Conference Paper

Digital transformation : analysis of economic impact and potential

Quaderni del Premio «Giorgio Rota»

N. 8, 2020

Digital transformation: analysis of economic impact and potential

Centro di Ricerca e Documentazione Luigi Einaudi

Con il sostegno di

Fondazione CRT

Centro
di Ricerca
e Documentazione
Luigi Einaudi

Quaderni del Premio «Giorgio Rota»
n. 8, 2020

Digital Transformation: Analysis of Economic Impact and Potential

Iniziativa realizzata con il sostegno di

Fondazione CRT

Gli autori di questo Quaderno:

Antonio Aloisi
IE Law School, IE University, Madrid

Moreno Frau
Università degli studi di Cagliari, Corvinus University of Budapest

Leonardo Madio
Toulouse School of Economics, University of Toulouse; also affiliated with CESifo (Center for Economic Studies).

Martin Quinn
Católica Lisbon School of Business & Economics, Lisboa; also affiliated with CVPIP (the Chair Value and Policies of Personal Information).

Pietro Terna
Già Professore ordinario di Economia dell'Università di Torino e consigliere Centro Einaudi

Centro di Ricerca e Documentazione Luigi Einaudi
Corso Re Umberto 1 • 10121 Torino • segreteria@centroeinaudi.it
www.centroeinaudi.it

Quaderni del Premio «Giorgio Rota»
n. 8, 2020

Indice

Il Premio «Giorgio Rota»

L'intento del Premio *«Giorgio Rota» Best Paper Award* è di riprendere l'attività di ricerca annualmente condotta dal Comitato / Fondazione Giorgio Rota prima della sua inclusione nel Centro Einaudi, sulla relazione tra il pensiero e l'agire economico e un aspetto (ogni anno diverso) del vivere in società, mantenendo vivo il ricordo e l'insegnamento dell'economista Giorgio Rota, uno dei primi animatori del Centro, prematuramente scomparso.

Dal 2012 il Cento Einaudi ha dunque raccolto questa eredità rinnovando la formula della ricerca: è stato perciò istituito questo premio annuale dedicato a giovani ricercatori, con una qualificazione accademica nei campi dell'economia, sociologia, geografia, scienza politica o altre scienze sociali. I paper possono essere presentati sia in italiano che in inglese, e non devono essere stati pubblicati prima della data della Conferenza Rota, l'evento pubblico nel quale i vincitori hanno modo di presentare il loro lavoro.

La prima edizione aveva per tema *Contemporary Economics and the Ethical Imperative* e la Conferenza Giorgio Rota 2013 si è tenuta presso il Centro Einaudi il 25 marzo 2013 con keynote speech di Alberto Petrucci, LUISS Guido Carli, Roma.

La seconda edizione, nel 2013, è stata su *Creative Entrepreneurship and New Media* con Conferenza Giorgio Rota presso il Centro Einaudi, 14 aprile 2014 e keynote speech di Mario Deaglio, Università di Torino.

La terza edizione ha analizzato il tema *The Economics of Illegal Activities and Corruption*, con Conferenza Giorgio Rota presso il Centro Einaudi, 15 giugno 2015. Keynote speech di Friedrich Schneider, Johannes Kepler University (Linz, Austria).

La quarta edizione verteva su *The Economics of Migration*. Il 20 giugno 2016 si è tenuta la Conferenza Giorgio Rota presso il Campus Luigi Einaudi, in collaborazione con FIERI. Keynote speech di Alessandra Venturini, Università di Torino. Dal 2016 inoltre il Premio è sostenuto dalla Fondazione CRT.

La quinta edizione, del 2017, trattava di *Economic Consequences of Inequality*, e i saggi vincitori sono stati presentati alla Conferenza Giorgio Rota del 4 maggio 2017, tenutasi presso il Campus Einaudi in collaborazione con il Dipartimento di Economia e Statistica "Cognetti de Martiis". L'Introduzione è di Andrea Brandolini, Banca d'Italia.

La sesta edizione del Premio, tenutasi nel 2018, è incentrata sul tema *The Economics of Health and Medical Care*. I paper vincitori sono stati presentati alla Conferenza Giorgio Rota tenutasi il 1° giugno 2018 presso il Campus Einaudi, in collaborazione con il Dipartimento di Economia e Statistica "Cognetti de Martiis". L'Introduzione è di Fabio Pammolli, Politecnico di Milano.

La settima edizione del Premio è incentrata sul tema *Rural Economies, Evolutionary Dynamics and New Paradigms*. I paper vincitori, riportati qui, sono stati presentati alla Conferenza Giorgio Rota il 6 maggio 2019 presso il Campus Einaudi, in collaborazione con il Dipartimento di Economia e Statistica "Cognetti de Martiis". Gli autori, Federico Fantechi, Georgios Manalis e Stefano Menegat, sono introdotti da un intervento di Donatella Saccone, docente di Economia politica all'Università di Scienze gastronomiche di Bra

Digital Transformation: analysis of Economic Impact and Potential è il titolo dell'ottava edizione del Premio. I paper vincitori sono stati presentati alla Conferenza Giorgio Rota l'11 maggio 2020 che quest'anno si è tenuta online, in collaborazione con il Dipartimento di Economia e Statistica "Cognetti de Martiis". Gli autori, Antonio Aloisi, Moreno Frau, Leonardo Madio e Martin Quinn, sono stati introdotti alla Conferenza e nel volume da un intervento di Pietro Terna, Ex Professore ordinario di Economia dell'Università di Torino e consigliere Centro Einaudi.

Chi era Giorgio Rota

Giorgio Rota (1943-1984) è stato professore di Economia politica presso l'Università di Torino e consulente economico. Per il Centro Einaudi, è stato coordinatore agli studi e membro del comitato di direzione di «Biblioteca della libertà».
Le sue pubblicazioni scientifiche abbracciano diversi temi: l'economia dei beni di consumo durevoli, l'economia del risparmio, il mercato monetario e finanziario, l'inflazione e la variazione dei prezzi relativi, il debito pubblico. Ricordiamo tra esse: *Struttura ed evoluzione dei flussi finanziari in Italia: 1964-73* (Torino, Editoriale Valentino, 1975); *L'inflazione in Italia 1952/1974* (Torino, Editoriale Valentino, 1975); nei «Quaderni di Biblioteca della libertà», *Passato e futuro dell'inflazione in Italia* (1976) e *Inflazione per chi?* (1978); *Che cosa si produce come e per chi. Manuale italiano di microeconomia*, con Onorato Castellino, Elsa Fornero, Mario Monti, Sergio Ricossa (Torino, Giappichelli, 1978; seconda edizione 1983); *Investimenti produttivi e risparmio delle famiglie* (Milano, Il Sole 24 Ore, 1983); *Obiettivi keynesiani e spesa pubblica non keynesiana* (Torino, 1983).

Tra le sue ricerche va particolarmente citato il primo *Rapporto sul risparmio e sui risparmiatori in Italia* (1982), risultato di un'indagine sul campo condotta da BNL-Doxa-Centro Einaudi, le cui conclusioni riscossero notevole attenzione da parte degli organi di stampa. Da allora il *Rapporto sul risparmio,* ora *Indagine sul risparmio*, continua a essere pubblicato ogni anno.

Quaderni del Premio «Giorgio Rota»
n. 8, 2020

Pietro Terna

Cyber Markets: What about Economic Freedom?

We start with a possible subtitle to clarify the contents of this presentation: "The good and the bad of the economic revolution coming from the web".

In the presentation, I will move from cyber markets to planning to agent-based simulation, artificial intelligence from the perspective of the markets' behavior.

First of all, the dawn of the second part of the last century's novelties, with the magic moment of the middle of the 40s, when great minds like John von Neumann, Oscar Morgenstern, and John Nash, lead to critical new emergencies. Morgenstern was an economist, not a computer scientist or a mathematician, but his role was vital in explaining to Neumann the economic reality.

We had there new calculation tools, and a new language for science and social science, with game theory and the new concept of complexity. But we have not to forget Wiener, who was a mathematician and a philosopher with the creation of cybernetics. Cybernetics it is the attempt of joining the analysis of machines and humans both behaving with connections, with similarities. This element is essential in my presentation because the idea of planning comes from there. After all, cybernetics was not only related to intelligence – now we name artificial intelligence most of the parts of the cybernetic studies – but it was also a fundamental organizational analysis.

From there, the idea of searching new tools for planning, putting together cybernetics, and the input-output tables construction to understand the connections operating within an economic system. Microeconomic data, or big data, are essential, but we were at the beginning of the second part of the last century, quite far from the current situation.

To plan an economy, we need the data and, most of all, to decentralize both the collection and the utilization of data. We have an excellent book and a superb paper of Gerovitch (2004, 2008) written at the beginning of this century, where we can find underlined the critical points towards transforming an economy. An important date was 1961 when the Cybernetics Council of the Soviet Academic of Science published a volume (Berg 1961-1962) whose title is essential: Cybernetics at the service of communism.

We repeat that we consider cybernetics both as the study of humans and machines' behavior and vice versa and both as an organizational tool. The book is significant, and you have a link to a large part of its contents at https://terna.to.it/CybCom/. From there, the Soviet government had the possibility of planning the economy. Still, the process never started working effectively.

Why? Did they lack computers? For technological problems? Probably not. The most significant obstacle to starting the new planning activity was in designing the technical and organizational choices and in power contrasts. To create a whole economy with central planning, it is necessary to proceed step by step, with trials and errors and learning adaptation, and not to operate top-down trying the create a unique applied planning system.

If you want to know deeply, what the proposal was in 1961, you can read the 1962 translation made by the Department of Commerce of the United States. All the world was paying a lot of attention to this massive experiment to these profound changes. The book is practically impossible to have. To my knowledge, there are few copies in the world in five libraries. I asked a European library to have the possibility of reading the book, and I obtained it for a few weeks. I don't think I have done a lousy action putting a part of its chapters online.

If you want to have a light knowledge of the same arguments, you can read the Spufford (2010) book: it is not a technical book, but it is fascinating to have a broad picture of that period, where to place also the starting activity of planning.

Why am I dedicating a large part of my introduction to Soviet planning to talk about the web and the economic activity? We are close to discovering why.

Planning cannot work in a world without valid prices. As a great economist, Enrico Barone, wrote in Italian in 1908 with the title *Il ministro della produzione nello stato collettivista* or *The ministry of production in the collectivistic state* (Barone 1908a; 1908b; 2012). All this was before the creation of the first collectivistic state. Barone was one of the "three of Lausanne", with Léon Walras and Vilfredo Pareto. He was a mathematician, and his work is formally grounded. His key sentence is that "The determination of the coefficients economically most advantageous can only be done in an experimental way: and not on a small scale, as could be done in a laboratory; but with experiments on a very large scale, because often the advantage of the variation has its origin precisely in a new and greater dimension of the undertaking". Pay attention, all this in 1908.

Now we have other tools as agent-based simulation and artificial intelligence. The agent-based simulation is my field of research. In the last 30 years, we started building models composed of small parts of code. These are computerized models, but, in my view, they are a part of the mathematical models. Each piece of the code represents an agent with articulated characteristics and capabilities in a heterogeneous construction. In this way, we can observe the artificial agents' behavior in a metaphorical space to analyze the emergent macro-level effects.

Another step: artificial intelligence. Is artificial intelligence already a concrete reality? My reply is yes. I use as support the wonderful incipit of a paper of Kasparov (2018), in «Science», after the world chess championship of 2018. In Kasparov's words, that was not the contest between the two strongest players of the planet but only between the two strongest humans.

For Kasparov, chess is the *drosophila*, the fruit fly, in some way easy to examine. From there, more complicated reasoning operations can be undertaken by artificial intelligence. Certainly, chess is *drosophila* as we cannot imagine using artificial intelligence to ask for philosophical constructions' basic answers. In a more straightforward but not so simpler field, such as economics, I guess that machine learning could quite soon produce direct analyses, and maybe it is already doing them.

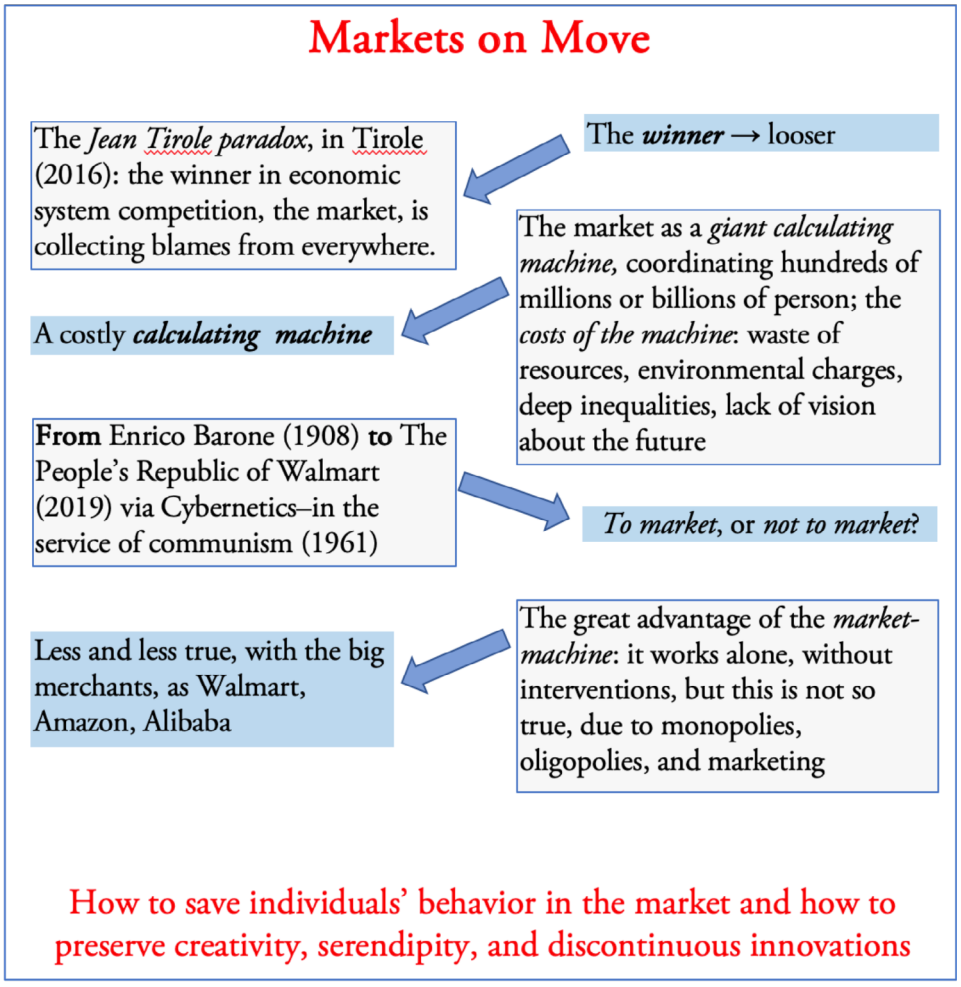
From cybernetics, agent-based simulation, artificial intelligence, are suggestions also arising for planning toward markets? Are we still living in a free market context? Maybe not. A consideration of mine: perhaps you have noticed that Amazon can send us in a few hours non-common use objects. In which way? Forecasting our probable decisions and feeding the warehouses whit the goods that we have not still ordered but that forecasts say we will buy. If a colossal operator decides what to buy, it is not far from determining how much to produce. This kind of action is already planning. If it also chooses those productions' selling prices, it is not far from planning the economy.

In which way? From «The Economist» (2019), we have another exciting incipit: each year, Amazon asks all the managers to explain how they plan to use machine learning, and a reply "not so much" is not appreciated. Machine learning to prepare what to do as merchants and a lot more than merchants. Amazon is starting to have preferred exclusive producers, so it is moving to the industry.

Summarizing in Fig. 1 about artificial intelligence power and markets, what is emerging?

A network of AI capabilities that helps reduce transportation overload, waste of resources, energy needs, environment damages, inequalities, or… a world of fighting or colluding oligopolies, managing markets, and directing consumptions?

FIGURE 1 • A FLOWCHART REPORTING THE DISCUSSION OF THIS NOTE IN A SCHEMATIC WAY



The reply is open.

References

Barone E. (1908a), *Il ministro della produzione nello stato collettivista*, «Giornale degli Economisti», 19, 37, pp. 267-293, http://www.jstor.org/stable/23222736
– (1908b), *Il ministro della produzione nello stato collettivista (continuazione)*, «Giornale degli Economisti», 19, 37, pp. 391-414, http://www.jstor.org/stable/23221778
– (2012), *The ministry of production in the collectivist state*, «Giornale degli Economisti e Annali di Economia», 125, 71, pp. 75-112 (Italian edition 1908), http://www.jstor.org/stable/43828055
Berg A.I. (ed.) (1961-1962), *Cybernetics at the service of communism* - USSR (translation published by Office of Technical Services, U.S. Department of Commerce, Washington D.C., 1962), vol. I, Moscow-Leningrad, Cybernetics Council of the Soviet Academy of Sciences, 1961
Gerovitch S. (2004), *From newspeak to cyberspeak: a history of Soviet cybernetics*. MIT Press.
– (2008), *InterNyet: Why the Soviet Union did not Build a Nationwide Computer Network*, «History and Technology», 24, 4, pp. 335-350, http://web.mit.edu/slava/homepage/articles/Gerovitch-InterNyet.pdf
Kasparov G. (2018), *Chess, a Drosophila of reasoning*, «Science», 362, 6419, pp. 1087, DOI: 10.1126/science.aaw2221, http://science.sciencemag.org/content/362/6419/1087.full
Spufford F. (2010), *Red Plenty*, London, Faber & Faber; (2012), Minneapolis, Graywolf Press.
«The Economist» (2019), *Amazon's empire rests on its low-key approach to AI*, Apr 11th 2019, https://www.economist.com/business/2019/04/13/amazons-empire-rests-on-its-low-key-approach-to-ai?frsc=dg%7Ce

ANTONIO ALOISI[1]

HIERARCHIES WITHOUT FIRMS? VERTICAL DISINTEGRATION, OUTSOURCING AND THE NATURE OF THE PLATFORM

Abstract. New forms of labour intermediation through digital platforms such as *Uber*, *Deliveroo* or *Amazon Mechanical Turk* can be conceptualised as the latest stage of a long-lasting process of disaggregation of the firm and "disorganisation of labour law." In particular, the rise of platform-mediated work can be seen as an instantiation of deliberate business strategies aimed at outsourcing labour while retaining intense and pervasive managerial prerogative. The phenomenon is exacerbating several unresolved tensions inherent in the contemporary world of work, let alone the perverse impact that "platformisation" is having on precariousness and social inequalities.

In short, new technologies allow platforms to abandon traditional methods of workplace governance and adopt a stronger version of the "command and control" logic. Direct interaction is replaced by a significant reliance on information communications technology: workers are monitored more closely and intimately than they ever used to be by means of tech tools, including algorithms, artificial intelligence and customers' reviews. This leads to the question whether the existing concept of "firm" is appropriate to face this new reality, whether minor or major adaptations may be necessary or whether we need a total re-invention of the underlying assumptions of the employment regulation.

After describing the theoretical antecedents of hierarchical outsourcing, the article explores the literature on the nature of "non-standard forms of firm" by applying transaction-cost economics. In an attempt to update the incomplete trichotomy among "hierarchies," "markets" and "networks," I present a complementary model combining pre-existing schemes. Finally, by building on theories unfolding the disarticulation of the formal employing entity and the pulverisation of work-related responsibilities, this paper demystifies the prototypical business model of rampant socio-economic actors in the on-demand economy.

Keywords. Labour platforms, transaction costs, business model, employment law, digital transformation.

[1] This paper builds on the first chapter of my doctoral dissertation completed for the Ph.D. in Legal Studies at Bocconi University, Milan. An earlier draft of this paper was presented at the 16th Marco Biagi International Conference, Modena, Marco Biagi Foundation, in March 2018. I am extremely grateful to Stefano Liebman, Valerio De Stefano and Miriam A. Cherry for great discussion and invaluable feedback. I am also grateful to Piera Loi, Brishen Rogers and Nastazja Potocka-Sionek for their helpful comments.

1. INTRODUCTION

We are all witnessing radical changes in the world of work (and in the corresponding legal fields), fuelled by globalisation, tertiarisation and digitalisation. Structural shifts have remodelled the internal structure of the firm and the work organisation. More importantly, they challenge the underlying assumptions of the employment relationship. Shifts in the labour market differ in their legal implications, yet in most cases they can be disentangled by looking at the interplay among new organisational patterns, contractual arrangements and, not least, power relationships. Therefore, it is worthwhile to complement and perhaps renew the copious studies on the "future of work" with a thorough analysis of changing forms of organisation and business models, a rather neglected topic.

To this end, one could use the illustrative case of labour intermediation through digital platforms, the most recent blatant manifestation of a long-lasting process of dissolution of the unitary firm and "disorganisation of labour law" (Valdés Dal-Ré 2002). The hallmark of the on-going digital transition is the use of technological channels to distribute one-off and low-income jobs through a local/global chain (Rogers 2015). The reliance on short-term assignments provided in a "just-in-time" fashion and compensated on a "pay-as-you-go" basis has a strong impact on the formal organisation of the employing entity and, above all, on the relationship between employers (requesters) and employees (providers).

As I demonstrate below, new technologies such as smart machines, artificial intelligence and online platforms allow abandoning the traditional method of workplace governance and adopting a stronger version of "command-and-control" logic (Taylor 1911). Direct control is replaced by a significant reliance on digital devices and software for coordination (Sprague 2007). Workers are supervised more closely and intimately than they ever used to be. Thanks to the "glue of the creation, monitoring, and enforcement of standards on product and service delivery, made available through new information and communication technologies" (Weil 2014: 9), online platforms impose stringent standards on nominally independent workers hired on the spot for specific tasks (Finkin 2016).

Commentators describe modern firms such as *Uber, Deliveroo* or *Amazon Mechanical Turk* as unparalleled or unprecedented organisations situated between hierarchies and markets or, even better, transcending these two orthodox options. Self-proclaimed "disruptive" companies act as online parasitic "middlemen" by

lowering information asymmetries as well as agents' opportunism, minimising organisational costs and engaging a pool of self-employed workers (virtually recruited, effectively organised and persistently disciplined) through instant commercial transactions with an authoritative attitude (De Stefano and Aloisi 2018). The combination of affordable broadband, algorithmic governance, geo-location widgets, machine learning and other wonders of information communications technology (ICT) has blurred the confines between the two classical alternatives – "make" or "buy" (Rubery and Wilkinson 1981) – more deeply than previous experimentations with corporate governance and lean organisations (Zarkadakis 2018).

This hybrid form is often used to avoid the obligations and costs associated with employment status. This latest wave of ICT-enabled outsourcing and deregulation calls for deeper scrutiny, since it is rebooting the firm-boundary problem and reshaping our conceptions. What is left out of the story is the impact of this ongoing revolution on the governance structure and the internal organisation of the firm. This article analyses the shared features of the most common business model adopted in the platform economy, by focusing on the disintegration of the employing entity and the "pulverisation" of employment-related obligations. The article is organised as follows. The next section reviews the fundamentals of transaction-cost economics (TCE). Section 3 analyses the principal trends in the gig economy, while Section 4 conceptualises the "Cerberus firm" as a combination of pre-existing models. Section 5 is concerned with the extent to which transaction cost theory can still explain the platform business model. Section 6 concludes.

2. ORTHODOX TAXONOMIES, TRANSACTION COSTS AND THE DIGITAL AGE

Economists, lawyers, organisational theorists, and business historians have long wrestled with the need to explain the firm's governance structure and internal workings (Salento 2003). Why should "islands of conscious power" arise in the surrounding "ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk" (Robertson 1923, cited in Coase 1937: 386)? This section re-examines some of the classical drivers of firms' decisions about internal organisation from a law-and-economics perspective, using Coase's and

Williamson's key insights on the "economic institutions of capitalism" to elucidate why firms can still derive full benefit from vertical integration in the "second machine age" (Brynjolfsson and McAfee 2014).

Transaction cost theory unpacks the decision-making processes determining the "efficient boundaries" of an organisation, defined as an optimal balance between activities completed within and outside the permeable borders of the firm (Stone 2004; Piore and Sabel 1984). According to Coase (1937), who first grasped the principle, transaction costs are minimised within the firm because formal bureaucratic power replaces time-consuming negotiation and price-mechanisms governance in the market. Transaction costs are defined as costs incurred for (i) obtaining reliable information, (ii) bargaining terms and conditions of the relevant contract and (iii) monitoring and enforcing the agreement (Williamson 1981). If these costs are prohibitive, firms bypass the markets by internalising production. According to Chandler (1977), businesses grow by bringing activities *within* the firm in order to optimise transaction costs and exercise upstream authority over resources. In classical and neoclassical economic theory it is well known that, when asset specificity, uncertainty and frequency are high (Simon 1955 and 1991), firms may find it more convenient to grow in a vertically integrated fashion, establishing a non-market governance system.

Coase (1937: 395) brilliantly observed that "[w]ithin a firm, market transactions are eliminated and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-coordinator, who directs production." The exercise of managerial prerogative is made possible by labour regulation and facilitated by the formal existence of an accepted hierarchy. Notably, this private governance structure can also explain the key economic functions of the employment relationship, a legal tool allowing firms to curb transaction costs by reducing the need to constantly search for and select providers, obtain their consent, negotiate terms and conditions and enforce them (in other words, the processes of resourcing, transacting and contracting). The notion of the firm as a "command hierarchy" implies the concept of employment, just as the concept of self-employment implies the notion of the market.

As Baronian emphasises (2020: 217), authority and hierarchy "lower transaction costs related to the contractual relation of employment." The increase in "subordination costs" (also known as organisational costs) is compensated by the possibility of exercising managerial prerogative and hierarchical power instead of

specifically negotiating each task through costly and lengthy transactions. Thanks to a unique scheme that "encapsulates" a set of developmental rules and conditions (Aloisi and De Stefano 2020), the employee accepts the authority of the firm and follows orders issued by managers in a given "zone of acceptance" (Simon 1951), thus avoiding the need for contracting every time from scratch. The employment relationship grants management essential organisational prerogatives: (i) the power to assign tasks and give instructions to workers; (ii) the power to control and assess the execution of such tasks; and (iii) the power to sanction non-compliant workers. The employment contract – which is the typical contractual scheme in the hierarchy model – is a perfect example of an "incomplete contract," an agreement that leaves some terms and conditions unspecified within a given framework of programmability. Its inherent flexibility represents a potent vehicle for integration.

To sum up, a firm is vertically integrated when market costs outweigh the internal costs of quick, robust administrative choices. As a result, the hierarchical firm can be far more efficient than the market structure, which may fail to coordinate production effectively and distribute resources optimally.

Undeniably, leading scholars have long since proposed unorthodox responses to the binary divide between "make" and "buy" – in particular, identifying networks as a very elastic way of coordinating economic activities. Networks – intermediate governance structures based on reciprocal, relational, mutually supportive actions – commonly involve aspects of dependency and indeterminacy in co-evolving ecosystems (Goetz and Scott 1981). Powell (1990: 301, 296) argues that the network model can be used "to make progress in understanding the extraordinary diversity of economic arrangements found in the industrial world" and that "the familiar market-hierarchy continuum does not do justice to the notion of network forms of organization."

Holmström and Roberts (1998) noted that many firms decide in favour of cooperation, rather than integration.[2] But seen through the prism of transaction cost economics (Lamoreaux *et al.* 2003), these relationships end up "imitating" the organisation of the centralised firm, or at least some of its defining characteristics (in particular, organisational power), thus building a hierarchy based on external resources rather than on internal ones (Hart and Moore 2005). In a context of

[2] In the Italian experience of *"distretti"* interactions among firms were aimed at promoting the development of specific ties of cooperation in a network (Moretti 2012).

formal independence and stable cooperation, one party dominates the other "interdependent" firm (Klein *et al.* 1978; De Stefano 2009).

