The Rise of Data Politics: Digital China and Evolving International Relations

Lizhi Liu[1]

This version: May 20, 2020

Abstract

Data has become one of the most valuable assets for governments and firms. Yet we know surprisingly little about how data reshapes international economic relations. This paper explores various aspects of data politics through the lens of China's digital rise and the country's global engagement. I start with the theoretical premise that data differs from traditional strategic assets (e.g., land, oil, and labor) in that it is nonrival and partially excludable. These characteristics have generated externality, commitment, and valuation problems, triggering three fundamental changes in China's external economic relations. First, data's *externality* problem makes it necessary for states to regulate data, or even to pursue data sovereignty. But clashes over data sovereignty can ignite conflicts between China and other countries. Second, the *commitment* problem in data use raises global concerns about privacy and foreign government surveillance. As data is easier to transfer across borders than physical commodities, Chinese tech companies' investment abroad is vulnerable to national security investigations by foreign regulators. Chinese tech companies therefore confront a "deep versus broad" dilemma: deep ties with the Chinese government help promote their domestic business but jeopardize their international expansion. Lastly, data's *valuation* problem makes traditional measures (e.g., GDP) ill suited to measure the relative strengths of the world's economies, which may distort perceptions of China and other states.

Key words: China; Data; Digital Economy; Political Economy

[1] Assistant Professor, McDonough School of Business, Georgetown University, lizhi.liu@georgetown.edu.

Introduction

A 2017 *Economist* article asserted that "The world's most valuable resource is no longer oil, but data."[2] Since then, world leaders including Xi, Abe, Modi and Merkel have all expressed similar ideas, reflecting the rising consensus on the strategic importance of data.[3] Data is vital not only for its economic value, but also for its growing influence on politics. For instance, Cambridge Analytica's illicit harvesting of Facebook data reportedly influenced voting in 2016 US presidential election as well as the Brexit referendum in the UK earlier that year.[4] The Trump administration has also frequently cited data to defend the ban on Chinese tech giant Huawei, claiming that using Huawei's 5G network would expose sensitive US data to the Chinese government.[5]

Despite its looming prominence, the politics over data remains vastly understudied. Prior research sheds light on subsets of data politics, such as censorship (Roberts 2018; King et al. 2013, 2014), online institutional building (Liu 2018; Liu and Weingast 2018, 2019), privacy (Farrell and Newman, 2019), digital authoritarianism, and surveillance (Xu 2020; Greitens 2019). Yet the scope of data politics reaches far beyond these domains. Previous studies also tend to focus on digital technologies rather than the data itself, which has overlooked many of data's unique features.

---

[2] https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.
[3] See https://asia.nikkei.com/Politics/G-20-summit-Osaka/Trump-and-Xi-trade-barbs-over-5G-before-summit and https://www.belfercenter.org/publication/geopolitics-information.
[4] https://www.newyorker.com/news/news-desk/new-evidence-emerges-of-steve-bannon-and-cambridge-analyticas-role-in-brexit.
[5] https://www.cnbc.com/2020/02/25/trump-official-calls-huawei-mafia-as-white-house-works-on-5g-battle-plan.html.

This paper aims to fill these gaps in the literature by exploring the nature of cross-national data politics, exploiting the case study of China's digital transformation and global engagement. China serves as a good analytical case for this study for two main reasons. First, China's mass adoption of digital technologies has made data the lifeblood of every aspect of society. The government quickly adopted a national big data strategy and embraced the concept of "data sovereignty." Second, China has become a key investor in the data-related sector on a global scale, which enables connectivity but also provokes geopolitical concerns worldwide.

The paper explores how the proliferation of data as a new strategic asset affects international politics, particularly China's engagement with the world. I first theorize how data differs from traditional economic assets such as land and oil, which are typically rival and excludable, and why this distinction makes data politics worth researching. Drawing on the nascent literature on the economics of data, I show that data as an economic input is a (1) *nonrival* (i.e., data can be infinitely used) and (2) *partially excludable* (i.e., it is not always possible to exclude non-paying individuals from having access to data) good.

I then argue that these two features generate three problems – externality, commitment and evaluation – that lay the foundation of data politics. Using examples from China's engagement with the world, I demonstrate how each problem challenges the stylized facts of Chinese politics, and creates new tensions and uncertainties in international relations.

The first problem is *externality*. As data is nonrival and only partially excludable, it is a quasi-public good (or quasi-public bad). In other words, data from those who choose to share it can generate both benefits for and damage to non-sharers. To balance the positive and negative externalities of data, many states – including China – have begun to pursue "data sovereignty" as a way to regulate cross-border data flows. Clashes over data sovereignty – for example, China's data localization requirements and U.S. extraterritorial access to data – can generate conflicts between governments. These data regulations also constitute a new form of non-tariff trade barrier for multinationals, which I illustrate by studying two cases in details: (1) how China's Cybersecurity Law affected Apple Inc. (2) and how Europe's General Data Protection Regulation (GDPR) influenced Chinese bike-sharing startup Mobike.

The second challenge generated by the nonrival and partially excludable nature of data is the *commitment* problem in data use and sharing. As data is only partially excludable, and it never perishes, there is great uncertainty regarding its future use. In general, tech firms that collect personal data cannot guarantee that they will not disclose their user data to the government, which may abuse the data for surveillance or coercion. This commitment problem is more severe for firms based in authoritarian countries such as China. Chinese tech firms therefore suffer from such a commitment problem in overseas markets, where they are constantly suspected of sharing foreign citizens' data with the Chinese government. Being a private company does not help assuage such doubts. As a result, Chinese tech firms confront a "deep versus broad" dilemma: deep ties with the Chinese government reduce *domestic* political risks but raise *overseas* regulatory risks. I analyze the US investigation into TikTok below to illustrate the "deep versus broad" dilemma, and to show how firms can use costly signaling to attenuate it.

