SCIENCE POLICY RESEARCH UNIT

SPRU Working Paper Series

SWPS 2019-14 (July)

The Impact of Increasing Returns on Knowledge and Big Data: From Adam Smith and Allyn Young to the Age of Machine Learning and Digital Platforms

Yao-Su Hu

Energy Policy

Sustainable Development

Science and Technology Policy

Innovation and Project Management

Economics of Innovation

UNIVERSITY OF SUSSEX | BUSINESS SCHOOL

SPRU Working Paper Series (ISSN 2057-6668)

The SPRU Working Paper Series aims to accelerate the public availability of the research undertaken by SPRU-associated people, and other research that is of considerable interest within SPRU, providing access to early copies of SPRU research.

Guidelines for authors

Papers should be submitted to swps@sussex.ac.uk as a PDF or Word file. The first page should include: title, abstract, keywords, and authors' names and affiliations. The paper will be considered for publication by an Associate Editor, who may ask two referees to provide a light review. We aim to send referee reports within three weeks from submission. Authors may be requested to submit a revised version of the paper with a reply to the referees' comments to swps@sussex.ac.uk. The Editors make the final decision on the inclusion of the paper in the series. When submitting, the authors should indicate if the paper has already undergone peer-review (in other series, journals, or books), in which case the Editors may decide to skip the review process. Once the paper is included in the SWPS, the authors maintain the copyright.

Websites

UoS: **www.sussex.ac.uk/spru/research/swps**
SSRN: **www.ssrn.com/link/SPRU-RES.html**
IDEAS:**ideas.repec.org/s/sru/ssewps.html**

The Impact of Increasing Returns on Knowledge and Big Data: From Adam Smith and Allyn Young to the Age of Machine Learning and Digital Platforms

Yao-Su Hu[1]

Hong Kong Shue Yan University[2] and SPRU, University of Sussex

Abstract

Allyn Young's concept of increasing returns, not to be confounded with static, equilibrium constructs of economies of scale and increasing returns to scale, is applied in this article to analyze how and why increasing returns arise in the production (generation) and use (application) of knowledge and of big data, thereby driving economic growth and progress. Knowledge is chosen as our focus because it is 'our most powerful engine of production' and big data is included to make the analysis more complete and up-to-date. We analyze four mechanisms or sources of increasing returns in the production of knowledge, and four in the use of knowledge. Turning to big data, increasing returns in the use thereof are examined in two spheres: the dominance resulting from the self-reinforcing functioning of digital platforms and machine learning through gigantic amounts of training data. Concluding remarks concern some key differences between knowledge and big data, some policy implications, and some of the social negative impacts from the ways in which big data is being used.

Note: Accepted in June 2019 for publication in *Prometheus.*

---

[1] I would like to express sincere thanks to Armar Archbold, Martin Bell, Mike Hobday, Yves Le Coadic, Stuart Macdonald, Lilach Nachum, Pari Patel, Ed Steinmueller, and Tony Yu for ideational encouragement.

[2] **Contact** yshu@hksyu.edu

Introduction

More than 120 years ago and foreshadowing the last forty years' widespread fascination[i] about the 'knowledge economy', Alfred Marshall put forward in the third edition of his *magnum opus*: 'Knowledge is our most powerful engine of production; it enables us to subdue Nature and force her to satisfy our wants. Organization aids knowledge…' (1895, Book IV, Ch.1). Although Marshall gave hints such as that the distinction between public and private property in knowledge was of great and growing importance or that capital was made up of knowledge and organization, he did not directly address the question of how knowledge was able to achieve its function as our most powerful engine of production. This is an important question,[ii] which does not appear to have not been adequately addressed, not even by 'new (endogenous) growth theory'. [iii] This paper is an attempt to grapple with the question, and to do so by adopting a dynamic (non-equilibrium) approach inspired by Allyn Young's conceptualization of increasing returns in his famous paper in *The Economic Journal* (Young, 1928).

In relation to the extensive existing literature on knowledge and on increasing returns, how is this paper different? First, although the characteristics of knowledge have been discussed (Hess and Ostrom, 2011, 8-10; Romer, 1990, S73-75; Foray, 2004, 91-101; Nonaka, Toyama and Hirata, 2008, 6-15; Zukerfeld, 2017), in neo-classical economics (including 'new growth theory') the connection between knowledge and economic growth/progress[iv] has not been addressed except indirectly in terms of spillovers/externalities, human capital or R&D (within complex mathematical models constrained by the requirement of Pareto-optimal general equilibrium); this paper relies squarely on increasing returns as the connecting principle between knowledge and economic growth/progress. Secondly, scholars in contemporary evolutionary economics have, following Adam Smith, identified the division of labour (and the concomitant division of knowledge) as the primary means of promoting increasing returns and the growth of knowledge;[v] in this paper, given that increasing returns in relation to knowledge and data is our principal focus, we analyze not only the division of labour but also (without aiming to be exhaustive) a range of other sources or mechanisms of knowledge-based and data-based increasing returns. Thirdly, although the expressions "increasing returns to scale" and "economies of scale" are frequently used in the relevant literature, they are static, equilibrium constructs, completely different from the dynamic concept of increasing returns as elaborated by Young and adapted in this paper.

Fourthly, we reject static, equilibrium economic theory in which change is reduced from being processes which unfold in historical time to comparisons of equilibrium positions in a timeless world.[vi] The implications of adopting, within a dynamic approach, the Youngian concept of increasing returns are profound. Nobel Economics Laureate Theodore Schultz wrote: 'It may be elementary but it is often overlooked that increasing return activities do not exist in the axiomatic core of general equilibrium theory, whereas each and every increasing returns event implies that there is disequilibrium…The idea of increasing returns has become a spoiler at this high table of theory" (1993, 18, 23-4). The determinacy (and apparent rigour and elegance) of static equilibrium analysis disappears or evaporates in a dynamic approach. 'Increasing returns generate not equilibrium but instability' (Arthur, 1996, 100) and the hallmarks of increasing returns in certain information-intensive sectors have included: market instability, multiple potential outcomes, unpredictability, the ability to lock in a market, the possible predominance of an inferior product, and fat profits for the winner (Ibid, 102). Finally, although many different kinds of knowledge processes/activities have been identified in different kinds of literature,[vii] this paper focuses on knowledge production/generation and knowledge utilization/application as the essence of how knowledge accomplishes its function as Marshall's 'most powerful engine of production' driving growth/progress.[viii]

In this paper, we apply the Youngian concept of increasing returns to knowledge, but will also examine increasing returns in relation to big data, because data, information and knowledge are often considered to be parts of the same continuum, and because big data, machine learning (using big data to learn) and digital platforms (capturing and exploiting big data) are currently at the cutting edge of much technological change. We will proceed in the following manner. First, we define knowledge, information and data and briefly explore their interrelationships and differences. We then discuss the notion of increasing returns according to Young, formulate a definition of increasing returns (which Young did not do), and articulate the linkage between Young's work and knowledge. Next, we analyze the mechanisms and processes whereby increasing returns arise in: (a) the generation and production of knowledge, and (b) the use and application of knowledge. Then we look at how increasing returns arise in the generation and capture of big data, and how they arise in the use/reuse of big data in two overlapping spheres: *machine learning* and *digital platforms* (such as Amazon and Google). We end with selective remarks concerning some key differences between big data and knowledge, some of their implications and some of the negative social impacts from the ways in which big data is being used.

What is knowledge? What about information and data?

The definition of knowledge, information and data and the specification of the differences and inter-relationships between the trio are not a simple, straightforward matter. An interactive study conducted in 2003- 2005 documented 130 definitions of data, information, and knowledge formulated by 45 scholars in the field (Zins, 2007). There is no universally agreed definition of data, nor of information, nor of knowledge. A detailed analysis of 16 influential textbooks in information systems and/or knowledge management (Rowley, 2007) found that typically information was defined in terms of data and knowledge in terms of information, but there was less consensus in the characterisation of the processes by which data is transformed into information, and information into knowledge, resulting in a lack of definitional clarity. For example, when data is transformed into information, it is a structure or is it meaning which is added? An influential perspective is that 'knowledge derives from information as information derives from data' (Davenport and Prusak, 1998, 6) but this runs into the book entitled '*Raw Data' Is An Oxymoron* (Gitelman, 2013) and also the fact that there are six approaches to the definition of information[ix] (Floridi, 2003).

According to Fritz Machlup knowledge has two meanings ---- that which is known and the state of knowing.[x] 'Knowledge as a state of knowing is produced by activities such as talking plus listening, writing plus reading, but also by activities such as discovering, inventing, intuiting. In the first group of activities, at least two persons are involved, a transmitter and a receiver, and the state of knowledge produced in the consciousness of the recipient refers to things or thoughts already known, at least to the transmitter' (Machlup, 1980, 28). Thus (and this is important for the purpose of this paper) the production of knowledge includes, by definition, the reproduction of knowledge, and need not refer only to the production of new knowledge (i.e. knowledge that is new to the world or society). The two meanings of knowledge also raise a perplexing problem for the measurement of knowledge: when more people come to know the same thing (i.e. when the same knowledge is diffused), does the social stock of knowledge increase or remain the same?

