

BIG DATA, COMPETITION AND RELEVANT MARKETS

MARCO GAMBARO

Working Paper 15/2017

OCTOBER 2017

 UNIVERSITÀ DEGLI STUDI DI MILANO



DIPARTIMENTO DI ECONOMIA, MANAGEMENT E METODI QUANTITATIVI

Via Conservatorio 7
20122 Milano

tel. ++39 02 503 21501 (21522) - fax ++39 02 503 21450 (21505)

<http://www.economia.unimi.it>

E Mail: dipeco@unimi.it

BIG DATA CONCORRENZA E MERCATI RILEVANTI¹

Marco Gambaro, Università degli Studi di Milano, Demm

ABSTRACT

Big data are considered at the same time a promising driver of economic development and a concern for possible manipulation and privacy intrusion. Data diffusion and their uncertain appropriability can make property rights for data less precise.

The paper reviews some economic features of data. In many digital markets data can be considered a relevant input but hardly an essential facility.

Many data are collected in two-sided market platforms and on one side they are used to personalize services and to add quality while on the other side of the platform they contribute to make advertising collection more efficient. So the transfer of personal data can be considered an implicit price for many free information services. Consumers usually are unaware of subsequent pervasive use of their personal data, and therefore give away easily their data. Big data can amplify competitive advantages and related dominant positions, leveraging on information asymmetries.

A dominant position obtained through collection and processing of big amounts of personal data allows practices such as first-degree price discrimination, personalized advertising, artificial degradation of services that sometimes can be considered competitive abuse, but it is difficult that data alone allow such a position.

Keywords: big data, data protection, personal data, privacy, value of data, value of privacy, relevant market for data

JEL Codes: K21, K24, L12, L86, L44

Contact: marco.gambaro@unimi.it

¹ Questo paper costituisce una versione preliminare di un capitolo del libro Falce V., Olivieri G., Big data e concorrenza nei mercati dell'innovazione, Giuffrè

BIG DATA CONCORRENZA E MERCATI RILEVANTI²

Marco Gambaro, Università degli Studi di Milano, Demm

1 INTRODUZIONE

Da alcuni anni i big data sono considerati uno dei più promettenti driver di sviluppo economico e allo stesso tempo, più recentemente, costituiscono fonte di preoccupazione per il possibile potere accumulato da chi raccoglie grandi quantità di dati e ha le capacità di elaborarli.

I big data sono al centro di un cluster di innovazioni che interagiscono tra di loro e che assieme favoriscono la digitalizzazione dei mercati e della società. Tutte si fondano sull'innovazione tecnologica e sulla riduzione dei costi degli apparati di elaborazione e di comunicazione.

Il cloud computing consente di fatto di trasformare in costi variabili quelli che una volta erano costi fissi di infrastruttura, elaborazione e memorie informatiche. Questo abbassa drasticamente le economie di scala per la nascita di nuove imprese tecnologiche e la scalabilità delle risorse accompagna la crescita senza richiedere ingenti investimenti specifici. Anche aziende di successo come Huber o Netflix sono partite utilizzando potenza di calcolo o mappe vettoriali fornite tra l'altro da Amazon e Google.

Gli open data si riferiscono per lo più ai dati relativi alle transazioni e alle procedure condotte dalla pubblica amministrazione e sono generalmente big (censuari, multe, accessi agli uffici), ma talvolta small (apertura centri commerciali, omicidi) e talvolta possono essere parte di internet of things (dati sui semafori).

Quando queste grandi quantità di dati sono disponibili per ricercatori e imprese emergono idee di incrocio e di elaborazioni che portano a nuove idee e a nuovi servizi, quindi si tratta di un driver potenzialmente di grande impatto. Ma non sempre la pubblica amministrazione si spoglia facilmente del controllo esclusivo dei dati.

Gran parte dei big data sono costituiti da rilevazioni e informazioni di macchine reti e apparati generalmente compresi in Internet of Things.

Per internet of things si intende generalmente la diffusione di sensori e regolatori in apparecchi ed oggetti che oltre a svolgere un'azione specifica di misura o di regolazione producono una gran quantità di dati sui fenomeni cui sono collegati. Le stime più ragionevoli parlano di 25-50 miliardi di apparati collegati o collegabili entro i prossimi 10 anni. Si tratta di un ordine di grandezza superiore al numero dei device attualmente connessi in rete. Prima che i dati siano utilizzabili occorre risolvere problemi di connettività (wifi e mashup), le stesse reti potrebbero aver problemi di saturazione non tanto di capacità perché la banda utilizzata è generalmente bassa, quanto di latenza e di risorse di segnalazione. Un ulteriore problema è in molti casi costituito dalle necessità elettriche di alimentazione, mentre questi dati potranno essere più utili se tutti questi dispositivi avranno un indirizzo IP che consente di gestirli come elementi connessi alla rete, anche per la raccolta e la distribuzione dei dati.

Oltre a questi vi sono tutti i dati di natura più personale tra cui dati generati dalle transazioni (finanziarie, amministrative e commerciali), tutti i dati collegati alla presenza e alla navigazione in rete e tutti gli user generated data (mail, social network, immagini, video, web content, testo libero) che generalmente sono in forma non strutturata e che richiedono analisi di tipo semantico per poter essere utilizzati.

