UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

“Digital Colonization" of Highly Regulated Industries: An Analysis of Big Tech Platforms' Entry into Healthcare and Education

Özalp, H.; Ozcan, P.; Dinckol, D.; Zachariadis, M.; Gawer, A.

DOI
10.1177/00081256221094307

Publication date
2022

Document Version
Final published version

Published in
California Management Review

License
CC BY

Link to publication

Citation for published version (APA):
Özalp, H., Ozcan, P., Dinckol, D., Zachariadis, M., & Gawer, A. (2022). “Digital Colonization" of Highly Regulated Industries: An Analysis of Big Tech Platforms' Entry into Healthcare and Education. *California Management Review*, *64*(4), 78-107. https://doi.org/10.1177/00081256221094307

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:10 Nov 2026

Check for updates

BerkeleyHaas
Haas School of Business
University of California Berkeley

Special Issue on Multi-Sided Platforms

California Management Review
2022, Vol. 64(4) 78–107
© The Regents of the
University of California 2022

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00081256221094307
journals.sagepub.com/home/cmr
SAGE

"Digital Colonization" of Highly Regulated Industries:

An Analysis of Big Tech Platforms' Entry into Health Care and Education

Hakan Ozalp[1], Pinar Ozcan[2], Dize Dinckol[2], Markos Zachariadis[3], and Annabelle Gawer[4]

SUMMARY

Digital platforms have disrupted many sectors but have not yet visibly transformed highly regulated industries. This study of Big Tech entry in healthcare and education explores how platforms have begun to enter highly regulated industries systematically and effectively. It presents a four-stage process model of platform entry, which we term as "digital colonization." This involves provision of data infrastructure services to regulated incumbents; data capture in the highly regulated industry; provision of data-driven insights; and design and commercialization of new products and services. The article clarifies platforms' sources of competitive advantage in highly regulated industries and concludes with managerial and policy recommendations.

KEYWORDS: platforms, entry modes, regulation, health care, education & industry, artificial intelligence

Over the past few years, digital platforms have disrupted competition and innovation across many sectors,[1] including retail, entertainment, hospitality, transportation, gaming, and music. Platform firms are now dominating the list of biggest firms

[1]University of Amsterdam, Amsterdam, The Netherlands
[2]University of Oxford, Oxford, UK
[3]The University of Manchester, Manchester, UK
[4]University of Surrey, Guildford, UK

by market capitalization, often referred to as "Big Tech" players. Recently, the prevalence of digital platforms has further increased in various industries as the COVID-19 pandemic amplified the role of digital services in people's lives, reshaping customer habits from how they shop, work, and entertain while skyrocketing the revenues of digital platforms.[2] These Big Tech firms are under scrutiny regarding how much value they return to end customers as they acquire, analyze, and take advantage of their data to boost profits and influence markets.

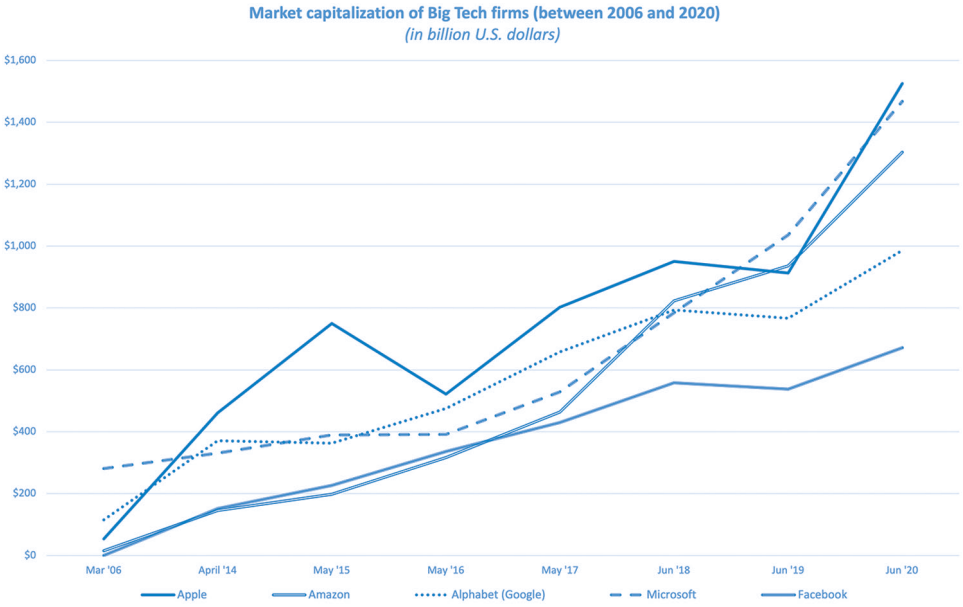
While platform firms have now become prevalent in many industries, highly regulated industries such as healthcare and education had lagged behind until recently, but there are clear signs that this has started to change. Considering these changes, we explore the entry paths of Big Tech platforms (more specifically Google (Alphabet), Amazon, Facebook (Meta), Apple, and Microsoft, also known as GAFAM) into highly regulated industries by looking at the prominent examples of healthcare and education in the context of the United States and United Kingdom, where they have been most active in these industries so far.

The Platform Business Model and the Role of Data

A platform creates value thanks to its advantages in connecting different users through enhanced matchmaking and facilitating transactions among them (e.g., by connecting customers and complementors).[3] Platforms can achieve rapid growth through highly scalable technological intermediation and reduction of various costs for transacting, matching, and innovating. Platform growth is further fueled by network effects, and this mechanism underpins how the value a user receives from a platform increases with each new user on the same side of the platform (i.e., direct network effects) and the other side of the platform (i.e., indirect network effects). More recently, there has also been the growing importance of data network effects, which refer to the increasing value users obtain from the platform in parallel with the amount of data the platform accumulates, such as better recommendations on Netflix. Thanks to their digital nature, platforms can connect various platform sides via digital interfaces[4] and, in the process, accumulate/leverage external resources (i.e., data) to develop relevant capabilities (i.e., algorithm-driven data analysis)[5] to improve further and expand their offering.

Due to their digital properties, use of data, and platform business models, certain technology companies have rapidly grown, becoming some of the largest and most influential firms globally (see Figure 1 for Big Tech firms' market caps). Big Tech firms started out as platforms with a single and focused intermediation activity (e.g., search engine). From there, they grew significantly in scope and entered new industries. Initially, they typically expanded into the space of their own complementors within their platform ecosystems (e.g., AmazonBasics competing with its own third-party sellers). Following this, they have entered related or adjacent sectors (e.g., Facebook acquisition of Instagram) or what may at first seem to be unrelated markets (e.g., Google acquisition of Waymo).[6]

FIGURE 1. Market capitalization of Big Tech firms.



Note: Based on data from: https://www.statista.com/statistics/216657/market-capitalization-of-us-tech-and-internet- companies/

Data sit at the heart of every digital platform. As such, the main logic underpinning the various market segment entries by platforms seems to aim to maximize data collection; enhance data network effects that they have already built across industries[7] to create more value; apply their data analysis capabilities; and take precedence over existing firms while improving products/services for consumers.[8]

This data-centric approach to platform growth and industry entry, however, regularly raises questions on data privacy, fair competition, and the balance of value creation and value capture in industries where platforms enter. These issues become even more critical in highly regulated industries where value creation becomes extremely important (e.g., patient lives saved by new technologies), and concerns around data privacy and fair competition are even more salient (e.g., medical or learning records already used by Google and others).

Platforms in Highly Regulated Industries

Despite the penetration and dominance of digital platforms in several industries, highly regulated sectors such as education, energy, finance, and healthcare appeared to have been left behind[9] due to high regulatory control creating barriers to entry for platforms.[10] Highly regulated industries typically have high entry barriers and high operational and compliance costs, as visible from the various regulations for the healthcare and education industries in Table 1.[11] Compared with other industries where regulatory interventions are typically "lighter" (e.g., taxi and transportation services), industries such as healthcare

TABLE 1. Healthcare and Education Industry Regulations.

