di posizione dominante: abusi di sfruttamento .....	66

2.4.1.	L'eccessivo sfruttamento dei dati .....	66
2.5.	La privacy sul mercato dei dati .....	69
2.6.	Implicazioni di policy .....	74
2.6.1.	Portabilità e interoperabilità.....	74
2.6.2.	Sistemi open standard e trasparenza .....	76
2.6.3.	Ribaltamento dell'onere probatorio .....	77
2.6.4.	General Data Protection Regulation .....	77
2.6.5.	La gestione da parte dell'Antitrust.....	83
2.6.6.	Le autorità regolatorie in Europa .....	85
<b>CAPITOLO III – ESTENSIONE DEL MODELLO “SELLING COOKIES” .</b>		<b>88</b>
3.1.	Il costo della privacy .....	89
3.2.	Ambiente binario.....	90
3.2.1.	Analisi della domanda .....	90
3.2.2.	Definizione del prezzo .....	93
3.3.	Ambiente continuo .....	94
3.3.1.	Analisi della domanda .....	94
3.3.2.	Definizione del prezzo .....	97
3.4.	Applicazione del modello .....	99
3.5.	Osservazioni conclusive.....	106
<b>CONCLUSIONI.....</b>		<b>108</b>
<b>BIBLIOGRAFIA .....</b>		<b>113</b>
<b>SITOGRAFIA .....</b>		<b>118</b>
<b>RINGRAZIAMENTI.....</b>		<b>119</b>



## INTRODUZIONE

Dati, big data e piattaforme digitali sono tra le parole più nominate e più lette all'interno di un ambiente tecnologico, ma non solo. Grazie ai decisivi progressi che lo stato dell'arte sta realizzando, e continuerà a realizzare, per digitalizzare gran parte della quotidianità di ciascun individuo, oggi stiamo vivendo in una realtà in cui il dato è la chiave per trarre benefici e profitti. Il mondo dei servizi sta attraversando una complessa fase di *servitization*, l'ambiente produttivo sta diventando sempre più tecnologicamente all'avanguardia e il settore dell'agricoltura si sta evolvendo verso l'agri-tech supportato da strumenti quali l'Internet of Things. Il denominatore in tutti questi ambienti è l'evoluzione tecnologica, la quale è supportata dall'acquisizione e dallo scambio dei dati. I dati, infatti, rappresentano il motore dell'economia digitale ed è fondamentale conoscerne le caratteristiche in profondità per poter definire i vantaggi e i rischi ad essi connessi.

Lo scopo del presente lavoro è, in un primo momento, quello di mostrare il funzionamento del *data market* mediante l'analisi dell'ambiente competitivo, e successivamente, di focalizzarsi sull'aspetto dell'acquisizione e del *pricing* dei dati all'interno di un contesto che simboleggia il rapporto impresa-consumatore odierno.

Il primo capitolo presenta una panoramica generale sul mercato dei dati, analizzando nel dettaglio i due attori principali, i big data e le piattaforme digitali. Per quanto riguarda i primi, uno sguardo speciale sarà riservato alle tecnologie abilitanti il processamento dei dati, le quali, oggi più che mai, sono in una fase di forte crescita e progresso tecnico (si pensi ad esempio a tutto ciò che riguarda il mondo del Cloud o del 5G). Allo stesso tempo si approfondiranno le caratteristiche delle piattaforme digitali in relazione ai diversi effetti, di livello macroeconomico e microeconomico, che esse scatenano nei confronti di attori economici quali regione geografica, imprese e consumatori. A tutto ciò segue poi un focus sulla struttura e sul funzionamento del *data market*, identificando le diverse caratteristiche di tale mercato tramite punti di forza, punti di debolezza, minacce e opportunità.

Il secondo capitolo si concentra sull'analisi della letteratura pubblicata fino ad oggi, prendendo in considerazione alcuni studi più datati a confronto con quanto

divulgato recentemente. Inizialmente, si evidenziano le principali fonti di problemi e inefficienze all'interno del mercato quali i fallimenti di mercato, la discriminazione tariffaria e gli abusi di posizione dominante. In secondo luogo si passa allo studio della tutela della privacy dei consumatori, diritto fondamentale che non può non essere considerato, e, in ultimo, si presentano quelle che, secondo l'odierna letteratura, sono le principali policy necessarie per attenuare gli effetti negativi generati dalle inefficienze descritte in precedenza. Inoltre, all'interno di questo capitolo verrà eseguito un focus sul modello "Selling Cookies" pubblicato nel 2015 da Dirk Bergemann e Alessandro Bonatti poiché tale modello sarà l'oggetto chiave del capitolo successivo.

Il terzo capitolo, infine, propone un'estensione del modello di Bergemann e Bonatti in cui, tra tutti i costi che devono sostenere gli attori del sistema economico (i.e. consumatori individuali, *advertisers*, *data seller/data provider*), si considera, in aggiunta, il costo per la privacy dei consumatori. Si sottolineerà l'importanza di garantire un tale diritto, quanto esso influisca sulle strategie di azione degli *advertisers* e sul prezzo per i cookies stabilito dal *data provider*. I risultati ottenuti saranno poi interpretati considerando le esternalità di rete negative presenti tra i diversi attori del mercato; essi si dimostreranno perfettamente in linea con quanto proposto dagli autori all'interno del modello base che non considera il costo della privacy e con ciò che un individuo razionale si aspetti che accada. Verrà presentata, infine, l'applicazione matematica del modello al caso della nota piattaforma YouTube, grazie alla quale è possibile comprendere numericamente gli effetti che l'introduzione dell'estensione comporta all'interno dell'ambiente competitivo.

# CAPITOLO I

## PANORAMICA GENERALE SUL DATA MARKET

Il progresso tecnologico e i nuovi strumenti di cui si dispone oggi hanno portato ad un cambiamento incentrato sulla digitalizzazione dell'economia e sulla competizione all'interno dei mercati e tra mercati di molteplici settori. Per comprendere meglio questo cambiamento è fondamentale analizzare le caratteristiche chiave di uno dei principali contesti in cui esso si inserisce: il *data market*.

In un mercato digitale la competizione tradizionale, i.e. la competizione tra un grande numero di imprese con un tasso di asimmetria informativa che varia a seconda dei casi, non è sempre possibile poiché le nuove tecnologie di informazione hanno portato allo sviluppo di dinamiche competitive profondamente diverse. La presenza di economie di scala e di esternalità di rete sono solo due dei fattori scatenanti questo cambiamento ma in realtà ne esistono molti altri che verranno analizzati in seguito.

Un ruolo non di secondo piano è occupato dalle piattaforme digitali che attuano comportamenti spesso anticoncorrenziali nei confronti di competitor e consumatori, dove per consumatori si intende sia il cliente finale sia gli intermediari, i.e. coloro che distribuiscono un certo bene o servizio. Per questo motivo è necessario che si evolva il diritto alla concorrenza e che vi sia innovazione positiva per il consumatore, il tutto supportato dal fatto che il diritto alla concorrenza accompagna e guida l'evoluzione delle piattaforme digitali.

L'elemento chiave di questa economia digitale sono però i big data e le modalità con cui vengono sfruttati.

### 1.1. I big data

La prima domanda che sorge spontanea è chiedersi quando è lecito parlare di dati o di big data.

I dati sono descrizioni codificate di informazioni e costituiscono l'input principale del *data market*. L'uso che viene fatto dei dati si distingue in uso primario,

quando il dato risponde a specifiche domande di ricerca, e uso secondario, quando si sfrutta il valore “opzionale” del dato di cui spesso non se ne conosce l’esistenza al momento della raccolta. I dati sono considerati un asset informativo fondamentale che richiede nuove forme di lavorazione ed elaborazione delle informazioni attraverso le quali è possibile migliorare le decisioni, intuire scoperte e ottimizzare i processi.

Quando si parla di big data, invece, si fa riferimento a grandi quantità di tipi diversi di dati prodotti da varie fonti, fra cui persone, macchine e sensori. I dati vengono classificati come “big” quando ad essi ci si riferisce in termini di volume, varietà, velocità e valore (Delmastro e Nicita, 2019).

1. Volume. È la caratteristica che fa riferimento alla mole di dati trattati, ovvero alla dimensione del fenomeno, e viene misurato mediante Zettabyte (1 Zettabyte corrisponde a 1 miliardo di Terabyte, ovvero ad una capacità di archiviazione di circa 250 miliardi di DVD, se si considera un DVD con capacità di 4 Gigabyte). La mole di dati archiviati è esposta ad una crescita esponenziale soprattutto grazie all’evolversi di tecnologie come il 5G o l’Internet of Things.
2. Varietà. È la caratteristica con cui si evidenzia l’eterogeneità delle fonti di acquisizione dei dati, delle modalità di rappresentazione, analisi e immagazzinamento e delle sorgenti di acquisizione dei dati. I dati acquisiti dalle sorgenti possono essere:
  - a. Non strutturati, quando non sono organizzati secondo una precisa struttura e richiedono tecniche sofisticate che li trasformino in informazione, come nel caso di immagini, foto, testi, video, social media e molti altri ancora;
  - b. Strutturati, quando contengono una grande densità di informazioni organizzate secondo uno schema ben definito;
  - c. Semistrutturati, quando sono composti sia da dati strutturati che da dati non strutturati e con formati assai diversi.

Alcuni esempi di varietà per l’analisi dei dati sono dati da strumenti quali cookies, pop-up e *online advertising*.

3. Velocità. È la caratteristica che si riferisce alla velocità delle tempistiche e alla frequenza con cui i dati circolano da un punto di origine ad un punto di raccolta, ma anche alla velocità di processamento che è molto rapida e consente di prendere decisioni in tempo reale.
4. Valore. Questa caratteristica si può considerare come diretta conseguenza delle tre appena citate ma è anche quella che ricopre maggior importanza. Con valore si intende la capacità di estrarre valore economico dai big data a favore delle imprese, con il lancio di nuovi prodotti, servizi e pubblicità targettizzate, e più in generale, a favore della società, sfruttando la sempre crescente mole di dati per incrementare il benessere complessivo. Tuttavia, definire un valore per i big data è molto complicato proprio a causa dell'interdipendenza tra molteplici fattori.

Esistono poi delle ulteriori caratteristiche da prendere in considerazione quando si trattano i big data:

- la veridicità, la fiducia che si può riporre in essi;
- la valenza, la presenza di connessioni sempre più dense e complesse;
- la visualizzazione, ciò che dà valore al dato rispetto al livello dei dati grezzi;
- la sensibilità, il contenuto dei dati riguarda informazioni sensibili, riservate, confidenziali;
- la qualità e gli aspetti complementari di un dato ovvero la sua completezza informativa e accuratezza.

Le principali differenze tra dati e big data sono evidenziate nella Tabella 1.

Tabella 1 – Differenza tra dati e big data - fonte: Kitchin e McArdle, 2016

	<b>DATI</b>	<b>BIG DATA</b>
<b>VOLUME</b>	Da dati a limitati a molti dati	Molti dati
<b>VELOCITA'</b>	Lenta, frequenza non continua	Veloce, frequenza continua
<b>VARIETA'</b>	Da limitata ad ampia	Ampia
<b>ESAUSTIVITA'</b>	Campioni	Tutta la popolazione
<b>INTERRELAZIONI TRA DATI</b>	Da debole a forte	Forte
<b>ESTENSIONE E SCALABILITA'</b>	Da bassa a media	Alta

I dati, inoltre, si dividono in due categorie, dati personali e dati non personali, al fine di identificare i vincoli normativi applicabili alla loro raccolta e all'uso di un certo tipo particolare di dato (Bourreau, Graef, de Stree, Valletti, 2017).

I dati personali riguardano qualsiasi informazione relazionata ad un soggetto identificabile o identificato, ovvero una persona. Fanno parte dei dati personali

- i dati sensibili: origini etniche, opinioni politiche, religiose o filosofiche, dati genetici o biometrici e molti altri;
- i dati pseudonimati: dati che non possono essere attribuiti ad un soggetto specifico se non vengono fornite informazioni aggiuntive, come i dati impiegati per indagini statistiche o scientifiche.

A differenza dei dati personali, la cui disciplina è regolata dal GDPR (General Data Protection Regulation), i dati non personali, invece, sono quei dati non protetti da nessuna normativa o legge. A loro volta essi si distinguono in

- dati anonimi: dati che con nessun mezzo, né di chi possiede il dato né da parte di terzi, possono venir ricondotti all'utente. L'utente non è identificabile;
- dati anonimizzati: dati personali che sono resi anonimi per rendere l'utente non identificabile. Sono dati che si riferiscono ad un individuo la cui identificazione

non è più possibile. Se il processo per rendere i dati anonimi è irreversibile, tali dati non rientrano nel campo di protezione delineato dal GDPR.

### **1.1.1. Il processamento dei dati**

Processare i dati è ciò che conferisce loro valore: un dato non trattato possiede un valore relativamente limitato. I big data vengono raccolti, scambiati, immagazzinati e se ne estrae valore attraverso un complesso ecosistema in cui coesistono molti mercati (*multi-sided market*). Poiché è importante per i possessori dei dati estrarre valore dai dati e incrementarne il potenziale, il processamento dei dati ricopre un ruolo chiave. Si possono distinguere due categorie di azioni: la gestione del processamento dei big data e l'analisi dei big data.

Il processamento dei big data consiste nell'implementazione di algoritmi per la raccolta e l'immagazzinamento dei dati (*data collection* e *data storage*), nell'applicazione di processi scalabili e flessibili che permettano l'acquisizione di un volume sempre crescente di dati, nella distinzione tra un processo di acquisizione dati online e uno offline.

Mediante l'analisi dei big data (*data analytics*) invece è possibile collezionare informazioni utili per le imprese. Questo è infatti uno dei principali obiettivi di molte organizzazioni.

### **1.1.2. Tecnologie abilitanti il processamento dei dati**

Dal punto di vista tecnico, sono molteplici le tecnologie che svolgono il ruolo di abilitatrici per il processamento dei dati<sup>1</sup>. In particolare per quanto riguarda la *data collection* si può parlare di rete 5G e Internet of Things, per il *data storage* di Cloud Computing e Web Caching, per la *data analytics* di Machine Learning e Cookies.

- Rete 5G. Il passaggio della rete dal 4G LTE (Long Term Evolution) al 5G ha lo scopo di implementare tecniche di trasmissione delle informazioni e dei dati ancora più evolute di quanto già presentato fino al giorno d'oggi. Gli obiettivi principali sostenuti dalla rete di quinta generazione riguardano l'aumento della

---

<sup>1</sup> Tratto da: “*Data Center, IoT, and Protocols, 5G, Web Services*. Slides del corso di Strumenti ICT per la gestione dei Servizi a cura del Professor Paolo Giaccone. Politecnico di Torino. A.A. 2019-2020.

velocità di trasmissione per utente che arriverà fino a 1Gbit/secondo, la diminuzione della latenza (i.e. il ritardo complessivo nella trasmissione) che deve aggirarsi tra 1 e 5 millisecondi, un aumento della copertura della rete riuscendo a servire 1 milione di dispositivi per chilometro quadrato grazie alla riduzione delle celle di copertura (questo abilita anche una trasmissione a minore potenza, evidenza scientifica del fatto che il 5G non è dannoso per la salute dell'uomo, come ribadito anche dall'Istituto Superiore di Sanità<sup>2</sup>), maggiore affidabilità e possibilità di comunicazione *device-to-device*. L'impiego di una rete più veloce facilita il processo di acquisizione dei dati degli utenti.

- Internet of Things. Un maggior numero di imprese e settori sta guardando con interesse crescente all'IoT, sia per la capacità delle applicazioni di poter collezionare dati sia poiché l'IoT contribuisce al processo di *servitization* al quale si sta assistendo oggi. Grazie ai sensori e alla loro possibilità di fare *sensing* degli oggetti connessi in rete che li circondano, è possibile oggi, e lo sarà ancora di più in futuro, acquisire con estrema facilità i dati e le informazioni dagli utenti.
- Cloud Computing. È un modello per permettere l'accesso tramite rete ad un insieme di risorse, quali servizi di *computing*, di *storage* e servizi di rete, che sono configurabili e condivise. Tali risorse possono essere procurate e rilasciate in breve tempo e con il minimo sforzo di gestione (Grance e Mell, 2011). Grazie al Cloud, quindi, si hanno a disposizione degli strumenti che lavorano in maniera elastica e che forniscono agli utenti e alle imprese, tra le altre cose, applicazioni e servizi di *data storage*. Un valido esempio è dato dal Web Service di Amazon "Simple Storage Service S3" il quale sfrutta il Cloud per fornire il servizio di archiviazione dei dati.
- Web Caching. Una web cache è un'entità di rete che, agendo come *proxy server*, soddisfa richieste al posto dei comuni *web server* e lo fa venendo

---

<sup>2</sup> Polichetti, A. (28 Aprile 2020). *5G e salute: cosa c'è da sapere*. Centro Nazionale per la Protezione dalle Radiazioni e Fisica Computazionale. Istituto Superiore di Sanità. [www.iss.it/primo-piano/-/asset\\_publisher/o4oGR9qmvUz9/content/id/5355363](http://www.iss.it/primo-piano/-/asset_publisher/o4oGR9qmvUz9/content/id/5355363).

implementata su pc di utenti diversi. Lo scopo di avere questa unità “centralizzata” è quello di diminuire il tempo in cui vengono fornite informazioni e soddisfatte le richieste dei *client* poiché la cache immagazzina localmente i dati degli utenti.

- Machine Learning. Con il termine Machine Learning si intende una specifica applicazione delle tecniche di Intelligenza Artificiale, le quali, tramite analisi statistiche di enormi quantità di dati, elaborano algoritmi volti all’apprendimento supervisionato, non-supervisionato o rinforzato. Il ML “si nutre” dei dati degli utenti per produrre applicazioni che, in alcuni casi, vanno a danneggiare le attività *labour/capital intensive* che possono essere facilmente standardizzate.
- Cookies. Quando si parla di cookies ci si riferisce a piccole porzioni di dati degli utenti. I cookies sono stati ideati da un punto di vista prettamente tecnico per sopperire al fatto che l’interazione di tipo *client-server* è senza memoria, *stateless*, infatti essi consentono al server di tenere traccia degli utenti, delle loro mosse. I cookies fungono da tecnologia abilitante per la *data analytics* poiché consentono fenomeni quali la *targetability*, ovvero l’abilità di collezionare e usare dati per determinare quale tipo di cliente sia più soggetto ad una particolare pubblicità, e la *measurability*, ovvero il riuscire a valutare e analizzare il comportamento dell’utente dopo che è stato sottoposto alla visione della pubblicità. Tutto ciò ha implicazioni di privacy non indifferenti; tale aspetto verrà approfondito nel capitolo seguente.

### **1.1.3. Il valore economico dei dati e la loro gestione**

Mediante le transazioni digitali i dati degli utenti vengono acquisiti e scambiati con l’obiettivo di utilizzarli come prodotto finale o come input per le piattaforme digitali che offrono nuovi beni e servizi.

I dati, considerati quasi come un bene pubblico, godono del fatto che non esiste rivalità nel consumo poiché l’utilizzo degli stessi da parte di un utente non preclude l’utilizzo ad altri soggetti e non esiste escludibilità, ovvero non se ne può impedire la fruizione ad utenti che non hanno pagato per averli. In assenza di diritti proprietari ben

definiti e interventi regolatori, si riscontrano quindi situazioni in cui il *digital market* è incapace di generare autonomamente un livello ottimale di produzione del bene e permette sovra-sfruttamento della risorsa da parte di terzi. Tutto ciò porta ad una sola conseguenza: un fallimento di mercato (Delmastro e Nicita, 2019).

Un'altra forma di fallimento di mercato presa in considerazione in presenza di dati è il potere di mercato della piattaforma che possiede tali dati. Per comprendere la relazione tra big data e potere di mercato è importante chiedersi come quest'ultimo è legato al valore economico dei dati. In particolare, è necessario determinare l'estensione delle economie di scala e il beneficio di cui risente la qualità del processo se si incrementasse la mole di dati archiviati, l'estensione delle economie di scopo e la svalutazione del dato legata alla sua "età". Anche la disponibilità dei dati influenza il potere di mercato. Le applicazioni e gli algoritmi hanno a disposizione una certa mole di dati e possono collezionarne ulteriori direttamente dall'utente, dalle macchine o dal mercato stesso. Maggiore è la mole di dati a loro disposizione maggiore sarà il potere di mercato. Tuttavia vi sono aspetti legali riguardanti la raccolta e l'uso dei big data che non si possono trascurare.

Il volume dei dati è una caratteristica che porta con sé ulteriori controversie. L'acquisizione e l'elaborazione di grandi volumi di dati richiede la necessità di adottare nuovi metodi di gestione e conseguenti ingenti investimenti in infrastrutture tecnologiche che solo le organizzazioni più grandi e strutturate possono permettersi. Le imprese più piccole affrontano questa transizione con maggiore cautela.

La reingegnerizzazione delle infrastrutture è necessaria poiché è indispensabile incrementare gli spazi fisici di *storage* e computazione dei dati, i data center. I data center, infatti, sono costituiti da un grande numero di dispositivi fisici quali hard disk, memoria SSD, server, router e switch che interconnettono le risorse di calcolo, racks (armadi con formati standard a seconda delle applicazioni) e sistemi di raffreddamento e approvvigionamento dell'energia. Incrementando questi spazi si contribuisce a velocizzare gli accessi alla rete, a bilanciare la velocità e la capacità di immagazzinare dati in dispositivi drive e ad aumentare la capacità di *backup* dei dati. In aggiunta, è importante che la gestione e il processamento dei dati avvenga in grandi database che si appoggiano su software e strutture hardware altamente performanti.

Infine, un fattore da non trascurare per quanto riguarda la gestione dei big data è l'aspetto connesso all'appropriazione di dati sensibili, l'aspetto legale, il quale deve rispettare le normative dei paesi in cui ci si trova ad operare.

Dopo aver presentato i principali aspetti relativi ai dati e ai big data, si analizzano ora le strutture per le quali i dati sono costituiscono l'asset principale per la creazione di valore.

## **1.2. Le piattaforme digitali**

Esistono numerose definizioni di piattaforme digitali. Quelle riportate di seguito evidenziano due punti di vista differenti tramite i quali tali piattaforme possono essere analizzate. Secondo la visione prettamente tecnica proposta da Spagnoletti, Resca, Lee (2015), le piattaforme digitali vengono definite come *“a building block that provides an essential function to a technological system and serves as a foundation upon which complementary products, technologies, or services can be developed”*. Nello specifico, ci si focalizza sullo sviluppo tecnico e sulle funzioni attraverso cui prodotti complementari e servizi possono essere sviluppati.

Rochet e Tirole (2003) hanno proposto, invece, il loro punto di vista non tecnico tramite cui le piattaforme digitali sono state definite come *“two-sided network that facilitate interactions between distinct but interdependent groups of users, such as buyers and suppliers”*. Il focus in questo caso avviene sull'interazione tra differenti gruppi di utenti. Inoltre, secondo la visione di Pagani (2013), *“multisided platform exists wherever a company brings together two or more distinct groups of customers (sides) that need each other in some way, and where the company builds an infrastructure (platform) that creates value by reducing distribution, transaction, and search costs incurred when these groups interact with one another”*.

Sono molte le ragioni per le quali il *data market* ha dato vita alle piattaforme digitali. In primo luogo, le piattaforme, viste come intermediari tra venditore e acquirente, facilitano il matching, riducono quindi le frizioni di mercato grazie ai bassi *searching cost* e costituiscono un driver per la produttività. Le piattaforme inoltre rendono gli scambi più efficienti riuscendo a servire un grande numero di utenti

velocemente e confidando sul fatto che per gli altri bisogni dei clienti ci si affida a piattaforme complementari. Le piattaforme digitali costituiscono, quindi, uno dei principali strumenti attraverso i quali l'uomo organizza le proprie attività, incluso le interazioni economiche, sociali e politiche. Con la nascita delle piattaforme digitali molti settori hanno iniziato ad evolvere e le imprese che hanno fatto leva su questo cambiamento hanno visto una crescita notevole in dimensioni e scala (Crémer, de Montjoye, Schweitzer, 2019).

Allo stesso tempo, le piattaforme digitali sono entrate a far parte di quella che viene definita una *disruptive innovation*, ovvero tutti quei meccanismi che hanno sbaragliato il tradizionale modello del *brick-and-mortar*<sup>3</sup>, con conseguente applicazione di approcci *disruptive* nel mondo dei servizi che hanno dato il via libera a controversie riguardanti la regolazione di questo settore.

### **1.2.1. Classificazione delle *digital platforms***

Le piattaforme digitali sono soggette a dinamiche comuni che però si manifestano diversamente a seconda di come le imprese sono organizzate nel *data market*. Per questo motivo, esse vengono classificate in quattro tipologie (Fig. 1):

1. Piattaforme di transazione. Sono le tecnologie, i prodotti o i servizi che agiscono come intermediari facilitando gli scambi o le transazioni tra diversi utenti, acquirenti o fornitori che altrimenti avrebbero avuto delle difficoltà nel matching reciproco. Le transazioni producono come risultato la riorganizzazione delle risorse e degli asset e questo porta con sé la tendenza a monopolizzare i clienti;
2. Piattaforme di innovazione. Sono le tecnologie, i prodotti o i servizi che costituiscono le fondamenta per le imprese con un'idea innovativa (ma scarsamente organizzate) che vogliono sviluppare prodotti, servizi o tecnologie complementari. Una caratteristica chiave delle piattaforme di innovazione è che qualsiasi altra piattaforma può entrare a fare parte del pool di potenziali innovatori, l'ecosistema innovativo. Ciò significa che qualsiasi persona può

---

<sup>3</sup> Con il termine *brick-and-mortar* si identifica un'attività legata all'economia reale, caratterizzata dall'esistenza di strutture aziendali fisiche in cui i clienti possono recarsi di persona per vedere ed acquistare i prodotti.

usare liberamente le risorse della piattaforma e i proprietari della stessa non hanno necessità di conoscere in anticipo chi o dove sono questi innovatori esterni che cercano di entrarne a far parte, diversamente da quanto avviene nei mercati tradizionali dove chi coordina la *supply-chain* cerca di conoscere in anticipo chi saranno i potenziali innovatori. Questo tipo di piattaforma costituisce quindi un magnete per innovatori di prodotti o servizi complementari;

3. Piattaforme integrate. Sono allo stesso tempo sia piattaforme di transazione sia di innovazione. Una caratteristica tipica delle piattaforme integrate è che cercano di attrarre il maggior numero di utenti possibili poiché traggono guadagno dagli effetti di rete;
4. Piattaforme di investimento. Sono un insieme di compagnie che hanno sviluppato un portfolio strategico di piattaforme e agiscono come una *holding*, una piattaforma attiva di investimento o entrambe con lo scopo di fornire un ventaglio ampio di soluzioni per le esigenze dei consumatori.

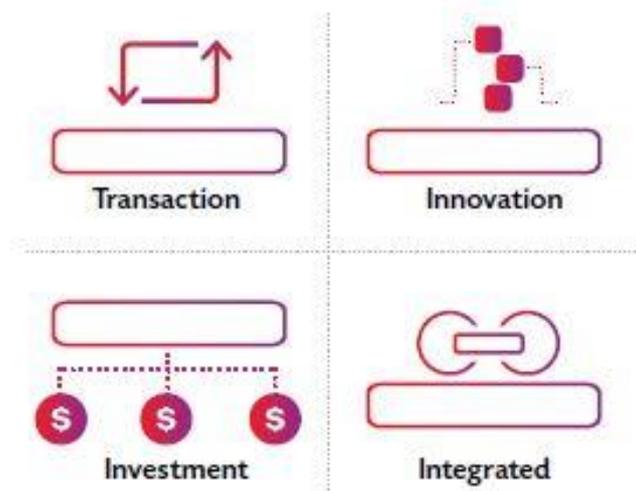


Figura 1 - Le diverse tipologie di piattaforme – fonte: Evans e Gawer (2016)

All'interno del mercato digitale, le piattaforme che ottengono maggior successo in termini di profitto sono le piattaforme integrate, il cui prestigio dipende da una forte presenza di effetti di rete e dall'ammontare non irrilevante di informazioni

di cui gode il cliente finale (scarsa asimmetria informativa) tale per cui si lascia che sia il mercato a guidare la competizione.

### 1.2.2. Le caratteristiche chiave

Le caratteristiche chiave di queste piattaforme sono intrinsecamente legate all'uso dei big data nel mercato. Infatti, le piattaforme digitali sono proprietarie dei dati in circolazione nella rete, i dati degli utenti, e questa è la ragione per la quale esse vengono anche chiamate piattaforme dominanti: i big data sono considerati l'asset fondamentale del mercato digitale e il loro possesso colloca tali piattaforme in una posizione di vantaggio competitivo.

Le piattaforme digitali, talvolta chiamate anche piattaforme online, presentano caratteristiche particolari non solo dal punto di vista tecnologico ma anche economico. Nello specifico, si riporta qui di seguito la visione di Evans e Gawer (2016) secondo cui le piattaforme digitali

- contribuiscono ad una significativa riduzione nei costi di transazione (costi di distribuzione, ricerca, contrattazione e monitoraggio);
- sono dominate dagli effetti di rete, diretti e indiretti (meglio specificati in seguito);
- sono molto performanti attraverso diverse dimensioni come efficienza, interazione produttiva, velocità di innovazione e scala;
- si fondano su obiettivi strategici dominati dai concetti di *lean* e *just-in-time delivery*;
- incentivano, attraverso adeguati strumenti tecnico-regolatori, utenti e service provider esterni alle piattaforme a partecipare allo sviluppo delle piattaforme stesse, fornendo anche prodotti complementari;
- sono utilizzate da un numero elevatissimo di utenti in grado di fornire nuovi contributi e generare nuovi *outcome* per problemi di diversa natura. Un esempio è dato dalla creazione delle piattaforme di crowdsourcing;
- permettono all'utente un elevatissimo *engagement*, ovvero la partecipazione e condivisione sociale, grazie alla quale è possibile tracciare con maggior velocità il profilo di gruppi di utenti che usufruiscono della piattaforma.

Inoltre, nei mercati in cui generalmente operano le piattaforme digitali si osservano strutture di costo caratterizzate da elevati costi fissi irrecuperabili (costi *sunk*) per la produzione dei beni informativi, ma costi marginali quasi nulli poiché il costo di un'unità aggiuntiva del bene è irrilevante. In aggiunta, i costi medi sono decrescenti.

### **1.2.3. Gli effetti delle *digital platforms* per gli attori economici**

La digitalizzazione dei mercati ha scatenato molteplici conseguenze dal punto di vista economico e le piattaforme online ne sono le protagoniste. A questo proposito si prendono in considerazione quattro attori economici attraverso i quali è possibile esplicitare tali conseguenze (Goldfarb e Tucker, 2019).

A livello macroeconomico si analizzano gli effetti sul paese e dal punto di vista regionale. Nel primo caso, come già mostrato in precedenza, Internet e le piattaforme digitali facilitano il commercio. In aggiunta, gli investimenti in tecnologie digitali influenzano i flussi del commercio online e fisico e la distanza tra intermediari e acquirenti (o più semplicemente tra *buyers* e *sellers*) non gioca più un ruolo chiave. A livello regionale invece, le tecnologie digitali hanno portato ad una redistribuzione dei benefici economici tra città e aree rurali: le città godono di vantaggi sproporzionati rispetto alle aree rurali ma i bassi costi di comunicazione favoriscono anche le zone più isolate. Generalmente però si considera che più una tecnologia è di difficile implementazione, maggiore sarà il beneficio dato dagli effetti di agglomerazione e quindi il trade-off si sposta verso le città.