Lastly, data has a *valuation* problem: It is difficult to properly assess the price of data because it is nonrival, and can generate new value through combining different datasets. The most commonly used measure of a country's economy, gross domestic product (GDP), does not fully capture the value of data and many digital products that are free to use. This will generate measurement error in evaluations of a country's economic strength, leading to perception distortions.

The paper proceeds as follows. I first present background information on China's digital transformation, outward foreign direct investment (FDI) in data-related sectors, and the Communist Party's recent policy shift towards data regulations. I then analyze the economic nature of data, and how the nonrivalry and partial excludability features of data leads to externality, commitment and valuation problems. I elaborate these three problems, discuss them in the Chinese context, and examine how they affect politics. I conclude by highlighting the growing prominence of data politics over time.

Background: Digital China and The Proliferation of Data

Digitization has transformed all aspects of Chinese society over the last decade. In 2006, the country's online retail sales were merely 3% of the US sales. China now hosts the world's largest e-commerce retail market with a 40% share of global sales. Every day, 1 billion active users on WeChat exchange an average of 45 billion messages. Mobile pay has taken the country by storm: even beggars are accepting alms through QR codes. China has also established the world's first

official digital currency ("digital yuan"), as well as the first cyber court, where defendants, plaintiffs, and judges meet over video conference. This fast-paced digitization process has launched an explosion of tech firms and startups. By market capitalization, Alibaba and Tencent were ranked in the top 10 publicly traded corporations globally as of early 2020. China is also home to 9 of the top 20 global internet firms by public or private market value; the US hosts the rest.[6]

Its domestic success has driven the Chinese digital sector to go global. In 2014–2016, China-based venture capital funds accounted for 14% of worldwide venture capital investment outside China, of which 75% invested in digital-related sectors such as artificial intelligence (AI), cloud computing, and digital payments (Woetzel et al. 2017). Chinese tech titans have also expanded their domestic rivalry to the global level through mergers and acquisitions. For example, two competing Southeast Asian online trading platforms, Lazada and Shopee, are backed by Alibaba and Tencent, respectively. The Chinese government has also been keen to export digital capacity. The Digital Silk Road,[7] part of the Belt and Road Initiative, aims to facilitate global digital connectivity and digital trade, and further strengthen China's technology power and influence.

Data serves as the lifeblood of China's digital transformation at home and abroad. Historically, data has always been valuable, for example to reduce information asymmetry in economic transactions. But data did not play such a central role until the development of digital technologies that enable the mass collection, storage and analysis of data. China produced approximately 7.8 trillion gigabytes (GB) of data in 2018, and this number is expected to reach

[6] https://www.marketwatch.com/story/china-has-9-of-the-worlds-20-biggest-tech-companies-2018-05-31
[7] http://en.ndrc.gov.cn/newsrelease/201503/t20150330_669367.html.

48.6 trillion GB by 2025, surpassing the US.[8] China's large quantity of data may constitute a competitive advantage in the age of AI, as "Data is what makes AI go; a very good scientist with a ton of data will beat a super scientist with a modest amount of data" (Lee 2018).

The Chinese government has long realized the strategic value of data. In November 2013, the National Bureau of Statistics signed agreements of strategic collaboration with 11 internet firms, incorporating the use of big data into government statistics. In 2014, China elevated the status of big data to a national strategy; action plans were released in 2015 that included the construction of a massive national data center in Guizhou Province. In April 2020, an important policy document listed data as one of the five basic factors of production, alongside land, labor, capital, and technology.[9] This goes along with China's 2020 "New Infrastructure" campaign, which intensively focuses on enhancing digital infrastructures such as 5G networks, data centers, and AI.[10]

China is hardly the only country that has registered the proliferation of data or acts on the policy front; digitization is a global mega trend. Many other governments have launched their own digital initiatives and data regulations.[11] By 2014, trans-border data flows accounted for roughly 3.5% of the global GDP, almost as valuable as the global trade in physical goods and services (Bughin and Lund 2017). How does the proliferation of data affect global politics?

[8] https://www.cnbc.com/2019/02/14/china-will-create-more-data-than-the-us-by-2025-idc-report.html.
[9] http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm
[10] https://news.cgtn.com/news/2020-03-17/China-launches-New-Infrastructure-a-innovation-driven-program-OVo9mgDvGg/index.html
[11] For example, the European governments have launched Gaia-X, the European cloud computing initiative, to build the data infrastructure for Europe. https://www.ft.com/content/956ccaa6-0537-11ea-9afa-d9e2401fa7ca.

The Nature of Data and the Foundations of Data Politics

To examine data politics, we first need to discern the economic nature of data, and in particular, how data differs from other economic inputs.

Data is a strategic asset, but so are many other economic factors. Past studies have extensively explored the global politics involved in the ownership and trade of traditional goods such as oil, land, and capital. One may attempt to apply existing theories on traditional goods to data.

Yet data inherently differs from physical assets. Thus, data politics rests on a unique foundation. As noted above, a small but burgeoning literature in economics suggests that data has two characteristics that distinguish it from many other economic goods that are rival and excludable, including oil: nonrivalry (Jones and Tonetti 2019; Varian 2018) and partial excludability (Carrière-Swallow and Haksar 2019).

Nonrivalry is a central feature of data (Jones and Tonetti 2019);[12] it refers to the fact that data can be used an infinite number of times and by many parties simultaneously. For instance, governments and firms can collect the same set of personal data; citizens who give out their data still have access to the same amount of data. By contrast, rival inputs such as capital, oil, and

[12] It is important to distinguish data from information and ideas, which are also nonrival. I follow Jones and Tonetti (2019: 3), who define information as "all economic goods that are nonrival." My definition of idea comes from Romer (1990): a set of instructions of working with raw materials. Both ideas and data are subsets of information; they are mutually exclusive. Idea is the method of making an economic good, whereas data is a factor of production (or an input). For example, for a self-driving car, idea is the machine-learning algorithm that guides the car, and data is the input to produce the idea (Jones and Tonetti 2019).

labor do not replenish once used: when one consumes a barrel of oil, it reduces the availability of oil to others.