Healthcare Industry Regulations		
Regulated Activities and Entities	**Main U.K. Regulations**	**Main U.S. Regulations**
Data Collection, Exchange, and Processing	• The Health Service (Control of Patient Information) Regulations • Data Protection Act • General Data Protection Regulations	• Health Insurance Portability and Accountability Act, Privacy and Security Rules • Health Information Technology for Economic and Clinical Health Act • Patient Safety and Quality Improvement Act
Providers (Hospitals, Care Homes, etc.)	• Health and Social Care Act • Care Quality Commission (Registration) Regulations • Care Act	• Centers for Medicare & Medicaid Services regulations • Joint Commission accreditation, certification, and standards • Emergency Treatment and Active Labor Act • Hospital Survey and Construction Act (Hill-Burton Act)
Services (e.g., Diagnosis, Treatments, Surgical Procedures)	• Health and Social Care Act • Care Quality Commission (Registration) Regulations	• Centers for Medicare & Medicaid Services regulations
Healthcare professionals	• Regulation of Healthcare Professionals • Regulation of Social Care Professionals	• State licensing boards • Centers for Medicare & Medicaid Services regulations • Stark Law
Medical Products (Pharmaceuticals, Medical Devices, etc.)	• Human Medicines Regulations • Food Information to Consumers Regulation • The Nutrition and Health Claims Regulation • Medical Devices Regulations • General Product Safety Regulations • Biocides Regulations	• FDA approval • Omnibus Budget Reconciliation Act • Prescription Drug Marketing Act • Medicine Equity and Drug Safety Act • Medicare Modernization Act
Insurance	• Provided by the state	• McCarran-Ferguson Act • Employee Retirement Income Security Act • Health Insurance Portability and Accountability Act • Affordable Care Act • Children's Health Insurance Authorization Act
Healthcare Market (Economic Regulation)	• Health and Social Care Act • Enterprise Act • Procurement, Patient Choice and Competition Regulations	• Prescription Drug Marketing Act • Affordable Care Act • Medicine Equity and Drug Safety Act • Medicare Modernization Act • National Health Planning Law

(continued)

TABLE 1. (continued)

Education Industry Regulations		
Regulated Activities and Entities	**Main U.K. Regulations**	**Main U.S. Regulations**
Data Collection, Exchange and Processing	• Data Protection Act • General Data Protection Regulations	Family Educational Rights and Privacy Act Children's Online Privacy Protection Rule Protection of Pupil Rights Amendment Electronic Communications Privacy Act Federal Information Security Management Act Gramm Leach Bliley Act Health Insurance Portability and Accountability Act Health Information Technology for Economic and Clinical Health
Providers (e.g., Schools, Universities)	• Education (Non-Maintained Special Schools) (England) Regulations • Further Education and Training Act • Disability Discrimination (Educational Institutions) Regulations • Higher Education and Research Act	Higher Education Act Americans with Disabilities Act Higher Education Opportunity Act Contracts with Third Party Servicer Federal Funding Accountability and Transparency Act Health Insurance Portability and Accountability Act
Services (e.g., Courses, Programs)	• Special Educational Needs and Disability Regulations • Apprenticeships, Skills, Children and Learning Act	Higher Education Act Age Discrimination Act
Professionals (e.g., Teachers)	• School Teachers (Recognition of Professional Qualifications) Regulations	State-level regulations for certification
Finance (e.g., Student Loans)	• Education (Student Loans) (Repayment) Regulations • Repayment of Student Loans and Postgraduate Master's Degree Loans Regulations	Higher Education Act Bankruptcy Abuse Prevention & Consumer Protection Act Truth in Lending Act
Education Market	• Higher Education and Research Act • Digital Economy Act	Higher Education Act Antitrust regulations (Sherman Antitrust Act, Clayton Antitrust Act, Federal Trade Commission Act)

and education are characterized by the heavy involvement of state and government actors. This is mainly because of the crucial strategic role these industries play in ensuring social welfare and boosting the country's economic growth

and development,[12] but also due to the associated social ramifications in terms of access, fairness, equality, privacy, and data sensitivity, as these factors directly tie to human and constitutional rights (e.g., "right to education"). Such state-controlled apparati, in addition to imposing a large set of rules and procedures upon private firms, often leave limited room for private actors to operate in, which presents a distinct challenge to market entrants.

Then, there is the thorny issue of data. Prospective digital platform entrants require data to develop new products or services, which calls for different strategies in highly regulated industries due to the need to capture and process *sensitive personal data*. If leaked or misused, such data can cause harm to individuals—for example, biometric data, genetic data, health-related data, race, or ethnicity data (typically held by healthcare providers), religious or philosophical beliefs (typically expressed in the context of education and recorded in essays, online educational platform discussions, and so on), and student education records.[13] This tends to raise the level of regulation further, thus exacerbating inhibition of new entry. Due to such considerations, digital platforms have, until recently, mostly been absent from highly regulated industries.

However, this is changing. Despite the challenges noted above, we observe that Big Tech firms are expanding their platforms into some of these highly regulated industries. Recent examples include Amazon acquiring U.S. online pharmacy Pillpack, Alphabet-Google partnering with the United Kingdom's National Health Service (NHS) for data sharing and developing AI-powered healthcare services, and U.S. universities partnering with Amazon to install Alexa in the dormitories and elsewhere. In 2020, the COVID-19 pandemic accelerated this trend further by causing the emergence of new initiatives. Examples include Google's subsidiary Verily offering COVID-19 testing and tracing, Google and Apple cooperating on mobile operating systems for COVID-19 contact tracing, Google Education expanding to support remote education, and Amazon offering COVID-19-specific Amazon Web Services (AWS) solutions for hospitals and research institutes.

Building on these trends, this article explores *how Big Tech platforms enter and compete in highly regulated industries.* Focusing on healthcare and education industries, we identify an entry pattern for these digital platforms, in which they typically begin as *suppliers of data-infrastructure services to incumbents* in the first phase. As incumbent service providers such as hospitals, schools, healthcare conglomerates typically lack capabilities in data management, they contract out these activities to Big Tech firms as technology service providers, aiming to reduce costs and improve services. In the second phase, Big Techs leverage their existing relationships as well as their data analysis capabilities (which they use to produce data-driven insights) to get access to the data already held by incumbent service providers. This indirect data capture (e.g., access to already collected data in a hospital), which they combine with their own direct data capture activities (e.g., through proprietary hardware such as Apple Watch, Google Tablet), then becomes an essential component of Big Tech firms' entry pathway into the targeted highly

regulated industry. As Big Tech firms combine the data they captured directly and indirectly, they can provide superior data-driven insights, which can add significant value to incumbent service providers (e.g., through saved lives, better learning outcomes, and lower costs). We find that a final component of entry for Big Tech firms is the design and commercialization of new products and services for the highly regulated industry target, where they may end up competing with their former clients over time.

Overall, our research suggests that Big Tech entry in highly regulated industries occurs via a process that we name "digital colonization," which we specify as composed of four stages: provision of data infrastructure services to incumbents; direct and indirect data capture in industry; provision of data-driven insights; and design and commercialization of new products and services. While Big Tech firms rarely end up directly offering the "primary service" (e.g., providing school education or becoming primary healthcare providers) in highly regulated industries, they change the power dynamics in these industries over time by commoditizing incumbent service providers, turning them into mere complementors while Big Tech firms control the data and become unique providers of critical, data-driven value.

Research Design and Methodology

To explore how Big Tech (platform) firms are entering highly regulated industries we have used a comparative case approach.[14] We focused on identifying entry patterns with particular attention to cross-platform, cross-country, and cross-industry differences. We relied on archival data collected online via LexisNexis searches with keywords including Big Tech firms' names and healthcare and education. Our search revealed over 3,500 articles, business reports, and company press releases for each sector. One researcher analyzed these documents using NVivo and formed a high-level case history for each industry, listing all the entries and activities of Big Tech firms, paying attention to the countries in which the activities took place.[15] Through this lens, we realized that state dominance in the delivery of services mattered for the entry strategies of Big Tech firms. This motivated us to choose the two country settings of the United Kingdom and the United States, where most of Big Tech firms' entry activities have happened so far. In the case of healthcare, the contrast between the United Kingdom and the United States in terms of the state involvement in offering primary services contributed to our findings. We observed that in education, Big Tech activity was further ahead in the United States compared with the United Kingdom (and the rest of the world). This uneven level of activity between the United States and United Kingdom led to an emphasis in our findings on Big Tech entry into U.S. education, where our insights arguably foreshadow the future of this industry in the United Kingdom and the rest of the countries.

Once the case histories were written, the research team started the cross-platform, cross-country, cross-industry comparisons to understand the similarities

and differences of Big Tech firms' entry and activities, as well as to develop a broader cross-industry understanding of each Big Tech firm. To compare and contrast cross-platform activities within a single industry (in healthcare and in education, separately), we started by mapping the main activities of each Big Tech firm, which were employed to disrupt or add value to the existing processes, products, or services in the specific industry. Then, we compared these activities to identify the similarities and variance across platforms. We grouped common approaches and activities across different platforms, also noting the differences, which enabled us to underline the relationship between platforms' specific capabilities and their entry decisions into highly regulated markets. Then, we moved on to cross-industry comparisons, where we accounted for the unique characteristics of education and healthcare, including customer needs and expectations, opportunities for complementarity/partnerships between Big Tech players and existing actors in the market, and the activities/products that are regulated in the value chain. Finally, we compared the United Kingdom and United States (especially with regards to healthcare) to identify the variance in Big Tech entry processes, depending on the level of regulation, the existence of central versus distributed systems, and the importance of private actors existing in the industry (i.e., hospital conglomerates in the United States). As the findings emerged, we engaged in iteration between theory and findings to arrive at the mid-range theory summarized below.

Findings

Empirical Observations on Big Tech Entry into Healthcare and Education

Table 2 provides a summary of the resources and capabilities of each Big Tech firm and their dominant mode of monetization across industries. Tables 3 and 4 (for healthcare) and 5 and 6 (for education) offer more detail into entry activities in each industry. In addition, Table 7 provides a further summary of Big Tech platform firms' entry activities, allowing comparison across entry pathways into both industries.

