Organisation for Economic Co-operation and Development

ECO/WKP(2022)24

Unclassified **English - Or. English**

30 August 2022

ECONOMICS DEPARTMENT

THE VALUE OF DATA IN DIGITAL-BASED BUSINESS MODELS: MEASUREMENT AND ECONOMIC POLICY IMPLICATIONS

ECONOMICS DEPARTMENT WORKING PAPERS No. 1723

By Carol Corrado, Jonathan Haskel, Massimiliano Iommi and Cecilia Jona-Lasinio

Authorised for publication by Luiz de Mello, Director, Policy Studies Branch, Economics Department.

Document available in pdf format only.

All Economics Department Working Papers are available at www.oecd.org/eco/workingpapers.

JT03501487

This document was produced with the financial assistance of the German Ministry of Justice and Consumer Protection. The views expressed herein do not necessarily represent the official views of the German Ministry of Justice and Consumer Protection, or of the Organisation for Economic Co-operation and Development and its member countries.

ABSTRACT / RESUME

The value of data in digital-based business models: Measurement and economic policy implications

A defining aspect of the digital age is data and its business use. Data have become an important input for firms (e.g., to train artificial intelligence algorithms) but data use is neither accounted for in macroeconomic statistics nor part of business contracts for goods and services provided to customers.

This paper puts data and data investments in a framework amenable to measurement and policy analysis aimed at sharpening our understanding of the modern economies. Data is conceptualized as an intangible asset: a storable, nonrival (yet excludable) factor input that is only partially captured in existing macroeconomic and financial statistics. We provide experimental estimates of data investment designed to encompass data and data intelligence for six major European countries (France, Germany, Italy, Spain, and the United Kingdom) and we found an average value of 5 to 6.5 percent of market sector gross value added in 2010-2018 (Corrado et al, 2022). We also develop a simulation exercise to test the potential growth contribution of data capital, and we find that even limited diffusion of data capital could raise labor productivity growth as much as ½ percentage point per year, but outcomes are highly dependent on factors influenced by policy settings.

JEL classification codes: O47, E22, E01

Keywords: intangible capital, data, innovation, productivity growth

****

La valeur des données dans les modèles commerciaux numériques : Mesure et implications en matière de politique économique

Les données et leur utilisation commerciale constituent un aspect essentiel de l'ère numérique. Les données sont devenues un élément important pour les entreprises (par exemple, pour former des algorithmes d'intelligence artificielle), mais leur utilisation n'est pas prise en compte dans les statistiques macroéconomiques et ne fait pas partie des contrats commerciaux pour les biens et services fournis aux clients.

Cet article intègre les données et l'Investment en données dans cadre analytique permettant de les mesurer et d'analyser les politiques en vue d'améliorer notre compréhension des économies modernes. Les données sont conceptualisées comme un actif immatériel : un facteur de production stockable, non rival (mais excluable) qui n'est que partiellement pris en compte dans les statistiques macroéconomiques et financières existantes. Nous fournissons des estimations expérimentales de l'investissement dans les données qui englobent à la fois la collection de données et l'intelligence des données (Data Intelligence) pour six grands pays européens (France, Allemagne, Italie, Espagne et Royaume-Uni), selon lesquelles l'investissement en données représenterait en moyenne de 5 à 6,5 pour cent de la valeur ajoutée brute du secteur marchand en 2010-2018 (Corrado et al, 2022). Nous développons également un exercice de simulation permettant de tester la contribution potentielle des actifs en données à la croissance. Nos résultats suggèrent que même une diffusion limitée de ces actifs pourrait augmenter la croissance de la productivité du travail jusqu'à ½ point de pourcentage par an, cependant ces résultats dépendent fortement de facteurs influencés par les politiques publiques.

Codes de classification JEL : O47, E22, E01

Mots clés: capital immatériel, données, innovation, croissance de la productivité

Table of contents

Tables

Figures

Boxes

The value of data in digital-based business models: Measurement and Economic Policy Implications

By Carol Corrado, Jonathan Haskel, Massimiliano Iommi and Cecilia Jona-Lasinio[1]

1. Introduction

1. A defining aspect of the digital age is data and its business use. Data have become an important input for firms (e.g., to train artificial intelligence algorithms) but data use is neither accounted for in macroeconomic statistics nor part of business contracts for goods and services provided to customers. Even though substantial data accumulation pertains to business operations, e.g., data used to run “factories of the future” or supply chain logistics, much data originate from consumer shopping patterns, lifestyle preferences, and locational habits.

2. The fast-increasing relevance of data is revealed by recent forecasts from the International Data Corporation (IDC). IDC projects that the volume of data generated by business enterprises is increasing more than 40 percent per year (as of this writing). Experimental estimates from Statistics Canada suggest that investments in data and data science were about 2-1/2 percent of the country’s GDP in 2017 (Statistics Canada 2019a, 2019b), and our own (still preliminary) estimates designed to encompass data and data intelligence more fully for six major European countries (France, Germany, Italy, Spain, and the United Kingdom) are larger, averaging 5 to 6.5 percent of market sector gross value added in 2010-2018 (Corrado et al, 2022).

3. These developments have created perceptions that official statistics, business practices and policy frameworks are not suitably adapted to the digital age. Data and investments in data are neither apparent in official macroeconomic statistics, nor reported by companies in public financial statements, nor accounted for in competition and other forms of policy analysis. Perhaps it is unsurprising that consumers are wary of businesses and public agencies that hold their personal information. Furthermore, many

[1] Corresponding authors : Carol Corrado (Carol.Corrado@conference-board.org) from LUISS Lab of European Economics (LLEE) and The Conference Board , Jonathan Haskel (j.haskel@imperial.ac.uk) from the Data Science Institute and the Imperial College Business School, Massimiliano Iommi (miommi@luiss.it) from the Italian Institute of Statistics (ISTAT) and Cecilia Jona-Lasinio (cjonalasinio@luiss.it) from LUISS Lab of European Economics and the Italian Institute of Statistics (ISTAT). The authors thank Giuseppe Nicoletti, Dan Andrews, Alain de Serres and Luiz de Mello (Economics Department) for fruitful discussions and insights, and John Mitchell (Statistics and Data Directorate), Mauro Pisu (Economics Department) and Simon Lange and Vincenzo Spiezia (Directorate for Science, Technology and Innovation), for helpful comments on preliminary versions of the paper. They also thank delegates and participants to the OECD meetings on the “Value of Data” (May 2020) and “Data Shaping Firms and Markets” (June 2021) for helpful comments and suggestions. The authors are indebted to Layla Khalaf for editorial assistance. All remaining errors and omissions are authors' responsibility.

observers view the fruits of data-intensive digital technologies to have been unevenly shared across firms and households and consider the slowdown in productivity growth, having occurred as artificial intelligence (AI) and data use has accelerated, as paradoxical.

4. What, then, is one to make of this post-industrial, “data-driven” age? This paper puts data and data investments in a framework amenable to measurement and policy analysis aimed at sharpening our understanding of the economic mechanisms behind these perceptions. Digital information, i.e., data, is conceptualized as an asset: a storable, nonrival (yet excludable) factor input to production. Data assets are inherently nonrival because they can be used elsewhere in economies, like a blueprint or generic drug formula.

5. Data processing long has been one of the primary functions of computing; indeed, at the dawn of computing technology (in the 1950s), modern industrial economies already contained significant information-processing activity (Machlup 1962). But the advances in computing technology and AI, combined with steep drops in storage costs and ubiquitous high-speed broadband, have led to the generation of vast amounts of stored data (so-called “big data”) in recent decades. Data (i.e., big data) as analysed in this paper is inextricably linked with the diffusion of modern digital technologies. These technologies are referred to as “data technologies” because organizations must exercise (or purchase) the capacity to exploit AI and machine learning tools to obtain value from the stores of data they have amassed (or purchased).

6. Privacy as a dimension of data assets creates new policy challenges for the current digital era compared with previous waves of information technologies. Consumer data privacy protection that restricts the uses of data assets presents a fundamental policy conundrum. Ensuring privacy engenders excludability, and excludability via policies that prohibit re-uses of data (e.g., lifestyle data collected by marketers used for precision medicine solutions) potentially affects the pace of digital innovation. Conversely, policy intervention may be needed to ensure both thriving market competition in data intensive markets (e.g., digital platforms) and protection of consumers.

7. All told, the impact of data in economies has multiple dimensions that raise new challenges for decision makers. Data has an impact on an economy’s innovation potential, and improved market intelligence (e.g., forecasts of product demand) potentially lowers losses due to macroeconomic shocks. The course of business investments in data and data technologies, and of policy interventions to protect consumer privacy in data markets, will determine the strength of these prospects for improved economic resilience.

8. This paper focusses on the economic and policy factors affecting consumer privacy and economic growth in light of the increased use of data in economies; that the increased use of data also engender changes to macroeconomic phenomena (e.g., resilience in the face of shocks) is left for future work. We find that data raised the complexity of the policy trade off between promoting innovation and maintaining competitive markets and that the stakes for decision makers to smoothly navigate the policy trade-offs are very high: Surveys reveal widespread concern among consumers (globally) regarding business use of their personal information, and inattention to these concerns seems impossible. And the potential for data capital to contribute to growth is substantial, e.g., as illustrated below, even limited diffusion could raise labor productivity growth as much as ½ percentage point per year, but outcomes are highly dependent on factors influenced by policy settings, including policies that build effective digital skills and capabilities and promote the inherent capacity for data to be shared will maximize productivity gains.

9. The next section (section 2) begins with the observation that data do not enhance productive activity simply because records of economic transactions accumulate at an astonishing pace at little to no cost, and it sets out the “data as an asset” framework that is used in the economic and policy analysis in the remainder of the paper. Section 3 delves more deeply into the analysis of consumer protection and the specific role of governments as a guardian of competitive markets and the privacy of personal data. Section

4 assesses the potential consequences of data capital on productivity growth, and section 5 offers directions for future research and concludes.

2. Data vs. Data value creation

10. Much work on big data (i.e., large stores of digitized information) focuses on business strategy and "the information value chain". In a best-selling book, Mayer-Schönberger and Cukier (2013) use a rich variety of case studies and document how companies have profited from data in various ways. Others assert further that data is the new oil (The Economist, 2017) and Google's Eric Schmidt is quoted as stating that as much data/information is being created every two days as was created from the dawn of civilization to 2003 (Wong 2012). Such statements suggest that data has significant impacts on economic activity.

11. New investment streams typically accompany the emergence of new technologies, e.g., the invention of the modern internal combustion (IC) engine was followed by a surge of spending on motorized equipment for transport. The seemingly sudden appearance of transport equipment stemmed from its many uses in consumption and production, e.g., personal travel, farming, goods delivery. The arrival of new data technologies such as AI might likewise cause a shift in the composition of investment towards "all things data"—data analytic tools, data stores, structured dataset development, data-derived business strategies—i.e., the appearance of data assets capable of further use in consumption or production.

12. Investments in data assets are neither apparent, nor fully included, in conventional macroeconomic statistics, an important issue reviewed in a companion paper (Corrado, Haskel, Iommi, and Jona-Lasinio 2021). Here, we summarize how data as discussed in much of the business and technology literature can be conceptualized in a framework capable of analysing consequences of the increased use of data on economic activity. We begin by discussing examples of data use, and then introduce our data value and data investment concepts, including their alignment with intangible capital and relation to data monetization and complementary investments in digital technologies. The section concludes with examples of how the framework may be "put to work" in data economy policy discussions.

2.1. Examples of Data Uses

13. Consider first examples of data use in modern economies. Table 1 lists examples of data uses, grouped according to whether the type of data is rival or nonrival. Data is inherently nonrival, as previously discussed. The classifications in the table represent the degree to which data are openly shared with the public or other organizations in the given example.

14. As may be seen, the uses listed on lines 1–5 mainly reflect applications of new digital technologies by firms, i.e., digital platform-based businesses and/or applications of machine learning and other AI-based algorithms to massive data. Product-led growth strategies (line 6) refers to marketing innovations based on user feedback data (also enabled by new technologies), and line 7 refers to the fact that customer lists and after-sales customer feedback long have been inputs to brand development, marketing, and customer retention strategies.

15. Examples of "nonrival" data use range from "new technology" marketers of personal data for B2C companies (line 8), to examples of industry-level data sharing, e.g., financial records held by credit bureaus and shared across financial institutions (line 9), vehicle accident and major repair records shared by buyers and sellers in used car markets (line 10), and personal medical records shared across providers of medical care services (line 11), to cross-platform and cross-purpose uses (lines 12 and 13). Finally, the table lists two examples of government open data.

Table 1. Examples of data and use

Rival
1. Product-level forecasting (e.g., Amazon)
2. A/B Internet testing and marketing (e.g., Google)
3. IoT factory systems (e.g., Siemens)
4. Targeted advertising on consumer content platforms
5. Fintech (e.g., algorithmic trading, digital lending, etc.)
6. Product-led growth strategies (e.g., Slack)
7. Customer lists/after sales services design
Nonrival
8. Financial records (FICO scores)
9. Vehicle records (CARFAX reports)
10. Personal medical records (across service providers)
11. Open-source data generated by web users (map data)
12. Private by-product data put to alternative uses (e.g., research)
13. Genomic and other public biomedical research data
14. Official statistics (economic, demographic, social)

Note: Data is inherently nonrival, and classifications reflect the degree to which owners share data with other organizations or the public.