Importantly, nonrivalry means that data can exhibit *increasing returns to scale* (Jones and Tonetti 2019; Agrawal et al. 2018). Firms that use data can benefit from a "data feedback loop" (Farboodi et al. 2019) or direct network externalities (Goldfarb and Trefler 2018), in which a firm's success attracts more users and user data, which raises the quality of products through AI, leading to yet more users and data. Data can also yield increasing returns when combined with other factors of production. For instance, to double the production of Tesla automobiles, one needs to double the rival inputs, such as iron, used to manufacture them. But the same data that empowers the self-driving algorithms in one Tesla can be used in two, three or millions of cars.

Data's nonrivalry feature, coupled with increasing returns to scale, has two economic implications. First, there may be substantial social gains if data is widely shared across firms and countries. Second, if data is not broadly shared, the quantity held by a firm or country can generate a competitive advantage. Bigger firms and countries that are associated with larger quantities of data can be more productive in the digital era, leading to greater market concentration.

Data is also *partially excludable*: it is possible to prevent non-paying individuals from consuming some data products. Traditional economic inputs are often excludable. For instance, a state can prevent individuals from occupying territory or consuming oil if they have not paid for it. However, it is not always possible to block non-paying individuals' access to data.

Data is only partially excludable because it is easy and inexpensive to duplicate it, and to transfer it over a long distance, even cross a country's border. This generates a high risk of unauthorized sales, usage or theft of data. Of course, data can still be excludable if it is fully encrypted, anonymized, or stored offline (Jones and Tonetti 2019). But encryption, anonymization, and offline isolation normally make the data less valuable. Many digital products, such as cloud computing and GPS navigation, can no longer function without real-time data sharing.

Table: Types of Economic Goods based on Rivalry and Excludability

	Excludable	**Nonexcludable**
Rival	*Private Goods*: Food; oil; common manufactured products; fish in a private pond, etc.	*Common Goods*: Forests; mines; fish in the ocean, etc.
Nonrival	*Club Goods*: Satellite TV; private parks, etc. **Data**	*Public Goods*: National defense; air; sunshine; news, etc.

As previewed above, the economic characteristics of data – nonrival and partial excludability – lead to three types of problems that lay the foundation for cross-national data politics: externality, commitment and valuation problems.

Data's Externality Problems and China's Move Towards Data Sovereignty

Externality Problems of Data

Data's partial excludability and nonrivalry features make it resemble a public good. It has both positive and negative externalities (Acquisti et al. 2016; Choi et al. 2019).

The sharing and creation of data can have positive externalities such as beneficial spillovers on a third party. For instance, the collection of personal data from data sharers can train a machine-learning algorithm that improves the quality of service provided to others, including non-sharers. But data also has negative externalities, especially on privacy. In some circumstances, data about the sharers can reveal some characteristics about non-sharers (MacCarthy 2011; Acemoglu et al, 2017). Erlich et al. (2018) found that their model only needed 2% of the target population's genetic data in order to identify almost everyone, including those who had never undergone genetic testing. Likewise, a famous project conducted by MIT students found it was easy to predict whether someone is gay based solely on the reported sexuality of their Facebook friends, leading to unintentionally "outing" (Jernigan and Mistree 2009).

For governments, it is important to balance the positive and negative externalities of data, especially when the externalities come from *international data flows*. The cross-national exchange of data can lead to huge trade benefits, but also expose citizen privacy to foreign surveillance, hacking and data breaches. To regulate data's externalities, market solutions are often not sufficient. For example, fearing the creative destruction caused by data sharing, firms

can hoard data for themselves (Jones and Tonetti 2019). This limits the positive externalities and welfare gains data can generate. In addition, firms have an incentive to over-collect data but under-invest in data security and privacy protection (Carrière-Swallow and Haksar 2019; Jin 2018), which amplifies the negative externalities of data.

Therefore, many governments have begun to claim "data sovereignty" as a way to regulate trans-border data flows. China is one of the early movers. The rest of the section discusses the Chinese approach to data sovereignty, and how the emergence of national digital borders affects tech firms within China and Chinese firms around the world.

Data Sovereignty, Chinese Style

The concept of data sovereignty means that data should be subject to the laws and regulations of the nation-state in which it is generated and processed. It is a political effort to restrict data services along national borders. China's data sovereignty is intricately linked to "cyber sovereignty," a broader concept that states should have control over the digital technologies, content and infrastructures within their jurisdictions, which President Xi called for in 2015.[13]

To establish "digital sovereignty," China has issued four key documents to regulate trans-border data flows. The foundation was the Cybersecurity Law (网络安全法), which was enacted in 2016 and came into effect in 2017. The central government then issued three additional documents to implement the law: the Draft Measures on Security Assessment of the Cross-

[13] https://www.bbc.com/news/world-asia-china-35109453

border Transfer of Personal Information of 2019 (个人信息出境安全评估办法), the Draft Measures for Data Security Management of 2019 (数据安全管理办法), and the Draft Information Security Technology-Guidelines for Data Cross-Border Transfer Security Assessment of 2017 (信息安全技术数据出境安全评估指南).

These documents comprise the core of China's approach to data sovereignty: the requirement of data localization. The key is "local storage, outbound assessment" (Liu 2020). It requires all businesses operating in China to store select data (e.g., Chinese citizens' personal data) on servers within China. If it is necessary to transfer locally stored data to another country, the Chinese government will conduct a security assessment.[14] The law also allows the Chinese government to conduct spot checks on foreign businesses.[15]

While many other countries have required data localization (Chander and Lê 2015) or struggled to maintain "data sovereignty,"[16] China's path is unique. So far, its data localization requirement is comprehensive but vague.[17] Initially in 2017, this requirement covered not only personal data but also important data regarding critical information infrastructure; "important data" is vaguely defined and can include any data, subject to the whims of the authorities.[18]

[14] https://medium.com/center-for-media-data-and-society/how-china-governs-data-ff71139b68d2
[15] https://thediplomat.com/2017/06/chinas-cybersecurity-law-what-you-need-to-know/
[16] Over the past few years, German Chancellor Angela Merkel has expressed deep concerns about Europe's "digital dependences" on US and China. She emphasized the importance of claiming digital/data sovereignty by developing Europe's own platforms to manage data. In a similar vein, French President Emmanuel Macron said in an interview, "the battle we're fighting is one of sovereignty…If we don't build our own champions in all areas—digital, artificial intelligence—our choices will be dictated by others" (see http://www.rfi.fr/en/france/20190918-macron-throws-5-billion-digitl-startups).
[17] https://jsis.washington.edu/news/chinese-data-localization-law-comprehensive-ambiguous/
[18] Ibid.