It could be argued that conventional theories of governance and organisation cannot capture hierarchical forms of outsourcing in the digital age. In particular, the traditional theory describes integration as an inevitable result of asset specificity, underestimating the future consequences of technological development (Holmström and Milgrom 1994). Arguably, digital transformation and market specialisation could challenge large-scale vertical integration and uphold interconnected forms of governance (Brynjolfsson *et al.* 1994). An influential article anticipated that "by reducing the costs of coordination, information technology will lead to an overall shift toward proportionately more use of markets – rather than hierarchies – to coordinate economic activity […]" (Malone, Yates and Benjamin 1987: 484). Muehlberger (2005: 4) concludes that inventive firms benefit from an ambiguous situation characterised by "incentives (typically linked to market transactions) and control (typically adopted in the bureaucratic model)." Downsides are less evident. Employers may face difficulties in dealing with a segmented, relatively uncommitted and inharmonious workforce, supervising isolated workers operating outside the firm's premises while meeting customers' needs for quality and reliability.

Against this background, the last decade is likely to be remembered for the rapid rise of "platform-mediated work," a "newer" form of employment in which a digital infrastructure facilitates matching labour demand with supply and organises work performance by means of guidelines, ratings and other internal proxies driven by algorithms and artificial intelligence (Ivanova *et al.* 2018). As Tomassetti (2016) explains, the result is an apparently "win-win situation" in which firms control resources without owning them, rapidly adapting to downturns in the market, thanks to "a set of calls on resources that are then assembled into a performance" (Davis 2015: 502). Thanks to multiple commercial contracts, the employer has access to a large workforce while avoiding obligations under labour law and social security.

Outsourcing can be deliberately used to disguise the need to obtain a large pool of workers, abating sunk costs and assembling a flexible organisation. Replacing the employment contract with commercial contracts significantly reduces salaries, turning the wage-setting issue into a mere contracting decision. Needless to say, this shift results in a failure to implement clauses laid down in the applicable

collective agreements. Therefore, these processes have been treated with fundamental disapproval, but also with insatiable curiosity and inevitable delay by labour lawyers (Corazza 2004).

If "pipeline businesses" built on departments, lines of authority, reporting mechanisms, and formal decision-making processes were well suited for production and distribution before the digital era (Van Alstyne, Parker and Choudary 2016), in an "always-connected" scenario the theory of an "economy in which firms [are] featured as islands of planned co-ordination in a sea of market relations" is called into question (Richardson 1972: 895). This is an issue of mounting importance in times of digital disruption: is the alternative between "market" and "hierarchy" still useful? To put it bluntly, the level of efficiency reached by new tech infrastructure can lower transaction costs and reduce frictions, making it easier and more convenient for firms to resort to complex and interdependent market relations to acquire "labour energies," instead of relying on vertical and accountable structures based on employment relationships (Williamson 1985). In short, the digital transformation is adding new impetus to the discussion on "what firms are and what they do" (Foss and Klein 2019), questioning the basic "make-or-buy" divide.

3. THE GIG-ECONOMY IS ANYTHING BUT COLLABORATIVE. TAKING "PLATFORMISATION" (MORE) SERIOUSLY

Are we on the verge of seeing the definitive eclipse of the firm as we know it? Is this the future of work to which we are headed? It is undeniable that the global labour market faces the threat of tremendous "platformisation" in all industries and latitudes (Corporaal and Lehdonvirta 2017). App- and platform-based firms have the potential to become dominant providers of a large number of services, shrinking the firm and redesigning its notion and shape (Allen, Root and Schwede 2017). This trend could lead to organisations that are "fluidly assembled and re-assembled from globally networked labor markets" (Kessler 2017).

Despite various differences, labour platforms share one common characteristic. They very effectively mobilise, organise and dispatch a flexible, volatile and scalable workforce, significantly reducing transaction costs and information asymmetries for both clients and firms thanks to the efficient use of digital tools (Edelman and Geradin 2016). Featuring an "at arm's length" pattern built "as-

needed," they break down jobs into small pieces and assign them to the lowest-bidding or, alternatively, highest-ranked worker – usually self-employed, with very limited access to labour and social security protection. Gig economy workers are excluded from many rights and benefits afforded to employees, including minimum wage, paid sick leave, parental leave, overtime pay, protection against unfair dismissal, compensation for occupational illness or injury, contributions to health insurance and retirement, and the freedom to organise and bargain collectively (De Stefano and Aloisi 2019).

By collecting a large amount of data and enforcing exclusivity clauses, many platform companies are adopting a *"fait accompli"* strategy, asking for forgiveness rather than permission (Garben 2017). Opportunistically, platforms select which rules they comply with and, often, contravene labour law principles or skip out on regulations, claiming that out-dated constraints should not hinder forward-looking innovation.

Platforms offer an indefinite "crowd" of precarious workers, making it cheap and easy to outsource; thus they fall into the vast category of tools tearing down the boundaries of the firm, promoting the engagement of external resources in lieu of stable employment relationships. By nature, they are built as "connecting hubs" ("brokers" or even "marketplace," according to their terms of service; see Hwang and Elish 2015; Aloisi 2016) where responsibilities are diluted. Collins's (1990) prediction about the transition "from mass production to networks of smaller business geared to rapid response to change in consumer taste" (356) is thus coming true.

From this viewpoint, the ability to create an "asset-light" enterprise out of existing relations is empowering a disintegrated form of organisation. Accordingly, researchers have generally agreed that the granitic notion of the firm has been redefined, to what has come to be known as the "*entreprise sans travailleurs*" ("firm without workers"), a temporary "network of individuals" specialised in coordinating funding, production and commercialisation (Malone and Laubacher 1998; Drahokoupil and Fabo 2016). In this respect, it is vital to differentiate between genuine innovations brought about by managerial decisions and restructuring processes that are merely aimed at circumventing labour and social security provisions.

Digital labour platforms represent a formidable example of centralised or hierarchical forms of outsourcing, because, "[b]y mixing governance structures, [they] are able to benefit from the advantages of outsourcing without losing

control over labour and assets" (Muehlberger 2005: 4). In addition, platforms seek total control even if they shed responsibility, by consolidating authority structures resembling those common in employment relationships such as setting goals and deliverables, monitoring and evaluating work, providing feedback and imposing sanctions on reluctant workers. Thus, digital devices "are being used in ways that are designed managerially and (il)legally to evade employment status and thereby social and legal entitlements" (Medland *et al.* 2019: 3).

First, matching infrastructures make it simple to recruit the best suited candidate; second, "taskification" aggravates an extreme substitutability of workers, as very little commitment is needed for performing one-off activities; third, these extemporaneous micro-tasks can be allocated efficiently and reassembled at a later stage, if needed. Technology, in fact, can decrease the unit costs of coordination, by extending technical control and making it more penetrating (Munger 2015; Aloisi and Gramano 2020). Transaction costs can be reduced drastically by using modern instruments: (i) information can be obtained through people analytics and consumer reviews (Bodie *et al.* 2017); (ii) fares and other terms are stipulated "algorithmically" on-the-spot by apps taking into account all relevant factors; (iii) the electronically observed failure to follow guidelines, recommendations and instructions may constitute a breach of the participation agreement, leading to automatic expulsion.

As Aloisi and De Stefano (2020) have argued, many modern firms want to have it both ways. They exercise an employer's degree of control over the workforce model without being held accountable as employers (Spicer 2018). New players in the platform economy have invented a rather distorted picture of flexible innovation, based on cost-cutting, risk-shifting and the selective application of legal provisions (regulatory and contractual arbitrage).

4. THE RISE OF THE *"CERBERUS"* FIRM, A PLURAL AND EFFECTIVE COMBINATION OF PRE-EXISTING MODELS

Even if transaction-cost economics has been foundational for most thinking about management, it might seem that "the business model of digital platforms has practically refuted the theoretical framework of TCE" (Baronian 2020: 229). However, TCE "still unites the thinking of academics, consultants and managers, and it still underpins most subjects taught in business schools. And there is a good

reason for that: the old narrative is still largely correct" (Foss and Klein 2019). While most of the existing explanations of the efficient perimeter of a firm have focused mainly on material items and commodities, the general principles of transaction-cost economics apply to both physical assets and workers.

Companies strive to be flexible, specialised and innovative to face unexpected changes. After defining the main sources of the competitive advantage (i.e., organisational strengths), theorists decompose organisations into their key components and subsegments. Platforms can be seen as aggregations of specialised entities with complementary interests – expanding and reconfiguring themselves in a way that best adapts to or even anticipates changing market dynamics. Their fragmented and "fissured" structure optimises contractual flows by adapting the zero-inventory model to workforce governance and slicing the organisation into its smallest components (DiMaggio 2009; Weil 2019).

While it is true that "traditional command-and-control management is becoming less common [since] decisions are increasingly being pushed lower down in organisations" (Malone and Laubacher 1998: 47), at the same time, firms are still relying on a centralised form of coordination and upstream power. Indeed, efficiencies are achieved "as a result of firm integration, of replacing the market exchange activities (or inter-firm transaction costs) […] with agency cost (intra-firm costs)" (Tomassetti 2016: 28). It could be aptly pointed out that platform companies have reduced transaction costs between the platform and its users, not between workers/providers and users. Like firms, they rely on labour to extract value and exercise their control power over their workforce;[3] like markets, they dispatch and connect nominally independent actors; like networks, they match and synchronise demand and supply of services by facilitating interdependence and creating value for both sides of the transaction (even if the vast bulk of the value is captured by the platform).[4]

Although the institutional taxonomy is an effective instrument for classifying the different models of (standard) firms, there are infinite intermediate options

[3] As is demonstrated in C-434/15 *Asociación Profesional Elite Taxi v. Uber Systems Spain, SL* (2015), there is a gap between rhetoric and reality. Indeed, Uber interferes in the discrete task by setting the price, arranging the trip and potentially excluding workers who are caught in breach of the relevant terms and conditions. In 2017, the Court of Justice of the European Union observed that "Uber determines at least the maximum fare by means of the eponymous application, […] receives that amount from the client before paying part of it to the non-professional driver of the vehicle, and […] it exercises a certain control over the quality of the vehicles, the drivers and their conduct, which can, in some circumstances, result in their exclusion" (ECLI:EU:C:2017:981). It is a system where "with great power comes virtual freedom" (Aloisi 2018).

[4] For a detailed analysis of network effects and multi-sided markets, see Zhu and Iansiti (2019).

along the spectrum from a centralised hierarchy to dispersed networks. Instead of an on-off toggle, we might think of a composite scale with movable switches (Grimshaw *et al.* 2005). This allows us to argue that platforms are (i) firms when it comes to exercising command-and-control prerogatives, as the authority mechanism can be enforced by vertical relational contracts; (ii) markets when it comes to treating workers as independent providers, avoiding subordination costs through commercial agreements; and (iii) an immaterial, modular infrastructure relying on "network effects" when it comes to allocating products and services by leveraging the number of users (Srnicek 2016; Cohen 2017).

This "non-standard form of firm" (Lo Faro 2017) can be also seen as a combination of elements pertaining to both "hierarchy" (a vertical structure with a traditional configuration and a classic organism based on "intra-firm contracts") and "the market" (inter-firm contracts). Platforms "replace the 'spontaneous' 'autonomous adjustments' of supply and demand from price signals with 'consciously coordinated adaptations' of centralised production" (Tomassetti 2016: 23). Like Cerberus, the mythological three-headed monstrous dog, platforms are multi-headed economic players that are likely to metastasise from transaction enablers to participation gatekeepers (Malone, Yates and Benjamin 1987). This is why I use the seemingly contradictory formula "hierarchies without a firm."

This sort of "hybrid" among market, hierarchy and network (the Cerberus firm, see Table 1) combines hierarchical organisations and interdependent models at the core area of the business, "while highly temporary market relations continue to predominate on the periphery," thus facilitating "a correspondingly (more) rapid change in the institutional arrangement" (Sydow and Helfen 2016: 2). The table shows how strong authority mechanisms and liquid responsibilities can go hand in hand.

TABLE 1 • STYLISED COMPARISON OF FORMS OF ECONOMIC ORGANIZATION

	Forms			
Key features	*Market*	*Hierarchy*	*Network*	*Cerberus firm*
– normative basis	Contract - property rights	Employment relationship	Complementary strengths	Contract - property rights
– means of communication	Prices	Routines	Relational	Relational
– methods of conflict resolution	Haggling - resort to courts for enforcement	Administrative fiat - supervision	Norm of reciprocity - reputational concerns	Supervision, norm of reciprocity - reputational tie
– degree of flexibility	High	Low	Medium	Low, nominally high
– amount of commitment among the parties	Low	Medium to high	Medium to high	Medium to low
– tone or climate	Precision and/or suspicion	Formal, bureaucratic	Open-ended, mutual benefits	Formal, bureaucratic
– actor preferences or choices	Independent	Dependent	Interdependent	Interdependent

How can "hierarchies without firms" be as effective as traditional highly integrated firms? The proliferation of vertical decomposition has cast doubt on whether entrepreneurs can succeed in running an efficient business while eschewing the powers granted to the formal employer (Marglin 1974). The simple answer is that they do not eschew those powers; they merely delegate them to algorithmic governance or automatic review mechanisms composed (mostly as 'work made for hire') by human programmers, at the direction of human bosses – software that can effectively manage, monitor, and consequently discipline performance execution. This organisational arrangement replaces middle managers with seemingly neutral, objective technology, thereby decoupling managerial power from protective obligations (Aloisi and De Stefano 2020).

5. THE PLATFORM BUSINESS MODEL DOES UBERISATION REDEFINE THE NOTION OF THE FIRM?

Undoubtedly, the increasing relevance of the service-based sector and the crucial role played by digitalisation may have heralded a new era of post-

industrialism; of course, authority is exercised in ways that differ from the caricature of the "command-and-control" approach.[5] But no, "Uberisation" does not redefine the notion of the firm – it merely hides the shift from a bureaucratic control to a more sophisticated, technocratic and invasive one (Yung 2005). As Foss and Klein (2019) explain, "the basic idea of a firm, the nature of ownership and responsibility, and how people coordinate tasks are the same as always." To this extent, the apparent success of the narrative describing the final eclipse of Taylorism is far from justified. Platforms rely on the sharp separation of design, management and execution. While claiming to definitively overcome strict protocols, they embody and fully exploit the principles of scientific management, implementing the crucial aspects of the traditional division of labour in a rather voracious and predatory way (Lomba 2005).

As I note above, it would be misleading to look at labour platforms as a unique monolith. Platforms have many dissimilarities; there is no such thing as a functional uniformity. Nevertheless, they share some hallmarks that are crucial for the design of a system of "organised irresponsibility" (Collins 2015; Countouris and Ratti 2018).

The interrelationship among actors could be described as triangular (or multi-party), as the platform (which controls intellectual property rights and governance) also connects between buyers ("requesters," according to the internal terminology) and workers ("sellers" or "providers"). Although the model resembles the one of temporary work agencies, this way of arranging a digital business blatantly denies the existence of an employment relationship, thus "undermining the regulatory framework envisaged for three-way relationships" (Potocka-Sionek 2020: 187). Unquestionably, it is more convenient for clients and employers to engage workers task by task rather than hiring them as employees. This peculiar model allows platforms to deploy managerial prerogatives over a contingent workforce mobilised by means of formal and informal contracts, thus responding to demand peaks and shifting the impact of fluctuations and uncertainty onto the worker's shoulders. This is the source of the platforms' considerable cost advantage.

Platforms exploit the massive use of advanced information technology, typically a combination of widespread broadband, a user-friendly digital application and increasingly effective tools, such as geo-localisation via GPS and management by algorithms, to facilitate transactions and keep the distribution lean (Womack, Jones

[5] This paragraph draws upon De Stefano and Aloisi (2018).

and Roos 1990; McGaughey 2018). In addition, platform work can be considered as a promising laboratory of new practices of people analytics, management by algorithm and gamification. By relying on customer-based feedback systems for quality checks that can be handled seamlessly through electronic interfaces, they externalise some control functions.

Platforms constitute a promising example of a two- or multi-sided market (Evans and Schmalensee 2016; Rochet and Tirole 2006; Evans 2003; Katz and Shapiro 1985). One side is made of clients who benefit from access to low-cost services while supplying the platform with data; the other side is made of clients who may also benefit from positive network externalities (Valenduc and Vendramin 2016). Platforms also benefit from the fact that workers must use their own equipment (personal computers, bicycles or cars, whether leased or owned) to provide a service (Telles 2016). On a closer inspection, this model of vertical outsourcing has existed for decades. What is new is the penetration of infrastructure that determines frictionless transactions, not to mention the quantitative leap and exponential growth in data and metrics that, collected, refined and analysed, can "train" the internal algorithm, making matching and governance even prompter and more successful (Valenduc and Vendramin 2016).

The basic structure can be found in completely different sectors, replicating the original model of a hiring hall or a virtual bulletin board such as Craiglist or eBay, which are advanced databases (Autor 2001). Platforms generate value by simplifying and supporting the interplay between providers and users/consumers. Each successful interaction guarantees a significant transaction fee to the platform. At the same time, platforms are able to avoid high fixed costs as well as to shed variable costs of production, which results in large economies of scale. These business relationships "also transform fixed costs into variable ones" (Muehlberger 2005: 3).[6] This is how economies of scope can be combined with economies of scale and specialisation (Golzio 2005), leading to a high-performing model of hierarchical outsourcing. At the same time, the firm expects employees to offer commitment without getting loyalty in return, thus changing "the implicit contract between the employee and the firm" (Stone 2005: 118).

Consequently, a small but growing body of research on the implications of non-standard firms has suggested that it is important to be able to orchestrate

[6] According to an article published in the magazine «TechCrunch», "Uber, the world's largest taxi company, owns no vehicles. Facebook, the world's most popular media owner, creates no content. Alibaba, the most valuable retailer, has no inventory. And Airbnb, the world's largest accommodation provider, owns no real estate" (Goodwin 2015).

processes, solicit participants, and interact fruitfully with the surrounding ecosystem. This is only partly true. In combination, the contingent nature of the relationship and the reliance on procuring (as opposed to developing internally) the skills that the firm needs significantly misalign the interests of employer and employees with regard to the development of key competences (in labour economics terms, "firm-specific human capital") and new skills. Marginal workers will remain so unless they develop "specialised not specific" skills that can be used outside the firm, assuming that firms do not require the same level of loyalty and commitment from all workers (Deckop, Mangel and Cirka 1999; Killick 1995). This may also have a negative and statistically significant effect on productivity (Lindbeck and Snower 1988; Boeri and Garibaldi 2007).

In short, according to business literature, platforms perform three specific functions: (i) match workers with employers/clients, (ii) provide a common set of tools and widgets that enable the delivery of work in exchange for money, (iii) set governance rules according to which good actors are rewarded and poor behaviour is discouraged. As Nick Srnicek (2017: 48) puts it, "[p]latforms, in sum, are a new type of firm; they are characterized by providing the infrastructure to intermediate between different user groups, by displaying monopoly tendencies driven by network effects, by employing cross-subsidization to draw in different user groups, and by having designed a core architecture that governs the interaction possibilities" in a fluid way.

Contrary to what usually happens in value chain models, platforms make profits as the ecosystem expands in a circular and iterative progression. Network effects increase proportionally with the growing number of participants on one side of the market (direct effects) or the opposite side (indirect effects); that is why online platforms may support one side of the network.[7] What makes a platform distinctive is the ability to capture and utilise information about its massive network of customers and suppliers (Birkinshaw 2018; Uber Technologies Inc. 2019). But platform companies "share with all other kinds of capitalist firms the relation of production based on the property and use of non-human assets by capitalists who extract living labor in order to valorize these social means of production" (Baronian 2020: 229). In short, and contrary to the widespread narrative on disruptive tech, there is no significant difference between the nature of the firm and the nature of the platform, at least from an organisational and

[7] Most platforms aim at quickly capturing network externalities and becoming monopolies (Schmidt 2017).

legal viewpoint (Henten and Windekilde 2015). Instead of advocating a partial abrogation of labour law to unleash innovation, we need to understand the broader picture in which "innovative firms" are situated.

6. FINAL REMARKS

Platforms should be understood as non-standard firms that style themselves as networks of market-based contracts, yet use both technological means and pure market power to dictate work rules in great detail, and to organise, control and discipline workers through distributed mechanisms (Edward 1980). Contrary to the industry's claims, by taking advantage of either new technology or new labour demographics or new patterns of production and consumption (Hyman 2018), these powers closely resemble managerial powers without being surrounded by the regulation essential to mitigate them (Aloisi 2018; Prassl 2018).

Vallas and Schor (2020: 10) have recently explained that, as distinctive organisations, platforms "incorporate many of the features of prior economic structures selectively" by retaining authority over important functions while ceding a little control over others. In order to stay competitive, they may look for ways to get rid of the presumed constraints of labour law and social security (Griswold 2016). On the one hand, this model represents a sort of "throwback to the industrial model, incorporating the efficiency and control of automatic management, without the industrial model's job security or stability" (Cherry 2016: 27). On the other, the use of non-standard arrangements makes it easier for platforms to gain a competitive advantage, as they face a much smaller regulatory burden than their competitors do. That is, despite the linguistic "sophistry" (Lobel 2018), the common business model in the platform economy combines features and functions belonging to classical models. Indeed, hierarchies, markets and networks are far from opposing and mutually exclusive forms of organisation.

In addition, "as platforms mature, vertical integration is growing" (Gapper 2019) to meet consumers' expectations of accountability and standardisation. Concomitantly, many gig companies rely on standard contracts and flexible schedules. Several cases demonstrate that the platform economy "can comfortably coexist with the legal determination of an employment status" and employee rights (Aloisi and De Stefano 2020: 56). Concerns that regulation will drive

platforms or new companies out of business would therefore seem to be overblown, much like earlier arguments that regulation would end various aspects of the "Fourth Industrial Revolution" (Cherry and Aloisi 2017). As a result, from a regulatory standpoint calls for special and differentiated treatment should not be supported or tolerated.

This spectacular intermingling of old and new challenges explains why platform-based work continues to catalyse so much attention. Transaction cost theory and the traditional repertory of coordinative mechanisms expose the unchanged power structure in the highly unstable gig-economy and explode the fervid myth of the novelty of online platforms. In analogy to what has already been said by Powell (1990), new types of coordination of economic activity represent a combination of existing models. Companies such as *Uber, Deliveroo* or *Amazon Mechanical Turk* retain authority, centralise power, consolidate control and develop ties among selected participants. The ascendancy of such new "geometry" in contemporary capitalism may prove misleading (Kornberger *et al.* 2017). Indeed, platforms do not disrupt the demarcation between alternative models; rather, they reinforce the implicit theory while proposing definitional hybrids that are not always a true reflection of reality.

REFERENCES

Allen J., Root J. and Schwede A. (2017), *The firm of the future*, Bain & Co., Brief

Aloisi A. (2016), *Commoditized workers. Case study research on labor law issues arising from a set of 'on-demand/gig economy' platforms*, «Comparative Labor Law & Policy Journal», 37, 3, pp. 653-690

– (2018), *'With great power comes virtual freedom'. A review of the first Italian case holding that (food-delivery) platform workers are not employees*, «Comparative Labor Law & Policy Journal», Dispatch n. 13

Aloisi A. and De Stefano V. (2020), *Regulation and the future of work: The employment relationship as an innovation facilitator*, «International Labour Review», 159, 1, pp. 1-23

Aloisi A. and Gramano E. (2020), *Artificial intelligence is watching you at work. Digital surveillance, employee monitoring and regulatory issues in the EU context*, «Comparative Labor Law & Policy Journal», 41, 1, pp. 101-127

Autor D.H. (2001), *Wiring the labor market*, «The Journal of Economic Perspectives», 15, 1, pp. 3-30

Baronian L. (2020), *Digital platforms and the nature of the firm*, «Journal of Economic Issues», 54, 1, pp. 214-232

Birkinshaw J. (2018), *How is technological change affecting the nature of the corporation?*, «Journal of the British Academy», 6, 1, pp. 185-214

Bodie M.T., Cherry M.A., McCormick M.L. and Tang J. (2017), *The law and policy of People Analytics*, « University of Colorado Law Review», 88, pp. 961-1042

Boeri T. and Garibaldi P. (2007), *Two tier reforms of employment protection: A honeymoon effect?*, «The Economic Journal», 1171, 521, pp. F357-F385

Brynjolfsson E. and McAfee A. (2014), *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York, W.W. Norton & Co.

Brynjolfsson E., Malone T.W., Gurbaxani V. and Kambil A. (1994), *Does information technology lead to smaller firms?*, «Management Science», 40, 12, pp. 1628-1644

Chandler A. (1977), *The Visible Hand: The Managerial Revolution in American Business*, Cambridge, MA, Harvard University Press.

Cherry M.A. (2016), *Beyond misclassification: The digital transformation of work*, «Comparative Labor Law & Policy Journal», 37, 3, pp. 544-577

Cherry M.A. and Aloisi A. (2017), *'Dependent contractors' in the gig economy: A comparative approach*, «American University Law Review», 66, 3, pp. 635-689

Coase R.H. (1937), *The nature of the firm*, «Economica», 4, 16, pp. 386-405

Cohen J.E. (2017), *Law for the platform economy,* «U.C. Davis Law Review», 51, 1, pp. 133-204

Collins H. (1990), *Independent contractors and the challenge of vertical disintegration to employment protection laws*, «Oxford Journal of Legal Studies», 10, 3, pp. 356-380

– (2015), *A Review of* The Concept of The Employer *by Dr Jeremias Prassl,* November 10, https://bit.ly/2Lp6PcG

Corazza L. (2004), "*Contractual integration" e rapporti di lavoro. Uno studio sulle tecniche di tutela del lavoratore*, Padova, Cedam

Corporaal G.F. and Lehdonvirta V. (2017), *Platform sourcing: How fortune 500 firms are adopting online freelancing platforms*, Oxford, Oxford Internet Institute, https://www.oii.ox.ac.uk/publications/platform-sourcing.pdf

Countouris N. and Ratti L. (2018), *The sharing economy and EU anti-discrimination law,* in N. Davidson, M. Finck and J. Infranca (eds), *The Cambridge Handbook of the Law of the Sharing Economy* (pp. 486-498), Cambridge (UK) - New York, Cambridge University Press

Davis G.F. (2015), *What might replace the modern corporation: Uberization and the web page enterprise*, «Seattle University Law Review», 39, 2, pp. 501-516

De Stefano V. (2009*), Smuggling-in flexibility: Temporary work contracts and the "implicit threat" mechanism*, ILO Working Document, 4

De Stefano V. and Aloisi A. (2018), *European legal framework for digital labour platforms.* Luxembourg, Publications Office of the European Union, https://publications.jrc.ec.europa.eu/repository/bitstream/JRC112243/jrc112243_legal_framework_digital_labour_platforms_final.pdf

– (2019), *Fundamental labour rights, platform work and human rights protection of non-standard workers*, in J.R. Bellace and B. ter Haar (eds), *Labour, Business and Human Rights Law* (pp. 359-379), Cheltenham, Edward Elgar

Deckop J.R., Mangel R. and Cirka C.C. (1999), *Getting more than you pay for: Organizational citizenship behavior and pay-for-performance plans*, «Academy of Management Journal», 42, 4, pp. 420-428

DiMaggio P. (ed.) (2009), *The Twenty-first-century Firm: Changing Economic Organization in International Perspective*, Princeton (NJ), Princeton University Press

Drahokoupil J. and Fabo B. (2016), *The platform economy and the disruption of the employment relationship*, ETUI Policy Brief, 5

Edelman B.G. and Geradin D. (2016), *Efficiencies and regulatory shortcuts: How should we regulate companies like Airbnb and Uber*, «Stanford Technology Law Review», 19, 2, pp. 293-328

Edwards R. (1980), *Contested Terrain: The Transformation of the Workplace in the Twentieth Century*, New York, Basic Books - London, Heinemann.