Big Tech in Healthcare

Google started its healthcare activities earlier than other platforms, offering services focused on collecting, standardizing, and analyzing health-related data for healthcare incumbents. Google's value proposition stemmed from its superior search and data-tracking capabilities applied to health-related queries. In 2008, Google announced plans to create a digital, customer-centric healthcare database (Google Health) for health insurers and doctors. Between 2008 and 2013, Google's focus remained on collecting, monitoring, and managing personal-level data as well as facilitating integration through interoperability across individuals and various healthcare organizations such as hospitals and pharmacies. From 2013, Google began a series of acquisitions in the sector and developed its subsidiaries' activities in technology development, product development, and services in therapeutic categories such as heart disease, diabetes, cancer, and Parkinson's.

TABLE 2. Big Tech Platforms' Resources, Capabilities, and Monetization.

		Big Tech Firms				
		Amazon	**Apple**	**Facebook**	**Google**	**Microsoft**
Pre-entry Resources and Capabilities	Existing User Base and Means of Data Capture	• Large user base in retail and cloud • Data captured relates to on-platform activity and trackers across the web. • A million employees to use as pilot or internal user base.	• Large user base in hardware • Data exclusively captured through proprietary hardware devices	• Very large user base across three dominant social media and messaging platforms • Additional data capture via widely used trackers in the web • Data captured limited to platforms and trackers.	• Very large user base • Data captured across most popular web infrastructural services (search engine, mail, advertising, geolocation) • Data captured through cloud services and mobile operating system	• Large user base in software (OS and Productivity) • Expanding user base in cloud services
	Capabilities in Data Analysis	• Moderate to strong capabilities, through: • Retail store (recommendation engines) • Services provided through cloud (text and image analysis)	Moderate capabilities, through: • Hardware devices using face recognition, fitness, and vitals data collection	Mixed capabilities • Strong capabilities in image, text, and video analysis • Weak capabilities in applied data analysis (e.g., in healthcare only through social media posts for mental health)	Strong capabilities • In multiple fields of AI • Including strong capabilities in healthcare and education applications	Strong applied capabilities, through: • Data access using existing corporate and government relationships • Growing capability in other/scientific applications of AI
	Existing Ties and Capability of Working with Institutional and State Actors	• Some ties through AWS services	• Few to moderate ties in education • Few ties in healthcare	• Negative ties for many state and institutional actors • Some legacy education ties	• Increasing ties in healthcare • Moderate ties in education	• Strong ties with institutional and state actors in healthcare and education
Main Monetization Model		• Sales of retail products • Marketplace of products • Provider of cloud services	• Sales of hardware devices and software	• Selling advertisement space and targeting data	• Selling advertisement space and targeting data, provider of cloud services	• Sales of hardware devices and software • Provider of cloud services

These activities exploited and leveraged Google's Artificial Intelligence (AI) and Machine Learning (ML) capabilities.

In 2015, Google (renamed as Alphabet) entered the medical devices sector by developing products and technologies such as needle-free blood draw devices, cloud-connected diabetics sensors, and robotic surgeons. In 2016, after acquiring U.S. API (application programming interface) management company Apigee, Google strengthened its portfolio of activities at the "data infrastructure" level. In 2018, it announced a healthcare data API to promote interoperability in fragmented healthcare providers. In 2019, Google announced various partnerships (e.g., with the Mayo Clinic) for storing patient data in Google Cloud. All these API and cloud-related activities leveraged Google's existing strengths in core infrastructural technologies. In parallel, from 2016 onward, Google formed partnerships with major pharmaceutical companies on R&D and commercialization of bioelectronic medicines. From 2018 onward, Google increased its level of

TABLE 3. Big Tech Platforms' Activities Timeline in Healthcare for Google and Microsoft.

Big Tech Activities in Healthcare: Google, Microsoft						
Year	**Commercial Partnerships**	**New Ventures**	**New Products or Services**	**Acquisitions**	**Research Partnerships**	**Service Terminations**
Google in Healthcare						
Pre-2010	1	1 (investment)	1 announced 1 launched			
2010	1		1 launched			
2011						1
2012						
2013	1	1				
2014	2		1 launched	1		
2015	1		1 announced			
2016	1		1 announced 1 launched	1		
2017			2 announced	2		
2018	1	1 (investment)	2 announced	1	2	
2019	4	3 (including 1 with Apple, investment)	2 announced 1 launched	1		
2020	1		2 announced 1 launched (Apple-Google Covid tracker)			1 (Buyer switched from Google Cloud to AWS)
Microsoft in Healthcare						
Pre-2010	1			2		
2010	2			1		
2017	2		1 announced		1	
2018	7				4	1
2019	8				3	

Note. AWS = Amazon Web Services.

investments in developing specific AI solutions for healthcare incumbents. Examples include the controversial "Project Nightingale" in the United States that has led to the creation of "Google Health Care Studio," which aimed to organize complex healthcare information so that clinicians could search, organize, and navigate patient data more easily and efficiently; the U.K. Deepmind (NHS agreements aiming to leverage patient data and AI technologies to improve diagnostics, patient care, and treatment); and the latest deal in May 2021 between Google and

TABLE 4. Big Tech Platforms' Activities Timeline in Healthcare for Apple, Facebook, and Amazon.

Big Tech Activities in Healthcare: Apple, Facebook, Amazon						
Year	**Commercial Partnerships**	**New Ventures**	**New Products or Services**	**Acquisitions**	**Research Partnerships**	**Service Terminations**
Apple in Healthcare						
2014			1 launched			
2015			1 launched			
2016			1 launched			
2017	2			1	1	
2018	2		3 announced 1 launched			
2019	1	1 (investment)	1 launched	1	1	
2020			3 launched (all Covid-related, including 1 in partnership with Google)		1	
Facebook in Healthcare						
2017	1		1 announced 1 launched		1	1 (after Cambridge Analytica)
2018			1 launched		1	
2019			1 announced		1	
2020			1 launched (Covid-related)		1	
Amazon in Healthcare						
2013			1 launched			
2017			1 launched			
2018			1 announced 1 launched		1	
2019	2		3 launched	1	1	
2020	1					

HCA Healthcare Inc. in the United States, which aims to leverage patient data in developing healthcare algorithms for operational efficiency, better patient monitoring, and doctor decision making.

In summary, over the 2008-2020 period, Google's role evolved from that of a peripheral IT service provider to healthcare incumbents to that of an

TABLE 5. Big Tech Platforms' Activities Timeline in Education for Google, and Apple.

Big Tech Activities in Education: Google, Apple						
Year	**Commercial Partnerships**	**New Ventures**	**New Products or Services**	**Acquisitions**	**Research Partnerships**	**Service Terminations**
Google in Education						
Pre-2010			1			
2018	1					
2019			1 launched			
2020			1 announced (Google Career Certificates)			
Apple in Education						
Pre-2010				1	1 (free computers initiative)	1
2014			3 launched		1 (Tech supply donation initiative in underserved schools)	
2015			1 launched			
2016			1 launched ("Apple Development Academy")			
2017	1		2 launched			

increasingly present and central actor in the industry. This entry pathway led to Google not only entering several of the sector's niches, but also becoming an essential partner to infrastructural projects for government agencies and state-controlled institutions, dominating the industry for diagnostics, electronic health records, enhancement of current devices and treatments, and development of new devices and treatments in healthcare.

Microsoft has been a long-time player in the healthcare system as an IT provider, with established long-term partnerships with several actors in or adjacent to the industry. This presence allowed the firm to collect data on a large scale and scope. To further this, Microsoft formed partnerships in 2011 to launch solutions for interconnectivity among hospitals, physicians, and patients. As digitalization started to take hold in many industries, Microsoft stepped in to support digital transformation of healthcare organizations. In 2017, Microsoft announced plans to expand its AI and cloud services for the healthcare sector, leveraging the

TABLE 6. Big Tech Platforms' Activities Timeline in Education for Facebook, Amazon, and Microsoft.

Big Tech Activities in Education: Facebook, Amazon, Microsoft						
Year	**Commercial Partnerships**	**New Ventures**	**New Products or Services**	**Acquisitions**	**Research Partnerships**	**Service Terminations**
Facebook in Education						
Pre-2010			2 launched ("school and workplace networks")			
2013				1		
2019			1 announced 1 launched			
Amazon in Education						
2012			1 launched	1		
2013						
2018						1
2020	1		1 launched ("Amazon Academy")			
Microsoft in Education						
2016	1					
2017	2					
2018	1		1 announced		1	
2019	1		1 launched			

capabilities developed in cloud services (Azure) and then AI/ML technologies. In 2018, Microsoft closed the HealthVault Insights app, showing their focus on B2B rather than B2C. In addition, Microsoft formed partnerships that focused on diagnostics of rare diseases. Microsoft has the same approach as Google regarding entry activities about AI and data strategy. One advantage it has is that it is already present in hospitals and healthcare organizations as a software/IT provider or consultant. This means it can use its established long-term relationships and develop customized solutions for these state and institutional actors. Compared with Google, Microsoft's approach has been limited to collaboration with healthcare organizations rather than starting new ventures. In summary, over the 2008-2020 period, Microsoft's role evolved from being an IT service provider to becoming an increasingly essential partner to many healthcare incumbents.

Apple maintains a dominance in the mobile and wearable devices market and leverages these devices to collect individual-level data for diagnosis. Apple's

TABLE 7. Big Tech Platforms' Points of Entry and Additional Value-Adding Activities Upon Entry.