16. The examples in the table suggest that the macroeconomic consequences of data are pervasive. Data is seen to drive business adoption of new technologies, which fosters new product development and lowers operation costs (production, distribution, marketing); data also reduce financial risk and information asymmetries in asset markets; and some organizations and governments apparently produce and provide significant stores of information assets for "free."

2.2. Data Value Creation: An economic framework

17. Data does not provide a flow of services to production simply because records of economic transactions accumulate at an astonishing pace at little to no cost. The accumulation of digital by-product data within the business sector has the potential to boost real output only when the sector also invests in transforming such records, possibly along with other available economic or social information, into analytical insights and actionable business intelligence.

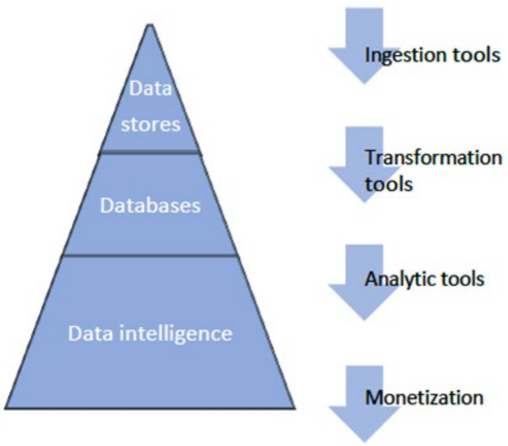
18. Rather than focus on features of data, i.e., big versus small amounts of them, or the speed of their accumulation, our framework treats data stores and knowledge gleaned from data stores via application of data technologies as long-lived assets that contribute to final production in an economy. The long-lived appropriability of accumulated stores of digitized information provides the basis for treating business spending on data as a capital investment and for including services of data assets as capital input in productivity analysis.

19. Our specific approach embraces both the technology and management literatures. Technologists characterize data according to a "data stack" that describes the transformation of raw data into usable data

structures and intelligence; their approach sequences data forms and tools in a single stack.[2] Business management strategists use a value chain construct that adds monetization, or market implementation, as an ingredient to (or tool of) data value creation.[3]

20. Both characterizations, i.e., of the technologists and the management strategists, are embedded in the framework for data value creation illustrated in figure 1. On the left, we identify three major data asset types based on a value chain notion. The sequencing of tools (depicted on the right) with data forms is implied (ingestion tools are used to create data stores, etc.). A key implication of the framework is that data value creation involves the application of layers of data technologies and monetization to create assets that generate productive value in an economy.

Figure 1. Data value creation: Data assets, digital tools, and monetization



Note: The stack to the left depicts the stages of data asset value creation based on applying layers of tools shown on the right.

21. The data asset stack has three layers of value—data stores, databases, and data intelligence—each corresponding to an asset type amenable to measurement.[4] The three data asset types are defined more precisely as follows:

- *Data stores* are raw records that have been stored but not yet cleaned, formatted, or transformed for analysis, e.g., data scraped from the web or sensor and economic data captured from production or transactions activity. Raw records also cover the raw data collected from experiments, statistical surveys, or administrative records.
- *Databases* consist of transformed raw data, records that have been cleaned, formatted, and structured such that they are suitable for some form of data analytics or visualization.

[2] See, e.g., Roca (2021a), for a recent depiction. The data stack has its roots in information science, which uses the concept of a "data pyramid" to depict the relationship between data, information, and knowledge (Varian 2018).

[3] See again Mayer-Schönberger and Cukier (2013), also PriceWaterhouseCoopers LLP (2019).

[4] A companion paper describes measurement of this framework. All told, our approach builds upon previous economic approaches to defining and measuring data, including McKinsey Global Institute (2016), Goodridge and Haskel (2016), Statistics Canada (2019a, 2019b), Nguyen and Paczos (2020), and Corrado (2021).

- *Data intelligence* reflects the further integration of data with advanced analytic tools (e.g., machine learning training algorithms); data intelligence is a set of quantitative inputs that provide actionable guidance for decision-makers, including solutions to scientific problems.

22. What separates the "modern" data stack from legacy systems is that modern systems are hosted in the cloud, requiring little technical configuration by users. According to technologists (e.g., Roca 2021b), "the modern data stack lowers the technical barrier to entry for data integration." And "components of the modern data stack are built with analysts and business users in mind, meaning that users of all backgrounds can not only easily use these tools, but also administer them without in-depth technical knowledge".

23. The data value chain framework, in which greater value added is created as raw data is processed and developed into insights and solutions, applies to data-driven R&D processes as well as data-driven development of new customer platforms and organizational practices.[5] From this perspective, i.e., a knowledge-based or intangible capital perspective, the increased use of data assets derived from modern digital technologies is an "innovation in the method of innovation."

24. Conceptually, modern data use fosters faster, more efficient experimentation and feedback in R&D processes, manufacturing production processes, marketing research, and business strategy and operating model development. This implies that the "productivity" of these activities improves, or that their unit cost (i.e., their resource cost per unit of final output) falls, the implications of which are examined in section 4 of this paper.

2.3. Personal Information in the Data Value Chain

25. As explained in greater detail in our companion paper on measurement, we develop estimates of data capital based on the costs incurred in creation, i.e., in the "work done" in each layer of the data stack shown in figure 1 (including complementary investments in digital tools). The estimates for the value of data capital that we generate thus reflect the resource cost value of all data processed, transformed, and used in an economy.

26. But because some of the largest and fastest growing internet companies (Alphabet, Google, Twitter etc.) are built mainly on the economics of transforming personal information into business and marketing intelligence, the valuation of personal data (versus all data) is viewed with keen interest. The World Economic Forum (2011) and OECD (2013) identify two broad categories for data—personal data and institutional data—based on the economic sector of origin of the information. These categories are not very amenable to measurement but help clarify conceptual issues regarding the valuation of personal information.

[5] Although the three asset types shown in figure 1 generally align with categories Statistics Canada set out in a conceptual framework for measuring data, Statistics Canada calls the third category "data science" and views it as unmeasured R&D, e.g., spending to develop new AI algorithms. Though data and data tools (AI) are inextricably bound via feedback and training data used to develop and refine AI tools, the data stack/data value chain notion of how value is created from data does not end with the development of algorithms.

Table 2. Classification of data by origin of raw information

	Institutional Data		
Personal Data	*Businesses*	*Governments*	*Non-Profits*
User-Generated	Personnel Files	Personnel Files	Personnel Files
Behavior	Accounting Records	Accounting Records	Accounting Records
Social	Legal Docs	Legal Docs	Legal Docs
Location	Financial Docs	Financial Docs	Financial Docs
Demographic	Customer Lists	Intelligence Records	Social Policy Programs
Official Identification	IoT Sensors	Diplomatic Cables	Public Policy Programs
		Defense Files	
		Statistical Surveys	
		Regulatory Records	
		Admin Records	
		Monitoring Tech	

Source: Adaptation of Kornfeld, Robert (2019), slide 8, which was drawn from WEF (2011) and OECD (2013).

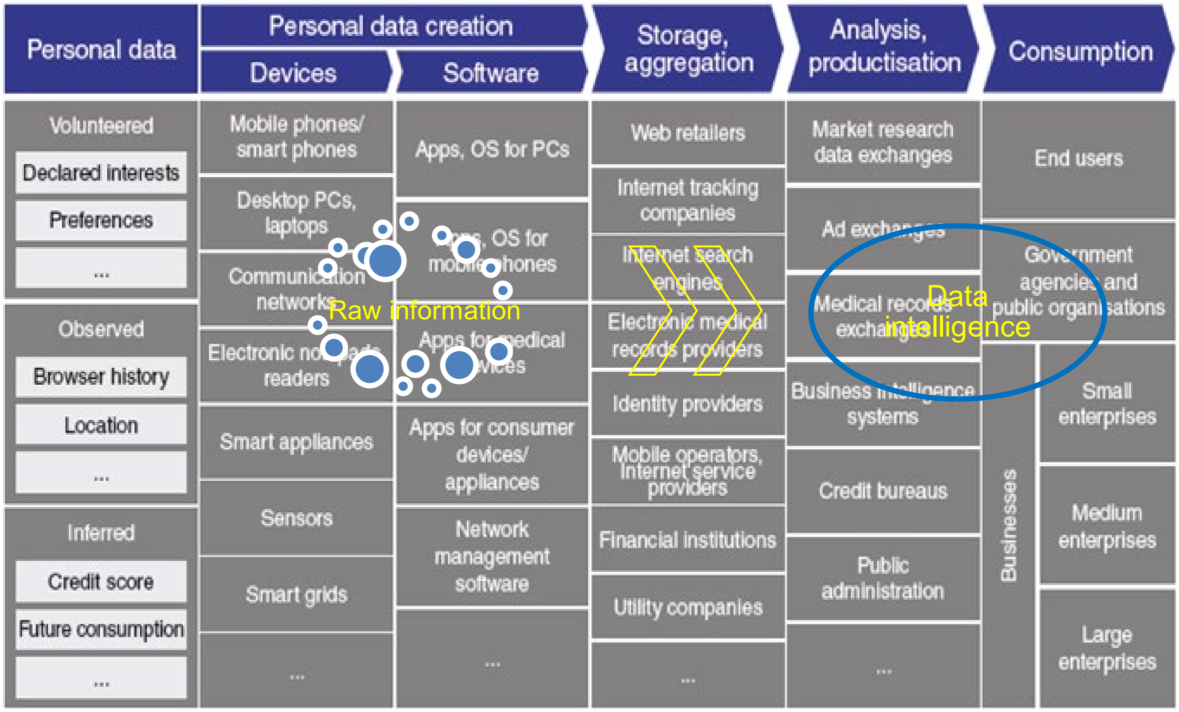
27. Table 2 sets out examples of "raw" data by economic sector origin based on the WEF classification. First, note that the collection, transformation, and use of the circled data items in the table—data on persons collected online and data collected by businesses via Internet of Things (IoT) sensors—are generally considered to be what is "new" in the modern digital age. Second, operational data common to all institutions (the shaded area) are also a source of digitized information, and they and the information systems that generate them long have been exploited by managers and operators for competitive advantage. Finally, as previously noted, nonmarket institutions (governments and non-profits) generate rather vast stores of information and are working to make the data they collect more "open", i.e., freely available for anyone to download, modify, and distribute without legal or financial restriction.[6]

28. Personal data stores are collected and transformed via an extensive ecosystem of digital tools and business processes illustrated in Figure 2. The figure may be seen to depict a value chain—the personal data value chain—with the pyramid of value creation placed on its side. Inflows of raw information are on the left and outcomes in the form of data intelligence are on the right (as indicated by the overlay).[7] The personal data value chain is thus a conceptual construct that sits within the overall data value chain in which public open data and business-specific information also reside and contribute to value creation.

[6] In fact, the Open Data Institute (ODI) in the UK estimates that the use of "core" public open data alone—data such as addresses, maps, weather, and land and property ownership records—currently contributes an additional ½ percent of the country's GDP in economic value every year (ODI, 2016).

[7] The first column in the figure uses descriptions set out in WEF (2011, p. 7), which categorizes personal data according to whether they are either volunteered, observed, or inferred.

Figure 2. The Personal Data Ecosystem/Value Chain



Source: Adaptation of figure 4 in WEF (2011, p. 15).

29. The value for data that we measure and analyse using the data stack will thus incorporate a broad range of digitized information: from the value of vehicular traffic and weather data processed for use in transportation logistics to the marketing value of consumers' personal information and online activity—value that may derive, at least in part, from its combination with institutional data. Seen from this perspective, the value of personal data as an economic resource cannot be readily disentangled from the value of other data records in an economy or the value added in processing data for final use.

2.4. *Data Capital≅ vs Intangible Capital*

30. Data investments and assets are closely related to activities captured in available estimates of intangible capital. The broad categories of intangible capital and its components by investment type are summarized in table 3, which displays the expanded investment framework due to Corrado, Hulten, and Sichel (2005, 2009). Intangible investment covers a wide class of investments, from databases to business processes, that would appear to be relevant for analysing the consequences of the increased use of data in economies.

Table 3. Intangible capital: Broad categories and investment types

Digitized information	• Software • Databases
Innovative property	• R&D • Mineral exploration • Artistic, entertainment, and literary originals • Attributed designs (industrial) • Financial product development
Economic competencies	• Market research and branding • Operating models, platforms, supply chains, and distribution networks • Employer-provided training

Source: Corrado, Hulten, and Sichel (2005, 2009).

31. A companion paper sets out a detailed assessment of the overlap between the processes displayed in figure 1 with estimates of intangible capital developed from measures of the investment streams listed in table 3.[8] It concludes that most expenditures on data-related capital formation—on data assets, on data technologies and their development via R&D—are included in available estimates of intangible investment, e.g., the INTAN-Invest database that provides industry-level estimates of intangible capital for many OECD countries. INTAN-Invest estimates of intangibles have been recently incorporated in an analytical production account module available via the ongoing EUKLEMS-INTANProd project.[9] The analytical module capitalizes all ten intangible investment streams listed in table 3. It uses official macro data for the subset of components that are capitalized in the macro data of most OECD countries: computer software (blended with databases), R&D, mineral exploration, and entertainment, artistic, and literary originals.

32. A major limitation of official macro data is that data intelligence is not fully captured, except to the extent it is included in R&D. National accounts thus miss much of the most valuable, final stage of the data value chain as it pertains to business practices. By contrast, the use of data to drive the development of operating models and platforms, new marketing strategies and product diversification is covered in INTAN-Invest.