Le previsioni sui possibili valori economici sono espresse in numeri molto grandi, ma per il momento appaiono più supposizioni che previsioni. Generalmente le previsioni sono fatte con la logica: se i costi di fabbricazione dell'intera economia scendessero dell'1% allora il valore sarebbe. Per cui ad esempio Deutsche Bank ipotizza nell'industria petrolifera una riduzione dell'1% in spese di capitale che si tradurrebbe in 90 miliardi; nelle smart utilities e nell'aviazione una riduzione dell'1% di carburante utilizzato si traduce rispettivamente in 66 miliardi e in 30 miliardi di dollari.

I settori più promettenti sono industria automobilistica, utilities, case connesse, settore industriale, sanità. Mentre le potenzialità teoriche sono spalmate orizzontalmente quasi sull'intera economia, gli sviluppi concreti saranno molto più localizzati e dipenderanno dalle specifiche condizioni competitive.

² Questo paper costituisce una versione preliminare di un capitolo del libro Falce V., Olivieri G., Big data e concorrenza nei mercati dell'innovazione, Giuffrè

In tutte le aree considerate vi sono rilevanti economie di rete che rendono non semplice l'avvio. In molti casi le imprese devono prima installare un gran numero di apparati, sensori o regolatori, poi porsi il problema di cosa fare con i dati, infine risolvere i problemi tecnici di elaborazione di grandi quantità di dati e solo alla fine possono sperare nei vantaggi.

La possibilità di poter internalizzare completamente i potenziali vantaggi costituisce un potente acceleratore, mentre in quei casi in cui il monitoraggio genera un possibile servizio migliore per terze parti o clienti finali, il disallineamento degli incentivi e i potenziali comportamenti opportunistici, possono rallentare molto l'adozione. Ad esempio se attraverso i sensori installati in un'automobile un costruttore è in grado di prevedere i possibili guasti e programmare interventi di manutenzione, dovrà anche convincere i clienti che non sta facendo dipendere questi interventi di manutenzione dalla necessità di ricavo delle sue officine o dal suo budget di vendita dei pezzi di ricambio.

Sempre con riferimento al caso delle automobili si pone il problema più generale, e potenzialmente molto rilevante, della proprietà del dato. Se sull'automobile che acquisto da un costruttore e di cui ho la piena proprietà vi sono dei sensori di monitoraggio i dati generati sono di mia proprietà oppure sono nella disponibilità del costruttore; e qualora il costruttore si rifiuti di riparare la vettura se non metto a disposizione i dati, sta abusando della sua posizione? Se inoltre sulla stessa automobile viene installato un gps di un'assicurazione che monitorizza l'utilizzo del mezzo, può utilizzare i dati dei sensori dell'automobile? Si tratta di problemi regolamentari non facili da risolvere.

I big data, assieme ad altre innovazioni digitali, costituiscono un prerequisito ed un driver per la nascita di nuovi mercati. Ad esempio nella sharing economy dove, se pure il punto di partenza è spesso la presenza di beni capitali sottoutilizzati (case e automobili), senza una grande diffusione degli smartphone, della disponibilità di dati di traffico, tecnologie di fotografia di alta qualità, mappe a basso costo, sistemi di localizzazione, servizi come Uber o Airbnb non sarebbero neppure pensabili.

Lo sviluppo dei big data comunque non è così tutto in discesa e occorre superare notevoli difficoltà sia nel caso di dati prodotti da oggetti che dai dati personali.

Quando sono utilizzati non per migliorare la specifica transazione ma per prevedere l'andamento dei mercati o delle attività i big data non sfuggono alle tradizionali trappole statistiche. Per quanto i volumi di dati siano molto grandi e a basso costo, tuttavia si tratta sempre di campioni e non dell'universo. In secondo luogo le azioni e le informazioni di oggi non possono essere estrapolate facilmente a domani e infine ciò che accade on line non è necessariamente uguale a ciò che accade offline.

In questo contesto è anche emerso il timore che queste grandi quantità di dati possano essere utilizzate in modi impropri e che si stia accentuando un'asimmetria di potere tra le imprese che raccolgono questi dati e i consumatori che sarebbero alla loro mercé. Alcune recenti come quella tra Facebook e Whatsapp, oppure quella tra Google e Double Click hanno accentuato queste preoccupazioni e favorito le richieste di intervento regolatorio sia agli organi legislativi che alle diverse autorità che sono coinvolte in quest'area variegata dall'antitrust alle autorità sulla privacy.

Il presente capitolo è organizzato come segue: il prossimo paragrafo descrive alcune caratteristiche economiche dei dati, il terzo paragrafo passa in rassegna le possibilità di intervento dell'antitrust, il quarto discute della definizione di mercato rilevante, il quinto esamina i problemi delle regolamentazioni ex ante e discute alcune implicazioni di policy.

2 CARATTERISTICHE ECONOMICHE DEI DATI

I dati condividono con l'informazione alcune caratteristiche economiche peculiari. Innanzitutto si tratta di un bene difficilmente appropriabile e non rivale. Il fatto dare a qualcuno un'informazione non impedisce di darla anche ad altri e non riduce le successive possibilità di utilizzo. In secondo luogo si tratta di un bene con elevati costi fissi e bassi o nulli costi variabili. I mercati che la riguardano sono quindi caratterizzati da economie di scala.