Additionally, the Chinese approach to data regulation is state centered, with "a unique combination of data protection and the government's control over data flows."[19] The state has begun to regulate businesses in order to protect citizens from the unnecessary collection of personal data. For instance, in May 2020 the government fined a gym that illegally collected facial images of users.[20] However, the state itself has almost unrestrained access to citizen data.

This state centered model makes China different from the EU and the US, where the data regulations are respectively citizen based and market based. As most of the digital platforms that collect EU citizen data are non-EU firms (e.g., Facebook, Google, Apple), Europe has offered regulatory options to individual users, empowering them through the robust GDPR. The US, however, has the world's most competitive digital companies and has largely relied on corporate self-regulation to govern data. While California recently adopted the strictest privacy law in the US – the California Consumer Privacy Act, which resembles GDPR, there is no federal-level privacy law. Compared to Europe, China's data regulations focus more on protecting national security rather than citizen privacy. Also, unlike the US, the Chinese state is very active in setting the agenda to regulate data.

The Inconvenient Consequences of Data Sovereignty

As China quickly erects borders for data, many countries will follow suit. The emergence of national digital borders will have at least four far-reaching political effects on China.

[19] https://medium.com/center-for-media-data-and-society/how-china-governs-data-ff71139b68d2
[20] https://posts.careerengine.us/p/5ec010b679e983564e43ee4a

First, China's move to enhance data sovereignty will further strengthen digital authoritarianism on top of its already-sophisticated system of censorship and surveillance. While censorship aims to keep information *outside* a country, data localization keeps all the citizen data *within* a country; the combination of censorship and data localization can lead to total control, further increasing the government's ability to coerce and surveil (Chander and Lê 2015).

Second, clashes over data sovereignty could become a new conflict hotspot for China and other countries. Many governments have erected data regulations with uncommon swiftness and little coordination, leading to incompatible rules. For instance, China's data localization requirement could clash with the US government's extraterritorial access to data. In 2018, the US Congress enacted the Clarifying Lawful Overseas Use of Data (CLOUD) Act, which authorizes law enforcement agencies to compel US-based tech firms (e.g., email and cloud service providers such as Amazon and Microsoft) to hand over data – whether it is stored on US or foreign servers.[21] The act generates tricky questions about enforcement. An op-ed by a former federal prosecutor discusses an interesting hypothetical scenario: "What if a 'wannabe 'America's Toughest Sheriff' serves a warrant on the Silicon Valley office of one of China's major service providers to demand a Chinese official's correspondence with party leaders stored in a Beijing server?"[22] In this way, data rules can generate conflicts that require diplomatic resolution.

[21] https://www.justice.gov/opa/press-release/file/1153446/download
[22] https://www.forbes.com/sites/roncheng/2018/05/29/seizing-data-overseas-from-foreign-internet-companies-under-the-cloud-act/#17a4278e16c0

Third, China's domestic legislation can generate problems when Chinese firms invest in countries with strict data regulations. While China has no equivalent to the US CLOUD Act,[23] its National Intelligence Law of 2017 sets broad requirements on organizations to "support, assist and cooperate with the state intelligence work."[24] This law is often cited in the security debate regarding whether Huawei presents an espionage threat to the West. Huawei critics interpret the law as a legal foundation to grant the Chinese government access to the company's data, even in the overseas market where data should be stored locally.[25]

Lastly, the emergence of national digital boarders will constitute a new trade barrier. Such borders will incur huge compliance costs for multinational tech firms, disrupting the global value chain and compromising economies of scale. Below I examine two cases to demonstrate this point: Apple in China and Mobike in Europe. They respectively demonstrate how Chinese Cybersecurity Law affects multinational operations within China, and how data regulations elsewhere penalize Chinese businesses in overseas markets.

Case I: Apple in China

China has been a crucial market for Apple Inc. – the second-largest market for iPhone sales and the key manufacturing base – but Apple's relationship with China has never been easy.

[23] Some legal experts argue that China may have quietly established extraterritorial access to data. https://www.lawfareblog.com/did-china-quietly-authorize-law-enforcement-access-data-anywhere-world
[24] https://cs.brown.edu/courses/csci1800/sources/2017_PRC_NationalIntelligenceLaw.pdf
[25] For example, https://www.reuters.com/article/us-eu-huawei-tech/chinas-intelligence-law-looms-over-eu-5g-safeguards-official-idUSKCN1UE18I

Many of Apple's skirmishes with Beijing have centered on user data. The company has frequently come under fire from China's state media, which suspected that it had provided user information to US intelligence agencies.[26] In 2014, China's state broadcaster accused Apple of posing a national security threat. It expressed concerns over iPhone's "Frequent Locations" feature, which tracks real-time user locations for Apple apps such as Weather and Maps. The state media claimed that the location data was extremely sensitive, as it could reveal China's economic conditions and "even state secrets."[27]

In response, Apple explained that location data was all encrypted and stored on individual phones, rather than on company servers. It also denied helping the US government spy on iPhone users.[28] To further assure the Chinese government, Apple began to store some Chinese user data on servers in mainland China. The data was moved from overseas to a domestic plant managed by the state-owned enterprise (SOE) China Telecom.