		Additional Value-Adding Activities Upon Entry	
	Point of Entry	**Design of New Products and Services**	**Leveraging Platform Business Model**
Google	• **Healthcare:** Data collection and capture through Google Health (early on), partnerships, user search data, as well as cloud services • **Education:** Data collection and capture through Chromebook, Google productivity apps, user search data, and cloud services	• **Healthcare:** DeepMind, Google Health, Medical devices via Verily, drug R&D collaborations with pharmaceutical companies • **Education:** Chromebook, Google Docs and Apps for Education, Google Careers Certificate	• **Healthcare:** Google Health; Fitbit; investments in industry-specific services (e.g., Oscar Health); Covid tracking system • **Education:** Google Docs and Apps for Education; Chromebook
Microsoft	• **Healthcare:** Data collection and capture through partnerships with institutional and state actors, user search data, and cloud services • **Education:** Data collection and capture through partnerships with institutional and state actors, and only later through hardware, software, user search data, and cloud services	• **Healthcare:** Partnerships with biotechnology firms • **Education:** Surface Hardware, Software and apps for operating system and productivity; cooperation with policy makers in designing products and services for the digital transformation of the education (Microsoft K12 Education Transformation Framework)	• **Healthcare:** Microsoft Genomics • **Education:** LinkedIn Learning; Microsoft Educator Center
Apple	• **Healthcare**: Data collection through hardware • **Education:** Data collection through hardware	• **Healthcare:** Collaborative R&D with research organizations (i.e., Biotech, universities) • **Education:** Hardware (esp. iPad) and software/apps; coding courses	• **Healthcare:** Apple Health; ResearchKit; Covid tracking system • **Education:** iTunesU (obsolete)
Facebook	• **Healthcare:** Data collection through users on social platform • **Education:** Data collection through users on social platform and through early partnerships with educational institutions and learning management software integration	• **Healthcare:** Research partnerships with non-profits (i.e., universities), experimental medical devices (e.g., a device enabling typing by thinking), Preventive Health • **Education:** Partnerships to provide "datafied" education program, and startup courses	• **Healthcare:** Using core Facebook platform for self-assessment tests, location-based suggestions for hospital • **Education:** Integration with the Blackboard Learning Platform (mostly obsolete)
Amazon	• **Healthcare:** Retail platform for product sales, cloud services, hardware (e.g., Alexa) • **Education:** Retail platform for product sales (e.g., textbook), cloud services, and hardware	• **Education:** Amazon Academy; Amazon Online Training Program and Amazon Career Choice (currently Internal)	• **Healthcare:** PillPack; Haven; existing marketplace for incumbent consumer medical device firms; Alexa; AmazonCare. • **Education:** Amazon Inspire Joint ventures (i.e., Haven)

most important resource is its proprietary mobile hardware offerings (e.g., iPhone, iPad), while its capability lies in data analysis through functionalities that are tightly integrated into its hardware (e.g., motion sensors). Apple's health-related endeavors became significant in 2014 when it launched its new health and fitness data app. In 2015, it launched ResearchKit, an open-source platform that lets researchers recruit participants for studies, and further pushed data collection via devices through acquisitions. Apple used ResearchKit to facilitate in-app engagement across parties and formed research-intensive partnerships to exchange and utilize these data. In 2018, Apple launched its "Health Records API" for developers and researchers to stimulate the growth of an app ecosystem utilizing health data. In 2020, it teamed up with Google for the development of a coronavirus-tracking technology for their respective mobile operating systems.

Overall, Apple's data analysis capabilities are more limited than Google's and more focused on functionalities directly supported by Apple hardware (such

as predicting a heart attack of an Apple device user). However, in line with its monetization model based on hardware (and software) sales, the firm exercises high control over users' data, touting its respect for users' privacy, which can help reduce state or institutional actors' privacy concerns. In summary, Apple's pathway to entry in healthcare is more narrowly focused compared with Google and Microsoft, aiming at developing an Apple-hardware-compatible app ecosystem with technologies that have a platform connector or infrastructural quality (particularly APIs and SDKs). Together with its installed base of devices and underlying platforms (iOS and App Store), Apple can capture a vast amount of health data which becomes an increasingly valuable asset.

Facebook has a limited presence in healthcare compared with other Big Tech firms. As of 2021, it has only one major activity in health diagnosis with its "Preventive Health" program, which uses social media data to suggest preventive tests and nearby health providers. Facebook's healthcare activities focus on combining healthcare data with social media data, using social media channels for healthcare-related communication, or utilizing its AI capabilities for healthcare. Facebook's main asset in these activities is its three main social platforms (WhatsApp, Instagram, and Facebook) and its reach through web trackers, which is almost as wide as Google's. Also, Facebook possesses advanced individualization and profiling capabilities via AI/ML. In 2018, Facebook started to deploy these to prevent suicides by scanning early signs in posts, partially to avoid suicides being streamed live on its platforms. In 2020, it launched a coronavirus information center to support staying at home. Overall, while Facebook's entry into healthcare is more limited, its data on individual users give the platform the potential to profile and track users over time to predict such things as mental health issues, diabetes, and pregnancy. The degree to which Facebook can use or share this data is currently a hotly contested debate in various regions of the world, particularly after the Cambridge Analytica scandal.

Amazon has a unique set of entry activities that are not employed by other platforms. It started with selling health products on its retail platform. After its 2018 acquisition of PillPack, Amazon began to offer prescription drugs on its platform in the United States, leveraging its unique physical assets (compared with other Big Tech firms) such as warehouses and logistics businesses. In parallel, AWS becoming a prominent infrastructural element in many industries, including healthcare, gave the platform a natural precedent for developing industry-specific products and services. AWS also helped Amazon build connections with institutional and state actors in the United States, where government agencies relied on its services to comply with security and secrecy needs. Amazon made several acquisitions to strengthen data interoperability across different actors, for example, Health Navigator, a venture with a popular API for online health services. Then, in 2018, Amazon made a partnership with Berkshire Hathaway and JPMorgan Chase by forming *Haven* to provide low-cost and high-quality healthcare for these firms' employees (Haven was shut down in 2021). Meanwhile, Amazon also leveraged Alexa for data collection and AI applications. In 2019, it formed a partnership with NHS UK to reduce the workload of NHS workers and to enable

users to access health advice via Alexa. A few months after this deal, news that the NHS had provided Amazon with free access to healthcare information (e.g., symptoms) caused a backlash in the press.[16] In 2019, Amazon launched the Amazon Care app for its employees, which was seen as a move to disrupt healthcare with a consumer-focused strategy, like in other industries.

In summary, Amazon's involvement in healthcare is increasing via multiple avenues in parallel. Its B2C online retail reach allows Amazon to distribute prescription drugs. With its infrastructural technology, AWS provides a critical entry point into most businesses, gaining access to enormous sources of data and monitoring businesses activities. Amazon is also involved in platform connector technologies such as APIs, providing another data capture point. Thanks to Alexa, Amazon can capture lifestyle data from individual users directly or via API-connected apps. Finally, its partnerships providing voice assistance and AI to government agencies and state-controlled institutions provide it with deep access to sensitive data. All in all, Amazon evolved from an infrastructure technology provider to an increasingly important player in the industry, especially considering the demand side of its prescription drug business.

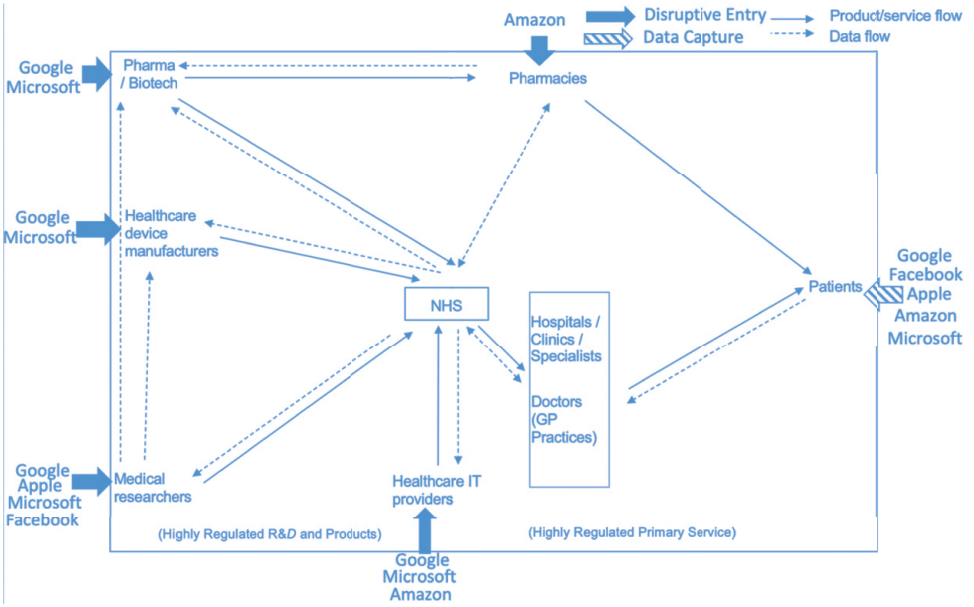
The entry activities of Big Tech platform firms into healthcare are illustrated in Figures 2 and 3.

