

Guan, Xingmiao; Qin, Xingfang

Article

The data factor's dual attribute and its interaction effects

China Political Economy (CPE)

Provided in Cooperation with:

School of Economics, Nanjing University

Suggested Citation: Guan, Xingmiao; Qin, Xingfang (2022) : The data factor's dual attribute and its interaction effects, China Political Economy (CPE), ISSN 2516-1652, Emerald, Leeds, Vol. 5, Iss. 1, pp. 40-51,
<https://doi.org/10.1108/CPE-09-2022-0009>

This Version is available at:

<https://hdl.handle.net/10419/319449>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

The data factor's dual attribute and its interaction effects

Xingmiao Guan

School of Marxism, Yangzhou University, Yangzhou, China, and

Xingfang Qin

Business School, Yangzhou University, Yangzhou, China

Abstract

Purpose – Data has become a factor of production. This occurs when history enters the era of big data, in which technologies such as artificial intelligence, cloud computing and blockchain are used to collect, manipulate, mine and process data. Data is a special product of labor, a sub-derivative of other production factors.

Design/methodology/approach – The data factor has a dual attribute: being physical (technical) and social. The social attribute of the data factor can not only materialize the technical attribute but also amplify it. In other words, the data has a multiplication effect on the allocation efficiency of other production factors. The social attribute of the data is brought out via the technical attribute as the medium. From a technical perspective, this medium is strongly adhesive, and after being bonded with other factors of production, it will only lead to a physical reaction and not change the nature of other factors.

Findings – However, once these two attributes interact with each other, especially when data is combined with capital, the most adhesive factor in the market economy, a series of new social relations will then be produced based on the technical attribute, resulting in significant adjustments in social relations, involving both positive and negative externalities.

Originality/value – Therefore, to get a scientific understanding of the dual attribute and its interaction effects on the data factor, it is necessary to take the following steps. We should promote institutional design that amplifies the positive externality, with a focus on facilitating public data sharing and improving the value of commercial data development. Also, we need to strengthen institutional arrangements that prevent and control the negative externality by emphasizing data supervision based on data types and levels as well as the rule of law.

Keywords Data factor, Special product of labor, Technical attribute, Social attribute, The interaction effect

Paper type Research paper

The Fourth Plenary Session of the Nineteenth Central Committee of the Communist Party of China (CPC) officially included data in the list of factors of production and put forward the basic principle and direction of market-oriented allocation in the “Opinions of the CPC Central Committee and the State Council on Constructing a More Complete Market-oriented Allocation System and Mechanism for Factors.” Compared with practice, academic research on the data factor is relatively backward, especially on the attributes of data. Researchers mainly focused on the technical attribute and the innovation and creation function of data with little attention to the social attribute of data and the externality arising from the interaction between social and technical attributes. This is a biased misinterpretation of the data factor and is likely to lead to deviations in the utilization of data. Thus, it is necessary to



make an in-depth analysis from the perspective of Marxist political economy to provide theoretical support for scientific utilization of the data factor.

Data factor's
dual attribute

1. Special product of labor in the big data era

Quantity and quality have coexisted ever since the existence of materials in the world. However, the signals reflecting a state of motion of objective things were left unprocessed and unexplained until the emergence of human society, where people gradually recorded the signals in forms like text, numerals, facts and images when they transformed the objective world. These initial impressions or the most primitive records perceived by the brain, sensory organs or observation instruments are called "data." However, as in the early period of human society marked by self-sufficient, small production activities, the production range was extremely limited and the productivity was very low, data then was very simple, which did not require complicated calculation and analysis and played a merely auxiliary role in human production and life. Even when human society evolved to the commodity economy, the importance of data continued to rise with ever-growing quantities and types of data and continually expanding scale and scope of exchange of commodities, though commodity producers and operators could record and measure their economic activity data precisely, those data were isolated, scattered, disorderly and of very low value for the whole society.

When individual producers and operators were facing the whole market, they were subject to not only data deficiency but also data asymmetry among subjects of economic activity. They had to pay high search costs to get relevant data. For scientific decision-making, statistical analysis became a professional activity to win the market. People continually collected, manipulated and collated data, extracting valuable information and forming conclusions helpful to their business. However, the statistical analysis was mainly about speculating on the real world with small quantities of sample data, judging the overall situation with average and predicting the future based on mathematical models. No matter how advanced the statistical means and statistical methods were (including computer-aided analysis) to improve efficiency, people still could not fundamentally solve problems including data deficiency, data asymmetry, data distortion and data transmission delay. Although data analysis was indispensable for scientific decision-making, and the position of data has significantly improved, data was still the source or object of statistical analysis rather than an independent factor of production then.