In an age in which 'open access' online to certain types of knowledge is proclaimed to be a universal human right, Hess and (the Nobel Economics Laureate) Ostrom propose the following definition: 'Knowledge…refers to all intelligible ideas, information, and data in whatever form in which it is expressed or obtained. Knowledge…refers to all types of understanding gained through experience or study, whether indigenous, scientific, scholarly, or otherwise nonacademic. It also includes creative works, such as music and the visual and theatrical arts' (Hess and Ostrom, 2011, 7-8). Furthermore, because of the complex and social nature of knowledge, it is useful to make a threefold distinction between knowledge facilities, knowledge artifacts and ideas. 'Facilities store artifacts and make them available[xi]… Artifacts are discrete, observable, nameable representations of ideas, such as articles, research notes, books, databases, maps, computer files, and web pages…Ideas are coherent thoughts, mental images, creative visions, and innovative information.

Ideas are the intangible content and the nonphysical flow units contained in artifacts' (Ibid, 2011, 47-8).

In neo-classical economic theory, it is standard to assume two key properties of knowledge: (1) that the consumption or use of knowledge is *non-rival* in the sense that consumption (use) of knowledge by A does not reduce the amount available to B or anyone else, and (2) that knowledge is heterogeneous with respect to *excludability* (i.e. the ability to exclude those who have not paid from access to and use of knowledge); in that excludability depends on both the nature of the knowledge in question and on the socially-constructed institutions and laws governing intellectual property rights and the ownership of (pieces of) knowledge. Non-rivalry (also known as non-subtractability) and non-excludability will be elaborated further as we proceed.

What about information? Again according to Machlup (1983, 642) the two traditional meanings of information are (1) the action of informing and (2) that which is being told (informed). Moreover, whereas information is acquired by being told, knowledge can be produced by thinking. "*Thus, new knowledge can be acquired without new information being received*" (Ibid, 644). In neo-classical economic theory, knowledge and information have been used as interchangeable terms; thus, in perfect competition, actors are assumed to have perfect information or knowledge, and both are subject to near zero transfer costs. Since roughly the start of this millennium, this equivalence between knowledge and information has been questioned. According to Foray (2004, 4), knowledge involves cognitive capability for action, whereas information remains passive and inert until used by those with the required knowledge. Cimoli, Dosi, Nelson and Stiglitz (2009, 22) identified two kinds of things which are included in knowledge but absent in information: the cognitive categories which allow information to be interpreted, and search and problem-solving heuristics which cannot be reduced to information. In our view, the key difference between knowledge and information stems, as Michael Polanyi pointed out more than half a century ago, from 'the fact that we can know more than we can tell' (Polanyi, 1966, 4), an instance of which is that 'the aim of a skilful performance is achieved by the observance of a set of rules which are not known as such to the person following them' (Polanyi, 1958, 49). Whereas knowledge includes tacit knowledge (i.e. knowledge which is non-codified or non-codifiable), information cannot encompass 'tacit' information, which would be a contradiction in terms. This difference is of fundamental importance because it affects the transferability of knowledge (Hu, 1995), its mode of transfer, and whether knowledge is 'sticky' or escapes freely when it is produced or used. Tacit knowledge is more difficult and costly to transfer (and to assimilate) than codified knowledge, information and data, but on the other hand, it is easier to exclude outsiders from accessing or using tacit knowledge. This means that tacit knowledge can be a source of a firm's competitive advantage since it (the tacit knowledge) is by definition not easily accessible, imitable or copyable.

What is data? According to yet again to Machlup: 'The use and misuses of the term *data* are due, in part, to linguistic ignorance. Many users do not know that this is a Latin word: *dare* means "to give"; *datum,* "the given" (singular); and *data*, "the givens" (plural). Data are the things given to the analyst, investigator, or problem-solver; they may be numbers, words, sentences, records, assumptions---just anything given, no matter in what form and of what origin' (Machlup, 1983, 646). With modern computers and the Internet, data and information are stored, processed, and transmitted in digital form (in bits, i.e. binary digits, and bytes, i.e. sets of 8 bits); data in analogue form (images, sounds, documents,etc.) can be converted into digital data using analogue-digital converters. One of the advantages of digitization is said to be the near zero cost, and the speed and accuracy, with which digital data can be reproduced or transmitted with no degradation in quality (fidelity) compared with analogue data/information. In the digital age, data can be defined as the representation of facts stored or transmitted in digital format (OECD, 2015, Glossary).[xii]

Given digital data, what is 'big data'? A series of techno-economic and social changes have, over the last twenty years, enabled and encouraged first a torrent, then an explosion, in the generation of all kinds of data which have been captured and stored in digital form. The definition of big data is not in terms of datasets being larger than a certain number of terabytes or petabytes, because technology is allowing the digital storage capacity to increase, and unit storage cost to fall, all the time. One way of defining big data is 'datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data' (McKinsey Global Institute, 2011, 1). To be useful, big data needs to be converted and structured into information; data analytics is defined as the techniques and tools used to extract information from data by revealing the context in which the data are embedded, their organization and their structure, and hidden patterns and correlations; data analytics overlaps with data cleaning, data mining, profiling, business intelligence and machine learning (OECD 2015, glossary, 452). The sources of big data encompass the following: (1) computer-mediated economic transactions; (2) data from billions of sensors embedded in a widening range of objects, bodies and places; (3) massive government and corporate databases, e.g. census, tax and financial records; (4) private and public surveillance cameras 'including everything from smartphones to satellites, Street View to Google Earth' (Zuboff, 2015, 78); and (5) non-market activities such as 'Facebook likes, Google searches, emails, texts, photos, songs and videos, location, communication patterns, networks, purchases, movements, every click, misspelled word, page view, and more. Such data are acquired, datafied, abstracted, aggregated, analyzed, packaged, sold, further analyzed and sold again' (Zuboff, 2015, 79). The ways in which big data is generated and collected account for much of the nature and characteristics of big data: 'big data tends to be heterogeneous, unstructured or semi-structured and agnostic. It is also trans-semiotic, entailing different combinations of data formats and communication modes that extend beyond alphanumeric systems to include tokens cast into the media of text, sound and image' (Constantiou and Kallinikos, 2015, 50). Big data is also real time and/or continuously updatable. We will return to increasing returns in relation to big data after analyzing increasing returns in the production and utilization of knowledge.

Youngian increasing returns

To grasp Allyn Young's approach it is useful to distinguish, in his 1928 paper, between the (non-mathematical) 'model' expounded and the methodological points interspersed[xiii] throughout the paper. At the risk of gross oversimplification, the essence of the model can be said to comprise three mechanisms or processes of increasing returns. First is 'the progressive division and specialisation of industries' (Young, 1928, 539). The archetypical example cited is as follows: 'The successor of the early printers…are not only the printers of today…but also the producers of wood pulp, of various kinds of paper, of inks and their different ingredients, of type-metal and of type, the group of industries concerned with the technical parts of the producing of illustrations, and the manufacturers of specialised tools and machines for use in printing and in these various auxiliary industries' (Ibid, 537-8). In this process of specialization or division of labour among industries, the representative firm loses its identity, which makes it meaningless to speak of the economies of scale of the firm. Secondly, this process of industrial differentiation is accompanied by the adoption of more 'roundabout' (i.e. more capital-intensive) methods of production because 'with the division of labour a group of complex processes is transformed into a succession of simpler processes, some which, at least, lend themselves to the use of machinery' (Ibid, 530). Thirdly, the increasing specialization of industries and increasing use of machinery drive the growth of productivity and output, which creates a virtuous circle between the growth of supply and the growth of demand at the aggregate level because, on the demand side, 'capacity to buy depends upon the capacity to produce…The size of the market is determined and defined by the volume of production' (Ibid, 533). Thus, 'the division of labour depends upon the extent of the market, but the extent of the market also depends upon the division of labour. In this circumstance lies the possibility of economic progress…' (Ibid, 539).

More important than the model are the methodological observations. First, the forces making for continuous changes are endogenous, 'engendered *within* the economic system' (Ibid, 535). This was more than sixty years before the emergence of 'endogenous growth theory.'[xiv] Secondly, as Charles Blitch demonstrates through a careful study of the correspondence between Young and his former student Frank Knight: 'In Young's view, competitive equilibrium methodology was not only inadequate but was a positively misleading vehicle for the examination of the returns phenomenon.' (Blitch, 1983, 360). Thus, we read in Young's paper: 'New products are appearing, firms are assuming new tasks, and new industries are coming into being. In short, change in this external field is qualitative as well as quantitative. No analysis of the forces making for economic equilibrium…will serve to illumine this field, for movements away from equilibrium, departures from previous trends, are characteristic of it…[T]he counter forces which are continually defeating the forces which make for economic equilibrium are more pervasive and more deeply rooted in the constitution of the modern economic system than we commonly realise' (Young, 1928, 528, 533).

Thirdly, in stark contrast to the *ceteris paribus* method of looking at one thing at a time, Young argues: 'What is required is that industrial operations be seen as an interrelated whole' (Ibid, 539). This led Schultz (1993, 8) to say: 'Young's approach to the origins of increasing returns entailed a proliferation of Smith's division of labour. Young stressed the *togetherness* of the various increasing returns activities. What he wanted to ascertain was the *sum* of the interacting effects of all of the various increasing returns activities.' Fourthly, change is cumulative: 'Every important advance in the organisation of production…alters the conditions of industrial activity and initiates responses elsewhere in the industrial structure which in turn have a further unsettling effect. Thus, change becomes progressive and propagates itself in a cumulative way' (Young, 1928, 533). Fifthly, the realising of increasing returns is 'a process requiring time' because 'new trades have to be learnt and new habits have to be acquired' and 'the accumulation of the necessary capital takes time' (Ibid, 533).