In terzo luogo nei mercati caratterizzati da un uso intenso dei dati, questi possono essere considerati un input rilevante, ma certamente non l'unico input. Per produrre conoscenza e valore occorrono anche altre risorse tra cui capacità di calcolo, software specializzato di elaborazione e risorse umane qualificate per

leggere e progettare i risultati, computer scientist e data analyst. Ma spesso la barriera più rilevante ad un uso proficuo dei dati disponibili, sono proprio i modelli manageriali consolidati e la scarsa capacità di rendere flessibili i processi organizzativi. Per questo le organizzazioni che nascono già con una cultura dell'uso dei dati hanno spesso un vantaggio di flessibilità sulle aziende consolidate che si trovano a dover intraprendere importanti e difficili processi di cambiamento.

Dal punto di vista economico un aspetto rilevante dei dati e dell'informazione consiste nella possibilità di sostituire altre risorse costose nei processi di produzione e di scambio. Gli esempi sono tanti e riguardano sia la sostituzione di risorse umane che di altri tipo di input. In tutta l'area commerciale i recommendation system sostituiscono negozi fisici e commessi, nei trasporti il monitoraggio e la pianificazione delle rotte fa risparmiare carburante, nell'industria petrolifera l'analisi preliminare delle potenzialità estrattive fa risparmiare molte trivellazioni. Si registra dunque una maggior efficienza di molti processi produttivi e allo stesso tempo una migliore appropriatezza dei prodotti e servizi offerti ai consumatori.

Un elemento di novità rispetto ai dati tradizionali è costituito dalla pervasività della raccolta e dal volume dei dati trattati. Infatti visto che molte attività umane sono realizzate con il supporto di applicazioni digitali, la raccolta di informazioni richiede pochissimi costi di monitoraggio aggiuntivi. In moltissimi casi la raccolta di informazioni è un byproduct di un'attività principale.

Molti dati personali sono raccolti attraverso piattaforme two sided che da un lato utilizzano i dati per realizzare e personalizzare il servizio offerto gratuitamente o fortemente sussidiato, e dall'altra parte usano i dati personali per rendere più efficace, più produttiva e quindi più redditizia la raccolta pubblicitaria. In questo contesto i dati potrebbero essere considerati un prezzo implicito per l'offerta di quei servizi sussidiati. Anzi secondo alcuni i dati starebbero addirittura diventando una nuova currency.

Innanzitutto va precisato che nelle piattaforme two sided il prezzo implicito è prevalentemente costituito dall'attenzione per la pubblicità ed i dati costituiscono un input accessorio per rendere più efficace la pubblicità stessa. In secondo luogo i dati non possono essere considerati una vera e propria valuta in senso economico perché non sono una risorsa scarsa e non hanno un valore uniforme riconosciuto da tutti i partecipanti. Naturalmente la perdita di privacy collegata alla cessione dei dati è una disutilità per il consumatore, ma non basta per parlare di currency.

Il valore dei dati è molto contingente e dipende da fattori diversi. Il singolo dato difficilmente ha molto valore mentre un'ampia raccolta di dati può essere preziosa. La tempestività del dato può essere importante perché spesso, soprattutto per gli usi operativi, l'utilità e il valore del dato decrescono col tempo. La disponibilità a cedere i dati cambia in funzione della fiducia nella controparte e della sua credibilità. Lo stesso utente può valutare diversamente la cessione dei suoi dati a seconda degli interlocutori e del contesto. Molti segnali sembrano indicare come il valore contingente dei propri dati personali sia per il consumatore abbastanza ridotto, anche perché la cessione del dato non è mai esclusiva. Tuttavia mancano delle valutazioni strutturate sulla dimensione del valore e sulle determinanti delle variazioni.

Da molte ricerche emerge come i consumatori siano preoccupati dell'uso dei propri dati personali. Secondo il Pew Research (2016) il 1% degli adulti pensa che i consumatori abbiano perso il controllo su come i loro dati sono raccolti e utilizzati dalle aziende e poche esprimono la convinzione che i loro dati rimangano privati e sicura una volta che sono stati raccolti.

Naturalmente il consumatore potrebbe essere in grado di scegliere quali dati cedere e in quali situazioni per cui la situazione che si osserva sarebbe semplicemente la fotografia delle preferenze dei consumatori. Ma questa possibilità di disciplina del mercato e questo ruolo del consumatore richiedono due condizioni importanti: la disponibilità di informazione completa sulle diverse contingenze e la presenza di alternative che possano far valere il gioco concorrenziale. In molte situazioni queste due condizioni non sono verificate. Molte ricerche empiriche, e anche l'esperienza quotidiana, mostrano come il consumatore non sia informato riguardo le implicazioni della cessione dei suoi dati personali sia perché leggere i lunghissimi disclaimer scritti a caratteri piccolissimi è praticamente impossibile per chiunque, sia perché non vengono esplicitati i passaggi che i dati faranno e i possibili usi successivi.

Anche da un punto di vista economico la preoccupazione dei consumatori non è irrazionale perché se l'informazione è potere l'asimmetria informativa che si crea con le imprese può spostare l'equilibrio di potere nelle transazioni. Se il venditore è in grado di sapere quanto poco so riguardo ad un prodotto che voglio acquistare cercherà di indirizzarmi verso un prodotto più conveniente per lui che per me. Anche

perché le imprese non hanno incentivi socialmente ottimali per puntare ad un matching corretto tra prodotti e clienti (JEL 2016).