The time that data truly became an independent factor of production was when human society entered the era of big data, in which modern technologies, including the Internet, artificial intelligence, cloud computing and blockchain, could be used to collect and treat mass data. According to the definition from [McKinsey Global Institute \(2011\)](#), big data refers to a collection of data whose scale greatly exceeds the capacity of traditional database software tools in terms of the acquisition, storage, management and analysis of data and is featured with massive data size, quick data transmission, diverse data types and low-value density. The transition from traditional data to big data resulted from the development of modern information communication technology and went through at least three stages. The first stage was extensive use of databases in operating systems, including production, sales and diagnosis and treatment, when data was generated during operating activities and recorded in databases passively. The second stage was the birth of the Internet. In particular, with the emergence of new social networks represented by blogs, Weibo and WeChat and new mobile devices represented by smartphones and tablet computers, data grew explosively, and the generation of data was proactive. The third stage was the extensive use of the perception system, which resulted in an explosive growth of data and finally led to the birth of big data and the automatic generation of data.

The emergence of big data was one of the causes of data becoming an independent factor of production. To be more specific, big data provides abundant original material for data as a factor of production. These massive items of data from various sources in different forms that contain different information are likely to be integrated and analyzed to identify new knowledge that could hardly be identified from traditional data and create new value thereby. What is more important, innovations in modern technologies have created the conditions for the gathering, manipulation, mining and treatment of big data into useful information, experience and knowledge, as well as the discovery of the laws, where the information perception and collection terminal provides a means to collect mass data; cloud computing technology provides a means for big data analysis; and blockchain technology provides a means to integrate various kinds of data and information. Compared with traditional data analysis, big data analysis can complete many tasks that traditional data analyses fail to complete within a tolerable time range. In particular, with the help of modern tools, big data analysis can capture, manage, manipulate and treat various full data to establish interrelations among data and generate texts that can answer specific problems and information that is interpreted to have a certain meaning in digital, factual, graphic and other forms, revealing a certain causal relationship implied in the facts and answering questions like who, what, where and when. Unlike traditional data, big data undergoing special treatment is an information asset that enables stronger decision-making, insight discovery and process optimization and thus becomes an essential factor of production in the new technological revolution and the new economy.

The historical evolution of data into a factor of production reveals that the data factor is a historical category and is essentially a special product of labor. On the one hand, unlike traditional data, the data factor is a product of labor resulting from people's collecting, manipulating, mining and treating big data through modern technologies and tools. Like an ordinary product of labor, the data factor is the result of concrete labor and has the use value. In a market economy, the data factor available for exchange is also a commodity and has a certain value. On the other hand, the data factor is not an ordinary product of labor or commodity and has its specificity. The specificity is not only reflected in the product's formation process and its form or value but mainly in the use value of the product or commodity. In the system of factors of production, labor and land are primary factors (basic factors) for the production of wealth, and the technological factor is essentially a factor of production derived during people's exploration of improving labor productivity and resource utilization efficiency. Capital is the product of a commodity economy or say market economy and is also a derivative factor, while the data factor is a sub-derivative factor derived from the interaction between primary factors including labor and land and derivative factors including technology and capital when socialized production develops into the era of big data. This sub-derivativeness determines that the data factor can neither exist independently of the functions of other factors of production nor play its own role independently of other factors of production. Only the extensive and in-depth integration of data with other factors of production can bring the use value of the data factor into full play and prove the value of the data factor.

The data factor can be divided from multiple perspectives. First, from the technological perspective, the data factor can be divided into structured, semi-structured and non-structured data. Second, from the perspective of data source, the data factor can be mainly divided into the following categories: traditional enterprise data, such as consumer data generated by the customer relationship management system, enterprise resources planning data and inventory data and account data; data generated on smart devices, mainly including machine and sensor data (such as industrial sensor or equipment log of smart meters, smart temperature controllers and smart instruments and data transmitted automatically by Internet-connected home appliances to central servers); individual behavioral data, such as

various data generated on App, blog, Wikipedia and other social media of mobile devices like smartphone and tablet, including individual transaction data, personal information and events in the status report; transaction data, including POS and e-commerce shopping data. Third, from the value orientation of utilizing the data factor, the data factor can be divided into commercial data and public data. Commercial data is data collected, mined, manipulated and utilized by various market entities through their data platforms for making profits mainly, which is a privately-owned product by nature. And public data is mainly data collected, manipulated and utilized by countries, governments and other non-profit organizations on their data platforms and is a public-owned product by nature. The disciplines related to natural sciences pay focus on the first and second perspectives, while the political economy attaches greater importance to the second and third perspectives.