What increasing returns are not

It is important to bear in mind that Youngian increasing returns are radically different from increasing returns to scale and also from economies of scale.[xv] Increasing returns to scale are based on the static equilibrium construct of the production function, in which an equiproportional increase in inputs may lead to an increase in output by the same proportion (this is labelled as constant returns to scale), a greater proportion (increasing returns to scale) or a smaller proportion (decreasing returns to scale). If 'knowledge' is represented by the symbol A and included with the other inputs (labour, capital, etc.) on the right-hand side of the equation, then an X% increase in output will require a less than X% increase in inputs, because A is postulated to be 'non-rival' so that a larger output does not require a larger amount of A; thus, output increases by more than inputs in proportional terms, and by definition this represents increasing returns to scale in relation to knowledge. This is unlikely to generate large and continuous impacts in terms of growth or progress. Moreover, whereas Youngian increasing returns are generated to a large extent through specialization (between industries), 'increasing returns to scale…are based on the notion of a production function defined over a set of factors of unchanging quality, while the very essence of the notion of specialization is that the quality of at least one factor involved in production changes' (Witt, 2000, 736). As static, equilibrium constructs, both increasing returns to scale and economies of scale assume away (treat as non-existent) the following: the passage of time, changes in the quality of inputs, changes in factor proportions, technological changes, organizational changes, inventions and innovations, the emergence of new products and services, etc.

Definition of increasing returns

To apply Youngian increasing returns to knowledge and data, it is necessary to define the expression (if only to help us to pull away from the mesmerizing hold of static, equilibrium thinking). It is noteworthy that in his 1928 paper, Young did not offer a definition of increasing returns. This was confirmed years later by one of his former Harvard students, Lauchlin Currie: 'Although, as remarked earlier, he [Young] did not offer a definition, it is clear that he was using the phrase in a rather novel manner' (Currie, 1981, 54).

Returns should not be confined to nor confounded with 'economies' i.e. savings on the cost side as in 'economies of scale.' [xvi] Returns include economies but extend much more widely to encompass all kinds of benefits or outcomes --- outputs, profits, gains, payoffs, utility, yields, efficiencies, ideas, designs, knowledge, solutions, etc. *Increasing* returns refer to continuous growth or change (as distinct from once-for-all increases)[xvii] in the returns other than in the short term, and may be divided into quantitative or qualitative increasing returns; the former include economic growth, growth in productivity, greater speed, lower costs, etc., while the latter come under the categories of development, progress, structural change, creativity, novelty, innovation, etc. Increasing returns may also be seen in an absolute or a relative manner. In absolute terms, increasing returns mean that the more something[xviii] is done (e.g. the greater the cumulative volume of the activity), the greater and/or better the returns tend to be over time. Relative increasing returns concern gaps, disparities or inequalities between persons, groups, firms, races or nations. They are 'the tendency for that which is ahead to get further ahead, for that which loses advantage to lose further advantage' (Arthur,1996, 100) or (in the Matthew effect)[xix] the tendency for initial advantage to beget further advantage, and disadvantage further disadvantage, creating widening gaps between the haves and the have-nots (Rigney, 2010, 1).

Finally, and perhaps most importantly for our definition, the realization of increasing returns is neither automatic nor effortless; increasing returns are not given (e.g. by the laws of economics or physics), they are not low-hanging fruits 'to be had merely for the taking' (Young, 1928, 541). As Loasby put it, 'increasing returns have to be worked for, and cannot be selected from a previously defined production set …' (1989, 52). A good example of this in the contemporary world is Moore's Law.[xx] 'Moore's law is very different from the laws of physics or Newtonian classical mechanics. Those laws are true no matter what we do. Moore's Law in contrast is a statement about the work of the computer industry's engineers and scientists; it's an observation about how constant and successful their efforts have been. The reason that Moore's Law has held up so well for so long is what we might call 'brilliant tinkering'---finding engineering detours around the roadblocks thrown up by physics…When it became difficult to cram integrated circuits more tightly together, for example, chip makers instead layered them on top of one another' (Brynjolfsson and McAfee, 2016, 41-42).

In sum, increasing returns can be defined, over time (i.e. other than in the short run), as either the tendency for any activity or work to generate better outcomes the more it is done (with effort and ingenuity) or the tendency for that which is ahead to get further ahead and for that which is behind to fall further behind. Increasing returns may well be more widespread than might appear at first: it goes under different labels whose contents are broadly similar – self-reinforcement, circular and cumulative causation, positive (but not negative)[xxi] feedback, virtuous and vicious circles, non-convexity, etc. (Arthur, 1988, 10).

The linkage with knowledge

What is the connection between Youngian increasing returns and knowledge? In this paper, increasing returns do not arise so much from knowledge *per se* as from the *production* and *use* of knowledge. Knowledge can be seen in terms of stocks or flows, but knowledge production and use are definitely activities, processes, things that are done with purpose, effort and ingenuity. In this, they accord with our definition of increasing returns in the first sense (the more something is done, the greater/better are the returns). Also, we follow Young's orientation in which the use of *both* existing and new knowledge drives economic progress. New knowledge refers to 'new inventions' (Young, 1928, 534) and 'the growth of scientific knowledge' (Ibid, 535) while existing knowledge is generated, applied and improved in the division of labour (in the production sector) and includes 'adaptations of known ways of doing things' (534). According to Roger Sandilands, an authority on Allyn Young: 'Youngian endogenous growth relates to that part of growth that depends on growth itself, and the way it permits and encourages the fuller exploitation and adaptation of existing knowledge' (Sandilands, 2000, 325).

Increasing returns in the production of knowledge

Increasing returns in the production of knowledge, seen at the social, group or team level, are due mainly to the increasing division of labour (or specialization)[xxii] in the production of knowledge. This was pointed out by Adam Smith some 240 years ago. In an early draft of *The Wealth of Nations*, he wrote: 'Philosophy or speculation, in the progress of society, naturally becomes, like every other employment, the sole occupation of a particular class of citizens. Like every other trade, it is subdivided into many different branches, and we have mechanical, chemical, astronomical, physical, metaphysical, moral, political, commercial and critical philosophers. In philosophy, as in every other business, this subdivision of employment improves dexterity and saves time. Each individual is more expert at his particular branch. More work is done upon the whole and the quantity of science is considerably increased by it' (cited in Rosenberg, 1965, 136). Moreover, according to Smith, the increasing returns from specialization are not merely allocative (according to the comparative advantage of each), they are creative (creating new knowledge and skills): 'The difference of natural talents in different men is, in reality, much less than we are aware of, and the very different genius which appears to distinguish men of different professions, when grown up to maturity, is not upon many occasions so much the causes, as the effect of the division of labour' (1776, Book I, Ch.2). Thus, the (increasing) returns from Smithian division of labour in the production of knowledge include: the greater individual expertise of the scientists, the creation of new knowledge, and growth in society's stock of scientific knowledge. These returns, however, come at a price which results from excessive fragmentation of the knowledge base – the dispersion of expertise, the difficulty of coordination and communication across areas of expertise, and the possible narrow-mindedness of some of the experts (Metcalfe, 2014, 11).

According to Pavitt (2000), the increasing division of labour in the production of knowledge since the time of Adam Smith has taken three complementary forms. First, there has been further specialization by discipline, involving the emergence of new engineering disciplines in the wake of new scientific disciplines. Thus chemical engineering emerged in the wake of applied chemistry, electronic engineering followed solid state physics, etc. Secondly, since the beginning of the twentieth century, in the business sector there has been specialization by corporate function within large firms and also a division of labour between large and small firms. Following the emergence of corporate R&D laboratories in large firms, the division and coordination between R&D, design, manufacturing and marketing within the large firm has been one of the most important factors differentiating successful from unsuccessful innovation, and a division of labour has emerged between large firms and a myriad of small, specialist suppliers of continuous improvements in capital goods. Thirdly, there has been a division of labour between the institutional actors in the 'national system of innovation,' with fundamental or basic research (and related post-graduate training) taking place in universities and public research institutes, and commercial R&D being the province of business firms. The *returns* from the three complementary forms of division of labour have also assumed various forms: creation of new knowledge in new disciplines; continuous improvements in the techniques of experimentation; improving capacity to design, develop and test prototypes and pilot plants; improving specialized instrumentation; continuous leaps in the ability to observe distant or minute phenomena which are out of reach for the naked eye; improving ability to experiment with a wider range of products and processes; faster technical change; and the 'roundabout and indirect' benefits of university research such as the provision of training to graduates, who may then act as carriers of new insights and as problem solvers for their employers. All these are not once-for-all events but are continuing sources of endogenous economic growth.