La cessione dei dati da parte dei consumatori avviene generalmente di fronte ad un beneficio immediato (un servizio o un prodotto) a fronte di costi incerti, spalmati in un periodo futuro e poco conosciuti. Non è detto dunque che il consumatore sia in grado di valutare correttamente i tradeoff associati alla cessione di informazioni.

Negli ultimi anni sono emerse in particolare le preoccupazioni relative alla raccolta e all'utilizzo delle informazioni personali, problema che come si è visto ha una dimensione economica ma che si allarga ad altri ambiti dell'interagire sociale. Infatti l'assenza di rilevanti asimmetrie informative, la disponibilità di informazioni affidabili e ragionevolmente obiettive costituisce un prerequisito per il buon funzionamento del sistema politico di presa delle decisioni, del mercato finanziario oltre che naturalmente per il funzionamento dei mercati dei beni e dei servizi.

E quindi se gran parte delle interazioni sociali sono mediate da piattaforme che raccolgono dati personali e monitorano i comportamenti e le scelte informative, emergono preoccupazioni proprio per queste dimensioni complessive della disuguaglianza.

In effetti è difficile sostenere che le informazioni raccolte sono relative e proporzionali ai servizi forniti. Le applicazioni sui dispositivi mobili delle grandi piattaforme richiedono tra l'altro la possibilità di modificare la rubrica e i contenuti del dispositivo, la possibilità di attivare il microfono e la telecamera, la lettura di tutti i contenuti delle memorie, il download dei file senza notifica, l'invio di comunicazioni all'insaputa dei proprietari. Si tratta di un grado di controllo potenziale particolarmente elevato di cui si cominciano a vedere i risultati. Come quell'utente che parlando con un amico di persona si lamentava dei parassiti alle piante e che poco dopo ha ricevuto la pubblicità personalizzata di un'insetticida. La conversazione evidentemente era stata intercettata dal microfono del telefonino e un qualche robot aveva riconosciuto qualche parola chiave.

Le grandi piattaforme, sulla base dell'analisi dei dati di traffico e di interazione, sono in grado di inviare contenuti personalizzati a loro volta capaci di influenzare l'umore e l'orientamento degli utenti. Ma cosa succede se chi controlla una piattaforma decide di candidarsi per una carica pubblica oppure decide di supportare un qualche candidato.

La raccolta, l'analisi e l'uso di grandi quantità di dati personali fanno emergere problemi che riguardano contemporaneamente il funzionamento economico dei mercati, la disuguaglianza, l'intrusione nella sfera personale.

Ad esempio quando un'automobile, ormai funzionante con una centralina digitale e con un gps, raccoglie dati sull'usura dei componenti, sulle modalità d'uso e sullo stile di guida, come assicurarsi che suggerisca interventi di manutenzione sulla base dell'effettiva necessità e non invece pensando alla massimizzazione dei ricavi dei propri centri assistenza. E quei dati raccolti sono proprietà del proprietario della macchina, del costruttore o della compagnia di assicurazione?

La privacy e la condivisione dei dati rappresentano dunque un terreno all'incrocio di diverse discipline e su cui è possibile pensare interventi di natura diversa che possono riguardare la promozione della concorrenza, la difesa della proprietà intellettuale dei dati personali, la protezione del consumatore, la difesa della privacy come diritto fondamentale.

3) IL RAGGIO D'AZIONE DELL'ANTITRUST

Alcuni recenti casi di acquisizioni da parte di grandi piattaforme, come quella di Whatsapp da parte di Facebook e di Doubleclick da parte di Google hanno generato molto dibattito sul ruolo possibile dell'Antitrust nel controllo di possibili abusi di posizioni dominanti da parte delle piattaforme che raccolgono grandi quantità di dati.

Il ragionamento sottostante grossomodo ipotizza che i meccanismi fondamentali di internet portino ad una progressiva concentrazione di mercato ad opera soprattutto dei dati personali raccolti che oltre una certa soglia renderebbero impossibile per un nuovo entrante replicare le capacità ed i dati dei leader.

La raccolta di dati personali sarebbe caratterizzata da economie di scala nel senso che maggiore è la raccolta di dati, maggiore è il valore che ne estraggo e da economie di scopo, dalla possibilità cioè di collegare dati di natura diversa ed insieme potrebbero ad effetti di rete che favoriscono una progressiva concentrazione. Secondo questa linea di ragionamento i dati potrebbero in prospettiva essere considerati come un essential facility.