A livello microeconomico invece si prendono in considerazione le imprese e i consumatori. Concentrandosi sulle imprese, le *digital platforms* sono state promotrici di un incremento di produttività e questo è il risultato del fenomeno di *reverse causality* che caratterizza il mondo dell'ICT. Infatti, l'influenza che le suddette piattaforme esercitano sulla produttività delle imprese viene considerata come un fattore endogeno: le imprese hanno cominciato ad investire nell'ICT con conseguenti incrementi di produttività e allo stesso tempo la crescente produttività delle imprese ha permesso di compiere diversi investimenti tecnologici. Ulteriori analisi sostengono, inoltre, che questo incremento di produttività abbia giovato solo a quelle imprese che già prima che si compissero gli investimenti godevano di risultati profittevoli. Dal punto di vista del consumatore gli effetti dettati dallo sviluppo delle *digital platforms*

si dividono in aspetti positivi legati a meccanismi quali effetti di rete, discriminazione di prezzo in base alla *willingness to pay*, incrementi di benessere, aspetti negativi riguardanti l'abuso dei meccanismi appena descritti e aspetti negativi che non vengono catturati dalle analisi sul surplus, come i danni di natura psicologica-sociale, i quali verranno discussi approfonditamente in seguito.

### 1.3. La struttura del *data market*

La competizione sul mercato dei dati ha portato alla definizione di una nuova struttura di mercato e all'incremento del potere di mercato delle imprese che vi fanno parte. Questo è dovuto non all'insorgere di nuove caratteristiche ma dalla combinazione di più fattori pre-esistenti.

Si analizzano ora questi fattori mediante lo strumento dell'analisi SWOT con cui è possibile valutare strategicamente i punti di forza (*Strengths*), i punti di debolezza (*Weaknesses*), le opportunità (*Opportunities*) e le minacce (*Threats*) che caratterizzano l'odierno mercato dei dati (Fig. 2).

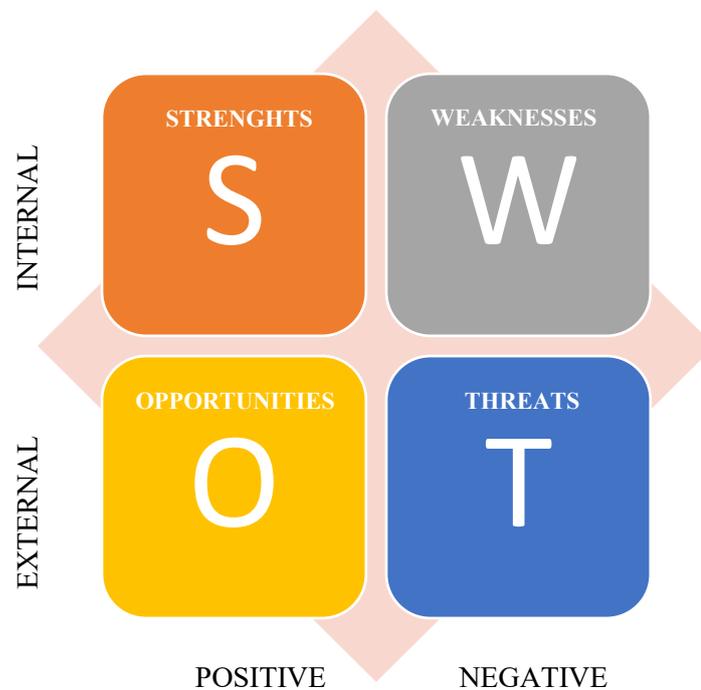


Figura 2 – Framework SWOT

## **PUNTI DI FORZA**

### *Economie di scala*

Con la presenza di economie di scala il costo di produzione dei beni e servizi digitali è decisamente meno che proporzionale rispetto al numero di utenti che usufruiscono di tali prodotti (i consumatori serviti) e questo è un vantaggio competitivo notevole per gli incumbent di un determinato settore all'interno del *data market* (Crémer, de Montjoye, Schweitzer, 2019). Le economie di scala hanno contribuito alla nascita delle piattaforme digitali, le quali hanno sfruttato il fatto che l'informazione o il dato, una volta creati, possono venir trasmessi ad un gran numero di persone a costi molto bassi. Infatti, nonostante il costo di creazione del servizio e della gestione fisica delle infrastrutture sia elevato, il costo marginale per servire un utente aggiuntivo è pressoché nullo e più la base di utenti si allarga, maggiore è la riduzione sui costi medi dell'impresa.

All'interno del *data market* le economie di scala costituiscono, quindi, la chiave iniziale del successo ma anche una spinta per la crescita futura poiché la creazione di valore nel mercato attira nuovi utenti. Questa è la dinamica che permette alle piattaforme digitali una crescita sostenibile nel tempo.

### *Economie di scopo e varietà*

Le economie di scopo e varietà rientrano tra quei fattori che spiegano la crescita del mondo digitale. Tramite le economie di scopo molti *service provider* di servizi online sfruttano la condivisione e il raggruppamento dei dati, delle competenze tecniche e delle relazioni esistenti tra utenti e fornitori del servizio per ridurre i costi e/o migliorare la qualità di un servizio. In aggiunta, grazie alle economie di varietà è possibile ridurre i costi relativi all'offerta congiunta di servizi.

Le economie di scopo e varietà sono quindi la ragione per cui un piccolo gruppo di imprese, tendenzialmente sempre le stesse, hanno potuto costruire un ecosistema di successo espandendo la loro presenza in mercati adiacenti.

### *Riduzione dei costi*

La riduzione delle seguenti tipologie di costo è promotrice dei cambiamenti che si verificano nei modelli economici standard (Goldfarb e Tucker, 2019).

- **Costi di ricerca.** Con l'avvento di Internet è diventato estremamente semplice trovare e comparare informazioni inerenti prodotti e servizi e questo ha facilitato le transazioni online. Empiricamente si è riscontrato che il costo dei beni venduti online è inferiore rispetto al costo dei beni venduti offline: l'abbassamento dei costi di ricerca ha intensificato la competizione tra *retailer*. Allo stesso tempo però la dispersione di prezzo tra imprese non è diminuita. Questo effetto si spiega con il fatto che diversi *retailer* applicheranno prezzi diversi poiché la qualità del bene o servizio che offrono spesso è diversa, com'è anche diversa l'esperienza di acquisto messa in atto. La riduzione dei *searching cost* è positiva poiché in presenza di tali costi le imprese manipolano il processo di ricerca in modo da sostenere alti margini e prezzi. I bassi costi di ricerca inoltre promuovono i prodotti superstar e quelli appartenenti alle code della distribuzione (i prodotti di nicchia) e incrementano il matching tra venditore e consumatore riuscendo a ricavare informazioni sulla domanda di mercato.
- **Costi di replicazione.** I beni in gioco sul mercato digitale sono i dati, i.e. i "beni fatti di bit", che si contrappongono ai classici prodotti e servizi dei tradizionali mercati. La grande differenza tra queste due categorie di beni non risiede nell'assenza di costi marginali (in oligopolio anche le imprese che hanno una domanda con  $MC > 0$  raggiungono l'equilibrio nel punto in cui  $MR = 0$ ) ma nel fatto che i beni-bits non sono rivali e possono essere "consumati" da un individuo senza ridurre la quantità di tali beni a disposizione di altri soggetti. Il costo di replicazione dei dati è nullo poiché, salvo limiti legali o tecnologici, i dati possono essere riprodotti da chiunque senza degradare la qualità iniziale del bene iniziale.
- **Costi di trasporto.** Se si considera nullo il costo di congestione della rete poiché legato ad aspetti puramente tecnologici, non esiste o è davvero basso un costo di trasporto dei beni digitali, i dati. In aggiunta, la differenza di costo tra comunicazioni vicine e distanti è quasi nulla. Uno dei fattori che in parte

influisce sul costo di trasporto delle informazioni è però il luogo in cui si trova l'utente/consumatore nel momento in cui usufruisce del bene: è vero che potenzialmente l'utente ha accesso agli stessi dati ovunque ma gli Stati, con le loro politiche di policy, possono influenzare il comportamento dei consumatori e veicolare l'accesso al dato.

- Costi di tracking. Il tracking delle informazioni permette fenomeni come personalizzazione dell'offerta, pubblicità targettizzata e discriminazione tariffaria. Se diminuiscono molto i costi di tracciamento dei dati, essi vengono registrati e organizzati sempre a costi minori da cui derivano personalizzazioni di offerta e creazione di mercati *one-to-one* con il consumatore. Questo fenomeno rende possibile esercitare discriminazione tariffaria e creare modelli potenzialmente pericolosi dal punto di vista dell'advertising. A seconda dei punti di vista, la diminuzione dei costi di tracking è un fattore che presenta aspetti positivi e negativi. Assumendo i panni di un consumatore, nonostante la discriminazione tariffaria sia un problema, ciò che potenzialmente è più dannoso è l'esposizione del consumatore alla ricezione di più appropriate, rilevanti e profittevoli pubblicità (*online targeting technology*).
- Costi di verifica. Tali costi sono associati alla verifica dell'identità e della reputazione del dato. Un modello di reputazione, ovvero ciò su cui si basano i costi di verifica, sono le recensioni online che gli utenti redigono a costo zero. Un modello di questo tipo permette a tutti gli utenti di azzerare il loro costo di verifica delle informazioni e permette a colui che vende un servizio di incrementare profitti e ricavi (in caso di recensioni positive). Il fallimento di questo modello risiede nel fatto che non tutti i consumatori impiegano strumenti di recensioni e nel caso in cui succedesse, non vi sono garanzie che le recensioni siano coerenti e oneste.

## PUNTI DI DEBOLEZZA

### *Consumer welfare*

In un'economia dominata dai dati il prezzo dei servizi non rappresenta il valore sociale che essi incarnano. Nei mercati tradizionali, dove ciò invece è vero, il prezzo viene usato per determinare le variazioni del benessere dei consumatori indotte dalle diverse politiche economiche. Nei mercati in cui sono presenti esternalità di rete, e più in generale all'interno del *data market*, il prezzo non riflette il valore sociale dei beni poiché non tiene conto dei benefici addizionali dati dalle esternalità e quindi determinare le variazioni di *welfare* per i consumatori è davvero difficile. Il benessere dei consumatori è fortemente influenzato dalla personalizzazione o discriminazione tariffaria e questo principalmente è causa di un incremento di benessere che però non può essere quantificato. Infatti, la *price discrimination* da un lato crea beneficio rendendo il mercato accessibile anche a quei consumatori con una bassa disponibilità a pagare, dall'altro lato può essere dannosa poiché può implicare il trasferimento del welfare dal consumatore al *service provider*.

Allo stesso modo, anche le informazioni che vengono estrapolate dai big data si sostiene possano generare un aumento del benessere dei consumatori e una crescita del PIL, e.g. le informazioni mirate alla creazione di *target advertising*, ma questi guadagni di efficienza sono difficili da misurare.

Conseguentemente a questa incertezza nella determinazione del benessere, il consumatore si affida all'applicazione di leggi antitrust che dovrebbero invogliare l'entrata in settori dominati dalle grandi piattaforme digitali e che sostenere la competizione, costituendo un punto a sostegno del *consumer welfare* ma un fattore di debolezza per gli incumbent del *data market*.

### *Danni*

Con il termine danni si intendono due casistiche differenti: i danni di natura psicologica-sociale (relativi all'uso di Internet) e i danni di natura economica nei confronti dei consumatori. Internet è entrata a far parte della vita quotidiana di moltissime persone. Purtroppo però costituisce anche un mezzo attraverso il quale contenuti illegali e inaccettabili vengono diffusi online. Il problema risiede nella scarsa

protezione dei dati personali degli utenti che ne usufruiscono e che sono quindi esposti a danni di molteplice natura: cyberbullismo, sfruttamento di minori, abusi e violenze online, terrorismo, vendita di sostanze illegali e molti altri (Fig. 3).

Harms with a clear definition	Harms with a less clear definition	Underage exposure to legal content
<ul style="list-style-type: none"> <li>• Child sexual exploitation and abuse.</li> <li>• Terrorist content and activity.</li> <li>• Organised immigration crime.</li> <li>• Modern slavery.</li> <li>• Extreme pornography.</li> <li>• Revenge pornography.</li> <li>• Harassment and cyberstalking.</li> <li>• Hate crime.</li> <li>• Encouraging or assisting suicide.</li> <li>• Incitement of violence.</li> <li>• Sale of illegal goods/ services, such as drugs and weapons (on the open internet).</li> <li>• Content illegally uploaded from prisons.</li> <li>• Sexting of indecent images by under 18s (creating, possessing, copying or distributing indecent or sexual images of children and young people under the age of 18).</li> </ul>	<ul style="list-style-type: none"> <li>• Cyberbullying and trolling.</li> <li>• Extremist content and activity.</li> <li>• Coercive behaviour.</li> <li>• Intimidation.</li> <li>• Disinformation.</li> <li>• Violent content.</li> <li>• Advocacy of self-harm.</li> <li>• Promotion of Female Genital Mutilation (FGM).</li> </ul>	<ul style="list-style-type: none"> <li>• Children accessing pornography.</li> <li>• Children accessing inappropriate material (including under 13s using social media and under 18s using dating apps; excessive screen time).</li> </ul>

Figura 3 - Principali danni on-line - fonte: HR Government (2019). Online Harms White Paper.

La debolezza del mercato dei dati affiora anche quando si affronta il tema dei danni nei confronti dei *business users*. Come già discusso precedentemente dal punto di vista del *consumer welfare*, le piattaforme o i fornitori di servizi possono attuare condotte che alterano la competizione o condotte abusive che sono causa di danni nei confronti dei consumatori, volti alla riduzione del loro benessere economico. Gli economisti reputano però difficile quantificare l'entità del danno al quale il consumatore è sottoposto e per questo, anche quando esso non può venir misurato con precisione, le strategie implementate dalle piattaforme, volte a ridurre la pressione

competitiva, dovrebbero essere vietate soprattutto in assenza di chiara documentazione sui guadagni (o perdite) di benessere per il consumatore.

## OPPORTUNITÀ

### *Network externalities*

I servizi nei mercati digitali sono caratterizzati da effetti di rete. È importante sottolineare che però le esternalità di rete esistevano già prima della nascita delle piattaforme digitali, basti pensare alla rete telefonica (Evans e Gawer, 2016). Prendendo come esempio un mercato a due versanti, ovvero un mercato in cui è presente una piattaforma che connette due gruppi di utenti che interagiscono tra di loro minimizzando i costi, si distinguono quattro tipologie di esternalità di rete.

In primo luogo tali effetti possono essere positivi se generano utilità per gli utenti/consumatori di un servizio oppure negativi se, al contrario, sono fonte di disutilità. Allo stesso tempo possono verificarsi:

- *Same-side effects* o effetti diretti, quando valore del servizio per l'utente aumenta con l'aumentare del numero di utenti attivi sulla stessa piattaforma;
- *Cross-side effects* o effetti indiretti, quando il valore per gli utenti incrementa con il numero di utenti appartenenti ad una piattaforma diversa, ad un altro versante.

Con le esternalità di rete la convenienza ad usare una certa tecnologia o a ricorrere ad un certo servizio aumenta con il numero di utenti che usufruiscono della rete, motivo per il quale l'incumbent non deve solo far leva sul prezzo o sulla qualità del servizio ma deve riuscire ad accaparrarsi anche un buon numero di utenti. In accordo con quanto scritto da Delmastro e Nicita (2019), le esternalità di rete rappresentano quindi un'opportunità per le piattaforme che appartengono al *data market* poiché permettono al *service provider* di conquistare gran parte del mercato (fenomeno "*winner takes all*") e inducono i consumatori ad essere attivi solo su una piattaforma (*single-homing*).

Le *network externalities* rappresentano uno dei fattori chiave per lo sviluppo dell'economia digitale e proprio poiché la portata di questi effetti è di grande entità, si ritiene necessario affiancare ai *service provider* alcuni servizi correlati come l'interoperabilità e la portabilità dei dati, i.e. il trasferimento dei dati da una piattaforma ad un'altra, come si vedrà in seguito, e la possibilità di fare *multi-homing*, i.e. la possibilità data ad un utente che desidera fare uno *switch* con un'altra piattaforma, di usufruire di entrambe le piattaforme per capire quale tra esse incontra maggiormente le sue necessità.

### *Barriere all'entrata*

I dati e la loro natura possono agire da barriere all'entrata nel mercato digitale e rappresentano, a seconda dei punti di vista, un'opportunità per le imprese o una minaccia.

Analizzando questo secondo caso, si è potuto stabilire che nonostante un mercato sia ritenuto profittevole, se è dominato dalle economie di scala e di scopo, non è considerato attrattivo per potenziali nuovi entranti poiché gli incumbent possono sfruttare il loro potere di mercato o ricorrere alla sussidiazione con altri business o settori correlati. In particolare, se il mercato è del tipo *single-homing* i potenziali entranti dovrebbero convincere un gran numero di utenti a cambiare piattaforma, al fine di raggiungere la scala e poter competere effettivamente con gli incumbent (Ofcom, 2019).

Concentrandosi però sul primo aspetto, ovvero sul lato opportunità, l'incumbent che possiede una grande quantità di dati in questo modo è in grado di consolidare la propria posizione e mantenere il proprio vantaggio competitivo sfruttando le asimmetrie informative, le quali vanno ad incrementare la qualità del servizio e abilitano il *target advertising* che, a sua volta, attira investimenti; il tutto va a beneficio del servizio che vedrà un aumento del numero di clienti (Fig. 4). Le tradizionali barriere geografiche o i costi di trasporto sono invece irrilevanti (Furman, Coyle, Fletcher, McAuley, Marsden, 2019).

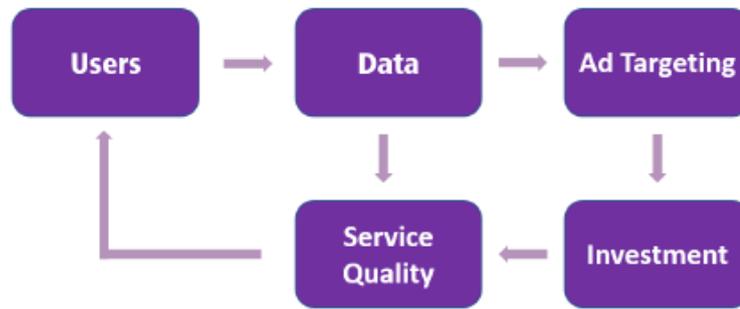


Figura 4 - Feedback loop: Rappresentazione del vantaggio competitivo di un incumbent - fonte: Furman, Coyle, Fletcher, McAuley, Marsden, (2019)

Anche la dimensione degli investimenti richiesti dal *data market* può costituire una barriera all'entrata. È noto che l'analisi e la gestione dei big data richiede notevoli investimenti infrastrutturali che possono costituire una barriera allo sviluppo di nuove imprese, rispetto invece agli incumbent, i.e. le piattaforme dominanti, che possono sfruttare le economie di scala e di apprendimento a loro vantaggio.

#### *Ruolo del dato*

Esplicitando la visione della Commissione Europea, proposta attraverso il report di Crémer, de Montjoye, Schweitzer (2019), i dati che le piattaforme possiedono e a cui esse hanno accesso si dividono in:

- dati volontari, i.e. dati volontariamente dall'utente alla piattaforma;
- dati osservati, ottenuti osservando il comportamento degli utenti, i quali lasciano una traccia nella rete;
- dati dedotti/inferiti, ottenuti trasformando le informazioni “non banali” dei dati forniti volontariamente od osservati a partire dal comportamento degli utenti.

L'opportunità fornita dai big data consiste nel riuscire a sfruttare le informazioni che essi contengono per offrire servizi nuovi e innovativi e prodotti sempre più competitivi. L'analisi e lo *storage* dei big data è il nucleo della *data*

*economy*, tuttavia esso solleva importanti problemi di *privacy*. Al giorno d'oggi, sono tre gli aspetti che si manifestano con maggiore insistenza<sup>4</sup>:

- *Data persistence*. Una volta creato il dato esso può perdurare più dell'uomo che lo ha generato, dati i bassi costi di *storage*.
- *Data repurposing*. È davvero difficile comprendere come un dato potrà venir sfruttato nel futuro, le applicazioni sono indefinite.
- *Data spillovers*. Possono esistere potenziali *spillovers* (negativi) che coinvolgono ignaramente diversi utenti.

Il ruolo del dato rappresenta, quindi, un'opportunità per l'ente che lo possiede che però al tempo stesso deve affrontare tutti quei problemi collegati all'incremento del potere di mercato. Il sistema europeo di *data protection* ha creduto fortemente, almeno fino a questo momento, nell'autoregolazione delle piattaforme e delle imprese e confida sul fatto che l'accesso ai dati di un utente (personali o non) possa avvenire solo se tale utente ne dà il consenso.

Vi sono altri fattori da considerare che permettono di sfruttare completamente le caratteristiche dei dati. In primo luogo potrebbe verificarsi che il possesso dei *big data* renda meno costosa l'esclusione degli utenti da tale possesso, quando proprio i singoli utenti non sono incentivati verso *network externalities*, i.e. alle piattaforme costa meno gestire i dati dell'utente. In aggiunta quando gli *switching cost* tra piattaforme o servizi sono alti il dato è fonte di vantaggio competitivo per la piattaforma o l'impresa che ne detiene il possesso.

## MINACCE

### *L'Antitrust*

L'autorità Antitrust, in Italia rappresentata dall'Autorità Garante della Concorrenza e del Mercato AGCM, è il principale strumento per la tutela dei consumatori e, di conseguenza, una potenziale minaccia per le imprese e piattaforme

---

<sup>4</sup> Tratto da: “*The Economics of Data, Privacy, and (a bit of) Machine Learning*”. Slide del corso di Economia e Management dei Servizi a cura del Professor Lorien Sabatino. Politecnico di Torino. A.A. 2019-2020.

appartenenti al *data market*. L'Antitrust ha il compito di comprendere e analizzare i problemi che sorgono in merito alla protezione dei consumatori e dei loro dati, si focalizza sull'applicazione delle leggi che possono influenzare il funzionamento del mercato dei big data e si fa carico della protezione della proprietà intellettuale.

In un mercato digitale sono le caratteristiche del settore stesso, come la presenza di esternalità di rete o di mercati *multi-sided*, che amplificano gli effetti sul controllo dei dati e rendono quindi necessaria l'attuazione della normativa Antitrust. Nello specifico, è fondamentale concordare in tutto il settore norme accettabili di comportamento concorrenziale per le imprese che possiedono uno status strategico di mercato, su come esse dovrebbero agire nei confronti di piccole imprese e consumatori.

Ciò nonostante, a favore delle imprese gioca il fatto che spesso l'applicazione di queste leggi è un processo che richiede tempo, si muove lentamente ed è strettamente confinato a casi specifici (si vedano i casi Antitrust europei e americani legati a Google e Facebook). L'Antitrust si trova a gestire sfide in mercati in rapida evoluzione e altamente complessi e una soluzione potrebbe essere data dal concordare le norme in anticipo, fornendo chiarezza alle imprese sul mercato circa le regole del gioco.

#### *Le Autorità Regolatorie*

In risposta ai potenziali danni causati dai comportamenti anti-competitivi delle piattaforme e imprese appartenenti al *data market*, alcuni paesi hanno proposto di istituire a tal proposito una nuova figura, un'Autorità Regolatoria (HR Government, 2019). Tale autorità si dovrebbe occupare di verificare i danni derivanti dalle attività online ai quali sono esposti i consumatori e conseguentemente fornire la tutela necessaria a contrastare i problemi di privacy. I governi hanno preso in carico due opzioni:

- Dare vita ad una nuova figura, un nuovo regolatore che si focalizzi sulla sicurezza online e sullo stabilire le linee guida per la sicurezza del settore e dei consumatori, anche se questo potrebbe essere molto costoso per i governi e

potrebbe complicare ulteriormente il panorama regolatorio, già notevolmente complesso;

- Impiegare un regolatore esistente che si assuma la responsabilità di supervisionare sulla sicurezza online degli utenti, sfruttando i vantaggi che derivano da una comprovata esperienza e credibilità sul campo.

Tra i paesi che più si stanno dibattendo su questo tema vi è il Regno Unito che negli anni ha prodotto vari documenti sul tema e ha istituito una Digital Markets Taskforce. La suddetta *taskforce* avrebbe dovuto istituire, da Luglio 2020, una Digital Market Unit. Purtroppo data la recente crisi pandemica Covid-19 attualmente non esiste niente di concreto. L'autorità in considerazione<sup>5</sup> dovrebbe avere il potere di regolare i modelli di business delle grandi piattaforme digitali (i.e. Google, Facebook) e dovrebbe poter assicurare che le riforme che verranno applicate per regolare i giganti della tecnologia siano pro-innovazione e coerenti.

Ma il Regno Unito non è il solo paese a dibattersi per ottenere un mercato digitale più regolato, altri esempi sono dati dall'Australia con l'Australian Competition & Consumer Commission oppure dalla stessa Europa con la Commissione Europea. Quando queste autorità regolatrici verranno coinvolte a pieno ritmo nel panorama regolatorio di ogni paese, costituiranno una minaccia non irrilevante per le attività all'interno dei mercati digitali.

---

<sup>5</sup> [www.corrierecomunicazioni.it/digital-economy/advertising-lantitrust-uk-pensa-a-una-task-force-sui-mercati-digitali](http://www.corrierecomunicazioni.it/digital-economy/advertising-lantitrust-uk-pensa-a-una-task-force-sui-mercati-digitali).

## CAPITOLO II

### RASSEGNA DELLA LETTERATURA

Dopo aver passato in rassegna le componenti e la struttura del *data market*, questo secondo capitolo ha l'obiettivo di presentare, mediante l'analisi della letteratura prodotta fino al giorno d'oggi, le principali dinamiche di tale mercato.

Si analizzeranno in primo luogo i fallimenti di mercato per poi passare al tema della discriminazione di prezzo nel mercato dei dati. Si proporrà un focus sulle due tipologie di abuso di posizione dominante (abuso di esclusione e abuso di sfruttamento) i quali riescono a trovare spazio a causa delle caratteristiche dell'asset in esame, i dati. Conseguentemente, si eseguirà un approfondimento sul tema della tutela della privacy. A tutti questi aspetti, caratterizzati da una connotazione prevalentemente negativa nei confronti del consumatore finale, segue l'approfondimento sulle implicazioni di policy ovvero sulle possibili strategie attuali e future volte alla mitigazione delle problematiche descritte.

La struttura di questo capitolo è riassunta nella Figura 5.

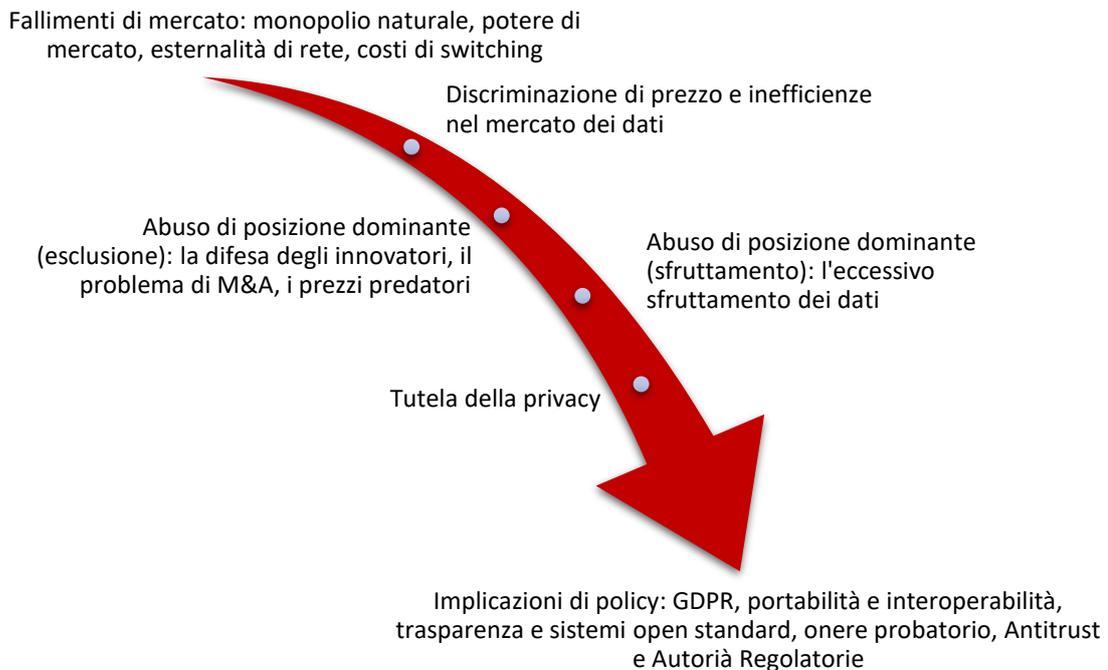


Figura 5 - Struttura della Rassegna della Letteratura

## 2.1. Fallimenti di mercato

Come evidenzia il report dell'OECD (2017), spesso si cerca di mitigare l'esistente competizione nel *data market* con le leggi sulla privacy, con i diritti sulla proprietà intellettuale e facendo appello ai diritti fondamentali dell'uomo. Nella maggior parte dei casi questo non è sufficiente a tutelare il consumatore dagli effetti negativi che scaturiscono dai fallimenti di mercato.

### 2.1.1. Monopolio naturale

Anche nel mercato dei dati si può parlare di possibile monopolio naturale, facendo riferimento alla situazione in cui vi è un monopolista che riesce a produrre a costi più bassi di quelli che si avrebbero con altre forme di mercato, cioè in caso di pluralità di imprese in concorrenza. Nello specifico, le piattaforme digitali potrebbero venire identificate come dei monopoli naturali in cui solo un piccolo numero di imprese riesce ad avere successo, operando con costi di produzione prevalentemente fissi (l'infrastruttura fisica sulla quale si fondano le piattaforme) e costi medi totali decrescenti.

Per comprendere meglio il concetto di monopolio naturale si propone ora un esempio riguardante la differenza tra la fruizione di un servizio creato con open standard e uno creato senza.

In un servizio costruito usando uno standard aperto le specifiche tecniche del sistema sono pubblicamente disponibili e implementabili tramite processi trasparenti. Ad esempio, HTTP (Hypertext Transfer Protocol) è il protocollo a livello di applicazione attraverso il quale si descrive come vengono scambiati i messaggi tra un *browser* (il sistema periferico che si comporta come *client*) e un *web server* (*l'host* che fa da server). Tale protocollo è completamente aperto e chiunque può creare un *web service*, cioè un servizio in cui l'interazione è del tipo *machine-to-machine*, basato su HTTP. Lo scopo di ciò è garantire un certo livello di interoperabilità che però in questo contesto assume il significato di compatibilità del servizio offerto con i servizi proposti da terze parti.