Moving from the social level of analysis to the level of individual agents and teams, there are also processes or mechanisms that generate increasing returns in the production of knowledge. Only three of these will be mentioned due to space limitations. First, the more relevant knowledge a person has accumulated, the more s/he can contribute in conversations with peers or in discussions in relevant 'epistemic' communities and/or communities of practice, and, because of the universal norm of reciprocity, the more knowledge and ideas s/he will receive in return (Macdonald, 1998, 21-23). Moreover, because prior relevant knowledge determines a person's 'absorptive capacity' (Cohen and Levinthal, 1990), s/he will benefit more from such informal information exchanges than a less knowledgeable person (a perfect example of the Matthew effect).

Secondly, the larger is the stock of existing knowledge in a given field, the greater is the number of possible recombinations and permutations of (bits or pieces of) this knowledge that can be made and the more likely it is that some of these would create further new knowledge. [xxiii] After a certain point, however, diminishing returns may set in because of the exhaustion of technical possibilities or opportunities in that field, but this may again be offset by radical new discoveries, inventions or breakthroughs (which may originate in other fields).

Thirdly, at the level of colleagues grouped into a project team within a firm, Nonaka and Takeuchi's (1995) SECI (socialization, externalization, combination and internalization) model can be understood as a circular and ascending spiral for generating increasing returns in the production of knowledge. *Socialization* encourages the transfer and sharing of tacit knowledge, leading to the creation of greater breadth and depth in the collective tacit knowledge; *externalization* transforms the tacit knowledge into explicit (codified) knowledge; *combination* of (different bits of) explicit knowledge may result in the creation of new ideas, designs or pieces of explicit knowledge; *internalization* (learning by doing) converts explicit knowledge into deeper tacit knowledge of the know-how kind. *The point is that the stock of knowledge increases* because the use of knowledge is non-subtractable (non-rival): if tacit knowledge is used to produce explicit knowledge, the amount of tacit knowledge is not diminished, but the amount of explicit knowledge at the level of the team may well increase so that the total amount of knowledge would also increase, and similarly for conversions from tacit to explicit, from one form of knowledge to another. Meanwhile, the quality or usefulness of the knowledge may also improve or become more finely tuned.

In sum, knowledge producing activities give rise to increasing returns because of the progressive division of labour in knowledge production and creation, because of the Matthew effect, because of the recombination of pieces of knowledge, and because of the conversion of knowledge from tacit to explicit and vice versa.

Increasing returns in the use of knowledge

The importance to society of its stock of knowledge was described by Allyn Young thus: 'From one point of view …a nations's real capital does not consist so much of buildings, machines, railways and the like, as of the industrial and technical knowledge which the members of the community possess, handed down from generation to generation, diffused by education and increased by the advances of science. Our knowledge and our science are infinitely more precious possessions that the whole aggregate of our accumulated capital. Destroy our material accumulations and, in a generation, they will have been replaced. Destroy our knowledge, annihilate our science, and man, reduced to the intellectual status of the savage, would not even know how to use the tools which he has made for himself' (Young, 1929/1999, 237). How are increasing returns generated in the use of knowledge? There are at least four mechanisms through which this is realized.

First, the use of knowledge not only does not *use up* knowledge but promotes its further growth and development. It is through *flows* of knowledge that the *stock* of knowledge comes to be *used*. By definition, flows of knowledge are flows between different persons, between transmitters or sources and receivers of knowledge.[xxiv] Now there is a critically important difference between flows of knowledge and flows of material goods: 'A flow of goods from one person to another reduces the stocks of the former and increases the stocks of the latter. By contrast, a flow of knowledge may increase the recipient's stock of knowledge without reducing the stock of the transmitter. This implies that every flow of knowledge may bring about an increase in the combined stock of knowledge' (Machlup, 1980, 237). This property of knowledge, which is due to its immaterial or intangible nature, has been labelled 'non-rivalry' in the neo-classical and endogenous growth theory literature, to reflect the fact that a 'purely non-rival good has the property that its use by one firm or person in no way limits its use by another' (Romer, 1990, S74). However, the expression 'non-subtractability', put forward by Hess and Ostrom (2011, 9) may be more accurate --- a flow of knowledge from A to B is non-subtractable even where A and B are not rivals or potential rivals. Non-subtractability means that the more knowledge is used, the more of it there is around, and the more of it there is around, the more it can be used further, and so on in a virtuous circle. Furthermore, 'ideas breed new ideas' (Davenport and Prusak, 1998, 17), partly because new ideas may arise from the recombination of existing ideas. Moreover, knowledge is seldom used as such but often requires modification and adaptation to be applied, which generate more and newer knowledge in the process. The sharing and exchange of ideas leading to new ideas are described in a timeless paragraph about industrial districts in Marshall's *Principles of Economics*: 'The mysteries of the trade become no mysteries; but are as it were in the air, and children learn many of them unconsciously. Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the source of further new ideas' (Marshall, 1920, Book IV, Ch.10).

Secondly, the increasing division of *labour* in the production of knowledge results in the increasing 'division of *knowledge*…The growth of knowledge proceeds through the differentiation and dispersion of knowledge…' (Loasby, 1998, 164-5). Consider first Adam Smith's analysis of the woollen coat as an example of the division of labour: 'The woollen coat, for example, which covers the day-labourer, as coarse and rough as it may appear, is the produce of the joint labour of a great multitude of workmen. The shepherd, the sorter of the wool, the wool-comber or carder, the dyer, the scribbler, the spinner, the weaver, the fuller, the dresser, with many others, must all join their different arts in order to complete even this homely production' (Smith, [1776] 1976), Book I, Ch. I). Now consider the application of divided knowledge or specialized knowledge in the construction of a modern house: 'Modern homes, with indoor plumbing, insulation, temperature control, full-service kitchens, and home entertainment systems, require a group effort. Consider the variety of trades that participate in building a modern home: surveyors, excavators, framers, bricklayers, roofers, plumbers, drywall and window installers, carpenters, painters, plasterers, electricians, cabinetmakers, landscapers, and carpet installers' (Sloman and Fernbach, 2017, 110-111). Thus the use of divided knowledge, given proper coordination, results in greater and better returns than the absence or a lower degree of specialization.

Thirdly, there is circular and cumulative causation between 'absorptive capacity' (the ability to evaluate and utilize 'external' knowledge) and application of knowledge. Psychological research suggests that memory development is subject to a virtuous circle ---- the more objects, patterns and concepts are used and stored in a person or team's memory, the better the absorptive capacity becomes; the better the absorptive capacity gets, the more readily is related new knowledge identified and assimilated and the easier it becomes for the person or team to use them in different settings (Cohen and Levinthal 1990). This helps to explain increasing returns to the use of knowledge/information 'in the sense that the more we use it, the easier it is, and dynamically, the higher is the likelihood of learning and producing ourselves "better", "novel", in some sense "innovative" further pieces of information' (Cimoli, Dosi, Nelson and Stiglitz, 2009, 22).

Last but not least, knowledge nurtures and supports creativity:[xxv] The more that knowledge is applied to creative ends, the higher becomes the probability of achieving creative outcomes. Creativity in its turn calls for more knowledge in order, for example, to overcome the technical obstacles of transforming a new concept into a workable prototype, and so creating a virtuous circle between the growth of knowledge and the growth of creativity. The application of relevant knowledge is a *necessary* but not *sufficient* condition for the emergence and development of creativity. The reasons are as follows: (a) Creative thinking and knowledge are not opposite forces but help each other. One of the first scientists to put forward this view was the original and influential Russian psychologist Lev Vygotsky in three papers originally written in 1930-1932 (Smolucha, 1992, 49). From a 2004 translation of his 1930 paper, we read: 'The creative activity of the imagination depends directly on the richness and variety of a person's previous experience. This experience provides the material from which the products of fantasy are constructed…All else being equal, the richer the experience, the richer the act of imagination…Fantasy is not the opposite of memory, but depends on it and utilizes its contents in ever new combinations' (Vygotsky, 2004, 14-16). (b) Prior knowledge guides the researcher in choosing between alternative paths. 'In the activity of invention, as in most goal-directed activities, the actor has a number of alternative paths among which he must choose. The greater his knowledge of the relevant fields, the more likely he will be eventually to find a satisfactory path, and the fewer the expected of tried alternatives before a satisfactory one is found. Thus, the greater the underlying knowledge, the lower the expected cost of making any particular invention' (Nelson, 1959, 300). Knowledge also helps to delineate the available tools and methods for problem solving in a domain. (c) knowledge can help an observer to recognize an opportunity ----to notice that something surprising has happened in a chance occurrence and that it is worthy of further investigation. 'Chance favours the prepared mind', Louis Pasteur is reported to have said. (d) Combinational or recombinant creativity, in which novelty results or emerges from the combination of existing ideas, requires a rich store of knowledge with sufficient variety therein. 'The prior possession of relevant knowledge and skill is what gives rise to creativity, permitting the sorts of associations and linkages that may have never been considered before' (Cohen and Levinthal, 1990, 130). (e) Knowledge is not a sufficient condition for creativity. An important reason is that, at the level of the individual, there is a significant difference in motivation and temperament between the creator and the expert: 'How does the creator differ from the expert? In my view, the difference is not principally cognitive…Tested on mastery of a domain, both kinds of individuals should perform equally well…The creator stands out in terms of temperament, personality, and stance' (Gardner, 2008, 82). The creative inventor strikes out in unfamiliar directions and enjoys going against the crowd, and accepts if not welcomes repeated failures and the prolonged solitude of the long-distance runner before any likelihood of a breakthrough.