Big Tech in Education

Our data show that Google has a broad influence and reach in education: it provides hardware and software to K-12 schools, starting with Google Education apps in 2010, then with Chromebook in 2011 (which is loaded with free educational apps and now constitutes most of the hardware shipped to schools), and finally integrating all the software with Google Classroom from 2014 onward. Google leveraged its monetization model (i.e., to charge advertisers while subsidizing users for software and hardware) to get access to data from more than 30 million U.S. children (who are also future customers) and build a data-driven business relationship with schools, using its AI/ML capabilities for tracking and predicting learning. Its collaboration with universities is similarly centered on digitalizing educational resources (books, documents) and providing digital tools to support education. In 2020, Google began to directly enter higher education as a primary service provider with "Google Career Certificates," a series of online courses for adults. It rallied a group of over 50 large employers (e.g., Walmart, Intel, and Bank of America), who claimed they would recognize these degrees and connect to certificate holders directly upon completion of their program. Overall, Google has entered the education industry as an IT and hardware provider and then moved into more central roles such as content provider and potential educator, preparing for a "skills" based future in which degree institutions are not the only primary service providers in the industry.

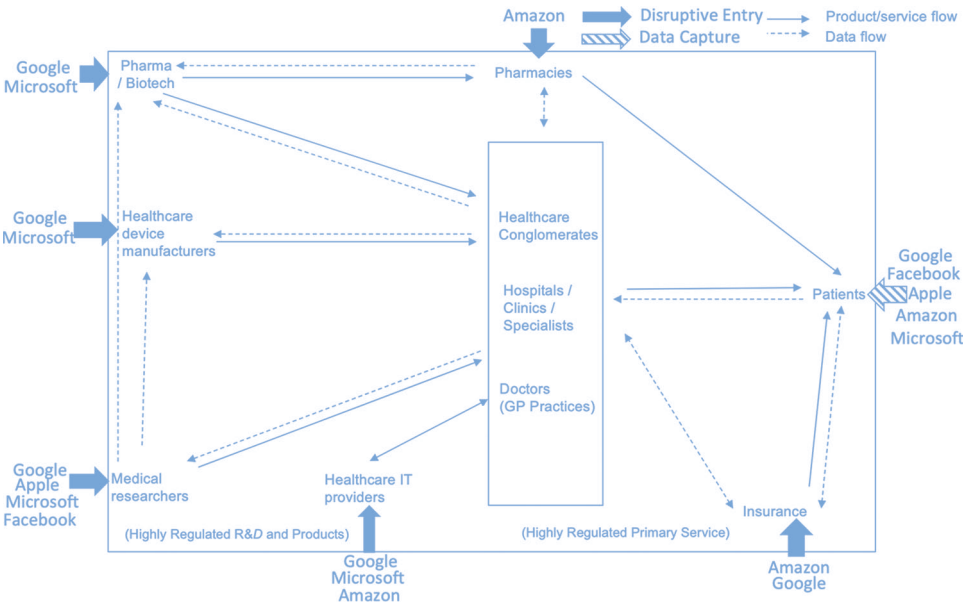
Microsoft's activities in education are slightly more diverse than other Big Tech firms. It offers more industry-specific services in education than other platforms, building on its (legacy) role as an operating system, infrastructure, and technology provider. First, through LinkedIn Learning (2017), it provides an online

FIGURE 2. U.K. healthcare industry structure and Big Tech entry activities.



Note: NHS = National Health Service, GP = General Practitioner.

FIGURE 3. U.S. healthcare industry structure and Big Tech entry activities.



Note: GP = General Practitioner.

learning platform (both in B2C and B2B segments, where clients include businesses, higher education institutions, and governments). Going one step further, Microsoft

started to cooperate directly with policymakers with the aim to "digitally transform education" (Microsoft K12 Education Transformation Framework, 2019), which leverages Microsoft's existing capabilities and ties with state and institutional actors. Microsoft also provides software through Microsoft Educator Center, which is an online tool that includes various software programs and applications such as Windows, Skype, and OneNote, as well as hardware such as Surface Tablets for teachers and students. Overall, Microsoft leverages its legacy ties with institutional and state actors to further strengthen its activities and expand them across the more digitalized education industry and beyond, moving into professional education. Its strategy in education seems similar to Google's, but it is lagging behind.

Apple was the first mover in disrupting education. It started by providing hardware (tablets, laptops) and software (on its own devices) to schools as early as 1978, and since then it made a "re-entry" with its iPad provision in 2010. Apple's entry activities into education included the acquisition of industry-specific IT providers (e.g., PowerSchool in 2001). Apple then moved on to offering courses in coding and software programming in 2015, leveraging its hardware and apps. In 2018, it announced plans to partner with educators and provide training for blind and deaf communities with accessible coding. Overall, despite an early entry, Apple appears to have lost its strong position in this industry to Google, which combined a supply of cheap hardware for data capture with a broader, hardware-agnostic supply of software to provide solutions for students and institutions.

Facebook leverages its own social platforms (Facebook, Instagram, and WhatsApp), as well as its subsidiaries, such as Oculus VR, for providing interactive teaching and communication among students. In 2005, Facebook expanded its network by targeting users/students in 800 colleges with these services. Since then, it has added high school and international school networks to grow its platform. In 2015, it started a partnership with California Public Schools to jointly develop a "datafied" education program for K-12 schools. In 2019, it announced plans to launch "Playground" in Israel to offer courses for startups and businesses. Facebook's potential to leverage all its platforms together is limited in some geographies, such as the EU (Germany, in particular), which puts constraints on how the data across its platforms can be combined. Over time, we observe Facebook's position in education evolving from mainly leveraging its social network to offering professional learning services and learning-related hardware (e.g., Oculus).

Amazon's activities in education started with online training programs for its own employees (e.g., Career Choice, 2012), with plans to roll this initiative out commercially in the future. The firm also recently launched Amazon Inspire, an open collaboration platform for K-12 teachers to share educational resources (currently in beta stage). Amazon also made significant business partnerships with U.S. universities, offering them IT infrastructure as well as AI-driven services. The partner universities pre-installed university accommodations with Alexa devices to provide student support to eventually replace some functions within the university administration and replace university databases (with critics worried about Alexa monitoring students against regulations). While outside our

geographical scope, it is interesting to note that Amazon has a highly successful test preparation app "Amazon Academy" in India, which offers mock tests, practice questions, and solutions to previous years' test papers to students. Overall, Amazon seems to be focused on higher and professional education, both to continuously reskill its own enormous human resource base as well as to commercialize the activity beyond its firm borders.

The entry activities of Big Tech firms in education are illustrated in Figure 4.

Synthesis

Figure 5 presents a synthesis of our findings in a process model of digital platform firms' entry in highly regulated industries and then analyzes their additional value-adding activities.

A clear pattern that emerges from our case histories is that Big Tech firms haven't so far been involved in offering primary services (e.g., providing healthcare, banking, or schooling).[17] Rather, they have focused on capturing data as a pathway to other value-adding activities. Capturing data from the highly regulated industry, combining it with data from various other industries, and analyzing it through co-specialized AI/ML technologies[18] give Big Tech firms a competitive advantage in the newly entered highly regulated industry. It allows them to generate data-driven insights and to design and commercialize new products and services for the industry, generating value above and beyond what incumbents can offer.

Big Tech Firms' Data Capture Activities in the Newly Entered Highly Regulated Industry

To capture data, Big Tech firms engage in two types of activities simultaneously: using their own hardware or software to build their own data sets (i.e., *direct data capture*); and/or forming partnerships with state or private actors (particularly primary service providers) in the industry for access to existing data (i.e., *indirect data capture*).

Direct data capture. Direct data capture represents these platforms' capture of user-level proprietary data, including those directly captured by the Big Tech firms in collaboration with institutional actors. For Big Tech firms, this activity augments or enriches the data obtained from the industry and also serves as an alternative path to data capture if attempts to obtain access to existing, often sensitive, industry data held by primary service providers fail. For instance, Amazon generates an alternate database for drug treatments by acquiring pharmacies and combining their database with its cloud service to standardize electronic records. While these alternate databases cannot replace the patient data residing in the healthcare system in terms of value, they can complement it, giving Big Tech firms an opportunity to enter partnerships with pharmaceutical companies, research institutions, and healthcare providers. For instance, fitness or mental

FIGURE 4. U.S. education industry structure and Big Tech entry activities.