Nella filiera dell'uso dei big data che comprende la raccolta, la memorizzazione e l'analisi, esistono molti effetti di rete diretti e indiretti che richiedono essere identificati dalle autorità antitrust. Questo richiede una comprensione delle interazioni e dei possibili feedback loop tra dati e altri input guardando all'insieme della filiera piuttosto che analizzando una singola parte separatamente (Bourreau, de Stree, Graef, 2017). Guardare alle quote di mercato è spesso fuorviante perché la concorrenza è principalmente dinamica e per il mercato. La recente storia di internet è piena di casi in cui un forte controllo sui dati rilevanti per quel mercato non hanno difeso la posizione di un incumbent contro un concorrente in grado di avere un'idea migliore. Così l'accumulo dei dati sulle search da parte del leader Yahoo non ha impedito che Google prevalesse grazie ad una migliore idea per il ranking dei siti. Facebook ha sbaragliato sia Friendster che Myspace perché organizzava il social network in modo più valido per gli utenti. I siti di incontri costituiscono un altro esempio perché si tratta di uno dei settori online più vecchi dove l'uso di dati personali è da sempre intenso per alimentare gli algoritmi di matching. Nel 2006 Yahoo! Personal era il leader indiscusso con un numero di utenti doppio di qualsiasi competitor. Ma invece di cavalcare una dominanza permanente abbandonò il mercato meno di quattro anni dopo (O'Connor, 2016). Secondo un sondaggio di Pew Research, 6 dei 10 siti di dating più popolari nel 2005 non esistevano più nel 2013. Oggi il mercato è dominato da pochi siti dominanti, ma ci sono oltre 2500 siti di dating online solo negli Stati Uniti. Gli effetti di rete sono stati superati dal multihoming degli utenti visto che molti mantengono il proprio profilo su parecchi siti, e dalla specializzazione verso una specifica fascia di pubblico.

Naturalmente non bisogna generalizzare questi esempi ed è possibile che in molte situazioni il controllo esclusivo di dati possa portare a forme di esclusione dei concorrenti.

Quando nel corso di una disputa commerciale tra Google e Hachette i risultati di ricerca relativi ai siti dell'editore sono spariti dalle prime posizioni del motore di ricerca è emerso il sospetto che anche i risultati naturali della search possano essere pilotati per interessi strategici.

Dalla rassegna della letteratura e dall'analisi dei diversi mercati coinvolti emerge come la valutazione degli effetti competitivi sia contingente e dipenda da molti fattori tra cui gli specifici effetti di rete e le interazioni tra online e offline. Si tratta dunque di valutare con attenzione caso per caso quale può essere il ruolo effettivo dei dati controllati e se questi costituiscano un ostacolo alla concorrenza.

Ma perché ci sia un vantaggio competitivo sostenibile nel tempo occorre che i concorrenti potenziali non siano in grado di replicare realisticamente i benefici di una strategia o di un input specifico (Lambert, Tucker 2015).

Questo si realizza solo raramente nel caso dei dati. Il fatto che i dati siano un bene non rivale fa sì che anche se il consumatore li ha ceduti ad un'impresa può senza problemi darli anche ad altri. Inoltre esiste un fiorente mercato intermedio per i dati di aggregatori che raccolgono dati da varie fonti e li rivendono alle imprese. Anche se dal punto di vista del consumatore questo mercato può essere discutibile e problematico, la sua presenza rende difficile un'azione pratiche escludenti attraverso i dati. L'uso dei dati richiede lo sviluppo di algoritmi specifici che iniziano ad avere forme di autoapprendimento e che talvolta possono rafforzare un vantaggio della prima mossa. Alcuni di questi algoritmi possono essere ottenute da terze parti, come avviene con il cloud computing per la potenza di calcolo, trasformando un costo fisso in un costo variabile e rendendo scalabile anche questo tipo di attività.

Un problema specifico riguarda il valore dei dati, sia per i consumatori che per le imprese che li raccolgono. Dai pochi lavori fatti su questo terreno risulta come per il consumatore il valore sia molto contingente e dipenda dalla fiducia nella controparte, dalle aspettative sull'uso futuro dei dati e dal framing con cui viene presa la decisione. Molti fattori sembrano indicare che la valutazione implicita è abbastanza bassa, vista la facilità con cui i dati vengono ceduti. Allo stesso tempo però la cessione dei dati come si è visto avviene senza che il consumatore abbia consapevolezza sugli usi successivi e sulle conseguenze anche su se stesso. Un esperimento recente mostra come la disponibilità a fornire dati sensibili si riduca drasticamente quando il termine privacy viene reso saliente, anche in un contesto positivo (Merriros, Tonin, Vlassopoulos, Schraefel, 2017). Questo sembrerebbe indicare che le persone sono spesso disattente e che le

preoccupazioni nascoste relative alla privacy emergono solo quando ai consumatori viene posta una domanda specifica.

Del resto non esiste un parametro oggettivo. Il riferimento dovrebbe essere il prezzo che il consumatore accetta per rilasciare i suoi dati, il prezzo che pagherebbe per proteggerli, o magari i costi nei quali il soggetto si aspetta di incorrere in caso di usi scorretto dei dati, oppure i profitti attesi generati dall'uso dei suoi dati personali?

Una posizione dominante ottenuta grazie al controllo dei dati personali può abilitare alcuni possibili abusi potenziali: la discriminazione di prezzo di primo grado, la pubblicità personalizzata e il degrado della qualità del servizio. Naturalmente nella valutazione di questi possibili comportamenti occorre anche valutare i vantaggi di efficienza e di miglioramenti dei prodotti ottenuti con l'uso dei big data.