2. Technical attribute materialized by virtue of social attribute and its multiplication effect

Since the data factor is a historical category and a special product of labor, it has a dual attribute – natural attribute or say, technical attribute, i.e. usefulness (hereinafter referred to as “technical attribute” or “usefulness”), and also social attribute, i.e. involving complicated interest relationships, of which the main is economic interest relationship. These two attributes are closely related and coexist within the data factor. For the duality of the data factor, natural sciences-related disciplines pay more attention to the technical attribute of the data factor; in comparison, the political economy pays more attention to the social attribute of the data factor, which is determined by the attributes and tasks of different disciplines. To date, most research in the fields of both natural sciences and social sciences, including political economy, focus more on the technical attribute of the data factor while there are obviously relatively few studies focusing on the social attribute of the data factor or analyzing technical attribute connecting with the social attribute. From the perspective of utilizing the data factor scientifically, the Marxist political economy must express its stance regarding it clearly.

Undoubtedly, the data factor is firstly a technological factor and has the technical attribute. Such technical attribute is endogenous in the process of data becoming a factor of production and does not vary with any difference in any social system and institutional background in the era of big data. Since the new-generation information communication technology produced new forms of application, such as mobile Internet, Internet of Things (IoT), social network, digital home and e-commerce, those applications have been continuously generating massive fragments of data, whose unit of measurement has evolved from Byte, KB, MB, GB and TB to PB, EB, ZB and YB and is now shifting to BB, NB and DB. However, those big data were not yet a ready-made factor of production, and their use value or value utilization density was very low. People can hardly collect or treat such data through traditional manual methods or even a single computer, not to mention find the intrinsic rules. All this can be carried out only on a big data platform, of which the basic framework consists of information resource management, the collection, transmission, storage, treatment and presentation of data, and system monitoring (Shidian Data, 2020), adopting distributed computing architecture and based on cloud and blockchain technologies, such as distributed processing of cloud computing, distributed database, cloud storage and virtualization technology (Zhai, 2020). Thereby, data fragments become useful information, and the role of data as a factor of production will be truly fulfilled after the useful information is applied to various applications through exchange or spread. These analyses show that the technical attribute of the data factor includes but is not limited to the following features: Data is not a natural factor of production but will certainly be useful once it becomes a factor of production. To enable data to be a factor of production depends on

technologies and methods of acquiring, manipulating, mining and treating big data. Those technologies and methods determine the quantity of the use value of the data factor and the quality of the use value of data. The data factor is a massive, dynamic, ever-changing data collection and is innately plural rather than singular.

As a special product of labor, the data factor's technical attribute cannot exist independently, which must coexist and interact with the social attribute. On the one hand, during the process when data becomes a factor of production, i.e. links including the collection, manipulation and treatment of data, the internal motivation of people to pursue the maximization of economic interest or social interest is the main driving force for data becoming a factor of production. During the data collection link, specific technologies and methods for the construction of big data platforms and data collection mainly involve the technical attribute of the data factor. However, what is the purpose of building the data platform, who should build it, who should be served who is entitled to collect data and what data can be collected (Tang, 2021) are complicated social problems that reflect the social attribute of the data factor and determine whether the data platform and the technical attribute of the data factor therein can be materialized, as well as the nature and utilization direction of the data factor. A public data platform and the data factor on this platform pursue the maximization of social interest and reflect public benefit and sharing, while a commercial data platform and the data factor on that platform pursue the maximization of profits and reflect private nature and for-profit purpose.

Also, the object, emphasis and scope of data collection exhibit technical differences driven by different interest relationships. In the links of data manipulating, mining and treatment, technologies and methods for data analysis and mining belong to technical problems. However, the treatment and presentation of the data also involve the issue of value orientation. Algorithms are a good case in point for it. The algorithm is a technical problem but also reflects complicated social relations in practice. In the era of big data, data platform companies can track user behaviors in real-time, including card swiping, web search, positioning and likes, know about users' emotional fluctuations and behaviors, and customize and continuously push commodities, services, reading material pertinent to particular users through further analyses. This will influence users' economic decision-making and their value judgments and political preferences, involving not only users' privacy protection but also the relation coordination of related interest subjects and even various safety issues.