In 2016, China issued the Cybersecurity Law, which requires all online data of Chinese citizens to be stored on domestic servers. The official purpose of the law is to help combat rising terrorism and hacking. But critics worry it could strengthen the censorship regime and raise the barriers to entry for foreign firms, which face higher adjustment costs under the law.[29]

Apple chose to comply. In 2017, it started to build a new data center operated by a company co-founded by the Guizhou provincial government. In 2018, Apple moved to China the keys needed

[26] https://www.bbc.com/zhongwen/simp/china/2014/07/140711_china_cctv_apple_iphone
[27] https://www.voanews.com/east-asia/cctv-iphone-threat-national-security
[28] https://www.bbc.com/news/technology-28292378
[29] https://www.reuters.com/article/us-china-parliament-cyber-idUSKBN132049

to unlock Chinese users' iCloud accounts.[30] This action provoked considerable criticism from human rights activists in the West.[31]

Apple's situation reflects a dilemma common to multinationals: not complying with Chinese data regulations risks a crackdown from the Chinese government, whereas compliance creates a liability in Western markets.

Case II: Mobike in Europe

Mobike is one of China's most renowned startups. It is the world's largest bike-sharing operator, which allows users to find a bike via an app, unlock it with a click, and park it anywhere in the city when they are done. The startup was founded in 2015 and quickly gained popularity in China. In 2017, Mobike launched its European operations. It had deployed 200,000 bicycles in 23 cities across Europe by 2018, accumulating 200 million registered users worldwide.

Mobike's real value is not in the bicycles it owns, but in the data it collects via the app. The data makes it possible to connect a user with a bicycle. It also reveals rich information about the user. Such information, if resold, can generate substantial revenue for the company. As an investor Mark Wiseman commented, "The owners of those bicycles (i.e., Mobike) know where they are being ridden, when and by whom…I'd be willing to pay a lot of money for that data. "[32]

[30] https://www.reuters.com/article/us-china-apple-icloud-insight/apple-moves-to-store-icloud-keys-in-china-raising-human-rights-fears-idUSKCN1G8060
[31] https://money.cnn.com/2018/02/28/technology/apple-icloud-data-china/index.html
[32] https://www.bloombergquint.com/markets/blackrock-s-wiseman-says-beijing-bike-rentals-show-data-is-king

The data-driven model – which is what makes Mobike successful – also makes it vulnerable to data regulations. In May 2018, shortly after Mobike entered the European market, the EU's GDPR came into force. GDPR is the world's toughest data regulation regime, which aims to protect EU citizens' personal data and privacy. Under GDPR, regulators can fine non-compliant companies up to 4% of their annual global turnover or €20 million, whichever is greater. In a prominent case, Google was fined $57 million in early 2019 for "lack of transparency, inadequate information and lack of valid consent regarding the ads personalization".[33]

In late 2018, after it had established a presence in multiple cities, Mobike was reported to be under investigation by Germany's privacy watchdog for a potential breach of GDPR. The details of the probe were not made public, but according to the Swedish privacy expert Alexander Hanff, the regulator's concerns were likely related to three aspects of its operations in Europe.[34]

The first is the excessive collection of user data. The app collects a significant amount of data, including precise location data, even when the user is not using the app. Second, Mobike's privacy policy allows it to share this data with third-party companies, and users have no control over the data handed over to these third parties. The third and biggest concern is that Mobike sends user data to China and Singapore, where it has its headquarters. This directly challenges GDPR, which imposes stringent restrictions on transfers of personal data outside the EU.

[33] https://www.cnbc.com/2019/01/21/french-watchdog-slaps-google-with-57-million-fine-under-new-eu-law.html
[34] https://medium.com/@a.hanff/chinas-surveillance-social-credit-system-alive-kicking-in-berlin-6c2b3b10b197

Three months after the probe was disclosed, Mobike started to seek to spin off its European arm. It eventually withdrew from the European market after a management buyout in 2019,[35] likely due at least in part to Europe's strict data regulation.

Commitment Problems and the "Deep versus Broad" Dilemma for Chinese Firms

Commitment Problems of Data

Data's nonrivalry feature means that it can be infinitely reused. The reuse of data is difficult to detect because it is only partially excludable. Therefore, once citizens exchange their personal data for free services from digital firms, they usually lose control over future use of the data (Jin 2018).

This uncertainty regarding future data use is a critical political matter, because it can generate concerns about foreign government surveillance. Such concerns arise from two commitment problems: (1) multinational firms cannot credibly commit to not share personal data with their home government, and (2) the home government cannot commit to not abuse personal data for surveillance or for other political purposes that encroach on individual liberty.

Both commitment problems deteriorate in an authoritarian country such as China. Regarding problem (1), firms are less likely to defy data requests from an authoritarian government. In

[35] https://www.chinaknowledge.com/News/DetailNews/86227/mobike-europe

2016, Apple openly resisted the FBI's request to unlock the iPhone of the shooter in a terrorist incident in San Bernardino, California. Making such an open refusal is not impossible for firms in authoritarian states, but far riskier. For example, Russia banned the messenger app Telegram for two years after it refused to share encryption keys with Roskomnadzor.[36] As for commitment problem (2), autocrats may have a greater incentive to engage in digital surveillance, as they have an information issue of not knowing their citizens' true preferences due to preference falsification (Xu 2020).

There are two important caveats regarding the commitment problems. First, the fact that a firm cannot credibly commit *ex ante* does not mean it will surely hand user data to its home government *ex post*, even for firms originated from an authoritarian state. Firms bear reputational costs if abusing their user data, and their interests are usually not strongly aligned with the interests of political authorities. Second, firms and governments in democracies are not immune from these commitment problems. Government mass surveillance projects are on the rise globally, as evident by the U.S. NSA's PRISM program, the British Tempora project, and many others. According to the annual transparency reports published by Facebook, Apple, and Google, a significant portion of the government data requests come from democracies, and these tech giants have fulfilled the vast majority of these requests in recent years.[37] Therefore, commitment problem in data use is by no means a unique issue for autocracies. One advantage for firms based in democracies, however, is that the institutional constrains – e.g., checks on government's

[36] https://www.theverge.com/2020/6/18/21295535/russia-telegram-ban-lifted-security.
[37] For example, Apple's 2019 transparency report can be found here https://www.apple.com/legal/transparency/pdf/requests-2019-H1-en.pdf. Facebook reports statistics of government data requests here https://transparency.facebook.com/government-data-requests, and Google's report is here https://transparencyreport.google.com/user-data/overview?hl=en.

arbitrary act (including accessing data), free press, and the rule of law – are stronger in democracies. These constraints on both firms and governments can serve as commitment devices, albeit imperfect ones, to bind future actions of data abuse.