Se invece un servizio non è offerto con open standard, il concetto di interoperabilità si riferisce solo alla possibilità di vedere trasferiti parte dei dati degli

utenti tra diverse applicazioni. Inoltre, non sarebbe possibile utilizzare *client* diversi per operare simultaneamente e per questo gli utenti si troverebbero a impiegare tutti uno stesso servizio, alimentando le economie di rete preesistenti e creando un servizio che assomiglia molto ad un monopolio naturale.

Tuttavia, come illustrato nel report di Furman, Coyle, Fletcher, McAuley, Marsden (2019), nonostante il mercato dei dati e le piattaforme digitali abbiano delle caratteristiche che possono ricondurre al monopolio naturale, è ancora troppo presto per decretare che in questo tipo di mercati non si arriverà mai a competere in modo efficiente. Lo scopo del report, infatti, è quello di esplorare le possibili policy a favore della competizione anche nei mercati digitali.

### **2.1.2. Potere di mercato**

Secondo quanto presentato da Crémer, de Montjoye, Schweitzer (2019) nel report della Commissione Europea “*Competition policy for the digital era*”, in primo luogo è fondamentale stabilire cosa concorre a definire la quota di mercato di un’impresa, i.e. la quota delle vendite rapportata alle vendite totali in un determinato mercato, poiché in base a tale quota un’impresa si dimostra più o meno dominante sul mercato. Nel calcolo del *market share* devono rientrare non solo le vendite di uno specifico prodotto o servizio ma anche i servizi che vengono offerti ai clienti o servizi di default. È bene sottolineare però che anche i servizi che in apparenza vengono offerti hanno come moneta l’appropriazione dei dati degli utenti: anch’essi, dunque, hanno un prezzo per chi ne usufruisce. In alcuni casi, quindi, risulta fondamentale valutare chi possiede i dati e quanti ne possiede poiché ciò è fondamentale per determinare il potere di mercato di un’impresa.

Il termine *data market* racchiude un gran numero di applicazioni e sottosettori (tutti quei settori in cui poter sfruttare i dati degli utenti porta ad un vantaggio competitivo) e, ancor prima di determinare il *market share* dell’impresa, è necessario comprendere a quale mercato o settore si sta facendo riferimento. Un valido strumento a questo proposito è valutare l’elasticità della domanda al prezzo di beni/servizi che per tipologia appartengono a due mercati diversi ma possono essere accomunati (Delmastro e Nicita, 2019). Se l’elasticità al prezzo è bassa, significa che ciò che succede in un mercato o settore non ha influenze sull’altra parte che si sta analizzando

e quindi i mercati rimangono distinti. Viceversa, se l'elasticità della domanda al prezzo è alta, i prezzi sono comparabili e un incremento marginale dei prezzi sposta la domanda, allora si può concludere che i due mercati sono concorrenti. Riuscire a identificare qual è il settore in cui si sta operando è molto importante poiché permette di identificare i concorrenti e permette di comprendere se vi è potere di mercato e di che entità.

Bourreau, De Streel, Graef (2017) propongono un *framework* basato su tre principi chiave per valutare l'importanza dei dati e come essi sono legati al potere di mercato:

1. I dati sono un input, molto importante ma non l'unico, per sviluppare applicazioni e algoritmi di successo.
2. La catena del valore legata ai big data (la collezione dei dati, lo *storage* e l'analisi) acquisisce importanza in concomitanza con gli effetti di rete, diretti e indiretti, ed è quindi fondamentale non trattare il *data market* come un mercato ad un versante ma come un *two-sided market*.
3. Ogni applicazione dei big data e ogni algoritmo è diverso e si deve analizzare situazione per situazione. Non si possono proporre delle linee guida generali poiché ogni caso è differente.

Possono quindi essere presenti situazioni in cui un'impresa riesce a sfruttare pienamente le economie di scala, di scopo e di rete e riesce ad appropriarsi dei dati ad un costo sostenibile mediante gli utenti, macchine o *data brokers*. In questo caso, si può parlare di potere di mercato e la relazione con il *data market* è evidente. Allo stesso tempo però, è necessario considerare che non tutte le imprese agiscono in questo modo e godono di queste caratteristiche.

Avendo in mente questi tre principi, le autorità che regolano la competizione e i fallimenti di mercato possono cercare di valutare il *market power* di un'impresa.

### **2.1.3. Esternalità di rete**

Le esternalità di rete, in particolare quelle indirette, rappresentano un fallimento di mercato nella misura in cui, come presentato da Bourreau, De Streel,

Graef (2017), la presenza di tali effetti genera competizione per il mercato e non nel mercato. Avere tanti dati concentrati tutti all'interno di un unico versante è sicuramente fonte di un aumento di *market share* e di potere di mercato per l'impresa che li detiene. Inoltre, ciò costituisce una barriera all'entrata per i *new entrants* nel mercato del digitale grazie anche agli effetti combinati con economie di scala e di scopo, i quali sono volti a ridurre la pressione competitiva e mantenere pochi e forti *players* sul mercato. Come presentato nel capitolo precedente, le *network externalities* costituiscono un'opportunità di crescita per le imprese ma se gli effetti sono decisamente forti, il mercato si orienta verso un competitor che diventerà quindi il monopolista (Scott Morton, Bouvier, Ezrachi, Jullien, Katz, Kimmelman, Morgenstern, 2019).

Crémer, de Montjoye, Schweitzer (2019) sostengono che uno dei modi per ridurre il potere di mercato delle imprese che operano in mercati *multi-sided* sia fare *bundling* tra prodotti per i quali è forte il controllo sui dati del consumatore e tra prodotti per i quali tali fattori giocano un ruolo minore (e.g. “Zero-Price problem”).

Allo stesso modo, si può dimostrare che è possibile contrastare le situazioni a fallimento di mercato causate dagli effetti di rete nei mercati *multi-sided* avendo solo una piattaforma che serve i diversi mercati e che internalizza gli effetti di rete<sup>6</sup>. Nello specifico, si considera un mercato a due versanti che compete sul prezzo e si pone in contrapposizione il caso in cui ciascuno dei due versanti sia servito da una piattaforma (a) rispetto al caso in cui una piattaforma serva entrambi i versanti (b).

- $D(p_1) = 1 - p_1$                       domanda nel versante 1
- $D(p_2) = 1 - p_2$                       domanda nel versante 2
- $q_1 = 1 + e_{21}D(p_2) - p_1$         quantità nel versante 1 con  $e_{21}$  che cattura l'effetto di rete dal versante 2 verso 1

---

<sup>6</sup> Tratto da: “*The Economics of Platforms*”. Slide del corso di Economia e Management dei Servizi a cura del Professor Lorien Sabatino. Politecnico di Torino. A.A. 2019-2020.

- $q_2 = 1 + e_{12}D(p_1) - p_2$       quantità nel versante 2 con  $e_{12}$  che cattura l'effetto di rete dal versante 1 verso 2

È importante sottolineare che se la domanda aumenta in uno dei due versanti, grazie ad una diminuzione del prezzo, si ha un effetto positivo anche nel versante opposto.

*Caso (a): piattaforme indipendenti*

L'obiettivo è massimizzare il profitto individuale in ciascun versante

$$\text{Max}_{p_1} \quad \pi = (1 + e_{21}D(p_2) - p_1) p_1$$

$$\text{Max}_{p_2} \quad \pi = (1 + e_{12}D(p_1) - p_2) p_2$$

da cui si ottiene:

$$p_1^{ind} = \frac{2 + e_{21}(1 - e_{12})}{4 - e_{12}e_{21}}$$

$$p_2^{ind} = \frac{2 + e_{12}(1 - e_{21})}{4 - e_{12}e_{21}}$$

$$\pi_1^{ind} = (p_1^{ind})^2$$

$$\pi_2^{ind} = (p_2^{ind})^2 .$$

*Caso (b): unica piattaforma*

L'obiettivo è massimizzare il profitto congiunto dei due versanti

$$\text{Max}_{p_1 p_2} \quad \pi = p_1 q_1 + p_2 q_2$$

da cui si ottiene:

$$\begin{cases} p_1(p_2) = \frac{1 + e_{21} - p_2 (e_{21} + e_{12})}{2} \\ p_2(p_1) = \frac{1 + e_{12} - p_1 (e_{12} + e_{21})}{2} \end{cases}$$

$$p_1^* = \frac{1 - e_{12}}{2 - (e_{12} + e_{21})}$$

$$p_2^* = \frac{1 - e_{21}}{2 - (e_{12} + e_{21})}$$

$$\pi^* = \frac{1 - e_{12}e_{21}}{2 - (e_{12} + e_{21})}.$$

Analizzando i prezzi e i profitti ottenuti si evince che, sia nel caso in cui gli effetti delle esternalità sono simmetrici  $e_{12} = e_{21}$  sia nel caso in cui vi è alta esternalità da un versante verso l'altro  $e_{12} \neq e_{21}$ ,  $p_1^* + p_2^* = 1$  mentre  $p_1^{ind} + p_2^{ind} = 1 + \frac{e_{12} + e_{21}(1 - e_{12})}{4 - e_{12}e_{21}} > 1$ . Nel caso (b) quindi la singola piattaforma produce prezzi più bassi, profitti più alti e maggiore *consumer surplus*. Ciò che avviene praticamente è che la singola piattaforma è in grado di internalizzare gli effetti di rete.

Attraverso fenomeni quali la sussidiazione incrociata, un'impresa può offrire in uno dei due versanti un bene/servizio a “Zero Price” innescando un fenomeno di incremento della domanda nel versante complementare. Se si è interessati ad interagire con entrambi i versanti allora è opportuno ricorrere ad un'unica piattaforma poiché, nel complesso, i prezzi sono più bassi. Come illustrato nella Figura 6, la sussidiazione causa un incremento di profitti che copre le perdite del versante in cui si abbassa (o si azzerà) il prezzo del bene e ciò non comporta problemi finché tale procedura non è volta a comportamenti escludenti o abuso di posizione dominante.

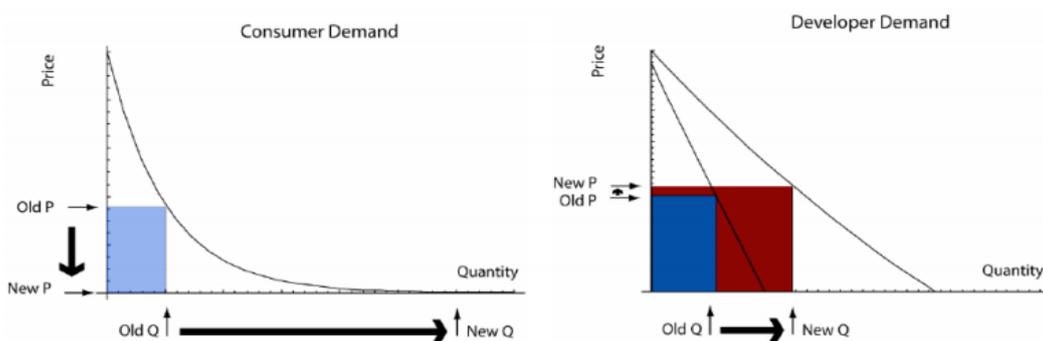


Figura 6 – Processo di sussidiazione – fonte: Slide del corso di Economia e Management dei Servizi a cura del Professor Lorien Sabatino. Politecnico di Torino. A.A. 2019-2020.

Nonostante la situazione di equilibrio ottenuta si possa definire Pareto-efficiente, spesso accade che, in mercati con esternalità di rete che possono essere internalizzate dagli operatori, la competizione sia molto agguerrita nella fase iniziale poiché si cerca di raggiungere la massa critica: l'innescarsi di questi effetti positivi può portare ad una posizione monopolistica delle imprese. Bisogna, quindi, prestare attenzione non solo alle esternalità positive ma anche alla velocità con cui un'impresa si espande all'interno del mercato.

A corredo dell'analisi appena effettuata si presenta ora una applicazione del concetto di esternalità di rete in relazione al valore della privacy per gli utenti e agli incentivi di acquisizione dei dati.

*Acemoglu, Makhdoumi, Malekian, Ozdaglar (2019)*

La letteratura generalmente sostiene che la disponibilità degli utenti nel permettere che i loro dati vengano utilizzati senza ricevere in cambio alcun beneficio (o un beneficio molto basso) sia sinonimo del fatto che gli utenti attribuiscono un basso valore alla loro privacy (i.e. per gli utenti i dati hanno un prezzo molto basso).

Al contrario, gli autori del modello in analisi ritengono che siano presenti forze maggiori che rendono i dati a livello individuale svalutati, sottoprezzati. Il mercato genera troppi dati e ciò è causa di esternalità negative. In alcuni casi sono gli stessi utenti che permettono la fuoriuscita dei dati di altri utenti. La svalutazione è quindi una conseguenza delle informazioni che trapelano: se i dati di un individuo sono già stati condivisi da altri utenti, c'è meno incentivo a proteggere la privacy e i dati.

L'analisi prevede la suddivisione degli utenti in due gruppi:

- utenti *high-value*, coloro che conferiscono molto valore alla loro privacy  $v$  e alla diffusione dei dati personali. Essi possiedono  $v > 1$ ;
- utenti *low-value*, coloro che non attribuiscono molto valore alla loro privacy possiedono  $v \leq 1$ .

Con il diminuire del valore che gli utenti attribuiscono all'acquisizione dei dati diminuisce il surplus (Fig. 7).

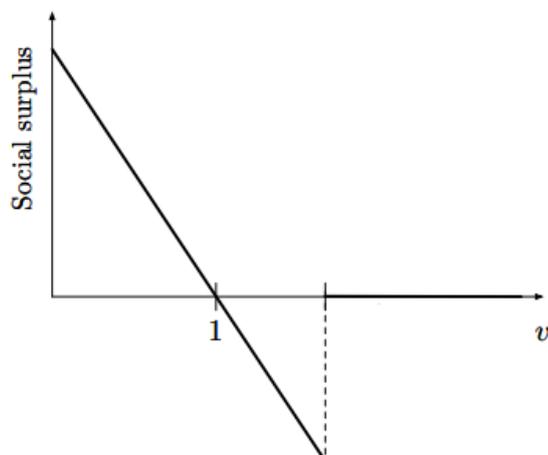


Figura 7 - Utenti High-Value e utenti Low-Value - fonte: Acemoglu, Makhdoumi, Malekian, Ozdaglar (2019)

Il modello analizza due situazioni differenti e due diverse casistiche di competizione: una piattaforma in monopolio e competizione tra diverse piattaforme.

*Caso in cui le piattaforme conoscono il valore della privacy dell'utente*

- Piattaforma in monopolio

La soluzione di un'unica piattaforma in monopolio è anche detta soluzione di *first-best* poiché l'intento degli utenti è di condividere i dati e di massimizzare il benessere sociale, anche se gli individui che condividono i loro dati causano esternalità negative sugli utenti che non condividono i loro dati.

Lo sviluppo del modello prevede in primo luogo di stabilire il punto di equilibrio e il prezzo dei dati poiché da ciò si comprende quanto il prezzo e la distribuzione del surplus dipendano dalle informazioni che trapelano dagli altri utenti. Successivamente si stabiliscono le condizioni che aumentano il benessere e le condizioni per cui il mercato è inefficiente. Se gli utenti *high-value* non sono correlati con gli utenti *low-value*, ci si trova in condizioni di equilibrio efficiente poiché l'assenza di correlazione è ciò che evita che tra i due gruppi di utenti possa esservi rilascio di informazioni non desiderate. Ciò nonostante, possono manifestarsi esternalità sui dati tra gli utenti *low-value* che influenzano i prezzi di mercato ma non creano carenze di privacy per coloro che non desiderano condividere i propri dati.

Al contrario, se le due classi di utenti sono correlate allora l'equilibrio è inefficiente. Le motivazioni sono molteplici: gli utenti *high-value* non condividono i loro dati ma la loro fuoriuscita dovuta alla correlazione con i *low-value* comunque causa mancanza di privacy; vista la quantità di dati che comunque sono già stati condivisi gli utenti *high-value* decidono di condividere i loro dati nonostante la riluttanza iniziale; se è presente correlazione all'interno dello stesso gruppo di utenti la piattaforma può invogliare qualcuno a dare i dati e qualcuno no.

Dopo aver identificato le (in)efficienze, gli autori propongono una nuova proposta di regolazione secondo cui i dati vengono scambiati in modo da ridurre le correlazioni tra utenti e minimizzando le informazioni che trapelano. Lo scopo, dunque, è rimuovere la correlazione tra utenti.

- Concorrenza tra multiple piattaforme

Per semplicità si assumono due piattaforme le quali cercano simultaneamente di attrarre utenti e acquisire i loro dati. Il modello è sviluppato a tre stadi:

1. Utenti decidono quale piattaforma usare;
2. Piattaforme offrono simultaneamente un prezzo per i dati;
3. L'utente decide se condividere i dati.

Anche in questo secondo caso si verificano delle inefficienze e proprio per questo il mercato risponde incrementando il prezzo per i dati o indirizzando gli utenti *high-value* verso piattaforme in cui vi è meno possibilità che trapelino le informazioni. Vengono inoltre proposte alcune estensioni come l'inversione della dinamica tra stadio 1 e stadio 2 in cui, ora, le piattaforme settano i prezzi per prime per attrarre gli utenti oppure come la riduzione della presenza delle *network externalities* solo tra utenti della stessa piattaforma (esternalità dirette).

#### *Caso in cui le piattaforme NON conoscono il valore della privacy dell'utente*

Il caso in cui il valore della privacy dell'utente, ovvero l'informazione privata, è conosciuto dagli utenti ma sconosciuto alle piattaforme rappresenta la soluzione di *second-best*. Il modello si focalizza sull'analisi di un'impresa monopolista il cui intento è di progettare un meccanismo per venire a conoscenza di tale valore. Il procedimento, conforme a quanto visto nell'analoga situazione di monopolio ma nel

caso precedente, dimostra che la soluzione di *second-best* coincide con la soluzione di *first-best*. Anche in questo caso, dunque, si verifica inefficienza e la condizione per ottenere efficienza richiede che gli utenti *low-value* abbiano  $v \leq 1$ .

### *Conclusioni*

Come si è visto, tutte le soluzioni nelle diverse casistiche trattate conducono a forme di equilibrio inefficiente nel determinare un prezzo per i dati. Una possibile soluzione potrebbe essere l'adozione di meccanismi di regolazione.

Il primo meccanismo riguarda la possibilità di introdurre un sistema di regolazione basato sulle tasse. Si propone il modello di tassazione Pigoviano dove il governo impone una tassa  $t_i$  sull'utente  $i$  che vende i propri dati alla piattaforma. Si dimostra, infatti, che non sottoporre al meccanismo di tassazione un utente che dovrebbe condividere i propri dati è sufficiente a garantire che esso poi li condivida, indipendentemente da ciò che fanno gli altri utenti. In aggiunta, la soluzione di *first-best* si può raggiungere se si impone una tassa sulle transazioni dei dati degli utenti che non dovrebbero condividere i propri dati.

Un secondo approccio consiste nella regolazione basata sulla "de-correlazione" delle informazioni. Si è visto che la correlazione tra gruppi di utenti riguarda il rilascio di informazioni non desiderate e quindi l'implementazione di esternalità di rete negative. Tale soluzione propone inizialmente la condivisione dei dati con un mediatore, piuttosto che con una piattaforma, il quale può decidere di non consegnare i dati alla piattaforma o di trasformarli prima della rivelazione. La chiave è proprio la trasformazione del dato: in questa fase, il mediatore rimuove ogni possibile correlazione tra utenti che non siano quelli da cui il dato è stato generato. A conclusione delle transazioni sui dati si rileva un aumento del surplus sociale.

#### **2.1.4. Costi di switching**

I costi che gli utenti si trovano ad affrontare quando vogliono fare *switch* tra diversi servizi possono rappresentare una forte barriera all'entrata con conseguente riduzione di competitività all'interno del mercato (riduzione dei *new entrants*) e incremento del potere di mercato.

I costi di switching che un utente sostiene nella maggior parte dei casi non sono dei costi sostenuti direttamente ma derivano indirettamente da pratiche o attività volte a fare *customer lock-in*. Secondo la visione dell'Ofcom (2019) esistono una serie di fattori che possono far sorgere delle difficoltà agli utenti quando si desidera cambiare piattaforma:

- Perdita dei dati personali. Spostandosi da una piattaforma ad un'altra spesso non è previsto che un consumatore possa portare con sé i propri dati personali. Riferendosi ad un servizio di comunicazione (e.g. WhatsApp) l'utente potrebbe perdere lo storico della chat.
- Costi di apprendimento. L'utilizzo di una certa piattaforma per l'utente ha richiesto un certo grado di apprendimento, in termini di sforzo e tempo. Se l'utente volesse cambiare piattaforma, e tale piattaforma fosse molto diversa dalla prima, si troverebbe nuovamente a sostenere dei costi di apprendimento dovuti alla non-familiarità con la nuova interfaccia (e.g. Android e iOS).
- Barriere tecniche e *tying*. Usufruire di un servizio può comportare l'acquisto di ulteriori servizi/beni correlati (e.g. acquisto di un servizio insieme ad un dispositivo hardware). Questa pratica commerciale, denominata *tying*, rappresenta un ostacolo allo *switch* tra piattaforme diverse.

Anche Calvano e Polo (2020) sostengono che la presenza di *switching cost* induce gli incumbent ad operare in maniera più aggressiva sul *data market*. Tuttavia, essi dividono il mercato in due fasi: nella fase *early stage* tali costi sono alti dovuti allo sviluppo di nuovi software e alla diffusione delle piattaforme; nella fase di *majority*, fase in cui i mercati digitali si trovano oggi, si manifestano bassi costi di switching poiché sottoscrivere un certo servizio non richiede grandi investimenti fisici, di tempo o di conoscenza. Ad esclusione del trattamento dei dati personali, le frizioni rappresentate dagli *switching cost* sono relativamente basse.

Gli studi condotti da Tucker (2019) portano a prendere una simile posizione riguardo gli *switching cost*. Tucker sostiene che quando si fa *switch* tra servizi o piattaforme (asset non tangibili), si va ad indebolire le esternalità di rete e il potere di mercato e non si crea un danno al consumatore. Gli effetti di rete, infatti, possono

risultare molto confinati ad una situazione locale (i.e. al consumatore non importa quanti utenti usufruiscono di una certa piattaforma ma solo se i suoi diretti collegamenti ne fanno uso) e gli utenti preferiscono contesti digitali in cui sono presenti meno persone per questioni legate alla privacy. Esistono, quindi, una serie di motivi tali per cui si può sostenere che il processo di digitalizzazione indebolisce gli *switching cost* creando situazioni favorevoli per gli utenti.

Ciò che però è certo è che alti costi di switching vengono impiegati per fare *lock-in* nei confronti degli utenti e riducono le possibilità dei competitors di accedere e impossessarsi di una risorsa molto preziosa per la crescita nel *data market*.

## **2.2. Discriminazione di prezzo e inefficienze nel mercato dei dati**

Per poter sfruttare al meglio tutte le potenzialità dei dati è importante distinguere per quale applicazione essi vengono impiegati: discriminazione di prezzo, advertising o design di nuovi prodotti e servizi.

Focalizzandosi in particolar modo sulla discriminazione di prezzo, essa trova spazio in ambienti in cui le imprese hanno potere di mercato (e.g. *data market*). Riprendendo la classificazione proposta da Pigou (1920), la discriminazione tariffaria si può distinguere in tre tipologie differenti:

- Discriminazione di primo grado. Anche chiamata discriminazione “perfetta” riguarda il caso in cui è possibile, per ciascun consumatore, fissare un prezzo separato per beni o servizi a seconda della singola disponibilità a pagare del consumatore.
- Discriminazione di secondo grado. In questo caso non si conosce il gruppo a cui appartengono i consumatori e quindi si propongono tariffe diverse per la stessa tipologia di bene. Sarà quindi il consumatore a scegliere quanto consumare e a collocarsi nel gruppo degli altospendenti o bassospendenti. I prezzi differenziati vengono applicati considerando le differenze tra gruppi ma non all’interno dei gruppi.

- Discriminazione di terzo grado. Le imprese possiedono le informazioni sufficienti per poter dividere i consumatori in base alla loro disponibilità a pagare. Questo tipo di discriminazione si considera efficiente poiché permette alle imprese di fornire beni/servizi a prezzi più bassi per gli utenti che manifestano una bassa *willingness to pay*.

In linea teorica la discriminazione di prezzo di primo grado comporta l'appropriazione da parte di un'impresa di tutto il surplus del consumatore e raggiunge l'efficienza tipica dei mercati perfettamente concorrenziali. Come sottolineato da Delmastro e Nicita (2019) è purtroppo impossibile osservare delle perfette discriminazioni di prezzo. Le imprese, infatti, non dispongono di un alto livello di dettaglio delle informazioni e spesso applicano un prezzo medio uguale per tutti creando situazioni di enorme vantaggio per i consumatori con una alta disponibilità a pagare. Potrebbero, inoltre, verificarsi situazioni in cui i consumatori siano incentivati ad acquistare beni a prezzi più bassi e a rivenderli a prezzi più alti, trasformandosi in diretti concorrenti delle imprese. Questo aspetto è stato affrontato e risolto con l'avvento delle piattaforme digitali, le quali, assorbendo informazioni private, sono in grado di proporre offerte differenziate basate sull'elaborazione dei big data da parte degli algoritmi di analisi e disincentivano quindi il processo di rivendita.

Spesso ci si è chiesti se la *price discrimination* comportasse effetti positivi o negativi a livello di *welfare*. Bourreau, De Strael, Graef (2017) sostengono che i prezzi personalizzati hanno effetti ambigui. Se dal punto di vista economico non esistono valide ragioni per bandire i prezzi personalizzati poiché non è sempre vero che si va ad erodere il *welfare* o il surplus del consumatore, è anche vero che un incumbent potrebbe sfruttare la quantità di big data che possiede per offrire prezzi scontati e customizzati al fine di raggiungere una posizione di monopolio. Nello specifico, Bergemann, Brooks, Morris (2015) mostrano che passando da un sistema a prezzi a uniformi ad un sistema che implica discriminazione tariffaria il benessere sociale e il surplus del consumatore non seguono un percorso predefinito: entrambi possono aumentare, diminuire oppure il benessere sociale può aumentare mentre il surplus del consumatore diminuisce. Al contrario, tramite un modello in cui si è reso effettivo per le imprese l'accesso alle informazioni derivanti dalla cronologia di ricerca dei

consumatori, Armstrong e Zhou (2010) hanno evidenziato che le imprese offrono prezzi più bassi a chi visita il sito per la prima volta e prezzi più alti dalla seconda visita in poi al fine di alzare complessivamente i prezzi e, allo stesso tempo, di catturare il consumatore; ciò implica una diminuzione del *consumer surplus* e del benessere totale.

Tutto questo si aggrava se poi l'impresa ha accesso esclusivo ai dati dei consumatori. I prezzi personalizzati possono quindi comportare

- una diminuzione della competizione poichè le imprese possono costruire prezzi così sofisticati da rendere quasi impossibile la comparazione con conseguente incremento dei *search costs* (Ellison e Ellison, 2009);
- un abbassamento della qualità dei servizi o beni offerti poiché le imprese, per invogliare i consumatori con una alta *willingness to pay* a dirigersi verso l'offerta di alta qualità, decidono strategicamente di degradare la qualità dei prodotti venduti ai bassospendenti (Deneckere e Preston McAfee, 1996);
- perdita di fiducia nei mercati online poiché i consumatori che vedono negativamente la discriminazione tariffaria ritengono insicuri gli acquisti online. Da uno studio di Turow, Feldman, Meltzer (2005) emerge, infatti, che i due terzi degli utenti di Internet ritengono illegale l'offerta di prezzi personalizzati da parte dei *retailers*;
- strategie di *customer lock-in*.

Per mitigare l'ambiguità di questi effetti una possibile soluzione potrebbe essere quella di garantire la trasparenza del meccanismo di formazione dei prezzi e l'istituzione di norme efficaci riguardo la protezione dei consumatori. Il possesso dei dati, anche in questo caso, è la chiave che facilita il processo di discriminazione: maggiori sono le informazioni che le imprese dispongono, più accurate sono le stime sulla disponibilità a pagare dei consumatori.

La posizione del consumatore di fronte alla discriminazione tariffaria necessita, quindi, di tutela poiché tali meccanismi comportano sì efficienza allocativa e

redistribuzione delle risorse nei singoli mercati ma, allo stesso tempo, potrebbero portare anche una diminuzione del benessere complessivo della società.

### **2.2.1. Modelli di pricing nel mercato dei dati**

#### *Montes, Sand-Zantman, Valletti (2019)*

All'interno dello studio di Montes, Sand-Zantman, Valletti (2019) sono stati investigati gli effetti della *price discrimination* sul *data market*. Il modello proposto analizza come cambiano i prezzi, i profitti e il surplus del consumatore quando le diverse imprese sul mercato hanno a disposizione le informazioni personali dei consumatori e quando invece i consumatori possono prevenire l'acquisizione delle informazioni personali da parte delle imprese.

Il modello si basa sul fatto che le imprese, in primo luogo, acquisiscono tramite un intermediario le informazioni sulle preferenze dei consumatori e, successivamente, offrono i loro prodotti o servizi ad un set di prezzi personalizzati. Gli attori che entrano in gioco sono quindi di tre tipi: i consumatori, le imprese in competizione (sono due) e il *data supplier* che colleziona informazioni sugli utenti. Nello specifico, si prendono in considerazione due gruppi di consumatori: i nuovi consumatori dai quali non possono venire estratte informazioni e i vecchi consumatori dei quali si dispongono informazioni poiché lasciano traccia della loro attività.

#### *Caso benchmark*

I consumatori non possono pagare per la loro privacy. Ciò significa che il costo della privacy  $c$ , dato dal costo delle azioni che i consumatori intraprendono per evitare che i loro dati vengano presi da imprese o terze parti oppure dato dai pagamenti che essi eseguono per impedire ciò, è molto alto. Inoltre, tale costo  $c$  aumenta se le imprese sono autorizzate a commerciare i dati poiché il costo di cancellare le tracce è più oneroso.

#### *Caso in cui il consumatore può pagare per la sua privacy*

I consumatori possono permettersi di sostenere il costo  $c$  ed evitare quindi di apparire nei database dei *data brokers*. Inoltre essi sono in grado di osservare quali

imprese acquistano le loro informazioni e possono farsi delle aspettative in merito ai prezzi.

Entrambi i casi appena presentati sono analizzati in situazioni di monopolio e di competizione in duopolio quando entrambe le imprese comprano le informazioni sui consumatori dagli intermediari e quando, invece, è solo una tra le due imprese ad acquisire tali informazioni.

#### *Estensioni del modello*

- Cambiamento di timing del modello. Mentre nella configurazione iniziale il consumatore decideva se era disposto a pagare  $c$  per la sua privacy e poi l'impresa settava i prezzi, ora è l'impresa che setta i prezzi, in primo luogo, e poi il consumatore paga per la privacy.
- Spacchettamento dei dati. Gli intermediari vendono solo una parte di dati ad una delle due imprese e la restante parte va all'altra impresa oppure decidono di vendere l'intero pacchetto di dati ad un'impresa, la quale rivende poi il pacchetto all'altra impresa riservandosi il diritto di scegliere quali dati cedere e il livello di precisione.
- Diverse dimensioni del mercato tra vecchi e nuovi consumatori. La massa di consumatori per ciascun mercato è diversa; quando i consumatori nuovi sono più piccoli dei vecchi si ha una migliore competizione.