On the other hand, the social attribute of the data factor is more fully demonstrated in the links of data exchange and utilization after data becomes a factor of production. The technical and social attributes of the data factor interact, and the technical attribute is materialized by virtue of the social attribute. At the link of data factor exchange, data exchange is also the exchange of interests. Such exchange needs to comply with the exchange of equivalents, the basic law of market exchange. It is an exchange of equal rights between the supply and demand of the data factor. The usefulness of the data factor cannot be truly reflected unless the foregoing exchange is successful. Otherwise, there is no difference between the data factor and data fragments. A successful exchange not only calls for effective matching of supply and demand for the data factor (Wang *et al.*, 2020) but also calls for a systematic, well-regulated data factor market and attribution of data ownership, including the definition of data property ownership before data entry into the market (Wang, 2020b) and protection of property rights after data exchange. Compared with other factors of production or industrial products, the data factor is highly replicable, which is one of its weaknesses. If the property rights on data are not clear, or data protection is inadequate, and data products are not subjected to classified or hierarchical management according to attributes, characteristics, quantity, quality, format, importance, sensitivity and other factors of data products, great disputes over interests will occur, thus suppressing the development of the data factor hugely. In real life, the main reason why numerous data cannot be integrated, treated and

shared is complicated interest relationships among regions, departments, enterprises and individuals rather than technical limitations. It indicates that the technical attribute or usefulness of the data factor cannot be materialized if social interest relationships within the data factor are not clarified.

In the link of data utilization, the combined action of the external competitive pressure and the social attribute of the data factor – people's internal motivation of pursuing the maximization of economic or social interests through the data factor not only materializes the technical attribute of the data factor but also triggers an amplification effect, which is mainly reflected in the multiplication effect of the data factor on the allocation efficiency of other factors of production. The multiplication effect is materialized through the following two approaches. One is directly derived from the data factor. That is, once the usefulness of the data factor is socially recognized, strong profit incentives will catalyze the rise of new technologies, products, services and business modes oriented toward the data market. For instance, in the field of hardware and integrated devices, the data factor will promote the development of chip and storage industries and then drive the birth of markets for integrated data storage processing servers, memory computing, etc. In terms of software and service, the data factor will trigger the development of technologies for rapid data treatment and analysis and data mining as well as software products. The other approach is indirect. To be specific, the owners of traditional factors of production combine the data factor with traditional factors of production, promoting the intellectualization of business and industry development and thereby multiplying the allocation efficiency of traditional factors of production, based on the potential of the data factor utilization to fulfill the goal of maximizing social interests or economic interests (Shi and Deng, 2020). For instance, in the agriculture and manufacturing industries, the combination of the data factor with traditional factors of production facilitates the transition from traditional agriculture to digital agriculture and from a traditional factory to a smart factory. And farmers and enterprises are empowered to respond to the market demand more effectively, as they can decide what to produce, how much to produce and how to produce the goods based on big data and select or adjust their production and operation modes in line with market trends. For business service, the combination of the data factor with traditional production factors provides support for decision-making in terms of accurately identifying target markets and customer preferences and selecting appropriate methods of circulation service and marketing strategies. For public services, the combination of the data factor with traditional factors advances the transmission of urban traffic management, security administration and community management towards intellectualization and promotes the development of smart urban transport, smart elderly care service, smart social management and smart medical industries or businesses.

The global outbreak of the COVID-19 pandemic in 2020 was a human catastrophe and a big test for the multiplication effect of the data factor. During this test, Chinese governments and social administration institutions (especially medical authorities) at all levels conducted pandemic monitoring and analysis, origin tracing, patient tracking and community management accurately and efficiently with the help of big data. Scientific researchers utilized big data to accelerate virus detection and diagnosis and the research and development of new vaccines. Enterprises and public institutions strengthened the management and application of data and ensured the normal, orderly operation of work and study during the pandemic through long-distance education, video conference, online ordering and so on. This partly explains why China controlled the pandemic successfully and was the only economy reporting positive growth among major global economies in 2020. In this sense, the data factor provides a new driver for high-quality development.

From this point of view, the technical attribute of the data factor cannot exist independently of the social attribute. It is necessary to highlight the role of the data factor as a

new driver for high-quality development in the era of big data, but this does not mean focusing on the technical attribute of the data factor while neglecting the social attribute. From the theoretical view, the technical attribute of the data factor is meaningless without the social attribute of the data factor. In practice, only when the conflicts or contradictions of social interests during the formation and utilization of the data factor have been solved, and the role of benefit motivation, competition restraint and government adjustment have been given full play, that the potential usefulness of the data factor can turn into practical usefulness (Xie *et al.*, 2020), bringing the multiplication effect on the allocation efficiency of other factors of production.