Evans D.S. (2003), *The antitrust economics of multi-sided platform markets*, «Yale Journal on Regulation», 20, 2, pp. 325-381

Evans D.S. and Schmalensee R. (2016), *Matchmakers. The New Economics of Multisided Platforms*, Boston, Harvard Business School Press

Finkin M. (2016), *Beclouded work in historical perspective*, «Comparative Labor Law & Policy Journal», 37, 3, pp. 16-12.

Foss N. and Klein P. (2019), *No boss? No thanks*, January 14, https://bit.ly/2WWVdVd

Gapper J. (2019), *Technology platforms are losing control*, May 22, https://on.ft.-com/2KwG3hX

Garben S. (2017), *Protecting workers in the online platform economy: An overview of regulatory and policy developments in the EU,* EU-OSHA Discussion Paper

Goetz C.J. and Scott R.E. (1981), *Principles of relational contracts*, «Virginia Law Review», 67, 6, pp. 1089-1150

Golzio L. (2005), *L'evoluzione dei modelli organizzativi d'impresa*, «Diritto delle Relazioni Industriali», 15, 2, pp. 312-323

Goodwin T. (2015), *The battle is for the customer interface,* March 4, https://tcrn.ch/-2QYBVeo

Grimshaw D., Marchington M., Rubery J. and Willmott H. (2005), *Introduction: Fragmenting work across organizational boundaries*, in M. Marchington, D. Grimshaw, J. Rubery and H. Wilmott (eds), *Fragmenting Work: Blurring Organizational Boundaries and Disordering Hierarchies* (pp. 2-38), Oxford (UK), Oxford University Press

Griswold A. (2016), *The Uber economy looks a lot like the pre-industrial economy*, October 1, https://qz.com/806117

Hart O.D. and Moore J. (2005), *On the design of hierarchies: Coordination versus specialization*, «Journal of Political Economy», 113, 4, pp. 675-702

Henten A. and Windekilde I. (2015), *Transaction costs and the sharing economy*, Paper presented at the 26th European Regional Conference of the International Telecommunications Society (ITS): 'What Next for European Telecommunications?', Madrid, Spain, 24th-27th June, 2015, International Telecommunications Society (ITS), Calgary

Holmström B. and Milgrom P. (1994), *The firm as an incentive system*, «The American Economic Review», 84, 4, pp. 972-991

Holmström B. and Roberts J. (1998), *The boundaries of the firm revisited*, «Journal of Economic perspectives», 12, 4, pp. 73-94

Hwang T. and Elish M.C. (2015), *The mirage of the marketplace*, July 27, https://bit.ly/-3bmIKgV

Hyman L. (2018), *Temp, How American Work, American Business, and the American Dream Became Temporary*, New York, Viking

Ivanova M., Bronowickav J., Kocher E., Degner A. (2018), *The app as a boss? Control and autonomy in application-based management,* Arbeit, Grenze, Fluss – Work in Progress interdisziplinärer Arbeitsforschung Nr. 2. Frankfurt, Viadrina European University

Katz M.L. and Shapiro C. (1985), *Network externalities, competition, and compatibility*, «American Economic Review», 75, 3, pp. 424-440

Kessler S. (2017), *Forget the on-demand worker: Stanford researchers built an entire on-demand organization*, July 13, https://qz.com/1027606

Killick T. (1995), *The Flexible Economy: Causes and Consequences of the Adaptability of National Economies,* London, Routledge

Klein B., Crawford R.G. and Alchian A.A. (1978), *Vertical integration, appropriable rents, and the competitive contracting process*, «Journal of Law and Economics», 21, 2, pp. 297-326

Kornberger M., Pflueger D. and Mouritsen J. (2017), *Evaluative infrastructures: Accounting for platform organization*, «Accounting, Organization and Society», 60, pp. 79-95

Lamoreaux N.R., Raff D.M. and Temin P. (2003), *Beyond markets and hierarchies: Toward a new synthesis of American business history*, «The American Historical Review», 108, 2, pp. 404-433

Lindbeck A. and Snower D.J. (1988), *The Insider-Outsider Theory of Employment and Unemployment*, Cambridge (MA), Harvard University Press.

Lo Faro A. (2017), *Core and contingent work: A theoretical framework*, in E. Ales, O. Deinert and J. Kenner (eds), *Core and Contingent Work in the European Union: A Comparative Analysis* (pp. 7-23), Oxford (UK), Hart Publishing.

Lobel O. (2018), *Coase and the platform economy*, in N. Davidson, M. Finck and J. Infranca (eds), *The Cambridge Handbook of the Law of the Sharing Economy* (pp. 67-77), Cambridge (UK) - New York, Cambridge University Press

Lomba C. (2005), *Beyond the debate over 'post-' vs 'neo-' Taylorism: The contrasting evolution of industrial work practices*, «International Sociology», 20, 1, pp. 1001-1074

Malone T.W. and Laubacher R.J. (1998), *The dawn of the e-lance economy*, «Harvard Business Review», 76, 5, pp. 145-152

Malone T.W., Yates J. and Benjamin R.I. (1987), *Electronic markets and electronic hierarchies*, «Communications of the ACM», 30, 6, pp. 484-497

Marglin S. (1974), *What do bosses do?*, «Review of Radical Political Economics», 6, 2, pp. 60-112

McGaughey E. (2018), *Taylorooism: When network technology meets corporate power*, «Industrial Relations Journal», 49, pp. 459-472

Medland I., Anderson B., Bales K., Bogg A., Novitz T., O'Connell Davidson J., Pitts Frederick H., and Turnbull, P. (2019), *The 'future' of work? A call for the recognition of continuities in challenges for conceptualising work and its regulation*, Law Research Paper 001

Moretti E. (2012), *The New Geography of Jobs*, Boston, Houghton Mifflin Harcourt

Muehlberger U. (2005), *Hierarchies, relational contracts and new forms of outsourcing*, ICER Working Paper n. 22

Munger M. (2015), *Coase and the 'sharing economy'*, in C.G. Veljanovski (ed.), *Forever Contemporary: The Economics of Ronald Coase* (pp. 187-208), London, The Institute of Economic Affairs

Piore M. and Sabel C. (1984), *The Second Industrial Divide: Possibilities for Prosperity*, New York, Basic Books

Potocka-Sionek N. (2020), *The changing nature of labour intermediation. Do algorithms redefine temporary agency work?* in J. Wratny and A. Ludera-Ruszel (eds), *New Forms of Employment*, (pp. 169-190), Wiesbaden, Springer VS.

Powell W.W. (1990), *Neither market nor hierarchy: Network form of organization,* in B.M. Staw and L.L. Cummings (eds), *Research in Organizational Behavior*, 12, pp. 295-336, Greenwich, JAI Press

Prassl J. (2018), *Humans as a Service: The Promise and Perils of Work in the Gig Economy*. Oxford (UK), Oxford University Press

Richardson G.B. (1972), *The organisation of industry*, «The Economic Journal», 82, 327, pp. 883-896

Robertson D.H. (1923), *The Control of Industry*, London, Nisbet & Co.

Rochet J.C. and Tirole J. (2006), *Two-sided markets: A progress report*, «Rand Journal of Economics», 37, 3, pp. 645-667

Rogers B. (2015), *The social costs of Uber*, «University of Chicago Law Review Dialogue», 82, pp. 85-102

Rubery J. and Wilkinson F. (1981), *Outwork and segmented labour markets*, in F. Wilkinson (ed.), *The Dynamics of Labour Market Segmentation* (pp. 115-132), London, Academic Press

Salento A. (2003), *Postfordismo e ideologie giuridiche: nuove forme d'impresa e crisi del diritto del lavoro*, Milano, Franco Angeli

Schmidt F.A. (2017), *digital labour markets in the platform economy, mapping the political challenges of crowd work and gig work*, Bonn, Friedrich-Ebert-Stiftung Study, http://library.fes.de/-pdf-files/wiso/13164.pdf

Simon H.A. (1951), *A formal theory of the employment relationship*, «Econometrica», 19, 3, pp. 293-305

– (1955), *A behavioral model of rational choice*, «The Quarterly Journal of Economics», 69, 1, pp. 99-118

– (1991), *Organizations and markets,* «Journal of Economic Perspectives», 5, 2, pp. 25-44

Spicer A. (2018), *No bosses, no managers: The truth behind the 'flat hierarchy' façade,* July 30, https://bit.ly/3dvX7RU

Sprague R. (2007), *From Taylorism to the Omnipticon: Expanding employee surveillance beyond the workplace*, «The John Marshall Journal of Computer & Information Law», 25, 1, pp. 1-36

Srnicek N. (2017), *Platform Capitalism*, New York, Polity Books

Stone K.V.W. (2004), *From Widgets to Digits: Employment Regulation for the Changing Workplace*, New York, Cambridge University Press

– (2005), *Flexibilization, globalization and privatization: The three challenges to labor rights in our time*, «Osgoode Hall Law Journal», 44, 1, pp. 77-104

Sydow J. and Helfen M. (2016), *Production as a service: Plural network organisation as a challenge for industrial relations*, Bonn, Friedrich-Ebert-Stiftung Study, http://library.fes.de/pdf-files/iez/12903.pdf

Taylor F.W. (1911), *Principles of Scientific Management*, New York, Harper & Brothers Publishers

Telles R. Jr. (2016), *Digital matching firms: A new definition in the "sharing economy",* Space, Economics and Statistics Administration Issue Brief, 1

Tomassetti J. (2016), *Does Uber redefine the firm? The postindustrial corporation and advanced information technology*, « The Hofstra Labor & Employment Law Journal», 34, 1, pp. 1-78

Uber Technologies, Inc. (2019), *Initial Public Offering*, Document S-1, April 11

Valdés Dal-Ré F. (2002), *Decentralisation productive et desorganisation du droit du travail*, in A. Lyon-Caen and A. Perulli (eds), *Trasformazione dell'impresa e rapporti di lavoro: atti del seminario dottorale internazionale*, *Venezia 17-21 giugno*, Padova, Cedam, pp. 53-71

Valenduc G. and Vendramin P. (2016), *Work in the digital economy: Sorting the old from the new*, WP ETUI, n. 3

Vallas S. and Schor J.B. (2020), *What do platforms do? Understanding the gig economy*, «Annual Review of Sociology», 46, pp. 16.1-16.2

Van Alstyne M.W., Parker G.G. and Choudary S.P. (2016), *Pipelines, platforms, and the new rules of strategy*, «Harvard Business Review», 94, 4, pp. 54-62

Weil D. (2014), *The Fissured Workplace: Why Work Became So Bad for So Many and What Can Be Done to Improve It*, Cambridge (MA), Harvard University Press

– (2019), *Call Uber and Lyft drivers what they are: employees*, July 5, https://lat.ms/2RlDG6t

Williamson O.E. (1981), *The economics of organization: The transaction cost approach*, «American Journal of Sociology», 87, 3, pp. 548-577

– (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, New York, Free Press.

Womack J.P., Jones D.T. and Roos D. (1990), *Machine That Changed the World*, New York, Rawson Associates.

Yung J. (2005), *Big Brother is watching: How employee monitoring in 2004 brought Orwell's 1984 to life and what the law should do about it*, «Seton Hall Law Review», 36, pp. 163-222

Zarkadakis G. (2018), *Do platforms work?,* May 28, https://bit.ly/2A8RBWJ

Zhu F. and Iansiti M. (2019), *Why some platforms thrive and others don't,* January-February https://hbr.org/2019/01/why-some-platforms-thrive-and-others-dont

MORENO FRAU

DIGITAL TRANSFORMATION BEHAVIORS IN THE AGRI-FOOD CONTEXT: AN EXPLORATORY ANALYSIS

Abstract. Firms cannot avoid digital transformation and nothing could be more important for them than exploit the digital data created by new technologies. Yet, we still lack a clear understanding of how firms, especially in agri-food industries, are transformed by digital technologies. This article contributes to an understanding of how agri-food firms behave in distinct stages of digital transformation and shows how digitalization enablers influence these behaviors. This research also empirically demonstrates that digital data exploitation behaviors change according to the data sources employed by the firms when trying to develop their products

Keywords. Digital transformation, digital data, firm behavior, multiple case study.

1. INTRODUCTION

Digital transformation is dramatically changing the face of the economy (Matzler *et al.* 2018). Digital transformation is the "application of new technologies […] [which] requires skills that involve the extraction and exchange of data as well as the analysis and conversion of that data into actionable information." (Schallmo, Williams and Boardman 2017: 4). Compared with other economic and social transformations, the digital transformation gives rise to an ever-growing quantity of native digital data which are *"born digital"* that means not subsequently entered in the information system by hand or digitized by computer tools (e.g. scanner) (Piccoli and Watson 2008). Therefore, digital transformation and digital data are having an increasing impact on the development of firms' competitive advantages (Piccoli and Ives 2005).

Previous studies on the digital transformation focused on some features of the phenomenon, such as digital transformation strategies (Ferreira, Fernandes and Ferreira 2019; Hess *et al.* 2016); alteration of the business model (Berman 2012); adoption of new technologies (Pankewitz 2017); data creation, collection, and analysis (Dremel *et al.* 2017). Although these studies have investigated the critical aspects of the digital transformation, they were mostly developed in contexts in

which the use of technologies was already established and in which the observation of the phenomenon was favored. Nowadays, the digital transformation is one of the strongest environmental influence that forces all industries to adapt to the changes it introduces, this is also true for the agri-food sector (Anastasiadis, Tsolakis and Srai 2018; Vlachos 2004). Despite the study of digital transformation in agri-food firms could provide new insights due to the traditional low level of technologies adoption, there is little empirical research that examines how agri-food firms are digitally transformed (Hess *et al.* 2016; Loonam *et al.* 2018).

Conversely, scholars who have intertwined technological innovation and food production (e.g. (Beckeman, Bourlakis and Olsson 2013; Grunert *et al.* 2008), on one hand, have significantly explained how agri-food firms made use of technologies and what new food products have been developed due to the new technologies (e.g. Leek, Szmigin and Carrigan 2001; Marette *et al.* 2009; Steenis and Fischer 2016). On the other hand, they missed the opportunity to investigate the increasing availability of digital data and how the information gathered by the elaboration of such data can foster product development (Schweitzer, Handrich and Heidenreich 2019).

These gaps inspire the research questions, "How are agri-food firms digitally transformed?" and "How do these firms exploit digital data to develop their products?".

By using an exploratory multiple case study design, this study provides two main contributions. First, this article contributes to an understanding of how agri-food firms behave in distinct stages of digital transformation and shows how digitalization enablers influence these behaviors. Second, this research empirically demonstrates that digital data exploitation behaviors change according to the data sources employed by the firms when trying to develop their products.

1.1 How digital transformation has been changing firms

Digital transformation is a complex phenomenon that affects several areas within a company. It has received greater research attention only recently so that it represents the salient topic in the current research agenda. Prior conceptualization of digital transformation are several (Olleros and Zhegu 2016; Schallmo, Williams and Boardman 2017), resulting in fragmented definitions across studies (Loonam *et al.* 2018). Digital transformation is defined as an organizational transformation

that integrates digital technologies and business processes in the digital economy (Liu, Chen and Chou 2011). Digital transformation requires reenergizing the business to benefit from digital technology (Bowersox, Closs and Drayer 2005). Specifically, this means shaping customer relationships, internal processes, and value propositions to exploit firms' main competence through the adoption of digital technology (e.g. analytics, mobility, social media, and smart devices) and gain competitive advantage (Brynjolfsson and Hitt 2000). While Schallmo and colleagues (2017) suggest a definition of digital transformation that recognizes the need to develop skills for collecting and analyzing digital data to convert them into information. So, how has the digital transformation been changing firms? Previous research mainly focused on: 1) digital transformation strategies (Ferreira, Fernandes and Ferreira 2019; Hess *et al.* 2016); 2) changes in the business model (Berman 2012); 3) adoption of new technologies (Pankewitz 2017); 4) data creation, collection, and analysis (data circle) (Dremel *et al.* 2017).

Regarding the *strategies for facing the digital transformation*, the factors that push firms to develop (or not) new digital processes and their implications in terms of innovation and performance received close attention (Ferreira, Fernandes and Ferreira 2019). Digital transformation is often associated with disruption. Even traditional and big old companies are not immune to the disrupting changes driven by digital transformation (Loonam *et al.* 2018). Thus, traditional firms need to learn from disruptive ventures and reimagine their business models, processes and products trying to strengthen them through the use of digital technology (Matzler *et al.* 2018; Sebastian *et al.* 2017). Other companies are exploiting the agile principles to facilitate the cultural and technical changes required by the digital transformation (Shaughnessy 2018). These strategy adaptations could avoid firms being affected by the disruptive effects of digital transformation (Matzler *et al.* 2018).

Digital transformation often triggers *changes in the business model*. Literature provides a morphology of the business model transformation before and after 2000, documenting the drivers of the changes (Kotarba 2018). However, a systematic approach for developing business models in the context of digital transformation seems missing (Schallmo, Williams and Boardman 2017). A possible plan for modeling the digital transformation consists of identifying existing products and services, deconstructing business models and discovering new configurations (Remane *et al.* 2017). Digital transformation creates an

opportunity to mold new customer-oriented business models grounded in the online customers' engagement at every link of the value chain (Berman 2012).

The *adoption of digital technology* presents differences due to the particular industry, in which it is applied. Examples of digital technology adoption can be the development of a simple e-platform for the digitalization of traditional services (Fisher *et al.* 2000; Liu, Chen and Chou 2011; Sunding and Zilberman 2001), or the use of more complex technologies such as social, mobile, analytics, cloud and Internet of things (IoT). However, the adoption of digital technology can provide unique opportunities as well as existential threats (Sebastian *et al.* 2017). Regarding the opportunities, a series of technologies (e.g. automation, robots, algorithms, and artificial intelligence) has great potential of disrupting not only the industry where they appear but also alike businesses and sectors (Pankewitz 2017).

Once equipped with digital technologies, firms can *generate, collect, and analyze digital data*. Digital transformation dramatically boosted the creation of native digital data which are "born digital" that means, not later entered in the information system manually or digitized after data creation by computer tools (e.g. scanner) (Piccoli and Watson 2008). Recognized antecedents of native digital data are the firm history, its organizational processes and assets (Vitari *et al.* 2012). This kind of data has a positive influence on IT-based competitive advantage, but the benefits are reduced by sudden digital transformation changes (Raguseo, Vitari and Piccoli 2012). In terms of better financial performance, those firms being most capable of exploiting native digital data also have higher financial performance (measured in terms of ROA, ROS and revenue growth) (Raguseo and Vitari 2014). Moreover, firms able to develop skills based on native digital data, obtain higher outputs in terms of data quality and accessibility (Raguseo, Vitari and Pozzi 2016). Firms also benefit from a direct relationship between data, information, and knowledge when supported by a growing number of organizational units that collect data and exploit data analytics (Thornley *et al.* 2016). Such dynamics project the firms in a data-rich environment in which they must develop analytics methods enough flexible to fit structured and unstructured digital data generated within or out of the firms' boundaries (Wedel and Kannan 2016). Data-richness can quickly lead to big data and extracting information from big data is a recognized competitive factor in the digital transformation (Krämer, Tachilzik and Bongaerts 2017). Thanks to a set of recommendations for how to

successfully introduce big data analytics, firms can master the related organizational renovations while facing the digital transformation (Dremel *et al.* 2017). On the other hand, sometimes digital data remain mostly untapped by firms, this implies that data availability not automatically means that firms are going to use them (Balducci and Marinova 2018). Thus, the main resource created by the digital transformation is occasionally ignored or underexploited (Balducci and Marinova 2018).

1.2 Digital technologies in the agri-food industry

The impact of digital transformation in the agri-food industry is greatly influencing raw materials supply chain, production, processing, distribution, and marketing (Wagner Weick 2001). Stimulated by the availability of novel technologies in the food industry, new products like fruit juices fortified with vitamins, yogurt enriched with prebiotics, and omega-3 eggs have radically revolutionized customers' food habits (Bigliardi and Galati 2013). The adoption of micro and nanotechnologies (Marette *et al.* 2009; Steenis and Fischer 2016) allowed, for example, the encapsulation of food active components (Roos *et al.* 2016). As a consequence, firms could introduce in the market a great number of innovative new "functional foods" (Bigliardi and Galati 2013; Tollin, Erz and Vej 2016). Still, the digital transformation in food production fosters the creation of new types of machinery, such as 3D food printers (Charlebois and Juhasz 2018).

The phenomenon of the digital transformation in the agri-food industry has divided customers into opened versus skeptics towards the adoption of new technologies. Looking at the relationship between consumers' age and product selection, elderly people are usually willing to pay a premium price for products treated with technologies that provide added health benefits (Leek, Szmigin and Carrigan 2001). Conversely, millennials who care about sustainability issues are skeptical regarding the positive contribution of technologies to produce more sustainable food products and consider technologies adopted to prolong food shelf life dangerous (Cavaliere and Ventura 2018; Steenis and Fischer 2016). However, a study on consumer preferences for "familiar" versus "novel" food products claims that age is not a determinant factor in consumption decisions with familiar products, while it plays a more decisive role in the structure of preference regarding novel food products, particularly in young consumers (Barrenar, García and Camarena 2015). Moreover, the growing use of technologies in the agri-food

industry requires an active role of a large variety of actors (Hoppe *et al.* 2014). Therefore, the literature focused also on retailers, which are described as a powerful actor in the food value chain (Beckeman and Olsson 2011). What is more, those retailers strong enough to develop internal tech-departments are also able to influence the agri-food chain, thanks to their technological capabilities (Ejye Omar 1995). Due to their key position, retailers can promote voluntary market regulation, as the case of adopting a front-of-the-package nutrition scheme, to which all manufacturers had to comply by modifying their labels (Van Camp, Hooker and Souza-Monteiro 2010).

Developing new food products is a hard task. In the attempt to managing food development, actors of the agri-food tie-up inter-organization collaborations. For example, technological centers collaborate with food manufacturers to gain new knowledge, while manufactures usually need experts to get support for product development (Hoppe *et al.* 2014). Regarding food manufacturers, little attention has been paid by researchers. Manufacturers have to mature endogenous capabilities (e.g. build relationships) if they want to develop and introduce new products in the marketplace (Capitanio, Coppola and Pascucci 2010). Some food manufacturers build peer collaborations with other producers, but there is a widespread lack of trust in the food industry which, in turn, leads to a limited sharing of data and information (Beckeman, Bourlakis and Olsson 2013).

So far, agri-food research has greatly explained how technologies have been employed by the food industry and what food products have been created by the intensive use of technologies (e.g. (Leek, Szmigin and Carrigan 2001; Steenis and Fischer 2016). These studies also deeply investigate the role of the customer in the food industry, providing interesting insights regarding the technological centers, suppliers, retailers, and manufacturers too (e.g. (Beckeman, Bourlakis and Olsson 2013; Beckeman and Olsson 2011). Notwithstanding, just a few scholars studied the adoption of new digital technologies and focused on a pivotal phenomenon such as digital transformation in the agri-food firms (Vlachos 2004; Anastasiadis, Tsolakis and Srai 2018). Thus, previous research missed the opportunity to study the digital transformation and the increasing availability of digital data generated by the application of new technologies in food production, and how the information provided by the processing of digital data can support product development (Schweitzer, Handrich and Heidenreich 2019).

2. METHODOLOGY

This paper aims to explore how agri-food firms are digitally transformed and to provide a theoretical framework concerning how digital data are employed for product development. An exploratory multiple case-study design was adopted (Eisenhardt and Graebner 2007) since agri-food firms' digital transformation is an empirically underexplored field of research.

2.1 Research sample and case selection

Case-study research involves collecting and comparing data from 14 cases at agri-food firms (see Table 1).

TABLE 1• OVERVIEW OF THE CASE STUDIES

Case Study	Business Area	Case Description	Size*	Respondent
1	Fruits and vegetable processing	The firm processes bio and local fruits to produce pulps, smoothies, juices, as well as vegetable products like tofu, tempeh, and seitan.	Medium	CEO
2	Fruits processing	The organic farm has a citrus and olive orientation. It produces and commercializes kiwis and citrus fruits, as well as jams, marmalade, juices and extra virgin olive oil.	Small	CEO
3	Olive oil production	The firm is a cooperative of 250 companies that produce different kinds of extra virgin olive oil.	Large	IT specialist
4	Dairy products	The primary activity of the firm is dairy production. Linked to this, there is the whole agricultural and cow breeding sector. The production of raw materials and the transformation of sewage into electricity is done by the firm.	Medium	CEO
5	Dairy products	The firm is a cooperative of shepherds that deals with the transformation of cow milk from the farms of members and the production and distribution of dairy products.	Medium	CEO; IT specialist
6	Dairy products	The firm takes the highest quality sheep milk and whey and processes it to obtain powdered products, combining the natural properties with the benefit of longer shelf life and high solubility.	Small	CEO
7	Dairy products	The firm processes milk and produces mainly mature sheep and goat cheeses.	Medium	Marketing Director

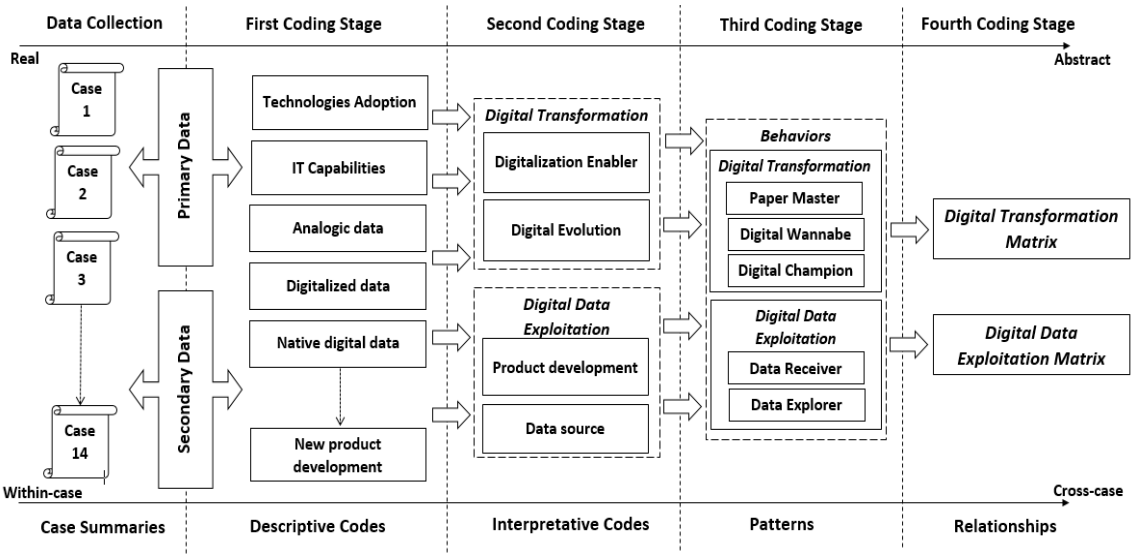
8	Poultry products	The firm is a specialist in the poultry market. It manages the entire integrated production cycle: the selection of raw materials, rearing units, hatcheries, feed facilities, food processing, packaging, and distribution.	Large	Head of IT & Digital Transformation; Head of R&D
9	Pasta and sweet products	The firm is operating for over 30 years in the production of regional fresh and dry pasta and local sweets.	Small	CEO
10	Fresh pasta	The firm produces fresh pasta such as tortellini, ravioli, and gnocchi, etc., for its shops.	Small	CEO
11	Fresh pasta	The firm produces fresh pasta such as tortellini, ravioli, and gnocchi, etc., for the organized large-scale distribution.	Small	CEO
12	Dry pasta and rusks	The firm produces several types and shapes of dry pasta as well as different kinds of rusks.	Large	Quality manager; Head of R&D
13	Food supplements	The firm develops, produces and markets food supplements mainly for athletes such as amino acids, creatine, protein, energy bars, etc.	Medium	Technical director
14	Cured meat	The firm processes and sells top-quality pork products and it is an important market player in several states of the European Union.	Large	Managing director

Firms were selected following a two-steps strategy. First, the main Italian organization and research centers dealing with digital transformation and the agri-food industry were contacted. A list of their food processing firm partners was required. Two organizations and a research center replay to the request suggesting a total of 25 firms to contact. These firms were emailed and five of them participated in this research. Second, the Italian chamber of commerce was phoned to ask if they could provide an equivalent list. They offered to call 228 food processing firms. 27 companies accepted to take part in this study. Collecting interviews was stopped at 14 cases when this study reached theoretical saturation that is "no additional data are being found […]. As he [the researcher] sees similar instances over and over again, the researcher becomes empirically confident that a category is saturated" (Glaser and Straus 2017: 61).