#### *Risultati*

Ciò che si evince dall'analisi di questo modello è che in situazioni di monopolio è più facile per i consumatori proteggere i propri dati. I risultati in termini di *consumer surplus* e profitti delle imprese sono però abbastanza ambigui poiché alcuni consumatori vincono e altri, invece, perdono. In condizioni di duopolio l'*outcome* è più chiaro: il surplus del consumatore aumenta con l'aumento del costo della privacy  $c$  mentre il benessere totale è massimizzato quando i consumatori non riescono a proteggere i loro dati. È quindi presente una forte competizione tra imprese per impossessarsi delle informazioni.

Nonostante i profitti siano maggiori nel caso del monopolio (Fig. 8) il prezzo per i dati, per ogni livello di  $c$ , è maggiore in duopolio (Fig. 9). Questo significa che per l'intermediario dei dati l'ottimo risiede nel creare situazioni di *winner-takes-all*, anche se ciò non esclude la vendita dei dati in pacchetti.

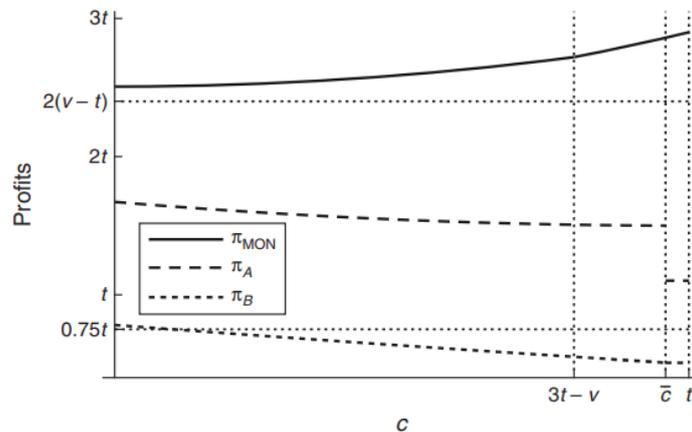


Figura 8 - Profitti delle imprese - fonte: Montes, Sand-Zantman, Valletti, T. (2019)

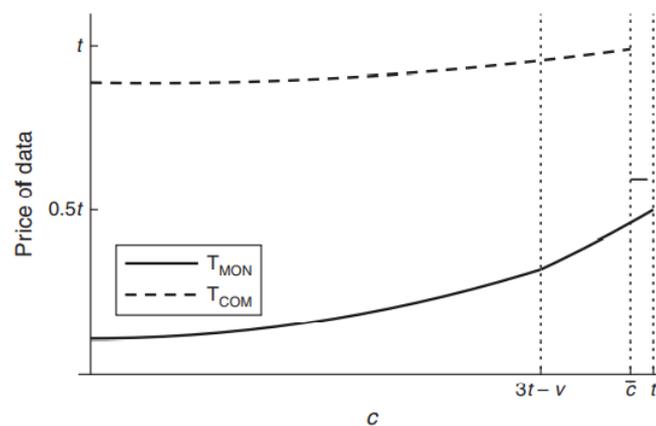


Figura 9 - Prezzo per i dati - fonte: Montes, Sand-Zantman, Valletti, T. (2019)

Lo strumento di lavoro che questo modello fornisce ai regolatori è volto a promuovere una allocazione non esclusiva dei dati tra le imprese che competono poiché questo, come visto, è causa di discriminazioni tariffarie. Il focus comunque dovrebbe rimanere su come avvengono le transazioni delle informazioni e sul rendere la privacy del consumatore più o meno costosa a seconda delle necessità.

### ***Bergemann e Bonatti (2015)***

Lo studio di Bergemann e Bonatti (2015) è orientato alla definizione di policy di *pricing* per l'acquisizione dei dati dagli utenti. In particolare, si fa riferimento alla possibilità di imprese quali *data sellers* di istituire politiche di prezzo per gli *advertisers* sulla base del target di utenti che essi vogliono raggiungere, determinato grazie ai cookies. Infatti, Gli *advertisers*, grazie alle informazioni acquisite, possono proporre pubblicità mirate ai diversi segmenti di consumatori. In generale, il modello di Bergemann e Bonatti viene applicato a qualsiasi ambiente impresa-consumatore in cui l'impresa possiede le informazioni sugli utenti e può offrire prezzi differenziati.

Si analizza ora tale modello con un maggiore livello di dettaglio poiché esso, come si vedrà nel Capitolo III, sarà oggetto di ulteriori approfondimenti.

La Figura 10 rappresenta la struttura del mercato, nella quale si evidenziano i rapporti tra i diversi attori che prendono parte al modello di *data provision* e *data pricing*. Il prezzo per i dati (i.e. le informazioni degli utenti, i cookies) pagato dall'*advertiser* riflette il trade-off tra *disclosure* e *privacy* che si trova a fronteggiare il consumatore nel momento dell'acquisizione di tali dati.



Figura 10 - La struttura del mercato nel modello di Bergemann e Bonatti (2015)

Le principali caratteristiche dell'*online market* in esame sono le seguenti:

- Gli *advertisers* possono segmentare i consumatori in gruppi diversi i quali ricevono pubblicità personalizzate. Potrebbe eventualmente esistere un gruppo residuale di consumatori che riceve un livello di pubblicità pari a zero.

- Il *data provider* imposta un prezzo uniforme per i dati degli utenti. In questo modo gli *advertisers* pagano un prezzo proporzionale alla quantità di cookies acquisita e questa strategia, la quale impedisce alle economie di scala di manifestarsi, prende il nome di “*per-unit-bit-pricing*”.

Il modello ideato dagli autori prevede che il surplus dello scambio sia dato da una funzione che assegna un valore ad ogni abbinamento realizzato tra consumatori e imprese, i.e. la funzione di *match value*. In aggiunta, è possibile evidenziare una corrispondenza con il modello presentato da Salop (1979). Bergemann e Bonatti, infatti, prendono in considerazione consumatori  $i$  e imprese eterogenee  $j$  distribuiti attorno ad un cerchio di lunghezza unitaria in cui i *match values* sono una funzione della distanza  $|i - j|$  (Fig. 11). In modo conforme a quanto presentato da Salop nel modello di differenziazione orizzontale, le informazioni dei consumatori differiscono solo attorno la loro dimensione orizzontale e ciascun *advertiser* sceglie di rivolgersi ai consumatori “più vicini” in base alle proprie preferenze.

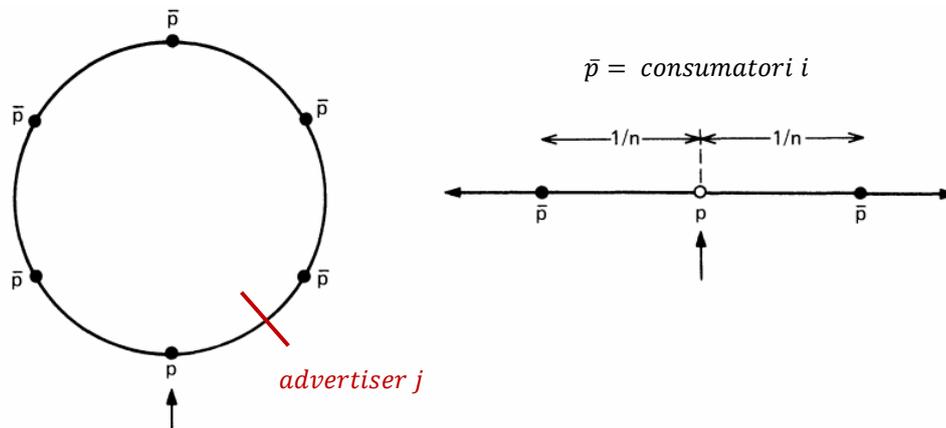


Figura 11 - Modello di Salop - fonte: Salop (1979)

Il timing del modello è rappresentato nella Figura 12. Per una migliore comprensione, si specifica che il termine  $A_j$  identifica il *targeted set*, ovvero il gruppo di consumatori a cui l'*advertiser* si rivolge ed è in grado di offrire un prodotto personalizzato. La dimensione di tale gruppo si determina considerando la disponibilità a pagare dell'impresa per un “cookie marginale” per ogni strategia di targeting. L'*advertiser*, infatti, può decidere di optare per due strategie diverse:

- *Positive targeting*, quando decide di comprare le informazioni dei consumatori con alto valore, indirizzando le proprie spese verso di essi ed escludendo qualsiasi altro gruppo di consumatori.
- *Negative targeting*, quando decide di comprare le informazioni dei consumatori con basso valore al fine di escluderli dall'essere contattati (*residual set*  $A^C$ ). L'*advertiser* indirizza quindi le proprie spese verso qualsiasi altro gruppo di consumatori di cui non possiede le informazioni.

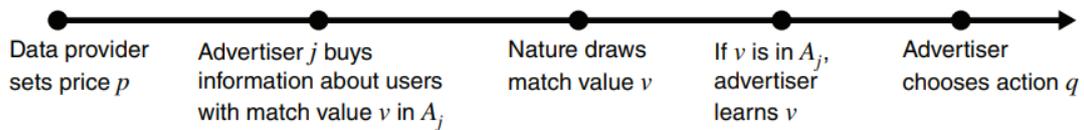


Figura 12 - Timing - fonte: Bergemann e Bonatti (2015)

L'azione  $q$  a cui si fa riferimento riguarda l'intensità di match, i.e. la probabilità che ha un'impresa di generare consapevolezza nel consumatore ed è legata alla scelta del targeting. Optare per una delle due strategie sopra citate inoltre consente di alleviare gli effetti negativi dell'advertising: sottoporre ad un livello di pubblicità eccessivo gli utenti a basso valore oppure riscontrare un livello di pubblicità insufficiente per i consumatori del *targeted set* che possiedono maggior valore (si possiedono più informazioni).

La funzione obiettivo del modello è data dal profitto dell'*advertiser* in caso di informazione completa:

$$\pi(v, q) \triangleq vq - c * m(q) \quad (1)$$

dove  $q$ , come visto poc'anzi, rappresenta l'azione intrapresa dall'*advertiser* per generare consapevolezza nel consumatore (i.e. contatta i clienti),  $v$  rappresenta il *match value*,  $m(q)$  è la *matching cost function* che rappresenta lo spazio pubblicitario e  $c$  il costo unitario di acquisto per tale spazio. L'impresa, infatti, intende customizzare l'azione  $q$  per ogni consumatore con caratteristica  $i$  che possiede un certo valore  $v$ .

Dal punto di vista del *data provider* il *match value* assume due significati: esso rappresenta la richiesta di cookies oppure il cookie stesso. Il *data provider*, inoltre, vende le informazioni dei consumatori al prezzo unitario  $p$  per ciascun cookie di cui dispone. Si nota che  $p$  non è riferito al singolo individuo poiché a ciascun utente possono appartenere più cookies.

In aggiunta, il *match value* può assumere diversi valori appartenenti all'insieme  $V$ . Per ogni sottoinsieme misurabile di valori di  $A$  (*targeted set*) in  $V$ , rappresentato da  $A^v \triangleq [\underline{v}, v]$  con l'inizio dell'intervallo indicato dal valore di *match value* più basso, la misura risultante corrisponde a  $\mu(A)$  e la funzione di distribuzione associata è definita da  $F(v) \triangleq \mu(A^v)$ . Il prezzo del *targeted set* invece è dato da  $p(A) \triangleq p * \mu(A)$ , i.e. dal prezzo unitario per i cookies per il valore del gruppo di consumatori.

#### *Gli advertisers e la domanda di informazioni*

Le imprese determinano il valore per le informazioni dall'incremento di profitto che si ottiene acquistando un numero maggiore di cookies. Per questo motivo è fondamentale comprendere come gli *advertisers* costruiscono il *targeted set* e qual è la loro domanda di cookies, in caso di piena informazione e in caso di disponibilità solo di informazioni preliminari. Il problema dell'impresa si riconduce quindi alla scelta del sottoinsieme misurabile  $A$  partendo dai valori di *match value* di  $V$ . Tale sottoinsieme raffigura il gruppo di consumatori a cui l'*advertiser* sceglie di rivolgersi conoscendo i *match value* ad essi associati ed è implicitamente legato alla domanda di informazioni: in caso di *positive targeting* è necessario possedere le informazioni del gruppo di consumatori a cui l'*advertiser* vuole rivolgersi; in caso di *negative targeting* l'*advertiser* richiede le informazioni del gruppo di consumatori che si sceglie di escludere.

Sotto le ipotesi di costi  $c$  di matching lineari, *match values* uniformemente distribuiti e ambiente binario in cui  $q$  può assumere solo i valori 0 o 1 (e.g. se  $v > c$  l'*advertiser* contatta il cliente  $q = 1$ ) il profitto per gli *advertisers* è pari a

$$\pi^*(v) \triangleq \max \{v - c, 0\} \quad (2)$$

e il *targeted set* ottimo corrisponde all'intervallo di valori dato da

$$A(c, p) = \begin{cases} [0, \max\{c - p, 0\}], & c < 1/2 \\ [\min\{c + p, 1\}, 1], & c \geq 1/2 \end{cases} \quad (3)$$

dove con  $c < 1/2$  l'impresa opta per il *negative targeting* mentre con  $c \geq 1/2$  si sceglie il *positive targeting*. La dimensione del *targeted set* dipende quindi dalla disponibilità a pagare per un cookie marginale sotto ogni strategia di *targeting*. Da ciò segue che il *residual set* è un insieme non vuoto, le imprese non optano necessariamente per l'acquisto dei cookies dei consumatori ad alto valore ed è il costo  $c$  dello spazio pubblicitario che guida la loro strategia.

Generalizzando le considerazioni precedenti, i.e. considerando un ambiente continuo di azioni  $q$  e una funzione di costo  $m(q)$  strettamente convessa, il valore dell'informazione per ogni *match value* è dato dalla differenza tra il profitto dell'*advertiser* in caso di informazione completa e in caso di informazione preliminare

$$\pi^*(v) - \pi(v, \bar{q}) \quad (4)$$

e i cookies acquistati dagli *advertisers* appartengono al set  $A = [\underline{v}, v_1] \cup [v_2, \bar{v}]$ . Il problema dell'impresa consiste quindi nell'identificare i valori di  $v_1$  e  $v_2$  che definiscono il *targeted* e il *residual set* imponendo l'uguaglianza tra il prezzo per i cookies e il valore dell'informazione (4). Ciò va a confermare quanto presentato in precedenza riguardo l'insieme di valori appartenenti al *residual set*, i.e. l'insieme è non-vuoto).

Sotto le condizioni di costi di matching quadratici ( $m(q) = \frac{q^b}{b}$  con  $b > 0$ ) e *match values* simmetrici, le imprese scelgono di indirizzarsi ad entrambi i tipi di consumatori (*joint targeting*) e quindi adottano sia *positive* che *negative targeting*. La composizione del *residual set* data da

$$A^c(c, p) = [E[v] - 2\sqrt{cp}, E[v] + 2\sqrt{cp}] \quad (5)$$

crebbe con l'aumentare del prezzo delle informazioni  $p$  e del parametro di costo  $c$ : se aumenta anche solo uno di essi diminuisce il numero di cookies acquistati e si riduce il *targeted set* (Fig. 13).

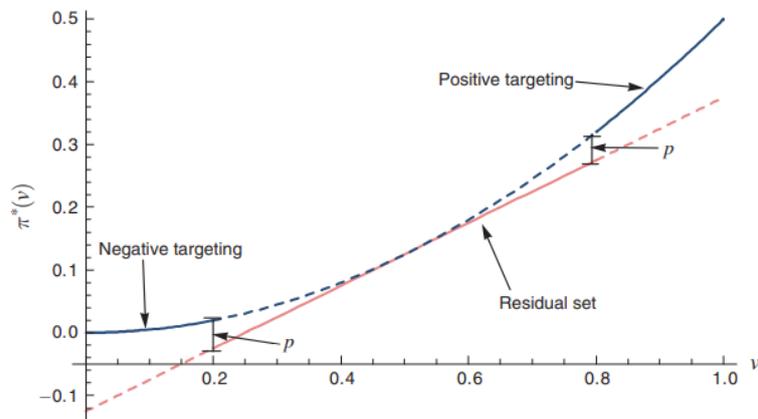


Figura 13 - Domanda per i cookies e livello di profitto risultante – fonte: Bergemann e Bonatti (2015)

Allo stesso modo, l'interazione tra i guadagni derivanti dal possesso delle informazioni e dalla distribuzione simmetrica lascia intravedere la possibilità di applicazione di *exclusive targeting* (Fig. 14). La scelta dipende dalle proprietà di monotonia del valore delle informazioni e della funzione di densità  $f(v)$ . Se la curvatura della funzione di matching cost  $m''(q)$  è decrescente la strategia ottima per l'impresa consiste nell'adottare il *positive targeting*; viceversa se la curvatura è crescente l'impresa adotta il *negative targeting*.

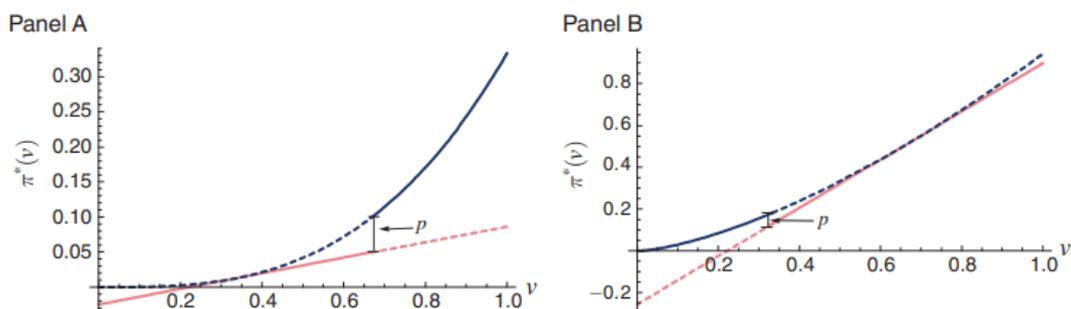


Figura 14 - Profitti di equilibrio sotto positive targeting (sinistra) e negative targeting (destra) - fonte: Bergemann e Bonatti (2015)

La domanda di informazioni degli *advertisers* si ottiene derivando la funzione di profitto ed è quindi pari a

$$q^*(v) = \arg \max [\pi(v, q)], \quad \text{con } q \in R_+ \quad (6)$$

mentre la domanda totale di advertising, che corrisponde alla somma delle domande nel *targeted set* e nel *residual set*, permette di comprendere gli effetti della quantità di dati venduti:

$$M(A) \triangleq \int_A m(q^*(v)) dF(v) + \int_{A^c} m(q^*(A^c)) dF(v). \quad (7)$$

Grazie all'analisi dell'interazione tra il mercato dei dati e l'advertising è inoltre possibile esprimere delle considerazioni sui *data providers* e i *publishers*, i.e. coloro che detengono lo spazio per la pubblicità. Le grandi imprese come Google o Facebook, infatti, giocano un doppio ruolo: forniscono informazioni e permettono la customizzazione dell'advertising e, allo stesso tempo, permettono agli *advertisers* di contattare i clienti (ruolo del *publisher*). Esse quindi possono abilitare il completo accesso o nessun accesso ai dati. Se la domanda per lo spazio pubblicitario decresce con  $p$ , i *publishers* traggono indirettamente dei benefici dalla vendita di informazioni facendo *bundling* di dati e spazio per advertising con lo scopo di far crescere anche la singola domanda inizialmente decrescente.

#### *Il data seller e il prezzo dei dati*

Focalizzandosi, invece, sul prezzo  $p$  fissato dal *data provider*, che si assume operi in un regime di monopolio, esso non influenza qualitativamente le strategie di *targeting* dell'impresa ma solo la dimensione del *targeted set*. Il dataset di cui dispone il monopolista potrebbe coprire tutti i consumatori ma si potrebbe avere accesso limitato ai dati con conseguente prezzo limitato sull'acquisto dei cookies. In ogni caso, è il monopolista che determina il prezzo per accedere alle preferenze del consumatore e tale prezzo è pari a

$$p^* = \arg \max [p \cdot \mu(A(c, p))] = \arg \max [p \cdot v^*] \quad (8)$$

dove  $\mu(A(c, p))$  rappresenta il valore dei dati nel *targeted set*. Le seguenti analisi sono tutte inerenti l'analisi del prezzo per i cookies assumendo sempre costante la dimensione di  $A(c, p)$ .

Ipotizzando di trovarsi in un ambiente binario, i.e. l'ambiente in cui si adotta il *positive targeting* se il costo  $c$  per l'*advertiser* è sufficientemente alto e viceversa si adotta il *negative targeting*, il prezzo di monopolio per i cookies corrisponde a

$$p^*(c) = (1/2) \min \{c, 1 - c\} . \quad (9)$$

Al contrario di ciò che suggerisce l'intuito, le nicchie di mercato con alti costi per lo spazio pubblicitario e pochi consumatori profittevoli non sono il migliore ambiente da cui un *data provider* può acquisire informazioni (i.e. la scelta del *positive targeting* non è sempre ottima). Le condizioni di mercato con grandi frazioni tra i consumatori o con consumatori sia profittevoli che non profittevoli, invece, implicano un maggior valore per le informazioni. Per il *data seller* questo vuol dire poter attribuire un prezzo superiore per i dati con conseguente incremento dei profitti.

Sotto l'ipotesi di competizione tra *data providers* e di ambiente continuo, si è in grado di valutare gli effetti delle esternalità di rete e di composizione del *residual set* che ricadono sui singoli *seller* quando ciascuno di essi fissa il proprio prezzo  $p$ :

$$p(v, v_2) \triangleq \pi^*(v) - \pi(v, q^*([v, v_2])) . \quad (10)$$

Tale prezzo per i cookies è superiore in caso di competizione tra *data sellers* rispetto al caso di monopolio. Infatti, se un monopolista espande la propria offerta di cookies si espande anche l'intervallo di valori che appartengono al *targeted set* e il prezzo per i cookies sarà più basso. Allo stesso tempo, anche il *residual set* cambia, si restringe, e il livello di pubblicità a cui è sottoposto sarà inferiore, attribuendo maggior valore alle informazioni dei consumatori marginali posizionati proprio sotto il valore soglia del *targeted set*. Questo effetto rappresenta un ulteriore incentivo ad abbassare i prezzi, cosa che invece non accade nel caso di competizione tra *data sellers* poiché

essi non internalizzano le esternalità positive tra i consumatori e il *data seller* che si manifestano dalla vendita dei cookies.

Considerando sempre un ambiente continuo ma tornando al caso di *data seller* in monopolio, si esplorano ora le conseguenze sul prezzo dei cookies quando si possiede solo una frazione  $\beta < 1$  di informazioni per tutti i consumatori. La misura di questi ultimi all'interno del dataset è pari a  $\beta F(v)$  dove  $F(v)$  indica la funzione di distribuzione dei valori dei quali si conosce il *match value*. Quando il database è frammentato, la disponibilità a pagare marginale dell'*advertiser* aumenta con l'aumentare della frazione di database posseduta dal *data provider*. Al contrario, il prezzo di monopolio per i cookies subisce una spinta verso il basso dovuta a due effetti:

- La disponibilità a pagare per i dati del *targeted set* aumenta e, inizialmente, il prezzo subisce un incremento. Con l'aumentare di tale prezzo diminuisce la domanda degli *advertisers* e si arriva ad una situazione di eccesso di offerta di informazioni (Fig. 15). Il mercato quindi reagisce e il prezzo per i cookies cala nuovamente.
- L'effetto di composizione, i.e. *targeted set* più *residual set*, fa sì che i consumatori al di fuori del primo sottoinsieme siano meno propensi ad avere un alto *match value* e ciò riduce la quantità di advertising domandata nel *residual set*. Se la domanda diminuisce anche il prezzo per i cookies si muove nella stessa direzione al fine di riattivare il mercato.

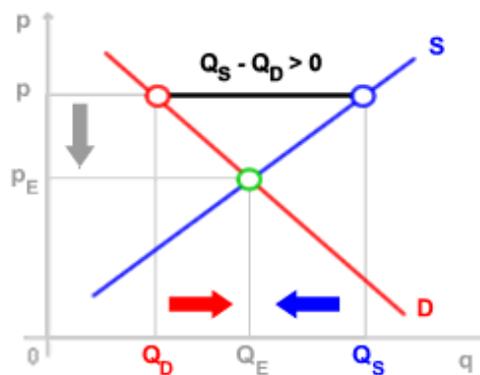


Figura 15 - Rappresentazione grafica dell'Eccesso di Offerta – fonte: [www.okpedia.it](http://www.okpedia.it).

Aggiungendo l'ipotesi che i dati raccolti dal *data seller* abbiano un diverso grado di precisione, rappresentato dal parametro di diffusione  $k$  che prende il significato anche di una varianza, è possibile, in ultimo, dimostrare che il prezzo per i cookies aumenta con l'incremento di  $k$ . Questa considerazione è valida però solo sotto la condizione di *joint targeting*, i.e. *positive targeting* e *negative targeting*, poiché gli effetti nel caso di targeting esclusivo sono intricati: il prezzo di monopolio subisce un impatto positivo al contrario di ciò che succede con le quantità di dati vendute.

*I consumatori e la loro marginal willingness to pay*

Estendendo le ipotesi di partenza e introducendo un parametro indicante la disponibilità a pagare marginale  $\theta$  dei consumatori varia la funzione obiettivo del modello in esame

$$\pi(v, q, \theta) \triangleq \theta v q - c * m(q) . \quad (11)$$

In caso di ambiente binario il valore netto del match, i.e. il profitto per gli *advertisers* è dato da

$$\pi^*(v, \theta) \triangleq \max \{ \theta v - c, 0 \} \quad (12)$$

e affinché l'advertising generi un valore positivo è necessario che la disponibilità a pagare dei consumatori sia maggiore del costo per lo spazio pubblicitario ( $v \in (0,1)$  poiché  $q \in (0,1)$ ). In questa estensione del modello base, gli autori suppongono che gli *advertisers* ricevano informazioni su ogni possibile realizzazione di  $v$  e ciò permette loro di contattare esclusivamente quei consumatori dei quali possiedono i dati, ignorando tutto il resto (*positive targeting*). Ogni impresa che comunica la propria disponibilità a pagare riceve in cambio un set di cookies

$$Q(\theta) \triangleq 1 - F(v^*(\theta)) \quad (13)$$

e la tariffa di pagamento per i cookies si indica con  $T(\theta)$ .  $Q(\theta)$  e  $T(\theta)$  sono strettamente crescenti con  $\theta$  e questo permette di definire uno schema di prezzo non-

lineare al quale gli *advertisers* devono sottostare. Il prezzo marginale che essi pagano per i cookies è dato da

$$\begin{aligned} p(Q) &\triangleq P'(Q) \\ P(Q) &\triangleq T(Q^{-1}(\theta)). \end{aligned} \tag{14}$$

Da ciò segue che il prezzo incrementale  $p(Q)$  permette di implementare una politica informativa ottima in cui tale prezzo segue un andamento decrescente all'aumentare di  $Q$  e in cui gli *advertisers* hanno accesso a delle porzioni di database (i.e. il meccanismo ottimo viene decentralizzato) secondo la logica degli sconti quantità, sconti per chi domanda una maggiore quantità di cookies.

#### *Conclusioni preliminari*

Dal modello si ottengono quindi due conclusioni. Il primo risultato riguarda la domanda per avere accesso alle informazioni dei consumatori e la customizzazione della spesa degli *advertisers* a tale proposito. Si evince che, in caso di informazione completa, le imprese possono raggiungere l'ottimo nelle sopracitate strategie di acquisto delle informazioni, sia con *joint targeting* che con *exclusive targeting*. Il secondo risultato, invece, stabilisce che il raggiungimento di un prezzo lineare per i cookies (i.e. l'ottimo del modello) dipende dall'abilità di influenzare la composizione del *targeted set* e del *residual set* di consumatori. Esiste un incentivo ad abbassare il prezzo per le informazioni e ciò dipende sia dal raggiungimento del database (i.e. con l'acquisizione di una quantità di dati sempre maggiore) sia dalla concentrazione delle vendite. Allo stesso modo, il prezzo per le informazioni aumenta se la vendita dei dati diventa frammentata. Inoltre, quando il consumatore possiede una certa disponibilità a pagare per i cookies, il *data provider* può decentralizzare il meccanismo ottimo di scambio offrendo un prezzo non-lineare e permettendo agli *advertisers* di fare auto-selezione del sottoinsieme di cookies.

## 2.3. Abuso di posizione dominante: abusi di esclusione

Con il termine abuso di posizione dominante ci si riferisce alla situazione in cui un'impresa è in grado di restringere la concorrenza se è in una posizione di forza su un determinato mercato, nel caso in esame nel *data market*. La posizione dominante non impedisce la concorrenza in quanto tale bensì deve sussistere uno sfruttamento di questa posizione che determini un abuso nei confronti di imprese o di consumatori. Nel caso di abuso di un'impresa nei confronti di altre imprese si parla di abusi di esclusione. Quando invece l'abuso è diretto ai consumatori ci si riferisce al caso di abusi di sfruttamento.

Secondo Delmastro e Nicita (2019) è altresì presente un caso intermedio tra il concetto di posizione dominante e abuso che però non trova spazio né in Europa né negli Stati Uniti poiché ritenuto un comportamento ragionevole: “il tentativo di monopolizzazione”. L'onere della prova, in questo caso, è in capo a chi denuncia e le condizioni che devono sussistere affinché si verifichi tale situazione sono tre:

- Deve esservi specifica intenzione di monopolizzare un mercato.
- Deve essere manifestato un comportamento volto a realizzare tale intenzione.
- Deve esistere una concreta probabilità di successo.

Nei seguenti due paragrafi verrà presentata un'analisi della letteratura sulle due principali tipologie di abuso, affrontando inizialmente l'abuso di esclusione.

### 2.3.1. La difesa degli innovatori

Uno dei quesiti chiave per l'Antitrust, messo in evidenza anche in un articolo del Sole 24 Ore<sup>7</sup>, riguarda la difesa delle imprese che decidono di collocarsi sul mercato nella posizione di innovatori. È necessario difendere gli innovatori da quelle aziende già forti sul mercato, con una forte immagine di marca, che quindi non solo

---

<sup>7</sup> De Blase, L. (4 Agosto 2019). *Il monopolio nell'era digitale. Un Antitrust che salvaguardi l'innovazione*. Il Sole 24 Ore. [www.datastamp.it/index.php/home/](http://www.datastamp.it/index.php/home/).

sfruttano le economie di rete, il *data spillover* ma anche il fatto di avere un brand consolidato?

Il mondo digitale è caduto definitivamente in mano alle piattaforme con l'invenzione dello smartphone poiché questo ha incrementato il processo di organizzazione delle informazioni e aggregazione di dati. L'Antitrust si dimostra, infatti, molto attento alla difesa non solo dei consumatori quanto più degli innovatori soprattutto in un mercato con caratteristiche così peculiari come il *data market*. In passato però questo è stato un tema affrontato anche in altri ambiti, basti pensare a quando Microsoft, per raggiungere la massa critica di utenti, ha iniziato a regalare il sistema operativo Windows favorendo così una penetrazione decisamente vasta a scapito di altre imprese che offrivano sistemi diversi.