In sum, the use and application of knowledge give rise to increasing returns because of the non-substractability of knowledge[xxvi] (resulting in the fact that use of knowledge generates more and/or new knowledge); because of the increasing division of knowledge; because of circular and cumulative causation between use and capacity; and because knowledge, though not a sufficient condition, helps and promotes creativity.

Since knowledge is used for a multiplicity of purposes, the increasing returns to its use can take a variety of forms, and encompass: better quality[xxvii] and/or functionality of products; higher productivity[xxviii] and lower costs of production; better decision-making; better problem-solving and the solving of more problems; continuous improvements in operations, processes, systems and services; the development of new materials, new products/services, new tasks/technologies; more creative R&D, design, engineering, prototyping and scaling up from pilot plants (activities which are essential in the process of creating and bringing new products to market); more effective marketing and targeting; new ideas, inventions, and innovations; sustained technological and organizational change; etc. These all are, as parts of 'an interrelated whole' (Young, 1928, 539), aspects or sources of economic growth/progress; they do not need to be translated into growth through complex mathematical manipulations involving production functions, diminishing marginal returns to capital, spillovers or externalities, R&D and/or human capital, monopoly, etc. all within the limits allowed by Pareto-optimal general equilibrium.

Increasing returns in the generation and capture of big data

The exponential growth over the last two decades in the quantities of big data that are generated, captured and stored has been due to circular and cumulative interactions between a number of techno-social developments. The digitization of 'all kinds of information and media --- text, sounds, photos, videos, data from instruments and sensors, and so on' (Brynjolfsson and McAfee, 2016, 61) plus the migration of social and economic activities to the Internet have acted as broad driving forces. Sensors are being embedded in mobile phones, cameras, smart energy meters, automobiles and industrial machinery which become interconnected in the Internet of Things. Any digital information or data that is connected to the Internet or telecommunications systems can be captured and stored, and this includes a tremendous amount of digital 'exhaust data,' (i.e. data that are created as a by-product of other activities). Social media sites, smartphones, and other consumer devices, including PCs and laptops, have allowed billions of individuals around the world to contribute to the amount of big data available. A certain proportion of the big data is stored for subsequent use and reuse (in the case of major players, in gigantic data centers located in deserts and inaccessible places).[xxix] The 'increasing returns' to all this take the form of the three Vs – volume, velocity and variety. The volume is huge and growing; the velocity is instant or real-time; and variety refers to unstructured data which can be linked to produce insights (OECD, 2015; McKinsey Global Institute, 2011).

Increasing returns in the use/reuse of big data

The returns from the use of big data (through increasingly powerful computers and telecommunications and ever improving software) have included: the development of new data-based goods and services; more efficient production or delivery processes; more efficient marketing (by providing targeted advertisements and personalized recommendations); faster and more efficient decision-making within existing practices; and more efficient and comprehensive search and research (e.g. by using new data-intensive methods for scientific exploration through the mining of vast, diverse data sets) (OECD, 2013, 327). Negative impacts associated with (rather than caused solely by) the utilization/exploitation of big data, however, have also emerged as serious threats to society and the individual, and will be discussed in the concluding section. Here we focus on two of the most dramatic examples of increasing returns in the use of big data --- increasing returns in machine learning from big data and increasing returns in the capture and use/reuse of big data by digital platforms.

Increasing returns in machine learning

‘Machine learning takes many different forms and goes by many different names: pattern recognition, statistical modeling, data mining, knowledge discovery, predictive analytics, data science, adaptive systems, self-organizing systems, and more. Each of these is used by different communities and has different associations…I use the term *machine learning* to refer broadly to all of them’ (Domingos, 2015, 8). Machine learning is related to big data in the following way. Instead of writing and using algorithms which specify detailed sequences of specific steps to be followed by the computer, the latest machine learning has become learning from examples, millions or billions of them, the reason being that many skills and tasks involve tacit knowledge where, by definition, we know more than we can tell, which means that what we cannot tell, we cannot instruct computers to do. Learning from examples then becomes the only way forward, and increasing returns refers to the fact that the more the system learns, over time, from gigantic and growing sets of examples (i.e. the training data also known as big data), the better or the more accurate the results, outcomes, predictions, insights, recommendations, etc. of machine learning become. Experts seem to agree that more data trumps better algorithms (Domingos, 2012, 6).

A few examples will make this clear. In language translation, IBM’s Candide project in the 1990s used ten years’ worth of Canadian parliamentary transcripts in French and English as training data; in 2006 Google used every translation it could find in the entire global Internet and more to train the computer system. Google’s translations were more accurate than those of Candide and other systems, and moreover by 2012 they covered more than 60 languages (Mayer-Schonberger and Cukier, 2013, 38). In chess, the world champion Garry Kasparov was defeated by IBM’s Deep Blue in 1997. ‘The reason computer chess programs play far better today than in the past is that… the systems have been fed more data. In fact, endgames, when six or fewer pieces are left on the chessboard, have been completely analyzed, and all possible moves have been represented in a massive table that fills more than a terabyte [1000 gigabytes or 1 million megabytes] of data. This enables chess computers to play the endgame flawlessly. No human will ever be able to outplay the system’ (Ibid, 36).

In the board game Go (which apparently is even more complex than chess), the world champion of Go, Lee Sedol, was defeated in 2016 by AlphaGo, built by Google subsidiary DeepMind. AlphaGo had been trained on huge libraries of Go matches between top players amassed over the game’s 2500- year history, and had also played millions of games against itself, using reinforcement learning to remember the moves that worked well (McAfee and Brynjolfsson, 2016). What has enabled modern machine learning has been the identification of massive troves of (potential) data, the huge increases in the power of computer hardware (following Moore's Law) and the paradigm of learning from (massive quantities of) examples. Where the data did not previously exist, it can be generated and captured through various kinds of sensors including mobile phones, and transmitted through the Internet of Things to giant data centers. Machine learning is apparently achieving increasing returns in the form of more accurate results, predictions, insights, recommendations, etc. in a number of specific application domains (chess, Go, self-drive cars, facial recognition, voice recognition, speech recognition, language translation, image tagging, automated financial trading, fraud detection, recommendations to customers from the likes of Amazon and Alibaba, etc.). This is 'narrow AI' (artificial intelligence).[xxx] Whether machine learning can ever completely substitute for human intelligence depends on whether it can achieve 'general AI,' the all-purpose technology that can do everything a human can (Lee, 2018, 10).

Increasing returns in the functioning of digital platforms

Digital platforms are businesses/markets that provide the infrastructure and rules that bring together producers and consumers (and also, in many cases, developers of new apps and advertisers) through online, IT-based exchanges and interactions (van Altsyne et al., 2016, 56-7). The two main types are matchmaking platforms (e.g. Amazon) and technology/innovation platforms (e.g. Microsoft). The functioning of digital platforms is one of the most powerful examples of increasing returns in the use of big data. The functioning of digital platforms gives rise to different types of virtuous circles (circular and cumulative causation), which are not mutually exclusive; the companies that own the platforms may realize several of the virtuous circles at the same time. (1) In the case of technology platforms (such as Windows, Linux, Android and iOS), in addition to network externalities,[xxxi] a growing number of users encourages more app developers to develop new apps for users of the platform; the more apps there are, the more existing users like to use their platform, and the more interested other people become in joining the bandwagon; growth in the number of users continues and accelerates, and the circle becomes cumulative. (2) An advertising platform attracts more users, which attracts more ads; the growth in ads and users increases the quantities of big data under the control of the platform, which improves the ability to target the ads to the users, which enhances the value of the ads and attracts *both* more users and more ads; and so forth (Iansiti and Lakhani, 2017, 90).[xxxii] (3) In the case of digital matchmakers, the larger the network becomes, the richer is the data set that can be used to find matches between supply and demand and the better the matches become (Van Alstyne et al., 2016, 58). This attracts more suppliers, while the data that has been automatically accumulated concerning each user (through click streams, traces of searches, records of past transactions, various forms of feedback, etc.) together with the use of powerful algorithms, enable personalized recommendations to be made to each user. This attracts more users, and so on. (4) The more feedback data is collected from users about the products and services offered by or through the platform, the more these products and services can be improved (through more effective machine learning), thus attracting more users from whom more data can be collected to grow the platform's stock of big data (Mayer-Schonberger and Ramge, 2018, 162-6).

Increasing returns in the functioning of the digital platforms means that the growth in the number of users and thereby of the amounts (and variety) of data collected can be used to further strengthen the competitive advantage of the platform in a self-reinforcing circle. If a digital platform gets ahead of rivals by being the first mover or by creative strategy or by chance, for whatever reason, increasing returns (defined in the relative sense) can operate to amplify this initial lead or advantage, which then grows exponentially relative to would-be rivals or entrants through a series of virtuous circles until the firm dominates or monopolizes the relevant market. Increasing returns, plus in some cases, customer lock-in or high switching costs, have resulted in the dominant positions of what is known as the Big Five: Amazon, Apple, Facebook, Google and Microsoft. Amazon has a share of more than 40 percent of online retailing revenues in the U.S. (Mayer-Schonberger and Ramge, 2018, 161). Google has 92 percent of the worldwide search market and Facebook has 70 percent of the social media market (Streitfeld, 2019). Windows (Microsoft) still dominates desktop PCs; Google (Android) and Apple (iOS) dominate their respective mobile phone operating systems and the apps that run on them; Facebook and Google dominate the Internet advertising business; and Amazon, Microsoft and Google dominate the "cloud" infrastructure (Manjoo, 2016).