Nella discriminazione di prezzo di primo grado il venditore, dotato di potere monopolistico, in grado di conoscere il prezzo di riserva di tutti i compratori e capace di prevenire l'arbitraggio è in grado di fissare per ogni consumatore un prezzo uguale al suo prezzo di riserva. L'effetto è tra l'altro quello di appropriarsi di tutto il surplus generato dalla transazione. Gli effetti di welfare di questa discriminazione di prezzo sono ambigui perché riuscendo a impedire che chi acquista a basso prezzo rivenda a chi è disposto a pagarne uno alto, il venditore può servire tutta la domanda con disponibilità a pagare superiore al costo marginale. D'altra parte però tutto il surplus del consumatore è incamerato dal venditore il che non pone direttamente problemi concorrenziali, ma sicuramente questioni sulla distribuzione del valore. D'altra parte la discriminazione di prezzo di primo grado implica anche un'espansione del mercato in quanto consumatori con bassa disponibilità a pagare vengono serviti. Sebbene in teoria una conoscenza approfondita dei consumatori e la scarsa trasparenza del mercato (ogni consumatore non vede i prezzi proposti agli altri) potrebbero consentire l'applicazione di prezzi personalizzati nella pratica questo strumento per il momento non sembra molto usato anche perché si scontra con un forte fastidio dei consumatori. Le esperienze sono spesso fatte in segretezza. Quando Amazon fu scoperta a proporre prezzi diversi a seconda dei consumatori dovette fare rapidamente marcia indietro. Generalmente vengono utilizzati sistemi più semplici, come la discriminazione temporale, nel caso delle linee aeree, oppure un pricing per gruppi ampi semplici da riconoscere.

Vista la contrarietà dei consumatori i prezzi personalizzati possono essere costruiti con meccanismi meno riconoscibili, come un prezzo uniforme e sconti personalizzati, oppure è possibile mostrare prodotti diversi ai consumatori di diversi gruppi in funzione della disponibilità a pagare. Naturalmente i sistemi di pricing complessi possono allentare la concorrenza e alzare i costi di ricerca dei consumatori.

La pubblicità personalizzata consiste nel mostrare un banner o un risultato di ricerca sponsorizzato ad un'audience specifica sulla base di caratteristiche o interessi personali. Dal punto di vista dell'inserzionista i vantaggi potenziali sono che l'annuncio viene mostrato ai clienti potenzialmente interessati riducendo sia i messaggi inutili che il fastidio del consumatore e in secondo luogo che il contenuto dell'annuncio può essere personalizzato, potenzialmente aumentando l'efficacia del messaggio pubblicitario (Bourreau, deStreel, Graef, 2017). Se la pubblicità personalizzata costituisce un miglioramento di efficienza e se non tutti i concorrenti sono in grado di realizzarla allo stesso livello allora ci si può aspettare un aumento della concentrazione e del potere di mercato, con un contestuale aumento di prezzo. La questione del prezzo però è un po' controversa perché normalmente la pubblicità targettizzata viene venduta a prezzi più elevati rispetto a quella tradizionale in funzione di quanto aumenta l'efficacia della pubblicità stessa. Per cui i potenziali guadagni di efficienza non vengono catturati dagli inserzionisti, che possono evitare di fare pubblicità inutile, ma dai venditori di contatti.

Dal punto di vista dei consumatori gli effetti di benessere possono essere differenti: da un lato ricevono annunci pubblicitari più vicini ai loro interessi e quindi potenzialmente più utili, ma d'altra parte i consumatori sembrano essere infastiditi perché si sentono sotto controllo. Spesso le piattaforme non consentono agli inserzionisti personalizzazioni molto spinte proprio per questo timore.

4 LA DEFINIZIONE DI MERCATO RILEVANTE

La definizione di mercato rilevante nella filiera dei big data presenta numerosi problemi per via della complessa articolazione verticale dell'attività e dei numerosi player coinvolti dove ogni operatore svolge contemporaneamente molti ruoli.

I dati possono avere molteplici ruoli nell'economia dei mercati digitali. Possono essere un prodotto in se, con un prezzo specifico, come avviene nella filiera dei mercati intermedi dove operatori e broker scambiano dati per completare le loro raccolte, oppure possono essere un input, che aiuta a definire le caratteristiche e i bisogni dei clienti oppure a ridurre il rischio, come ad esempio in molte transazioni finanziarie.

Quindi la definizione del mercato rilevante può prendere come riferimento solo i dati quando questi sono effettivamente il prodotto principale che viene scambiato. In questo contesto il mercato sembra operare in modo ragionevolmente concorrenziale con molti operatori, sebbene il tema dell'integrazione verticale, cioè di operatori che operano in strati multipli della catena del valore possa essere rilevante.

Nei mercati digitali la semplice analisi delle quote di mercato non dice molto sul potere di mercato consolidato dei principali operatori perché la concorrenza ha una forte componente dinamica e spesso la concorrenza avviene per il mercato piuttosto che nel mercato e come i primi decenni di vita di internet hanno mostrato i cicli di vita possono essere estremamente brevi con leader consolidati che vengono scalzati nel giro di pochi anni. La concorrenza potenziale spesso emerge da nuovi entranti in grado di rifinire i modelli di business con l'aggiunta di ingredienti a cui prima gli incumbent non pensavano, come nel caso di Huber, oppure semplicemente ridefinendo la funzione di produzione con una combinazione diversa di input, come nel caso di Google. L'ambiente digitale facilita questa flessibilità sia attraverso lo sviluppo di numerosi mercati intermedi per servizi specifici, come servizi di pagamento, data center, software specializzato, sia attraverso una generale riduzione dei costi di elaborazione e comunicazione dei dati, che nell'insieme hanno l'effetto di abbassare le economie di scala.