2.2 Data collection

Data analysis was conducted in 4 cumulative stages of coding, starting with the within-case analysis of each case, moving from the specific case context to the overall phenomenon (Saldaña 2015) (see Figure 1).

FIGURE 1 • DATA ANALYSIS PROCESS



The process started with a preliminary within-case analysis of the 14 cases and their characteristics by reconstructing the summaries of individual case studies. Summaries were created by reviewing interview transcripts, archive data, the firms' websites, and social network profiles.

During the first coding process, data were segmented and grouped following a data-driven coding scheme. A set of 11 descriptive codes was identified (Miles and Huberman 1994). Accordingly, the outcome of this stage of coding was a list of codes as observed in the single-considered cases (e.g. technologies adoption, IT capabilities, analogic data, internal data, incremental innovation, new product development).

TABLE 2

Descriptive code	Interpretative code	Definition	Description	Illustrative quote
Technologies adoption	Digitalization enabler	A digitalization enabler helps firms in accomplishing their digital transformation.	The adoption of digital technologies, acquisition of digital capabilities, the effort to increase firms' efficacy and agility in decision-making push firms towards the digital transformation.	The creation of data and information was much more difficult before the adoption of new technologies. Once we had a system that was not as precise as this about milk conductivity. The previous system had a much higher degree of error. Instead [technology name] is very precise. Before we had to rely on the monthly samples we took from the herd. *IT specialist, Case-5.*
Digital data capabilities				
Efficiency pursuing				
Agility seeking				
Analogic data	Digital evolution	Digital evolution is the firms' transition from the creation of analogic data to native digital ones.	Firms in distinct stages of digital evolution create different kinds of data.	"There is a processing sheet in which the operators write all the necessary data, for example, if there have been machinery downtimes, machinery consumptions, etc. Then, the coordinators input the data into the information system. While the most recent technologies are capable of producing digital data in the outgoing phase of the warehouse." *Technical director, Case-13.*
Digitalized data				
Native digital data				

Incremental innovation	Product development	The creation of products with new or different characteristics that offer new or additional benefits to the customer.	Process in which data are involved to create new products or improve the currentlyproduced ones.	"We have another benefit from data analysis. For example, we have shops where we directly sell our products, one of them is next to the dairy building. So, if we want to create some new product or test variations of the original product, we usually do these tests in our stores and collect data from customers." *Marketing Director, Case-7.*
New product development				
Internally created data	Data source	A data source is a location where data that are being used come from.	Firms use various data sources for separate tasks of processes.	"When a product is particularly performing, it could push us to improve its characteristics or those of products that could interest the same consumer. In this case, we very often base our analysis on the trend of internal data. While to find an indirect customer need, the best way is to try to interpret the sectoral market data. Market data are provided by the trade association and are national data." *Technical director, Case-13.*
Externally created data				

TABLE 3 • CROSS-CASE SUMMARY OF THE INTERPRETATIVE CODES

Case Study / ***Interpretative code***	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Digitalization enabler	X		X	X	X	X	X	X			X	X		X
Digital evolution	X	X	X	X	X	X		X	X	X		X	X	X
Product development	X	X	X	X			X	X	X			X	X	X
Data source			X	X	X	X	X	X		X	X	X	X	X

Then, the third stage of coding led the analysis to a further level of abstraction. Starting from the previously identified interpretative codes, patterns were pinpointed (Miles and Huberman 1994). In doing so, 5 behaviors were identified which, according to the analysis, reveal the digital transformation of agri-food firms and explain how digital data are employed for product development (see Table 4). Also at this stage, another cross-case analysis was performed to verify whether any construct was repeated in the 14 cases (Table 5).

TABLE 4• SUMMARY OF THE BEHAVIORS RELATED TO DIGITAL TRANSFORMATION AND DIGITAL DATA EMPLOYMENT

Phenomenon	Behavior	Definition	Description	Illustrative quote
Digital transformation	Paper master	A "Paper master" uses paper supports to take notes of data related to food processing.	Behavior performed by firms that mostly collect data on paper.	"At the end of the day, every person who works in a certain phase of the processing, must fill in the worksheets and take them to a production manager who files them. There is a whole paper system, we are not yet digitizing anything." *CEO, Case-11.*
	Digital wannabe	A "Digital wannabe" digitalizes food production data thanks to computer tools (e.g. scanner) or manually inputs data into the information system.	Behavior performed by firms that digitalize analogic data with the aim of benefit from having available digital data.	"Data are collected manually on product sheets that are stored in physical archives. Lately, we are scanning the product sheets. We do this not only because product sheets can be lost, but also because it is much simpler to code and group them by product families. As a result, product sheets are available on a computer to retrieve the data we need." *CEO, Case-2.*
	Digital champion	A "Digital champion" employs machinery able to create and send production data in digital format straight to the information system.	Behavior performed by firms which prefer food processing technologies able to create native digital data and communicate with the information system.	"Data are acquired thanks to sensors located in different points of the production process and transmitted to the information system. Data, directly in digital format, are stored on servers owned by the company." *Head of R&D, Case-12.*

Digital data exploitation	Data receiver	A "Data receiver" waits for prearranged production information to make decisions.	Behavior performed by firms that passively create and collect data while the analysis is done to produce prearranged information.	"Then there are a bunch of analyzes of product quality and productivity. We don't need to do additional analysis. From the data collected by the machinery, the information system generates information for us. Then, the man has to interpret the information, but we have already available all the analyses we need." *Managing director, Case-14*
	Data explorer	A "Data explorer" critically examines data to find new pieces of information.	Behavior performed by firms that actively explore production data by deepening data analysis.	"Machinery suppliers partially limit the autonomy of the company to carry out independent analyses. We are a bit forced to use and follow their models. We can carry out independent analyzes thanks to the additional sensors that we insert in the machinery to produce an autonomous and parallel data collection." *Head of IT & Digital Transformation, Case-8.*

TABLE 5 • CROSS-CASE SUMMARY OF THE BEHAVERS

Case Study / ***Behavior***	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Paper master		X							X	X	X		X	
Digital wannabe		X	X		X		X	X		X			X	
Digital champion	X		X	X	X	X		X				X		X
Data receiver		X	X		X	X	X	X	X	X	X	X	X	X
Data explorer	X			X				X				X		X

The final stage of data analysis involved assessing the relationships among them. This final coding aimed at connecting the constructs and transformed them from static and standalone behaviors into dynamic and integrated theoretical frameworks (see Figure 2 and Figure 3). The coding process was supported by Nvivo 10 software.

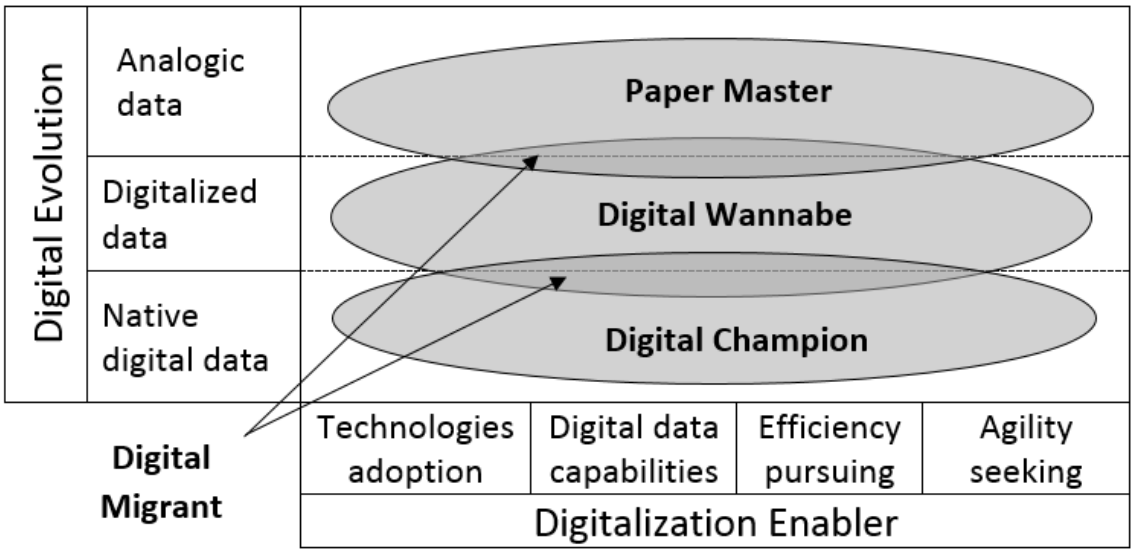
3. Digital transformation and digital data exploitation behaviors in agri-food firms

The findings of this study show that agri-food firms adopt five main behaviors when dealing with digital technologies that were labeled as Paper master, Digital wannabe, Digital champion, Data receiver, and Data explorer. Two groups of behaviors emerge from the analysis. The first one is related to the phenomenon of digital transformation and illustrates what firms do at several stages of digital evolution. The second has behaviors of digital data exploitation. Here, firms generate information from different digital data sources to develop their products

3.1 Digital transformation behaviors

In each of the 14 cases, data are created and managed. However, data creation can dramatically defer from case to case depicting how evolved is a firm in terms of digital transformation. Despite the kind of data used by the firms, the analysis of the cases unveils digital transformation enablers that indiscriminately push firms to adopt digital solutions. By combining the degrees of digital evolution and the digitalization enablers, three main behaviors related to digital transformation were pinpointed (see Figure 2).

FIGURE 2 • DIGITAL TRANSFORMATION MATRIX



Paper Master. Firms enacting this behavior are far from a complete digital transformation. The main characteristic that discerns Paper masters' behavior is the extensive and pervasive use of paper. The case-study analysis elicits that Paper masters employ paper in data collection because their technological equipment produces analogic data. Thus, Paper masters have no alternative that uses physical support to keep track of some aspects of food processing (e.g. quantities of raw material, temperatures, electricity consumption). However, the case-study analysis displays Paper masters pursuing efficient and fast decision making. The search for higher levels of efficiency and decision agility push firms in planning to acquire digital technologies: "At the moment, we have no technologies that allow us to weigh and mix the ingredients automatically. From this point of view, we are anchored to the older part of the factory technology. We planned to acquire better technologies because every year we invest to increase our efficiency and production capacity." (CEO, Case-11). Yet, Paper masters prefer to employ paper even when machinery can generate digital data, as the CEO Case-9 put it: "Data collection is manual, there are several data that are detected by the machine, which can be downloaded onto a USB stick and then transferred to a computer, but… these measurements are written on paper." Among the reasons for the pervasive use of paper, there is, for example, the employees' lack of digital capabilities: "even if you have very good workers, if you ask them to turn on a PC and open an Excel sheet for uploading some data, the panic starts!". This means that the technological equipment, purposes of improving efficiency and decision agility are not enough to make firms behave as more evolved ones in the digital transformation. The dataset analyst unveils that digital capabilities are needed too.

Digital Wannabe. Firms that belong to this category are more aware of the benefits of having digital technologies, especially as regards the availability of digital data: "I come from the ICT sector, I am perfectly aware of the importance of the data. Even when data do not seem useful, after a while or when certain things happen […] data turn useful" (CEO, Case-10). This awareness pushes Digital wannabes to collect a wider range of data compared with the Paper masters. Furthermore, the Digital wannabes make great effort for digitalizing data and it is a feature that differentiates their behavior. Form the analysis emerges that Digital wannabes' employees have at least basic digital data capabilities. These workers digitalize data employing computer tools such as keyboards by inputting

manually analogic data into an information system or scanning paper sheets to have a digital copy. As a result, Digital wannabes benefit from some of the digital transformation advantages. For example, they get useful information from the digitalized data analysis in operation management such as better standardization of the production and the improvement of quality control accuracy: "We have several milk suppliers, all the data related to the milk analysis are [manually] uploaded to files and stored. We use these data to evaluate milk quality and estimate what price to pay for it" (Marketing director, Case-7). Nonetheless, the analysis reveals data digitalization's side effects like high data collection costs, time-consuming data collection activities, poor data quality which is affected by human errors, and missing information. Concluding, Digital wannabes yearn to improve their digital conditions and they are halfway in the digital transformation. However, their technologies and capabilities limit a full transformation.

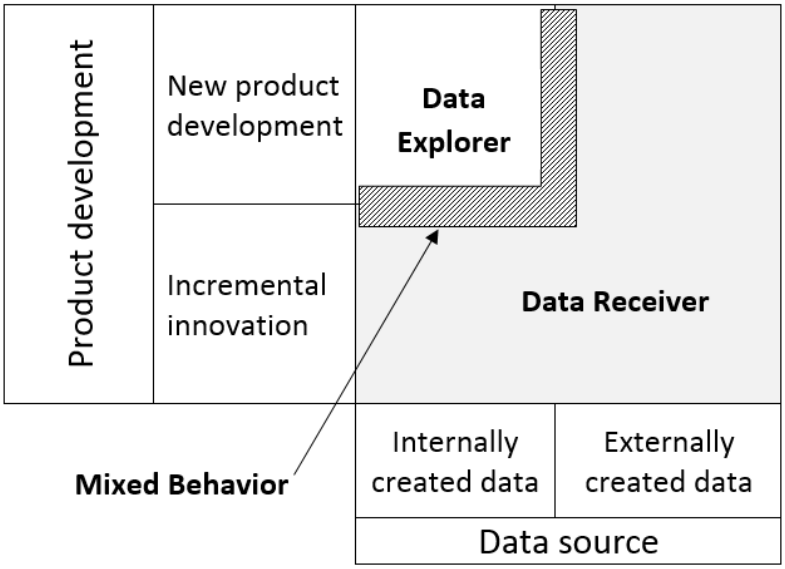
Digital Champion. Here, a pivotal role is played by the firms' technologies adoption: "the rusks factory is the most recent group's facility. No one in our company had ever run a facility with such recent production technologies." (Head of the R&D, Case-12). Digital champions' machinery creates data straight in digital format which are saved on servers connected to an information system. Firms have available a great variety of data regarding the details of the whole food processing chain, from the supply of raw materials to the sales results: "To give you some examples… the number of hectares cultivated; the real-time quantity of product harvested by machine; product humidity; in which warehouse the product must be stored; etc." (CEO, Case-4). Thanks to the digital data analysis, Digital champions use the available information to make real-time decisions, as the CEO Case-4 keep explaining: "Comfortably seated in our office, we receive a variety of information available in real-time. Based on this information, we advise the employee who is using the machine." Digital data employed by the information system quickly generate precise and ease to access information which in turn fuel agile decision making. The analysis suggests that a real-time decision-making process is the distinguishing feature of the Digital champions' behavior. It helps Digital champions reach a high level of efficiency since greater control of the production process improves the quantity and speed the production and decreases costs: "Based on the data collected during the production process, we look for a correlation between the flour mixture and the finished product yield and quality. If we see that there is a negative trend, we can strategically choose to modify the

flours mixture. On the contrary, if a production process has a very constant trend, we can decide to make a longer production getting a bigger quantity of a better outcome." (Quality manager, Case-12). Nevertheless, in some cases, the potential of digital technologies is not fully exploited (e.g. Case-6 and -14). Lack of digital data capabilities has a negative role in the digital transformation: "Unfortunately, we use approximately 30-40% of the potential of the technologies we have available because we do not have the right people to do this." (CEO Case-14). This means that, while technology adoption, efficiency pursuing and decision agility seeking seems to have mainly positive effects in the firms' digital transformation, the digital data capabilities can positively or negatively influence the transformation.

3.2 Digital data exploitation behaviors

Firms analyze data for several reasons. The research focuses on the exploitation of digital data for product development in terms of additional benefits to the customer (incremental innovation) or the creation of new products. By combining the types of product development processes with different data sources two behaviors were found both connected to digital data exploitation: Data receiver and Data explorer (see Figure 3).

FIGURE 3 • DIGITAL DATA EXPLOITATION MATRIX



Data Receiver. Data receivers have a passive attitude towards data analysis which is mainly done to produce standardize pieces of information to foster product development. The Data receiver is the most common behavior among firms. It appears in three of the four quadrants of the matrix. According to the analysis, Data receivers located in the bottom-left quadrant conduct analysis on data concerning internal aspects of the firm. Here, firms aim at improving products that are already produced exploiting internally created data: "Each production is a test, the data are analyzed to identify strengths and weaknesses of the product. Over time, we improve our products" (Quality manager, Case-12). The firms placed in the bottom-right quadrant are Data receivers too. They also want to improve their current products (or create variations of them), but they do that by analyzing firms' external dynamics (e.g. sales or mark trends): "We track everything we sell and we usually invest in products that sell the most to create variations of them." (IT specialist, Case-3). Lastly, Data receivers populate the top-left quadrant too. There, firms are exploiting externally collected data to create new products. For instance, by analyzing competitors' products: "We are followers as regards the creation of new products. We observe large companies that can make important investments in R&D. Then, we analyze their products and we try to adapt to what the biggest companies do." (CEO, Case-6).

Data Explorer. Such behavior is adopted by firms that explore digital data by examining them with a critical eye, and deepening data analysis, find novel pieces of information. Digital explorer is a minority and it is located in the top-left quadrant of the matrix. These firms query their database to create new products and understand whether they can do so, as the CEO Case-1 up it: "Production data are analyzed to create new products. Therefore, the opening of new markets is done by analyzing the production data to understand if the production plant is capable of producing a product that presents new characteristics that make the products more interesting to the final customer".

4. Discussion

Building on prior research on digital transformation and digital technologies in the agri-food industry, this study contributes by exploring the digital transformation in the agri-food sector and providing first insights about how such

firms exploit digital data for product development. First, this article contributes to an understanding of how agri-food firms behave in different stages of digital transformation and shows how distinct kinds of data and digitalization enablers influence these behaviors. Second, the research empirically demonstrates that digital data exploitation behaviors change according to the data sources employed by the firms when trying to develop their products.

4.1. Digital transformation behaviors mirror firms' data evolution

To date, former studies on the digital transformation focused on, e.g., changes in the firm digital transformation strategies (Ferreira, Fernandes and Ferreira 2019; Hess *et al.* 2016); alteration of the business model (Berman 2012); adoption of new technologies (Pankewitz 2017); data creation, collection, and analysis (Dremel *et al.* 2017). Even though these studies have examined significant digital transformation features, they were mostly developed in high-tech industries. Nevertheless, digital transformation is a priority in agri-food industries too (Anastasiadis, Tsolakis and Srai 2018; Vlachos 2004). Thus, previous studies live us without an explanation about how agri-food firms are digitally transformed (Hess *et al.* 2016; Loonam *et al.* 2018). The study contributes integrating previous literature by pinpointing three behaviors adopted by agri-food firms during the digital transformation. In particular, the results reveal that firms behave depending on the kind of data they use to operate and on digital enablers (e.g., digital data capabilities, technology adoption). For example, firms that make great use of paper (Paper master) are doing that because they deal with analogic data or have not capabilities to collect and use digital data. While aware of the advantage of managing digital data, firms try to digitalize their data (Digital wannabe), even if these firms get some pros of digitalizing data, they also face its cons (e.g., high costs of data collection). The most advanced firms in the digital transformation are those wich use native digital data (Digital champion). Oslo in this case, their digital capabilities can limit the advantages they can obtain from the utilization of digital data.

4.2. Digital data exploitation behaviors and the lack of new product miners

Earlier research identifies digital data as the key outcome of the digital transformation (Dremel *et al.* 2017). Studies on technological innovation in the food production (e.g., Beckeman, Bourlakis and Olsson 2013; Grunert et al. 2008),

on one hand, shed light on how agri-food firms utilize technologies and what new products have been developed due to the new technologies (e.g., Leek, Szmigin and Carrigan 2001; Marette *et al.* 2009; Steenis and Fischer 2016). On the other hand, they do not investigate the increasing availability of digital data in the agri-food sector, and how the information gathered by the analysis of digital data can affect product development (Schweitzer, Handrich and Heidenreich 2019). The study extends previous research by identifying digital data behaviors and theorizing how these behaviors change according to the data source and the kinds of the product development process. More specifically, while previous research demonstrates that new technologies adoption has positive implications in terms of product development, this article suggests that firms mostly adopt a passive behavior, (Data receiver) when exploiting digital data both for incremental innovation and for new product development. However, this study also shows that active behavior (Data explorer) is needed to exploit internally created data (e.g., production data) with the intent of creating new products. Despite the importance of exploiting digital data for new product development, Data explores are a minority. They display an active will in finding correlation between their available data and the possibility to create new producers.

Acknowledgements

The author offers his sincere thanks to professor Francesca Cabiddu for her invaluable help and guidance. The author thanks the autonomous region of Sardinia (Italy) for funding the project "Digital data e food innovation per lo sviluppo della filiera agroindustriale del Sulcis (DigIFood)" code: SULCIS-821194. The author also thanks the CIAS (Corvinus Institute for Advanced Studies) for funding his current position as visiting research fellow so that he can keep working on Digital Transformation in the agri-food industry.

REFERENCES

Anastasiadis F, Tsolakis N. and Srai J. (2018), *Digital technologies towards resource efficiency in the agrifood sector: Key challenges in developing countries*, «Sustainability», 10, 12, 4850

Balducci B. and Marinova D. (2018), *Unstructured data in marketing*, «Journal of the Academy of Marketing Science», 46, 4, pp. 557-590

Barrenar R., García T. and Camarena D.M. (2015), *An Analysis of the Decision Structure for Food Innovation on the Basis of Consumer Age*, «International Food and Agribusiness Management Review», 18, 3, pp. 149-170

Beckeman M., Bourlakis M. and Olsson A. (2013), *The role of manufacturers in food innovations in Sweden*, «British Food Journal», 115, 7, pp. 953-974

Beckeman M. and Olsson A. (2011), *The role of Swedish retailers in food innovations*, «The International Review of Retail, Distribution and Consumer Research», 21, 1, pp. 51-70

Berman S.J. (2012), *Digital transformation: opportunities to create new business models*, «Strategy & Leadership», 40, 2, pp. 16-24

Bigliardi B. and Galati F. (2013), *Innovation trends in the food industry: the case of functional foods*, «Trends in Food Science & Technology», 31, 2, pp. 118-129

Bowersox D.J., Closs D.J. and Drayer R.W. (2005), *The digital transformation: technology and beyond*, «Supply Chain Management Review», 9, 1, pp. 22-29

Brynjolfsson E. and Hitt L.M. (2000), *Beyond computation: Information technology, organizational transformation and business performance*, «Journal of Economic perspectives», 14, 4, pp. 23-48

Capitanio F., Coppola A. and Pascucci S. (2010), *Product and process innovation in the Italian food industry*, «Agribusiness», 26, 4, pp. 503-518

Cavaliere A. and Ventura V. (2018), *Mismatch between food sustainability and consumer acceptance toward innovation technologies among Millennial students: The case of Shelf Life Extension*, «Journal of Cleaner Production», 175, pp. 641-650

Charlebois S. and Juhasz M. (2018), *Food Futures and 3D Printing: Strategic Market Foresight and the Case of Structur3D*, «International Journal on Food System Dynamics», 9, 2, pp. 138-148

Dremel C., Wulf J., Herterich M.M., Waizmann J.-C. and Brenner W. (2017), *How AUDI AG Established Big Data Analytics in Its Digital Transformation*, «MIS Quarterly Executive», 16, 2

Eisenhardt K.M. and Graebner M.E. (2007), *Theory building from cases: Opportunities and challenges*, «Academy of management journal», 50, 1, pp. 25-32

Ejye Omar O. (1995), *Retail influence on food technology and innovation*, «International Journal of Retail & Distribution Management», 23, 3, pp. 11-16

Ferreira J.J.M., Fernandes C.I. and Ferreira F.A.F. (2019), *To be or not to be digital, that is the question: Firm innovation and performance*, «Journal of Business Research», 101, pp. 583-590

Fisher D.K., Norvell J., Sonka S. and Nelson M.J. (2000), *Understanding technology adoption through system dynamics modeling: implications for agribusiness management*, «The International Food and Agribusiness Management Review», 3, 3, pp. 281-296

Glaser B.G. and Strauss A.L. (2017), *Discovery of grounded theory: Strategies for qualitative research*, London - New York, Routledge.

Grunert K.G., Boutrup Jensen B., Sonne A.-M., Brunsø K., Byrne D.V., Clausen C., Friis A., Holm L., Hyldig G. and Kristensen N.H. (2008), *User-oriented innovation in the food sector: relevant streams of research and an agenda for future work*, «Trends in Food Science & Technology», 19, 11, pp. 590-602

Hess T., Matt C., Benlian A. and Wiesböck F. (2016), *Options for formulating a digital transformation strategy*, «MIS Quarterly Executive», 15, 2

Hoppe A., Marques Vieira L., Dutra de Barcellos M. and Rodrigues Oliveira G. (2014), *Research and development project of innovative food products from an inter-organizational relationship perspective*, «Journal on Chain and Network Science», 14, 2, pp. 137-147

Huber G.P. (1985), *Temporal stability and response-order biases in participant descriptions of organizational decisions*, «Academy of Management Journal», 28, 4, pp. 943-950

Kotarba M. (2018), *Digital transformation of business models*, «Foundations of Management», 10, 1, pp. 123-142

Krämer A., Tachilzik T. and Bongaerts R. (2017), *Technology and disruption: How the new customer relationship influences the corporate strategy*, in A. Khare, B. Stewart, R. Schatz (eds), *Phantom Ex Machina: Digital Disruption's Role in Business Model Transformation*, Cham, Springer, pp. 53-70

Leek S., Szmigin I. and M. Carrigan (2001), *Older consumers and food innovation*, «Journal of International Food & Agribusiness Marketing», 12, 1, pp. 71-89

Liu D.-Y., Chen S.W. and Chou T.-Z. (2011), *Resource fit in digital transformation: Lessons learned from the CBC Bank global e-banking project*, «Management Decision», 49, 10, pp. 1728-1742

Loonam J., Eaves S., Kumar V. and Parry G. (2018), *Towards digital transformation: Lessons learned from traditional organizations*, «Strategic Change», 27, 2, pp. 101-109

Marette S., Roosen J., Bieberstein A., Blanchemanche S. and Vandermoere F. (2009), *Impact of environmental, societal and health information on consumers' choices for nanofood*, «Journal of Agricultural & Food Industrial Organization», 7, 2

Matzler K., Von den Eichen S.F., Anschober M. and Kohler T. (2018), *The crusade of digital disruption*, «Journal of Business Strategy», 39, 6, pp. 13-20

Miles M.B. and Huberman A.M. (1994), *Qualitative data analysis: An expanded sourcebook*, Thousand Oaks, Sage

Olleros F.X. and Zhegu M. (2016), *Research handbook on digital transformations*, Cheltenham, Edward Elgar

Pankewitz C. (2017), *Automation, Robots, and Algorithms Will Drive the Next Stage of Digital Disruption*, in A. Khare, B. Stewart, R. Schatz (eds), *Phantom Ex Machina: Digital Disruption's Role in Business Model Transformation*, Cham, Springer, pp. 185-196

Piccoli G. and Ives B. (2005), *IT-dependent strategic initiatives and sustained competitive advantage: a review and synthesis of the literature*, «MIS quarterly», 29, 4, pp. 747-776

Piccoli G. and Watson R.T. (2008), *Profit from Customer Data by Identifying Strategic Opportunities and Adopting the 'Born Digital' Approach*, «MIS Quarterly Executive», 7, 3

Raguseo E. and Vitari C. (2014), *The development of the DDG-capability in firms: An evaluation of its impact on firm financial performance*, in L. Caporarello, B. Di Martino and M. Martinez (eds), *Smart Organizations and Smart Artifacts. Fostering Interaction Between People, Technologies and Processes*, Springer International, pp. 97-104.