3. Social attribute catalyzed with technical attribute and its externality

The social attribute of the data factor is the manifestation of specific methods, rules, orders and other relations on the collection, manipulation, treatment, exchange, distribution and utilization of data; in other words, various social relations are dependent on the technical attribute or catalyzed with the technical attribute as a medium during the materialization of the technical attribute. In this sense, the social attribute of the data factor is determined and achieved by the technical attribute. Compared with other factors of production, as the technical attribute of the data factor has its specificity – the usefulness of the data factor occurs only when the data factor interacts and is bonded with other factors of production. And once the data factor is effectively bonded with other factors of production, a multiplication effect will be produced on the allocation efficiency of other factors of production. Therefore, this specificity catalyzes new characteristics of the social attribute of the data factor. The task of the political economy is not only to reveal the law that social attribute is determined by the technical attribute of the data factor but also to explore new characteristics and rules of the social attribute.

In the system of factors of production, factors of production are bonded in two forms. The first one is that a factor of production proactively bonds with another or several other factors of production and then enters into the production process. In the market economy, regarding the momentum and ability of a factor of production to bond with other factors, the capital factor has the highest bonding strength among all factors and can bond with almost all other factors of production. Before bonding with capital, various factors of production have their independence and own productivity, such as labor, land or resource and technology. However, after bonding with capital, those factors of production become a capital factor, and their productivity all becomes the productivity of capital. Karl Marx indicates that “all strength of labor is manifested as the strength of capital” (Marx, 2009b), “. . . like social labor productivity developed historically, labor productivity restricted by natural conditions is also manifested as the productivity of capital integrating with labor” (Marx, 2009a). Likewise, technological productivity is also converted into capital productivity, thus contributing to higher productivity of capital. According to Karl Marx (Marx, 2009b), “science and technology cause functional capital to have an expansion capability independent of its certain amount. Meanwhile, this expansion capability reacts upon the portion of the original capital that has entered the update stage. Capital absorbs the social progress materialized behind its old form in a new form with no price.” The role of capital therein is not merely bonding the factors of production scattered in the hands of owners and investing them into the production process, and more importantly, causing these factors of production to react chemically and thereby creating higher productivity. In this sense, both Marx and Engels argued that capital has a great civilizing role, and thus the capitalist society has more remarkable historical progress than any preceding society.

The second form is that a factor of production is bonded with another or several other factors of production. Within a considerable period before the birth of the new technological

revolution and the new economy, apart from two original factors of production – labor and land, the technological factor was the factor most bonded with other factors of production. In the era of big data, the data factor will become the most factor of production bonded with other factors of production. It can be bonded with the labor factor to improve labor productivity. Land owners, contractors and operators can bond the data factor with land to promote the intensive, refined, scientific utilization of land and improve land productivity. Owners of capital and technological factors can bond the data with capital and technologies to improve the productivity of capital and technology. Besides, the data factor can be bonded with entrepreneurs' ability to improve the management and institutional innovation efficiency of enterprises. Obviously, different from the bond between capital and other factors of production, that is, the productivity produced by capital in the purchase of other factors of production, all of which are inputted into the production process, is the productivity of capital, the bond between the data factor and other factors of production does not change the nature of other factors of production and their productivity after the data factor is bonded with other production factors organically by owners of other production factors through various approaches such as data sharing, self-study and purchase. In this sense, capital is a chemical bonding agent of other factors of production, and data is a physical bonding agent of other factors of production.

Under the precondition that the data factor is not bonded with capital, the technical feature of the data factor as a physical bonding agent will catalyze a series of new social relations. At the microcosmic level, the new technical feature of the data factor promotes the direct bond between data factor and labor, technology or knowledge, promoting social activities of entrepreneurship and innovation and the emergence of new forms of employment, like network marketing, live streaming and network technology service. This makes it possible for people to realize their interests with knowledge and skills in a market economy and jump out of the paradigm of realizing individual interests by either owning capital or selling their labor, which is a new relationship between labor and knowledge, labor and technology and among laborers. At the macrocosmic level, public data resource is a national basic strategic resource in the era of big data and is also an indicator for effective adjustment of the national economy and society. The wide application of data can overcome the problems of inaccuracy, delay and incompleteness of information existing in macro decision-making to the largest extent, which not only improves the possibility of a more conscious application of the law of proportional distribution of social labor in the operation of the national economy but also provides the possibility for the government to observe, analyze and grasp the occurrence, development and evolution trends of social conflicts in a more comprehensive, accurate and timely manner, thus forming a new type of plan–market relationship and government–enterprise relationship.