Uno degli obiettivi perseguiti da chi opera nel *data market*, e sostenuto dall'Antitrust, consiste nel cercare di proteggere le imprese e in particolar modo chi intraprende un percorso da innovatore dalle condotte abusive delle grandi imprese (e.g. le *over-the-top*: Google, Facebook, Amazon). Resta quindi focale il fatto di permettere l'innovazione, di far emergere chi riesce ad interpretare meglio delle imprese dominanti i vantaggi che offre la tecnologia e scongiurare il più possibile un eventuale abuso di posizione dominante.

Un esempio di successo, a tal proposito, è dato da settore del FinTech e da imprese quali Satispay<sup>8</sup> o N26<sup>9</sup> le quali hanno saputo sfruttare al meglio i vantaggi forniti dalle piattaforme rispetto alle tradizionali banche (i.e. gli incumbent del settore): questo è un caso in cui l'innovazione, che non è partita dall'incumbent ma da chi si affaccia sul mercato per la prima volta, ha avuto successo senza che le imprese con maggior *market power* potessero ostacolare tale trend evolutivo.

---

<sup>8</sup> App per gestire pagamenti tramite Smartphone, la quale permette di gestire transazioni solo mediante l'inserimento del numero di cellulare e del codice IBAN e senza carta di credito o di debito. [www.satispay.com/it-it](http://www.satispay.com/it-it).

<sup>9</sup> È la prima banca diretta completamente adibita e progettata per smartphone. Offre diversi servizi finanziari e presenta funzionalità innovative (e.g. la gestione degli Spaces). [n26.com/it](http://n26.com/it).

Al contrario, come presentato nel report finale dell' Australian Competition & Consumer Commission (2019), la diatriba<sup>10</sup> tra l'impresa innovatrice Spotify ed Apple, che ha coinvolto anche la Commissione Europea la quale attualmente sta indagando a tal proposito, dimostra come la necessità di protezione nei confronti degli innovatori sia un bisogno sempre più sentito al giorno d'oggi. Nello specifico, si è riscontrato che Apple, proprietaria sia della piattaforma iOS sia di Apple Store, applica comportamenti che portano all'azienda stessa vantaggi anti-competitivi. Ad esempio, Apple impone a Spotify e ad altri servizi digitali, i.e. gli sviluppatori delle applicazioni, una tassazione del 30% sugli acquisti che i consumatori eseguono mediante sistema di pagamento Apple. Allo stesso tempo, se gli sviluppatori non prevedono l'impiego di tale sistema di pagamento Apple applica alle imprese numerose limitazioni tecniche e restrizioni. Questo comportamento, definito abusivo dal CEO di Spotify, è proprio ciò che è necessario scongiurare in un mondo in cui l'innovazione è il motore del successo.

### **2.3.2. Il problema di Mergers and Acquisitions**

Le strategie di Mergers and Acquisitions (M&A) sono generalmente volte a ridurre il numero di competitor sul mercato. Il confine tra una sana competizione e ciò che invece rientra in abuso escludente è sottile.

Come evidenziato da Dubus (2019), il numero di imprese che decide di ricorrere allo strumento di M&A è in aumento, in particolar modo sul mercato dei dati. I *policy maker*, infatti, guardano con occhio critico alla recente ondata di merger che ha visto protagoniste molte imprese che svolgono il ruolo di *data broker*. Tuttavia, non sono ancora ben note le conseguenze causate da un aumento di concentrazione nel *data market* sul potere di mercato. Ad esempio, la nota azienda Oracle nel 2014 ha acquistato due grandi compagnie che svolgono da intermediarie nello scambio di dati quali Bluekai<sup>11</sup> and Datalogix<sup>12</sup> e ciò ha causato un incremento di concentrazione nel settore il cui impatto non è stato ben definito ma che di certo non è passato inosservato.

---

<sup>10</sup> [newsroom.spotify.com/2019-03-13/consumers-and-innovators-win-on-a-level-playing-field](https://newsroom.spotify.com/2019-03-13/consumers-and-innovators-win-on-a-level-playing-field).

<sup>11</sup> Piattaforma per la raccolta e l'immagazzinamento dei big data basata su Cloud. [www.oracle.com/data-cloud/products/data-management-platform](https://www.oracle.com/data-cloud/products/data-management-platform).

<sup>12</sup> Compagnia che colleziona dati sui consumatori. [www.oracle.com/corporate/acquisitions/datalogix](https://www.oracle.com/corporate/acquisitions/datalogix).

La visione di Crémer, de Montjoye, Schweitzer (2019) si focalizza sui M&A delle imprese che limitano la concorrenza acquisendo i piccoli competitor. Tale analisi si fonda sul fatto che, generalmente, le start-up sfuggono al tradizionale controllo sui *mergers* poiché non generano un sufficiente *turnover* e poiché spesso è complicato distinguere una condotta anti-competitiva da una pro-competizione o neutra (e.g. merger verticali: le imprese integrano prodotti o attività complementari che non fanno parte del loro *core business*). La Commissione Europea ritiene, quindi, necessario che rientrino sotto indagine anche le imprese con fatturati decisamente inferiori a quelli che sono gli attuali valori previsti delle regole europee.

L'aspetto positivo delle strategie di M&A è che possono portare allo sviluppo di nuovi servizi grazie all'accesso ad un dataset più ricco. Allo stesso tempo, possono essere altresì la causa di un controllo maggiormente concentrato, a tal punto da non rendere possibile per i competitor la replicazione delle risorse di cui le imprese che fanno *merger* si servono (i.e. il dataset). Tutto ciò ha conseguenze sull'incremento del potere di mercato e sull'insorgere di un abuso di tipo escludente nei confronti delle imprese nel settore.

L'acquisizione di un'impresa ad alto tasso innovativo da parte di una piattaforma dominante che beneficia di forti esternalità di rete, non rientra nella categoria delle cosiddette *killer acquisition*<sup>13</sup> ma è da tenere sotto osservazione. In particolare, è necessario analizzare le situazioni in cui l'ecosistema con cui si vuole fare *merger* ha un basso fatturato ma una base di utenti che cresce velocemente e, in aggiunta, un potenziale futuro di mercato alto (i.e. ci si aspetta uno sviluppo positivo).

### **2.3.3. I prezzi predatori: “Zero-Price problem”**

Nei mercati digitali si riscontrano spesso beni o servizi caratterizzati da un prezzo nullo. Tale fenomeno è ancor più evidenziato dalla presenza di esternalità di rete che coinvolgono una piattaforma e i consumatori. Crémer, de Montjoye, Schweitzer (2019) identificano questo fenomeno con il nome di “Zero-Price problem”.

---

<sup>13</sup> L'acquisizione di un'impresa innovativa avviene per evitare che tale impresa emergente porti a termine il proprio obiettivo, potenzialmente dannoso per l'incumbent che sente minacciato il suo potere di mercato.

Si è riscontrato, infatti, che alcune imprese applicano la strategia dei prezzi predatori, ovvero decidono di abbassare i prezzi dei beni o servizi venduti così tanto da escludere completamente i rivali dal mercato, i quali si trovano a non poter in alcun modo competere. All'interno del *data market*, generalmente, le piattaforme applicano un prezzo nullo ai consumatori ma sussidiano il versante non a pagamento ottenendo profitti dall'advertising. Inoltre, è importante considerare che, la piattaforma può guadagnare anche dal lato "Zero-Price": essa acquisisce i dati dei consumatori e offre loro un miglior servizio.

Una parziale soluzione potrebbe derivare dall'applicazione del test SSNIP, un test atto a verificare il potere di mercato di un'impresa e quindi strumento idoneo, spesso impiegato anche dall'Antitrust, a constatare se è effettivamente presente un abuso di esclusione. Tale test consiste nel verificare se un aumento dei prezzi del bene/servizio spinge il consumatore verso un altro prodotto. Se ciò non si verifica significa che potrebbe sussistere realmente una situazione di abuso di posizione dominante.

Nel caso in esame, però, gli autori del report non ritengono opportuno l'impiego di questo metodo poiché sostengono che non abbia senso incrementare i prezzi da un lato del versante senza modificare ciò che succede nell'altro lato. La soluzione migliore sarebbe stabilire delle chiare linee guida ma al momento le autorità che regolano la competizione non hanno ancora delineato un profilo da seguire.

## **2.4. Abuso di posizione dominante: abusi di sfruttamento**

La seconda tipologia di abuso presa in considerazione è volta ad esplorare le condotte anti-competitive nei confronti dei consumatori e le condotte di mercato che inducono gli stessi a compiere scelte non consapevoli. Come descritto nel paragrafo precedente, tale abuso prende il nome di abuso di sfruttamento.

### **2.4.1. L'eccessivo sfruttamento dei dati**

L'introduzione dei big data ha determinato un nuovo approccio al trattamento dei dati: essi sono allo stesso tempo uno strumento di indagine utile per soddisfare

esigenze o comprendere la domanda di beni/servizi ma anche un prodotto capace di generare valore in sé (Delmastro e Nicita, 2019). Tutto ciò comporta innumerevoli aspetti positivi per il mercato e, in sintesi, si può stabilire che i dati degli utenti abilitano le piattaforme digitali a creare profili molto dettagliati dei consumatori.

Il consumatore, invece, prende una posizione diversa: le operazioni sui dati sono permesse fintanto che non comportano una violazione della sfera personale. Esiste, quindi, un'intricata relazione tra privacy, consumatore e dati. L'eccessivo sfruttamento dei dati dei consumatori e l'acquisizione di enormi quantità di informazioni, infatti, può portare a conseguenze quali prezzi discriminati negativi (i.e. inducono il consumatore a consumare meno), prezzi gravosi per beni e servizi e scelte di acquisto e consumo non consapevoli; uno dei mezzi attraverso cui le imprese facilmente riescono a mettere in atto tali condotte abusive è il *target advertising*.

Il mercato del *target advertising* si è decisamente ampliato con l'avvento del mondo digitale grazie a tecnologie che hanno permesso la raccolta dei cookies per offrire prodotti personalizzati e per condurre analisi sui comportamenti degli utenti in seguito all'esposizione di pubblicità mirate (i.e. *targetability* e *measurability*). Nello specifico, il problema della pubblicità targetizzata è già stato affrontato nel Paragrafo 2.2.1 ma sotto un punto di vista differente: è stato preso in considerazione lo studio di Bergemann e Bonatti (2015) inerente il rapporto consumatore-*advertiser* focalizzandosi sull'aspetto *pricing*. Ora, invece, l'intento è di approfondire il comportamento delle imprese che sfruttano la pubblicità targetizzata per operare scorrettamente nei confronti dei consumatori che vedono violata la propria privacy.

Secondo la visione dell'Australian Competition & Consumer Commission (2019) sono due i motivi che influenzano la capacità del consumatore di compiere scelte consapevoli:

- Le asimmetrie informative presenti tra consumatori e piattaforme sono molto forti. Spesso il consumatore è del tutto inconsapevole di quanti dati sono stati collezionati o come e questo è una conseguenza del fatto che le policy di privacy sono spesso ambigue o non permettono al consumatore un intervento adeguato.

- Le piattaforme digitali, rispetto ai consumatori, detengono un potere contrattuale molto alto. Le piattaforme, infatti, chiedono il consenso per i dati nella forma del tipo “prendere o lasciare” oppure sottopongono l’utente ad un *bundle* di consensi, per cui aderendo ad una richiesta, in automatico si esprime una preferenza anche per le richieste collegate.

Quest’ultimo tipo di condotta rientra in ciò che è stato definito come abuso di sfruttamento: il consumatore non possiede l’abilità di operare liberamente se intende usufruire del servizio e non è in grado di agire in modo informato.

Un ulteriore aspetto è chiamato in causa dal governo del Regno Unito (HR Government, 2019), il quale identifica il *target advertising* come responsabile di danni morali e psicologici ai consumatori quali manipolazioni messe in atto mediante propaganda e false informazioni. Internet, i social media e l’intelligenza artificiale hanno permesso l’acquisizione di grandissime moli di dati che sono state rese effettive e protagoniste della manipolazione di gran parte dell’opinione dei consumatori. A questo proposito è, infatti, inevitabile non citare il recente scandalo politico Facebook-Cambridge Analytica<sup>14</sup>.

Esiste, però, anche un altro risvolto sulla questione evidenziato da Goldfarb e Tucker (2011). In Europa, più che in altre regioni, sono presenti meccanismi di regolazione sull’uso e sfruttamento dei dati degli utenti e ciò è conforme con quanto voluto dai consumatori a protezione della privacy e per scongiurare qualsiasi tipo di condotta abusiva. Tali meccanismi hanno quindi limitato la possibilità degli

---

<sup>14</sup> Cambridge Analytica è una società statunitense che ha sviluppato un sistema di microtargeting comportamentale, i.e. pubblicità altamente personalizzata su ogni singola persona che fa leva non solo sui gusti, come fanno già altri sistemi analoghi per il marketing, ma anche sulle emozioni degli utenti. Tale impresa ha acquistato dal proprietario di un’applicazione collegata a Facebook i dati degli utenti (interessi, fotografie, aggiornamenti di stato, etc.) archiviati dall’applicazione stessa a causa di una falla nel sistema di Facebook che lo rendeva possibile. Essendo Cambridge Analytica la società che ha gestito la campagna elettorale di Donald Trump nel 2016, la società è stata accusata di aver influenzato l’opinione pubblica mediante creazione di bot (account fasulli gestiti automaticamente) che hanno diffuso post, notizie false e altri contenuti a sfavore della rivale Hillary Clinton. Inoltre, anche Facebook è stata messa a processo poiché accusata di essere a conoscenza del problema di archiviazione che ha coinvolti milioni di utenti e ancor più le loro informazioni private. A luglio 2019 la Federal Trade Commission ha imposto a Facebook una multa pari a circa 5 miliardi di dollari. Fonte: [www.ilpost.it/2018/03/19/facebook-cambridge-analytica](http://www.ilpost.it/2018/03/19/facebook-cambridge-analytica).

*advertisers* di collezionare dati. L'effetto negativo di queste limitazioni, come evidenziato dagli autori, consiste nella perdita di efficacia della pubblicità targettizzata in quanto pensata e diretta al singolo consumatore. Infatti, la presenza di maggior regolazione non impedisce agli *advertisers* di offrire i propri prodotti (i.e. i consumatori sono ugualmente sottoposti a pubblicità) ma ricade sul consumatore che riceve pubblicità non consistenti con le proprie preferenze.

## 2.5. La privacy sul mercato dei dati

La protezione dei dati personali e della privacy è un diritto umano. Proprio per questo, le varie autorità regolatrici dei paesi devono impegnarsi per imporre un chiaro e consistente sistema legale per la protezione dei dati.

L'ecosistema che ruota attorno al mondo della privacy è molto vasto: esistono, infatti, una moltitudine di definizioni e scenari diversi sul tema della privacy e la Figura 16 ne è un chiaro esempio (Acquisti, Taylor, Wagman, 2016). Tuttavia, nonostante le evidenti diversità, ciascuno di questi aspetti implica la gestione e la negoziazione dei confini tra gli aspetti pubblici e privati.

L'unione tra studi teorici e studi empirici proposta dagli autori porta alla luce tre aspetti. In primo luogo, essi evidenziano che la protezione della privacy ha effetti ambigui sul surplus economico e sull'efficienza allocativa: fare *disclosure* di informazioni personali porta a benefici sia per i possessori di dati sia per i soggetti che all'origine ne erano i proprietari (e.g. risparmi, guadagni di efficienza, incremento nei ricavi, customizzazione dei prodotti, offerte e promozioni); allo stesso tempo, però, tale *disclosure* può risultare costosa per tutti gli attori coinvolti nel processo poiché implica costi dovuti ad una acquisizione decisamente intrusiva da parte delle imprese oppure costi meno tangibili legati alla discriminazione di prezzo o ai danni psicologici che subiscono i consumatori. Il secondo aspetto rilevante riguarda l'abilità dei consumatori di prendere decisioni informate e consapevoli riguardo la propria privacy: essi, sovente, si trovano in situazioni di informazione imperfetta con riferimento al quando, a che scopo e con quali conseguenze i dati sono stati collezionati. L'ultimo aspetto prende in considerazione specifiche euristiche che possono influenzare

profondamente le decisioni sulla privacy dal momento che il trade-off a cui è soggetto un individuo nel momento in cui deve scegliere se fare *disclosure* è spesso ambiguo.

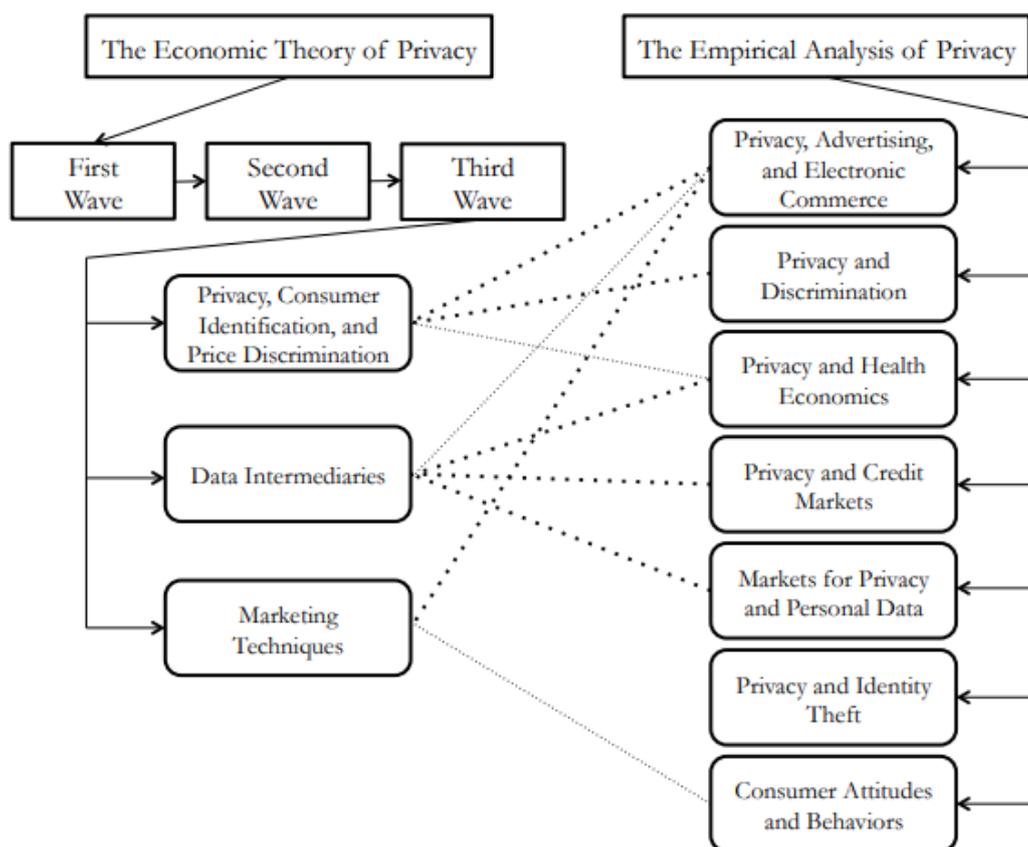


Figura 16 – Rappresentazione delle connessioni tra teoria e risultati empirici - fonte: Acquisti, Taylor, Wagman (2016)

Sempre tramite uno studio di Acquisti, questa volta in compagnia di Brandimarte e Loewenstein, vengono presentate alcune verifiche empiriche sui comportamenti legati alla privacy. In tale studio il focus è spostato su quanto gli individui siano consapevoli di come i loro dati vengono usati in rete, su come e perché essi decidono di rivelare informazioni e su quali fattori influenzano una maggiore/minore attenzione ai temi della privacy (Acquisti, Brandimarte, Loewenstein, 2015). In particolare, gli aspetti presi in considerazione riguardano i seguenti temi:

- Incertezza. Una prima forma di incertezza deriva dall'asimmetria informativa, che, come si è visto anche in precedenza, sorge a causa dell'avanzamento

tecnologico il quale, molto spesso, rende la *data collection* un processo invisibile agli occhi dei consumatori. Come conseguenza, l'individuo è molto incerto sulla quantità di informazioni con la quale vuole fare *disclosure*.

- La dipendenza dal contesto in cui ci si trova. Uno stesso individuo può in alcuni casi attribuire molta importanza al tema della privacy e quindi interessarsi molto al percorso cui sono soggette le proprie informazioni; in altri casi può completamente disinteressarsi della questione. La difficoltà maggiore per gli utenti, inoltre, si trova nella definizione dei confini del contesto. Quando le interazioni avvengono online, l'utente non possiede una chiara rappresentazione dei confini spaziali (i.e. i lettori) cui i messaggi e le informazioni che genera sono diretti. Diventa quindi difficile distinguere ciò che è pubblico da ciò che è privato e si genera ulteriore complessità al processo di *decision-making*.
- La malleabilità e l'influenza di esterni (governi, piattaforme digitali). Il comportamento degli individui e le preferenze inerenti la privacy sono in gran parte malleabili poiché tali decisioni sono dipendenti dal contesto e il potere dei soggetti che hanno interesse a vedere una divulgazione delle informazioni è molto forte (e.g. *advertisers*). La manipolazione cui sono soggetti gli individui si manifesta spesso attraverso le decisioni di condivisione/non condivisione di informazioni sui social network oppure fornendo più controllo all'utente, il quale si illude di poter condividere le informazioni con modalità più sicure.

A conferma di quanto presentato sinora, è stato condotto, inoltre, un esperimento online in cui agli utenti sono state poste domande intrusive sul comportamento personale in determinate situazioni. Ciò che si è riscontrato è che, a seconda dell'interfaccia grafica con cui si presentavano le domande (in un caso molto professionale, nell'altro meno), gli utenti hanno fornito un livello di *disclosure* di informazioni diverso. Nello specifico, il questionario con il design più abbozzato e banale è stato quello a cui i soggetti hanno rivelato maggiori informazioni private, nonostante il design del questionario professionale fosse percepito dai rispondenti come più sicuro (Fig.17).

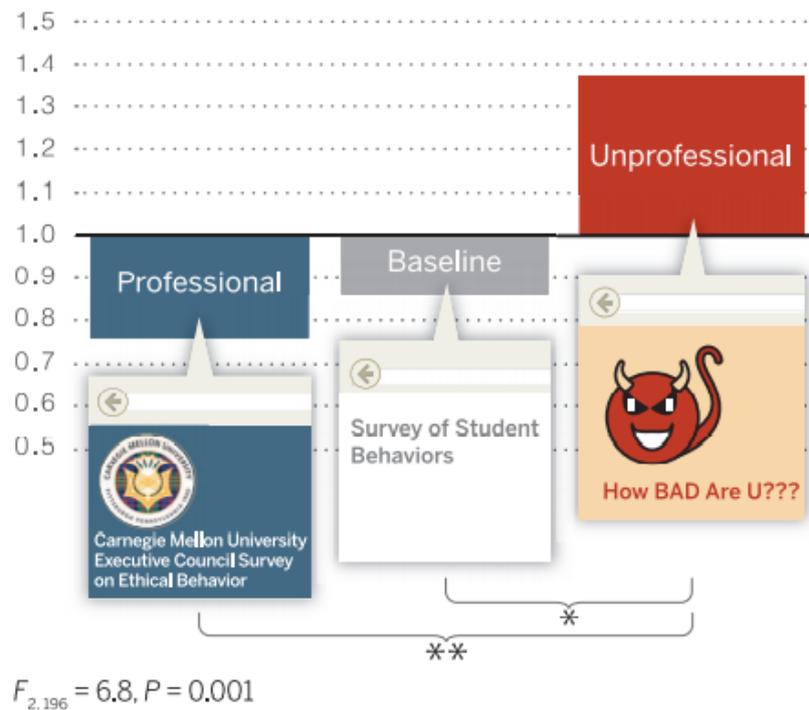


Figura 17 – Risultati dell’esperimento di privacy-disclosure – fonte: Acquisti, Brandimarte, Loewenstein (2015)

Da questo esperimento sociale si può concludere che, nonostante il comportamento dei soggetti vari a seconda della cultura di origine, il dilemma su cosa condividere o cosa mantenere privato permane e dipende fortemente dai tre elementi presentati in precedenza. La dipendenza dal contesto e la malleabilità di cui sono stati protagonisti gli individui rispondenti al questionario si è manifestata attraverso semplici meccanismi grafici. Appare, quindi, evidente il potere che risiede nelle mani delle piattaforme digitali le quali hanno accesso in modo più o meno autorizzato alle informazioni degli individui.

Guardando al tema della privacy dal punto di vista economico, lo studio di Jentzsch, Preibusch, Harasser (2012) offre un interessante spunto di analisi. Il progetto messo in atto dagli autori si focalizza in particolar modo sul ruolo della privacy nelle transazioni economiche e prevede due fasi: la prima fase consiste nella creazione e applicazione di un modello economico, la seconda fase riguarda la validazione di tale modello attraverso alcuni esperimenti pratici.

Il modello è stato sviluppato assumendo che vi siano due *service providers* in competizione e che i consumatori debbano scegliere a quali di essi affidarsi per compiere un acquisto in un ambiente caratterizzato da prodotti differenziati (differenziazione di prezzo), e.g. prodotti customizzati o differenziati per tipo di dati personali richiesti. Per semplificare l'analisi sono stati presi in considerazione solo due gruppi di consumatori: coloro che attribuiscono grande importanza alla privacy e, al contrario, coloro che ne attribuiscono un basso valore. Dall'applicazione di questo modello, presentato in versione ad uno stadio e successivamente a due stadi con interazione tra consumatore e fornitore di servizi, si ottengono valutazioni e approfondimenti sul comportamento dei *service providers* riguardo l'acquisizione di dati personali dei consumatori in un ambiente molto competitivo. Inoltre, è anche possibile analizzare come il consumatore reagisce quando viene a conoscenza dell'acquisizione dei propri dati.

Gli esperimenti pratici (esperimenti di laboratorio e studi sul campo) vengono messi in atto, infatti, proprio a questo scopo, i.e. applicare praticamente quanto stabilito dal modello. Dall'esperimento di laboratorio, eseguito in un ambiente controllato in cui i partecipanti sono al corrente dello studio, si testano le teorie e le assunzioni microeconomiche, in particolare si analizzano gli incentivi economici e le decisioni degli individui coinvolgendoli direttamente in alcuni compiti e azioni di acquisto del prodotto; al termine dell'esperimento essi dovranno poi rispondere ad alcune domande inerenti la privacy e la protezione dei dati personali. Lo studio sul campo, invece, ha coinvolto persone che non sapevano di far parte di un esperimento (studenti di una grande università tedesca) le quali dovevano portare a termine la transazione di acquisto mediante un sito web.

Gli autori hanno raccolto i dati e le risposte fornite da tutti i partecipanti agli esperimenti e hanno concluso che, in generale, gli individui attribuiscono molta importanza alle questioni di privacy e sono molto interessati all'argomento. In aggiunta, si è riscontrato che

- circa il 90% dei partecipanti, in caso di acquisti ripetuti, non *cambia service provider*;

- se il prezzo offerto dai due *service providers* non cambia gli utenti preferiscono acquistare da colui che si dimostra maggiormente *privacy-friendly*;
- se tale prezzo invece cambia, in particolare se il fornitore *privacy-unfriendly* applica un prezzo inferiore, si verifica una perdita di quote di mercato del fornitore *privacy-friendly*;
- circa un terzo degli individui sono disposti a pagare di più per il servizio offerto dal fornitore *privacy-friendly*.

In conclusione, si può stabilire che in caso di poche differenze di prezzo tra i due *service providers*, un buon competitor che riesce a ridurre l'acquisizione di dati personali degli utenti possiede un vantaggio competitivo.

Tuttavia, dallo studio di Chiou e Tucker (2017) emerge che la chiave per garantire la protezione della privacy degli utenti e la sicurezza dei dati degli individui è il tempo per cui tali dati vengono conservati. Ridurre la quantità di *storage* di dati, i.e. ridurre lo sfruttamento, conferisce agli individui un beneficio inferiore a quello ipotizzato. Tale provvedimento influenza solo debolmente l'accuratezza delle ricerche mediante *web browsing* e ciò significa che le piattaforme rimangono comunque in possesso di una quantità di dati sufficiente per poter attuare condotte abusive. I motori di ricerca più potenti mantengono quindi la posizione di dominanza sul lungo periodo e non vi sono garanzie che ciò porti ad una maggiore tutela sulla privacy.

## **2.6. Implicazioni di policy**

Dopo aver analizzato le principali questioni e controversie riportate dalla letteratura sul funzionamento del *data market*, ora si presenteranno le più importanti politiche di tale mercato necessarie per garantire il rispetto della privacy e la tutela del consumatore. Ci si focalizzerà, in conclusione, su alcuni tra i recenti interventi in merito al tema e sugli interventi di policy attuati dai Governi dei diversi stati.

### **2.6.1. Portabilità e interoperabilità**

All'interno del *data market* è necessario lo sviluppo di un ambiente che supporti le innovazioni guidate dai dati e stimoli la domanda per prodotti e servizi che

si appoggiano ai dati come fattore rilevante per la produzione. L'innovazione tecnologica gioca quindi un ruolo chiave nella crescita di nuovi business model per i quali è fondamentale fare affidamento sui concetti di interoperabilità e portabilità dei dati come strumenti di policy a tutela del consumatore.

In accordo con quanto sostenuto da Crémer, de Montjoye, Schweitzer (2019), agli individui deve essere garantita la portabilità dei dati, i.e. i dati che un'azienda raccoglie sui soggetti possono essere utilizzati dagli stessi o trasferiti rapidamente anche ad eventuali aziende concorrenti al fine di evitare situazioni opportunistiche, il tutto nel rispetto della privacy. Ma non solo, le piattaforme digitali devono assicurare al consumatore anche l'interoperabilità, la quale può essere suddivisa in

- interoperabilità dei protocolli, quando deve essere garantito che i due sistemi tra cui avviene lo scambio di dati lavorino insieme e in modo complementare per fornire i servizi richiesti dal cliente;
- interoperabilità dei dati, quando tali dati devono essere accessibili allo stesso tempo da utenti e macchine rispettando le norme di sicurezza.

In questo modo l'impresa che possiede i dati, l'asset fondamentale, è meno incentivata ad esercitare *hold-up* sul consumatore poiché è il consumatore stesso il proprietario ultimo dell'asset e in qualsiasi momento può deciderne il trasferimento. Infatti, come ribadito anche da Delmastro e Nicita (2019), la portabilità dei dati costituisce un valido meccanismo di disciplina per le principali *digital platforms*: si riduce il costo opportunità di abbandonare una piattaforma e i nuovi entranti sul mercato, avendo accesso alle informazioni, possono offrire servizi customizzati per il cliente. Tuttavia, bisogna considerare che l'efficacia della portabilità si ha solo nel momento in cui un nuovo entrante possiede la scala minima di utenti, i.e. riesce a sfruttare adeguatamente le esternalità di rete. Ma soprattutto, la portabilità del dato deve essere accompagnata da meccanismi di aggregazione degli utenti (e.g. aste alle quali le piattaforme partecipano per trattenere o conquistare utenti) e di cancellazione dei dati da parte delle piattaforme cedenti, il tutto con lo scopo di aumentare la competizione all'interno del mercato.