Raguseo E. and Vitari C. and Piccoli G. (2012), *Gaining Competitive Advantage from Digital Data Genesis Dynamic Capability: the Moderating Role of Environmental Turbulence*, in IX Conference of the Italian Chapter of AIS, 28-29 September, Rome

Raguseo E. and Vitari C. and Pozzi G. (2016), *Investigating the impact of digital data genesis dynamic capability on data quality and data accessibility*, in C. Rossignoli, M. Gatti, R. Agrifoglio, *Organizational Innovation and Change*, Springer International, pp. 251-262

Remane G., Hanelt A., Nickerson R.C. and Kolbe L.M. (2017), *Discovering digital business models in traditional industries*, «Journal of Business Strategy», 38, 2, pp. 41-51

Roos Y.H., Fryer P.J., Knorr D., Schuchmann H.P., Schroën K., Schutyser M.A.I., Trystram G. and Windhab E.J. (2016), *Food engineering at multiple scales: case studies, challenges and the future: A European perspective*, «Food Engineering Reviews», 8, 2, pp. 91-115

Saldaña J. (2015), *The coding manual for qualitative researchers*, London, Sage

Schallmo D., Williams C.A. and Boardman L. (2017), *Digital transformation of business models: Best practice, enablers, and roadmap*, «International Journal of Innovation Management», 21, 08, 1740014

Schweitzer F.M., Handrich M. and Heidenreich S. (2019), *Digital transformation in the new product development process: the role of it-enabled PLM systems for relational, structural, and NPD performance*, «International Journal of Innovation Management», 23, 7, 1950067

Sebastian I., Ross J., Beath C., Mocker M., Moloney K. and Fonstad N. (2017), *How big old companies navigate digital transformation*, «MIS Quarterly Executive», 16, 3

Shaughnessy H. (2018), *Creating digital transformation: strategies and steps*, «Strategy & Leadership», 46, 2, pp. 19-25

Steenis N.D. and Fischer A.R.H. (2016), *Consumer attitudes towards nanotechnology in food products: an attribute-based analysis*, «British Food Journal», 118, 5, pp. 1254-1267

Sunding D. and Zilberman D. (2001), *The agricultural innovation process: Research and technology adoption in a changing agricultural sector*, in B. Gardner and G. Rausser, *Handbook of Agricultural Economics*, vol. I, part I, Elsevier, pp. 207-261

Thornley C., Carcary M., Connolly N., O'Duffy M. and Pierce J. (2016), *Developing a maturity model for knowledge management (KM) in the digital age*, in *16th European Conference on Knowledge Management*, 1-2 September 2016, University of Ulster, Northern Ireland

Tollin K., Erz A. and J. Vej (2016), *The strategic viewpoints of innovation and marketing teams on the development of novel functional foods*, in D. Bagchi and S. Nair, *Developing New Functional Food and Nutraceutical Products*, Academic Press, pp. 63-83

Van Camp D.J., Hooker N.H. and Souza-Monteiro D.M. (2010), *Adoption of voluntary front of package nutrition schemes in UK food innovations*, «British Food Journal», 112, 6, pp. 580-591

Vitari C., Piccoli G., Mola L. and Rossignoli C. (2012), *Antecedents of IT dynamic capabilities in the context of the digital data genesis*, in *Proceedings of ECIS 2012: The 20th European Conference on Information Systems*, 10-13 June, Barcelona

Vlachos I.P. (2004), *Adoption of electronic data interchange by agribusiness organizations*, «Journal of International Food & Agribusiness Marketing», 16, 1, pp. 19-42

Wedel M. and Kannan P.K. (2016), *Marketing analytics for data-rich environments*, «Journal of Marketing», 80, 6, pp. 97-121

Wagner Weick C. (2001), *Agribusiness technology in 2010: Directions and challenges*, «Technology in Society», 23, 1, pp. 59-72

LEONARDO MADIO AND MARTIN QUINN

USER-GENERATED CONTENT, STRATEGIC MODERATION, AND ADVERTISING[1]

Abstract. Social networks act as "attention brokers" and stimulate the production of user-generated content to increase user activity on a platform. When ads are displayed in unsuitable environments (e.g., disputed material), advertisers may face a backlash. This article studies the incentive for an ad-funded platform to invest in content moderation and its impact on market outcome. We find that if moderation costs are sufficiently small (large), the ad price is U-shaped (decreasing) in brand risks and the optimal content moderation always increases (is inverted U-shaped). When platforms compete for user attention, content moderation decreases as competition intensifies and this constitutes a market failure. Finally, well-intended policy measures, such as taxation of platform ad revenues, alter incentives to invest in content moderation and this might lead to the spread of harmful content.

Keywords. Advertising; content moderation; user-generated content; platforms.

1. INTRODUCTION

Online activities represent nowadays an essential part of citizens' life. In 2018 alone, Internet users spent 2.8 million years online, and most of this traffic (33% of the total time spent online) was generated by social media accounts (GlobalWebIndex 2019). Social media websites such as Facebook, YouTube, Instagram, Snapchat, TikTok, and many others, act as "attention brokers": they encourage users to spend more time online and monetize their attention with advertisements (ads). The more time spent on a social media website, the higher the number of profitable interactions with advertisers, the higher the platform's profit.

[1] The authors thank Luis Abreu, Malin Arve, Luca Ferrari, David Henriques, Laurent Linnemer, Christian Peukert, Carlo Reggiani, PatrickWaelbroeck for helpful comments on a previous draft. We are also grateful to seminar participants in Lisbon, Paris Saclay, Telecom ParisTech, UK OFCOM, at the Workshop on *Platforms E-commerce and Digital Economics* (CREST, 2019), at the Conference on *Auctions, Competition, Regulation, and Public Policy* (Lancaster, 2019), the *17th ZEW ICT Conference* (Mannheim, 2019), the *EARIE* (Barcelona, 2019), and the *Giorgio Rota Best Paper Award Conference* (Centro Einaudi, Turin, 2020). Leonardo acknowledges nancial support from the "MOVE-IN Louvain" Incoming Fellowship Programme and the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 670494). The usual disclaimer applies.

Advertisers' exposure on these platforms is not risk-free. As most contents are generated or uploaded by users, it lacks external and professional validation (Allcott and Gentzkow 2017). As a result, it is often the case that online material is inappropriate, harmful, or even illegal. Recent estimates suggest that approximately 4-10% of display advertising does not meet brand safety requirements and the majority of content can be classied at a moderate risk level (Plum 2019). The recent story of social media platforms is full of examples and scandals, which raised several concerns on howplatforms deal with what is posted online. In June 2020, several inuential brands and advertisers, ranging from Adidas to BestBuy, from Unilever to Coca-Cola, started boycotting – pulling their ads from – Facebook for its failure to create a safe environment for advertisers.[2]

Facebook was not the only platform dealing with protests for failure over content moderation. Between 2017 and 2019, YouTube went through the so-called "The Adpocalypse". Big advertisers such as Clorox, Disney, Epic Games, Hasbro, McDonald's, Nestlé, PepsiCo, Walmart, Starbucks, AT&T, Verizon, Volkswagen appeared just next to inappropriate usergenerated content, e.g., racist, extremist, and unsafe content.[3] Subsequently, they suspended their marketing campaign: some reduced their ad expenditure up to 70% in light of the extensive user market coverage the platform had. Others, instead, returned to the platform after a temporary pullback. The reason was distinctly expressed by the Association of National Advertisers, who argued that that because of such scandals, "reputation […] can be damaged or severely disrupted".[4]

To contain the scandals, YouTubewas forced to intervene by tightening its moderation policy, by shutting down 400 channels (including popular YouTubers such as PewDiePie), and by removing thousands of comments and videos. These interventions were part of a new program launched by YouTube in 2017 to allow the monetization of advertiser-friendly content only.[5] Other platforms, like Facebook and Instagram, followed suit, In November 2019, Facebook announced a "brand safety" tool for advertisers and, in May 2020, the creation of an independent body –

[2] See *The Brands Pulling Ads From Facebook Over Hate Speech*, «The New York Times» (https://www.nytimes.com/2020/06/26/business/media/Facebook-advertising-boycott.html).
[3] See *YouTube Adpocalypse*, «Fandom» (https://youtube.fandom.com/wiki/YouTubeAdpocalypse). See also *A timeline of the YouTube brand safety debacle*, «Digitalcontentnext», March 31, 2017 (https://digitalcontentnext.org/blog/2017/03/31/timeline-youtube-brand-safety-debacle/).
[4] See *Statement from ANA CEO on Suspending Advertising on YouTube*, March 24, 2017: https://www.ana.net/blogs/show/id/mm-blog-2017-03-statement-from-ana-ceo.
[5] See e.g., https://support.google.com/youtube/answer/9194476.

Oversight Board – to decide which content could be allowed to remain on the platform.[6]

This article explores the incentives of platforms to invest in content moderation and its interlink with the prices that advertisers pay to reach users. When content is not manifestly unlawful (e.g., hate speech, illegal content, whose presence may make the platform liable), a platform faces a challenging trade-of. On the one hand, the platform has incentives to invest in content moderation to create a safe environment for advertisers. As the risk of being associated with unsafe content decreases with stronger moderation enforcement, advertisers' willingness to pay increases and the platform can extract more revenue. On the other hand, the platform may want to safeguard individuals' fundamental freedom of speech, and please users not willing to be monitored. This may increase advertiser's risk of being displayed next to unsafe content, but it also allows to reach a larger audience. For instance, recent evidence showed that Tumblr, Yahoo's micro-blogging social network acquired by Verizon and later sold to WordPress, once with a high tolerance for not-safe-for-work (NSFW) content, lost nearly 30% traffic after banning porn in late 2018, and almost 99% of its market value. The ban was designed to keep "content that is not brand-safe away from ads".[7]

We find that the marginal gains from moderation depend on the direct and indirect effects that a stronger moderation policy entails. The direct (positive) impact leads to more impressions, which may create a disutility for users if ads are not informative. The indirect (negative) effect leads to fewer users on the platform and, as a result, fewer impressions. Interestingly, such a trade-of depicts a non-monotonic relationship between the optimal content moderation policy and the price advertisers pay to be on the platform. When the cost of moderating content is sufficiently small, the platform always increases its moderation effort if advertiser sensitiveness to brand risk

[6] In February 2019, Dune, Marks and Spencer, the Post Office and the British Heart Foundation charity experienced brand safety issues with Instagram as their ads appeared next to self-harm and suicide videos. See e.g., *Facebook' sorry' for distressing suicide posts on Instagram*, BBC, January 23, 2019 (https://www.bbc.com/news/uk-46976753). To tackle the problem, Facebook and Instagram increased content moderation efforts. For instance, Facebook claimed actions on 3.4 million content, including terrorist propaganda, graphic violence, adult nudity, and sexual activity, hate speech, and fake accounts in the first quarter of 2018. See *Facebook Community Standards Enforcement Preliminary Report*, 2018. In November 2019, Facebook announced a partnership with Integral Ad Science to help advertisers create a list of possibly sensitive videos.

[7] In other cases, such as YouTube, strict regulation on cannabis and rearm-related content fuelled new niche platforms such as TheWeedTube.com and Full30.com. See *After the porn ban, Tumblr users have ditched the platform as promised*, «The Verge», March 14, 2019 (https://www.theverge.com/2019/3/14/18266013/tumblrporn-ban-lost-users-down-traffic). See also *The road to becoming a weedtuber isn't easy*, «Leafbuyer», November 10, 2018 (https://www.leafbuyer.com/blog/weedtube/).

increases. Notwithstanding, the ad price is U-shaped in the brand risk and the highest price is set for very high or minimal brand risk. The reason is that when brand risk is small, advertisers care more about the customer reach and, hence, the platform can set a very high price. On the contrary, when brand risk is very high, the platform prefers to moderate all content and set a very high price to compensate for its investment.

The relevance of moderation costs in shaping platform behaviour also emerges when these costs are very large. This is the case – for example – of small entrant platforms which may face signicantly high cost for moderating content due to scarcity of past data or lack of state-of-the start equipment. Likewise, it can also be the case of language barriers or when the manifestly unlawful content and not-manifestly unlawful – but still harmful for advertisers – content becomes narrow. We find that when moderation costs are sufficiently high, instead, content moderation decreases has an inverted U-shaped relationship, such that it initially increases up to the point in which moderation becomes so costly that the platform finds it optimal to disinvest. In other words, the platform stops moderating content because it gets too expensive to accommodate advertiser preferences without losing customers. In this case, the ad price always decreases with brand risk.

Our analysis builds on a two-sided market model in which a platform (i.e., a social media website) provides meaningful interactions between Internet users (who consume online content) and advertisers.[8] Users join the platform free of charge, while advertisers pay an ad price to the platform. There are two types of content hosted on the platform: safe and unsafe ones. The first type always benefits users and advertisers. The second type can have some controversial effects: these contents can be valuable for (some) users while entailing a negative externality on advertisers. In other words, the presence of unsafe content creates "brand safety" issues for advertisers.[9] We model the presence of brand safety issues as the net value that advertisers obtain from joining a platform with a certain amount of unsafe content. However, the platform can indirectly control their presence and virality of unsafe content by investing in content moderation (and changing their terms and conditions for its users) such as hiring human content moderators and investing in monitoring

[8] See the pioneering works on two-sided markets of Rochet and Tirole (2003); Armstrong (2006). For a comprehensive discussion on the advertising-financed business model, see Anderson *et al.* (2016).
[9] SmartyAds defines brand safety as "the set of measures that aim to protect the brand's image from the negative or harmful influence of inappropriate or questionable content on the publisher's site where the ad impression is served" (https://smartyads.com/glossary/brand-safety-definition).

and AI-based content moderation. The stricter a platform content moderation policy, the lower the share of inappropriate content, the smaller the brand risk advertisers may face.

Our main analysis is performed by looking at the strategies of a monopolist platform and results hold in very general settings. A natural variation of our model is to consider how platform competition influences the incentives to invest in content moderation. We therefore present a Hotelling setup in which two (horizontally) differentiated platforms compete for user attention. In such a scenario, as platforms become more substitutable from the consumer perspective (e.g., more intense competition, lower switching costs), platforms react accordingly by lowering their content moderation effort and increasing or reducing the price advertisers pay to place their ads. The rationale is that as competition intensies, the marginal users become more valuable from the consumer perspective which can be attracted by lowering content moderation and reducing the nuisance they face in the presence of ad impression. If content moderation is sufficiently costly, platforms prefer to be more lenient with unsafe content and charge more advertisers because of the larger customer audience ensured. On the contrary, a more tolerant content moderation policy is associated with a lower ad price if content moderation is not very expensive. Indeed, this would compensate advertisers for possible brand safety issues.

In Section 4, we provide several variants of our model. Above all, we study the effect of a tax on ad revenues on the platform's optimal content moderation policy. In 2019, France adopted the so-called "GAFA tax", whereas the 2018's Nobel Prize laureate in economics put forward a proposal to tax digital ads "to protect and restore this public common" in light of dangerous misinformation and hate speech circulating on social media platforms.[10] Specifically, we find that the introduction of a fixed tax per ad placed on the platform has twofold effects. First, it reduces the incentives to invest in content moderation. Second, it can lead to a higher or lower price than in an environment with tax-free ads. The reason is that there is a first-order pass-through of the tax on the ad price. However, due to the lower content moderation, there is a second-order effect such that the ad price decreases (to compensate advertisers for the increased brand risk). Depending on the prevailing effect – which is linked to the cost function's convexity – the price can either increase or decrease.

[10] P. Romer, *A tax that could fix Big Tech*, «New York Times», May 7, 2019 (https://www.nytimes.com/2019/05/06/opinion/tax-facebook-google.html).

Our results provide implications for marketers and policymakers. As discussed, brand safety has become paramount in recent years and major brands coordinated their actions to induce platforms to tackle the problems of content moderation. However, these actions are unlikely to be successful if, on the other side of the market, there is a demand for controversial, viral, or potentially harmful content. The same problem would arise in the presence of users reluctant to forms of control of their expression online, especially for content whose identification can be challenging for automated tools. While our model accounts for the direct negative externality that the presence of potentially harmful content entails, our results can be relevant to discuss the platforms' incentives when the harmed party is external to the platform environment. For example, this can be the case of inappropriate content that may cause long-term negative externalities for society, e.g., fake news impacting election outcomes (Allcott and Gentzkow 2017) or leading to vaccine hesitancy (Carrieri *et al.* 2019). In European Union and in the United States, policymakers have started considering upgrades of the current liability regimes applied to online intermediaries and stricter regulation may impose to platforms procedural obligations and duties at least concerning manifestly unlawful content and hate speech.[11] Similarly, code of conducts on disinformation may reduce the extent to which advertisers may be exposed to unsafe content.

A second result drawing policy implications concerns the typical concern characterizing markets with strong network externalities and the winner-takes-all scenarios (see, e.g., Furman *et al.* 2019). Our results suggest that absent regulatory tools or changes in platform liability regimes, stimulating more competition in the market may lead platforms not to internalize fully the negative externalities linked to unsafe content. As a result, competition would introduce distortion regarding both ad pricing and content type and configure a market failure.

RELATED LITERATURE. This study contributes to the scant literature on user-generated content (UGC). Most of this literature features UGC as a media problem (Yildirim *et al.* 2013; Zhang and Sarvary 2014; Luca 2015; de Corniere and Sarvary 2018) and concerns the media outlet provisions of news and other content. Other

[11] In the US, platforms are considered hosting service providers and, hence, exempted from liability (US Communication Decency Act Section 230). Under the European E-Commerce Directive (2000), platforms can benefit from a conditioned liability exemption, depending on the knowledge standard of the illegal activity carried out on the platform and their passive role in the distribution of the information. In 2020, the European Commission has launched the Digital Services Act to upgrade liability rules for platforms.

studies in the marketing literature looks at UGC in their forms of online reviews and their impact on sales (Chevalier and Mayzlin 2006; Chintagunta *et al.* 2010; Proserpio and Zervas 2017; Chevalier *et al.* 2018). This literature falls short of explaining the possible side-effects of UGC on advertisers. Instead, this paper studies how brand safety influences advertisers' behavior and shows that heterogeneity in advertisers' aversions to brand-risk has signicant consequences for the platform optimal content moderation and ad prices.

We also add to the literature on advertising and media, which has, so far, addressed different types of questions.[12] The ad-targeting literature is perhaps the closest to the spirit of our study. This literature generally assumes a better match between the user's preference and the advertisers' type. This way, the likelihood of wasteful advertising campaigns is reduced, and each customer becomes a proper market. In this article, instead, targeting is not customer-specific. Investments in moderation allow a platform to decide which segment to serve and, as a result, it attracts users and advertisers more favorable to the type of content hosted by the platform.

Moreover, this article bears some similarities with the literature on media bias, which has mainly dealt with news bias originated in the supply side or the demand side of the market. The former deals with a bias originated by advertisers, political orientations, government pressures, and lobbies (see e.g., Ellman and Germano 2009; Besley and Prat 2006). The latter depends on beliefs of targeted audiences (see e.g., Gentzkow and Shapiro 2006; Mullainathan and Shleifer 2005; Xiang and Sarvary 2007; Gal-Or *et al.* 2012). A major feature of this literature is that a content provider decides about the distortion of the news.[13] Our approach differs from it in at least two dimensions. First, a platform acts as a content aggregator. This implies that it is not directly involved in content creation and in choosing the direction of the bias. On the contrary, it chooses which sides of the market to please the most. Second, the platform can gain control over a content only by exercising costly moderation effort. To this end, it trades-off the benefits of ensuring a higher brand safety to advertisers with a costly effort and a potential demand contraction on the user side. This way, the platform decision can entail either a supply-side or demand-side bias depending on its moderation effort.

[12] The literature on advertising and media has mainly focused on the different types of ads displayed to users (Anderson and De Palma 2013), targeting technologies and matching (Bergemann and Bonatti 2011; Peitz and Reisinger 2015), overlaps in the customer base and homing decision (Ambrus *et al.* 2016; Athey *et al.* 2016; Anderson *et al.* 2017), ad-avoidance (Anderson and Gans 2011; Johnson 2013), and more generally to the media see-saws (Anderson and Peitz 2020).
[13] For a review, see e.g., Gentzkow *et al.* 2015.

The above aspects allow us to differentiate this contribution from that of some closely related studies on media bias. For instance, Van Long *et al.* (2019) study competition on content quality (real or fake news) between media outlets and find that competition increases user polarisation. Although this underlines how content providers tailor their material and bias their news, the paper does not feature advertisers' preferences and UGC. Ellman and Germano (2009) investigate media bias in a market in which platforms sell content to readers and profit from advertisers. They give the power to platforms to change the accuracy of the news. Such a lever can have a signicant effect as a lack of accuracy in the reporting of violent or shocking news may allow the platform to generate a better match with ads. Our article underlines a similar mechanism when considering the impact of UGC on platform profits. In this case, the platform might influence that match by moderating content (more) carefully.

In the framework of media bias, Mullainathan and Shleifer (2005) show that when newspapers compete for user demand, there is an incentive to exaggerate media bias. Similarly to ours, Gal-Or *et al.* (2012) study the competition between ad-based media outlets in the presence of heterogeneous readers and endogenous homing decisions of advertisers. Although our mechanism is reminiscent of theirs, they show that when a media outlet relies on ad revenues, there are more incentives to moderate content as this results in a higher ad price. In this way, advertisers multihome and attract moderate readers. However, the authors also show that when advertisers singlehome, newspapers become a bottleneck, and competition intensies. This results in more slanting to soften competition and greater polarization of readers. In our model, instead, when competition intensifies, content moderation becomes more tolerant and the number of impressions users are exposed to decreases.

Finally, recent empirical studies support our results and show how different platforms engage in different moderation policies. For instance, Chiou and Tucker (2018) study Facebook's decision in 2016 to ban ads linking to external websites fabricating fake news. They find that the ban was effective: fake news declined more on Facebook than on Twitter after the policy. Rao (2018) documents the effectiveness of the US Federal Trade Commission enforcement on fake news websites, showing that when these websites were shut down, consumer interest for fake news declined and was displaced by the interest for regular advertisements. Their study alongside Allcott *et al.* (2019) motivate our analysis on platform heterogeneity in moderation policies. They show that Facebook was more prone than Twitter in banning fake and false news, underlying platform heterogeneity.

ARTICLE STRUCTURE. The article unfolds as follows. In Section 2, we present a fairly general model with a platform monopolist. The effect of platform competition on content moderation is studied in Section 3. In Section 4, we present a number of extensions. Section 5 provides concluding remarks and policy implications.

2. THE MODEL

Consider a platform environment in which an online intermediary (e.g., social media website) connects users and advertisers. Users consume UGC available on the platform, and their attention is catered to advertisers. For simplicity, let us assume that users only consume UGC and do not engage in their production. Such an assumption can be justied by the fact that a few very popular content creators generate typically viral content (e.g., popular YouTubers, influencers on Instagram) and there is a long-tail of unpopular creators with a little number of views. For instance, on YouTube, content creators can only monetize views when reaching at least 1,000 subscribers and have streamed at least 4000 hours in the last 12 months.[14]

Users can consume two types of content: a mass 1 of safe content and a mass (m) of unsafe content. The former, which identifies professional videos and news, pictures of vacations and pets, entail positive benefits for both users and advertisers. For advertisers, one can imagine a positive match value when impressions are just next to these contents. The latter, instead, identifies controversial and possibly harmful content. For instance, these can be borderline comments which userswant to protect in light of their freedom of speech but can create brand safety issues for advertisers. The mass of this content depends on the moderation policy the platform selects and which is identified by the parameter $m \in [0, 1]$, with $\theta(0) = 1 \; and \; \theta(1) = 0$. When $m = 0$, there is a unit mass of unsafe content (and hence 50% of the entire platform content is potentially dangerous), whereas with $m = 1$the platform moderates all content.

THE PLATFORM. There is an ad-funded platform that charges a zero price to users and lets advertisers (acting on behalf of brands) pay for launching an ad campaign at price *p*. We assume that advertisers do not compete for ad space and they launch at

[14] See e.g., *Additional changes to YouTube partner*, YouTube (https://youtubecreators.googleblog.com/2018/01/additional-changes-to-youtube-partner.html.

most one ad campaign. We denote the number of advertisers joining the platform by $a(m,p)$. The platform maximizes profits by choosing the price p and investing in costly content moderation $C(m)$. We assume that content moderation is suciently convex, such that $C'(m) > 0$, $C''(m) > 0$, and $C(0) = 0$. While it can be argued that there are economies of scale, one must consider that moderation can be increasingly challenging when the content type to be monitored becomes larger. To see why, consider a very mild content moderation policy that only checks whether a content promotes terrorism. In this case, content moderation may require a certain degree of investment $C(m)$. However, if the platform wants to a enforce a much stricter moderation policy, also including conspiracy theories and borderline comments - for which categorization can require more effort and capabilities than with manifestly harmful content - then platform costs are likely to be much larger as requiring additional investments in text analysis.[15] Similarly, while AI tools and filters based on tags and keywords can have benefits, some content may require post-human moderation, therefore leading to much higher prices. All these costs are taken into account by a platform when choosing ad prices and content moderation policies. Platform's profits can then be summarized as follows:

$$\prod = a(p,m)\,p - C(m). \tag{1}$$

INTERNET USERS. There is a unit mass of Internet users. Each user is identied by the duple (u,ϕ) that captures her taste for "safe", u, and "unsafe" content, $\varnothing$. Specically, we assume that the preference for safe content is distributed according to the following parameter $u \in [0,\overline{u}]$. Users are also differentiated according to their taste ϕ for unsafe content, with $\phi \in [u,\overline{u}], \overline{u} > 0$. Note that while the sign of ϕ is positive, the sign of ϕ is unspecified. When this is negative, (some) users gain from content moderation, whereas when it is positive, all users dislike content moderation. Moreover, at this stage, we do not put any restrictions on the distribution function form of u and ϕ and assume both distributions are independent of one another. Moreover, we also assume that users dislike ads as being perceived as a nuisance cost, with $\gamma > 0$ identifying this parameter. The total nuisance cost to which users are

[15] Moderation can be ex-ante or ex-post. When ex-ante, for instance, all content must be validated and approved by a moderator. When ex-post concerns moderation performed after the content has circulated. Content moderation can have type-I and type-II errors, thereby leading to removal of genuine content and errors in moderating harmful content. The study of these effects would not change the main trade-off faced by the platform.

therefore exposed is then equal to $\Upsilon \times a(p, m)$. The utility of the users when joining a platform is

$$U = u + \phi\theta(m) - \Upsilon. \tag{2}$$

ADVERTISERS. There is also a unit mass of advertisers that decide whether to launch their ad campaign on the platform depending on their long- and short-term profitability. The utility of an advertiser can be expressed as follows:

$$V = \pi(n, \Omega) - p, \tag{3}$$

where $\pi(n, \Omega)$ captures the profitability of the ad campaign and p is the price paid to the platform. For short-term profitability, $\pi(n, \cdot)$, we intend revenues obtained from the interaction with the n users the platform attracts. Online interactions yield a stream of (exogenous) revenue r. The higher r, the larger the advertisers' cross-network externalities. Thisway, $r \times n$ can represent revenues from individual clicks or the possibility to obtain short-term after-market transactions when users buy products in-store or online. For long-term profitability, $\pi(\cdot, \Omega)$, we denote the impact of a brand's (long-term) reputation. As discussed in the introduction, advertisers are increasingly concerned about the impact of scandals on their reputation. As many marketers argued when urging digital platforms to tackle misinformation, racism, and hate speech, this impact is not directly channeled through a reduction in sales or click rates but via reputation which contributes to a significant share of a firm's value (Jovanovic 2020).