Increasing returns, however, do not answer the question --- why can't rivals to any of the big five or would-be new entrants use big data to generate similar virtuous circles, albeit (initially) at a smaller scale? The answer is in two parts. First, access to and use of big data is perfectly *excludable* (this will be explained in the next section). Secondly, if a new entrant has to generate and collect its own big data from scratch, the cumulative advantage that any of the big five derives from its cumulative volume of big data is such that the new entrant would have no chance of surviving in the same market, let alone catching up.

Some concluding remarks

Public goods, private goods, common-pool resources and club (also known as toll) goods are differentiable from each other according to whether (or the degree to which) each is non-subtractable (this is, non-rivalrous) versus subtractable, and non-excludable[xxxiii] versus excludable. The hallmarks of a 'pure' public good are that it is perfectly non-subtractable (e.g. knowledge) and perfectly non-excludable (e.g. law and order, good weather).[xxxiv] Data, information and knowledge are all non-subtractable because they are intangible (to the extent that they are embodied in physical form, say in books, they lose their non-rivalry or non-subtractability: A cannot use a book if B is using it). Non-subtractability is one of foundational principles or forces that underpin some, if not all, mechanisms of increasing returns in relation to knowledge. Earlier in this paper we saw that non-subtractability of knowledge helped to give rise to increasing returns in the use or utilization or application of knowledge in general (see page 16) and that, in relation to the production of knowledge, increasing returns in Nonaka's SECI cycle are made possible by the non-subtractabity of knowledge (see page 12). The use of big data is also non-subtractable, which means that it can be used an infinite number of times or by an infinite number of agents without using it up (in other words, at zero marginal cost), thus enabling and facilitating continuous experimentation *ad infinitum* in data extraction and analysis, in data combination and recombination, and in data personalization and customization (Zuboff, 2015, 75).

When it comes to the degree of excludability (sometimes rendered as exclusivity) versus non-excludability, there is a crucial difference between big data and knowledge. While knowledge has a natural tendency to expand or spread and is non-excludable except where it is covered by intellectual property rights (IPRs) or kept secret from the public, the big data that is systematically collected from the users, stored and used/reused by the Big Five is kept under their exclusive control, even though legally speaking they do not own it.[xxxv] It is not only perfectly excludable, most of it is in fact excluded from the individuals (the users of the platforms) from whom the data is collected (usually without payment but often in return for 'free' services) and also excluded from competitors and would-be competitors, suppliers to the platforms, government and the general public ---- they are all given little access to the data. Excludability is the key to sustaining competitive advantage: if a firm or platform's knowledge, data, innovation or critical resource can be imitated (reproduced or replicated), the advantage (and monopoly rent and barriers to entry against other players) is likely to be competed away sooner rather than later. In the case of big data, the collected and stored data is kept for the sole exclusive use of the owners of the platforms; thus the increasing returns from the use of big data are appropriated by them, and are not available to would-be competitors and new entrants.[xxxvi]

Differences between knowledge and data, arising from the fact that knowledge includes tacit knowledge while there is no such thing as 'tacit data', can give rise to interesting policy implications. The spreading, combination and creation of new ideas, which often involve elements of tacit knowledge and/or knowledge which is so novel that it has not yet been codified, requires interactions and face-to-face contacts between knowledgeable and creative people who usually live in cities. This means that, to promote creativity and innovation, cities should be liveable, connected and convivial in terms of facilities and infrastructure, and expandable in terms of housing and workplaces (Haskel and Westlake, 2018, 215). The aggregation and use of big data do not give rise to clear policy implications because they do not require the same face-to-face interactions and do not create many jobs.

This paper would be incomplete if we failed to touch upon the negative impacts that are emerging from the ways in which big data is being used. These impacts, which are associated with rather than caused solely by big data, include the following: massive collection of personal data by the giant platforms without the targets' authorization in any meaniningful sense or (often) without the targets even being aware of the fact and certainly without compensation or renumeration for the private data that is used to generate revenues on the "other side" of the platforms; intrusions into and threats to privacy; massive, secret and (often) illegal surveillance of and spying on millions of citizens by national security agencies (as revealed by Edward Snowden in 2013 and after); the theft and illicit use of big data (e.g. the use of Facebook data by Cambridge Analytica to target and influence voters in 2018); the vulnerability of almost any large organization against hacking and theft of data and information; commercial espionage; sabotage through the use of viruses and malware to paralyse or wreck large computerized systems; the climate of fear and suspicion that all this engenders; the international mobility of 'intangible' assets such as big data and the increased scope for tax evasion/avoidance that this affords; and the possibility of Artificial Intelligence displacing labour to such an extent that calls are being made to seriously consider the possibility of introducing a national 'basic income' (Brynjolfsson and McAfee, 2016, 232-3) or even a national 'social investment stipend' for care work, community service and education (Lee, 2018, 220). Unless and until better safeguards, regulation and governance are introduced and put into effect, 'surveillance capitalism' (Zuboff, 2019), based on the unprecedented asymmetries in knowledge and power between the watchers and the watched, will continue to cast an all-pervasive and oppressive darkness on the functioning of democracy and capitalism and on human rights and human dignity. The collective struggle for better and more effective safeguards, regulation and governance, which promises to be long and arduous, is less about data or technology than about the social construction of the systems in which we live.

Notes

[i] For example, see OECD (1996).
[ii] The "residual", assumed to consist mainly of knowledge or technology, has been found by growth accounting exercises to account for more than two-thirds of long-term economic growth.
[iii] 'Whether it be the older literature on research and development or the modern New Growth Theory, the mainstream account runs along the following lines. Knowledge is produced privately using a sausage-machine called research and development that takes in inputs and gives off technological knowledge, which then immediately augments the production function for other goods' (Langlois, 1999, 249). This characterisation applies to Paul Romer's path-breaking paper (1990) in which knowledge and new designs are generated in the R&D department of firms, and the spillover of such knowledge to all other firms is reflected by including K (the total stock of knowledge in the economy) in the production function of the representative firm (see also Kurz, 2012, 95-97). The growing literature on 'non-R&D' sources of knowledge production and/or innovation (Barge-Gil et al, 2011; Lee and Walsh, 2016), together with the difficulty of drawing the line, within firms, between R&D, design, engineering, prototyping and scaling up from pilot plants (Freeman and Soete, 2009) should put question marks over such an approach.
[iv] Economic growth is quantitative, progress qualitative.
[v] According to Loasby, 'The division of labour is the primary means of increasing the division of knowledge, and thereby of promoting the growth of knowledge. Knowledge grows by division: each of us can increase our knowledge only by accepting limits on what we can know' (Loasby, 1999, 135). According to Metcalfe: 'The division of labour is a division of knowing and, moreover, the division of labour applies to the development of knowledge as well as to its application' (Metcalfe, 2014, 17).
[vi] Joan Robinson repeatedly pointed out that "a confusion between comparisons of imagined equilibrium positions and a process of accumulation going on through history" was "an error in methodology" on the part of neoclassical economists (Robinson, 1979, 58).
[vii] The knowledge management literature identifies at least four kinds of "knowledge processes" at the organizational level ---knowledge creation, knowledge application, knowledge integration and knowledge retention (Kraaijenbrink 2012). The sociology of knowledge distinguishes between processes of knowledge production; knowledge organization, dissemination-distribution, and storage-retrieval; and knowledge application (Holzner, 1979). The history of knowledge identifies at least 32 processes that can be grouped under the 4 main stages of knowledge gathering, analysing, disseminating and employing (Burke, 2016).
[viii] Pavitt (2005) identified the production of scientific and technological knowledge and the transformation of knowledge into working artefacts as two of the three key sub-processes in the process of innovation.
[ix] They are: the communication theory approach, the probabilistic approach, the modal approach, the systemic approach, the inferential approach, and the semantic approach.
[x] This was quite a departure from the tradition in epistemology, which defines knowledge as justified true belief.
xi In the modern age they include libraries, archives, Internet-related items such as optical fibre, copper-wire switches, routers, host computers and end-user workstations, bandwidth, free-space optics and wireless systems, etc.
[xii] Shapiro and Varian (1999, 3) defined *information* as anything that can be digitized (encoded as a stream of bits).
[xiii] The paper was originally a lecture, delivered as the Presidential Address to Section F of the British Association for the Advancement of Science, Glasgow, Sept. 10, 1928.
[xiv] Paul Romer's paper 'Endogenous Technological Change' was not published in the *Journal of Political Economy* until October 1990.
[xv] For a detailed critique of economies of scale and increasing returns to scale, see Chandra and Sandilands (2006, 200-202).
[xvi] It is for this reason that we do not adopt the expression 'economies of specialization and division of labour' put forward by Yang and Ng (1998, 8).
[xvii] One of the founders of economic growth theory wrote more than 70 years ago: 'Problems arising from a once-over change can, I believe, be satisfactorily handled by the apparatus of static theory. It is when we come to a steadily continuing change that we have to consider a different technique…Dynamics will specifically be concerned with the effects of continuing changes and with rates of change...' (Harrod, 1948, 7, 8).