33. All told, we believe that the available estimates of intangible investment cover most investments in data technologies via software-driven data tools and structured databases, as well as data-derived business intelligence. Though these empirics have measurement limitations discussed in the companion paper, their comprehensive coverage of the data value chain and of complementary digital tools, underscores their utility for analysing the implications of the increase in the use of data in economies.

[8] The assessment is based on the analysis in Corrado, Haskel, Iommi, and Jona-Lasinio (2019, pp 375-76) and Statistics Canada (2019a, 2019b) and in reviews by Edquist, Goodridge, and Haskel (2020) and Corrado (2021).

[9] This update/expansion is being funded by a grant from the European Commission's Directorate-General for Economic and Financial affairs (DG ECFIN). Data and further information are available at https://euklems-intanprod-llee.luiss.it/.

2.5. Putting the Framework to Work

34. Here we consider data capital as intangible capital to summarize some implications relevant to data policy discussions.

2.5.1. Macro perspective on market power

35. The intangibles framework helps to interpret commonly used indicators of market power, namely, price markups, share of rents in profits, and average rates of return. Studies that use conventional measures of variable costs (labor and materials) to calculate markups will erroneously depict the degree to which price markups may be rising to the extent data capital costs are ignored. This is because, if data capital is growing in relative importance in industries, incomes to its owners will rise in relation to the total value of production. This fact alone will bias conventional measures of unit costs and profits attributable to rents. (Unit costs will fall to an erroneous degree, and rents will rise correspondingly.)

36. Most firm-level datasets do not include amortization costs for intangibles, which suggests market power indicators derived from these datasets are misleading unless the omission has been mitigated.[10] This is more than a measurement issue in that markups calculated using firm-level datasets are widely used and discussed in policy settings (e.g., along with other indicators, they were included in the IMF's April 2019 World Economic Outlook chapter on the rise of corporate power).

37. In macro data, the price markup is the inverse of labor's share of total factor income. National accounts cover 40 to 50 percent of total intangible investment, based on estimates from INTAN-Invest, which suggests that studies using macro data—whether national accounts or the expanded investment measures in INTAN-Invest—will arrive at different conclusions regarding trends in market power than studies based on firm-level data. In fact, a much more moderate rise in markups in the United States is found using national accounts data (Eggertsson, Robbins, and Wold 2018) than is suggested studies that rely on firm-level data (e.g., De Loecker, Eeckhout, and Unger 2019).[11]

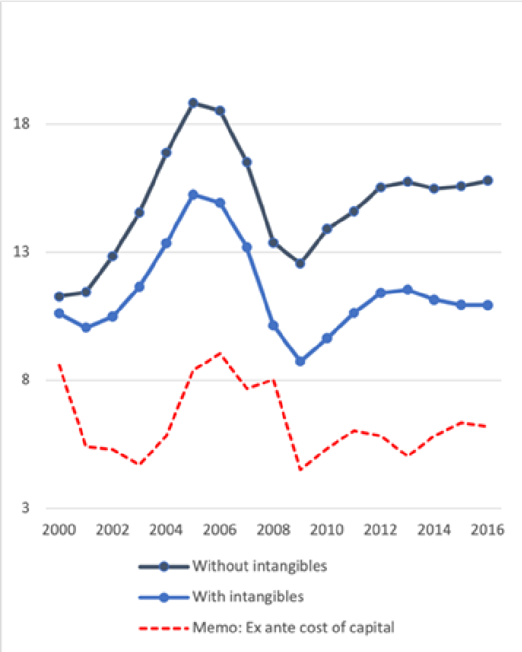
38. Figure 3 shows that the ex post after-tax rate of return implied by macro productivity data extended to cover all intangibles—based on the INTAN-Invest extended EUKLEMS productivity dataset for the United States—is dramatically affected by the inclusion of intangibles. Ex post average rates of return calculated using macro data reflect trends in competitive capital costs as well as price markups, and capital markets have their own dynamics in terms of risk premia and trends in market rates. An aggregate ex ante capital cost calculated as a weighted average of the expected return on stocks and after-tax cost of debt also is shown.[12] While the gap between the ex-ante rate and the ex post rate calculated using all intangibles fluctuates, it does not materially widen in the post financial crisis period, suggesting it is unlikely that the competitive intensity of the US economy changed significantly since then. Note that the markup story in the EU is rather different (Gutierrez 2017, Battiati et al. 2021). EU countries are less intangible intensive and labor shares have not declined to the extent they have in the United Kingdom and the United States.

[10] Details of how intangibles are recorded in firm-level data based on company financial reports and establishment-level data are reviewed in our companion paper.

[11] Koh, Santaeulàlia-Llopis, and Zheng (2020) contend the capitalization of IPP products (software, R&D, and artistic, literary, and entertainment originals) in the US national accounts is the entire explanation for the country's decline in the labor share, though note that a trend toward investments in tangible assets with higher depreciation rates also will contribute to a decline in labor's share of total factor income.

[12] Specifically, the weights reflect leverage for the nonfinancial corporate sector as a whole; the return on stocks is measured is the sum a risk free rate (10-year Treasury bond rate) and equity risk premium from (Damodaran 2014, updated here and cost of debt is the after tax corporate BAA bond rate.

Figure 3. Rate of return (after tax), United States



2.5.2. Industry perspective on digitization and Covid-affected industries

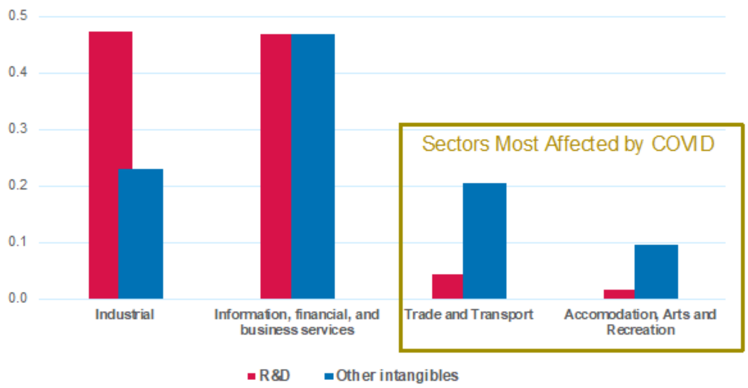
39. OECD research finds that the intangible intensity of an industry is correlated with independent indicators of digitization at the industry-sector level, a link that is not driven by R&D and that occurs largely in industries outside manufacturing.[13] Additionally, intangible investment and output growth are correlated at the sector level in the INTAN-Invest database. These linkages suggest that intangible investment captures firms' efforts to use digital technologies to change a business model and provide new revenue and value-producing opportunities.

40. The linkages also imply that intangible investment offers a perspective relevant to stimulating sectors most affected by the COVID pandemic. As may be seen in figure 4, the sectors most affected by the pandemic invest disproportionately in non-R&D intangibles. As a group, these assets largely represent the data-driven creation of business intelligence. The figure thus underscores that to "see" how digital technologies are diffusing throughout the economy, we need to look at investment expanded to include the full complement of intangibles.

[13] Corrado, Criscuolo, Haskel, Himbert, and Jona-Lasinio (2021). Indicators include ICT equipment, purchases of IT services, and employment of workers in "tech" occupations.

Figure 4. Sector distribution of intangible investment, UK 2015

Sector share of investment



Note: R&D is combined with mineral exploration for the industrial sector. Estimates cover market sector industries as defined in EUKLEMS.
Source: ONS experimental intangible investment data.

2.5.3. Firm-level perspective on increased productivity dispersion

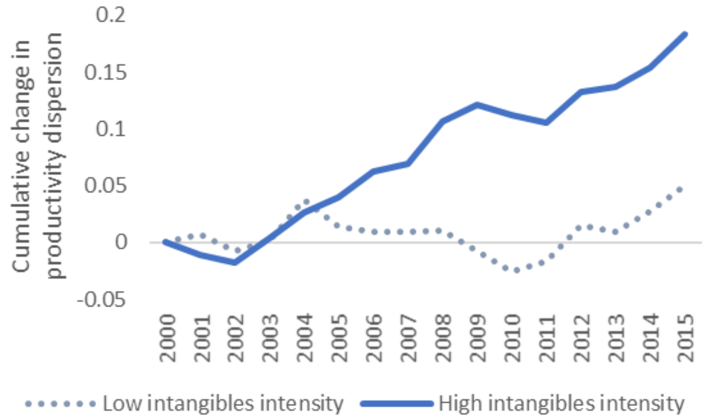
41. Andrews, Criscuolo, and Gal (2016) documented substantial increases in productivity dispersion using cross-country company financial data, suggesting a worrisome decline in the diffusion of new ideas and technologies in the 21st century. The inability of this and related studies to account for intangibles as factors of production in generating (or explaining) the uneven within-industry firm performance outcomes raises questions about interpretation and appropriate policy responses, as noted by, e.g., Crouzet and Eberly (2018, 2019) and Haskel and Westlake (2018). Corrado, Martin and Wu (2020) provided evidence that the deployment of intangible assets explains a good part of the unevenness in corporate performance via examining financial market outcomes of large public multinational companies, but country-industry controls and policy analysis were not part of this study.

42. Research by the OECD using its cross-country, establishment-level MultiProd database in conjunction with INTAN-Invest industry-level investment data examines determinants of the within-industry increase in productivity dispersion considering data capital and digital technologies use.[14] The approach and data used are more comprehensive than many microdata-based studies that have examined competition and productivity performance in view of modern digital capabilities. The OECD research demonstrates, e.g., that the intangibles/data capital intensity of an industry is positively correlated with both (a) firm size and rising industry concentration and (b) increased within-industry productivity dispersion.[15]

[14] MultiProd is a database developed via a distributed microdata approach pioneered in the early 2000s for cross-country analysis of productivity (Bartelsman, Scarpetta, and Schivardi 2003, Bartelsman, Haltiwanger, and Scarpetta 2013).

[15] These findings refer to research reported in Bajar, Criscuolo, and Timmis (forthcoming) and Corrado, Criscuolo, Haskel, Himbert, and Jona-Lasinio (2021). See also Gal, Nicoletti, von Rueden, Sorbe and Renault (2020) and Nicoletti, von Rueden, and Andrews (2020).

Figure 5. Evolution of Within-Industry Productivity Dispersion by Intangibles Intensity, 2000 to 2015



Note: Estimation is based on the OECD MultiProd database in combination with INTAN-Invest. Countries included are AUT, BEL, DEU, DNK, FIN, FRA, IRL, ITA, NLD, PRT. The graph plots the evolution of productivity dispersion over time within manufacturing and market services industries. Unweighted averages across two-digit industries are shown, normalized to 0 in the starting year. Productivity dispersion is measured as the 90-10 difference in multi-factor productivity a la Woolridge, i.e., the difference in productivity between firms at the 90th percentile of the productivity distribution in a country-industry and firms at the 10th percentile. The vertical axes represent log-point differences from the starting year: for instance, productivity dispersion in market services has increased by about 0.11 in the final year, which corresponds to approximately 11% higher productivity dispersion in 2015 compared to 2000.
Source: Corrado, Criscuolo, Haskel, Himbert, and Jona-Lasinio (2021).

43. Figure 5 displays the direct relationship between productivity dispersion within narrowly defined industries and the intangible intensity of the industry found in this work. Productivity is measured using the MultiProd microdata, *which does not account for intangibles in capital input*. This research suggests that the mechanism behind increased productivity dispersion is mainly the growing relative importance of intangible investment, especially investments in economic competencies by market services industries.

44. That digital transformation is an important driver of these results is consistent with the framework set out above where software and economic competencies are the most data intensive components of intangible investment. The study attributed its findings largely to scale economies associated with intangibles, though it could not identify whether the economies reflected increasing returns due to, e.g., data agglomeration effects within firms or externally driven network effects of technology platforms. Even so, policies to alleviate the uneven gains from digital transformation (e.g., by easing access to intangibles for disadvantaged firms) and ensure that its benefits are shared more widely would appear to be needed.

2.6. Recap

45. This section argued that data and data capabilities are intangible assets, and that the intangible investment framework is useful for analysing the consequence of increased data use in economies.

46. Our primary finding is that the economic activity associated with the modern “data stack” concept is generally captured by available measures of intangible investment, and that this has several policy implications. One implication is that data amortization costs or the size of data assets can be included in competition analysis of data-intensive industries and sectors via accounting for intangible assets. Another is that commonly used indicators of digital transformation and intangible investment are correlated at the industry level, suggesting that policies designed to stimulate productivity growth via digital transformation can be modelled using the framework.

47. Finally, it appears that the unevenness in productivity outcomes within industries observed since the turn of the century is associated with the growth of data/intangible capital. This suggests that

technology (not monopoly per se) is the dominant force behind rising market concentration and market power in data-intensive industries. Thus, there would appear to be a need for policies to both stimulate digital capabilities in laggard firms and sectors of economies while maintaining or spurring competition in digitally driven markets.

3. Consumers Data markets and privacy

48. Data as sharable economic and social information on consumers raises complex issues about data markets and their efficiency and impact on consumer welfare. The privacy dimension characterizing consumer data set it apart from the other intangibles, and the treatment of individual consumer information poses new challenges to competition law and consumer law regulators looking at data as a source of market power and their use a potential risk to privacy. This section explores questions surrounding these issues.