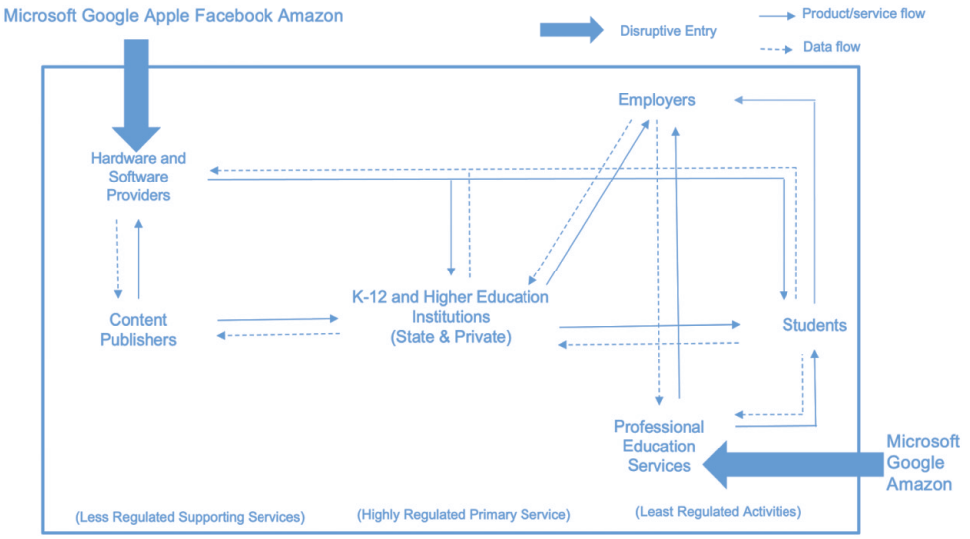
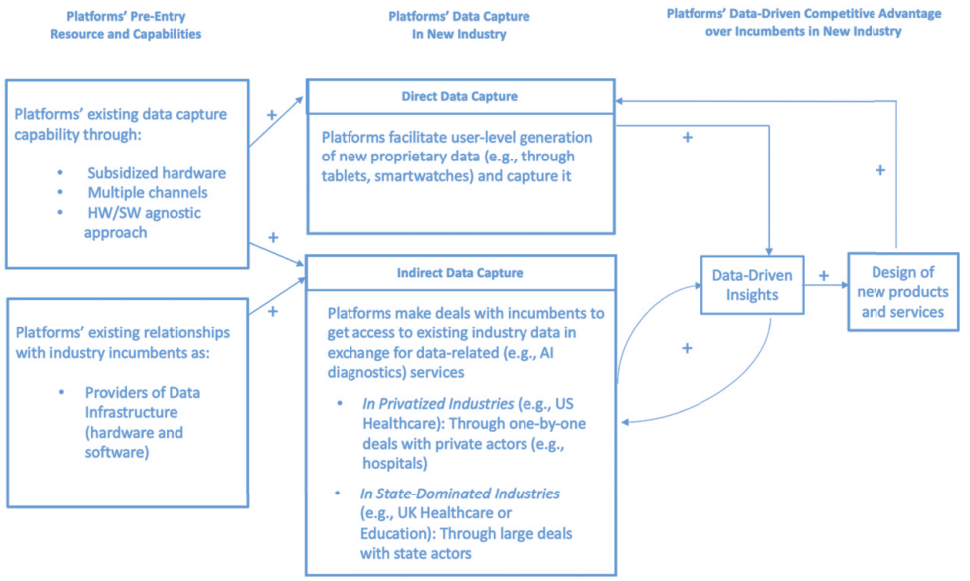


FIGURE 5. A process model of Big Tech entry into highly regulated industries.



Note: AI = Artificial Intelligence, HW/SW = Hardware/Software.

health information is typically lacking in healthcare systems, but such real-time data can have high value for health services and medical research.

In a race to capture the most data in a new industry, a question arises: What drives Big Tech firms' success? First, we observe that those with already strong data

capture capabilities—using subsidized hardware, often with other multiple channels (apps, websites)—can benefit from a head-start in entry. This is particularly relevant when the highly regulated industry has limited or disparate individual-level data. In K-12 schools, for instance, student data historically consisted of exam results, with most other data (e.g., class participation, homework) remaining with teachers. Google's overtaking of Apple in K-12 classrooms thus can at least partially be explained through their willingness to subsidize hardware (Chromebook) together with their software (Classroom) to widen data capture, which was attractive to education providers in improving learning in a cost-effective manner. Similarly, in healthcare, Google and others' partnerships and investments with medical hardware developers can be partially explained by their intent to deploy hardware to capture user data. We therefore posit that the subsidized hardware investments of platforms (used in tandem with other channels) strengthen the direct data capture of platforms entering the highly regulated industry.

We also observe that platforms that have already built superior data sets and capabilities in data capture without depending on particular hardware or software have a competitive advantage (over competitors whose data capture is primarily dependent on particular hardware or software). These firms can flexibly leverage whichever channel works best for direct data capture, which eventually leads to superior data-driven insights. In this regard, Google has the comparative advantage that it has a large user base through which it captures direct data from the casual? uses of digital technologies. On the other hand, Apple and Facebook are most disadvantaged in this aspect due to their high hardware (Apple) and software/app (Facebook) dependency.

Indirect data capture. A second type of activity through which Big Tech firms capture data in a highly regulated industry is through partnerships with primary service providers to get data access. We find that to get such access, *providing data infrastructure services* is a common complementary value-added activity that they offer to primary service providers. Big Tech firms can capture value through these services by charging for these services, receiving data that can be monetized outside the industry (e.g., more targeted advertising services), and monetizing data inside the industry by designing new products or services.

The nature of Big Tech firms' indirect data capture within the highly regulated industry depends on the extent to which the industry is privatized. In healthcare, countries greatly vary in how privatized their healthcare services are. In countries such as the United States, where most healthcare services are offered by private actors, data exist inside healthcare providers, who are largely disconnected and may keep data in different formats. As the data in each of these providers are a portion of the entire data in the system, having a connector build a universal database and then provide data analysis (e.g., AI diagnostics) is a complementary activity to primary service providers with high added value. However, this fragmented system constitutes a coordination challenge for platforms in that they need to convince these providers one-by-one and harmonize (i.e., combine in an interoperable way) the dispersed data sets. This then becomes a winner-takes-most

landscape, with the player convincing enough service providers gaining momentum in data access and being in the best position to provide AI services.

A state-dominated industry (e.g., U.K. healthcare, U.K. and U.S. education), on the other hand, has the feature that data are typically centrally kept.[19] However, data capture may still be at a low level as the state may collect limited data (e.g., lack of hardware/software) or it may be difficult to implement statewide initiatives for data collection (e.g., in education). In addition, state-owned databases may be harder to access due to ethical, technical, and security reasons. This makes access to data an essential element of a tough winner-take-all game.

Considering indirect data capture under this circumstance, we observe that Big Tech firms' existing relationships play an important role in successfully making deals to get access to existing industry data in exchange for feeding back data-driven insights that can add high value to incumbents' services. Three players stand out: First, Microsoft has a long history of working with state and private actors across many industries, including education and healthcare, and can leverage its cloud services and customized solutions for institutional actors. Second, Amazon is the leading cloud provider, with strong ties to the U.S. state actors. Finally, Google uses cloud services to expand its indirect data capture by making deals with incumbents (e.g., the NHS-Google deal).

In addition, platforms' existing data capture capabilities strengthen their ability to make deals with incumbents to get access to existing industry data. Again, platforms that use subsidized hardware, leverage multiple channels, and can capture data without dependence on particular hardware or software are better positioned to make deals with industry incumbents due to their superior capability to offer data-driven insights.

Overall, we observe that platform firms attempt to make deals with service providers, offering data infrastructure and analysis services in exchange for data access. This entry path is typically accompanied by an effort to build additional data sets through direct data capture, which becomes particularly crucial when the platform is unsuccessful in striking a deal with incumbents.

Big Tech Firms' Data-Driven Competitive Advantage over Incumbents

We observe that beyond capturing data both directly and indirectly, Big Tech firms engage in two types of activities that add and capture value in a highly regulated industry: they generate data-driven insights that improve existing products and services in the industry; and/or participate in the design of new products or services for the industry.

Data-driven insights. Data-driven insights, which are insights that rely on the data analysis capabilities of the platforms leveraging AI/ML, are the "powerhouse" driving Big Tech entry into highly regulated industries. A data analysis service that is particularly noteworthy in these regulated sectors is *AI-based diagnostics*, that is, early diagnosis of diseases or responses to treatment. These services can significantly cut costs and improve results for healthcare and educational service

providers.[20] These data-driven insights benefit Big Tech firms in multitude ways: initiating or advancing indirect data capture by offering these insights to primary service providers, thus entering a positive feedback loop; designing and commercializing new products and services that are superior to incumbents' alternatives; and utilizing the insights in their existing platforms, thus strengthening their dominant position.

New products and services. Once they collect data, Big Tech firms also start to add value to the value chain in healthcare and education by leveraging their data-driven insights to participate in the innovation process for new products and services that complement primary services. As with the examples of Facebook partnering with California Public Schools to develop a K-12 education program or Google partnering with GlaxoSmith Kline to make and sell bioelectronic medicines, Big Tech firms use data analysis to cut costs and speed up the R&D process for new products and services. Furthermore, they can target lucrative areas where their own ventures can develop more tailored products and services, such as in the case of Amazon acquiring TenMarks to design web-based math curricula or Google's Verily developing wearable non-invasive devices to track blood sugar levels. These new products and services that are based on cloud and/or AI capabilities contribute to a change in the core products or services in the industry, moving Big Tech platforms toward the core of value creation in the industry.