[xviii] 'Something' here means anything, at the micro or macro level, that requires purposeful *doing* and encompasses work, actions, activities, tasks, production, operations, processes, routines, functions, investments, learning, projects, programmes, applications, adaptations, etc.
[xix] The Gospel of Matthew has a passage that says: 'For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath' (Matthew 25:29); the term 'Matthew effect' was introduced by sociologist Robert Merton to describe the reward and communication systems of science (Merton, 1968). I am grateful to Stuart Macdonald for pointing out this equivalence with the concept of increasing returns.
[xx] The law, named after Gordon Moore of Intel Corp., has been expressed in several different ways: that the number of transistors incorporated in an electronic chip would approximately double every two years, that the performance of microchips would double every 18 months, or that the price of integrated circuits halves as the number of transistors therein doubles. I am grateful to Stuart Macdonald for putting me right on this.
[xxi] Strictly speaking, negative feedback processes lead to a successive reduction in the amplitude of the deviations towards equilibrium, whereas vicious circles are the symmetrical opposite of virtuous circles.
[xxii] Division of labour and specialization are synonymous except in the case where the division of labour is increased to the point of greatly simplifying the job to be done, which nullifies the element of job-specific skills or specialization. See M. Morroni (1992).
[xxiii] See, inter alia, Weitzman (1998, 335) and Fagerberg (2005, 10).
[xxiv] In the unlikely case of a piece of knowledge being used by the same person who generated it in isolation from the world, one can say that use of that knowledge would involve a flow of knowledge from that person's conscious or unconscious memory to other parts of his or her brain and neuro-motor system.
[xxv] According to Amabile (1986, 1): 'Creativity is the production of novel and useful ideas in any domain**.** In order to be considered creative, a product or an idea must be different from what has been done before. (Few creativity theorists hold the strong position that a creative idea must be completely unique.) But the product or idea cannot be merely different for difference's sake; it must also be appropriate to the goal at hand, correct, valuable, or expressive of meaning. Innovation is the successful implementation of creative ideas within an organization.'
[xxvi] The use of information and data is also non-subtractable.
[xxvii] The measurement of GDP growth does not adequately reflect quality improvements because none of the price indexes used to separate real growth from price changes reflect improved quality (Samuelson and Nordhaus, 2010, 102).
[xxviii] There used to be a productivity paradox in the U.S., where between 1973 and 1995 labour productivity growth was poor (1.3% p.a.) despite the huge investments that had been made in Information Technology. This could have been due in part to mismeasurement of outputs and inputs, and also time lags in learning and adjustment (Brynjolfsson, 1993). "The impact of the computer revolution became apparent in the productivity statistics beginning around 1995. Having grown slowly during the 1973-1995 period, labour productivity then surged ahead at 2.6% per year from 1995 to 2008" (Samuelson and Nordhaus, 2010, 233).
[xxix] 'In 2017, 24 hyperscale firms operated 320 data centres with anywhere between thousands and millions of servers' (Zuboff, 2019, 501). Google's warehouse-sized data centres, spanning 15 locations, had in 2016 an estimated 2.5 million servers in 4 continents (Ibid, 188).
[xxx] Artificial intelligence can be defined as 'intelligence' demonstrated by machines rather than human beings or animals.
[xxxi] Network externality means that the value of connecting to a network depends on the number of other people already connected to it. A tenfold increase in the size of the network leads to a hundredfold increase in its value. See Shapiro and Varian (1999, 183-184).
[xxxii] It is worth noting that, although Google is primarily an (if not the) search engine, most of its revenues are derived from advertisements. In 2016, 89 percent of the revenues (over US$90 billions) of Google's parent company (Alphabet) derived from Google's targeted advertising programs (Zuboff, 2019, 93).
[xxxiii] Non-excludability of knowledge means that other people (people other than the owner, holder or user of knowledge and other than those who have paid to access and/or use it) cannot be prevented from accessing and/or using it.
[xxxiv] Thus, knowledge is a partial public good because, although non-subtractable, pieces of knowledge can be made excludable, to a greater or smaller extent, through intellectual property rights (IPRs), secrecy, and the difficulty of transferring and/or acquiring tacit knowledge.
[xxxv] The concept of ownership of knowledge is well-established in and through the law of IPRs, whereas ownership of data is not yet firmly established in law. The EU's General Data Protection Regulation (GDPR),

which came into effect in May 2018, is a first step towards establishing that personal data should be owned by the persons from whom the data is collected.

[xxxvi] In the context of the the big data gathered and used by the Big Five, it has been suggested that, since this poses an insuperable barrier to entry for would-be new entrants, regulators should consider imposing a compulsory sharing of a certain proportion of the dominant firms' big data; the data would be anonymized, and the percentage would rise with the market share of the dominant platform owner (Mayer-Schonberge and Ramge, 2018, 166-68). While we do not know (without further appropriate research) whether this would work (e.g. to increase real competition or new entry), it is almost certain that 'surveillance capitalism' would fight this tooth and nail because it goes against the core of their business models.

Disclosure statement

No potential conflict of interest was reported by the author

References

Amabile, T. (1986). Creativity and innovation in organizations. *Harvard Business School case no. 9-396-239*. January. Cambridge, Massachusetts: Harvard Business School.

Arthur, W. (1988). Self-reinforcing mechanisms in economics. In Anderson, P. et al. (eds), *The Economy As An Evolving Complex System* (pp. 9-32). Reading, Massachusetts: Addison-Wesley.

Arthur, W. (1996). Increasing returns and the new world of business. *Harvard Business Review*, July, vol. 74, no. 4, 100 - 119.

Barge-Gil, A., Nieto, M. and Santamaria, L. (2011). Hidden innovators: the role of non-R&D activities. *Technology Analysis and Strategic Management*, vol. 23, issue 4, 415-432.

Blitch, C. (1983). Allyn Young on increasing returns. *Journal of Post-Keynesian Economics,* spring, vol. 5, issue 3, 359-372.

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communication of the ACM,* December, vol. 36, no. 12, 67-77.

Brynjolfsson, E. and McAfee, A. (2016). *The Second Machine Age.* New York: Norton.

Burke, P. (2016). *What is the History of Knowledge?* Cambridge: Polity Press.

Chandra, R. and Sandilands, R. (2006). 'The role of pecuniary external economies and economies of scale in the theory of increasing returns'. *Review of Political Economy*, vol. 18, issue 2, 193-208.

Cimoli, M., Dosi, G., Nelson, R. and Stiglitz, J. (2009). Institutions and policies shaping industrial development. In Cimoli, M. et al. (eds), *Industrial Policy and Development* (pp. 20-38). New York: Oxford University Press.

Cohen, W. and Levinthal, D. (1990, March). Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly*, vol. 35, no. 1, 128-152.

Constantiou, I. and Kallinikos, J. (2015). 'New games, new rules: big data and the changing context of strategy'. *Journal of Information Technology,* vol. 30, issue 1, 44-57.

Currie, L. (1981). Allyn Young and the development of growth theory. *Journal of Economic Studies,* vol.8 no.1, 52 - 60.

Davenport, T. and Prusak, L. (1998). *Working Knowledge: How Organizations Manage What They Know.* Boston: Harvard Business School Press.

Domingos, P. (2012). 'A few useful things to know about machine learning'. *Communications of the ACM,* October, vol. 55 issue, 10, 1-10.

Domingos, P. (2015). *The Master Algorithm.* London: Allen Lane.

Fagerberg, J. (2005). 'Innovation, a guide to the literature'. In Fagerberg, J., Mowery, D. and Nelson, R. (eds), *The Oxford Handbook of Innovation* (pp. 1-28). Oxford: Oxford University Press.

Floridi, L. (2003). 'Two approaches to the philosophy of information'. *Minds and Machines,* vol. 13, issue 4, 459-69.

Foray, D. (2004). *The Economics of Knowledge.* Cambridge, Massachusetts.: MIT Press.

Freeman, C. and Soete, L. (2009). Developing science, technology and innovation indicators: what we can learn from the past. *Research Policy*, vol. 38, issue 4, 583-589.

Gardner, H. (2008). *Five Minds for the Future.* Boston: Harvard Business School Press.