3.1. Personal Information and Policy Trade-offs

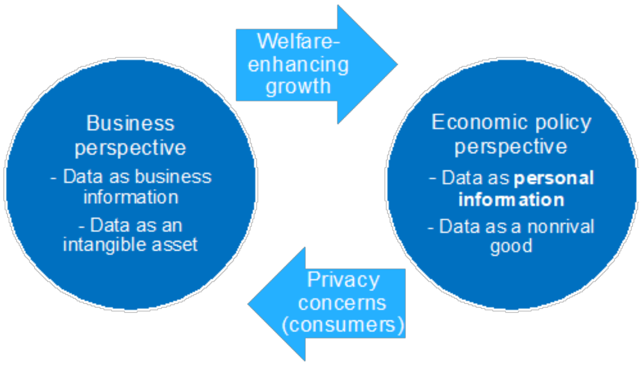
The trading of personal information such as credit card ratings or Social Security numbers in data markets raises many concerns among consumers. Companies collecting and using personal data are responsible for their protection and need to mount effective cyber security systems to prevent exposing consumers to undue risks from the loss of privacy. But even if companies are perfect on this score (and they are not), individual companies are not able to protect (or know the uses of) the data that they share or trade with other companies.

Thus, policymakers face a trade-off: encouraging data sharing as a factor affecting the development of the digital economy while ensuring the privacy and protection of personal information. To balance this trade off, policy makers need more information, preferably quantitative, on the welfare benefits of data sharing and assessments of the degree to which data privacy protection impinges on these benefits. The literature on data markets and economics of privacy is flourishing, but the evidence base on policy implications is still emerging and guidelines for assessing conflicts and trade-offs are still unsettled.

Theoretical contributions on the economic impacts of data sharing propose some relevant and interesting insights, however. Jones and Tonetti (2020) argue that data sharing enables firms to produce and deliver greater variety and choices to consumers, which is welfare-improving. Acemoglu, Makhdoumi, Malekian, and Ozdaglar (forthcoming) argue that the costless replicability of data and other externalities depress the price of data and lead to excessive data sharing; they conclude that shutting down data markets improves welfare under certain conditions. Both studies treat data acquisition as costless and do not model the data value chain, in which raw data obtains economic value once transformed into intelligence and actionable guidance for decision makers.

It is likely too early to solve the policy trade-off as there are still many dimensions of data use and their economic consequences to be investigated. However, when assessing the economic effect of data regulation, it is relevant to distinguish between business and consumer perspectives on data (Figure 6) as they raise different policy challenges. Companies demand data and view them as one of their core intangible assets critical to business performance and to welfare enhancing growth (see sections 2 and 4). From the consumers perspective, the challenge is about solving the conflict between their privacy preferences and data-sharing needs to satisfy their service demands.

Figure 6. Perspectives on Data



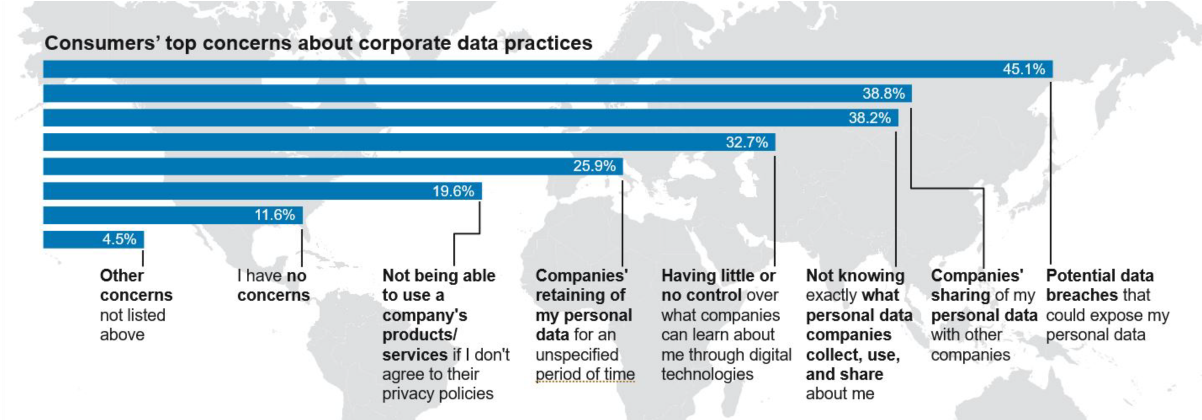
49. Jones and Tonetti (2020), suggest focusing the attention on the ownership of data to work out the trade-off. They investigate different allocations depending on firm or consumer ownership of personal information considering the nonrival nature of data. When firms own the data, they have a competitive advantage and are not stimulated to share their data with competitors thus imposing a social cost that limits competition and discourages innovation. If instead consumers control their personal data, there are efficiency gains as they would protect their own privacy and would be motivated to sell their data to various companies and organisations. Jones and Tonetti conclude that the welfare gains will be larger when consumers own their data, suggesting that a market-based approach to privacy regulation might improve social welfare. The digital pioneer/visionary Jaron Lanier (2010, 2013) calls for an exchange, an economy of micropayments that compensate consumers for business use of their personal data, including their creation of user content (Lanier 2010, 2013; see also Arrieta-Ibarra et al. 2018).

50. Information on consumers preferences concerning data privacy is a relevant input for policymakers who need to answer to simple questions such as: What information do consumers wish to protect? What information are consumers willing to share? And who do consumers trust to oversee their privacy? What are the actual attitudes of consumers on data practice and regulation?

51. A global survey of consumer attitudes about data practices was recently conducted as part of The Conference Board® Global Consumer Confidence Survey, an online survey of almost 32,000 consumers in 63 markets, conducted in collaboration with Nielsen.[16] The sample included 500 respondents per market, representative of each market's online population by age and gender. The survey revealed that consumers value free content *significantly more* than personalized content as a benefit of sharing their data. In the sample, about 40 percent of consumers are willing to be tracked in exchange for discounts, certain services, or information related to their current location. Discounts on car and health insurance are among the most convincing benefits to data sharing. Finally, the survey revealed that one in five consumers have bought a brand less, or abandoned it, over a company's data practices.

[16] The survey was conducted in February 2020; see Dalhoff (2020).

Figure 7. Consumers' Attitudes about Data Practices



Source: Consumers' Attitudes about Data Practices, The Conference Board
Question (Q1): What are your biggest concerns about companies' data practices regarding your personal data? Please select up to 3 choices. N = 500 per market, 63 markets.

52. The survey's overall results show that individuals are increasingly willing to share their personal information, but they remain concerned about data security, third-party sharing, and lack of transparency (figure 7). The findings support the "privacy paradox," namely, that most individuals care about their privacy but at the same time they readily share their personal information for free or for a small compensation. Consumers seem to be much more sensitive about the use of their medical, financial, and strictly personal information (gender, age, address) than about their brand consumption or media usage attitudes; apparently, people do not wish to lose control of data that are strongly linked to their personal character (Dahloff 2020; see also McKinsey Global Institute 2021).

53. Regulators are frequently asked to renew/update privacy laws in view of high-profile exposures of personal information and growing demand from consumers for stronger protection. These may be seen as calls for a shift of privacy regulation toward better consumer protection (McKinsey Global Institute 2021), an interpretation underscored by consumers attitudes reported in The Conference Board's survey. Consumers expressed a preference for external oversight of business data practices of some sort.[17]

54. Moves in this direction include the General Data Protection Regulation (GDPR) in Europe, the California Consumer Privacy Act (CCPA) in the US, and the Consumer Data Right (CDR) in Australia. The GDPR and CCPA both safeguard large markets and have extraterritorial effects, as they cover businesses incorporated beyond their respective jurisdictions, with GDPR's provisions being stricter than CCPAs. Still, the nature of regulation differs significantly across jurisdictions, which poses operational challenges for multinational companies.

55. Overall, the main policy challenge is to promote better privacy protection and harness digital dividends for consumers while encouraging sustainable development of the data economy. Data privacy legislation is at a very early stage, however, and it is difficult to assess its effectiveness. While GDPR has had some unintended consequences according to some studies (e.g., Aridor, Che and Salz 2021), most reviews see GDPR as the first step toward a transparent and effective data privacy system for which the complete and ideal course of action has yet to be identified (e.g., Accessnow Progress Report, 2020).

[17] Most consumers globally reported preference for a government data watch dog, though private approaches—a private consumer advocacy/protection group or organizations consisting of company and consumer representatives—seem to be favored by consumers in the United States.

3.2. Data Markets and Competition Policy

56. The digitalization of the economy and the consequent growing volume of data shaped entirely new markets for data in different domains, from financial markets to advertising, navigation services and many others (Spiekerman, Acquisti, Böhme, and Hui 2015). Already in 2011, the World Economic Forum recognized the rise of complex ecosystems of entities collecting, analyzing, and trading personal information, i.e., the operation of data markets. The identification and definition of data markets usually considers two main elements, data collection and use, in recognition of the fact that data gain value both through ingestion and eventual use.

57. In principle, there are many data markets because data trading involves different dimensions (asset types, services/use types). Typically, consumers supply their own data to firms in a primary business to consumers (B2C) market. Firms acquire this data either to build data intelligence for themselves, or to exchange the data (perhaps after adding value through processing and transformation) with other firms in a secondary business to business (B2B) market. These markets interact posing regulatory concerns as the consumer's attitude for sharing individual information with the collectors (firms) will depend on the sharing conditions in the primary market as well as on the use of these data by the firm in the secondary market.

58. Also, trade transactions of customer data activities (personal information) remain mostly in the domain of companies' proprietary/private custody thus generating large benefits to big network and technological advanced corporations and favouring dominant business positions. Therefore, the search for the best balance between privacy protection and data sharing has important implications for competition policy.

59. In this setting, firm's competitiveness is going to increasingly depend on timely access to relevant data and other intangible assets to be used to develop new innovative products (EC Report, 2019). The main concern is that under the current competition policy framework, the firm's propensity to data hoarding could be unlimited, with the firm's data treasure becoming a source of competitive advantage impossible to be overcome by its competitors (Haucap, J., 2019). In turn, this would preclude market contestability and the related social benefits from data use.

3.2.1. Abuse of dominance

60. What matters for competition policy is whether a firm is abusing its market position. The analysis of abuse of dominance typically focuses on the option for consumers to substitute away from the products supplied by the firm under examination or the option for the firm's competitors to obtain access to the scarce asset, in this case data, necessary to supply the product. Thus, a natural question is whether defining data as an asset can provide a key to approach such competition policy issues.

61. In this regard, the data characteristic of non-rivalry is pivotal. Non-rivalry favors economies of scale and network externalities. The emergence of increasing returns to scale from the use of data facilitates market concentration, thereby raising concerns by competition regulators that incumbent firms may use their control over data in an abusive manner.

62. As discussed in section 2.5, recent empirical studies suggest that intangibles are associated with higher productivity dispersion between the most productive firms and the rest of the firms predominantly through their complementarity with digital technologies. As digital technologies necessitate intangible investment, laggard firms that are unable to carry out the necessary intangible investment fall behind in digital intensive sectors. This might suggest that possession of intangible assets, such as data, creates an unassailable advantage for incumbents. For instance, De Ridder (2019) argues that firms using intangibles more efficiently can undercut their competitors on price, as they can scale up their production and divide their fixed costs over more units of production. The ensuing persistence in market leadership can deter innovation from entrants or even prevent potentially innovative firms from entering the market altogether.

Intangible assets therefore favor an increase in productivity dispersion via the rising market power of firms that are efficient at exploiting them. This phenomenon is reflected in the "winner take most" characteristics of super star firms (Autor et al 2020), which can potentially be exacerbated by the failure to carefully revise standard competition rules.

3.2.2. Merger control

63. Recently, the approach of antitrust authorities has been questioned as they generally focus only on short term impacts of M&A between digital companies while it is more relevant to monitor long run effects that are likely harmful both to competition and consumers. Scholars agree on the need of adopting a more dynamic merger control but disagree on the revision of antitrust rules. For example, Cabral (2020) claims that current antitrust regulation predominantly focused on consumer welfare is suitable to protect both consumers and competition thus not requiring substantial revisions besides the adoption of a more dynamic approach. Furman et al (2019) instead see a complete update of merger policy as necessary to guarantee consumer and innovation protection while preserving market competition. They clarify that large digital companies may be beneficial for consumers or businesses but only in a framework where they do not abuse their position by unfairly expanding or protecting it. Antitrust laws must be revised enabling faster actions directly targeting and addressing abusive behavior and creating space for new businesses in a more pro-competitive context. This would entail generating new opportunities for competition innovation, and consumer choice.

64. At this stage the debate is still unresolved, but it is widely agreed that there is a need to update antitrust policy to foster competition while ensuring consumer protection and privacy. It is also critical not to deter innovation in digital markets promoting pro-competitive measures favoring new entrants. Overall, UNCTAD (2019) suggests a holistic approach favoring close cooperation among authorities to promote consumer and data protection as well as competition.

3.2.3. Data as an essential facility and regulation

65. Another issue currently discussed by competition regulators is whether data should be considered as an essential facility. Where, using Graef's terminology," an *essential facility* is an asset or infrastructure to which a third party needs access to offer its own product or service on a market Graef (2016). A facility is essential if no reasonable alternatives are available, and duplication of the facility is not feasible due to legal, economic or technical obstacles." More specifically, Abrahamson (2014) advocates that data can be considered as an essential facility if the firm controlling access to data, adopts an exclusionary anticompetitive conduct by refusing to give access to them thus impeding competition from rivals and impairing consumers.