Una visione maggiormente critica sul tema viene proposta dall'Australian Competition & Consumer Commission (2019). La commissione, infatti, sostiene che

spingere il mercato verso la portabilità dei dati non comporti necessariamente una diminuzione nei costi di switching o maggiori entrate nel settore, i.e. gli effetti sono ambigui. Nello stato attuale in cui si trova il *data market*, non esistono piattaforme in grado di competere con le *over-the-top* e i consumatori non riconoscono quindi nessun potenziale sostituito con cui fare *switch*. Inoltre, alcuni servizi quali i servizi bancari, Internet e i social network sono forniti gratuitamente agli utenti: sorge il problema della mancanza di incentivi nel cercare un servizio sostituto a cui applicare le policy analizzate. In conclusione, l'unico vero scopo dell'implementazione della portabilità e interoperabilità dei dati rimane la possibilità di sviluppare nuovi servizi e la creazione di un mercato molto più competitivo rispetto allo stato attuale.

### **2.6.2. Sistemi open standard e trasparenza**

Una forte spinta all'interoperabilità e alla consistenza tra diversi settori arriva dall'impiego di sistemi open standard. I sistemi open standard sono costruiti usando specifiche tecniche condivise, liberamente disponibili sul mercato e mediante processi trasparenti e aperti alla partecipazione di tutte le imprese del settore.

Furman, Coyle, Fletcher, McAuley, Marsden, (2019) ritengono che siano presenti molte similitudini tra i concetti di portabilità e standard aperti poiché entrambi permettono di raggiungere un certo livello di interoperabilità. Tuttavia, la portabilità è riferita alla mobilità di specifici pezzi di dati mentre gli standard aperti implicano la compatibilità dell'intero servizio con quanto offerto da terze parti.

Poter costruire un sistema con open standard porta benefici al mercato: abilita gli innovatori a costruire servizi compatibili e/o complementari con i servizi già esistenti, facilita la competizione poiché è possibile operare in contemporanea su sistemi diversi e scongiurare pericoli quali monopoli naturali ed eccessivi effetti di rete (e.g. se tutti gli individui impiegano lo stesso mezzo di comunicazione) ed evita il *lock-in* dei consumatori all'interno di una stessa piattaforma. Un esempio di quanto i sistemi open standard siano necessari viene dato dai sistemi di comunicazione che funzionano attraverso Internet: un utente, infatti, può solo raggiungere i propri contatti che possiedono un account con lo stesso *service provider* e questo direziona gli individui ad optare verso la piattaforma che gode di maggiori esternalità di rete (e.g. WhatsApp). Allo stesso modo, anche le applicazioni basate sull'Internet of Things dovrebbero

poter essere costruite con sistemi open standard per incentivare l'innovazione tecnologica.

Anche la trasparenza rientra tra le iniziative di policy maggiormente considerate. I consumatori, infatti, si aspettano che la privacy diventi uno dei temi ai quali le piattaforme digitali guardano con interesse, in particolar modo per migliorare l'offerta di mercato e per garantire maggiore trasparenza.

Secondo Crémer, de Montjoye, Schweitzer (2019), la mancanza di trasparenza sul funzionamento del mercato e delle piattaforme permette alle piattaforme stesse di distorcere la competizione. Le autorità pubbliche dovrebbero trovare un modo per assicurare ai consumatori precise informazioni sul funzionamento delle piattaforme digitali, rispettando i segreti del business. Inoltre, ciò permetterebbe agli utenti di comprendere meglio come le informazioni personali e i dati vengono collezionati e sfruttati nel mercato.

### **2.6.3. Ribaltamento dell'onere probatorio**

Quando un incumbent all'interno del *data market* viene accusato di comportamenti anti-competitivi, ciò che succede attualmente è che l'imputato viene dichiarato "innocente" salvo prova della sua colpevolezza, il che può richiedere anni di indagine.

A fronte di un decisivo risparmio di tempo nella gestione delle indagini, Crémer, de Montjoye, Schweitzer (2019) propongono di ribaltare l'onere probatorio a carico dell'incumbent e non a carico di chi accusa. In tale modo l'imputato verrebbe dichiarato "colpevole" fino a prova contraria, i.e. fino alla prova che il servizio o prodotto che limita la competizione è però in grado di migliorare la possibilità di scelta dei cittadini. Questa opzione risulterebbe particolarmente favorevole in tutti i settori altamente concentrati caratterizzati da forti esternalità di rete e conseguenti alte barriere all'entrata in cui è fondamentale disincentivare le condotte anti-competitive.

### **2.6.4. General Data Protection Regulation**

Un valido tentativo di tutela nei confronti delle persone (i consumatori, gli utenti della rete) che sono costantemente sottoposti ad abusi da parte di società che si

collocano in una posizione dominante è dato dal Regolamento Generale sulla Protezione dei Dati (GDPR).

Il GDPR, frutto di diversi anni di lavoro, è entrato in vigore il 25 Maggio 2018 ed è valido per tutti gli stati membri dell'Unione Europea, nello specifico per tutte le società che trattano o gestiscono dati e che monitorano il comportamento di individui residenti nell'Unione Europea. Gli elementi su cui si concentra tale regolamento sono molti; in seguito vengono citati i più rilevanti<sup>15</sup>:

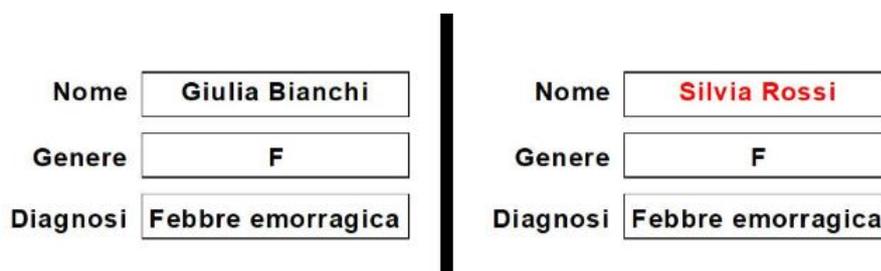
- Il consenso esplicito dell'utente. I dati personali, particolarmente sensibili sotto il profilo dei diritti e delle libertà fondamentali, non devono essere oggetto di trattamento a meno che il diretto interessato non esprima un consenso esplicito oppure si trovi in situazioni in cui si manifestano esigenze specifiche (e.g. attività che necessitano dell'esercizio delle libertà fondamentali). L'individuo deve, inoltre, possedere il diritto di non essere sottoposto a nessuna decisione né tantomeno profilazione salvo tale decisione non sia di interesse per lo Stato (o un Ente che lo rappresenta) o non vi sia, anche in questo caso, il consenso esplicito.
- Il diritto alla cancellazione del dato. Tale diritto, anche chiamato "diritto all'oblio", non deve venir pregiudicato da altri diritti e l'individuo interessato deve poter, in qualsiasi momento, chiedere al titolare del trattamento dei dati che essi vengano cancellati, che il suddetto trattamento venga limitato oppure può opporsi al trattamento. Il titolare del trattamento ha l'obbligo, quindi, cancellare il dato senza ingiustificati ritardi ed è anche sottoposto all'obbligo di notifica, i.e. deve comunicare a ciascun destinatario cui sono stati trasmessi i dati la cancellazione e può comunicare al diretto interessato, se richiesto, i destinatari dei propri dati.
- La riduzione al minimo dell'asimmetria informativa nei confronti dell'utente. L'utente infatti deve essere chiaramente informato sulla dimensione e su quali dati le imprese collezionano. In particolare, l'utente deve essere consapevole

---

<sup>15</sup> Tratto da: "*The Economics of Data, Privacy, and (a bit of) Machine Learning*". Slide del corso di Economia e Management dei Servizi a cura del Professor Lorien Sabatino. Politecnico di Torino. A.A. 2019-2020.

delle altre finalità di impiego dei dati e dei suoi diritti (compreso il diritto di opporsi).

- Portabilità dei dati. Per rafforzare il controllo sui dati è opportuno anche che l'interessato abbia il diritto di ricevere in un formato strutturato, di uso comune, leggibile da dispositivo automatico (e.g. smartphone) e interoperabile, i dati personali che lo riguardano. È fondamentale che i dati vengano trattati con formati interoperabili che ne consentano la portabilità. Come analizzato anche in precedenza, i dati devono poter essere trasmessi senza impedimenti al diretto interessato o ad un altro titolare da egli indicato.
- Pseudonimizzazione del dato. I dati personali, riconducibili ad una persona identificata o identificabile, devono essere protetti e uno dei meccanismi è la pseudonimizzazione. È bene non confondere un dato pseudonimato con un dato anonimizzato. Nel primo caso, con molte difficoltà si può risalire al proprietario originale del dato poiché ciò che avviene è un tentativo di diffondere un'informazione quanto più completa possibile; nel secondo caso invece è impossibile risalire all'individuo di partenza. Per una migliore comprensione del concetto di pseudonimizzazione si propone un esempio inerente una schermata applicativa vista a scopo sanitario (Fig. 18).



*Figura 18 - Pseudonimizzazione del dato (destra) rispetto al dato di partenza (sinistra) – fonte: [www.cybersecurity360.it/legal/privacy-dati-personali/pseudonimizzazione-e-anonimizzazione-dei-dati-differenze-tecniche-e-applicative](http://www.cybersecurity360.it/legal/privacy-dati-personali/pseudonimizzazione-e-anonimizzazione-dei-dati-differenze-tecniche-e-applicative).*

Con l'introduzione delle regole imposte dal GDPR molte imprese hanno dovuto quindi rivedere le proprie policy e adeguarsi alle richieste del regolamento. In particolare è stato imposto l'obbligo di notificare alle autorità competenti e ai diretti

interessati, entro 72 ore dalla presa d'atto, le violazioni di dati che possono mettere a rischio i diritti e le libertà degli individui.

Il grande pregio del GDPR è stato quello di aver trasformato in qualcosa di concreto, in legge, un concetto che fino a qualche anno fa era solo considerato come “buon senso” o “pratica di uso comune”. Quando si progetta o si ripensa un prodotto o un servizio è fondamentale tenere già in considerazione la nuova regolamentazione europea.

Lo studio di Aridor, Che, Salz (2020) mette in pratica i concetti analizzati fino ad ora. Attraverso una verifica empirica è possibile analizzare l'effetto dell'introduzione del GDPR sui consensi espliciti che i siti richiedono ai consumatori, ottenuti mediante *opt-in*<sup>16</sup>, per l'uso dei dati. L'analisi è un'analisi di regressione che si focalizza sulle politiche complessive piuttosto che su specifici effetti e viene realizzata attraverso il metodo *difference-in-difference* utilizzando un dataset di imprese quali intermediari appartenenti al settore dell'*online travel*. È possibile, quindi, identificare i seguenti fattori chiave per lo sviluppo del modello diff-in-diff.

- Shock esogeno: introduzione in tutti gli stati dell'Unione Europea del GDPR il 25 Maggio 2018 che corrisponde alla settimana 22 dell'anno.
- Gruppo di trattamento: siti di viaggio e piattaforme dei più importanti paesi europei (Italia, Regno Unito, Francia, Germania e Spagna).
- Gruppo di controllo: siti di viaggio e piattaforme negli Stati Uniti, Canada e Russia. Questi paesi, infatti, sono stati scelti come gruppo di controllo poiché non sono influenzati direttamente dal GDPR ma la stagionalità dei viaggi è molto simile a quella europea (simile clima, simili festività durante l'anno).

La regressione si presenta, invece, con la seguente forma:

$$y_{tcjobp} = \alpha_{(t)} + \delta_{jc} + \kappa_c + \varepsilon_j + \gamma_o + \zeta_b + \omega_p + \beta(EU_j \times after) + \epsilon_{tcjobp}$$

---

<sup>16</sup> Si definisce *opt-in* la comunicazione pubblicitaria e commerciale inviata soltanto a chi abbia preventivamente manifestato il proprio consenso a riceverla. Il metodo *opt-in* si contrappone alla comunicazione spam che, invece, è inviata ai destinatari senza alcun consenso preventivo. Fonte: [www.okpedia.it/opt-in](http://www.okpedia.it/opt-in).

$t =$  settimana dell'anno;  $c =$  Stato;  $j =$  sito web;  $o =$  sistema operativo;  $b =$  browser web;  $p =$  tipo di prodotto (volo o hotel)

Il coefficiente che cattura l'effetto dell'introduzione del GDPR è  $\beta$ . Esso infatti è associato alla variabile  $EU_j$ , una dummy che denota quando un sito web è soggetto alla regolazione, e alla variabile  $after$ , una dummy che denota se la settimana in cui si compie l'analisi è dopo l'introduzione del GDPR. Gli altri elementi presenti nella regressione corrispondono alle variabili che catturano gli effetti fissi dovuti al tempo (i.e. la settimana), alle caratteristiche specifiche dei diversi paesi, dei siti web e dei siti web associati ai paesi, ai diversi sistemi operativi e browser web.

Mediante la regressione si possono condurre tre analisi differenti. La prima, il cui *outcome* è rappresentato nella Tabella 2, misura come la possibilità di fare *opt-out* introdotta dal GDPR, i.e. negare l'accesso alla raccolta dei cookies, impatta sul numero di cookies e ricerche osservate dagli intermediari. Dalla regressione log-lin delle colonne (1) e (3) si evince che è stato fatto *opt-out* per il 12.5% dei cookies e le ricerche registrate vedono un decremento del 10.7%; è, inoltre, importante evidenziare che non si parla di consumatori poiché un consumatore può apparire sotto multipli cookies e comportarsi diversamente a seconda dei siti web. La regressione delle colonne (2) e (4) mette in rilievo gli stessi risultati ma dal punto di vista numerico e non percentuale.

Tabella 2- Primo outcome della regressione – fonte: Aridor, Che, Salz (2020)

	(1)	(2)	(3)	(4)
	log(Unique Cookies)	Unique Cookies	log(Recorded Searches)	Recorded Searches
DiD Coefficient	-0.125** (-2.43)	-1378.1* (-1.71)	-0.107* (-1.87)	-9618.3** (-2.24)
Product Type Controls	✓	✓	✓	✓
OS + Browser Controls	✓	✓	✓	✓
Week FE	✓	✓	✓	✓
Website × Country FE	✓	✓	✓	✓
Observations	63840	63840	63840	63840

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Notes: t-statistics are reported in parentheses. The standard errors for every regression are clustered at the website-country level. We aggregate every dependent variable to the website-country-product type-week level between weeks 16 and 29, including both weeks 16 and 29 (April 13th - July 20th). The dependent variables in the regression reported in the first and second column are the log and overall level of the number of unique cookies observed. The dependent variables in the regression reported in the third and fourth column are the log and overall level of the number of total recorded searches.

La seconda analisi misura, invece, la dimensione con la quale il consumatore che fa *opt-out* e l'aumento della tracciabilità influenzano il valore dei consumatori per gli *advertisers* e i ricavi complessivi per *advertisers* e siti web (Tab. 3).

Tabella 3 - Secondo outcome della regressione – fonte: Aridor, Che, Salz (2020)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	asinh(Total Clicks)	Total Clicks	asinh(Distinct Clicks)	Distinct Clicks	asinh(Revenue)	Revenue	Average Transfer	Average Bid
DiD Coefficient	-0.135** (-2.32)	-251.9* (-1.91)	-0.133** (-2.33)	-214.9* (-1.84)	-0.168 (-1.54)	-32972.3 (-0.75)	28.97** (2.12)	15.41*** (2.90)
OS + Browser Controls	✓	✓	✓	✓	✓	✓	✓	✓
Product Category Controls	✓	✓	✓	✓	✓	✓	✓	✓
Website × Country FE	✓	✓	✓	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	62328	62328	62328	62328	62328	62328	62328	62328

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Notes: t-statistics are reported in parentheses. The standard errors for every regression are clustered at the website-country level. We aggregate every dependent variable to the website-country-product type-week level between weeks 16 and 29, including both weeks 16 and 29 (April 13th - July 20th). The dependent variable in the regression reported in the first column is the total number of clicks associated with each observation and the second column is the inverse hyperbolic sine transform of this value. Likewise, the dependent variables in the third and fourth columns are the total number and inverse hyperbolic sine transform of the total number of unique cookies who interacted with advertisements. The dependent variables in the fifth and sixth column are the total number and inverse hyperbolic sine transform of the total revenue. The dependent variable in the seventh column is the average transfer between the intermediary and advertisers and in the eighth column it is the average bid by advertisers. Since some of the outcome variables can take zero values, in order to preserve these observations we utilize a common transformation in the applied microeconomics literature and use the inverse hyperbolic sine transform instead of the natural logarithm of the outcome variables (Bellemare and Wichman, 2019). The resulting transformed outcome variable,  $\bar{y}$ , is given by  $\bar{y} = \text{arcsinh}(y) = \ln(y + \sqrt{y^2 + 1})$

Dalle colonne (1) e (2) si nota un decremento statisticamente significativo del 13.5% nel numero totale dei click; le colonne (3) e (4) mostrano che anche il numero di click distinti è diminuito; le colonne (5) e (6) forniscono la stima dell'impatto sui ricavi, la quale è negativa ma non statisticamente significativa; le colonne (7) e (8) mostrano che i trasferimenti medi di dati e le offerte da *advertisers* a intermediari aumentano e ciò si percepisce come un incremento del valore che gli *advertisers* attribuiscono ai consumatori.

La terza analisi si focalizza sul verificare le performance delle predizioni fatte con l'introduzione del GDPR e, quindi, si studia se lo shock esogeno ha influenzato la capacità dell'intermediario di prevedere il comportamento dei consumatori. Dai risultati mostrati in Tabella 4, complessivamente si può affermare che l'introduzione del GDPR non ha influenzato negativamente l'abilità di predizione degli intermediari ed esiste qualche evidenza che, invece, tale influenza sia positiva.

Tabella 4 - Terzo outcome della regressione – fonte: Aridor, Che, Salz (2020)

	(1) Class Proportion	(2) Average Predicted Probability	(3) MSE	(4) AUC	(5) Purchaser MSE	(6) Non-Purchaser MSE
DiD Coefficient	0.00915* (1.77)	0.00129 (0.17)	0.0130*** (3.74)	0.0124 (1.12)	-0.00579 (-0.43)	-0.00126 (-0.45)
Product Type Controls	✓	✓	✓	✓	✓	✓
OS + Browser Controls	✓	✓	✓	✓	✓	✓
Week FE	✓	✓	✓	✓	✓	✓
Website × Country FE	✓	✓	✓	✓	✓	✓
Observations	15470	15470	15470	15470	14298	15470

\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

Notes: t-statistics are reported in parentheses. The standard errors for every regression are clustered at the website-country level. We aggregate every dependent variable to the website-country-product type-week level between weeks 16 and 29, including both weeks 16 and 29 (April 13th - July 20th). The dependent variable in the regression reported in the first column is the proportion of purchasers associated with each observation and the second column is the average predicted probability. The dependent variables in the third and fourth column are the MSE and AUC, respectively. Finally, in the fifth and sixth columns the dependent variables are the MSE conditional on the true class of the observation.

Dallo studio di tale analisi di regressione, si può concludere che le decisioni individuali dei consumatori, volte alla protezione della privacy, possono impattare su tutta l'economia, incluso gli altri consumatori, le imprese e gli *advertisers* che basano il proprio business sui dati degli utenti.

### 2.6.5. La gestione da parte dell'Antitrust

La regolazione da parte dell'Antitrust esiste per prevenire un eccessivo potere centralizzato all'interno di un settore. I regolatori e le Commissioni dei vari stati o regioni possono intervenire per promuovere la competizione sanzionando le imprese che si comportano in modo anti-competitivo, spezzettando grandi compagnie o prevenendo situazioni di M&A potenzialmente dannose. È stato dimostrato che lasciare alle imprese il compito di autocontrollarsi non è un'alternativa funzionante. Sono quindi necessarie regole Antitrust che impediscano il controllo privato dei big data, che obblighino le piattaforme ad essere interoperabili, che difendano la neutralità della rete. È altresì certo è che le leggi Antitrust non saranno mai sufficienti fintanto che non esistono alternative a ciò che offrono le piattaforme dominanti (De Blase, 2019).

Un esempio molto rappresentativo di come l'Antitrust intervenga a tutela dei consumatori con le proprie politiche di policy è dato dai recenti avvenimenti del Luglio 2020 che vedono coinvolte le quattro *over-the-top* Amazon, Apple, Google e Facebook<sup>17</sup>.

Tali imprese, infatti, rientrano all'interno dell'indagine aperta dall'Antitrust americano sul potere di mercato e sul soffocamento della concorrenza. La Commissione ha ritenuto che ciascuna delle suddette piattaforme costituisca un problema per la concorrenza poiché esse rappresentano un “collo di bottiglia” per un canale chiave della distribuzione, impiegano il loro controllo sull'infrastruttura digitale per sorvegliare altre aziende e sfruttare le informazioni che ottengono per consolidare il potere di mercato, abusano del controllo sulla tecnologia per dare ai loro prodotti un vantaggio su quelli dei concorrenti che utilizzano la stessa piattaforma. Con riferimento alle specifiche imprese, i comportamenti giudicati anti-competitivi sono stati i seguenti:

- Amazon mette in piedi strategie di *killer acquisition*<sup>18</sup> per eliminare dei competitor dal mercato. Alcuni dipendenti di Amazon, invece, hanno preso possesso dei dati dei competitor per entrare sul mercato con prodotti concorrenti.
- Apple, come visto nel Paragrafo 2.3.1., applica eccessive commissioni agli sviluppatori che vendono le loro app sull'App Store e rimuove o limita fortemente le app simili a quelle sviluppate dalla piattaforma (e.g. Screen Time).
- Facebook è accusato di acquisire i concorrenti piuttosto che competervi. L'acquisizione di Instagram, infatti, è stata vista dal fondatore del nuovo social

---

<sup>17</sup> [www.dday.it/redazione/36400/amazon-apple-google-facebook-antitrust-usa](http://www.dday.it/redazione/36400/amazon-apple-google-facebook-antitrust-usa).  
[www.repubblica.it/tecnologia/2020/07/29/news/l\\_antitrust\\_americano\\_mette\\_sotto\\_accusa\\_amazon\\_apple\\_facebook\\_google-263198696](http://www.repubblica.it/tecnologia/2020/07/29/news/l_antitrust_americano_mette_sotto_accusa_amazon_apple_facebook_google-263198696).

<sup>18</sup> Tali strategie consistono nell'acquistare un competitor per l'impresa ed entrare nello stesso business, tagliare i prezzi per lo stesso prodotto/servizio per incrementare la competizione nel tentativo di far fallire l'impresa acquisita e, una volta riusciti nell'intento avendo eliminato il competitor, rialzare i prezzi poiché si gode di maggior *market share* e presumibilmente di quasi-monopolio.

network come un passaggio obbligato poiché era forte il timore che, in caso di non cessione, Facebook sarebbe entrata in “modalità distruttrice”.

- Google attacca duramente i motori di ricerca verticale suoi concorrenti, minacciando il *delisting* nel caso in cui essi non ospitino le recensioni della piattaforma in esame. Inoltre, Google sorveglia gli altri siti web per formulare le proprie strategie.

Come si evince da Garofalo (2019), facendo particolare riferimento alle piattaforme Facebook e Google, i comportamenti anti-competitivi riguardano maggiormente il settore dell'*online advertising*. Secondo l'autore dell'articolo non sarebbe più presente un libero mercato e una libera competizione e ciò causa un aumento dei prezzi con conseguente danno per i consumatori. Anche in Europa sono aperte indagini su Google volte ad accertare quali siano i confini di mercato dell'impresa e se la stessa detenga o meno una posizione dominante nel campo della pubblicità online. Decisione più recente, datata fine Luglio 2020, è quella presa dall'Australian Competition & Consumer Commission secondo cui Google e Facebook dovrebbero pagare i media tradizionali attraverso cui le *big-firms* pubblicano i loro annunci; si tratta, quindi, di una condivisione di proventi che però le due piattaforme, per ora, non hanno intenzione di appoggiare (Boccellato, 2020).

In merito all'indagine aperta contro le quattro *over-the-top* della Silicon Valley, la Commissione Antitrust americana ha, quindi, ritenuto le piattaforme responsabili di comportamenti monopolistici e tra le possibili soluzioni che si stanno valutando rientrano una più adeguata regolamentazione del mercato digitale ed eventualmente la suddivisione delle compagnie sul modello AT&T.

#### **2.6.6. Le autorità regolatorie in Europa**

Come conseguenza del crescente interesse verso il tema della privacy all'interno del *data market*, i diversi paesi hanno cercato di adeguarsi e di rispondere alle esigenze dei consumatori che richiedevano maggior tutela. Di seguito vengono analizzati le principali decisioni prese a livello europeo, con particolare focus alla nuova figura del regolatore del *data market*.

L'Europa crede fortemente di poter diventare un modello da seguire per le società costituite sui dati, aiutandole a prendere migliori decisioni di business e in ambito pubblico. Tale obiettivo, raggiungibile non senza sforzi, ha conseguenze sul rafforzamento della *governance*, sul miglioramento della salute e del benessere dell'ambiente nonché sulla rivalutazione della qualità dei servizi pubblici. L'organo che si occupa di delineare gli step da perseguire è la Commissione Europea la quale programmerà interventi di policy e strategie di investimento per far progredire la *data economy* per i prossimi cinque anni. Nello specifico, l'UE intende creare un unico spazio europeo per i dati, dove vengono custoditi i dati personali e non-personali e dove si può avere accesso ad un infinito ammontare di informazioni di alta qualità appartenenti alle principali realtà industriali, promuovendo la crescita economica e generando valore (European Commission, 2020).

Focalizzandosi a livello di singolo paese, in Italia autorità quali Antitrust, Agcom e Autorità Garante per la Privacy hanno elaborato congiuntamente delle linee guida presentate in “*Big Data. Indagine conoscitiva congiunta. Linee guida e raccomandazioni di policy*” (Luglio 2019) in cui pongono l'attenzione sui diversi temi trattati dai vari stati europei all'interno dei propri report o *white paper* ma, in particolar modo, ritengono fondamentale la cooperazione tra le tre autorità che potrebbe culminare con un coordinamento permanente. Inoltre, il report evidenzia la necessità di rafforzamento del potere istruttorio già prima che si apra un procedimento e di aumento delle sanzioni.

Anche il Regno Unito, nonostante presenti già alti standard di *law enforcement* (i.e. GDPR e Data Protection Act 2018), ha sottolineato quanto l'intervento dell'Antitrust sia tardivo: sarebbe, infatti, opportuno istituire un'Autorità Regolatrice del mercato digitale che abbia poteri generali di coordinamento nei confronti delle singole istituzioni di regolazione, e.g. privacy, antitrust, comunicazioni elettroniche, poiché è evidente che vi siano in gioco numerosi e diversi temi. (Authority of the House of Lords, 2019).

Attraverso il *white paper* redatto dal HR Governments (2019) si possono delineare, invece, una serie di azioni che il Governo inglese deve mettere in pratica al fine di tutelare gli individui e di determinare responsabilità e diritti dei soggetti che

usufruiscono di Internet. Il focus è volto ai danni che subiscono i bambini e gli adolescenti causati da abusi anonimi, cyberbullismo, disinformazione online, manipolazione e altri problemi di natura relazionale/tecnologica (e.g. tempo speso dietro ad un computer o un telefono). Il report ha quindi lo scopo di delineare una serie di linee guida, *duty of care*, che devono essere necessariamente rispettate dalle piattaforme per mantenere gli utenti della rete al sicuro e per tracciare eventuali comportamenti dannosi inerenti i loro servizi. Tale dovere di diligenza viene supervisionato da un'Autorità Regolatrice Indipendente, alla quale le imprese dovranno dei chiarimenti in caso di non rispetto delle regole. Il regolatore giudicherà l'operato delle imprese con il criterio della proporzionalità (nel rispetto di dimensione di impresa, età degli utenti, risorse, etc.) e renderà chiaro alle imprese con che mezzi esse possono perseguire i loro obiettivi di profitto in accordo con le nuove disposizioni. Inoltre, il regolatore farà completo affidamento sull'approccio *risk-based*, dando priorità all'analisi delle piattaforme i cui danni hanno un impatto maggiore sulla società, in termini di diffusione e tipologia di servizi. A tutto ciò, seguiranno delle verifiche empiriche in accordo con alcune compagnie. È evidente che i concetti di *monitoring* generale, trasparenza, fiducia e responsabilità ricoprono grande importanza. Altri due attori partecipano al progetto di tutela: il governo inglese, il quale ha il ruolo di supervisionare l'Autorità Regolatoria e la tecnologia. Le compagnie, infatti, dovrebbero investire nello sviluppo di tecnologie sicure che permettano di aumentare la sicurezza online (e.g. tra le iniziative già esistenti si citano SuperAwesome, Crisp, Yoti e applicazioni come LegoLife).

In conclusione, ciò a cui guardano i paesi europei è la miglior gestione della sicurezza online degli utenti. I governi, attualmente, sono impegnati nel supportare l'alfabetizzazione digitale e le piattaforme digitali appoggiano tali iniziative positive ma, purtroppo, è presente una mancanza di trasparenza su più fronti che impedisce al mercato di operare con uno spirito maggiormente competitivo e ai consumatori di sentirsi protetti e tutelati durante le transazioni.

## CAPITOLO III

### ESTENSIONE DEL MODELLO “SELLING COOKIES”

L'intenzione del presente capitolo è quella di proporre un'ulteriore estensione del modello di *data provision* e *data pricing* “Selling Cookies” pubblicato da Bergemann e Bonatti (2015) e descritto nel paragrafo 2.2.1.

A tal proposito, si ritiene fondamentale sottolineare che tale modello, così come presentato dagli autori, è frutto di numerose semplificazioni. In primo luogo, bisogna considerare che il potere di impostare i prezzi per le informazioni e la natura stessa delle informazioni vendute dipendono dall'allocazione iniziale dei dati tra i diversi attori del *data market*. Fino ad ora si è analizzato l'andamento del prezzo per l'acquisto dei cookies imposto dal *data provider* agli *advertisers* in diversi ambienti competitivi, senza prendere in considerazione il costo di acquisizione delle informazioni dai singoli individui, dai *publisher* o dagli *advertisers*.

Lo scopo della presente analisi, invece, è proprio di mettere in evidenza come cambia il prezzo per i cookies quando il *data seller* è soggetto ad un costo di acquisizione delle informazioni. Tale costo corrisponde al valore della privacy che possiede ciascun proprietario dei dati, in particolar modo il consumatore individuale.

Come già sottolineato in precedenza, in caso di informazione completa, le imprese possono raggiungere l'ottimo nelle strategie di acquisto delle informazioni e nella customizzazione della spesa per gli *advertisers*, sia con *joint targeting* che con *exclusive targeting*. Inoltre, il raggiungimento di un prezzo lineare per i cookies dipende dall'abilità di influenzare la composizione del *targeted set* e del *residual set* di consumatori: il prezzo per le informazioni diminuisce quando si raggiunge il database e si hanno vendite concentrate; al contrario, esso aumenta se la vendita dei dati diventa frammentata.

L'estensione del modello verrà proposta analizzando, teoricamente e mediante un'applicazione ad un caso reale, due scenari diversi, l'ambiente binario e l'ambiente continuo, assumendo però sempre la presenza di un solo *data provider* monopolista e non di diversi *data providers* in concorrenza l'uno con l'altro.