Gitelman, L. (2013). *"Raw Data" Is An Oxymoron.* Cambridge, Massachusetts: MIT Press.
Harrod, R. (1948). *Towards a Dynamic Economics.* London: Macmillan.
Haskel, J. and Westlake, S. (2018). *Capitalism Without Capital, The Rise of the Intangible Economy.* Princeton: Princeton University Press.
Hess, C. and Ostrom, E. (ed). (2011). *Understanding Knowledge as a Commons.* Cambridge, Massachusetts: MIT Press.
Holzner, B. and Marx, J. (1979). *Knowledge Application: The Knowledge System in Society.* Boston: Allyn and Bacon.
Hu, Y. (1995). The international transferability of the firm's advantages. *California Management Review*, summer, vol. 37 no. 4, 73 - 88.
Iansiti, M. and Lakhani, K. (2017). Managing our hub economy. *Harvard Business Review,* Sept., vol. 95, issue 5, 84-92.
Kurz, H. (2012). On the growth of knowledge about the role of knowledge in economic growth: a critical assessment of recent literature on growth theory. In H. Kurz, *Innovation, Knowledge and Growth: Adam Smith, Schumpeter and the Moderns* (pp. 87 - 115). Abingdon: Routledge.
Langlois, R. (1999). Scale, scope and the reuse of knowledge. In Dow, S. and Earl, P. (eds.), *Economic Organization and Economic Knowledge, Essays in honour of Brian J. Loasby,* Vol. I (Chapter 13). Cheltenham: Edward Elgar.
Lee, K. (2018). *AI Superpowers: China, Silicon Valley, and The New World Order.* New York: Houghton Mifflin Harcourt.
Lee, Y. and Walsh, J. (2016). Inventing while you work: Knowledge, non-R&D learning and innovation. *Research Policy*, vol. 45, issue 1, 345-359.
Loasby, B. (1989). *The Mind and Method of the Economist.* Cheltenham: Edward Elgar.
Loasby, B. (1998). 'The concept of capabilities'. In Foss, N. and Loasby. B (eds). *Economic Organization, Capabilities and Co-ordination* (pp.163-182). London: Routledge.
Loasby, B. (1999). *Knowledge, Institutions and Evolution in Economics.* Abingdon: Routledge.
Macdonald, S. (1998). *Information for Innovation, Managing Change From An Information Perspective.* Oxford: Oxford University Press.
Machlup, F. (1980). *Knowledge: Its Creation, Distribution and Economic Significance, Volume I: Knowledge and Knowledge Production.* Princeton: Princeton University Press.
Machlup, F. (1983). 'Semantic quirks in studies of information'. In Machlup, F. and Mansfield, U. (eds). *The Study of Information: Interdisciplinary Message* (pp. 641- 662). New York: Wiley.
Manjoo, F. (2016). Frightful 5's domination appears to be unshakable. *New York Times*, Jan. 21.
Marshall, A. (1895). *Principles of Economics, 3rd edition.* London: Macmillan.
Marshall, A. (1920). *Principles of Economics, 8th edition.* London: Macmillan.
Mayer-Schonberger, V. and Cukier, K. (2013). *Big Data.* London: John Murray.
Mayer-Schonberger, V. and Ramge, T. (2018). *Reinventing Capitalism in the Age of Big Data.* New York: Basic Books.
McAfee, A. and Brynjolfsson, E. (2016). A computer wins by learning like humans. *New York Times, 16 March*.
McKinsey Global Institute. (2011). *Big Data: The Next Frontier for Innovation, Competition and Productivity.* New York: McKinsey.
Mehrling, P. and Sandilands, R. (1999). *Money and Growth, Selected papers of Allyn Abbott Young.* London: Routledge.
Merton, R. (1968). The Matthew effect in science. *Science,* vol. 159, no. 3810, 56-63.
Metcalfe, S. (2014). Capitalism and evolution. *Journal of Evolutionary Economics,* vol. 24, issue 1, 11-34.
Morroni, M. (1992). *Production Process and Technical Change.* Cambridge: Cambridge Unversity Press.
Nelson, R. (1959). The simple economics of basic scientific research. *Journal of Political Economy*. June, vol 67, no. 3, 297-306.
Nonaka, I. and Takeuchi, H. (1995). *The Knowledge-Creating Company.* New York: Oxford University Press.
Nonaka, I. et al. (2008). 'The characteristics of knowledge'. In Nonaka, I. et al. *A Process Theory of the Knowledge-Based Firm* (pp. 6-15). London: Palgrave Macmillan.
OECD. (1996). *The Knowledge Economy.* Paris: OECD.
OECD. (2013). *Supporting Investment in Knowledge Capital, Growth and Innovation.* Paris: OECD.
OECD. (2015). *Data-Driven Innovation: Big Data for Growth and Well-Being.* Paris: OECD.
Pavitt, K. (2000). *Research and development, SEWPS no. 44.* Falmer: SPRU.

Pavitt, K. (2005). Innovation processes. In Fagerberg, J., Mowery, D. and Nelson, R. (eds), *The Oxford Handbook of Innovation* (pp. 86-114). Oxford: Oxford University Press.

Polanyi, M. (1958). *Personal Knowledge.* London: Routledge.

Polanyi, M. (1966). *The Tacit Dimension.* Chicago: University of Chicago Press.

Rigney, R. (2010). *The Matthew Effect, How Advantage Begets Further Advantage.* New York: Columbia University Press.

Robinson, J. (1979). *Collected Economic Papers,* vol. 5. Oxford: Blackwell.

Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, October, S71-S102.

Rosenberg, N. (1965). Adam Smith on the division of labour. *Economica,* May, vol. 32, no. 12, 127 - 139.

Rowley, J. (2007). 'The wisdom hierarchy: representations of the DIWK hierarchy'. *Journal of Information Science,* vol. 33 no. 2, 163-180.

Samuelson, P. and Nordhaus, W. (2010). *Macroeconomics.* New York: McGraw-Hill.

Sandilands, R. (2000). Perspectives on Allyn Young in theories of endogenous growth. *Journal of the History of Economic Thought*, vol. 22, issue 3, 309-328.

Schultz, T. (1993). *Origins of Increasing Returns.* Oxford: Blackwell.

Shapiro, C. and Varian, H. (1999). *Information Rules.* Boston: HBS Press.

Sloman, S. and Fernbach, P. (2017). *The Knowledge Illusion: Why We Never Think Alone.* London: Macmillan.

Smith, A. ([1776] 1976). *An Inquiry into the Nature and Causes of the Wealth of Nations.* Chicago: University of Chicago Press.

Smolucha, F. (1992). 'A reconstruction of Vygotsky's theory of creativity'. *Creativity Research Journal*, vol. 5, issue 1, 49-67.

Streitfeld, D. (2019) 'Taking on the tech giants,' in *International New York Times,* 24 June 2019.

Van Alstyne, M., Parker, G. and Choudary, S. (2016). Pipelines, platforms and the new rules of strategy.

Harvard Business Review, April, vol. 94, issue 4, 54-62.

Vygotsky, L. (2004). Imagination and creativity in childhood. *Journal of Russian and East European Psychology*, Jan.-Feb., vol. 42, no. 1, 7-97.

Weitzman, M. (1998). Recombinant growth. *Quarterly Journal of Economics*, vo. 113, issue 2, 331-360.

Witt, U. (2000). Changing cognitive frames, changing organizational forms. *Industrial and Corporate Change,* vol. 9, no.4, 733-755.

Yang, X. and Ng, S. (1998). 'Specialization and division of labour: a survey'. In Arrow, K. et al. (eds). *Increasing Returns and Economic Analysis* (pp. 3-63)*.* London: Macmillan.

Young, A. (1928). Increasing returns and economic progress. *The Economic Journal,* Dec., Vol. 38, no. 152, 527-542.

Young, A. (1929). The sources of wealth: The necessary parts played by land, capital and labor in the production of wealth. In *The Book of Popular Science*. New York: The Grolier Society (pp.3503-3513). Reprinted in Mehrling and Sandilands, 1999, 231-238.

Zins, C. (2007). 'Conceptual approaches for defining data, information, and knowledge'. *Journal of the American Society for Information Science and Technology*, vol. 58, issue 4, 479-493.

Zuboff, S. (2015). 'Big other: surveillance capitalism and the prospects of an information civilization'. *Journal of Information Technology,* vol. 30, issue 1, 75-89.

Zuboff, S. (2019). *The Age of Surveillance Capitalism.* London: Profile Books.

Zukerfeld, M. (2017). Typologies of knowledge: a reexamination from the perspective of cognitive materialism. *Prometheus,* vol 35, no 1, 3-20.

Recent papers in the SPRU Working Paper Series:

July

Eco-Innovation and Firm Growth in the Circular Economy: Evidence from European SMEs. Pelin Demirel and Gamze Ozturk Danisman

June

The Rush for Patents in the Fourth Industrial Revolution: An Exploration of Patenting Activity at the European Patent Office. Mario Benassi, Elena Grinza and Francesco Rentocchini.

Power and Export Sophistication in Buyer-Supplier Relationships: Insights from Colombian Customs Data. Filippo Bontadini.

Trade Specialisation and Performance in Global Value Chains. Filippo Bontadini.

DARE to be Different? Applying Diversity Heuristics to the Analysis of Collaborative Research. Frédérique Bone, Michael M. Hopkins, Ismael Ràfols, Jordi Molas-Gallart, Puay Tang, Gail Davey and Antony M. Carr.

May

Multiple Regime Interactions, Conversion, and South Africa's Liquefied Natural Gas. Marie Blanche Ting.

February

Specialisation, Diversification and the Ladder of Green Technology Development. François Perruchas, Davide Consoli and Nicolò Barbieri.

Suggested citation:

Yao-Su Hu (2019). The Impact of Increasing Returns on Knowledge and Big Data: From Adam Smith and Allyn Young to the Age of Machine Learning and Digital Platforms. SPRU Working Paper Series (SWPS), 2019-14: 1-27. ISSN 2057-6668. Available at: www.sussex.ac.uk/spru/swps2019-14

Science Policy Research Unit
University of Sussex, Falmer
Brighton BN1 9SL
United Kingdom

SPRU website: **www.sussex.ac.uk/business-school/spru**
SWPS website: **www.sussex.ac.uk/business-school/spru/research/swps**
Twitter: **@spru**