Under the condition that the data factor is bonded with capital, the combination of data as the factor with the most demand to be bonded with other factors of production and capital as the factor having the strongest power to bond other factors of production will change the nature of the data factor and its productivity fundamentally. In other words, data will belong to capital, and its productivity will be converted into the productivity of capital. In this condition, the social attribute determined by the technical attribute of the data factor will have new changes. On the one hand, driven by profit maximization and competitive advantages, capital will introduce more advanced technologies and facilities as well as high-level professional and technical professionals, build higher-level data platforms, continuously improve the capabilities of collecting, manipulating, mining and treating data and further promote scale, specialization, integration and service precision of data elements. Thereby, the technical attribute and usefulness of the data factor, especially the tendency to be bonded and multiplication effect on the allocation efficiency of other factors of production, will be continuously amplified. On the other hand, after the data factor is bonded with capital, the

nature of the data factor will change accordingly. To be more specific, the data factor is converted into digital capital or network capital usually, and the expansion force of the data force will be converted into the expansion force of digital capital or network capital. Since productivity decides relations of production, the expansion force of the combination of data and capital in the productivity field will inevitably trigger huge adjustments in social relations. Such adjustment may have positive externality but also negative externality as well.

From the supply side of the data factor, the expansion effect and interest incentive resulting from the combination of data and capital in the productivity field will attract increasing capital to the data field, promoting a geometric increase in data development and utilization and thereby accelerating the data processing on economic entities' behaviors, transparency of market operation, and intellectualization of government administration, further contributing to the development of new social relations, including labor relationship and government–market relationship. However, the profit-seeking nature of capital will not change in a socialist market economy. If no effective regulation applies, the negative externality of the combination of capital and data may not be underestimated. The most typical case is interest conflicts and safety hazards resulting from data monopoly. As major commercial data platforms gather knowledge and technology, talents and capital and are usually established by relatively large network capital or other industrial capital, their competitive advantages are increasingly prominent with an increase of data in their charge, and data monopoly is likely to occur over time.

In the field of economy, companies of data platforms may adopt monopoly agreements to acquire excess profits, i.e. based on algorithm or data feedback mechanism, jointly making business decisions beneficial to each other and eliminating other players deviating from such agreements. Also, they may abuse their dominant position in markets, such as price discrimination based on consumption data analysis, obtaining users' privacy information through unfair terms of agreements, "either-or" choice, blocking or shielding behavior and gathering players (that is, gathering data-driven players that have the goal of increasing the available data and strengthening the control of data) to amplify the "winner-takes-all" competition effect in the internet field (Hong, 2020). These monopolistic behaviors may cause negative externalities, including destroying market competition order, impairing legitimate rights and interests of consumers and users, suppressing the innovation ability of small- and medium-sized enterprises, weakening the driving forces of economic and social development and reducing social welfare.

When it comes to politics, data monopoly dominated by capital may influence public opinions and political choice and induce people to make value judgments consistent with the intention of capital. The documentary *The Great Hack* revealed that the Trump team used the artificial intelligence and big data technologies provided by Cambridge Analytica to smear Hillary Clinton and manipulate swing voters. In the presidential election, the Trump team bought approximately 50 million pieces of personal information through Cambridge Analytica from Facebook and carried out data analysis. After that, they selected a batch of swing voters and customized and advertised personalized and relatively biased content to them through numerous blog articles, videos and ads. Those people were "brainwashed" unwittingly and elected Donald Trump whom Cambridge Analytica wanted them to choose. In the 2020 US presidential campaign, the Democratic Party retaliated by leveraging companies leaning towards Democratic Party, such as Google, Twitter and Facebook, to push pro-Biden content to voters repeatedly by virtue of algorithms, which influenced the election result to some extent.

Regarding the demand side of the data factor, people are well aware of the multiplication effect of the bonding of capital and data, i.e. enhancing the effect of the data factor on increasing the allocation efficiency of other factors of production. Also, they have recognized the optimization effect on social relations brought about by the development of productivity.

However, certain conditions must be met to convert the potential demand for the data factor into real effective demand and produce the multiplication or optimization effect. Otherwise, effective demands may fail to be generated, and data poverty and the negative externalities of the bonding of data and capital on social relations may be caused.