66. Conversely, Tucker (2019) argues that it is unlikely for digital data to be considered as an essential facility as this requires data to be unique and that there is no alternative input. Rather data are non-rival and ubiquitous as their consumption by someone does not prevent their availability to others.[18] This debate is inconclusive, however, and it is not clear whether the perspective offered by the 'essential facilities doctrine' could be useful in identifying proper data regulation.

3.3. Summing up

67. The COVID-19 pandemic has demonstrated that an organized and controlled use of data can generate high quality outcomes and strongly increase the efficiency of processes to generate social gain. As an example, sharing data among hospitals and health research institutions across the globe has

[18] See: https://www.ftc.gov/news-events/press-releases/2014/05/ftc-recommends-congress-require-data-broker-industry-be-more?utm_source=govdelivery

accelerated the discovery of a vaccine for COVID-19 producing a big value for the society. But for making the interactions between companies and consumers generate positive effects on social welfare and competition it is necessary to coordinate regulatory and competition interventions. Companies collecting consumer data should create more trustworthy relationship with their customers carefully respecting privacy protection rules. Firms should actively engage in competition on service and product quality adopting a clearer approach in customer's transactions favoring transparent data sharing while increasing consumer's trust.

68. Survey results show that consumers' perceptions and concerns about data sharing are rather heterogeneous across countries, thus potentially affected by data privacy policies in individual markets. This would suggest that for example European consumers would be relatively less worried than Latin American consumers thanks to the GDPR regulation. The same is likely to apply to Australian consumers protected by the CDR regulation expected to improve consumer convenience by allowing safe data sharing with trusted recipients. Both regulations focus on the collection, access and usage of personal data thus being easier to implement compared to the laws on data ownership.

69. Summing up, it seems that a prompt policy response to the new data driven challenges should be focused on the definition of clear rules favoring data sharing to foster competition and innovation while ensuring consumers more choices and protection about how their data are used.

4. Data assets and productivity

70. This section examines factors and policy settings affecting the potential for investments in data capital to boost productivity growth.

4.1. Increasing Returns: Macro vs micro considerations

71. The innovative potential of investments in data rests in their ability to yield competitive returns to owners while also generating pecuniary externalities ("spillovers") elsewhere in an economy. Spillovers are produced when a technology or business idea is deployed in multiple firms in an economy, e.g., a blueprint or original software tool (Romer 1990, Jones 2002, 2005). This mechanism characterizes the diffusion of innovations embodied in intangible assets across firms rather than the existence of increasing returns at the firm level. Indeed, owing to their nonrival property, intangibles are only partially appropriable by their owners/creators but also provide an extra "kick" to productivity growth when replicated at low cost for use elsewhere in an economy.

72. At the micro level, data is assumed to have diminishing returns. For instance, as Varian (2018) points out, there are diminishing returns to more and more training data fed to AI algorithms. Accordingly, in their aggregate model of data in an economy, Jones and Tonetti (2020) assume that data is a productive intermediate input with diminishing returns, not a "technology" that leads to increasing returns. In the intangible capital model set out in section 2, data are productive long-lived assets whose value stems in part from the application of data technologies. There are obvious differences between these approaches (e.g., data as an intermediate vs data and data technologies as capital). This is because the stylized Jones and Tonetti model is designed to highlight the macro impacts of data sharing, while the intangible capital framework is designed to better represent data value creation via business investment. Ultimately, the data/intangible capital approach of section 2 combined with the existence of productivity spillovers is a close representation of the processes theorized by Jones and Tonetti in that (a) data assets have diminishing returns in production but (b) returns to data asset ownership may spill over to other firms to the extent they are shared within an industry or economy.

73. However, there are also many possible sources of scale economies to data at the firm and industry level. So, in innovative data-intensive firms, diminishing returns to data assets can co-exist with rising

market power due to scope economies and local scale effects.[19] R&D, managerial expertise, and skilled labor (e.g. data scientists) are costly inputs, and companies may realize local economies when such costs are spread over multiple product lines. Firms may also amass market power due to agglomeration effects that weaken the law of diminishing returns, e.g. by recombining data for different uses.[20] And sometimes an industry's scope of operations expands due to outside developments that create external economies of scale, e.g. network externalities enjoyed by social media and other sharing platforms.

74. The aggregate effect of the increased use of data reflects, then, a mix of innovation-related mechanisms that pull in opposite directions. Costless diffusion confers benefits to economies as new products and new business models diffuse throughout the economy. On the other hand, innovating firms may accumulate market power when competing firms are unable to replicate innovations at low cost, possibly due to IP rights or network effects, while at the same the innovations are being "scaled" within the innovating company.

75. We have already indicated the presence of scale effects due to data/intangible capital at the firm level (subsection 2.5). To assess how they are playing out in the aggregate versus the growth-promoting aspects of digital transformation, the available evidence on productivity spillovers and factors affecting the course of labor productivity growth are examined.

4.2. Productivity spillovers to intangible capital: The evidence

76. Like the productivity spillovers to R&D established by Griliches in a series of studies (e.g., Griliches 1992), cross-country evidence suggests that there are productivity spillovers to other intangible assets as well. This is illustrated in figure 8 below, which shows rates of growth in total factor productivity and intangible capital services for the market sector of 10 European economies and the United States before and after the global financial crisis.

77. The figure points to a proportional relationship between productivity change and growth in intangible capital services, consistent with the spillover relationship reported in an econometric analysis that controlled for the endogeneity of inputs in cross-country data (Corrado, Haskel, and Jona-Lasinio, 2017a). The results of that study conform with the simple linear regression shown as the red dotted line in the figure. The econometric study used data prior to the Great Recession, but the correlation between productivity and services from intangibles appears to persist into the recession's aftermath (i.e., to 2016), a correlation that deserves further scrutiny given that the data-intensity of economies grew rapidly since then.[21]

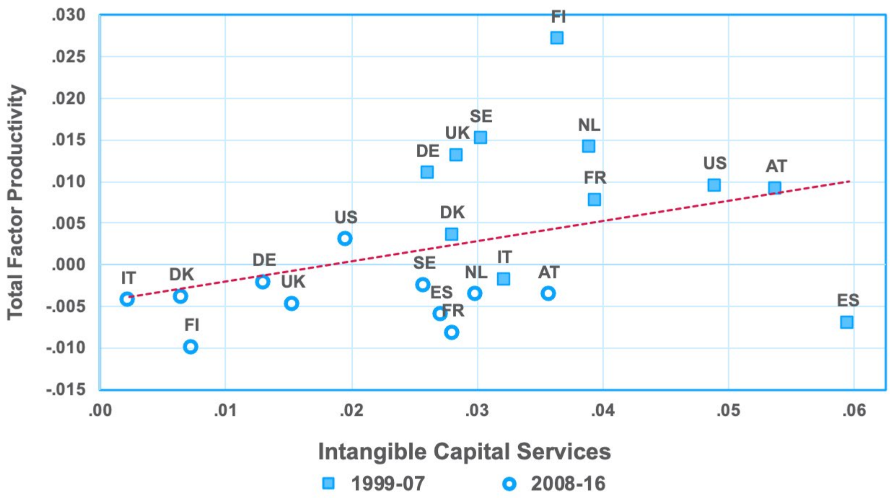
[19] Unlike economies of scale, where unit costs fall as the volume of production rises, economies of scope are efficiencies that arise from variety, not volume, creating a situation where a company's average cost of production falls with product diversification. Economies of scope are often characterized by local cost complementarities among factors of production as well as the existence of fixed costs, especially in large enterprises (e.g., supply-chains, promotion activities).

[20] As used here, agglomeration effects refer to the fact that proprietary data assets of one type may be combined with another type to generate whole new uses or solutions, and to the extent this occurs within a single firm, it weakens the effect of diminishing returns to data.

[21] In fact, the dots plotted for the United States in figure 8 lie closely along the fitted line, suggesting that fewer spillovers to intangible capital services almost fully account for the post-2007 drop in U.S. productivity growth.

Figure 8. Productivity spillovers to intangible capital services (Market sector industries only)

Log differences



Source: Corrado, Haskel, Iommi, and Jona-Lasinio (2021). Country abbreviations are as follows: Austria (AT), Germany (DE), Denmark (DK), Spain (ES), Finland (FI), France (FR), Italy (IT), Netherlands (NL), Sweden (SE), United Kingdom (UK), United States (US).

78. In comparison, the evidence of productivity spillovers to IT capital has been more elusive (e.g., Stiroh 2002), though most research was based on earlier generations of IT and pertains to equipment; there is evidence of network effects associated with communications, however. Regarding software, the underlying econometric evidence for the spillover relationship for intangibles displayed in figure 8 does *not* derive from the significance of spillovers to software investments; in fact, when software is excluded, the estimated spillover relationship is stronger.

79. IT has important complementarities with intangibles, however. Using intangibles-extended productivity datasets, a cross-country analysis demonstrates complementarities between IT (defined to include software) and intangibles (excluding software and R&D; see Corrado, Haskel, and Jona-Lasinio 2016). These findings are consistent with previously established evidence of complementarities between intangible and IT capital based on firm-level data (Bresnahan, Brynjolfsson, and Hitt 2002).

80. Figure 9, Panel A, displays the main findings regarding the indirect effects of intangibles on productivity in the literature as described above, i.e., that non-R&D intangibles and IT (defined to include software) are complements in production, as are product R&D and marketing (e.g., Vinod and Rao 2000), and that R&D and non-R&D intangible capital (excluding software) each have productivity spillovers.

81. What do these results imply for spillovers to investments in data assets? First, though section 2 used an intangible asset approach to conceptualize how to analyse data assets, it also pointed out that intangible assets are themselves becoming more data intensive. This suggests that the spillover and complementary linkages displayed in panel A pertain to a less data-intensive period that the present and recent past, and we can only conjecture how to project them in terms of newer data assets and data technologies.

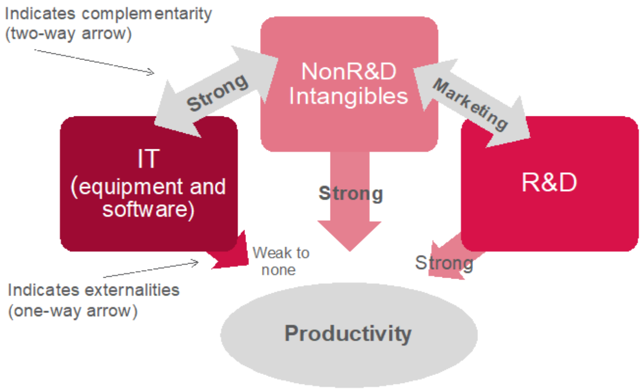
82. Panel B sets out a mapping for thinking about these linkages for assessments of the indirect effects of data assets on productivity going forward. The data value chain inherent in our framework is based on complementarities among data technologies and data assets. Implicit in this framework are complementarities with the digital infrastructure that makes the application of modern data technologies and storage of data possible, i.e., prevalence of cloud services, etc. Panel B depicts the likely potential for

these complementarities across all framework elements (depicted by the three boxes), but indicates uncertainty regarding productivity spillovers.

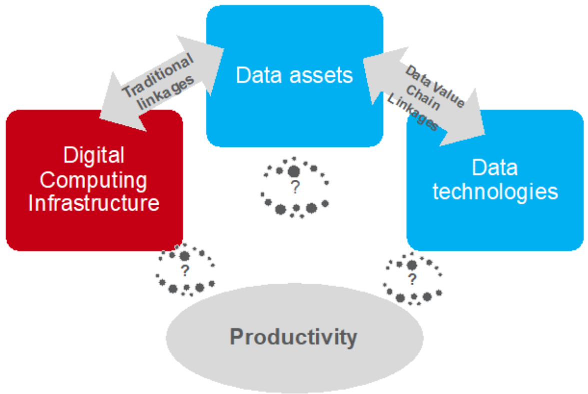
83. While complementary relationships are inherent in the data/intangible framework (and its relationship with IT), arguments for expecting productivity spillovers to investments in data assets are more difficult to make and, ultimately, the matter is empirical. It seems reasonable, however, that data intelligence investments generate productivity spillovers. But to the same extent that copying a business model in the past meant restructuring an organization and/or training employees, copying a digitally enabled business model also requires developing the requisite data stores and databases to power it. This may present a significant hurdle for firms, even though the hurdle should be weighed against the emergence of B2B markets in data. On the other hand, owing to the growing open-source content of many AI tools, it seems reasonable to consider the emergence of new spillovers to investments in software whose investments did not exhibit spillovers in earlier macro productivity data.[22]

Figure 9. Indirect impacts of intangibles on productivity

A. The evidence to date



B. …Extended to Data



[22] Open-source software also potentially contributes directly to productivity, however. See Murciano-Goroff, Raviv, Ran Zhuo and Shane Greenstein (2021).

Note: Indirect impacts refers to contributions of measured productivity after attributing the direct (i.e., appropriable) impacts of intangibles via growth accounting. Panel A depicts the existing literature, i.e., that non-R&D intangibles and IT are complements in production, as are product R&D and marketing, and that R&D and non-R&D intangible capital (excluding software) have productivity spillovers. Network effects are excluded.

4.3. Data, AI, and the Modern Productivity Paradox

84. The idea that information technology and artificial intelligence has or will cross some boundary after which economic growth again rises rapidly is advanced by technologists and some economists, e.g., Brynjolfsson and McAfee (2014). Productivity growth has slowed, however, while AI technology and data use has been accelerating.