When Chinese tech companies expand globally, they suffer from these commitment problems, which make them vulnerable to national security investigations worldwide. As data is core to tech firms' business operations, they have to collect personal data of foreign users. A growing number of foreign governments, however, maintain that China-based tech firms may expose foreign citizen data to the Chinese government. Such distrust in Chinese tech firms is certainly related to – but not solely driven by – the rising skepticism towards China itself. The underlying issue is a general problem: the commitment problems in data use facing every multinational tech firm and every government.

The rest of this section specifies how the commitment problem affects Chinese private tech firms through a "deep versus broad" dilemma by examining the case study of the US investigation of TikTok.

Ownership Bias and Blurred State–Business Boundaries

All governments seek to maximize the economic benefits and, meanwhile, minimize national security risks associated with inward FDI. When dealing with Chinese investors, foreign governments in the past had "ownership bias": they used to be more cautious with investments from Chinse SOEs than those from private firms. This is consistent with existing research on the

Chinese corporate sector, which has overwhelmingly relied on firm ownership to demarcate the boundary of the state. SOEs are portrayed as vehicles for the state to fulfill domestic political goals and to exert economic statecraft globally. By contrast, private enterprises are excluded from, or only peripherally connected to, the political process.[38] While private firms are not entirely autonomous from the state, they are not as connected to the state as SOEs.

However, ownership has long been a fuzzy concept in China.[39] More importantly, the rise of tech companies has blurred the state–business boundary, making it porous, elusive, and fluid. Unlike traditional private companies that are marginalized in state initiatives, privately held Big Techs have taken many strategic roles that only SOEs could assume in the past.

Within China, there are extensive collaborations between the state and digital companies over a wide range of issues. Underlying this is a phenomenon of what I call *institutional outsourcing*[40]: as the authoritarian state is unable or unwilling to reform formal institutions, it has implicitly or explicitly outsourced some institutional functions to key private actors, in particular digital platforms (Liu 2018). To clarify, institutional outsourcing does not mean that these private Big Techs have become part of the state, or mainly act upon the state's political directives. Big Techs

[38] Pearson (1997) and Dickson (2003) find that private entrepreneurs are not regime challengers despite having little political representation. Tsai (2006) argues that when formal institutions are too rigid, private businesses use creative informal ways to bypass the restrictions of formal institutions.

[39] At the dawn of the reform era, when there was no legal framework for private ownership, many "collectively owned" rural enterprises were private in disguise (Huang 2008). In later periods, SOE reforms created many mixed-ownership firms, in which state and private shareholders share ownership and management responsibilities, making it difficult to classify whether a firm is public or private. For instance, ZTE, the telecommunications giant that paid a $1 billion fine for violating US sanctions, is state owned in terms of ownership structure but private from the perspective of management control. While the state holds 51% of ZTE shares, private shareholders take sole responsibility for managing the firm (Milhaupt and Zheng 2015).

[40] In the late stage of this project, I discovered that the term had been previously used by other scholars in different contexts, including China's integration into the global economic system (Steinfeld 2010), the study of comparative capitalism (Allen 2013), and elite-mass relations in Russia (Polishchuk 2013). I use the term in a distinctly different way from the prior literature. See Liu (2018) for more details.

are still primarily profits-driven, pursuing commercial interests that are not fully aligned with the state's goal. What features institutional outsourcing is the extensive collaboration between the state and Big Techs to solve various governance issues that the state falls short of.

In the economic realm, online trading platforms assist the state to create market institutions (e.g., contract enforcement, fraud prevention), enforce laws (e.g., Alibaba and JD help the Supreme Court enforce debt payment), conduct policy experiments (e.g., Alibaba helps local governments build "smart cities"), and facilitate rural development (Liu and Weingast 2018, 2019; Couture et al. 2020).

In the political realm, private tech companies such as social media platforms help the state perform censorship (Gallagher and Miller 2019) and conduct surveillance.[41] In the social area, during the Covid-19 outbreak, a high-level official at the Ministry of Civil Affairs "earnestly requested" Big Techs to develop community apps to contain the spread of the virus.[42] Alibaba helped the Hangzhou government build the "health code," using big data to assign a color code to indicate individuals' health status, which was then quickly rolled out nationwide.[43]

The intimate domestic collaboration between the Chinese state and private tech companies, especially on the data front, can become a liability of these firms in the overseas market. While many of these collaborations are nonpolitical and Big Techs reportedly have turned down

[41] https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html
[42] https://www.thepaper.cn/newsDetail_forward_5896831
[43] https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html

government data requests occasionally, [44] the lack of institutionalized checks on Chinese state's arbitrary power makes it difficult for private techs to prove their independence. Overseas regulators no longer view private ownership as a credible signal of a firm's relative autonomy from the Chinese state, which, for example, constitutes a major hurdle facing the Chinese tech giant Huawei when rolling out its 5G network worldwide.[45]

Private Techs' "Deep versus Broad" Dilemma in Global Expansion

As foreign governments lack complete information about the real type of a Chinese firm – whether it will pose a national security risk – and firm ownership no longer suffices to predict the firm type,[46] foreign regulators tend to infer the Chinese firm's type from its business activities within China.