Unlike the previous literature dealing with the ad market, we assume that brands (via advertisers) care about the suitability of the environment in which impressions appear. Hence, advertisers benefit from the presence of a mass of safe content according to a parameter $\upsilon \in [0, \overline{\upsilon}]$ but face a disutility, $\lambda \in [0, \overline{\lambda}]$ from the presence of a mass (m) of unsafe content. Formally, can be expressed as

$$\Omega = 1 \times \upsilon - \lambda\theta(m). \tag{4}$$

Note that this very general specification captures the large heterogeneity across advertisers' benefit from being displayed just next to a safe/unsafe content. For instance, a large λ may represent advertisers promoting luxury goods or charities,

that would have a lot to lose when associated with extreme content (i.e., $\frac{\partial \Omega}{\partial m}$ is high). On the contrary, unsafe content can have a small impact on advertisers promoting gambling websites (i.e., a small λ). Hence, advertisers' utility in Equation (3) can be written as follows:

$$V = rn + \upsilon - \lambda\theta(m) - p. \tag{5}$$

TIMING. The timing of the game is as follows. In the first stage, the platform maximizes profits and chooses both ad price and the content moderation policy. In the second stage, users choose whether to visit the platform and advertisers decide whether to place their ad. These decisions are made simultaneously and we assume that users and advertisers have fulfilled expectations on the number of participants on the opposite side of the market. The game is solved backward and the equilibrium concept is subgame perfect.

2.1 Optimal content moderation

We first compute the level of activity on the platform. Following Rochet and Tirole (2003) and using equations (2-5), the number of users joining the platform can be written as and the number of advertisers placing their ads as $n = \Pr(U \geq 0)$. Formally, this implies

$$\begin{aligned} a &= \Pr(\upsilon - \lambda\theta(m) + rn - p \geq 0) \equiv D^a(p,n,m), \\ n &= \Pr(u - \gamma a + \phi\theta(m) \geq 0) \equiv D^n(a,m), \end{aligned} \tag{6}$$

assuming that the above system of equations admits a unique solution that defines a and n depending on (p,m) such that $a = d^a(p,m)$ and $n = d^n(p,m)$.[16] In the first stage, the platform chooses m and p to maximize $\Pi = d^a(p,m)p - C(m)$. Denote by Ψ the *elasticity of profit with respect to moderation* such that

$$\Psi = \underbrace{\frac{\partial D^a}{\partial m}}_{\text{Brand safety effect}} + \overbrace{\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial m}}^{\text{Eyeball effect}}. \tag{7}$$

[16] For more details, see Rochet and Tirole (2003).

From the first-order conditions, we obtain:

$$\begin{aligned} p =& -\frac{d^a(p,m)}{\frac{\partial d^a}{\partial p}} = -\frac{d^a(p,m)(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})}{\frac{\partial D^a}{\partial p}}, \\ C'(m) =& p\frac{\partial d^a}{\partial m} = p\Big(\frac{\Psi}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\Big), \end{aligned} \tag{8}$$

and the marginal gain from moderation, which implicitly defines the optimal level of content moderation, is summarized by the lemma below.

Lemma 1. Define the following as the platform marginal revenue from content moderation

$$MR(m^*,p^*) = -\frac{d^a(m^*,p^*)\Psi}{\frac{\partial D^a}{\partial p}}. \tag{9}$$

The optimal content moderation is implicitly defined by the following expression such that marginal costs equal marginal revenue:

$$\begin{aligned} C'(m^*) =& -\frac{d^a(m^*,p^*)\Psi}{\frac{\partial D^a}{\partial p}}. \\ MC(m^*) =& MR(m^*,p^*). \end{aligned} \tag{10}$$

Note that the optimal moderation policy is chosen in away such that the marginal gains/revenues (MR) from moderation equal the marginal costs (MC). Importantly, due to the multisidedness of the market, the latter account for the price the platform selects, the effect that content moderation has on advertiser demand, and how consumer and advertiser react to increased brand safety (via $\boldsymbol{\Psi}$). One can easily see that the larger $\boldsymbol{\Psi}$, the larger the gains from content moderation. In the limit case in which $\boldsymbol{\Psi < 0}$, the platform sets $m^* = 0$ and advertisers are associated with potentially harmful content.

To shed some further light on how the advertiser sensitiveness to brand risk impacts equilibrium outcomes, we present some simple comparative statics on how the optimal price and moderation react to an increase in brand risk (a higher $\bar{\lambda}$). As brand risk is only contained in $\boldsymbol{\Psi}$ and demand forms, we investigate how p and m changes with $\boldsymbol{\Psi}$ – the elasticity of platform profits with respect to content moderation. The next proposition summarizes the main findings and highlights the relevance of moderation costs.

Proposition 1. *There exists a cut-off*

$$\tilde{C} \equiv -2\left(1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}\right)\frac{d^a(m^*, p^*)^2}{(\frac{\partial D^a}{\partial \Psi})^2 \frac{\partial D^a}{\partial p}},$$

such that if moderation costs are sufficiently small ($C''(m^*) < \tilde{C}$), then p* is U-shaped while m* is increasing in Ψ. Else, p* is decreasing while m* is inverted U-shaped in Ψ.

Proof. See Appendix A.

Specifically, when moderation costs are suciently small, a platform can easily adjust its moderation effort depending on how many users and advertisers can attract and the surplus to be extracted. This way, when brand risk increases – and the advertiser's willingness-to-pay to advertise decreases – the platform optimally adjusts its moderation effort. As a result, the moderation effort m^* monotonically increases with Ψ up to the point in which full moderation $m^* = 1$ is ensured. When Ψ is large enough, the brand safety effect largely outweighs the eyeball effect. In practice, the platform prefers to have fewer users to safeguard advertisers' reputation. For instance, platforms may be more meticulous when attracting advertisers promoting luxury goods or charities, whose reputational losses from scandals might be substantial.

By contrast, when moderation costs are sufficiently high, if advertiser brand risk increases, the platform faces increasingly high costs to satisfy their requests and this is not compensated by a large market capture on the consumer side. As a result, the moderation effort is inverted U-shaped: it monotonically increases to the point in which accommodating advertiser preferences becomes too expensive in terms of investments and in the number of consumers exiting the platform.

Then, the moderation effort m^* starts decreasing with Ψ to the point in which contents are no longer moderated, $m^* = 0$. Interestingly, for very high brand risk with high moderation cost, the platform prefers not to moderate content at all.

A similar discussion also applies to the effect of Ψ on ad prices, which presents some nonmonotonicity. To see why, we first analyze the situation where moderation costs are small. In this case, moderation increases with Ψ and the ad price is U-shaped. Namely, the platform relatively high prices for low and high values of Ψ and these correspond to when no moderation or full moderation is enforced. The reason is that at one extreme, the platform sells a very large number of user eyeballs to advertisers and given the low risk of being exposed to harmful material, the price can be high aswell. At the other extreme, the platform sacrices some audience and meets

the moderation requests of advertisers which - given their high willingness-to-pay for content moderation - also pay a very high price. For intermediate values of Ψ, that is, when the brand safety effect is not much more significant than the eyeball effect, the platform sets an intermediate level of moderation. This mild content moderation may feature controls of flags of some disputed content, such as 'hate speech', violence, nudity and sexual content, intellectual property rights violation as well as the veracity of the news. The ad price reaches a minimum when

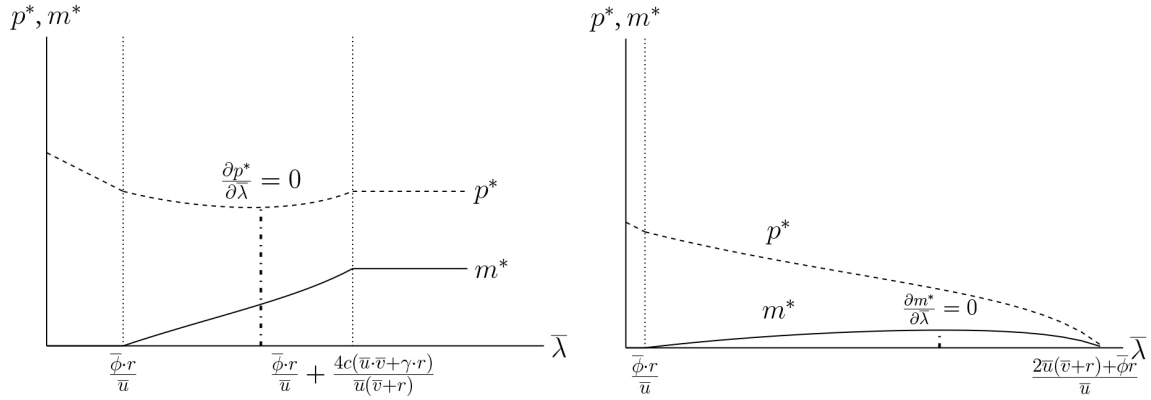
$$-2\frac{\partial D^a}{\partial p}d^a(m^*,p^*)=\frac{\Psi\dfrac{\partial D^a}{\partial \Psi}\dfrac{\partial D^a}{\partial p}}{1-\dfrac{\partial D^a}{\partial n}\dfrac{\partial Dn}{\partial a}},$$

where p^* reaches a minimum in $\overline{\lambda}$ and so the platform mediates the divergence between the two sides of the market by granting advertisers a price discount. In turn, p^* is convex in Ψ. On the contrary, when moderation costs become too large the optimal ad price is always decreasing in Ψ. The reason is twofold. First, when brand risk is suciently small, the platform increases moderation, but the way it increases does not compensate advertisers for the consumers who exit the platform. As a result, the price goes down.

However, suppose advertisers become too sensitive to unmoderated content. In that case, the marginal revenues from increased moderation become lower than the marginal costs of moderation (in terms of intrinsic moderation costs and consumers exit), so the platform starts reducing moderation and compensates the advertisers for the very high risk of being exposed to unsuitable content. In turn, the ad price decreases.

The above discussion emerges prominently in Figure 1 - where advertiser and user preferences follow a uniform distribution (see Appendix B). The two figures present how the optimal price and content moderation react to advertiser aversion to brand risk when moderating costs are small (left) and large (right).

FIGURE 1 • EXAMPLE WITH A UNIFORM DISTRIBUTION OF PREFERENCE: EFfECT OF $\overline{\lambda}$ ON p^* AND m^* WHEN c IS SMALL (LEFT) AND LARGE (RIGHT)



3. PLATFORM COMPETITION

A natural variation of the benchmark model is the introduction of platform competition. Whereas platforms often exhibit forms of monopolization in their natural market, they also compete for user attention in several other markets. For instance, although their services can be regarded as sufficiently differentiated from the user perspective, YouTube competes with Facebook for advertising revenues and on the provision of UGC. In this subsection, we study a model of platform competition in the presence of full market coverage. We then study how the intensity of competition influences content moderation policy and ad pricing strategies.

Platforms are located at the endpoints of a Hotelling-line of unit distance. Platform 1 is located at coordinate 0, whereas Platform 2 at coordinate 1. A platform i sets a price p_i with $i = 1, 2$ for the entire ad campaign and a_i represents the number of advertisers deciding to buy a space on the website. Hence, platforms' profits are defined as follows

$$\Pi_i = a_i p_i - C(m_i).$$

Consistently with the previous literature (Anderson *et al.* 2016), we let advertisers multihome. Hence, platforms characterize a competitive bottleneck and each platform becomes the only way to reach unique users. This implies that platforms compete for attracting consumers. We assume that platforms are symmetric and we look for a

symmetric equilibrium. As in the presence of a monopolist, advertisers are defined by a duple $(\upsilon, \lambda) \in [0, \overline{\upsilon}] \times [0, \lambda]$, with a uniform distribution of υ and λ. This setting is an adaptation of a two-dimensional differentiation as in Anderson and Gans (2011), Economides (1986) and Vandenbosch and Weinberg (1995). Their utility when patronizing platform *i* is

$$V_i = \upsilon + rm_i - \lambda\theta(m_i) - p_i. \tag{11}$$

There is a unit mass of users and we assume that the market is fully covered. Users are are uniformly and independently distributed on a line of unit length; they are identied by a duple relative to their relative preference for platform *i(j)*, defined by their position x on the Hotelling line and by their aversion for moderation ϕ. The latter is assumed to be uniformly distributed in the interval $[0, \overline{\phi}]$ Note that, for tractability, we only consider the case in which users dislike moderation, but their tastes are heterogeneous, i.e., ϕ in Section 2, is set equal to zero.[17]

The utility of a user located at x and joining platform *i* is as follows:

$$U_i = u + \phi\theta(m_i) - \gamma a_i - \tau \mid x - l_i \tag{12}$$

where $l_i \in \mid \{0,1\}$ indicates the location of the platform.

The timing of the game is as before. In the first stage of the game, platforms compete by simultaneously and non-cooperatively choosing ad prices and content moderation policies. In the second stage of the game, advertisers decide whether to place an ad on both platforms or stay out of the market, whereas Internet users decide which platform to join. We look for a symmetric equilibrium.

To provide clear insights on the optimal ad price *p** and moderation policy *m** and to compute the equilibrium, we assume that the mass of unsafe content is a linear and decreasing function of *m*. Similarly, we assume quadratic moderation costs.

$$\theta(m) = 1 - m \quad \text{and} \quad C(m) = c\frac{m^2}{2}, \quad \text{with} \quad c > 0.$$

[17] This simplifying assumption allows us to focus on the case in which advertisers' and users' preferences over moderation are conicting - which is the most insightful case. As a result, if some users had a negative ϕ, the platform would have a slightly higher incentive to increase content moderation effort as users' preferences would converge towards those of advertisers.

By solving the model backward, the number of advertisers is

$$a_i(n_i, n_j) = 1 + \frac{rn_i - \overline{\lambda}\theta(m_i) - p_i}{\overline{v}}, \tag{13}$$

whereas the number of users 'exclusive' to each platform is

$$n_i(a_i, a_j) = \frac{1}{2} + \frac{\overline{\phi}(\theta(m_i) - \theta((m_j)) - 2\gamma(a_i - a_j)}{4\tau}, \qquad n_j = 1 - n_i. \tag{14}$$

Rearranging the above expression, we obtain the second-period market shares of platform *i* for both advertisers and users, respectively denoted by d_{ai} and d_{ni}. In the first stage, platforms make their decision on moderation and ad price simultaneously and non-cooperatively to maximize

$$\max_{p_i, m_i} \Pi_i = d_{a^i}(p_i, p_j, m_i, m_j)p_i - C(m_i). \tag{15}$$

We assume that profits are well-behaved as long as $C(m_i)$. is suffciently convex. For the sake of simplicity, we let moderation costs be quadratic (i.e., $C(m_i) = cm_i^2 / 2$). As in the benchmark case, it is also assumed that $\theta(m) = 1 - m$.

$$\lambda_{11} \equiv \frac{\overline{\phi} r \overline{v}}{(\tau \overline{v} + \gamma r)}, \quad \lambda_{12} \equiv \frac{4c(4\tau\overline{v} + 3\overline{\gamma}r)}{(2\overline{v} + r)(2\overline{v}\tau + \gamma r)} + \frac{\overline{\phi} r \overline{v}}{2\tau\overline{v} + \gamma r},$$

Define the following cut-of value of λ:

$$\lambda_{11} \equiv \frac{\overline{\phi} r \overline{v}}{(\tau \overline{v} + \gamma r)}, \quad \lambda_{12} \equiv \frac{4c(4\tau\overline{v} + 3\overline{\gamma}r)}{(2\overline{v} + r)(2\overline{v}\tau + \gamma r)} + \frac{\overline{\phi} r \overline{v}}{2\tau\overline{v} + \gamma r},$$

and by solving for the symmetric equilibrium, we can state the following proposition.

Proposition 2. *When competition takes place, a symmetric equilibrium exists as follows.*

(a) When $\overline{\lambda} \leq \lambda_{11}$, *the platform enforces no moderation, m* = 0, and sets the following ad price*

$$p_i^* = p_j^* = \frac{(\tau\overline{v} + r\gamma)(r + 2\overline{v} - \overline{\lambda})}{4\tau\overline{v} + 3r\gamma}.$$

(b) For any $\lambda_{11} < \lambda < \lambda_{12}$*, Nash equilibrium outcomes are an interior solution such that*

$$p_i^* = p_j^* = \frac{4c\bar{v}(2\bar{v} + r - \bar{\lambda})(\tau\bar{v} + \gamma r)}{4c\bar{v}(4\tau\bar{v} + 3\gamma r) + \bar{\lambda}(\bar{\phi} r\bar{v} - \bar{\lambda}(2\bar{v}\tau + \gamma r))},$$
$$m_i^* = m_j^* = \frac{(2\bar{v} + r - \bar{\lambda})(\bar{\lambda}(2\tau\bar{v} + \gamma r) - \bar{\phi} r\bar{v})}{4c\bar{v}(4\tau\bar{v} + 3\gamma r) + \bar{\lambda}(\bar{\phi} r\bar{v} - \bar{\lambda}(2\bar{v}\tau + \gamma r))}.$$

c) When $\bar{\lambda} \geq \lambda_{12}$*, platforms set* $m^*_i = m^*_j = 1$ *and*

$$p_i^* = p_j^* = \frac{(2\bar{v} + r)(\tau\bar{v} + \gamma r)}{4\tau\bar{v} + 3\gamma r}.$$

Proof. See Appendix A.

The analysis of the symmetric equilibrium shows that when there exists a sufficiently small brand risk, that is, whenever $\bar{\lambda}$ is sufficiently low, platforms enforce no moderation policies - i.e. $m_i^* = 0$ - and set low ad prices to attract as many advertisers as possible. For intermediate brand risk, $\lambda_{11} < \lambda < \lambda_{12}$, the optimal moderation policy is an equilibrium solution, $m_i^* \in [0,1]$. In this case, platforms can increase their ad price as advertisers are willing to pay more given the reduced brand risk.[18] Finally, when brand risk is high, the platforms respectively enforce a full moderation policy – $m_i^* = 1$ – and set a very high ad price.

However, such results may depend on the intensity of competition in the market, which in our case is proxied by the degree of platform differentiation. To better grasp this effect, we compute derivatives of m_i^* and p_i^* with respect to τ. When τ decreases, product differentiation reduces and, in turn, competition intensies. The following proposition summarizes this result.

Proposition 3. *Let* $\tilde{c} := \frac{\bar{\lambda}(\bar{\lambda}\gamma + \bar{v}\bar{\phi})}{4\bar{v}\gamma}$ *, for any* $m_i^* \in [0,1]$*, fiercer market competition on the user side leads to (a) lower content moderation; (b) a price reduction (increase) for suciently small (large) moderation costs (*$c < (>)\tilde{c}$ *) and (c) fewer ads displayed to users.*

[18] Note that, in both cases, a condition to ensure a non-negative price is such that $\bar{\lambda} < r + 2\bar{v}$. This implies that advertiser's brand risk is not as high (i.e., low enough $\bar{\lambda}$) relative to advertiser gain from being on the platform, i.e., $r + 2\bar{v}$ is large). This ensures that the market exists.

Proof. See Appendix A.

Proposition 3 shows that when competition for users becomes ercer, platforms have two ways to attract more users. On the one hand, they can relax their moderation policy and, hence, please users with a strong aversion to content moderation. Onthe other hand, they can reduce the number of ads and, therefore, their nuisance. In equilibrium, the mechanism works as follows. When moderation is sufficiently expensive, content moderation is already low. In this case, the onlyway to attract users is to reduce the number of ads by increasing the ad price. This increases platforms' profits and attracts additional users. When moderation is less expensive, the moderation policy is already quite strict. As competition intensifies, the platform prefers to reduce content moderation to attract more users and compensate advertisers with a reduction in the price. In turn, this mitigates the advertisers' exit. These two forces are complements to reach the goal of attracting users when competition intensifies but, due to the symmetry of the market, in equilibrium, it does not bring about additional users and the platforms obtain equal market shares. Different is the effect on the advertiser side: as competition gets fiercer, the number of ads placed on each platform decreases regardless of the pricing strategy. As ads are considered a nuisance cost, this turns out to increase the user welfare. The above proposition also has a relevant implication. Typically, fostering more competition in the market is advocated by policy-makers and regulatory agencies. For instance, this could translate in lowering barriers to entry, reducing switching costs, facilitating data portability, larger compatibility across platforms, or having non-exclusive access to essential inputs.

Similarly, authorities other than competition ones are concerned with potential societal externalities stemming from the uncontrolled presence of UGC. For instance, negative externalities can result from misinformation, hate speech. Absent other interventions or the possibility to enforce platform liability given the existing framework in Europe and the US, well-intended policy measures aimed at increasing competition in the market are likely to generate negative externalities. First, advertisers would face a higher brand risk without observing any demand expansion.[19] Second, negative societal externalities may arise if UGCs are perceived as harmful from policy-makers, though not always persecutable in courts.

[19] In our framework, this depends on the Hotelling structure of the model and of the full market coverage assumption.

4. DISCUSSIONS AND EXTENSIONS

4.1 Impact of policy tools: a tax on digital revenues

In recent years, several countries in Europe (e.g., France, Germany, Italy) have started cosidering the introduction of a tax on online ads to create a fairer environment. More related to the aim of this paper, in 2019, the Nobel Prize laureate Paul Romer proposed the introduction of a tax on digital ads as a measure to induce social media platforms to limit misinformation.

To shed some light on the possible unintended effects of the introduction of such a tax, we modify our benchmark model, and we assume that the Government (exogenously) imposes a tax f on each ad. As a result, the Government can raise af, which drives platform profits to be $\Pi = d^a(p,m)(p-f) - C(m)$. As taxes impact the platform's marginal profits, we expect it to affect the price advertisers pay and accordingly, the content moderation decided by the platform.

Using the Implicit Function Theorem and the Cramer's rule, the introduction of a tax on digital ads determines the following results.

Proposition 4. *The ad price increases (decreases) with a tax on platform's revenues*

$$\frac{\partial p^*}{\partial f} = \frac{\frac{\partial D^a}{\partial p} C''(m)(1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) + \Psi^2}{2\frac{\partial D^a}{\partial p} C''(m)(1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) + \Psi^2} > (<) 0$$

The optimal moderation policy also decreases with f such that

$$\frac{\partial m^*}{\partial f} = -\frac{\frac{\partial D^a}{\partial p}\Psi(1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})^{-1}}{2C''(m)\frac{\partial D^a}{\partial p}(1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) + \Psi^2} < 0.$$

Proof. See Appendix A.

Proposition 4 underlines very interesting results. First, the ad price may increase or decrease with the tax depending on the cost of moderation. When the tax increases, a first-order effect drives the ad price up. This is common in optimal taxation theory as the platform is just passing the tax into the price.[20] However, a second-order effect implies a reduction in the platform moderation effort, which, in turn, decreases the ad price. One effect dominates the other depending on moderation cost. Specifically,

[20] See Mankiw *et al.* (2009) for a complete review on optimal taxation theory.

when moderation costs are suciently low, the negative indirect effect dominates as moderation decreases faster with a tax. In this case, ad price also decreases faster that it increases with the direct effect. When moderation costs are high enough, the opposite is true.

Second, the moderation policy always decreases with a tax. This is because the tax directly reduces the marginal revenues from advertisers. Hence, the higher the tax, the lower the marginal revenue from moderation enforcement, the lower the moderation effort. All in all, as in the benchmark model, the effect on moderation is aligned with advertisers' interests. As content moderation is relaxed, also advertisers place fewer ads. To better understand the above mechanisms, in Appendix B, we provide an example with a uniform distribution of preferences.

4.2 Endogenous content creation

In the main specication, we have assumed exogenous content creation. In this section, we relax this assumption and consider the case in which agents can also create content. This implies endogenizing the volume of both safe (i.e., 1) and unsafe content (i.e., $\theta(m)$). As before, we assume that the platform accepts all safe materials, whereas it moderates the unsafe ones.

Indeed, we explicitly model the presence of content creators among the users, who obtains a utility $U_\theta = u_\theta + nk - m$ when creating unsafe content,[21] with m being the platform moderation policy, u_θ his willingness to create an inappropriate content, and nk payoffs being the network effect from being exposed to n users on the platform. Such utility from content creation u_θ may be heterogeneous on the support $[0, \overline{u}_\theta]$ with $u_\theta < 1 - k$ such that, if $m = 1$ (full moderation), all content creators make negative utility. Hence, the number of endogenously created content, θ, would be equal to $\theta = P(u_\theta + nk - m > 0)$. As in the benchmark model, the marginal gains from moderation are equal to

$$MR(m^*, p^*) = -\frac{d^a(m^*, p^*)\Psi}{\frac{\partial D^a}{\partial p}}$$

[21] Unsafe materials often generate virality. These can be any sensationalist or attention-grabbing content produced by creators in social networks and community platforms like Youtube, e.g., Conspiracy Theories, No-Vax comments, etc.

with

$$\Psi = \frac{\partial D^a}{\partial m}(1 - \frac{\partial D^n}{\partial \theta}\frac{\partial D^\theta}{\partial n}) + \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial \theta}\frac{\partial D^\theta}{\partial m}$$

The above expression is an augmented version of the one presented in Lemma 1. However, it differs in the inclusion of indirect network externalities stemming from content creators' presence and their interest for a broad audience base. Such a result suggests that a platform is pursuing a stricter content moderation policy while pleasing risk-averse advertisers may dissatisfy both creators and users. This is entirely in line with the "Tumblr spiral". After the acquisition of Yahoo from Verizon and the policy on content moderation to make the platform brand-safe, many creators and users decided to leave the platform, and the stock value of the former $1.1 billion-platform plunged to only $3 million - the price paid by Automatic, the owner of Wordpress.

4.3 Targeting

So far, we have not considered the possibility that the platform(s) can target users. Targeting can arise in different ways. First, ads can be targeted to users in a way that does not cause any distress and nuisance. This implies that the platform can eventually control γ. If γ were considered equal to 0, such that ads are neutral to users, our main results would go through as in the benchmark model. An important difference, however, would be present when considering the case of competing platforms: ad prices would always be reduced when competition intensifies. When the competition for users becomes more intense, the platform no longer needs to compete by lowering the nuisance costs to users by increasing the ad price. As a result, a more intense competition leads to a reduction of both content moderation and ad price.

The second form of targeting can be related to better matching between advertisers and content. In this market, advertisers typically create lists of keywords they want (or do not) to be associated with. For instance, according to IAS Insider, the most blocked keywords by advertisers in November 2019 included "shooting, explosion, dead, bombs, etc".[22] This may ensure some forms of safeguards for brands and marketers. However, targeting is far from perfect (Nielsen 2018), and better

[22] See IAS Insider (https://insider.integralads.com/the-20-most-blocked-keywords-in-november-2019/).

precision may require investment costs which are very similar to the one used in our model. As the main trade-off remains unchanged, our model also encompasses a setup in which targeted moderation is imperfect.