For instance, at the level of individuals or specific groups, the data factor is profoundly changing people's way of employment, working, and living. However, individuals or groups vary in the capability of applying data. In the era of big data, those individuals or groups lacking knowledge or skills related to data utilization or the capability to apply data despite having knowledge or skills face the plight of data poverty, which will consequently affect their career choices and the sources and level of their income and cause them to suffer from severe discomfort in modern life and even both material and spiritual poverty, especially for those lacking basic material resources. If this effect is superimposed with the role of the market, it will further amplify the polarization effect on social classes with relatively low education and income and intensify the differentiation of occupations, incomes and lifestyles, forming new obstacles to the construction of a harmonious society.

At the regional level, the extensive utilization of data must be predicted on the availability of matching infrastructure and a certain talent pool, and the gap in conditions will amplify the development gap among regions. Among others, the disparity of conditions for data application between urban and rural regions, such as the lagging construction of new infrastructure, the polarization of central cities to rural talents and the relatively limited purchasing power and weak capability of utilizing the data factor of the rural population, will become a new trigger for the widening urban-rural development gap. If the data factor is monopolized by capital, the position of farmers in the market will be more vulnerable. For a long time, one of the main reasons why farmers make a good harvest but do not enjoy income growth is that farmers fail to share the profits from the distribution of agricultural products. While in the traditional model, most of the proceeds generated in this link are controlled by agricultural intermediaries, such profits are likely to be divided by platform enterprises in the era of big data.

Therefore, it is not enough to consider the data factor issue purely from the technical level. To maximize the positive externality of the data factor and effectively prevent and control the negative externality at the same time, it requires both scientific exploration at the technical level and institutional design from the perspective of social relation coordination.

On the one hand, to give play to the positive externality of the data factor, public data and commercial data should be governed separately. With regard to public data, it is necessary to persist in the concept of prioritizing the promotion of public data sharing, formulate and release a list of responsibilities for public data sharing (Dai, 2020), construct platforms for sharing public data and use 5G, IoT, artificial intelligence and other information technologies to innovate the development and application of data in various fields, including agriculture, industry, transportation, education, security, urban management and public resource transactions (Tang, 2021) and further realize the comprehensive collection, all-process and all-scenario coverage of public data (Wang, 2020a). Also, it is important to promote data sharing and exchange among regions and departments, facilitate the further integration of data on government affairs into the market of factors and take differentiated security control measures to ensure the monitorability and traceability of government data during the sharing and opening process. For commercial data, it is suggested to support the construction of multi-field data opening and application scenarios and construct the data factor market according to categories and levels for enhancing the development and utilization value of data; to support the establishment and improvement of the resources allocation mechanism and the governance rule system integrating data development, attribution of data ownership, data transaction, contribution assessment and pricing, distribution of remuneration and data

security protection; and to encourage enterprises to participate in formulating international rules and standards regarding the digital field (Liu, 2020).

On the other hand, to prevent and control the externality of the data factor, data categorization and classification standards (Li *et al.*, 2021) and security management rules should be formulated from the collection to utilization of data. The rule of law regarding data should be vigorously promoted, with focuses on defining the data development subject's ownership of data and the rights of independently using, sharing and opening data as well as transacting data business; specifying the qualifications of the data transaction subjects, the list of rights and responsibilities and the transaction rules and inhibiting data monopolization; ensuring that data and the transaction process are traceable and auditable and strengthening the supervision and security review system regarding data controllers' collecting, using, manipulating and transmitting data of citizens or enterprises to protect personal privacy, enterprises' commercial secrets and national security. With regard to the governance of data poverty, government investment and social support should be increased for backward regions and rural areas in new infrastructure construction, education and training, and so on, and scientific popularization, technology promotion and volunteer services for the application of data factor should be strengthened, to make the data factor not become a new cause of the return of poverty-eradicated regions and groups to poverty but become a new driver for common prosperity and modernization.