85. Some view this productivity paradox as a measurement problem, an opinion expressed for instance by Hal Varian, Chief Economist at Google, who opined there was “a lack of appreciation of what’s happening in Silicon Valley because we don’t have a good way to measure [the digital innovations coming out of] it” (Aeppel 2015). On the other hand, Gordon (2016) argues that productivity growth since 2004 is merely a reversal of a one-off aberration from 1995 to 2004, whereas Syverson (2013) likens recent experience to prior diffusions of general-purpose technologies, e.g., electrification and the internal combustion engine, which came in multiple waves over decades and from which he concludes that another productivity acceleration due to information technology is not out of sight.

86. Timing of effects—short-run dynamics vs longer run effects—are then subjects for further examination using the data capital framework. Selected measurement issues are also discussed.

4.3.1. Short-run productivity dynamics

87. A framework for assessing the short-run dynamics of the impact of AI on productivity growth as a “missing investment” stream is set out in Brynjolfsson, Rock, and Syverson (2021). Their framework addresses missing investment issues much as the data/intangible asset framework outlined in Section 2 does, but the emphasis is on the dynamics of AI alone. Their main finding is that there is a “J-curve” pattern to the impacts of AI with the height of the impact of the technology on *total factor productivity* growth not coming until a decade or so after introduction.

88. Corrado, Haskel, and Jona-Lasinio (2021) discussed the effects of AI and missing investment using an intangible capital approach. They treated intangibles assets as a “missing” investment stream in productivity calculations and found that, though there is plenty of missing investment, the upward “swoosh” of the J-curve pattern, which reflects the salutary effects of returns to missing investments, is not there.

89. Perhaps the differences in these findings can be explained by the service lives used to estimate depreciation rates for intangible assets, which tend be short in the latter study but were rather longer in the first. Perhaps, too, the real quantities of investments in relevant categories are misstated. The impacts calibrated in these studies were based largely on extrapolating existing relative prices, e.g., non-AI software prices for modern software tools and traditional business services prices for investments in business intelligence.

4.3.2. Long-term growth contribution of data capital

90. Though we are still in the early stages of pinning down data capital in macroeconomic statistics, the potential contribution of data capital to *long-term* labor productivity growth can be calibrated with a minimum of very explicit assumptions.

91. The long-term growth-promoting potential of data-driven inputs depends on the extent to which their volume rises more rapidly than their relative price falls (i.e., that the input shares continue to rise).

This is ultimately an empirical question about the degree of substitutability between data/AI and human efforts, and there are limits to this substitutability as discussed in Nordhaus (2021).

92. Apart from spillovers (discussed above), which are difficult to model and quantify, the impacts of data capital can be calibrated using estimates of two effects, a "use" effect determined by input cost shares and relative prices, and a "production" effect determined by production shares and relative prices. This approach has been a mainstay of productivity analysis with IT capital.

93. As business services derived from IT equipment have shifted to the cloud, domestic production effects (via IT services production) have become more pronounced in calibrations of IT impacts on an economy (Byrne and Corrado 2017). When thinking about data capital, then, production effects are also likely to loom large because much production of data capital occurs within using firms.[23]

94. This approach can be applied to data capital to calibrate ranges for thinking about its longer-term contribution to *labor productivity* growth. The results of this exercise are reported in table 4, which shows alternative scenarios for the productivity-enhancing impact of data capital.[24] The scenarios vary according to the breadth of investments in data capital in an economy (broad or limited diffusion), the extent to which data assets are domestically produced, and the productivity advantage of data assets and data technologies (based on their relative price). The capital income share of data capital is used to measure diffusion via use and is assumed to be less than the corresponding total intangible capital income; the ranges used in the table are based on actual shares in high vs low intangibles-intensive countries. The production share is assumed to be the capital income share +/- 10 percent, roughly the range for net exports of corresponding intangible investment services in high vs low intangibles-intensive countries (Corrado et al, 2022).

[23] The available estimates suggest that about one-half of intangibles are produced for use with the same organization in the United States (Corrado 2021).

[24] The calculations are based on the steady-state solution to a two-sector model that consists of a data capital producing sector and all other goods and services. In this simple model, the contribution of the data sector to labor productivity equals the sum of the use effect, ${\bar{v}^D}/{\bar{v}^L}\ (-\dot{p})$ plus the production effect, $\bar{\omega}^D(-\dot{p})$, where $\bar{v}^D$ and $\bar{v}^L$ are the income shares of data capital and labor, $\bar{\omega}^D$ is the production share of data investments, and $(-\dot{p})$ is the relative productivity of data measured as the rate of decline in the relative price of data assets (sign reversed). The calculations in the table assume labor's share of total income, $\bar{v}^L$, equals .7. For a derivation, see Oulton (2012) or Byrne and Corrado (2017).

Table 4. Productivity Scenarios: Contribution of data capital to annual labor productivity growth (percentage points)

	Productivity advantage (relative asset price growth differential)	
	Narrow edge 1 percentage point	Large edge 5 percentage points
Broad use (and net exporter of data services) 10 percent capital income share 11 percent production share	0.25	1.26
Limited use (and net importer of data services) 5 percent capital income share 4.5 percent production share	0.12	0.58

Note: Contributions include the sum of the use and production effects of data capital.

95. The lower bound for the productivity advantage is drawn from recent evidence on the relative price differential implied by an intangible investment price index designed to capture the impacts of digitization on investments in brand and the IT consulting and marketing subcomponents of organizational capital, about 1 percentage point per year (Corrado 2021). This deflator takes national accounts asset prices for R&D and software as given, however, and thus is a lower bound in that these deflators do not incorporate efficiency gains due to increased application of AI methods or use of open-source content. The upper bound is guided by the long-term relative price decline of conventionally defined IT capital of about 15 percent per year (based on the estimates reported in Byrne and Corrado, 2017). It is conservatively set at one third of that, i.e., 5 percentage points per year.

96. All told, estimates of the contribution of data capital to labor productivity growth range by more than a factor of 10—from 0.12 percentage point per year to 1.26 percentage point per year. The range highlights the synergies among data capital efficiency and an economy's breadth of use and capability for digital transformation, which implies much scope for policies to affect outcomes. Promoting diffusion through the use effect (i.e., encouraging both data investments and data sharing) is very important, and a typical focus of traditional IT policies.

97. The table further implies that the course of data productivity is "doubly" important, operating as it does through both the use and production channels given that much data value creation occurs within firms and international trade in data assets and data asset services remains limited. This calls for policies that foster innovation in data technologies, the creation of new scope economies within firms (more data-driven business functions), and, to the extent possible, the development of well-functioning markets for data assets and data asset services.

4.4. Digitization, consumer prices, and productivity

98. The previous subsection underscored that accurate data asset prices are needed to estimate the contribution of data capital to growth. Measures of other prices were taken as given. Our assessment of price mismeasurement more generally concludes that—save for digital services prices, discussed below—there is little evidence that persistent sources of biases, such as biases due to new goods and increased

varieties have changed markedly in the recent, data-driven digital era (Moulton 2018; Aghion, Bergeaud, Boppart, Klenow, and Li 2019).[25]

4.4.1. Consumer digital services prices

99. Research on consumer digital services in the United Kingdom (Coyle et al., 2019) and the United States (Byrne and Corrado, 2020, 2021) suggests that the dynamic nature of these markets is not fully picked up in official prices to the extent that sampled prices are contract prices not based on actual usage, e.g., the price per month for a streaming service is not typically tied to hours per month spent streaming. Contract prices in most countries have trended up slightly since the early 2000s—even those that are quality-adjusted—whereas usage rates have risen dramatically. How important is this services usage bias? And how long might it be expected to persist? The presumption is that the bias may be large and long-lived: indeed, we have been in its midst for nearly 20 years.

100. Payments for consumer digital access services—internet, mobile phone, cable TV, and streaming—accounted for 2-1/4 percent of U.S. household consumption in 2018, having grown sharply between 2000 to 2015 from an initial base of less than ½ percent at the dawn of the internet era. A price index constructed for these services using direct measures of volume (data transmitted, talk time, and hours of programming) fell 12 percent per year from 1988 to 2018, while official prices moved up modestly. Using this price index for consumer digital services, total personal consumption expenditure (PCE) prices are estimated to have risen nearly 1/2 percentage point slower than the official index since 2008 in the United States, and the spread between this alternative and official PCE price inflation increased noticeably with time.

101. This bias in output growth is likely to be smaller for countries that are less technologically advanced and with less access to connectivity and less widespread use of the Internet (and vice versa).

4.4.2. "Free" goods

102. The impact of "free" consumer content delivered digitally is not included in the benefits implied by the alternative price indexes for consumption described above. Estimates of consumer surplus stemming from innovations in digital content delivery suggest benefits to consumers have been large. Byrne and Corrado (2021) estimated that U.S. consumers enjoyed a cumulative gain in economic welfare to the tune of nearly $25,000 *per user* from 2004 to 2016 from digital services, including the consumption of free content. Their estimates are comparable to figures derived by Brynjolfsson, Collis, and Eggers (2019) based on massive online choice experiments (see their supplementary information table S8, sum of median WTA values for 2016). Massive online choice experiments are a "stated preference" approach to measurement that could be applied to valuing data as discussed in our companion paper.

4.5. Summing Up

103. The microeconomic literature on data looks at dimensions along which digitized information lowers costs and reduces information asymmetries. This suggests that the increased supply and use of data

[25] Furthermore, too little is known about recent "sourcing bias," especially in the wake of disruptions to physical economic activity caused by the COVID 19 pandemic. Sourcing bias occurs when consumers shift to lower priced outlets (e.g., retail online platforms) or buyers shift to low-cost suppliers (e.g., Chinese suppliers) for the same quality-adjusted product; see Nakamura, Diewert, Greenlees, Nakamura and Reinsdorf (2014) for an analysis of this bias. While sourcing changes impart an upward bias to price indexes, imports are a subtraction from GDP, so the net effect of sourcing bias on GDP prices cannot be reliably signed. All told, while we do not believe this bias affects the labor productivity scenarios presented in the previous subsection, it will affect measures of household income deflated by consumer prices.

potentially benefits the aggregate economy via productivity growth, but productivity growth slowed in most advanced economies prior to the onset of the global pandemic.

104. It is possible that productivity growth slowed because the spillover effects of data-intensive intangibles are weaker than the assets of a decade or so ago, e.g., data-enabled or digital platform business models may be costly to replicate if the assets that drive them are more like proprietary trade secrets than freely available engineering solutions or blueprints. This notion runs counter to claims of substantial benefits from AI and data-driven business models, suggesting that policies to encourage the development of business models based on data that are freely available (e.g., public open data) or that may be generated from machine-to-machine operations (e.g., server requests from Kayak's travel search engine to counterpart servers at airlines) will raise the odds that economies realize productivity spillovers from investments in data capital.

105. Faster productivity growth due to more data-driven organizations may be yet to come. But decision makers need to realize that a wide range of outcomes are possible, depending on the extent to which data technologies are adopted and the significance of the gains/innovations they engender. Based on available research, factors determining these synergies were provided in terms of ranges, and prospects for the contribution of data capital to growth in labor productivity were estimated to vary widely—from just a tad more than .1 percentage point to a boost of 1-1/4 percentage point per year. This suggests there is both much room, and much at stake, for policies designed to set economies on a productive path.

106. These estimates, though based on clear assumptions, are not meant to be exact in terms of framing prospective outcomes for labor productivity growth, nor by extension, growth in household real incomes (abstracting from distribution issues). Recall also that the estimates in table 4 did not factor in the benefits that households derive from their consumption of digital content obtained via subscription services; these benefits are understated in official statistics and have been estimated to have already provided substantial boosts to consumption (and GDP/productivity) growth in the United Kingdom and the United States during the past two decades.

5. Directions for future research and conclusions

107. This paper began by noting that the impact of data on economic activity has multiple dimensions that raise new challenges for decision makers. Its primary objective was to highlight some of the economic and policy implications of treating data as an asset. The paper considered policy settings affecting consumer privacy and long-term productivity growth due to the increased use of data capital in economies, leaving the analysis of some key macroeconomic mechanisms that also may be affected by data capital, e.g., changes to the dynamics of product prices and factor input adjustments (see Box), for future work.

108. The analysis in this paper used an intangible capital framework to consider economic activity around the creation of value as raw data is processed and developed into marketing insights and business solutions using modern digital technologies (the "data stack"). The paper further argued that most expenditures related to data capital formation—on data assets and their monetization, and on data technologies and their development via R&D—are included in the measures of intangible investment available via the INTAN-Invest database.

109. The salient conclusion of the analysis considered in the paper is that data capital raises the complexity of the policy trade-off between promoting innovation and maintaining competitive markets. Data adds complexity to the extent that competition policies must also act as guardians of personal information on consumers held by business. And because data-intensive industries tend to produce uneven outcomes due, at least in part, to the ease with which some firms use data technologies to accumulate data assets, competition policies also need to work to ensure a more even distribution of outcomes in these industries. We pointed to at least three main reasons behind these added complexities of data:

- Data is different from other intangible assets in that it has a consumer privacy dimension.
- Data is nonrival and, like other intangible assets, capable of improving economic welfare via sharing, either as a within industry or general-purpose commons.
- Data, though nonrival, is frequently used exclusively. Business owners may treat data assets as trade secrets. Data privacy laws often mandate exclusivity.

110. We also argued that data markets are asset markets and inherently dynamic, and procedures for evaluating fairness applied to day-to-day consumer/business “trades” can easily be confounded by false equivalences (assets vs services flows) and network effects. The dynamic nature of data asset markets also suggests that protecting personal information by restricting future re-use is a restraint on innovation.