Therefore, China-originated private firms increasingly confront a "deep versus broad" dilemma: deep embeddedness in the Chinese market – where building strong political connections is necessary – takes a toll on their global expansion. The rub is, a deep tie with the Chinese government reduces domestic political risks but increases overseas regulatory risks.

[44] Alibaba and Tencent once refused to submit their customer loan data to Baihang, a credit scoring business launched by the Chinese central bank. See https://www.ft.com/content/93451b98-da12-11e9-8f9b-77216ebe1f17.
[45] https://www.wsj.com/articles/state-support-helped-fuel-huaweis-global-rise-11577280736
[46] However, foregoing ownership as an analytical tool runs another risk: the risk of assuming the state's ubiquitous presence in all private firms. Essentially, there is a trade-off between Type I and Type II errors. While the use of ownership may underestimate the state's reach within the private sector (i.e., the error of having "false negatives"), ignoring ownership can generate many "false positives" by assuming all private firms are agents of the state. This trade-off makes it difficult to analyze the Chinese state's reach within a Chinese company.

This dilemma is particularly salient for Chinese firms operating in countries with strong anti-China sentiment. It is also exceptionally tricky for digital companies, which collect personal data for business operation. As discussed previously, data is nonrival and only partially excludable. These features make the firm difficult to commit *ex ante* to not share data with the Chinese government *ex post*. As data is easier to transfer across borders than physical commodities, Chinese tech firms are particularly vulnerable to national security investigations by foreign governments.

Is there a way to break, or at least attenuate, the dilemma? Some tech firms have engaged in costly signaling by creating "walls" to separate their domestic markets from overseas businesses. In the following, I examine the case study of the US national security investigation into TikTok to explicate the logic of data politics: how the collection of personal data becomes a sensitive issue, how the firm's inability to commit generates suspicions, and how TikTok seeks to address the commitment problem by costly signaling.

Case: TikTok's Commitment Problems and Costly Signaling

TikTok, owned by the Beijing-based tech giant ByteDance, is the first Chinese social app to take the world by storm.[47] This AI-empowered platform allows users to create and share short videos, often of themselves dancing, lip-syncing, or cooking. By April 2020, merely three years after its

[47] Other Chinese social media platforms such as WeChat and Weibo have struggled to win overseas users.

launch, the app hit 2 billion downloads worldwide, amassing 800 million monthly active users,[48] 26.5 million of whom reside in the US.[49]

TikTok's meteoric rise has prompted U.S. lawmakers to panic. Although it is privately owned and operates exclusively outside of China, it has been suspected of being under the control of the Chinese government. In an open letter released in October 2019, US Senators Tom Cotton and Chuck Schumer contended that "TikTok is a potential counterintelligence threat we cannot ignore," requesting intelligence officials to assess the national security risk posed by it and other China-owned platforms.[50] In November 2019, the US government launched national security investigation into TikTok.[51] As of March 2020, several US government agencies, including the Transportation Security Administration and the US Army, had banned the app from employee devices.[52]

How could an app famous for lip-syncing and dance challenges raise national security concerns? The investigation is not simply driven by the growing US–China rift that has politicized business matters. Data politics is also at play, which can take place between any country dyad, and regarding any multinational firm that collects personal data.

[48] https://www.scmp.com/lifestyle/entertainment/article/3079011/why-tiktok-such-online-success-ask-drake-jane-fonda-and-its
[49] https://www.reuters.com/article/us-tiktok-cfius-exclusive/exclusive-u-s-opens-national-security-investigation-into-tiktok-sources-idUSKBN1XB4IL
[50] https://www.cotton.senate.gov/?p=press_release&id=1239
[51] https://www.cnbc.com/2019/11/01/us-to-investigate-tiktok-over-national-security-concerns-sources-say.html
[52] https://www.businessinsider.com/us-government-agencies-have-banned-tiktok-app-2020-2

The senators' open letter specified two data-related concerns over TikTok. The first was how the company stores and handles US user data, and whether it will be exposed to the Chinese government. The second is related to the potential for disinformation and content manipulation. They argue that TikTok has reportedly censored political content disliked by the Chinese Communist Party, including material linked to the Hong Kong protests, Tibet, and Taiwan. It also mentioned that TikTok may serve as a potential forum for foreign meddling in US elections, as Facebook did in 2016.[53]

TikTok denied all of these charges. In a statement, it emphasized that all US consumer data is stored in the US and backed up in Singapore, arguing that "our data centers are located entirely outside of China, and none of our data is subject to Chinese law." It also claimed that it would not censor content based on political sensitivities and was not influenced by any government.[54]

Some lawmakers disparaged TikTok's denials as cheap talk. Senator Josh Hawley commented in a hearing that "TikTok claims they don't store American user data in China. That's nice. But all it takes is one knock on the door of their parent company, based in China, from a Communist Party official for that data to be transferred to the Chinese government's hands, whenever they need it."[55]

This reflects a commitment problem facing TikTok: the firm cannot credibly commit *ex ante* that it will not share user data with the Chinese government *ex post*. While every firm has a similar

[53] https://www.cotton.senate.gov/?p=press_release&id=1239
[54] https://newsroom.tiktok.com/en-us/statement-on-tiktoks-content-moderation-and-data-security-practices
[55] https://www.nbcnews.com/politics/congress/hawley-takes-aim-tiktok-china-congressional-hearing-n1076586

commitment problem, firms from an authoritarian country suffer more from it, because authoritarian countries do not have strong institutional constraints on the ruler's arbitrary power. As the senators' open letter claims, "without an independent judiciary to review requests made by the Chinese government for data or other actions, there is no legal mechanism for Chinese companies to appeal if they disagree with a request."[56]

This commitment problem leads to the "deep versus broad" dilemma. TikTok's parent company, ByteDance, is deeply embedded in the Chinese market. To some extent, ByteDance's deep success within China becomes a toll on TikTok's international expansion.