### 3.1. Il costo della privacy

Lo studio di Bergemann e Bonatti si prefigge come obiettivo quello di mostrare i cambiamenti nell'andamento del prezzo per i cookies in base al contesto in cui avviene lo scambio.

Il *data seller*, e in alcuni casi anche il *publisher*, acquisisce i dati dei consumatori e li rivende agli *advertisers*. La novità apportata con il presente capitolo riguarda l'introduzione del costo della privacy  $\tau$ . Ciascun consumatore individuale o soggetto che genera informazioni attribuisce un valore alle proprie preferenze, gusti personali e informazioni di carattere generale, e proprio per questo potrebbe decidere di limitare l'accesso ai propri dati. D'altro canto il *data provider*, che necessita di queste informazioni per eseguire transazioni con gli *advertisers* i quali, a loro volta, desiderano adottare la migliore strategia di advertising, si trova quindi a dover sostenere un costo per l'acquisizione dei dati che tiene conto del diritto alla privacy del consumatore. Tale costo, in linea generale, si riflette nell'aumento del prezzo per i cookies e viene sostenuto dal *data provider* e, successivamente, ribaltato agli *advertisers*.

Come anche nel modello nella sua accezione base, il *data seller* monopolista non è in grado di influenzare qualitativamente la composizione del *targeted set* attraverso la definizione del prezzo unitario di acquisto dei cookies  $p$ . Egli infatti, prende per data la composizione del set di consumatori e sceglie il prezzo che massimizza i propri ricavi secondo l'equazione (8). Nel caso in esame si estende l'equazione (8) e il prezzo per i cookies determinato dal *data provider* è pari a

$$\begin{aligned} p^N &= p^* + \tau = \\ &= \arg \max [p \cdot \mu(A(c, p, \tau))] + \tau = \arg \max [p \cdot v] + \tau . \end{aligned} \quad (15)$$

Gli *advertisers* si trovano quindi a pagare un prezzo per i cookies del *targeted set* pari a  $p \cdot \mu(A(c, p, \tau)) + \tau \cdot q(v)$ , dove  $\tau \cdot q(v)$  rappresenta il costo unitario della privacy per la quantità di dati acquistata.

La funzione obiettivo del modello, i.e. il profitto per l'*advertiser*, non varia e, a tal proposito, si richiama l'equazione (1):

$$\pi(v, q) \triangleq vq - c * m(q). \quad (1)$$

### 3.2. Ambiente binario

Si ipotizza ora di trovarsi in un ambiente binario in cui, come nel modello base, si hanno costi  $c$  di matching lineari, *match values* uniformemente distribuiti con  $q$  che può assumere solo i valori 0 o 1 (e.g. se  $v > c$  l'*advertiser* contatta il cliente  $q = 1$ ).

#### 3.2.1. Analisi della domanda

Partendo dall'analisi della domanda di informazioni degli *advertisers*, il valore dell'informazione acquisita dagli stessi per ogni *match value*, si ottiene partendo dall'equazione (2) e facendo la differenza tra il profitto degli stessi in caso di esercizio dell'azione  $q$  per generare consapevolezza nel consumatore e in caso di scelta di non esercizio:

$$\Delta\pi^*(v) \triangleq \pi(q = 1) - \pi(q = 0) = (v - c)q - 0 = v - c. \quad (16)$$

Per simmetria con quanto presentato nel modello base, se gli *advertisers* adottano il *positive targeting* acquisteranno le informazioni dai consumatori che lasciano quindi ai primi un'utilità netta non negativa pari a  $v^N = c + p + \tau$  (modello base  $v^* = c + p$ ). Allo stesso modo se essi adottano il *negative targeting* l'utilità degli *advertisers* sarà pari a  $v^N = c - p - \tau$  (modello base  $v^N = c - p$ ). Anche in questa situazione risulta fondamentale identificare il gruppo di consumatori a cui l'*advertiser* si rivolge: in caso di *positive targeting* la richiesta di informazioni è diretta al *targeted set*; in caso di *negative targeting* la richiesta di informazioni è diretta al gruppo di consumatori che si sceglie di escludere. In accordo con quanto definito dall'equazione (3), il valore del *targeted set* ottimo corrisponde all'intervallo di valori dato da

$$A(c, p) = \begin{cases} [0, \max\{v^N, 0\}], & c < ? \\ [\min\{v^N, 1\}, 1], & c \geq ? \end{cases} \quad (17)$$

È fondamentale sottolineare che introducendo il costo unitario per la privacy  $\tau$ , il valore soglia che delimita la scelta della strategia di targeting dell'*advertiser*, corrispondendo al costo di acquisto dello spazio per l'advertising e unico costo preso in considerazione nel modello base, non è più pari a  $c = 1/2$  poiché è necessario considerare anche l'influenza di  $\tau$  nella scelta dell'impresa. Di seguito viene illustrata la dimostrazione per giungere al nuovo valore.

Il primo passo consiste nel considerare i profitti dell'*advertiser* sotto *positive targeting* e *negative targeting*

$$\pi_+(c, p, \tau) \triangleq \max \int_{v^N}^1 (x - v^N) dF(x) = \quad (18)$$

$$= \int_{c+p+\tau}^1 (x - c - p - \tau) dF(x) \quad 19$$

$$\pi_-(c, p, \tau) \triangleq \max \int_{v^N}^1 (x - c) dF(x) - pF(v^N) - \tau q = \quad (19)$$

$$= \int_{c-p-\tau}^1 (x - c) dF(x) - pF(c - p - \tau) - \tau \quad 20$$

e, successivamente, nel fare la differenza tra gli stessi. Si sottolinea che i passaggi di calcolo sono stati omissi.

$$\begin{aligned} & \pi_+(c, p, \tau) - \pi_-(c, p, \tau) = \\ & = p[F(c - p - \tau) + F(c + p + 2\tau) - 1 + c\tau] - \int_{c-p-\tau}^{c+p+\tau} v - c - \tau \quad (20) \end{aligned}$$

Per simmetria con quanto presentato nel modello base, sotto l'ipotesi di distribuzione uniforme il secondo termine in (20) è nullo. Dal primo termine invece si ottiene il valore soglia di  $c$  che è pari a  $c^N = \frac{1-\tau}{2+\tau} < c = \frac{1}{2}$ .

---

<sup>19</sup> Con  $x$  si fa riferimento ad un qualsiasi valore di match.

<sup>20</sup> Il costo della privacy, come il prezzo  $p$ , si considera direttamente sostenuto dal data anche in caso di *negative targeting*: per decidere quale gruppo di consumatori escludere è stata comunque fatta *disclosure* dei dati.

È ora possibile scrivere il valore del *targeted set* con il nuovo valore soglia:

$$A(c, p, \tau) = \begin{cases} [0, \max\{c - p - \tau, 0\}], & c < \frac{1 - \tau}{2 + \tau} \\ [\min\{c + p + \tau, 1\}, 1], & c \geq \frac{1 - \tau}{2 + \tau} \end{cases} \quad (21)$$

dove con  $c < \frac{1 - \tau}{2 + \tau}$  l'impresa opta per il *negative targeting* mentre con  $c \geq \frac{1 - \tau}{2 + \tau}$  si sceglie il *positive targeting*. È fondamentale evidenziare che rispetto al modello base dove  $c = \frac{1}{2}$ , il valore soglia si è spostato verso il basso. Infatti, introducendo il costo della privacy si adotta il *negative targeting* quando il costo di acquisto dello spazio  $c$  è molto basso e, allo stesso modo, non è necessario che  $c$  sia sufficientemente alto per optare per una strategia di *positive targeting*, come rappresentato in Figura 19.

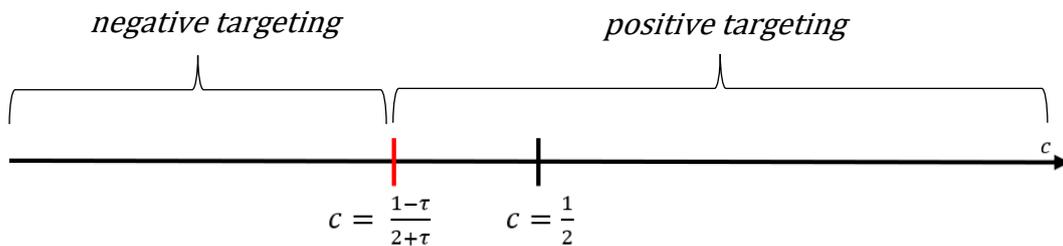


Figura 19 – Rappresentazione del valore soglia del costo dello spazio per advertising nel modello base e nel modello esteso

Per capire il risvolto economico di questo risultato è importante fare la premessa che, sia nel modello base che nel modello esteso, il gruppo di consumatori residuale è considerato più ampio del *targeted set*. A questo proposito, spesso al *targeted set* viene accostato il nome di “mercato di nicchia”. Partendo da ciò, si può quindi notare che gli *advertisers* optano per strategie di *negative targeting* (i.e. contattano il *residual set* con l'intento di esclusione) quando  $c$  è basso proprio poiché introducendo il costo della privacy  $\tau$  si trovano a dover fronteggiare maggiori spese, essendo il *residual set* di dimensione superiore rispetto al *targeted set*. Al contrario, quando  $c$  è alto, o comunque più alto del valore soglia, gli *advertisers* si dirigono verso

il *positive targeting* poiché è più profittevole contattare direttamente la nicchia dei consumatori ad alto valore.

Trattandosi di ambiente binario, anche nel modello esteso, la domanda di informazioni  $q \in [0,1]$  dipende dai valori di  $v$  e  $c$ : se  $v > c$  allora  $q = 1$ , se  $v < c$  allora  $q = 0$ . Da ciò è possibile ricavare il profitto per gli *advertisers* dato dall'equazione (2) che si ricorda essere pari a:

$$\pi^*(v) \triangleq \max \{v - c, 0\}. \quad (2)$$

### 3.2.2. Definizione del prezzo

Considerando invece il prezzo per i cookies stabilito dal *data provider*, anch'esso varia. Sviluppando quanto previsto dall'equazione (15) si ha che

$$p^N(c, \tau) = \begin{cases} \arg \max [p \cdot v^N] + \tau, & c < \frac{1 - \tau}{2 + \tau} \\ \arg \max [p \cdot (1 - v^N)] + \tau, & c \geq \frac{1 - \tau}{2 + \tau} \end{cases} \quad (22)$$

corrisponde al punto di partenza per determinare il prezzo delle informazioni. Tenendo presente che sia con *positive targeting* che con *negative targeting* si acquistano informazioni (dal *targeted set* e dal *residual set* rispettivamente) e si sostiene quindi il costo  $\tau$ , si ottiene che

$$p^N(c, \tau) = \begin{cases} \arg \max [p \cdot (c - p - \tau)] + \tau, & c < \frac{1 - \tau}{2 + \tau} \\ \arg \max [p \cdot (1 - c - p - \tau)] + \tau, & c \geq \frac{1 - \tau}{2 + \tau} \end{cases} \quad (23)$$

da cui, massimizzando il sistema si ricava il prezzo unitario per i cookies in caso di *negative targeting* e *positive targeting*

$$\frac{dp^N}{dp} = c - 2p - \tau = 0 \rightarrow p = \frac{c - \tau}{2}$$

$$p^N = p + \tau = \frac{c + \tau}{2}$$

$$\frac{dp_+^N}{dp} = 1 - c - 2p - \tau = 0 \rightarrow p = \frac{1 - c - \tau}{2}$$

$$p^N = p + \tau = \frac{1 - c + \tau}{2} \quad 21$$

$$p^N(c, \tau) = \frac{1}{2} \min \{c + \tau; 1 - c + \tau\}. \quad (24)$$

Confrontando l'equazione (24) con la (9) si può notare che  $c + \tau > c$  e  $1 - c + \tau > 1 - c$ . Da ciò segue che, con l'introduzione del costo della privacy il prezzo unitario ottimo di acquisto dei cookies proposto dal *data provider* subisce un incremento. Infatti, il *data seller* si trova a dover fronteggiare un maggior costo di acquisizione delle informazioni e questo si riflette nell'aumento del prezzo delle stesse. Gli *advertisers* vedono, infine, un decremento del loro profitto al netto del prezzo di acquisto dei cookies. Un'analisi dettagliata sul significato economico di tale prezzo  $p^N$  seguirà in conclusione al Paragrafo 3.3.2.

### 3.3. Ambiente continuo

La seconda parte dell'approfondimento del modello di *data provision* e *data pricing* si focalizza sull'estensione riguardante la presenza del costo della privacy  $\tau$  in un ambiente continuo, sempre considerando l'esistenza di un *data seller* monopolista. Come visto anche nel modello base, all'interno dell'ambiente continuo  $q$  può assumere qualsiasi valore, la funzione di costo  $m(q)$  è strettamente crescente e convessa e, inoltre, si assume che  $m'(0) = 0$ .

#### 3.3.1. Analisi della domanda

Anche nel caso di ambiente continuo la funzione di profitto per gli *advertisers* è pari all'equazione (1). Per simmetria con quanto visto nel modello base, è fondamentale sottolineare la presenza di due tipi di domanda: la domanda sotto

---

<sup>21</sup> Si nota che il termine  $\tau$  non rientra nell'equazione che viene massimizzata. Tale termine infatti viene sommato dopo aver trovato il prezzo ottimo "parziale", come previsto dall'equazione (15).

informazione completa e la domanda sotto informazioni preliminari. Nello specifico, con informazioni preliminari si intendono le informazioni alle quali ha accesso l'*advertiser* quando conosce solamente  $F(v)$ , i.e. la distribuzione dei *match values* del *targeted set*, ma non possiede informazioni sul valore dei consumatori al di fuori di esso.

Ricordando l'equazione (6) che presenta la domanda di informazioni per gli *advertisers* in caso di informazione completa

$$q^*(v) = \arg \max [\pi(v, q)], \quad \text{con } q \in R_+ \quad (6)$$

e applicandola al caso in esame derivando la funzione di profitto, si ottiene

$$\frac{d\pi(v, q)}{dq} = v - c \frac{dm(q)}{dq} = 0 \quad (25)$$

da cui si ricava il *match value*  $v$  che determina la domanda di spazio per advertising in caso di informazione completa:

$$v = c \cdot m'(q^*(v)) . \quad (26)$$

Allo stesso modo, la domanda di informazioni in caso di informazioni preliminari risulta essere pari a

$$\bar{q} \triangleq q^*(A^c = V) = q^*(E[v]) \quad (27)$$

e si può ricavare, ragionando in parallelo, il *match value*  $v$  che determina la domanda di informazioni nel *residual set*:

$$E[v|v \in A^c] = c \cdot m'(q^*(v)) . \quad (28)$$

Come mostrato in precedenza, in primo luogo il *data seller* fissa un prezzo per i cookies e in base a questo gli *advertisers* stabiliscono la loro disponibilità a pagare per le informazioni. Ricordando che il prezzo per i cookies possiede la componente

aggiuntiva riguardante il costo della privacy dei consumatori, il valore delle informazioni per ogni *match value* è rappresentato da

$$\Delta\pi^*(v) \triangleq \pi^*(v) - \pi(v, \bar{q}) = p + \tau = p^N \quad (29)$$

che si identifica anche con la domanda inversa del *data seller*, con  $p$  e  $\tau$  valori unitari.

Dalla condizione espressa con (29) l'*advertiser* è in grado di scegliere la dimensione del *targeted set* e del *residual set*. In precedenza, si è discusso a più riprese dell'importanza della definizione dei valori soglia che delimitano l'appartenenza dei consumatori ad un set predefinito. Anche in questo caso quindi rimangono valide le considerazioni espresse. Nello specifico, è fondamentale comprendere come determinare i valori  $v_1$  e  $v_2$  che delimitano il *targeted set*  $A = [\underline{v}, v_1] \cup [v_2, \bar{v}]$  e il *residual set*  $A^C = [v_1, v_2]$ , tenendo conto che  $v_1$  e  $v_2$  dipendono da  $c, p$  e  $\tau$ ; inoltre, si considera quindi implicito il fatto che il *residual set* sia un insieme non vuoto.

La scelta di  $v_1$  e  $v_2$ , che viene associata al problema di acquisizione di informazioni dell'*advertiser* è data da

$$\begin{aligned} \max \int_{v_1}^{v_2} [\pi(v, q^*([v_1, v_2])) - \pi^*(v) + p + \tau] dF(v) \\ \text{s.t. } c \cdot m'(q^*([v_1, v_2])) = E[v|v \in [v_1, v_2]]. \end{aligned} \quad (30)$$

Da (30) si può ricavare che, se i confini del *residual set* vengono ampliati,  $p + \tau$  rappresenta il beneficio marginale per l'*advertiser* mentre la differenza tra i profitti viene associata ad un costo marginale. Questo accade poiché allargando  $A^C$ , il *targeted set* invece diventa più piccolo (o quantomeno non varia) e le imprese che si dirigono verso di esso mediante advertising si trovano quindi a poter contattare un numero inferiore di consumatori (nel caso di *positive targeting*) o consumatori a basso valore (nel caso di *negative targeting*), risparmiando  $p + \tau$  che avrebbero speso se avessero contattato un insieme più grande di consumatori. Per lo stesso motivo, si riduce anche il profitto dell'*advertiser*, i.e.  $\Delta\pi$  si identifica con un costo marginale.

Dopo aver definito la dimensione del *targeted set* e *residual set* è necessario determinare la strategia ottima dell'*advertiser* che, tenendo conto della curvatura della funzione di costo di match, può optare per

- *joint targeting*, i.e. decidere di adottare congiuntamente *positive targeting* e *negative targeting*;
- *exclusive targeting*, i.e. decidere di adottare solo *positive targeting* o *negative targeting*.

I risultati di tale analisi sono perfettamente in linea con quanto visto nel modello base all'interno dell'ambiente che però prende in considerazione i costi di matching quadratici, i.e.  $m(q) = \frac{q^b}{b}$  con  $b > 0$ .

### 3.3.2. Definizione del prezzo

Procedendo in parallelo con quanto visto all'interno dell'ambiente binario esteso, anche in questo caso il *data provider* stabilisce il prezzo per i cookies in base all'equazione

$$\begin{aligned} p^N &= p^* + \tau = \\ &= \arg \max [p \cdot \mu(A(c, p, \tau))] + \tau = \arg \max [p \cdot v] + \tau. \end{aligned} \quad (15)$$

In primo luogo, risulta necessario distinguere il prezzo  $p^N$  in base alle strategie di targeting dell'*advertiser*, influenzate dalla dimensione del set verso cui l'impresa sceglie di indirizzarsi. All'interno dell'ambiente binario tale prezzo è stato presentato mediante sistema di equazioni poiché la scelta tra *positive targeting* o *negative targeting* dipende dal valore del costo per l'acquisto dello spazio per advertising  $c$ , che gioca un ruolo chiave nel definire la dimensione del *targeted set* e *residual set* (si veda per completezza l'equazione (21)). Trattandosi ora di un contesto continuo invece, i drivers del prezzo per i cookies sono i valori  $v_1$  e  $v_2$ . Per questo motivo, è possibile esplicitare l'equazione (15) come

$$p^N(v, q, \tau) = \arg \max [p \cdot F(v)] + \tau \quad (31)$$

$$p_+^N(v, q, \tau) = \arg \max [p \cdot (1 - F(v))] + \tau. \quad (32)$$

Massimizzando i prezzi presentati con le equazioni (31) e (32) si ottiene:

$$\frac{dp_-^N}{dp} = F(v) \quad (33)$$

$$\frac{dp_+^N}{dp} = 1 - F(v) \quad (34)$$

Ricordando che  $F(v) = \mu(A(c, p, \tau))$  e ponendo le derivate uguali a zero si può esplicitare il prezzo unitario  $p$ , il quale inserito all'interno dell'equazione (15) dà come risultato il prezzo per i cookies  $p^N$ .

Tale prezzo può essere studiato solo da un punto di vista qualitativo poiché per svolgere analisi quantitative è necessario conoscere l'utilità netta degli *advertisers* e il *match value*, il quale, a differenza del caso binario, non è più pari a  $F(v) = v$  con  $v \in (0,1)$ .

Come anche all'interno dell'ambiente binario, è fondamentale attribuire al costo  $\tau$ , il quale indubbiamente fa parte di  $p^N$ , la corretta interpretazione a seconda delle caratteristiche del *data provider*. Quando colui che acquisisce le informazioni degli utenti è un ente privato, un'impresa, una piattaforma che non possiede alcun legame con lo Stato, sorge spontaneo chiedersi se esiste un motivo per il quale egli debba pagare un prezzo maggiore per l'acquisizione di tali informazioni. È noto che, in rapporto ai singoli consumatori, un qualsiasi *data provider* possieda un potere contrattuale tale da evitare di sostenere costi aggiuntivi per i cookies quali l'aggiunta di  $\tau$ . Ciò nonostante, i consumatori riescono ad attribuire maggior valore alle loro preferenze e alle informazioni private grazie a tutte le iniziative di policy e normative che stanno iniziando a prender forma all'interno del *data market* e che consentono maggiore monitoraggio sulle attività di scambio dei dati. Un esempio è dato dall'introduzione del GDPR. Il rispetto del diritto alla privacy del consumatore si traduce quindi in un incremento di sforzo per il *data provider*

nell'approvvigionamento dei dati, proprio per questioni tecnicamente legate alle restrizioni e accortezze che vengono imposte nel momento dello scambio delle informazioni. Sotto questo punto di vista  $\tau$  rappresenta quindi il costo dello sforzo per l'acquisizione dei dati che un *data provider* deve sostenere, chiamato generalmente costo per la privacy del consumatore.

Similarmente, se colui che acquisisce le informazioni è un ente pubblico, e.g. lo Stato, la protezione delle informazioni dei consumatori rientra sotto dei precisi doveri. È quindi di interesse pubblico tutelare il consumatore dalle esternalità di rete negative che si generano all'interno del mercato, poiché, come già analizzato nei precedenti capitoli, gli utenti possono divenire oggetto di comportamenti discriminatori o di abusi di posizione dominante (e.g. abusi di sfruttamento). Il costo  $\tau$  a carico del *data provider* pubblico, riflette il costo delle esternalità negative a cui i consumatori sono soggetti e viene instaurato con l'interno di controbilanciare tali effetti negativi creando un beneficio per la collettività.

Indipendentemente dalla natura del *data provider* (pubblico o privato) non è quindi possibile sottrarsi dal sostenere il costo per la privacy del consumatore  $\tau$  e il prezzo per i cookies, in ogni caso, subisce un incremento. Tale incremento di prezzo viene, in primo luogo, addebitato al *data provider* e poi ribaltato agli *advertisers* che, come già presentato all'interno del contesto binario, anche in questo caso si vedono costretti a pagare un prezzo maggiore per le informazioni dei consumatori.

### **3.4. Applicazione del modello**

L'estensione del modello di *data provision* e *data pricing* può venir applicata ad un caso concreto. A questo proposito, si propone come ambiente di riferimento la nota piattaforma YouTube, che si ipotizza operare in monopolio, e se ne studiano le determinanti di prezzo per i cookies, la domanda di informazioni degli utenti e la funzione di profitto degli *advertisers* all'interno del contesto binario e continuo. Ricalcando lo schema della Figura 10, la struttura della presente applicazione, nella quale si evidenziano i rapporti tra i diversi attori, è rappresentata in Figura 20.



Figura 20 - Rappresentazione della struttura del modello applicata al caso di YouTube

In primo luogo è necessario determinare il valore dei parametri di costo  $c$  e  $\tau$ . Si ricorda che la fonte principale per guadagnare su YouTube sono le pubblicità: più visualizzazioni si hanno su un video e maggiori sono gli introiti che quest'ultimo porta al suo autore. Tali introiti vengono poi spartiti, in percentuali predeterminate, con la piattaforma. Generalmente, quando ci si riferisce ad un costo legato alle visualizzazioni, esso si può indicare come *cost per thousand impressions* (CPM). Poiché il modello in esame prende come unità di riferimento il cookie e non la visualizzazione è necessario comprendere quanti cookies vengono lasciati dall'utente quando visualizza i video all'interno della piattaforma, i.e. si comporta come uno spettatore. A questo proposito si può fare riferimento allo strumento Blacklight<sup>22</sup>, grazie al quale inserendo l'URL del sito che si intende analizzare è possibile ricavare il numero di cookies che l'utente lascia quando visita un sito. Ispezionando la piattaforma YouTube tale strumento rileva la presenza di due cookies: un cookie che si comporta da *ad tracker* (i.e. acquisisce i dati per fare *target advertising* mediante codice Javascript o piccole e invisibili immagini) e un cookie di terze parti impiegato per tracciare il profilo dell'utente, entrambi appartenenti ad Alphabet, Inc.

Il costo per acquistare lo spazio pubblicitario  $c$  viene determinato in base ai ricavi che la piattaforma si aspetta di ottenere. In media, si rileva che YouTube ricava

<sup>22</sup> Il sito preso in esame è: [themarkup.org/blacklight](http://themarkup.org/blacklight).

Tale informazione è stata estrapolata da: Luna, R. (29 Settembre 2020). *Scopri quanti ti spiano mentre visiti il tuo sito preferito*. La Repubblica. [www.repubblica.it/dossier/stazione-futuro-riccardo-luna/2020/09/29/news/scopri\\_quanti\\_ti\\_spiano\\_mentre\\_visiti\\_il\\_tuo\\_sito\\_preferito-268846248/](http://www.repubblica.it/dossier/stazione-futuro-riccardo-luna/2020/09/29/news/scopri_quanti_ti_spiano_mentre_visiti_il_tuo_sito_preferito-268846248/).

sui 7€ lordi ogni 1000 visualizzazioni pagati dagli *advertisers* per inserire le pubblicità prima dei video<sup>23</sup>. Semplificando l'ambiente, si può ipotizzare che un utente entri sul sito, guardi un video ed esca dal sito, i.e. si raccolgono i cookies ad ogni visualizzazione poiché essa è unica (un accesso = un video). Questo si rende necessario per poter continuare con lo sviluppo dell'applicazione poiché altrimenti sarebbe necessario conoscere, in media, quanti video guarda un utente quando accede alla piattaforma (un accesso = ? video); questa informazione è decisamente variabile e dipende da un numero elevato di fattori. Tenendo presente ciò, la piattaforma ottiene un ricavo pari a 0.007€/visualizzazione, e considerando una media di due cookies raccolti a visualizzazione, guadagna 0.0023€/cookie, che trasformato in CPM è pari a 2.33€ per mille cookies. La piattaforma, dunque, non può settare il costo di acquisto dello spazio pubblicitario  $c$  a meno di 0.0023€/cookie.

Per individuare il costo della privacy del consumatore  $\tau$ , invece, è necessario comprendere per quale valore l'utente è disposto a cedere parte delle sue informazioni personali. Ipotizzando la presenza di un utente attento alla privacy, si può definire  $\tau$  pari a 0.05€/cookie, i.e. 50€ per mille “pezzi” di informazione rilasciati.

Il passo successivo consiste nell'identificare gli altri parametri del modello. La *matching cost function* che rappresenta lo spazio pubblicitario  $m(q)$ , all'interno dell'ambiente binario, come visto in precedenza, si ipotizza lineare e pari a  $q$ ; all'interno dell'ambiente continuo si ipotizza quadratica e pari a  $\frac{q^2}{2}$ .

In ultimo, la funzione di distribuzione dei valori del *targeted set*  $F(v)$  si può ricavare dall'utilità netta degli *advertisers* considerando che tale modello di *data provision* e *data pricing* viene associato ad un modello di Salop e, eventualmente, anche ad un modello di Hotelling, identificando i consumatori non come singoli individui ma solo come appartenenti ai due gruppi *targeted set* e *residual set* situati agli estremi di una “città lineare”. Proprio per questo, in riferimento al modello di Hotelling, si può scrivere che

$$F(x)_+ = v - c(v - 0) - p - \tau$$

---

<sup>23</sup> Informazione tratta da: [www.aranzulla.it/come-youtube-paga-i-video-66228.html](http://www.aranzulla.it/come-youtube-paga-i-video-66228.html).

$$F(x)_- = v - c(1 - v) - p - \tau$$

dove nella posizione 0 si colloca il *targeted set* e nella posizione 1 il *residual set*.

Le variabili in esame all'interno del modello sono  $v$ , che dipende dal gruppo di consumatori che si prende in considerazione, e  $q$ , determinata come conseguenza del valore di match.

Avendo ora a disposizione tutte le informazioni, riassunte nella Tabella 5, è possibile procedere con l'applicazione numerica.

Tabella 5 - Dati del modello applicati al caso YouTube

$c$	0.0023€/cookie
$\tau$	0.05€/cookie
$m(q)$	$q$ ambiente binario $\frac{q^2}{2}$ ambiente continuo
$F(x)_+$	$v - c(v - 0) - p - \tau$
$F(x)_-$	$v - c(1 - v) - p - \tau$
$v$	variabile
$q$	conseguenza di $v$

#### Ambiente binario

Il prezzo per i cookies stabilito dal *data provider*, i.e. YouTube, all'interno dell'ambiente binario si ricava dall'equazione (24) ed è pari a

$$p^N(c, \tau) = \frac{1}{2} \min \{c + \tau; 1 - c + \tau\}. \quad (24)$$

$$p^N(c, \tau) = \frac{1}{2} \min \{0.0523; 0.9477\} = 0.0262 \text{ €/cookie}.$$

Confrontando tale prezzo con il caso di assenza di costo della privacy, i.e. con l'equazione (9), appare evidente che il prezzo appena determinato sia superiore. Le ragioni di tale incremento sono da imputare alla presenza del *data provider* privato,

i.e. YouTube, il quale, in seguito all'introduzione di nuove normative più stringenti nel campo della protezione dei dati, si vede costretto ad operare un maggior sforzo per l'approvvigionamento dei dati degli utenti.

L'equazione (21) invece fornisce il valore del *targeted set*:

$$A(c, p, \tau) = \begin{cases} [0, \max\{c - p - \tau, 0\}], & c < \frac{1 - \tau}{2 + \tau} \\ [\min\{c + p + \tau, 1\}, 1], & c \geq \frac{1 - \tau}{2 + \tau} \end{cases} \quad (21)$$

$$A(c, p, \tau) = \begin{cases} [0, \max\{-0.0739, 0\}], & c < 0.4634 \\ [\min\{0.0785, 1\}, 1], & c \geq 0.4634 \end{cases}$$

$$A(c, p, \tau) = \begin{cases} 0 & c < 0.4634 \\ 0.0785 & c \geq 0.4634 \end{cases}$$

Da cui si evince che il valore soglia  $c$  si abbassa rispetto a quanto previsto in assenza del costo della privacy ( $c = 1/2$ ). Ciò significa che gli *advertisers* scelgono di contattare i clienti a basso valore, i.e. *negative targeting*, solo quando  $c$  è basso proprio perché, per definizione, il *residual set* è di dimensione maggiore rispetto al *targeted set* e si deve sostenere il costo  $\tau$  per le informazioni di un maggior numero di individui. Nel caso in esame, essendo  $c = 0.0023\text{€}/\text{cookie}$ , si adotta quindi la strategia del *negative targeting* e il valore del *targeted set* è pari a 0. Quando il costo per la privacy degli utenti è alto come nel caso in esame, gli *advertisers* preferiscono contattare una grande massa di utenti a basso valore per poi escluderli dalla personalizzazione dell'advertising.