111. The paper finds the stakes to be very high for decision makers to smoothly navigate the policy trade-offs. Surveys reveal widespread concern among consumers (globally) regarding business use of their personal information, and inattention to these concerns seems impossible. The potential for data capital to contribute to growth is substantial and highly dependent on factors influenced by policy settings, e.g., policies that build digital skills and capabilities, as well as the factors affecting the inherent capacity for data to be shared.

112. At the very least policy makers need timely assessments of data capital penetration, data capital sharing, and data capital contributions to industry output growth to monitor the balances that policies must strike. The costs of data/intangible capital also are an ongoing concern in the analysis of data-intensive industries for competition policies. The framework set out in the paper is designed to facilitate these assessments, though for those assessments to be effective, timely and accessible measurements of data capital are required.

Box 1. Data and the macroeconomy

Data has the potential for altering the responsiveness of product prices to economic conditions, and the rise of data/intangible capital as a strategic factor input has the potential for altering the cyclical pattern of investment and, consequently, factor input demand in the short run. These are important subjects for future research.

This box addresses these topics from the perspective of empirical macro puzzles surrounding the behavior of the aggregate price mechanism, e.g., the apparent breakdown of the relationship between unemployment and inflation since the 1990s (the Phillips curve), the suppression of wage growth across many countries for many decades, and persistent low inflation following the global financial crisis despite a strengthening in aggregate demand (until the COVID crisis). Research has demonstrated that reductions in product and labor market regulations and spread of global supply chains contributed to the disinflation (e.g., Andrews, Gal, and Witheridge 2018, who provide a review). Here we consider that the rise of data capital might be another factor that has been affecting the aggregate price mechanism.

Price flexibility in high-frequency microdata

The evidence is mixed on whether prices have become broadly more flexible in the digital age. Nakamura, Steinsson, Sun and Villar (2018) examined micro-CPI data in the United States from 1975 to 2015 and concluded that the frequency of (nonsale) price change had not materially changed over the 40-year period they studied. The prices they studied were primarily from brick-and-mortar retail outlets and results highly sensitive to the treatment of temporary price discounts (sales).

The frequency of sales discounts and the e-commerce segment of markets have trended up along with the increased use of data capital, however, and research on the duration of price rigidity remains an active field of inquiry. A recent microdata study sheds some light on frequency of adjustment of online sales prices. Gorodnichenko and Talavera (2017) constructed a massive dataset on online price quotes for tradable consumer durable goods between the United States and Canada covering nearly five years (November 2008 to September 2013). Based on U.S. goods comparable to those studied in Nakamura and Steinsson (2018), mainly consumer electronic equipment, they find that online prices are significantly more flexible than prices in regular retail outlets.[26] The products compared are less than 5 percent of market-based goods spending by households in the United States, however.

Regarding online tradable goods prices, Gorodnichenko and Talavera (2017) find that prices (a) display equal flexibility in terms of increases and decreases while (b) generally exhibiting larger responses in Canadian markets than those in the United States, consistent with the view that price adjustment is likely to be larger in smaller markets. The sensitivity of prices to changes in the nominal exchange rate is found to be systematically correlated with the characteristics of goods and their markets (e.g., the degree of competition), suggesting that industry concentration creates frictions that partially offset the increased responsiveness of pricing to changes in market conditions induced by digital transformation.

Price flexibility and aggregate supply

To the extent that business is more data driven and economies are more intangible-intensive, businesses also are less likely to be capacity constrained, which, if true, suggests that economies will have lower inflation and flatter aggregate supply curves. A recent IMF study (Lail and Zang 2020) examined the flattening of the short-run aggregate supply relationship for 20 countries and found evidence for structural breaks between inflation and the output gap for most of them. They controlled for globalization and found that the intangible intensity of an economy—a proxy for data capital intensity—contributed to explaining and quantifying the degree of the flattening.

The same issue is considered through the lens of a micro-founded new Keynesian Phillips curve framework by Haskel (2019). The slope of the Phillips Curve (e.g., as in Walsh, 2017, p. 310) is the sum of the inverse of the wage elasticity of labor supply and the coefficient of relative risk aversion times a term that is increasing in the fraction of firms who can adjust prices each period. If more firms adjust prices, then the slope of the Phillips Curve should rise, which seems inconsistent with the finding that the aggregate supply curve is now flatter. This may be reconciled, however, by noting that in this literature, there is typically only one input, namely labor, and if a firm wishes to expand output it must hire more labor, which will raise wages and prices. In an intangibles-intensive economy, however, firms may find that they can also raise output by using data capital more intensively (or by exploiting its low-cost replicability in new ways), in which case the effects of output variation on wages and prices will become muted.

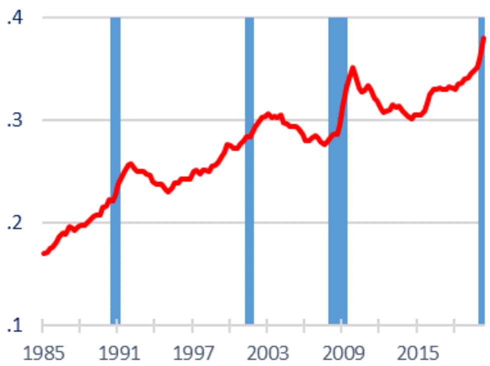
Intangible investment and aggregate demand

Though data as intangible capital (as a short-run substitute for labor) might contribute to a flattening of the aggregate supply curve through restraining prices, there are equally plausible arguments that certain intangibles are more costly to adjust than tangible investments, suggesting the responsiveness of aggregate demand to changes in interest rates may also be part of the story.

Recent studies look at the effectiveness of monetary policy rate-setting tools vis a vis intangible investment. One study used firm-level data to look at firm's investment response to changes in short-term Treasury rates and found that intangible investment reacted much less than did tangible investment (Döttling and Ratnovski 2020). The authors put forth evidence suggesting two mechanisms that might explain their finding: a credit channel (intangibles are less reliant on secured debt financing) and a depreciation channel (intangible assets have higher depreciation rates than tangible assets, and the same interest rate changes change their user cost proportionately less). They also considered an adjustment cost channel (intangible investment is costlier to scale up or down compared with tangible investment) but could not support (or rule out) this mechanism as an explanation for the disparate behavior of intangibles vs tangibles in response to changes in interest rates.

Figure 10. Intellectual property products investment in the United States, 1985Q1- 2020Q2

Share of Gross Private Nonresidential Investment



Source: Elaboration of BEA data; shaded areas are periods of business recession as defined by the NBER.

[26] Specifically, they find that the size of price changes in online stores is less than one-half the size of price changes in regular stores (approximately 4 versus 10 percent) and that price changes occur much more frequently in online stores (approximately once every three weeks or less versus only once every four to five months or more in regular stores).

The figure to the right displays the fluctuations in the intellectual property products (IPP) share of nonresidential investment in a cyclical context using quarterly data from the U.S. national accounts; IPP includes software, R&D, and entertainment originals. As may be seen, these investments are the last category of capital spending cut during downturns. This suggests that businesses may view the acquisition of software (and other intangibles) as moves to increase efficiency and dampen the net revenue impact of cutbacks in customer demand.

This interpretation is in line with the reasoning that data capital might contribute to a flattening of the aggregate supply curve in response to positive demand shocks, there are plausible arguments that certain intangibles are more costly to adjust during a downturn, suggesting possible (and difficult-to-detect) asymmetries in the relationship between intangible investment and aggregate demand. Investments often are forward-looking, multi-period decisions that are not easily reversible, e.g., R&D experiments require sequential outlays and are rarely shut-down mid-stream. Furthermore, the creation of data driven business intelligence tends to be done internally within firms using highly skilled human capital as a key production factor. Hiring and firing scarce talent is difficult, costly, and risky. Indeed, Eisfeldt and Papanikolaou (2013) document that firms with more organizational capital are more likely to list the loss of talent as a risk in their annual reports, suggesting differential adjustments to purchased data investments versus those that are produced and consumed within the same firm.

<u>Summary</u>

Data capital has impacts on both aggregate supply and aggregate demand, and impacts may be asymmetric. Monetary nonneutrality depends on aggregate price rigidity, and if data and data feedback systems are leading to dynamic price adjustment (i.e., greater price flexibility in product markets) then the greater data-intensity of intangible capital might be contributing to a flattening of the short-run aggregate supply curve, either directly through affecting price flexibility or indirectly through altering the composition of aggregate investment and cyclical patterns of factor input adjustments. Diminished competition intensity, whether due to the rise of data capital or not, are an offset to these effects.

References

Abrahamson, Zachary (2014), *"Essential Data" Yale Law Journal* 124. Available at: https://digitalcommons.law.yale.edu/ylj/vol124/iss3/7

Accessnow, (2020), "Two Years Under the EU GDPR: An Implementation progress report." Available at: https://www.accessnow.org/cms/assets/uploads/2020/05/Two-Years-Under-GDPR.pdf

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar (forthcoming), "Too Much Data: Prices and Inefficiencies in Data Markets," *American Economic Journal: Microeconomics*. Preview available at: https://www.aeaweb.org/content/file?id=13871

Aeppel, Timothy (2015), "Silicon Valley Doesn't Believe U.S. Productivity Is Down," *Wall Street Journal*, July 16.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li (2019), "Missing Growth from Creative Destruction", *American Economic Review*, 109 (8): 2795-2822.

Akcigit, Ufuk and Sina T. Ates (2019), "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory", *NBER Working Paper Series* No. 25755 (April).

Andrews, Dan, Peter Gal and William Witheridge (2018), "A Genie in a Bottle? Globalisation, Competition and Inflatio", *Economics Department Working Paper* No 1462. Available here.

Andrews, Dan, Chiara Criscuolo, and Peter Gal (2016), "The Global Productivity Slowdown, Technology Divergence and Public Policy: A Firm Level Perspective", *OECD background paper*. Available here.

Arrieta-Ibarra, Imanol, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier, and E. Glen Weyl (2018), "Should We Treat Data as Labor? Moving beyond 'Free'", *AEA Papers and Proceedings*, 108: 38-42.

Autor, David, Claudia Goldin and Lawrence F. Katz (2020), "Extending the Race between Education and Technology," *AEA Papers and Proceedings* 100: 347-351.

Babina, Tania, Anastassia Fedyk, Alex He, and James Hodson (2021), "Artificial Intelligence, Firm Growth, and Industry Concentration". Available at SSRN: https://ssrn.com/abstract=3651052 or http://dx.doi.org/10.2139/ssrn.3651052

Bajgar, Matej & Giuseppe Berlingieri & Sara Calligaris & Chiara Criscuolo & Jonathan Timmis (2019), "Industry Concentration in Europe and North America," *OECD Productivity Working Papers* No.18, OECD Publishing.

Bartelsman, Eric, Stefano Scarpetta and Fabiano Schivardi (2003), "Comparative Analysis of Firm Demographics and Survival: Micro-Level Evidence for the OECD Countries", *OECD Economics Department Working Papers* No.348, OECD Publishing.

Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta (2013), "Cross-Country Differences in Productivity: The Role of Allocation and Selection', *American Economic Review*, American Economic Association, vol. 103(1), pages 305-334, February.

Battiati, C., C. Jona-Lasinio, E. Marvasi and S. Sopranzetti (2021), "Market power and productivity trends in the European economies: A macroeconomic perspective", *Luiss School of European Political Economy Working Paper* 1/2021.

Bresnahan, Timothy F., Erik Brynjolfsson and Lorin M. Hitt (2002), “Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence”, *Quarterly Journal of Economics* 117(1): 339-396.

Brynjolfsson, Erik, Avinash Collis and Felix Eggers (2019), “Using Massive Online Choice Experiments to Measure Changes in Well-being”, *Proceedings of the National Academy of Sciences* 116(15): 7250-7255.

Brynjolfsson, Erik and Lorin M. Hitt (2000), “Beyond Computation: Information Technology, Organizational Transformation and Business Performance”, *Journal of Economic Perspectives* 14 (4): 23–48.

Brynjolfsson, Erik., Lorin M. Hitt, and Shinkyu Yang (2002), “Intangible Assets: Computers and Organizational Capital”, *Brookings Papers on Economic Activity*: Macroeconomics (1): 137-199.

Brynjolfsson, Erik, and Andrew McAfee (2014), *The Second Machine Age—Work, Progress, and Prosperity in a Time of Brilliant Technologies*, New York: W.W. Norton and Company.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson (2021), "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies", *American Economic Journal: Macroeconomics*, 13 (1): 333-72.

Byrne, David and Carol Corrado (2017), “ICT Services and their Prices: What do they tell us about Productivity and Technology?”, *International Productivity Monitor*, 33 (Fall), 150-181.

Byrne, David and Carol Corrado (2020), “The Increasing Deflationary Influence of Consumer Digital Access Services”, *Economic Letters* 196 (November), 109447.

Byrne and Corrado (2021), “Accounting for Innovations in Consumer Digital Services: IT still matters”, In Corrado, Haskel, Miranda, and Sichel, eds., *Measuring and Accounting for Innovations in the 21st Century*, pp 471-517, *NBER Studies in Income and Wealth Volume 78, Chicago: University of Chicago Press.*

Cabral, Luís M B (2020), "Merger Policy in Digital Industries”, *CEPR Discussion Papers* 14785, C.E.P.R. Discussion Papers.