When developing and commercializing these new products and services, Big Tech firms also benefit from their existing user base, data, and distribution channels to efficiently deliver the new products and services. This is an additional reason why these new products and services become superior to those offered by the incumbents. For example, in the case of Google Health Care Studio, Google was able to leverage its search data, data infrastructure/API, and data analysis capabilities to overlay on top of the existing electronic health records and provide real-time information across complex patient records,[21] which is out of the reach of incumbents. In addition, Big Tech platforms can deal with high-cost regulatory licenses and other requirements much better compared with many other potential entrants, due to their vast user base and financial resources. For example, following the implementation of GDPR, Google increased its market share in web technology services (e.g., online advertising)—and consequently, the amount of data collected—in favor of smaller firms since Google was able to gather user consent at a mass scale.[22]

The value that is captured by Big Tech firms in new product and service design is threefold. First, in the case of a partnership with another party (e.g., pharma), value capture focuses on charging for providing data for R&D purposes and/or capturing the newly created data in the partnership for fueling AI inside and outside the industry. Second, various kinds of new data (e.g., from transaction or hardware) are captured and can be monetized through data analysis inside and outside the industry. Third, in the case of a direct investment into a new venture, the platform receives direct profits from sales of the products and services.

While discussing how Big Tech firms add and capture value in highly regulated industries, it is important to note the synergies between different types of activities. In our description of activities above, we differentiated between provision of data infrastructure services, generation of data-driven insights, and design and commercialization of new products and services as primary ways for Big Tech to capture value. In that context, we observe strong complementarities between these activities, that is, a strong presence in one activity can give a firm advantage in growing its presence in the others. For example, Google and Microsoft, which have a strong presence in the infrastructure provision and data-driven insights, are particularly active in new product and service R&D through spin-offs and partnerships, as listed in Tables 3, 4, 5, and 6.

Overall, in highly regulated industries where access to primary service data (e.g., individual-level clinical data or educational records) is a bottleneck due to the sensitivity of this data, those firms that overcome this bottleneck, generally through indirect data capture combined with data-driven insights,[23] gain a unique competitive advantage in the design of new products and services where incumbents (e.g., medical device manufacturers, textbook publishers) have historically operated without such data capture or analysis capabilities. This allows Big Tech firms to "digitally colonize" a highly regulated industry without providing the primary services (e.g., education, healthcare) that are highly unprofitable by providing unique added value through data-driven products and services that other industry players increasingly rely on, which in turn commoditizes these players over time.

Contributions to Theory and Practice

In this study, we analyzed how Big Tech firms enter highly regulated industries that are characterized by high barriers to entry and typically high state involvement. After expanding rapidly across sectors in the last decade, Big Tech firms, who currently enjoy a combined market capitalization of almost $7.4 trillion, have begun to target highly regulated industries with "high-worth" data using their data analysis capabilities. As the ongoing exogenous trend of global digitalization (accelerated during the COVID-19 pandemic) renders the digital provision of services more essential, the data-fueled, foundational, and infrastructural services Big Tech firms offer have become increasingly more central to the new functioning of these highly regulated industries.

Our findings reveal the patterns of Big Tech entry into education and healthcare, two highly regulated industries that compose a significant part of the global economic activity. We highlight that the crux of platform entry into highly regulated markets is access to sensitive data. Successful platform entry occurs via partnerships with incumbent firms and institutions by first providing data infrastructure and, later, data analysis services. Once platforms establish access to sensitive data in exchange for providing incumbents with superior data analysis (e.g., saved lives, better learning outcomes), they pivot into offering products and services that increasingly rely on platforms' data-driven capabilities. We label this

pattern of entry as the "digital colonization" of highly regulated industries, where the primary service providers in the industry do not change, but they increasingly rely on data-driven products and services provided by digital platforms. This pattern of entry is a different mode of platform entry that is distinct from the well-known "platform envelopment,"[24] which describes how platform firms overtake competing platforms in related or unrelated markets through bundling of features and sharing of user bases across activities.

Our findings are also generalizable to other, less-regulated industries. While data access can be more challenging in highly regulated industries, and replacing primary service providers can be unprofitable, the way that data capture opens doors to digital colonization of the industry is a plausible pattern of disruption for many industries (e.g., banking, energy, insurance) where incumbents have lagged behind in leveraging data they possess for service provision and product innovation.

A critical point to consider in the digital colonization of highly regulated industries is the balance between the increase in value creation, which will serve the whole industry, and the rise in value capture by Big Tech firms. Our study illustrates the two opposing forces quite vividly. On one hand, Big Tech firms that can add more value to existing products and services through better data capture, data analysis (and therefore data-driven insights), and data infrastructure services are more likely to get access to data collected by incumbents. On the other hand, these Big Tech firms, once they get access to data, can develop a unique competitive advantage over incumbents and become increasingly powerful within the highly regulated industry as well as in other industries (e.g., advertising, retail) where the obtained data can be leveraged.

Our study has important managerial implications. For platform strategy, we suggest that superior data analytics capabilities that enable the generation of data-driven insights matter even more in highly regulated industries, as they offer a pathway to break the high entry barriers in these industries and solve the bottleneck of data access. We also find that, rather than replacing existing actors and competing head-on to offer primary services, successful platforms add value to the existing value chain in highly regulated industries through data infrastructure services, data-driven insights, and finally new products and services. This suggests that for firms with a platform-based business model, entry into highly regulated, data-intensive industries, where the cost of capturing value decreases as the accumulation and generation of data increases,[25] requires a multi-stage strategy. First, platform managers need to negotiate access to primary service data. At the same time, they need to make efforts to capture novel data through proprietary hardware and software user interfaces. In this context, platform firms may find that subsidizing hardware and access to services can work effectively to "buy their way" into data access. After this initial stage, they can focus on capturing value through a variety of data-related industry activities, such as selling data-driven insights or designing new products and services.

Our findings also highlight the paramount importance of a platform's policies and special procedures in dealing with the usually sensitive data within highly regulated industries. While taking precautions in the use of sensitive data may seem limiting for value creation (and capture) for a platform, this is necessary as eventually, a platform's value for its users is also driven by how well it balances its diverse stakeholders' interests, especially in terms of data privacy and security, but also in terms of the explainability and fairness of its AI-driven activities.[26] This is an increasingly important element of the balance between value creation and value capture from the perspective of Big Tech firms.

Our study also holds managerial implications for incumbent firms who need to respond to Big Tech firms' uniquely powerful form of competition. First, we show that entry of established platforms is not impossible but takes a different form in highly regulated industries. We clarify the possible pathways of entry and explain how these depend on Big Tech firms' data-related capabilities. Our findings suggest that incumbent firm managers need to formulate their own data capture and analysis strategies and decide quickly whether they will compete or partner with entering platforms.

Furthermore, if partnering with an entrant Big Tech firm seems a better strategy than competing, our findings also give managers tools to distinguish between potential platform partners based on their pre-entry resources and capabilities. We identify, for instance, the importance of hardware/software-independent data capture strategy as a basis for competitive advantage in a regulated industry. Eventually, incumbent firms need to consider the long-term implications of initially value-creating actions of Big Tech firms, such as providing data-driven insights, and whether these will, over time, lead to the commoditization of incumbents' own activities.

Our findings also provide precautions for new, non-platform entrants into highly regulated industries. First, they highlight the role of data access in gaining competitive advantage in the newly entered industry. For many new entrants (e.g., biotech or education startups), partnerships with large incumbents and/or Big Tech platforms will be essential in gaining this access. In addition, our findings indicate that the locus of competition shifting toward data-driven products and services in these industries will mean a change in entrepreneurial activity in these industries, most likely organized within ecosystems around Big Tech platforms. Understanding this new competitive landscape will be important for the entrepreneurs of tomorrow to develop their growth strategies.

Finally, understanding platform entry patterns into regulated markets is important for policymakers and regulators, who want to ensure that digital markets remain competitive while protecting the privacy of consumers. Dominant platforms such as Big Tech have been accused of taking undue advantage of their ability to harvest huge amounts of data from users across various industries and regions.[27] While arguing that most of their activities are in the best interest of their users, digital platforms often provide access to services and activities that rely on users giving up their valuable data. As a result, a recent focus of

regulatory activity in the United States, Europe, and Asia has been to prevent antitrust violations of Big Tech firms.[28] The concentration of private platform power has generated concerns that are even more salient for data-sensitive industries: in such sectors, violations of users' privacy and mishandling of users' data can have important consequences for human rights and civil liberties, in addition to possibly limiting access to essential services such as healthcare or education.[29] The findings of this study suggest that Big Tech entry into highly regulated industries is driven by capturing data and generating data-driven insights and activities that provide value to the industry. However, by eschewing direct involvement in primary activities that are highly regulated, Big Tech firms can become core actors in these industries while escaping conventional sectoral regulation. Table 1 shows that Big Tech firms are only affected by a handful of data-related regulations in healthcare and education (which are generally outdated in handling concerns related to Big Tech firms).

Considering again the balance between value creation and capture, we highlight an important dilemma for policymakers and regulators: platforms with more advanced data capture and analysis capabilities can provide more significant added value through better prediction and therefore lower costs, saved lives, and more effective education systems. But these are platforms that pose additional privacy and security concerns in a highly regulated industry with sensitive personal data. On one hand, policymakers and regulators need to serve the public through superior technologies and a much-needed efficiency in healthcare and education. On the other hand, relying on Big Tech will make it difficult for these actors to protect the personal and social rights of their citizens from the data-related practices of these platforms. As a result, our recommendation to policymakers and regulators is that special consideration should be given to regulating access to data and the usage of technology providers in highly regulated industries. This is also in line with recent work that suggests platform regulation should consider issues like "sharing (in situ) platform data" and "data mobility/portability" above and beyond a purely market power-based approach that is utilized in the utilities sector.[30] A particular balance needs to be established between giving enough space for platforms to bring their data-driven innovations to various industries and setting clear guidelines as to what data they can access, use, and combine, and what additional responsibilities they must carry while operating in highly regulated industries.