Questa flessibilità e questa articolazione verticale fanno sì che molti operatori svolgano molteplici ruoli sia verticalmente che orizzontalmente e che la loro collocazione strategica evolva rapidamente. Un tipico esempio è quello di Amazon che pochi avrebbero pensato dieci anni fa come concorrente di Microsoft o Google, ma che giocando sulla flessibilità della propria integrazione verticale, è entrata in mercati intermedi come il cloud computing diventandone uno dei leader. Oppure Apple che è contemporaneamente una piattaforma multiversante, un venditore di prodotti tecnologici e un fornitore di servizi IT, un operatore di contenuti multimediali.

Tra tutte queste attività ci sono frequentemente economie di scopo che talvolta sono company specific. La definizione del mercato rilevante va quindi fatta in funzione del caso specifico e non sempre si presta a generalizzazioni.

Molti mercati digitali sono caratterizzati da una struttura multiversante con piattaforme sia di tipo informativo che transazionale. Sebbene il concetto dei two sided market sia ormai entrato nella dotazione standard delle autorità antitrust, non sempre le analisi relative a casi che le riguardano vengono condotte tenendo conto fino in fondo delle implicazioni di questa caratteristica.

L'elemento chiave delle piattaforme è che costituito dalle esternalità che una transazione su un lato del mercato esercita sugli operatori e sulle transazioni negli altri versanti. Nell'esempio più tradizionale dei media un ipotetico innalzamento dei prezzi sul versante redazione abbassa il numero di lettori, ma allo stesso tempo questo rende meno attrattiva la piattaforma per gli investitori pubblicitari.

L'individuazione dei mercati rilevanti nei diversi versanti può essere complicata quando le piattaforme realizzano servizi gratuiti in cambio di dati che vengono utilizzati prevalentemente nel versante pubblicitario. Altrettanto complesso è la definizione quando i diversi operatori considerati hanno un portafoglio di attività multisided diverso tra loro che si sovrappongono solo parzialmente e che realizzano specifiche economie di scopo e sono caratterizzate da esternalità reciproche differenti.

Attraverso l'uso di dati personali le piattaforme possono offrire ai consumatori una molteplicità di servizi personalizzati, sussidiati e talvolta totalmente gratuiti, che rende difficile applicare lo SSNIP test che rimane uno degli strumenti standard per definire i confini del mercato.

La soluzione può essere quella di valutare il possibile degrado di qualità (ossia di una riduzione dei costi di produzione) che un operatore con potere di mercato può attuare, attraverso un test modificato come il SSQNDQ (small but significant non-transitory decrease in quality).

Il problema è che il degrado della qualità può essere attuato anche per altre ragioni, alcune sanzionabili dal punto di vista della disciplina della concorrenza.

Il degrado può avvenire in modo selettivo su alcuni gruppi di consumatori per sostenere politiche di discriminazione di prezzo. Per spingere gli utenti con alta disponibilità a pagare a prendere la versione più cara è possibile degradare la qualità del servizio per gli altri utenti, cosa relativamente semplice e poco costosa in ambiente digitale.

Oppure si può attuare un degrado di servizio verso clienti che utilizzano i servizi di concorrenti in altri versanti del mercato. Ad esempio un motore di ricerca potrebbe eliminare dalle prime posizioni i risultati di ricerca naturali di fornitori di contenuti con i quali non si è messo d'accordo per la ripartizione dei ricavi, ma questo sarebbe un caso di innalzamento dei costi di un concorrente, oppure di esclusione che si tratta bene con strumenti tradizionali delle autorità antitrust.

5 LO SPAZIO PER LE POLITICHE DI REGOLAZIONE

Insomma i big data possono essere facilmente degli amplificatori di altri vantaggi competitivi e pongono diversi problemi nella relazione tra cittadino e imprese, tra cittadino e pubblici poteri e nella creazione di asimmetrie informative che possono avere conseguenze negative sul funzionamento della società.

Vista la trasversalità del tema dei big data si pone il problema se occorrono sistemi di regolazione nuovi e di quali siano i contesti adeguati per costruirli. In particolare si ragiona se la protezione della concorrenza debba essere espansa per perseguire obiettivi di politiche pubbliche che riguardano l'equità degli scambi, la giustizia economica o diritti fondamentali come la privacy.

Credo che questa sia una strada poco promettente sia perché significherebbe distorcere sia strumenti che obiettivi dell'antitrust, sia perché operando ex-post su casi specifici non si presta a governare scelte di policy che potrebbero avere un carattere più generale.

Molte delle transazioni con gli operatori digitali sono caratterizzate da asimmetrie informative e i consumatori non sempre sono consapevoli del valore dei loro dati personali e dell'uso che ne verrà fatto successivamente.

D'altra parte negli ultimi anni la posizione contrattuale del consumatore è costantemente peggiorata nel senso che progressivamente si è trovato sulle spalle gran parte del rischio associato ad una qualsiasi transazione economica, dopo l'ondata di sviluppo di normative a difesa del consumatore che aveva sviluppato l'informazione sui prodotti, i diritti di garanzia e i diritti di recesso, limitando un po' il potere contrattuale arbitrario delle imprese.