4.4 Other Applications

OFFLINE NEWS OUTLETS. Our setting can provide more general insights into content moderation also arising in other markets. For instance, consider a (traditional) media outlet hosting content. Typically, these outlets have full control over the type of content they display. Such a practice differs from platforms that do not control content production. However, even professional content can feature a divergence between the interests of the users and those of the advertisers. In September 2016, following the online campaign "Stop Funding Hate" related to the presence of disputed content on migrants, several advertisers such as The Body Shop, Plusnet, Walkers, and many others announced that they would stop advertising on *The Daily Mail* and *The Sun*. Others, like the Co-operative Group, preferred to maintain their adverts as driving up sales.[23] Such a story well fits the trade-off that traditional media outlets may face when producing content. We discuss this by making two contributions.

Consider a news outlet that only produces professional content that is sufficiently attention grabbing to be attractive for users but also allows advertisers to place their ads. Hence, this outlet would strategically choose the sensitivity of materials to produce to balance user attraction and advertisers' exit. Whereas investments in content moderation are not needed in this case as there are no UGC available on the platform, content production may still be costly. The better (or, the more professional) the content, the higher the cost, the safer it can be for advertisers. However, we may imagine that producing professional content is cheaper than moderating thousands of comments and posts online. In this case, our framework indicates that competition between outlets would make content quality going down while the ad price goes up.

CONTENT AGGREGATORS. Our study can also provide applications for content aggregators that host both first-party (i.e., professional content) and third-party (i.e., UGC) content. In such a case, the aggregator would directly balance user and

[23] See The Co-operative Group, *An update on our advertising policy* (https://blog.coop.co.uk/2017/03/23/anupdate-on-our-advertising-policy/).

advertiser preferences when choosing the type of content to produce and display to (safely) monetize users' eyeballs. Indeed, a content aggregator would need to balance users' attraction strategically and advertisers' brand safety concerns. Such a set-up allows us to endogenize the platform's design choice that consists of accepting or not UGC to be displayed on the platform. Depending on platform moderation costs and production costs, an outlet may be keener on introducing UGC or not on the platform. For instance, a high-end fashion website may only attract advertisers with high brand safety. In this case, we conjecture that when moderation cost is higher than production cost, such a website would prefer to produce its content rather than allowing moderate too costly UGC.

TV REALITY SHOWS. The framework we depict can also be applied to TV reality shows, such as the famous The Big Brother, which are sponsored by advertisers and feature the presence of a group of (unprofessional) contestants. While viewers might like some of the houseguests' scandals, which keep the reality game alive after years, this might not always be the case of advertisers that sponsor the program with their products. For instance, in Italy, in 2018, several different sponsors, including Nintendo, decided to give up their partnership with the TV show after bullying in the house.[24]

5. MAIN HIGHLIGHTS AND CONCLUSIONS

The digital revolution has changed the production of media content. Whereas in the past, thesewere mostly produced by professionals (e.g., journalists), the advent of social mediawebsites has given users control over production and difusion of content. In most cases, this happened without any external and professional validation and created concerns among advertisers and marketers. This article studies the economic implications of such a situation and underlines the trade-of faced by a social media platform when strategically enforcing content moderation. In the following, we disentangle the importance of our results for both managers and policymakers.

MANAGERIAL IMPLICATIONS. This article provides a rationale for the signicant heterogeneity across platforms in tackling illegal, harmful, or disputed content. We argue that content moderation policies are rather platform-specific as depending on

[24] *Grande Fratello, la grande fuga degli sponsor: niente acqua, shampoo e Nintendo*, «Blitzquotidiamo.com», May 4, 2018, (https://archivio.blitzquotidiano.it/tv/grande-fratello-fuga-sponsor-acqua-nintendo-2876635/).

the overall platform elasticity concerning moderation. This is due to the type of users and advertisers each social media attracts. Hence, we provide managerial implications for both brands and platforms.

First, the two-sidedness of the market is crucial for both advertisers and platforms. On the one hand, a platform should consider its moderation cost, audience type, and advertisers type when deciding to invest in content moderation. Only in this case, the platform will be able to balance advertising price and content moderation strategy that together maximize advertisers' willingness to pay. On the other hand, advertisers must be able to detect platform choice in scouting both moderation technology and audience type depending on their nature. For instance, old brands with inherited reputation should pay more attention to platforms pursuing lax content moderation and may decide to advertise only if the short term revenues do not jeopardize their brand image. On the contrary, young brands may care less about brand safety, hence maximizing short term revenue without imperiling their long term strategy.

Second, our results underline the importance of moderation costs in ad price and content moderation decisions. As shown in its moderation report, Facebook admits facing a cost to moderate that is idiosyncratic to countries, depending on language, culture, and characteristics.[25] Our analysis shows that for a monopolist moderation, costs can lead the platform to react differently to increase in brand risk - like the one presented in recent protests by advertisers.

More importantly, the advertisers push for more brand safety may not be supported by Big Tech if moderation costs are very large. This may lead to reduced content moderation if content moderation becomes very costly, perhaps because of too many content to be analyzed or different languages to be considered. Paradoxically, such an outcome is more likely to arise the more advertisers become concerned about brand risk. On the contrary, it is in the interest of the platform to accommodate advertisers' requests if moderation costs are sufficiently small.

Third, our results showthat absent platform liability, competition between platforms plays a crucial role. Specically, we find that as competition intensies, a social media platform would always decrease its content moderation, but must be more sophisticated about its price strategy. A first situation arises when moderation is expensive. In that case, content moderation is already low and the only way for the platform to compete for user attention is by reducing the nuisance costs users face.

[25] A summary of the report can be found here https://transparency.facebook.com/community-standards enforcement.

One way to do this is to increase the ad price, which reduces advertiser demand. Alternatively, the platform could also invest in targeting technologies. However, such a solution is unlikely to provide satisfactory results as content moderation and targeting typically exhibit type-I and type-II errors. A second situation emerges when moderation can be achieved at a low cost. In that case, content moderation is high and the platform may allow for a lot more unsafe content as it is very efficient in attracting new users. In this case, as advertisers bear some risks, the platform may be willing to reduce the ad price.

POLICY IMPLICATIONS. The above-described results are also of paramount relevance not only for marketers but also for policymakers. Social media moderation policies are not neutral and this article highlights that their decisions depend on the trade-of between generating revenues from advertisers and capturing user attention. At different institutional levels, it is widely debated what platforms should do to prevent the difusion of illegal content and misinformation going on social mediawebsites as their effects could be detrimental to society. For instance, the European Commission recently issued a recommendation on how tackling effectively illegal content online (EU 2018) stressing how platforms need to "exercise a greater responsibility in content governance" and, in 2020, it launches the Digital Services Act with plans to revise the EU E-Commerce Directive, change the liability regimes of online intermediaries, and regulate content moderation and algorithms.[26] In 2018 the German Bundestag passed a law requiring platforms to remove hate speech within 24 hours or face nes of up to 50 million euro (see e.g., CEPS 2018). In this respect, we also discuss more broadly how well-intended policies aimed at stimulating more competition in digital markets might have the (unintended) effect of lowering platform incentives to invest in content moderation. Our results show that increasing competition between platforms is likely to reduce their moderation effort and distort pricing strategies on the advertising market.

In addition, we study the impact of an often advocated policy measure like the digital tax on advertising revenues. This was adopted in France, Germany, Italy, and recently supported by the Nobel Prize laureate Paul Romer. This article shows that these well-intended measures may have the perverse effect of reducing moderation

[26] See *Illegal content on online platforms*, «Digital Single Market» (https://ec.europa.eu/digital-singlemarket /en/illegal-content-online-platforms). Similarly, see e.g., T. Kaeseberg on «VoxEu», December 12, 2019, *Promoting competition in platform ecosystems* (https://voxeu.org/article/promoting-competitionplatform-ecosystems).

effort for the platform, thereby increasing the relevance of the current problem faced by democracies and advertisers. More complex settings of our setup may provide further insights. For instance, platform reputation may represent away to mitigate negative externalities fromUGC and induce more responsible actions. Similarly, the recent Cambridge Analytica scandal pushed Facebook to intervene to regain its user's trust. Moreover, an extension of this work may also consider a mixed business model and the incentives of these platforms to deal with content moderation when also users pay a subscription price.

References

Allcott H. and Gentzkow M. (2017), *Social media and fake news in the 2016 election*, «Journal of Economic Perspectives», 31, 2, pp. 211-236

Allcott H., Gentzkow M. and Yu C. (2019), *Trends in the difusion of misinformation on social media*, NBER Working Paper No. 25500

Ambrus A., Calvano E. and Reisinger M. (2016), *Either or both competition: A 'two-sided' theory of advertising with overlapping viewerships*, «American Economic Journal: Microeconomics», 8, 3, pp. 189-222

Anderson S.P. and De Palma A. (2013), *Shouting to be heard in advertising*, «Management Science», 59, 7, pp. 1545-1556

Anderson S.P., Foros Ø. and Kind H.J. (2017), *Competition for advertisers and for viewers in media markets*, «The Economic Journal», 128, 608, pp. 34-54

Anderson S.P. and Gans J.S. (2011), *Platform siphoning: Ad-avoidance and media content*, «American Economic Journal: Microeconomics», 3, 4, pp. 1-34

Anderson S.P. and Peitz M. (2020), *Media see-saws: winners and losers on media platforms*, «Journal of Economic Theory», forthcoming

Anderson S.P., Waldfogel J. and Stromberg D. (2016), *Handbook of Media Economics*, vol 1A, Elsevier

Armstrong M. (2006), *Competition in two-sided markets*, «The RAND Journal of Economics», 37, 3, pp. 668-691

Athey S., Calvano E. and Gans J.S. (2016), *The impact of consumer multi-homing on advertising markets and media competition*, «Management Science», 64, 4, pp. 1574-1590

Bergemann D. and Bonatti A. (2011), *Targeting in advertising markets: implications for off-line versus online media*, «The RAND Journal of Economics», 42, 3, pp. 417-443

Besley T. and Prat A. (2006), *Handcufs for the grabbing hand? media capture and government accountability*, «American Economic Review», 96, 3, pp. 720-736

Carrieri V., Madio L. and Principe F. (2019), *Vaccine hesitancy and (fake) news: quasi-experimental evidence from Italy*, «Health Economics», 28, 1377-1382, pp. 417-443

CEPS (2018), *Germany's netzDG: A key test for combatting online hate*, Centre for European Policy Studies (CEPS)

Chevalier J.A., Dover Y. and Mayzlin D. (2018), *Channels of impact: User reviews when quality is dynamic and managers respond*, «Marketing Science», 37, 5, pp. 688-709

Chevalier J.A. and Mayzlin D. (2006), *The effect of word of mouth on sales: Online book reviews*, «Journal of Marketing Research», 43, 3, pp. 345-354

Chintagunta P.K., Gopinath S. and Venkataraman S. (2010), *The effects of online user reviews on movie box oce performance: Accounting for sequential rollout and aggregation across local markets*, «Marketing science», 29, 5, 944-957

Chiou L. and Tucker C.E. (2018), *Fake news and advertising on social media: A study of the anti-vaccination movement*, NBER Working Paper No. 25223

de Corniere A. and Sarvary M. (2018), *Social media and news: Attention capture via content bundling*, Mimeo

Economides N. (1986), *Nash equilibrium in duopoly with products defined by two characteristics*, «The RAND Journal of Economics», 17, 3, pp. 431-439

Ellman M. and Germano F. (2009), *What do the papers sell? a model of advertising and media bias*, «The Economic Journal», 119, 537, pp. 680-704

EU (2018), *Commission recommendation of 1.3.2018 on measures to effectively tackle illegal content online*, C(2018) 1177 final

Furman J., Coyle D., Fletcher A., Marsden P. and McAuley D. (2019), *Unlocking digital competition*, Report of the Digital Competition Expert Panel, March

Gal-Or E., Geylani T. and Yildirim T.P. (2012), *The impact of advertising on media bias*, «Journal of Marketing Research», 49, 1, pp. 92-99

Gentzkow M. and Shapiro J.M. (2006), *Media bias and reputation*, «Journal of Political Economy», 114, 2, pp. 280-316

Gentzkow M., Shapiro J.M. and Stone D.F. (2015), *Media bias in the marketplace: Theory*, in *Handbook of Media Economics*, Elsevier, vol. 1, pp. 623-645

GlobalWebIndex (2019), *Social media flagship report* (https://www.globalwebindex.com/hubfs/Downloads/Social-H2-2018-report.pdf)

Johnson J.P. (2013), *Targeted advertising and advertising avoidance*, «The RAND Journal of Economics», 44, 1, pp. 128-144

Jovanovic B. (2020), *Product recalls and firm reputation*, «American Economic Journal: Microeconomics», forthcoming

Liu Y.-H. (2018), *The impact of consumer multi-homing behavior on ad prices: Evidence from an online marketplace*, Mimeo

Luca M. (2015), *User-generated content and social media*, in *Handbook of Media Economics*, Elsevier, vol. 1, pp. 563-592

Mankiw N.G., Weinzierl M. and Yagan D. (2009), *Optimal taxation in theory and practice*, «Journal of Economic Perspectives», 23, 4, pp. 147-174

Mullainathan S. and Shleifer A. (2005), *The market for news*, «American Economic Review», 95, 4, pp. 1031-1053

Nielsen (2018), *Nielsen digital ad ratings: Benchmarks and findings through 2h 2016, Europe*.

Peitz M. and Reisinger M. (2015), *The economics of Internet media*, in *Handbook of Media Economics*, Elsevier, vol. 1, pp. 445-530

Plum (2019), *Online advertising in the UK*, Report commissioned by the UK Department of Digital, Culture, Media Sport

Proserpio D. and Zervas G. (2017), *Online reputation management: Estimating the impact of management responses on consumer reviews*, «Marketing Science», 36, 5, pp. 645-665

Rao A. (2018), *Deceptive claims using fake news marketing: The impact on consumers*, available at SSRN 3248770

Rochet J.-C. and Tirole J. (2003), *Platform competition in two-sided markets*, «Journal of the European Economic Association», 1, 4, pp. 990-1029

Van Long N., Richardson M. and Stähler F. (2019), *Media, fake news, and de-bunking*, «Economic Record», 95, 310, pp. 312-324

Vandenbosch M.B. and Weinberg C.B. (1995), *Product and price competition in a twodimensional vertical differentiation model*, Marketing Science, 14(2):224–249.

Xiang Y. and Sarvary M. (2007), *News consumption and media bias*, «Marketing Science», 26, 5, pp. 611-628

Yildirim P., Gal-Or E. and Geylani T. (2013), *User-generated content and bias in news media*, «Management Science», 59, 12, pp. 2655-2666

Zhang K. and Sarvary M. (2014), *Differentiation with user-generated content*, «Management Science», 61, 4, pp. 898-914

Appendix A

Proof of Lemma 1

The number of users who decide to join the platform is $n = \Pr(U \geq 0)$, whereas the number of advertisers is $a = \Pr(V \geq 0)$. Following Rochet and Tirole (2003), the demands can be expressed as follows:

$$\begin{aligned} a &= \Pr(v - \lambda\theta(m) + rn - p \geq 0) \\ n &= \Pr(u - \gamma a + \phi\theta(m) \geq 0) \end{aligned}$$

Assume that the above system of equations admits a unique solution that defines a and n depending on (p, m) such that $a = d^a(p, m) \equiv D^a(p, m)$ and $n = d^n(m, p) \equiv D^n(m, p)$. Then, we can solve the model in the first stage of the game whereby the platform chooses m and p to maximize profits:

$$\Pi_i = d^a(p, m)p - C(m).$$

In what follows, we first look at how demands on both sides of the market change with ad prices and moderation. The derivatives of d^a and d^n with respect to p and m can be deduced from those of D^n and D^a as in the following expressions

$$\begin{aligned} \frac{\partial d^a}{\partial p} &= \frac{\partial a}{\partial p} + \frac{\partial D^a}{\partial n}\frac{\partial d^n}{\partial p} \\ \frac{\partial d^n}{\partial p} &= \frac{\partial D^n}{\partial p} + \frac{\partial D^n}{\partial a}\frac{\partial d^a}{\partial p} \end{aligned}$$

and with content moderation

$$\begin{aligned} \frac{\partial d^a}{\partial m} &= \frac{\partial D^a}{\partial m} + \frac{\partial D^a}{\partial n}\frac{\partial d^n}{\partial m} \\ \frac{\partial d^n}{\partial m} &= \frac{\partial D^n}{\partial m} + \frac{\partial D^n}{\partial a}\frac{\partial d^a}{\partial m} \end{aligned}$$

The above expressions can be rearranged to obtain

$$\begin{aligned} \frac{\partial d^a}{\partial p} &= \frac{\frac{\partial D^a}{\partial p}}{1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} < 0, \quad \frac{\partial d^n}{\partial p} = \frac{\frac{\partial D^n}{\partial a}\frac{\partial D^a}{\partial p}}{1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} < 0 \\ \frac{\partial d^a}{\partial m} &= \frac{\frac{\partial D^a}{\partial m} + \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial m}}{1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}0, \quad \frac{\partial d^n}{\partial m} = \frac{\frac{\partial D^n}{\partial m} + \frac{\partial D^n}{\partial a}\frac{\partial D^a}{\partial m}}{1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}0. \end{aligned} \tag{18}$$

Consider the maximization problem of the platform when choosing m and p simultaneously. From the

first-order conditions, and using the above expressions, it follows that

$$p = -\frac{d^a(p,m)}{\frac{\partial d^a}{\partial p}} = -\frac{d^a(m,p)(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})}{\frac{\partial D^a}{\partial p}},$$
$$C'(m) = p\frac{\partial d^a}{\partial m} = p\left(\frac{\frac{\partial D^a}{\partial m}+\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial m}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right).$$

Denote MR the marginal gain from moderation such that $MR = p^* \frac{\partial d^a}{\partial m}|_{m=m^*}$. By using p above, we have the following expression

$$MR(m^*) = -\frac{d^a(m^*,p^*)(\frac{\partial D^a}{\partial m}+\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial m})}{\frac{\partial D^a}{\partial p}} = -\frac{d^a(m^*,p^*)\Psi}{\frac{\partial D^a}{\partial p}}, \tag{19}$$

where

$$\Psi = \frac{\partial D^a}{\partial m}+\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial m}$$

represents the platform's elasticity to moderation. The optimal level of content moderation is implicitly defined by the following expression $MR(m^*) = C'(m^*)$, where the latter term accounts for the marginal costs.

Proof of Proposition 1

Consider how the equilibrium variables change with marginal gains from moderation (i.e either changes in brand risk or changes in user's preference for moderated contents), let us first consider how demands on both sides of the market react. This boils down to understand how market parameters change in Ψ.

Now, consider the problem of a platform and recall the first-order conditions such that there exists a duple $(p,m) = (p^*,m^*)$ satisfying the following two expressions:

$$p^*\frac{\partial D^a}{\partial p} + d^a(m^*,p^*)(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) = 0,$$
$$C'(m^*)\frac{\partial D^a}{\partial p} + d^a(m^*,p^*)\Psi = 0.$$

By differentiating the above expressions with respect to Ψ, we have the following

$$0 = \frac{\partial D^a}{\partial p}\frac{\partial p^*}{\partial \Psi} + \frac{\partial d^a(m^*,p^*)}{\partial \Psi}(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}),$$
$$0 = C''(m)\frac{\partial m^*}{\partial \Psi}\frac{\partial D^a}{\partial p} + \frac{\partial d^a(m^*,p^*)}{\partial \Psi}\Psi + d^a(m^*,p^*).$$

Using the chain rule, $\frac{\partial d^a(p^*,m^*)}{\partial \Psi} = \frac{\partial d^a}{\partial \Psi} + \frac{\partial d^a}{\partial m}\frac{\partial m^*}{\partial \Psi} + \frac{\partial d^a}{\partial p}\frac{\partial p^*}{\partial \Psi}$, we then have

$$0 = \frac{\partial D^a}{\partial p}\frac{\partial p^*}{\partial \Psi} + (1 - \frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})\left(\frac{\partial d^a}{\partial \Psi} + \frac{\partial d^a}{\partial m}\frac{\partial m^*}{\partial \Psi} + \frac{\partial d^a}{\partial p}\frac{\partial p^*}{\partial \Psi}\right),$$
$$0 = C''(m)\frac{\partial m^*}{\partial \Psi}\frac{\partial D^a}{\partial p} + \left(\frac{\partial d^a}{\partial \Psi} + \frac{\partial d^a}{\partial m}\frac{\partial m^*}{\partial \Psi} + \frac{\partial d^a}{\partial p}\frac{\partial p^*}{\partial \Psi}\right)\Psi + d^a(m^*,p^*).$$

Using (18) and exploiting $\frac{\partial d^a}{\partial \Psi} = \frac{\frac{\partial D^a}{\partial \Psi}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} < 0$, we have

$$0 = 2\frac{\partial D^a}{\partial p}\frac{\partial p^*}{\partial \Psi} + \frac{\partial m^*}{\partial \Psi}\Psi + \frac{\partial D^a}{\partial \Psi},$$
$$0 = C''(m)\frac{\partial m^*}{\partial \Psi}\frac{\partial D^a}{\partial p} + \left(\frac{\partial D^a}{\partial \Psi} + \Psi\frac{\partial m^*}{\partial \Psi} + \frac{\partial D^a}{\partial p}\frac{\partial p^*}{\partial \Psi}\right)\frac{\Psi}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} + d^a(m^*,p^*).$$

Rearranging the above expressions, we have

$$-\frac{\partial D^a}{\partial \Psi} = 2\frac{\partial D^a}{\partial p}\frac{\partial p^*}{\partial \Psi} + \frac{\partial m^*}{\partial \Psi}\Psi,$$

and

$$-d^a(m^*,p^*) - \frac{\partial D^a}{\partial \Psi}\frac{\Psi}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} = \frac{\Psi}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\frac{\partial p^*}{\partial \Psi}\frac{\partial D^a}{\partial p} + \frac{\partial m^*}{\partial \Psi}\left(C''(m)\frac{\partial D^a}{\partial p} + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right),$$

Following the Implicit Function Theorem, then have

$$\begin{pmatrix} 2\frac{\partial D^a}{\partial p} & \Psi \\ \frac{\Psi\frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} & \frac{\partial D^a}{\partial p}C''(m) + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} \end{pmatrix}\begin{pmatrix} \frac{\partial p^*}{\partial \Psi} \\ \frac{\partial m^*}{\partial \Psi} \end{pmatrix} = \begin{pmatrix} -\frac{\partial D^a}{\partial \Psi} \\ -d^a(m^*,p^*) - \frac{\Psi\frac{\partial D^a}{\partial \Psi}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} \end{pmatrix}. \tag{20}$$

Using the Cramer's rule, we then get

$$\frac{\partial p^*}{\partial \Psi} = \frac{\det\begin{pmatrix} -\frac{\partial D^a}{\partial \Psi} & \Psi \\ -d^a(m^*,p^*) - \frac{\Psi\frac{\partial D^a}{\partial \Psi}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} & \frac{\partial D^a}{\partial p}C''(m) + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} \end{pmatrix}}{\det\begin{pmatrix} 2\frac{\partial D^a}{\partial p} & \Psi \\ \frac{\Psi\frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} & \frac{\partial D^a}{\partial p}C''(m) + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} \end{pmatrix}}$$

$$\frac{\partial m^*}{\partial \Psi} = \frac{\det\begin{pmatrix} 2\frac{\partial D^a}{\partial p} & -\frac{\partial D^a}{\partial \Psi} \\ \frac{\Psi \frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} & -d^a(m^*,p^*) - \frac{\Psi \frac{\partial D^a}{\partial \Psi}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} \end{pmatrix}}{\det\begin{pmatrix} 2\frac{\partial D^a}{\partial p} & \Psi \\ \frac{\Psi \frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} & \frac{\partial D^a}{\partial p}C''(m) + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}} \end{pmatrix}}$$

The denominator of both terms is equal to

$$\frac{\partial D^a}{\partial p}\left(2C''(m)\frac{\partial D^a}{\partial p} + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right),$$

which is positive (ensuring convexity in costs) if, and only if,

$$C''(m) > -\frac{\Psi^2}{(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})2\frac{\partial D^a}{\partial p}}. \tag{21}$$

The numerator of $\frac{\partial p^*}{\partial \Psi}$ is

$$-\frac{\partial D^a}{\partial \Psi}\left(\frac{\partial D^a}{\partial p}C''(m)+\frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right)+\Psi\left(d^a(m^*,p^*)+\frac{\Psi\frac{\partial D^a}{\partial \Psi}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right) = \Psi d^a(m^*,p^*)-\frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}C''(m).$$

So that

$$\frac{\partial p^*}{\partial \Psi} = \frac{\Psi d^a(m^*,p^*) - \frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}C''(m)}{\frac{\partial D^a}{\partial p}(2C''(m)\frac{\partial D^a}{\partial p}+\frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}})}.$$

As the denominator is positive, the effect depends on the numerator, such that $sign(\frac{\partial p^*}{\partial \Psi}) > 0$ if $\Psi d^a(m^*,p^*) > \frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}C''(m)$. Note that the LHS and the RHS are both positive. Hence, $\frac{\partial p^*}{\partial \Psi}$ is positive if Ψ and $d^a(m^*,p^*)$ is large enough while $C''(m)$ is low enough. Note that for $\Psi < 0$, we have $m^* = 0, C(m^* = 0) = 0$, which drives $\frac{\partial p^*}{\partial \Psi} < 0$.

Turning to the numerator of $\frac{\partial m^*}{\partial \bar{\lambda}}$, we then have

$$-2\frac{\partial D^a}{\partial p}\left(d^a(m^*,p^*)+\frac{\Psi\frac{\partial D^a}{\partial \Psi}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right)+\frac{\Psi\frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\frac{\partial D^a}{\partial \Psi} = -2\frac{\partial D^a}{\partial p}d^a(m^*,p^*)-\frac{\Psi\frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}.$$

So that

$$\frac{\partial m^*}{\partial \Psi} = -\frac{2\frac{\partial D^a}{\partial p}d^a(m^*,p^*) + \frac{\Psi\frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}}{\frac{\partial D^a}{\partial p}(2C''(m)\frac{\partial D^a}{\partial p}+\frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}})}.$$

As the denominator is positive, the effect depends on the numerator, such that $sign(\frac{\partial m^*}{\partial \Psi}) > 0$ if $-2\frac{\partial D^a}{\partial p}d^a(m^*,p^*) > \frac{\Psi \frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}$. Hence, $\frac{\partial m^*}{\partial \Psi} > 0$ if Ψ is low enough while $d^a(m^*,p^*)$ is large enough. Recall that convexity in costs need to be satisfied. Rearranging (21), this requires

$$\Psi < \Psi^c \equiv \sqrt{-2C''(m)\frac{\partial D^a}{\partial p}(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})}.$$

Call Ψ^p and Ψ^m the critical value of Ψ such that the numerators of $\frac{\partial p^*}{\partial \Psi}$ and $\frac{\partial m^*}{\partial \Psi}$ go to zero, then

$$\Psi^p \equiv \frac{\partial D^a}{\partial \Psi}\frac{\partial D^a}{\partial p}C''(m)/d^a(m^*,p^*)$$

$$\Psi^m \equiv= -2d^a(m^*,p^*)(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})/\frac{\partial D^a}{\partial \Psi}.$$

Denote by

$$\tilde{C} \equiv -2\Big(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}\Big)\frac{d^a(m^*,p^*)^2}{(\frac{\partial D^a}{\partial \Psi})^2\frac{\partial D^a}{\partial p}}$$

We find that when :

- $C''(m) > \tilde{C} \Longrightarrow \Psi^p > \Psi^c > \Psi^m$
- $C''(m) < \tilde{C} \Longrightarrow \Psi^p < \Psi^c < \Psi^m$

As a result, if $C''(m) < \tilde{C}$:

- If $0 < \Psi \leq \Psi^p$, then $\frac{\partial p^*}{\partial \Psi} < 0$ and $\frac{\partial m^*}{\partial \Psi} > 0$;
- If $\Psi^p < \Psi \leq \Psi^c$, then $\frac{\partial p^*}{\partial \Psi} > 0$ and $\frac{\partial m^*}{\partial \Psi} > 0$;

Hence, m^* always increases with Ψ in the relevant parameter space, whereas p^* is U-shaped in Ψ.