In conclusion, this study enhances our knowledge of firm strategy and platform growth by extending the traditional focus of platform research beyond low-regulation industries such as retail or entertainment to highly regulated industries. The phenomenon we focus on is an important one, as highly regulated industries play a vital role in the economy, and the entry of powerful platforms bears high value but also high risks. Highly regulated industries are also notoriously inefficient, with many actors in these sectors expressing a dire need for efficiencies and insights brought by data analytics. But the logic of efficiency brought about by digital platform firms can clash with the logic of public service and the necessity of protecting user data and privacy as a fundamental right.

Finding ways to combine the benefits brought by digital platforms with respectful consideration of personal data will become an increasingly important problem to solve in the years to come.

Authors Note

Hakan Ozalp is also affiliated with KIN Center for Digital Innoation, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.

Author Biographies

Hakan Ozalp is an Assistant Professor in Strategy at the Amsterdam Business School, University of Amsterdam (email: h.ozalp@uva.nl).

Pinar Ozcan is Professor of Entrepreneurship and Innovation at Saïd Business School, Oxford University, and the academic director of the Oxford Future of Finance and Technology (FinTech) Initiative at Oxford (email: pinar.ozcan@sbs.ox.ac.uk).

Dize Dinckol is a Postdoctoral Research Fellow at Oxford Centre for Corporate Reputation, Saïd Business School, Oxford University (email: dize.dinckol@sbs.ox.ac.uk).

Markos Zachariadis is the Chair of Financial Technology (FinTech) and is Full Professor of Information Systems at Alliance Manchester Business School (AMBS) at the University of Manchester (email: markos.zachariadis@manchester.ac.uk).

Annabelle Gawer is Chaired Professor in Digital Economy at the University of Surrey and is the Director of the Surrey Centre of Digital Economy (CoDE) (email: a.gawer@surrey.ac.uk).

Notes

1. M. W. van Alstyne, G. G. Parker, and S. P. Choudary, (2016). "Pipelines, Platforms, and the New Rules of Strategy," *Harvard Business Review*, 94(4): 54-62; A. Hagiu and E. J. Altman, "Finding the Platform in Your Product," *Harvard Business Review*, 95/4 (July/August 2017): 94-100; M. A. Cusumano, A. Gawer, and D. B. Yoffie, *The Business of Platforms: Strategy in the Age of Digital Competition, Innovation, and Power* (New York, NY: HarperBusiness, 2019).
2. "TMT News & Research: 2020 Recap," January 11, 2021, https://www.spglobal.com/marketintelligence/en/news-insights/blog/tmt-news-research-2020-recap; "Big Tech's Domination of Business Reaches New Heights," https://www.nytimes.com/2020/08/19/technology/big-tech-business-domination.html
3. A complementor is defined as a company that sells a product or service that complements the products or services of another company. https://www.merriam-webster.com/dictionary/complementor. G. G. Parker, M. W. Van Alstyne, and S. P. Choudary, *Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You* (New York, NY: W.W. Norton, 2016).
4. A. Gawer, "Digital Platforms' Boundaries: The Interplay of Firm Scope, Platform Sides, and Digital Interfaces," *Long Range Planning*, 54/5 (October 2020): 102045, doi:10.1016/j.lrp.2020.102045.
5. G. Parker, G. Petropoulos, and M. W. Van Alstyne, "Digital Platforms and Antitrust," Antitrust Writing Awards 2021, https://awards.concurrences.com/en/awards/2021/academic

-articles/digital-platforms-and-antitrust; G. Parker, G. Petropoulos, and M. Van Alstyne (2021). "Platform Mergers and Antitrust," *Industrial and Corporate Change*, 30(5): 1307-1336.

6. Ibid.; G. Biglaiser, E. Calvano, and J. Crémer, "Incumbency Advantage and Its Value," *Journal of Economics & Management Strategy*, 28/1 (Spring 2019): 41-48.
7. These large-scale data effects, together with the traditional network effects, have resulted in Big Tech firms becoming super-platforms that have also been a main contributor to the dramatic increase in regulatory efforts on how to handle these firms and their growing dominance.
8. Biglaiser et al. (2019), op. cit.; J. Prüfer and C. Schottmüller, "Competing with Big Data," *The Journal of Industrial Economics*, 69/4 (December 2021): 967-1008.
9. Besides industries with high regulatory control, Parker et al. ([2016], op. cit.) also note that industries with high costs of failure (e.g., a mismatched doctor and patient, in comparison to a mismatch between an Uber driver and rider) and resource-intensive industries (e.g., mining) have not fully experienced platform disruption completely.
10. Parker et al. (2016), op. cit.
11. J. S. Bain, "Advantages of the Large Firm: Production, Distribution, and Sales Promotion," *Journal of Marketing*, 20/4 (April 1956): 336-346; M. E. Porter, "Market Structure, Strategy Formulation, and Firm Profitability: The Theory of Strategic Groups and Mobility Barriers," Division of Research, Graduate School of Business Administration, Harvard University, 1977.
12. H. A. Sultz and K. M. Young, *Health Care USA: Understanding Its Organization and Delivery* (Burlington, MA: Jones & Bartlett Learning, 2006); T. H. Marshall, "The Right to Welfare," *The Sociological Review*, 13/3 (November 1965): 261-272; A. Acquisti, C. Taylor, and L. Wagman, "The Economics of Privacy," *Journal of Economic Literature*, 54/2 (June 2016): 442-492.
13. K. M. Eisenhardt and M. E. Graebner, "Theory Building from Cases: Opportunities and Challenges," *Academy of Management Journal*, 50/1 (February 2007): 25-32.
14. K. M. Eisenhardt, "Building Theories from Case Study Research," *Academy of Management Review*, 14/4 (October 1989): 532-550.
15. P. Hartmann and J. Henkel, "The Rise of Corporate Science in AI: Data as a Strategic Resource," *Academy of Management Discoveries*, 6/3 (September 2020): 359-381.
16. "NHS Gives Amazon Free Use of Health Data under Alexa Advice Deal," *The Guardian*, December 8, https://www.theguardian.com/society/2019/dec/08/nhs-gives-amazon-free-use-of-health-data-under-alexa-advice-deal.
17. One exception to this is the case of professional education, where due to less-stringent regulation, various Big Tech firms are making an entry to become a primary service (education) provider.
18. A. Gambardella, S. Heaton, E. Novelli, and D. J. Teece, "Profiting from Enabling Technologies?" *Strategy Science*, 6/1 (March 2020): 1-109; D. J. Teece, "Profiting from Innovation in the Digital Economy: Enabling Technologies, Standards, and Licensing Models in the Wireless World," *Research Policy*, 47/8 (October 2018): 1367-1387; R. Á. Pinto, "Digital Sovereignty or Digital Colonialism?" *Sur: International Journal on Human Rights*, 15/27 (2018): 15-27.
19. If we consider that regulation is in a continuum from state to the national level, we expect state-level regulation to be similar to the case with multiple private actors and one-by-one deals, except this time the counterparty being state actors (and have their sensibilities over the private firms). We thank an anonymous reviewer for helping us to clarify this point.
20. For an example for education, see a report by Accenture in 2019: https://www.accenture.com/_acnmedia/PDF-94/Accenture-Western-Digital-Value-of-Data-Teaching-the-Future-with-AI.pdf.
21. V. Kuraitis, E. G. Anderson, and G. Parker, "Will Google Health Platformize the Electronic Health Record Market?" The Health Care Blog, May 26, 2021, https://thehealthcareblog.com/blog/2021/05/26/will-google-health-platformize-the-electronic-health-record-market/.
22. C. Peukert, S. Bechtold, M. Batikas, and T. Kretschmer (2022). "Regulatory Spillovers and Data Governance: Evidence from the GDPR," *Marketing Science*.
23. While synergies also allow a platform like Amazon to leverage its alternate data captured from its retail platform to form lucrative partnerships for data and AI-based diagnostic services, these synergies are smaller compared with platforms that get exclusive access to individual data from primary services. In healthcare research, for instance, drug and hardware sales data are arguably less valuable than patients' diagnostic and treatment records, but the former can still be useful for entities that do not have direct partnerships with healthcare providers (e.g., Amazon itself).

24. T. Eisenmann, G. Parker, and M. Van Alstyne, "Platform Envelopment," *Strategic Management Journal*, 32/12 (December 2011): 1270-1285.
25. Prufer and Schottmüller (2021), op. cit.
26. Peukert et al. (2022), op. cit.
27. A. Gawer and N. Srnicek (2021). "Online Platforms: Economic and Societal Effects," *European Parliament*.
28. M. Moore and D. Tambini, *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple* (Oxford, UK: Oxford University Press, 2018).
29. Ibid.
30. R. W. Gregory, O. Henfridsson, E. Kaganer, and H. Kyriakou, "The Role of Artificial Intelligence and Data Network Effects for Creating User Value," *Academy of Management Review*, 46/3 (July 2020): 534-551.