References

- Dai, S.X. (2020), "Shuju yaosu: zhuyao tezheng, tuidong xiaoying ji fazhan lujing [Data factor: main attributes, driving effect, and development path]", *Makesi zhuyi yu xianshi [Marxism and Reality]*, Vol. 31 No. 6, pp. 171-177.
- Hong, Y.X. (2020), "Hulianwang '+' shichang de jingjixue fenxi [Economic analysis of the "Internet + Market]", *Jiaoxue yu yanjiu [Teaching and Research]*, Vol. 55 No. 3, pp. 5-12.
- Li, G., Zhang, Q.K. and Zhu, K.X. (2021), "Shuju yaosu quequan jiaoyi de xiandai chanquan lilun silu [The determination of data elements and its transaction: a thought based on modern property rights theory]", *Shandong daxue xuebao (zhixue shehui kexueban) [Journal of Shandong University (Philosophy and Social Sciences)]*, Vol. 47 No. 1, pp. 87-97.
- Liu, D. (2020), "Jiakuai shuju yaosu shichang yunxing jizhi jianshe [accelerating the construction of data element market operation mechanism]", available at: <http://theory.people.com.cn/n1/2020/0904/c40531-31848888.html> (accessed 21 December 2021).
- Marx, K. (2009a), "Absolute and relative surplus-value", in Central Compilation and Translation Bureau (Ed.), *Makesi engesi wenji: diwujuan [Marx and Engels Collected Works: Volume V]*, Chinese translation edition translated by Central Compilation and Translation Bureau, People's Publishing House, Beijing, p. 589.
- Marx, K. (2009b), "Conversion of surplus-value into capital", in Central Compilation and Translation Bureau (Ed.), *Makesi engesi wenji: diwujuan [Marx and Engels Collected Works: Volume V]*, Chinese translation edition translated by Central Compilation and Translation Bureau, People's Publishing House, Beijing, pp. 699-701.
- McKinsey Global Institute (2011), "Big data: the next frontier for innovation, competition, and productivity", available at: https://www.mckinsey.com/~media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Big%20data%20The%20next%20frontier%20for%20innovation/MGI_big_data_full_report.pdf (accessed 21 December 2021).
- Shi, D. and Deng, Z. (2020), "Cujin shuju yaosu youxiao canyu jiazhi chuangzao he fenpei [Promoting effective participation of data factor in value creation and distribution]", available at: <http://opinion.people.com.cn/n1/2020/0122/c1003-31559156.html> (accessed 21 December 2021).
- Shidian Data (2020), "Jiyu dashuju pingtai de hulianwang shuju caiji pingtai jiben jiagou jieshao [Introduction to the basic architecture of internet data acquisition platform based on big data

- platform]", available at: <https://baijiahao.baidu.com/s?id=1663218855141262957&wfr=spider&for=pc> (accessed 21 December 2021).
- Tang, Y.J. (2021), "Shuju chanquan de jingji fenxi [Economic analysis on the property right on data]", *Shehui kexue jikan [Social Science Journal]*, Vol. 43 No. 1, pp. 98-106.
- Wang, L., Liu, Q.H. and Zeng, Z. (2020), "Jianquan jichuxing zhidu, peiyu shuju yaosu shichang [Improving the basic system and cultivating the data factor market]", available at: <http://theory.people.com.cn/n1/2020/1028/c40531-31908627.html> (accessed 21 December 2021).
- Wang, M. (2020a), "Shuzi jingji fazhan de falv guizhi—yantaohui zhuanjia guandian zongshu [Regulation of digital economy development—a summit review]", *Zhongguo liutong jingji [China Business and Market]*, Vol. 34 No. 12, pp. 114-124.
- Wang, Y.M. (2020b), "Shenhua yaosu peizhi gaige shifang zengzhang xinqianneng [Deepening the reform of factor Allocation to release the new potential for growth]", available at: http://dz.jjckb.cn/www/pages/webpage2009/html/2020-08/04/content_66380.htm (accessed 21 December 2021).
- Xie, K., Xia, Z.H. and Xiao, J.H. (2020), "Dashuju chengwei xianshi shengchan yaosu de qiye shixian jizhi: chanping chuangxin shijiao [The Enterprise Realization Mechanism of Big Data Becoming a Real Production Factor: from the Product Innovation Perspective]", *Zhongguo gongye jingji [China Industrial Economics]*, Vol. 37 No. 5, pp. 42-60.
- Zhai, S.Y. (2020), "Xiangxi jiedu dashuju, rengong zhineng, qukuailian, yuncunchu, qiye xinxihua deng jishu [Detailed explanation of big data, artificial intelligence, blockchain, cloud storage, enterprise informatization and other technologies]", available at: https://www.sohu.com/a/374109551_120521409 (accessed 21 December 2021).

About the authors

Xingmiao Guan is PhD student at the School of Marxism, Yangzhou University.

Xingfang Qin is Professor and Doctoral Supervisor at the Business School, Yangzhou University and Director of Jiangsu Urban-Rural Integration Development Research Center. Xingfang Qin is the corresponding author and can be contacted at: xfqin@yzu.edu.cn