Nel commercio tradizionale il consumatore seleziona il prodotto nel negozio, lo prova, ne prende possesso e a quel punto lo paga, anzi talvolta, col credito al consumo lo paga a rate successivamente.

Nel commercio elettronico è diventata pratica standard il pagamento anticipato dove il consumatore a quel punto subisce il rischio della mancata consegna, del prodotto difettoso, del prodotto non conforme di tutti i possibili ritardi. Lo scambio parte solo quando è realizzato il pagamento. Si tratta di uno spostamento di potere molto rilevante che non sempre è adeguatamente sottolineato.

Con la raccolta massiccia di informazioni personali la bilancia contrattuale può spostarsi ulteriormente a favore dei venditori che possono scegliere se attuare discriminazione di prezzo, mentre il consumatore fornisce inconsapevolmente un'informazione, privata sul prezzo di riserva, che poi sarà sfruttata per estrarre tutto il surplus generato dallo scambio.

Se le aziende hanno tutto il potere contrattuale e l'unica possibilità che ha il consumatore è quella di accettare le condizioni oppure non utilizzare servizi che spesso sono quasi indispensabili per le relazioni sociali.

Credo che nei prossimi anni si svilupperà una nuova ondata di misure di protezione del consumatore che tengano conto del nuovo orizzonte digitale e tendano a ribilanciare leggermente le asimmetrie informative e di potere.

BIBLIOGRAFIA

- Acquisti A., C.Taylor, L.Wagman, 2016, The economics of privacy, *Journal of Economic Literature* 54(2) pp 442-292
- Balto D., M.Lane, 2016, Monopolizing water in a Tsunami: finding sensible antitrust rules for big data, Burreau M., de Streel A., I.Graef, 2017, Big data and competition policy:market power, personalized prices and advertising, Project report, Centre on Regulation in Europe
- Colangelo G., 2016, Big data, piattaforme digitali e antitrust, *Mercato Concorrenza e Regole* XVIII N.3 pp 425-460
- Evans D., 2009, The online advertising industry: Economics evolution and privacy, *Journal of Economics Perspectives*, 23(3) 404-450
- Evans D. R.Schmalensee, 2012, The antitrust analysis of multisided platform business, Coase Sandor Working Paper Series, University of Chicago Law School
- Feri F., C.Giannetti, N.Jentzch, 2016, Disclosure of personal information undr risk of privacy shocks, *Journal of Economic Behavior & Organization*, 123 pp136-148
- Gambaro M. 2017, Concorrenza e pluralismo nel mercato di Internet: la prospettiva economica, in M.Bassani, O.Pollicino, T.Frosini, *Le libertà fondamentali nell'era di internet*, Milano, Mondadori,
- Gambaro M., R.Puglisi, 2015,What do ads buy? Daily coverage of listed companies on the Italian press, *European Journal of Political Economy*, N.39
- Gambaro M. 2013, Some economics of new media content production and consumption, and strategic implication for media companies, _in M. Friedrichsen and W. Mühl-Benninghaus, *Handbook of social media management*, Springer
- Goldfarb A. 2014, What is different about online advertising, *Review of Industrial Organization* 44(2) pp 366-376
- Henten A., I. Windekilde, 2015, Transaction costs and the sharing economy, 26th Conference of the International Telecommunication Society, San Lorenzo de El Escorial 24-27 June 2015
- Horton J., R.Zeckhauser, 2016, Owning, using and renting: some simple economics of the “sharing economy” NBER Working Per 22029
- Huberty M., 2015, Awaiting the second big data revolution: from digital noise to value creation, *Journal of Industry Competition and Trade* 15 pp 35-47
- O'Connor D., 2016, Understanding online platform competition: common misunderstandings,...
- Jamin R., A.O'Hara, 2015, Big data and the transformation of public policy analysis, *Journal of Polity Analysis and Management*
- Kerber W., 2016, Digital markets, data, and privacy: competition law, consumer law, and data protection, Joint Discussion PaPER Series in Economics N. 14-2016
- Marrienos H., M.Tonin, M.Vlassopoulos, M.C.Schraefel, 2017, “Now that you mention it” A survey experiment in information , inattention and online privacy, *Journal of Behavior & Organization* 140 pp 1-17
- Miller A., 2015, The dawn of big data monopolists, ...

Newman N., 2016, Search, antitrust and the economics of the control of user data, Yale Journal of Regulation vol 31.2 pp401-454

OECD, 2016, Big data: bringing competition policy to the digital era, Paris, Oecd

Pew Research, 2016, The state of privacy in post Snowden America,

Polo M., 2017, Consumer's search in the era of big data, Competition Policy International

Schudy S., V.Utkal, 2017, You must know about me – On the willingness to share personal data, Journal of Economic Behaviour & Organization 141 pp 1-13

Sokol D. R. Comereford, 2016, Does antitrust have a role in regulating big data?

Varian H. 2014, Big data: new tricks for econometrics, Journal of Economics Perspectives 28(2) pp3-28

Zervas G., D.Proserpio, J.Byers, 2016, The rise of the sharing economy: estimating the impact of Airbnb on the hotel industry, working paper