Next, if $C''(m) > \tilde{C}$, then :

- If $0 \leq \Psi^m$, then $\frac{\partial p^*}{\partial \Psi} < 0$ and $\frac{\partial m^*}{\partial \Psi} > 0$;
- If $\Psi^m < \Psi \leq \Psi^c$, then $\frac{\partial p^*}{\partial \Psi} < 0$ and $\frac{\partial m^*}{\partial \Psi} < 0$;

Hence, in the relevant space, p^* always declines with Ψ while m^* is inverted U-shaped in Ψ.

Proof of Proposition 2

Consider the case in which platforms compete. First, we must assume $c > \underline{c} = \frac{(\overline{\lambda}(2\tau\overline{v}+r\gamma)-r\overline{v}\overline{\phi})^2}{16\overline{v}(\tau\overline{v}+r\gamma)(2\tau\overline{v}+r\gamma)}$. In the second period of the game, by using (11) and (12), we can determine the number of users and

advertisers in platform i as

$$n_i = \frac{1}{2} + \frac{\overline{v}\overline{\phi}(\theta(m_i) - \theta(m_j)) + \gamma(2(p_j - p_i) + \overline{\lambda}(l(m_i) - l(m_j)))}{4(\tau\overline{v} + \gamma \cdot r)} \tag{22}$$

$$a_i = 1 - \frac{\overline{v}(r\overline{\phi}(\theta(m_i) - \theta(m_j)) + 2\tau(r + 2p_i - l(m_i)\overline{\lambda})) + r\gamma(2(r - p_i - p_j) - \overline{\lambda}(l(m_i) + l(m_j))}{4\overline{v}(\tau\overline{v} + \gamma r)} \tag{23}$$

Using these expressions in (15), from the first-order conditions we can solve for the symmetric equilibrium prices and moderation policies $p_i^* = p_j^*$ and $m_i^* = m_j^*$. For ease of exposition, let us define the following critical values of λ.

$$\lambda_{11} \equiv \frac{\overline{\phi} r \overline{v}}{(\tau\overline{v} + \gamma r)}, \quad \lambda_{12} \equiv \frac{4c(4\tau\overline{v} + 3\overline{\gamma}r)}{(2\overline{v} + r)(2\overline{v}\tau + \gamma r)} + \frac{\overline{\phi} r \overline{v}}{2\tau\overline{v} + \gamma r}.$$

The optimal values of m^* and p^* for a symmetric equilibrium. Formally, when $\overline{\lambda} < \lambda_{11}$, $m_i^* = m_j^* = 0$, and

$$p_i^* = p_j^* = \frac{(2\overline{v} + r - \overline{\lambda})(\tau\overline{v} + \gamma r)}{4\tau\overline{v} + 3\gamma r}.$$

When $\lambda_{11} < \overline{\lambda} < \lambda_{12}$, an interior solution exists, with $m_i^* = m_j^* \in [0, 1]$ defined as follows

$$p_i^* = p_j^* = \frac{4c\overline{v}(2\overline{v} + r - \overline{\lambda})(\tau\overline{v} + \gamma r)}{16c\tau\overline{v}^2 + ((\overline{\lambda\phi} + 12c\gamma)r - 2\overline{\lambda}^2\tau)\overline{v} - \gamma\overline{\lambda}^2 r},$$
$$m_i^* = m_j^* = \frac{(2\overline{v} + r - \overline{\lambda})(\overline{\lambda}(2\tau\overline{v} + \gamma r) - \overline{\phi} r \overline{v})}{16c\tau\overline{v}^2 + ((\overline{\lambda\phi} + 12c\gamma)r - 2\overline{\lambda}^2\tau)\overline{v} - \gamma\overline{\lambda}^2 r}.$$

When $\overline{\lambda} \geq \lambda_{12}$, $m_i^* = m_j^* = 1$ and

$$p_i^* = p_j^* = \frac{(2\overline{v} + r)(\tau\overline{v} + \gamma r)}{4\tau\overline{v} + 3\gamma r},$$

Proof of Proposition 3

Denote $\tilde{c} := \frac{\overline{\lambda}(\overline{\lambda}\gamma + \overline{v}\overline{\phi})}{4\overline{v}\gamma}$ an intermediate moderation cost which satisfies the assumption of convexity, we compute the (negative) derivatives of m_i^* and p_i^* with respect to τ. As a remark, we know that $\overline{\lambda} < r + 2\overline{v}$ to ensure non-negative prices when m^*. Then, call $p^{SH} = p_i^* = p_j^*$, $m^{SH} = m_i^* = m_j^*$. It follows that

$$-\frac{\partial m^{SH}}{\partial \tau} < 0, \qquad -\frac{\partial a^{SH}}{\partial \tau} < 0,$$

and

$$-\frac{\partial p^{SH}}{\partial \tau} = \frac{4cr\overline{v}^2(r+2\overline{v}-\overline{\lambda})(4c\overline{v}\gamma-\overline{\lambda}(\gamma\overline{\lambda}+\overline{v}\overline{\phi}))}{(4c\overline{v}(4\tau\overline{v}+3r\gamma)-\overline{\lambda}(2\tau\overline{v}\overline{\lambda}+r\gamma\overline{\lambda}-r\overline{v}\overline{\phi}))^2},$$

with the latter expression being negative for $c < \tilde{c}$, and positive otherwise.

Note that if $\gamma = 0$ (no nuisance from ads), then $-\frac{\partial p^{SH}}{\partial \tau} < 0$ and $-\frac{\partial m^{SH}}{\partial \tau} < 0$.

Proof of Proposition 4

Once again, consider the problem of a platform for a given tax $f < p^*$. Assuming interior solutions, the first-order conditions are:

$$\begin{aligned}\frac{\partial \Pi}{\partial p}(p^*, m^*) =&(p^* - f)\frac{\partial d^a}{\partial p} + d^a(p^*, m^*) = 0,\\ \frac{\partial \Pi}{\partial m}(p^*, m^*) =&(p^* - f)\frac{\partial d^a}{\partial m} - C'(m^*) = 0\end{aligned} \tag{24}$$

for some $(p, m) = (p^*, m^*)$. To understand the effect of f on p^* and m^*, differentiate the above system of equations with respect to f so as to obtain:

$$\begin{aligned}\frac{\partial d^a}{\partial p}(\frac{\partial p^*}{\partial f} - 1) + \frac{\partial d^a}{\partial f} =&0,\\ \frac{\partial d^a}{\partial m}(\frac{\partial p^*}{\partial f} - 1) - C''(m)\frac{\partial m^*}{\partial f} =&0\end{aligned} \tag{25}$$

As d^a depends on f only through p^* and m^*, using the chain rule, we can define the following $\frac{\partial d^a}{\partial f} = \frac{\partial d^a}{\partial m}\frac{\partial m^*}{\partial f} + \frac{\partial d^a}{\partial p}\frac{\partial p^*}{\partial f}$. As a result, we can define the above system of equations as

$$\begin{pmatrix} 2\frac{\partial d^a}{\partial p} & \frac{\partial d^a}{\partial m} \\ \frac{\partial d^a}{\partial m} & -C''(m) \end{pmatrix} \begin{pmatrix} \frac{\partial p^*}{\partial f} \\ \frac{\partial m^*}{\partial f} \end{pmatrix} = \begin{pmatrix} \frac{\partial d^a}{\partial p} \\ \frac{\partial d^a}{\partial m} \end{pmatrix} \tag{26}$$

By the Implicit Function Theorem and the Cramer's Rule, we get

$$\frac{\partial p^*}{\partial f} = \frac{\det\begin{pmatrix} \frac{\partial d^a}{\partial p} & \frac{\partial d^a}{\partial m} \\ \frac{\partial d^a}{\partial m} & -C''(m) \end{pmatrix}}{\det\begin{pmatrix} 2\frac{\partial d^a}{\partial p} & \frac{\partial d^a}{\partial m} \\ \frac{\partial d^a}{\partial m} & -C''(m) \end{pmatrix}} \tag{27}$$

$$\frac{\partial m^*}{\partial f} = \frac{\det\begin{pmatrix} 2\frac{\partial d^a}{\partial p} & \frac{\partial d^a}{\partial p} \\ \frac{\partial d^a}{\partial m} & \frac{\partial d^a}{\partial m} \end{pmatrix}}{\det\begin{pmatrix} 2\frac{\partial d^a}{\partial p} & \frac{\partial d^a}{\partial m} \\ \frac{\partial d^a}{\partial m} & -C''(m) \end{pmatrix}}. \tag{28}$$

Using $\frac{\partial d^a}{\partial m}$ and $\frac{\partial d^a}{\partial p}$ as previously defined, we can rewrite the denominator of both expressions as equal to

$$-2C''(m)\frac{\partial d^a}{\partial p} - \left(\frac{\partial d^a}{\partial m}\right)^2 = -\frac{1}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\left(2C''(m)\frac{\partial D^a}{\partial p} + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right).$$

The numerator of $\frac{\partial p^*}{\partial f}$ is equal to

$$-C''(m)\frac{\partial d^a}{\partial p} - \left(\frac{\partial d^a}{\partial m}\right)^2 = -\frac{1}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\left(C''(m)\frac{\partial D^a}{\partial p} + \frac{\Psi^2}{1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}}\right).$$

Hence, we obtain

$$\frac{\partial p^*}{\partial f} = \frac{C''(m)\frac{\partial D^a}{\partial p}(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) + \Psi^2}{2C''(m)\frac{\partial D^a}{\partial p}(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) + \Psi^2},$$

which is positive if, and only if, the numerator is negative, that is, when

$$\frac{\partial p^*}{\partial f} > 0 \iff -C''(m)\frac{\partial D^a}{\partial p}(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a}) > \Psi^2.$$

The numerator of $\frac{\partial p^*}{\partial f}$ is equal to

$$2\frac{\partial d^a}{\partial p}\frac{\partial d^a}{\partial m} - \frac{\partial d^a}{\partial m}\frac{\partial d^a}{\partial p} = \frac{\frac{\partial D^a}{\partial p}\Psi}{(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})^2}.$$

Hence, we obtain

$$\frac{\partial m^*}{\partial f} = -\frac{\frac{\partial D^a}{\partial p}\Psi(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})^{-1}}{2C''(m)\frac{\partial D^a}{\partial p}(1-\frac{\partial D^a}{\partial n}\frac{\partial D^n}{\partial a})+\Psi^2} < 0,$$

as the numerator is always negative.

Competition with multihoming users.

In social networks, given the absence of a real monetary cost paid by users, multihoming decisions are most dominant. In this section, we relax the assumption of singlehoming users to verify how content moderation changes when this is the case. One can easily note that when some users multihome, multihoming advertisers might interact with the same users twice, thereby placing wasteful ads. This might force the platform to reduce the ad price or please advertisers with tight content moderation policy. On the other hand, the presence of multihoming users relaxes the competition between platforms for the marginal consumer and this reduces the incentive to engage in a lax moderation policy.

To understand the optimal business strategies, we present the following variations of the model with singlehoming users. The utility of singlehoming users is as the benchmark model, whereas the utility of the multihoming consumer on the platform i is $U_i^{mh} = (1+\sigma)u+\phi(\theta(m_i)+\theta(m_j))-\gamma(a_i+a_j)-\tau$, with $\sigma \in (0,1)$ represents the marginal utility that consumers gain when joining the second platform. By equating the utility of the consumer singlehoming on platform i and multihoming, we can derive the following critical values which identify the users indifferent between singlehoming on platform 1 and multihoming and singlehoming on platform 2 and multihoming.

$$\tilde{x}_1 \equiv 1 - \frac{u\sigma + \phi\theta(m_2) - \gamma a_2}{\tau}, \qquad \tilde{x}_2 \equiv \frac{u\sigma + \phi\theta(m_1) - \gamma a_1}{\tau}$$

which implies that the demand of platform 1 is equal to $n_1 := \frac{1}{\phi}\int_0^{\overline{\phi}} \tilde{x}_2 \mathrm{d}\phi$, and $n_2 := \frac{1}{\phi}(\overline{\phi} - \int_0^{\overline{\phi}} \tilde{x}_1 \mathrm{d}\phi)$. However, as the multihoming advertisers observe some users twice, the value of their interactions is generated only once. This implies that multihoming users, n^{MH} account for only half of the value on each platform. Such an assumption finds justification in a recent work by Liu (2018). Indeed, an advertiser joining platform i obtains the following utility

$$V_i = v + r(n_i^{SH} + \frac{n^{MH}}{2}) - \lambda\theta(m_i) - p_i$$

where $n_1^{SH} = \frac{1}{\phi}\int_0^{\overline{\phi}} \tilde{x}_1 \mathrm{d}\phi$, $n_2^{SH} = \frac{1}{\phi}(\overline{\phi} - \int_0^{\overline{\phi}} \tilde{x}_2 \mathrm{d}\phi)$, and $n^{MH} = \frac{1}{\phi}\int_0^{\overline{\phi}} (\tilde{x}_2 - \tilde{x}_1)\mathrm{d}\phi$.

We show that allowing users to multihome does not change the platform strategies. As a result, prices and content moderation efforts are identical to those arising with singlehoming users and, as such, the

effect of more intense competition on equilibrium outcome does not change. In practice, increasing the fierceness of the competition by reducing transportation costs leads to less moderation in equilibrium. Similarly, we also verify that the effect of intense competition the equilibrium ad price depends on moderation cost and can go both ways.

Appendix B

To get a better picture of the trade-offs present in the benchmark model with a monopolist platform, in this Appendix we present an example with a uniform distribution of preferences of advertisers and users.

Monopolist Platform

First, let us specify the utility of users and advertisers, respectively. An Internet user is defined by a duple $(u, \phi) \sim U[0, \overline{u}] \times U[0, \overline{\phi}]$, such that both u and ϕ are uniformly and independently distributed. We only consider the case in which users dislike moderation, but their tastes are heterogeneous, i.e., $\underline{\phi}$ in Section 2, is set equal to zero. This simplifying assumption allows us to focus on the case in which advertisers' and users' preferences over moderation are conflicting - which is the most insightful case. As a result, if some users had a negative ϕ, the platform would have a slightly higher incentive to increase content moderation effort as users' preferences would converge towards those of advertisers. We also assume that advertisers are heterogeneous: an advertiser is a duple $(v, \lambda) \sim U[0, \overline{v}] \times U[0, \overline{\lambda}]$, such that both v and λ are uniformly and independently distributed. This allows us to capture differences in the net gain from purchasing an ad campaign on a social network, depending on the long-term net gain from being on the platform (see equation (4)). To make the model tractable, we assume that preferences for safe content have a stronger effect on both users and advertisers than the ones for unsafe content. Here, we present the assumptions on users and advertisers' preferences that allow us to compute equilibria. On the user side, $\overline{u}$ needs to be sufficiently large such that users who do not value unsafe content (say $\phi = 0$) do not refrain from using platform when their intrinsic preference for the platform, u, is large enough. Conversely, we also assume that $\overline{\phi}$ is low enough such that users who largely enjoy unsafe content ($\phi = \overline{\phi}$) may decide not to visit the social network if the intrinsic utility they get from the platform (i.e., access to a mass 1 of safe content) is not high enough. More generally, the above assumptions imply that preferences for safe content have a stronger effect on user decisions than aversion for the moderation of unsafe content.

On the advertiser side, we dig into the nature of the net gains from brand safety, Ω, as defined by (4). Hence, we let $\overline{v}$ be such that advertisers not concerned with brand safety (e.g., $\lambda = 0$) will not display ads when not benefiting enough from interactions with the unit mass of safe content. We also assume that the $\overline{\lambda}$ is low enough, such that some advertisers with a strong preference for brand safety ($\lambda = \overline{\lambda}$) may still display ads when deriving high gain from being exposed to the mass 1 of safe content. More generally, such an assumption implies that v has a stronger impact on advertisers' decisions to buy an ad than brand risk issues alone.

To provide clear insights on the optimal ad price p^* and moderation policy m^* and to compute

the equilibrium, we assume that the mass of unsafe content is a linear and decreasing function of m. Similarly, we assume quadratic moderation costs.

$$\theta(m) = 1 - m \quad \text{and} \quad C(m) = c\frac{m^2}{2}, \text{ with } \quad c > 0.$$

With the above specifications on hold, on the second stage, advertisers decide whether to display an ad and users whether to join the platform. This results in the following demand:

$$d^a(m,p) = \frac{\theta(m)(r\overline{\phi} - \overline{u}\overline{\lambda}) + 2\overline{u}(r + \overline{v} - p)}{2(\overline{uv} + \gamma r)}. \tag{29}$$

$$d^n(m,p) = \frac{(2\overline{u} + \theta(m)\overline{\phi} - 2\gamma)\overline{v} + 2\gamma p + \theta(m)\gamma\overline{\lambda}}{2(\overline{uv} + \gamma r)}.$$

By maximizing (1) and rearranging it, we obtain

$$\begin{aligned} p(m) =& \frac{2\overline{u}(\overline{v} + r) + r\overline{\phi}\theta(m) - \overline{u}\overline{\lambda}\theta(m)}{4\overline{u}}, \\ C'(m) =& p\frac{(r\overline{\phi}\theta'(m) - \overline{u}\overline{\lambda}l'(m)}{2(\overline{uv} + \gamma r)}. \end{aligned}$$

Denote $\Psi = \overline{\lambda}\overline{u} - \overline{\phi}r$ the platform profit elasticity with respect to moderation, convexity in costs is ensured by the following expression

$$c > \frac{\Psi^2}{8\overline{u}(\overline{uv} + \gamma r\overline{u})} \tag{30}$$

which is equivalent to the one in (21). Using $\theta(m) = 1 - m$, we obtain the following equilibrium outcomes

$$m^* = \begin{cases} 0 \text{ if } \overline{\lambda}\overline{u} - \overline{\phi}r \leq 0, \\ \frac{\Psi(2\overline{u}(\overline{v}+r)-\Psi)}{8c\overline{u}(\overline{uv}+\gamma r)-\Psi^2} \text{ if } \frac{4c(\overline{uv}+\gamma r)}{\overline{v}+r} > \overline{\lambda}\overline{u} - \overline{\phi}r \geq 0. \\ 1 \text{ if } \overline{\lambda}\overline{u} - \overline{\phi}r \geq \frac{4c(\overline{uv}+\gamma r)}{\overline{v}+r}. \end{cases} \tag{31}$$

$$p^* = \begin{cases} \frac{\overline{u}(\overline{v}+r)-\Psi}{2\overline{u}} \text{ if } \overline{\lambda}\overline{u} - \overline{\phi}r \leq 0, \\ \frac{2c(\overline{uv}+\gamma r)(2\overline{u}(\overline{v}+r)-\Psi)}{8c\overline{u}(\overline{uv}+\gamma r)-\Psi^2} \text{ if } \frac{4c(\overline{uv}+\gamma r)}{\overline{v}+r} > \overline{\lambda}\overline{u} - \overline{\phi}r \geq 0. \\ \frac{\overline{v}+r}{2} \text{ if } \overline{\lambda}\overline{u} - \overline{\phi}r \geq \frac{4c(\overline{uv}+\gamma r)}{\overline{v}+r}. \end{cases} \tag{32}$$

First, we begin to prove the corner solutions. Suppose the platform chooses a no moderation policy such that $m^* = 0$. This happens whenever $2\overline{u}(\overline{v}+r) < \Psi$ and $\Psi > 0$ or $2\overline{u}(\overline{v}+r) > \Psi$ and $\Psi < 0$. One can easily note that former would imply negative price, which is not possible, while the latter is instead compatible with positive user and advertisers demand. Hence $m^* = 0$ when $\Psi \leq 0$. Suppose now the platform enforces full moderation, $m^* = 1$. This happens for any $\frac{4c(\overline{uv}+\gamma r)}{\overline{v}+r} > \overline{\lambda}\overline{u} - \overline{\phi}$. In both case, p^* is retrieved by substituting $m = \{0, 1\}$ into $p(m)$ to obtain the expression defined in the first and the third lines in (32).

Second, consider an interior solution such that m^* belongs to $(0, 1)$. This happens if $\frac{4c(\overline{uv}+\gamma r)}{\overline{v}+r} > \overline{\lambda}\overline{u} - \overline{\phi}r \geq 0$, and so the price is the one defined in the second line in (32).

Using the equilibrium outcomes, we can then derive the profits of the platform as follows.

$$\Pi^* = \begin{cases} \frac{(\overline{v}+r)(\overline{u}(\overline{v}+r)-\Psi)}{4(\overline{uv}+\gamma r)} \text{ if } \overline{\lambda}\overline{u} < \overline{\phi}r, \\ \\ \frac{c(2\overline{u}(\overline{v}+r)-\Psi^2)}{2(8c\overline{u}(\overline{uv}+\gamma r)-\Psi^2)} \text{ if } \frac{4c(\overline{uv}+\gamma r)+\overline{\phi}r(\overline{v}+r)}{\overline{v}+r} > \overline{\lambda}\overline{u} \geq \overline{\phi}r. \\ \\ \frac{\overline{u}(\overline{v}^2+r(2\overline{v}+r))-2c(\overline{uv}+\gamma r)}{4(\overline{uv}+\gamma r)} \text{ if } \overline{\lambda}\overline{u} \geq \frac{4c(\overline{uv}+\gamma r)+\overline{\phi}r(\overline{v}+r)}{\overline{v}+r}. \end{cases}$$

In what follows, we provide support for the results in the general model and, more specifically, for those in Proposition 1. Hence, we study how p^* and m^* vary with Ψ.

First, notice that numerators for both m^* and p^* cancels out when $\Psi = 2\overline{u}(\overline{v}+r)$. This happens as brand risk is sufficiently large such that the advertiser willingness to pay decreases. Second, rewrite condition (30) with respect to Ψ. We find that for both m^* and p^*, the numerator cancels out for higher value of Ψ than the denominator if

$$c < \frac{\overline{u}(\overline{v}+r)^2}{2(\overline{uv}+\gamma r)} \equiv \tilde{c} \tag{33}$$

Importantly, this leads to the following results. If moderation costs are sufficiently small, $c < \tilde{c}$, we find that $\frac{\partial^2 p^*}{\partial^2 \Psi} > 0$. This then implies that p^* admits a minimum in

$$\Psi = 2\overline{u}(\overline{v}-r) - 2\sqrt{\overline{u}(\overline{uv}(\overline{v}+2r) - 2c(\overline{vu}-\gamma r))}.$$

Hence, p^* is convex and U-shaped in Ψ. Similarly, we it can be easily verified that $\frac{\partial m^*}{\partial \Psi} > 0$ when $c < \frac{\Psi^2(\overline{v}+r)}{8(\overline{uv}+\gamma r)(\Psi-\overline{u}(\overline{v}+r))}$ - this is ensured by (33). To see it, consider that m^* admits an inflexion point in $\Psi = \overline{u}(\overline{v}+r)$, which is also the maximum of the numerator. Hence, for $\Psi < \overline{u}(\overline{v}+r)$, $\frac{\partial m^*}{\partial \Psi} > 0$ as the denominator decreases and the numerator converges to the maximum. For a higher value of Ψ, the numerator may pass its maximum, turning the shape of m from concave to convex. However, as the

denominator is shrinking faster than the numerator, we then have $\frac{\partial m^*}{\partial \Psi} > 0$

If moderation costs are sufficiently large, $c > \tilde{c}$, the numerator of p^* and m^* cancels out for existing value of Ψ that ensure profit concavity. In this case, we find $\frac{\partial p^*}{\partial \Psi} < 0$, until a point where Ψ is so high such that no the consumer demand disappears. By the same mechanism as the previous case, we find that $\frac{\partial^2 m^*}{\partial^2 \Psi} < 0$, and that m^* admits a maximum in

$$\Psi = \frac{4c(\overline{uv} + \gamma r) - 2^{\frac{3}{2}}\sqrt{c(\overline{uv} + \gamma r)(2c(\overline{uv} + \gamma r) + \overline{u}(\overline{v} + r)^2)}}{\overline{v} + r}.$$

Hence, m^* is inverted U-shaped in Ψ.

To sum up,

- if $c \leq \tilde{c}$, p^* is convex in Ψ and $\frac{\partial m^*}{\partial \Psi} > 0$.
- if $c > \tilde{c}$, m^* is concave in Ψ and $\frac{\partial p^*}{\partial \Psi} < 0$.

Effect of a tax on ad-revenues

Assume a uniform distribution of preferences. The following results are presented. First, the effect on advertisers is such that

$$\frac{da}{df} = \frac{1}{2(\overline{uv} + \gamma r)}\left\{\Psi \frac{\partial m}{\partial f} - 2\overline{u}\frac{\partial p}{\partial f}\right\}$$

where $\Psi = \overline{\lambda}\overline{u} - \overline{\phi}r$ the *elasticity of profit with respect to moderation*. Its sign depends on the sign of $\frac{dm}{df}$ and $\frac{dp}{df}$. To see it, consider the second stage of the game. The platform chooses m and p to maximize profits. Rearranging the first order conditions, we can see that the total effect on moderation effort and prices comes from the solution of the following system of equations.

$$\begin{aligned} \frac{dm}{df} =& \frac{1}{2(\overline{uv} + \gamma r)}\left\{-1 + \frac{\partial p}{\partial f}\right\}, \\ \frac{dp}{df} =& \frac{1}{2} + \frac{1}{4}\left(\overline{\lambda} - \frac{r\overline{\phi}}{\overline{u}}\right)\frac{\partial m}{\partial f}, \end{aligned} \tag{34}$$

which then can be solved as follows:

$$\begin{aligned} \frac{dm}{df} =& -\frac{2\overline{u}f}{8c\overline{u}(\overline{uv} + \gamma r) - \Psi^2}, \\ \frac{dp}{df} =& \frac{4c\overline{u}(\overline{uv} + \gamma r) - \Psi^2}{8c\overline{u}(\overline{uv} + \gamma r) - \Psi^2}, \end{aligned} \tag{35}$$

The above results indicate that $\frac{dm}{df} < 0$, whereas $\frac{dp}{df} > (<)0$ depending on the sign of the numerator. When c is sufficiently large, i.e., $c > \overline{c} := \frac{\Psi^2}{4\overline{u}(\overline{uv}+\gamma r)}$, the price increases with more taxation. Else it decreases.

By substituting (35) into (5), the total effect on advertisers is negative, i.e.,

$$\frac{da}{df} = -\frac{4c\overline{u}}{8c\overline{u}(\overline{uv}+\gamma r) - \Psi^2} < 0. \tag{36}$$

I tre lavori qui pubblicati sono i vincitori dell'ottava edizione (2020) del Giorgio Rota Best Paper Award for Young Researchers *sul tema "Digital Transformation: analysis of Economic Impact and Potential" e sono stati presentati alla* Giorgio Rota Conference *tenutasi online l'11 maggio 2020.*
Il Premio, istituito dal Centro Einaudi in memoria dell'economista torinese che ne era stato uno degli animatori, è sostenuto dalla Fondazione CRT.

Il Centro di Ricerca e Documentazione Luigi Einaudi conduce attività di ricerca in proprio o su commissione di enti nazionali o internazionali, cura siti web e la pubblicazione di periodici, svolge opera di formazione di giovani studiosi e ricercatori, organizza seminari, conferenze, convegni. L'attività è finanziata attraverso il contributo di soci e sostenitori, nonché con i proventi delle ricerche. Il Centro, che non ha fini di lucro e dal 2006 gode del riconoscimento della Regione Piemonte, è nato a Torino nel 1963 come libera associazione di imprenditori e intellettuali, grazie all'iniziativa e all'impegno di Fulvio Guerrini.

ISBN 978-88-94960-17-4