Infine, trattandosi di ambiente binario, è importante sottolineare che gli *advertisers* contattano il gruppo di consumatori guardando il *match value*  $v$  da cui è possibile ricavare il valore della domanda di informazioni  $q$  degli utenti che visitano e usufruiscono di YouTube

$$q^*(v) = \begin{cases} q = 1 & v > 0.4634 \\ q = 0 & v \leq 0.4634 \end{cases}$$

e ottenere un profitto, al netto del prezzo di acquisto dei cookies pari a (2) – (24)

$$\pi^*(v) \triangleq \max \{v - c, 0\} - p^N = \max \{v - 0.0023, 0\} - 0.0262 .$$

Poiché  $p^N > p^*$ , è implicito determinare che il profitto netto degli *advertisers* in presenza del costo  $\tau$  è inferiore rispetto al caso di assenza di costo della privacy.

#### *Ambiente continuo*

Procedendo per simmetria con quanto appena analizzato, il prezzo per i cookies stabilito dal *data provider*, i.e. YouTube, all'interno dell'ambiente continuo si ricava dalle equazioni (31) e (32) ed è pari a

$$p_-^N(v, q, \tau) = \arg \max [p \cdot F(v)] + \tau \quad (31)$$

$$p_+^N(v, q, \tau) = \arg \max [p \cdot (1 - F(v))] + \tau \quad (32)$$

$$p_-^N(v, q, \tau) = \arg \max [p \cdot (v - c(1 - v) - p - \tau)] + \tau$$

$$p_+^N(v, q, \tau) = \arg \max [p \cdot (1 - (v - c(v - 0) - p - \tau))] + \tau$$

e derivando ciascuna equazione si ottiene

$$\frac{dp_-^N}{dp} = v - c + cv - 2p - \tau = 0 \rightarrow p = \frac{v(1 + c) - c - \tau}{2}$$

$$\frac{dp_+^N}{dp} = 1 - v + cv + 2p + \tau = 0 \rightarrow p = \frac{v(1 - c) - 1 - \tau}{2}$$

$$p_-^N(v, q, \tau) = p + \tau = \frac{v(1 + c) - c + \tau}{2} = 0.5012v + 0.0239 \text{ €/cookie}$$

$$p_+^N(v, q, \tau) = p + \tau = \frac{v(1 - c) - 1 + \tau}{2} = 0.4989v - 0.4750 \text{ €/cookie}$$

da cui si evince che il prezzo per i cookies di YouTube dipende dal tipo di utente per il quale la piattaforma intende acquisire i dati. Anche in questo secondo ambiente tale prezzo risulta maggiore rispetto al caso in cui gli utenti non attribuiscono alcun valore alla loro privacy: per dimostrarlo è sufficiente non considerare il termine  $\tau$  che compare all'interno di  $F(x)$  e di  $p^N$ .

All'interno dell'ambiente continuo, la domanda di informazioni per i cookies di YouTube non corrisponde più a  $q = 0$  o  $q = 1$  ma si ricava a partire dall'equazione (6) tale per cui

$$q^*(v) = \arg \max [\pi(v, q)], \text{ con } q \in R_+. \quad (6)$$

Considerando che il profitto per gli *advertisers* è pari a  $\pi(v, q) \triangleq vq - c * \frac{q^2}{2}$ , derivando tale funzione ed uguagliandola a zero si ottiene che, in caso di informazione completa la domanda di informazioni degli utenti corrisponde a

$$\frac{d\pi(v, q)}{dq} = v - cq = 0 \rightarrow q^* = \frac{v}{c} = \frac{v}{0.0023}$$

mentre in caso di informazioni preliminari l'equazione (27) indica che la domanda dipende dal valore atteso del *match value* per cui

$$\bar{q} \triangleq q^*(A^C = V) = q^*(E[v]) \quad (27)$$

$$\bar{q} = \frac{(E[v])}{c} = \frac{(E[v])}{0.0023}.$$

Il profitto netto per gli *advertisers* si determina sostituendo all'interno della funzione di profitto la domanda appena ricavata. Prendendo come riferimento il caso di informazione completa, si può determinare un guadagno netto corrispondente a

$$\pi(v) \triangleq v \frac{v}{0.0023} - c * \frac{\left(\frac{v}{0.0023}\right)^2}{2} - p^N \cdot \frac{v}{0.0023} = 652v^2 - \frac{v}{0.0023} p^N.$$

Infine, trattando separatamente il caso di *negative targeting* e *positive targeting* si può concludere che

$$\pi_-^N(v) = 652v^2 - \frac{v}{0.0023} (0.5012v + 0.0239) = 434v^2 - 10.39v$$

$$\pi_+^N(v) = 652v^2 - \frac{v}{0.0023} (0.4989v - 0.4750) = 435v^2 + 206.52v .$$

Anche in questo caso, il profitto per gli *advertisers* dipende dal valore degli utenti dai quali la piattaforma YouTube acquisisce i dati mediante i cookies e ai quali le imprese che desiderano fare *target advertising* scelgono di rivolgersi.

### 3.5. Osservazioni conclusive

All'interno del presente capitolo è stata studiata un'estensione del modello di *data provision* e *data pricing* riguardante l'introduzione del costo della privacy del consumatore  $\tau$ . Sono stati presi in considerazione diversi attori quali gli *advertisers*, i consumatori e un *data seller/data provider* monopolista all'interno di un ambiente binario e un ambiente continuo.

Nello specifico, per entrambi è stata analizzata la composizione del *targeted set* e del *residual set* poiché da essa è possibile dedurre la strategia di targeting ottima per l'impresa. Inoltre, dopo aver approfondito la domanda di informazioni degli *advertisers*, l'analisi si è spostata sul determinare il prezzo ottimo per i cookies stabilito dal *data provider* tenendo conto dell'introduzione di  $\tau$ . Per concludere l'analisi, infine, è stata presentata un'applicazione del modello ad un caso reale.

I risultati ottenuti grazie all'estensione proposta nel capitolo vengono riassunti con le seguenti proposizioni.

PROPOSIZIONE 1 (Ambiente binario): *Con l'introduzione del costo della privacy  $\tau$  il valore del targeted set per gli advertisers è pari a*

$$A(c, p, \tau) = \begin{cases} [0, \max\{c - p - \tau, 0\}], & c < \frac{1 - \tau}{2 + \tau} \\ [\min\{c + p + \tau, 1\}, 1], & c \geq \frac{1 - \tau}{2 + \tau} \end{cases} \quad (21)$$

dove con  $c < \frac{1 - \tau}{2 + \tau}$  l'impresa opta per il negative targeting mentre con  $c \geq \frac{1 - \tau}{2 + \tau}$  essa sceglie il positive targeting.

PROPOSIZIONE 2 (Ambiente binario): *Il prezzo per i cookies stabilito dal data provider in monopolio è pari a*

$$p^N(c, \tau) = \frac{1}{2} \min\{c + \tau; 1 - c + \tau\}. \quad (24)$$

PROPOSIZIONE 3 (Ambiente continuo): *Per ogni valore di  $c, p, \tau > 0$ , il residual set ottimo è un insieme non vuoto dato dall'intervallo  $A^C = [v_1, v_2]$ . Grazie a questa considerazione è possibile scrivere il problema dell'impresa di scelta dei valori soglia  $v_1$  e  $v_2$  come*

$$\begin{aligned} \max \int_{v_1}^{v_2} [\pi(v, q^*([v_1, v_2])) - \pi^*(v) + p + \tau] dF(v) \\ \text{s.t. } c \cdot m'(q^*([v_1, v_2])) = E[v|v \in [v_1, v_2]]. \end{aligned} \quad (30)$$

PROPOSIZIONE 4 (Ambiente continuo): *Il prezzo per i cookies stabilito dal data provider (pubblico o privato) in monopolio risente delle esternalità di rete negative. Il costo della privacy si identifica come un costo per i data sellers. Tale costo viene ribaltato agli advertisers che sostengono un prezzo maggiore per l'acquisizione delle informazioni, rispetto al caso di assenza di  $\tau$ .*

## CONCLUSIONI

Tutti noi comprendiamo perfettamente quanto sia attuale al giorno d'oggi parlare di digitalizzazione e di dati, in funzione soprattutto dei cambiamenti apportati allo stile di vita di ogni cittadino dovuti alla pandemia Covid-19.

In questi ultimi mesi, tutte le realtà aziendali sono state dirottate sullo *smart working*, l'educazione di qualsiasi grado e livello si è spostata sulle piattaforme online, talvolta create *ad hoc* e il Politecnico di Torino ne rappresenta un lampante esempio, lo sport, dove possibile, è stato messo in pratica singolarmente e in alcuni casi ha subito una battuta d'arresto. In poche parole, la vita di ciascun individuo ha subito una svolta.

Ciò che, in modo più o meno cosciente, è sempre cresciuto e ha continuato a creare valore all'interno del mercato è stato l'uso e l'acquisizione di dati. Il mercato dei dati sta vivendo attualmente un momento di forte espansione grazie alle innumerevoli quantità di processi che sono stati resi digitali e questo crea sicuramente un beneficio per la collettività. Sorge spontaneo però chiedersi quanto gli utenti, i consumatori, i cittadini siano consapevoli delle attività di chi invece all'interno di tale mercato necessita di dati per originare profitti.

All'interno del presente lavoro di tesi è stato, infatti, analizzato il mercato dei dati sotto due punti di vista differenti. Mediante l'analisi SWOT sono stati messi a confronto gli aspetti positivi e negativi derivanti dalle attività legate ai dati e alle piattaforme digitali, i.e. *data collection*, *data storage* e *data analytics*. Sono state quindi analizzate le economie di scala, di scopo e di varietà, le quali sfruttando i bassi costi di gestione del dato, e in particolare la diminuzione dei costi di ricerca, replicazione, trasporto, tracking e verifica, hanno permesso la crescita del mondo digitale; le esternalità di rete e le barriere all'entrata che, grazie al ruolo che svolge il dato all'interno dell'economia, permettono un consolidamento della posizione degli incumbent nel mercato; l'incertezza nel determinare il *consumer welfare* e i danni di natura psicologica-sociale degli utenti e consumatori; il ruolo giocato dall'Antitrust e dalle Autorità Regolatorie le quali costituiscono uno tra i principali strumenti a tutela dei consumatori.

In accordo con lo stato della letteratura odierna, l'intento è stato quello di evidenziare i principali problemi presenti all'interno del *data market* e come si può intervenire per cercare di colmare le inefficienze.

I fallimenti di mercato rappresentano sicuramente un grosso ostacolo verso una sana e leale competizione e il monopolio naturale incarnato dalle piattaforme digitali, il potere di mercato delle grandi imprese, gli *switching costs* imposti agli utenti e le esternalità di rete che invogliano lo scambio di dati tra due versanti gestito da un'unica piattaforma digitale, ne sono un chiarissimo esempio. Anche la discriminazione tariffaria è motivo di inefficienze poiché la possibilità di offrire prezzi personalizzati causa effetti quali la diminuzione della competizione, l'abbassamento della qualità dei beni e servizi offerti, la perdita di fiducia nei mercati online e il *customer lock-in*. Inoltre, mediante lo studio di due modelli di pricing dei dati, si evince che è importante promuovere una allocazione non esclusiva dei dati tra le imprese che competono. Da non sottovalutare sono poi tutte le questioni riguardanti gli abusi di posizione dominante nei confronti delle stesse imprese appartenenti al *data market*, e.g. gli innovatori, e nei confronti dei consumatori, e.g. *target advertising*, sfruttando la poca consapevolezza sul funzionamento del mondo digitale.

La tutela della privacy ricopre quindi un ruolo chiave tra le dinamiche competitive: in particolare, il tempo per il quale i dati degli utenti vengono conservati viene definito come il driver più importante per garantire la sicurezza degli individui. I problemi e le improduttività mostrate possono venir attutate e, in parte, anche colmate da alcune operazioni di policy quali la garanzia della portabilità e interoperabilità dei dati, l'istituzione di sistemi *open standard* e trasparenti, il ribaltamento dell'onere probatorio a carico dell'incumbent in caso di accusa di comportamenti anti-competitivi, la messa in atto del GDPR e l'attenta gestione delle situazioni problematiche da parte di Antitrust e Autorità Regolatorie.

La letteratura presentata finora mostra, inoltre, interessanti spunti da cui possono sorgere ulteriori analisi, come è stato fatto poi all'interno dell'ultimo capitolo del presente lavoro. Partendo dal modello di *data provision* e *data pricing* di Bergemann e Bonatti (2015) si è esaminato come cambia l'ambiente competitivo quando si monetizza il diritto alla privacy dei consumatori all'interno del mercato dei

dati. L'analisi degli aspetti chiave del modello, i.e. dei cookies e delle dinamiche mediante le quali essi vengono scambiati, è stata svolta sia teoricamente, approfondendo l'ambiente binario e continuo, sia praticamente e numericamente, applicando tale modello al caso reale di YouTube. Si è potuto concludere che se si dovessero prezzare i dati che vengono scambiati tra i diversi attori del mercato, il prezzo possiederebbe un sovraccarico dato dal costo della privacy. Tale costo, inoltre, influenza le scelte delle imprese che acquistano le informazioni private poiché esse decidono di rivolgersi ad un certo segmento di consumatori, contattando gli utenti ad alto valore o quelli a basso valore, a seconda del punto in cui è massimo il loro profitto.

### *Uno sguardo tra passato e futuro*

A conclusione di tutto questo percorso, è quindi fondamentale domandarsi se e in che misura stiamo cominciando a prendere coscienza del ruolo che le grandi imprese e piattaforme digitali ricoprono oggi all'interno dell'economia mondiale. Da uno studio pubblicato ne "Il Sole 24 Ore" basato sui dati di Standard&Poor's (Franceschi De Marchi, 2018) si evince che tra le dieci aziende più potenti al mondo per capitalizzazione di mercato sono inserite le più famose over-the-top (Fig. 21).



Figura 21 - Le top 10 imprese più potenti al mondo per capitalizzazione di mercato (anno 2018) - fonte: lab24.ilsole24ore.com/aziende-top/.

A differenza di quanto avveniva in passato, quando le imprese che traevano maggior successo erano le compagnie petrolifere o inerenti comunque al settore dell'energia, oggi il driver dell'economia mondiale è rappresentato dal settore della tecnologia (Fig. 22).

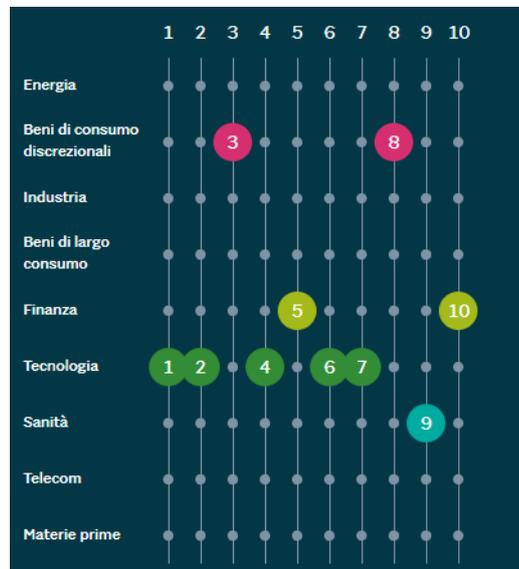


Figura 22 - Suddivisione delle top 10 per settore (anno 2018) - fonte: [lab24.ilsole24ore.com/aziende-top/](http://lab24.ilsole24ore.com/aziende-top/).



Figura 23 - Evoluzione nel tempo delle top 10 più tecnologiche - fonte: [lab24.ilsole24ore.com/aziende-top/](http://lab24.ilsole24ore.com/aziende-top/).

Il cambiamento messo in atto negli ultimi vent'anni, che ha portato alla definizione di nuovi standard quali gli smartphone e l'impiego di Internet in mobilità,

non è affatto intenzionato ad arrestarsi e, come si può sperimentare quotidianamente, il progresso tecnologico viaggia ad una velocità estremamente competitiva (Fig. 23).

Non è chiaro il punto di arresto di un mercato così in ampio divenire. Ciò che però è degno di nota è che è estremamente necessario tutelare gli utenti e i consumatori che operano transazioni al suo interno e risulta altrettanto fondamentale regolare il comportamento di tutte quelle imprese che, per i differenti motivi analizzati nel presente elaborato di tesi, si trovano ad avere una posizione che impedisce al *data market* di essere concorrenziale. Come visto anche all'interno dell'estensione del modello "Selling Cookies", incrementando il prezzo di acquisizione delle informazioni dei consumatori è possibile offrire i corretti incentivi ai *data provider* privati affinché siano meglio regolate le transazioni sui dati.

Ritengo però di altrettanta importanza la posizione che lo Stato occupa nel mercato. Potrebbe esso trasformarsi nel possessore dei dati degli utenti? Sarebbe in grado di garantire la gestione delle transazioni agendo come unico *data provider* pubblico? Accade tutt'ora in altri settori che lo Stato, istituendo la figura dell'Autorità Regolatoria, intervenga a tutela dei consumatori mediante l'introduzione di tariffe che limitano il potere di azione delle imprese private (si veda a questo proposito il ruolo ricoperto dall'Autorità di Regolazione per Energia, Reti e Ambienti oppure dall'Autorità di Regolazione dei Trasporti). Potrebbe quindi lo Stato istituire una figura che ricopra un ruolo analogo anche per il *data market*? È proprio questo, ciò che mostra in conclusione l'estensione del modello analizzato. Se lo Stato decidesse di agire come *data provider* pubblico, esso potrebbe assicurare ai consumatori la massimizzazione dei benefici per la collettività. Anche in questo caso, entrerebbe in gioco un meccanismo tariffario *ad-hoc* per compensare gli utenti che si vedono espropriati delle loro informazioni personali. Infine, lo Stato potrebbe essere in grado di garantire e, soprattutto, di rispettare il diritto alla privacy dei consumatori? A quest'ultima questione purtroppo non è possibile rispondere univocamente. Lo Stato è costituito da un insieme di cittadini, enti e governatori eletti dai cittadini e, in parte, dipende da ogni singolo individuo assicurarsi uno Stato in grado di soddisfare le esigenze e le necessità collettive, in questo caso il diritto alla privacy. In conclusione, i dati rappresentano per la comunità un asset preziosissimo e delicato e per tale ragione è imprescindibile, sin da ora, maneggiarli con la dovuta attenzione.

## BIBLIOGRAFIA

- Acemoglu, D., Makhdoumi, A., Malekian, A., & Ozdaglar, A. (2019). *Too much data: Prices and inefficiencies in data markets*. National Bureau of Economic Research.
- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). *Privacy and human behavior in the age of information*. Sciencemag.org.
- Acquisti, A., Taylor, C., & Wagman, L. (2016). *The economics of privacy*. Journal of economic Literature.
- AGCM, AGCOM, Garante per la protezione dei dati personali. (Luglio 2019). *Big Data. Indagine conoscitiva congiunta. Linee guida e raccomandazioni di policy*. www.garanteprivacy.it.
- Aridor, G., Che, Y. K., & Salz, T. (2020). *The Economic Consequences of Data Privacy Regulation: Empirical Evidence from GDPR*. National Bureau of Economic Research.
- Armstrong, M., & Zhou, J. (2010). *Conditioning prices on search behaviour*. Mimeo.
- Asadullah, A., Faik, I., & Kankanhalli, A. (2018, June). *Digital Platforms: A Review and Future Directions*. PACIS.
- Australian Competition & Consumer Commission. (2019). *Digital Platforms Inquiry. Final Report*. ACCC.
- Authority of the House of Lords. (March 2019). *Regulating in a digital world*. Select Committee on Communications.
- Bergemann, D., & Bonatti, A. (2015). *Selling cookies*. American Economic Journal: Microeconomics.
- Bergemann, D., Brooks, B., & Morris, S. (2015). *The limits of price discrimination*. American Economic Review.
- Boccellato, P. (31 Luglio 2020). *L'Australia si schiera con gli editori. "Google e Facebook pagheranno le news"*. Key4biz. www.key4biz.it.

- Bourreau, M., De Streel, A., & Graef, I. (2017). *Big Data and Competition Policy: Market power, personalised pricing and advertising*. Centre on Regulation in Europe.
- Calvano, E., & Polo, M. (2020). *Market Power, Competition and Innovation in digital markets: A survey*. Information Economics and Policy, 100853.
- Crémer, J., de Montjoye, Y. A., & Schweitzer, H. (2019). *Competition policy for the digital era*. Report for the European Commission.
- Chiou, L., & Tucker, C. (2017). *Search engines and data retention: Implications for privacy and antitrust*. National Bureau of Economic Research.
- Delmastro, M. & Nicita, A. (2019). *Big data. Come stanno cambiando il nostro mondo*. Il Mulino.
- De Blase, L. (4 Agosto 2019). *Il monopolio nell'era digitale. Un Antitrust che salvaguardi l'innovazione*. Il Sole 24 Ore.
- Deneckere, R. J., & Preston McAfee, R. (1996). *Damaged goods*. Journal of Economics & Management Strategy.
- Dubus, A. (2019). *Strategic information and competition in digital markets*. *Economics and Finance*. Université Paris-Saclay. NNT: 2019SACL037.
- Ek, D. (13 Marzo 2019). *Consumer and Innovators Win on a Level Playing Field*. [newsroom.spotify.com/2019-03-13/consumers-and-innovators-win-on-a-level-playing-field](https://newsroom.spotify.com/2019-03-13/consumers-and-innovators-win-on-a-level-playing-field).
- Ellison, G., & Ellison, S. F. (2009). *Search, obfuscation, and price elasticities on the internet*. *Econometrica*.
- European Commission. (2020). *Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the regions. A European strategy for data*. EU Commission.
- Evans, P. C., & Gawer, A. (2016). *The rise of the platform enterprise: a global survey*.

- Franceschi De Marchi, A. (2018). *Le prime 10 aziende al mondo negli ultimi 25 anni*. Il Sole 24 Ore. S&P Market Intelligence. [lab24.ilsole24ore.com/aziende-top/](http://lab24.ilsole24ore.com/aziende-top/).
- Furman, J., Coyle, D., Fletcher, A., McAuley, D., & Marsden, P. (2019). *Unlocking digital competition: Report of the digital competition expert panel*. UK government publication, HM Treasury.
- Giaccone, P. (2020). *Data Center, IoT and Protocols, 5G, Web Services*. Slide del corso di Strumenti ICT per la gestione dei Servizi. Politecnico di Torino.
- Garofalo, L. (10 Settembre 2019). *Agcom: "Procedimento per accertare posizioni dominanti nella pubblicità online" (Cosa rischiano Google e Facebook)*. Key4biz. [www.key4biz.it](http://www.key4biz.it).
- Garofalo, L. (10 Settembre 2019). *Google e Facebook, indagini antitrust anche negli Usa. Dall'entusiasmo alla stretta sulla Silicon Valley*. Key4biz. [www.key4biz.it](http://www.key4biz.it).
- Goldfarb, A., & Tucker, C. E. (2019). *Digital economics*. Journal of Economic Literature.
- Goldfarb, A., & Tucker, C. E. (2011). *Privacy regulation and online advertising*. Management science.
- Grance, T., Mell, P. (2011). *The NIST definition of Cloud Computing*. National Institute of Standards and Technologies NIST.
- HR Governments (2019, April). *Online Harms White Paper*.
- Jentsch, N., Preibusch, S., & Harasser, A. (2012). *Study on monetising privacy: An economic model for pricing personal information*. ENISA.
- Kitchin, R., & McArdle, G. (2016). *What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets*. Big Data & Society.
- Luna, R. (29 Settembre 2020). *Scopri quanti ti spiano mentre visiti il tuo sito preferito*. La Repubblica. [www.repubblica.it/dossier/stazione-futuro-riccardo-luna/2020/09/29/news/scopri\\_quanti\\_ti\\_spiano\\_mentre\\_visitati\\_il\\_tuo\\_sito\\_preferito-268846248/](http://www.repubblica.it/dossier/stazione-futuro-riccardo-luna/2020/09/29/news/scopri_quanti_ti_spiano_mentre_visitati_il_tuo_sito_preferito-268846248/).

- Montes, R., Sand-Zantman, W., & Valletti, T. (2019). *The value of personal information in online markets with endogenous privacy*. Management Science.
- OECD. (2017). *Algorithms and Collusion: Competition Policy in the Digital Age*. [www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.html](http://www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.html).
- Ofcom (2019). *Online market failures and harms. An economic perspective on the challenges and opportunities in regulating online services*. Office of Communications.
- Pagani, M. (2013). *Digital business strategy and value creation: Framing the dynamic cycle of control points*. Mis Quarterly.
- Pigou, A. C. (1920). *The economics of welfare*. Macmillan & Co.
- Polichetti, A. (28 Aprile 2020). *5G e salute: cosa c'è da sapere*. Centro Nazionale per la Protezione dalle Radiazioni e Fisica Computazionale. Istituto Superiore di Sanità. [www.iss.it/primopiano/asset\\_publisher/o4oGR9qmvUz9/content/id/5355363](http://www.iss.it/primopiano/asset_publisher/o4oGR9qmvUz9/content/id/5355363).
- Rochet, J. C., & Tirole, J. (2003). *Platform competition in two-sided markets*. Journal of the european economic association.
- Sabatino, L. (2020). *The Economics of Data, Privacy, and (a bit of) Machine Learning*. Slide del corso di Economia e Management dei Servizi. Politecnico di Torino.
- Sabatino, L. (2020). *The Economics of Platforms*. Slide del corso di Economia e Management dei Servizi. Politecnico di Torino.
- Salop, S. C. (1979). *Monopolistic competition with outside goods*. The Bell Journal of Economics.
- Scott Morton, F., Bouvier, P., Ezrachi, A., Jullien, B., Katz, R., Kimmelman, G., Morgenstern, J. (2019). *Committee for the study of digital platforms: Market structure and antitrust subcommittee report*. Stigler Center for the

Study of the Economy and the State, University of Chicago Booth School of Business.

Spagnoletti, P., Resca, A., & Lee, G. (2015). *A design theory for digital platforms supporting online communities: a multiple case study*. Journal of Information Technology.

Turow, J., Feldman, L., & Meltzer, K. (2005). *Open to exploitation: America's shoppers online and offline*. Departmental Papers (ASC).

Unione Europea. (25 Maggio 2018). *Regolamento (UE) n. 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati)*. Gazzetta Ufficiale dell'Unione Europea.

## SITOGRAFIA

[ec.europa.eu/competition/consumers/abuse\\_it](https://ec.europa.eu/competition/consumers/abuse_it).

[lab24.ilsole24ore.com/aziende-top/](https://lab24.ilsole24ore.com/aziende-top/).

[n26.com/it](https://n26.com/it).

[newsroom.spotify.com/2019-03-13/consumers-and-innovators-win-on-a-level-playing-field](https://newsroom.spotify.com/2019-03-13/consumers-and-innovators-win-on-a-level-playing-field).

[ventiblog.com/gli-effetti-della-rete-il-concetto-di-esternalita-di-rete](https://ventiblog.com/gli-effetti-della-rete-il-concetto-di-esternalita-di-rete).

[www.aranzulla.it/come-youtube-paga-i-video-66228.html](https://www.aranzulla.it/come-youtube-paga-i-video-66228.html).

[www.cybersecurity360.it/legal/privacy-dati-personali/pseudonimizzazione-e-anonimizzazione-dei-dati-differenze-tecniche-e-applicative](https://www.cybersecurity360.it/legal/privacy-dati-personali/pseudonimizzazione-e-anonimizzazione-dei-dati-differenze-tecniche-e-applicative).

[www.dday.it/redazione/36400/amazon-apple-google-facebook-antitrust-usa](https://www.dday.it/redazione/36400/amazon-apple-google-facebook-antitrust-usa).

[www.ilpost.it/2018/03/19/facebook-cambridge-analytica](https://www.ilpost.it/2018/03/19/facebook-cambridge-analytica).

[www.okpedia.it/opt-in](https://www.okpedia.it/opt-in).

[www.oracle.com/data-cloud/products/data-management-platform](https://www.oracle.com/data-cloud/products/data-management-platform).

[www.oracle.com/corporate/acquisitions/datalogix](https://www.oracle.com/corporate/acquisitions/datalogix).

[www.repubblica.it/tecnologia/2020/07/29/news/1\\_antitrust\\_americano\\_mette\\_sotto\\_a\\_ccusa\\_amazon\\_apple\\_facebook\\_google-263198696](https://www.repubblica.it/tecnologia/2020/07/29/news/1_antitrust_americano_mette_sotto_a_ccusa_amazon_apple_facebook_google-263198696).

[www.satispay.com/it-it](https://www.satispay.com/it-it).

## RINGRAZIAMENTI

*Durante gli ultimi mesi ho avuto modo di comprendere quanto io sia cresciuta grazie a Voi: siete stati per me una preziosa fonte di insegnamenti e sarebbe scontato dire solamente “grazie”. Proprio per questo ho deciso di esprimere la mia gratitudine a mio modo, scrivendo ciò avete rappresentato per me in questi anni.*

*Professor Cambini, mediante i Suoi insegnamenti, Lei mi ha ridato motivazione e voglia di arrivare. Ha sempre mostrato la Sua infinita diponibilità e gentilezza prima di tutto e queste sono qualità rare non solo per un Professore ma proprio come Persona. Sono molto grata di aver incontrato un Professore come Lei.*

*Amici-colleghi-compagni-maestre di Twirling (Cri, Fra Fiore, Gabri, Giada, Luca, Robi, Silvia, Stefi), siete la mia seconda famiglia da quando avevo sei anni. Ogni giorno siete una certezza, un appoggio, un sostegno e un conforto.*

*Sciao Dimi Crew (Ale Dolce, Ani, Annina, Bira, Checco, Franci, Giachi, Matti, Paolino e Robi), in questi cinque anni siete stati per me il vero motore delle mie giornate al Poli, le risate in aula studio o nelle pause, le sfuriate post esame e le imprecazioni per le insoddisfazioni, le feste Isef prima e gli aperitivi su Skype poi. Da voi ho imparato quella leggerezza che tanto mi mancava e sono davvero felice di quanto ci siamo divertiti durante la nostra breve-ma-intensa convivenza a Porto.*

*Jack, tu lo sai. Il tuo modo di essere, di comportarti, di affrontare le mie stesse situazioni nel medesimo contesto è stato per me la strada a cui guardare. Mi hai insegnato ad avere pazienza, ad apprezzare i miei traguardi e ad essere contenta di me stessa, e sai che faccio fatica. Sei stato l'unico che ha sempre condiviso i miei obiettivi, ma soprattutto, che ha supportato e sopportato come io ho scelto di raggiungerli, anche se tu non avresti fatto lo stesso. Sei prezioso.*

*Mamma, papà e Lollo, avete da sempre rappresentato per me il luogo in cui potevo disperarmi perché quello che facevo non andava secondo i miei piani. Col passare degli anni, poco per volta, avete accompagnato i miei piccoli successi alleviando un po' la tensione degli attimi più intensi. E oggi, nonostante tutto, possiamo goderci insieme questo momento felice. Sarete per sempre la mia casa, in qualsiasi luogo io deciderò di andare. Grazie.*