Charbonneau, Karyne, Alexa Evans, Subrata Sarker, and Lena Suchanek (2017), “Digitalization and Inflation: A Review of the Literature”, Staff Analytical Note, 2017-20. Ottawa: Bank of Canada.

Corrado, Carol, Jonathan Haskel, Massimiliano Iommi, Cecilia Jona-Lasinio and Filippo Bontadini (2022, “Data, Intangible Capital, and Productivity”, paper presented at the CRIW Conference on Technology, Productivity, and Economic Growth (March)

Corrado, Carol, Chiara Criscuolo, Jonathan Haskel, Alexander Himbert, Cecilia Jona-Lasinio, (2021), "New evidence on intangibles, diffusion and productivity," *OECD Science, Technology and Industry Working Papers* 2021/10, OECD Publishing.

Corrado, Carol, Jonathan Haskel, and Cecilia Jona-Lasinio (2016), “Intangibles, ICT and Industry Productivity Growth: Evidence from the EU.” In D. Jorgenson, K. Fukao, and M. Timmer, eds., *The World Economy: Growth or Stagnation?* Cambridge, U.K.: Cambridge University Press.

Corrado, Carol, Jonathan Haskel, and Cecilia Jona-Lasinio (2017a), “Knowledge Spillovers, ICT, and Productivity Growth,” *Oxford Bulletin of Economics and Statistics* 79:4 (August), 592-618.

Corrado, Carol, Charles R. Hulten, and Daniel Sichel (2005), “Measuring Capital and Technology: An Expanded Framework,” in C. Corrado, J. Haltiwanger, and D. Sichel (eds.), *Measuring Capital in the New Economy*, Volume 66, Studies in Income and Wealth, 11–46. Chicago: University of Chicago Press.

Corrado, Carol, Charles R. Hulten, and Daniel Sichel (2009), “Intangible Capital and U.S. Economic Growth,” *Review of Income and Wealth* 55 (3): 661–685.

Corrado, Carol and Charles Hulten (2010), “How Do You Measure a ‘Technological Revolution’?” *American Economic Review*: 100 (2): 99-104.

Corrado C., J. Haskel, C. Jona-Lasinio, M.Iommi (2019), ”Intangible Capital, Innovation, and Productivity à la Jorgenson", in Fraumeni, Barbara M. (ed.) Measuring Economic Growth and Productivity: Foundations, KLEMS Production Models, and Extensions, Academic Press, Cambridge, MA. ISBN: 9780128175965, 15 November 2019.

Corrado, Carol, David Martin, and Qianfan Wu (2020), "Innovation α: What do IP-Intensive Stock Price Indexes tell us about Innovation?", *AEP Papers and Proceedings* 110: 31-35.

Corrado, Carol (2021), "Intangible Capital", Chapter in forthcoming Brookings volume (available upon request from author).

Corrado C, Haskel J and Jona-Lasinio C, (2021), "Artificial intelligence and Productivity," forthcoming in *Oxford Review of Economic Policy.*

Corrado C., J. Haskel, C. Jona-Lasinio, M.Iommi (2021), "The Value of Data in Digital-based Business Models: Measurement", Mimeo.

Coyle, Diane and David Nguyen (2018), "Cloud Computing and National Accounting," *Economic Statistics Centre of Excellence (ESCoE) Discussion Papers* ESCoE DP-2018-19.

Coyle, Diane, Mohamed Abdirahman, Richard Heys, and Will Stewart (2019), "A Comparison of Deflators for Telecommunications Services Output", *Discussion Papers*.

Crouzet, Nicolas and Janice C. Eberly (2018), "Investment, Intangibles, and Efficiency", *AEA Papers and Proceedings*, 108: 426–431

Crouzet, Nicolas and Janice C. Eberly (2019), "Understanding Weak Capital Investment: The Role of Market Concentration and Intangibles", *NBER Working Paper Series* No. 25869 (May).

Dahlhoff, Denise (2020), "Consumers' Attitudes about Data Practices", The Conference Board (October). https://conference-board.org/publications/consumers-attitudes-about-data-practices-ES.

De Loecker, Jan, Jan Eeckhout, Gabriel Unger (2020), "The Rise of Market Power and the Macroeconomic Implications", *The Quarterly Journal of Economics*, Volume 135, Issue 2, May, Pages 561–644, https://doi.org/10.1093/qje/qjz041

De Ridder, Maarten (2019), "Market Power and Innovation in the Intangible Economy", *Discussion Papers* 1907, Centre for Macroeconomics (CFM).

Döttling, Robin and Lev Ratnovski (2020), "Monetary Policy and Intangible Investment", IMF working paper WP/20/160 (August).

Eggertsson, Gauti B., Jacob A. Robbins, Ella Getz Wold (2018), "Kaldor and Piketty's Facts: The Rise of Monopoly Power in the United States", *NBER Working Papers* 24287, National Bureau of Economic Research, Inc.

Eisfeldt, Andrea and Dimetris Papanikolaou (2013), "Organization capital and the cross-section of expected returns", *Journal of Finance* 68 (4), 1365–1406.

Eisfeldt, Andrea and Dimitris Papanikolaou (2014), "The Value and Ownership of Intangible Capital", *AER: Papers and Proceedings* 104 (5), 189–194.

European Commission, International Monetary Fund, Organization for Economic Cooperation and Development, United Nations, and World Bank (2009), *System of National Accounts 2008.* New York: United Nations.

European Commission (2019), "Competition Policy for the Digital Era", EC Report.

Furman, J, D Coyle, A Fletcher, D McAules, P Marsden (2019), "Unlocking digital competition: Report of the digital competition expert panel", available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf. Report prepared for the Government of the United Kingdom, March

Gal, Peter, Nicoletti, Giuseppe, von Rüden, Christina, Sorbe, Stéphane and Renault, Théodore, (2019), "Digitalization and Productivity: In Search of the Holy Grail - Firm-level Empirical Evidence from

European Countries", *International Productivity Monitor*, 37, issue, p. 39-71, https://EconPapers.repec.org/RePEc:sls:ipmsls:v:37:y:2019:2.

Goodridge, Peter and Jonathan Haskel (2016), "Big Data in UK industries: An intangible investment approach", *Imperial College Discussion Paper* 2016/01.

Goolsbee, Austan D. and Peter J. Klenow (2018), "Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce", *NBER Working Paper* 24649 (May).

Gordon, Robert J. (2016), *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton, N.J.: Princeton University Press.

Gorodnichenko, Yuriy and Oleksandr Talavera (2017), "Price Setting in Online Markets: Basic Facts, International Comparisons, and Cross-Border Integration", *American Economic Review* 107(1): 249–282.

Graef I. (2016), EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility, Kluwer Law International. Graef I., Husovec, M. and N. Purtova (2019).

Griliches Z: (1992), The Search for R&D Spillovers, *Scandinavian Journal of Economics*, 94(Supplement): S29–47.

Gutiérrez, Germán, and Thomas Philippon (2017), "Declining Competition and Investment in the U.S", *NBER Working Paper* 23583.

Hall, Robert E. (2004), "Measuring Factor Adjustment Costs", *Quarterly Journal of Economics* 119 (3), 899–927.

Haskel, Jonathan and Stian Westlake (2018), *Capitalism without Capital: The Rise of the Intangible Economy*, Princeton, N.J.: Princeton University Press.

Haucap, Justus, Rasch, Alexander and Stiebale, Joel (2019), How mergers affect innovation: Theory and evidence, *International Journal of Industrial Organization*, 63, issue C, p. 283-325,

International Monetary Fund (2019), The Rise of Corporate Market Power and its Macroeconomic Effects. Chapter 1 in *World Economic Outlook: Growth Slowdown, Precarious Recovery.* Washington, DC, April.

Jones, Charles I. (1995), "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103, No. 4, 759-784.

Jones, Charles I. (2002), "Sources of U.S. Economic Growth in a World of Ideas," *American Economic Review*, American Economic Association, vol. 92(1), pages 220-239, March.

Jones, Charles I. (2005), "Growth and Ideas," Handbook of Economic Growth, in: Philippe Aghion & Steven Durlauf (ed.), *Handbook of Economic Growth*, edition 1, volume 1, chapter 16, pages 1063-1111, Elsevier.

Jones, Charles I. and Christopher Tonetti (2020), "Nonrivalry and the Economics of Data", American Economic Review 2020, 110(9): 2819–2858

Koh, Dongya, Raül Santaeulàlia-Llopis and Yu Zheng (2020), "Labor Share Decline and Intellectual Property Products Capital", *Econometrica* 88(6): 2609-2628, November.

Lanier, Jaron (2010), *You Are Not a Gadget: A Manifesto*. New York: Alfred A. Knopf.

Lanier, Jaron (2013), *Who Own's the Future?*, San Jose: Simon & Schuster.

Machlup, Fritz (1962), *The Production and Distribution of Knowledge in the United States*. Princeton, N.J.: Princeton University Press.

Mayer-Schönberger, V., & Cukier, K. (2013), *Big data: A revolution that will transform how we live, work, and think.* Houghton Mifflin Harcourt.

McKinsey Global Institute (2016), "The Age of Analytics: Competing in a Data-Driven World", *MGI Research Report* (December). Available here.

McKinsey Global Institute (2021), The consumer-data opportunity and the privacy imperative", Available: https://www.mckinsey.com/~/media/mckinsey/business%20functions/risk/our%20insights/the%20consumer%20data%20opportunity%20and%20the%20privacy%20imperative/the-consumer-data-opportunity-and-the-privacy-imperative.pdf?shouldIndex=false

Mouton, Brent (2018), "The Measurement of output, prices and productivity." Brookings Institution. Available: https://www.brookings.edu/wp-content/uploads/2018/07/Moulton-report-v2.pdf

Muciano-Goroff, Raviv, Ran Zhuo and Shane Greenstein (2021), "Hidden Software and Veiled Value Creation: Illustrations from Server Software Usage." Mimeo.

Nakamura, Alice Orcutt and Diewert, W. Erwin and Greenlees, John and Nakamura, Leonard I. and Reinsdorf, Marshall B., Sourcing Substitution and Related Price Index Biases (March 29, 2014), *FRB of Philadelphia Working Paper* No. 14-34, Available at SSRN: https://ssrn.com/abstract=2524022 or http://dx.doi.org/10.2139/ssrn.2524022

Nakamura, Emi, Jon Steinsson, Patrick Sun, Daniel Villar (2018), "The Elusive Costs of Inflation: Price Dispersion during the U.S. Great Inflation", *Quarterly Journal of Economics* 133(4): 1933–1980.

Nicoletti, Giuseppe, Christina von Rueden, Dan Andrews, "Digital technology diffusion: A matter of capabilities, incentives or both?", *European Economic Review*, Volume 128, 2020, 103513, ISSN 0014-2921, https://doi.org/10.1016/j.euroecorev.2020.103513.

Nguyen, David and Marta Paczos (2020), "Measuring the Economic Value of Data and Cross-border Data flows: A business Perspective", *OECD Digital Economy Papers* No. 297. OECD Publishing, Paris.

Organization for Economic Cooperation and Development (2013). "Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value", *OECD Digital Economy Papers* No. 220, OECD Publishing, Paris.

Open Data Institute (2016), "Research: The economic value of open versus paid data." Available at: https://theodi.org/article/research-the-economicvalue-of-open-versus-paid-data

PriceWaterhouseCoopers LLP (2019), "Putting a value on data." Available at: https://www.pwc.co.uk/data-analytics/documents/putting-value-on-data.pdf.

Roca, Ciara (2021a), "How to build a modern data stock from the ground up". Available at: https://fivetran.com/blog/how-to-build-a-modern-data-stack-from-the-ground-up

Roca, Ciara (2021b), "What is the modern data stack?" https://fivetran.com/blog/what-is-the-modern-data-stack (last accessed August 25, 2021).

Romer, Paul M (1990), Endogenous Technological Change, *Journal of Political Economy* 95(5): S71–S102.

Spiekermann, S., A. Acquisti, Rainer Böhme and K. Hui (2015), "The challenges of personal data markets and privacy." *Electronic Markets* 25: 161-167.

Statistics Canada (2019a), "Measuring Investment in Data, Databases and Data Science: Conceptual Framework." (June)

Statistics Canada (2019b), "The Value of Data in Canada: Experimental Estimates." (July).

Stiroh, Kevin J. (2002), "Are ICT Spillovers Driving the New Economy?",*Review of Income and Wealth* 48(1):33-57.

Syverson, Chad (2013), "Will History Repeat Itself? Comments on 'Is the Information Technology Revolution Over?", *International Productivity Monitor* 25 (Spring): 37–40.

The Economist (2017), "Leader: The world's most valuable resource is no longer oil, but data." The Economist Group Limited (May 6).

Tucker, Catherine (2019), "Digital Data, Platforms and the Usual [Antitrust] Suspects: Network Effects, Switching Costs, Essential Facility," *Review of Industrial Organization* 54(4): 683-694.

Varian, Hal (2018), “Artificial Intelligence, Economics, and Industrial Organization”, *NBER working paper* 24839.

Vinod, H.D. and P.M. Rao (2000), “R&D and Promotion in Pharmaceuticals: A Conceptual Framework and Empirical Exploration.”, *Journal of Marketing Theory and Practice* 8 (4): 10-20.

UNCTAD (2019), Competition Law, Policy and Regulation in the Digital Era. Available at: https://unctad.org/system/files/official-document/ciclpd57_en.pdf