To expand globally, TikTok has to signal its independence from ByteDance's China business. So far it has made three costly moves to send this signal. First, TikTok limited its own revenue sources by not allowing any paid political advertising on its platform.[57] Second, it started to look for global headquarters outside China, in either Singapore, London, or Dublin.[58] Third, TikTok announced it would no longer use China-based moderators to monitor overseas content.[59] It is too soon to tell whether this costly signaling strategy will succeed, and whether these signals will be regarded as credible by the U.S. law enforcement.

[56] https://www.cotton.senate.gov/?p=press_release&id=1239
[57] https://www.engadget.com/2019-10-03-tiktok-political-ads.html
[58] https://www.wsj.com/articles/tiktok-searches-for-global-headquarters-outside-of-china-11577097150
[59] https://www.wsj.com/articles/tiktok-to-stop-using-china-based-moderators-to-monitor-overseas-content-11584300597

Data's Valuation Problems and the Measurement Bias on States' Power

Valuation Problems of Data

Data proliferates as an economic input in the modern economy, but its economic value is difficult to measure for two reasons. First, data is nonrival and intangible. It therefore cannot be depleted or worn in the way that many tangible goods are. Second, its value does not necessarily depreciate over time. Quite often, "data fusion" – the recombination of different independent datasets – generates new value (Li et al. 2019).

This valuation problem makes it challenging to evaluate a country's real economic strength. With the rise of the digital economy, a country's operation increasingly relies on the collection, exchange, and sale of data, for which we lack proper valuation. What particularly complicates the measurement problem is the prevalence of free digital products such as search engines, social media, and email. As the use of digital products involve no or few monetary payments, their value is substantially underestimated by current measures of economic power, such as GDP.

The valuation of data is not simply an economic problem. It also matters substantially in politics. Politicians often develop foreign policies based on the economic conditions of their own and foreign states. For a long time, states have used GDP and GDP growth to measure countries' relative economic power. As data plays an increasingly important role in the world economy, incomplete measurements of countries' economies, such as GDP, will produce inaccurate assessments of world powers, and thus distorted foreign policies.

The rest of this section discusses the problems of the GDP measure in the digital era, and the difficulty of measuring the welfare gain generated by digital services, using China's rural e-commerce project as an example. I also discuss how this measurement problem can distort our view of the relative strengths of states.

GDP's Measurement Error in the Digital Era

Most political analyses use the GDP statistic to measure the size of an economy and states' relative economic power. China's miraculous decades-long GDP growth has motivated a rich literature that examines the political roots of growth (e.g., Shirk 1993; Oi 1999; Landry 2008) and the performance/loyalty-based nomenklatura system (e.g., Shi et al. 2012; Landry et al. 2018).

The rise of the digital economy, however, raises a cautionary note regarding the use of GDP to measure power. Prior research has questioned the reliability of China's GDP figures, citing evidence of data manipulation in official statistics (e.g., Wallace 2016). But there is a deeper concern: by design, the GDP statistic cannot capture economic welfare, particularly the welfare gains from free or nearly-free digital products. This problem is not unique to China; it is universal.

GDP measures the market value of all final products and services produced in an economy. It has two inherent limitations. First, it does not include zero-priced products in the market (Aitken

2019). Second, GDP primarily covers market production, which provides a clear set of quantities and prices. It does not include nonmarket production – e.g., household production for self-use such as cooking and maintenance, or unpaid volunteer work (IMF 2018).

These measurement problems become increasingly thorny in the digital era. Internet services such as Facebook and Google are often *free to use* or strikingly underpriced. For example, the Chinese messaging app WeChat allows migrant workers to have free video calls with their left-behind kids. This welfare gain is not fully captured by the GDP measure, as WeChat users *do not pay* for the service. Instead the users barter their personal data for such digital products, enabling the digital companies to monetize by selling targeted ads or reselling data. While these advertising profits can be counted as part of the GDP, they represent only a fraction of the value of data. These advertising profits constitute a fraction of the value of these services (Brynjolfsson and Collis 2019). Therefore, many digital services have reduced the cost of living and improved the quality of citizens' lives, yet the GDP measure underestimates their effects.

Another feature of the digital economy is that voluntary contributions – e.g., Wikipedia, its Chinese counterpart Baidu Baike, and other open-source software – generate substantial productivity. Again, GDP does not capture this part of the welfare gain. A good example shown by Brynjolfsson and Collis (2019) illustrates how the GDP measure can be *negatively* correlated with actual economic well-being: "Consider *Encyclopedia Britannica* and Wikipedia. Britannica used to cost several thousand dollars, meaning its customers considered it to be worth at least that amount. Wikipedia, a free service, has far more articles, at comparable quality, than

Britannica ever did. Measured by consumer spending, the industry is shrinking...But measured by benefits, consumers have never been better off" (Brynjolfsson and Collis 2019: 145).

The following case of China's rural e-commerce initiative indicates the difficulty of fully capturing the welfare gains from digital services, and how consumption gains can take place without changes in output or nominal income.

Case: Quantifying Household Welfare Gain from Rural E-Commerce

Around 2014, most e-commerce activities in China still took place in urban areas, leading to a national policy initiative to expand e-commerce to rural areas. To implement this policy, the Chinese government partnered with a large online trading platform. The program aimed to introduce e-commerce into 100,000 villages by installing an e-commerce service point in each participating village. Each service point hired a terminal manager to help villagers buy and sell products on the platform. The firm and the government also connected the participating village with the urban center, and fully subsidized the shipping costs of packaging going into and out of the village.

To study how e-commerce connectivity causally affects rural household welfare, Couture et al. (2020) conducted a field experiment in collaboration with the firm. They randomized the location of e-commerce service points across 100 villages in three provinces of China and found substantial consumption welfare gains for adopters of e-commerce. E-commerce provided rural households with cheaper, more, and higher-quality products, as well as improved shopping

amenities. E-commerce thus raised household purchasing power without significantly increasing villagers' retail expenditure. However, the study did not find that e-commerce generated significant short- to medium-term production gains (e.g., nominal income change, entrepreneurship).

This research does not aim to propose a new method of measuring the digital economy, but it demonstrates the difficulty of capturing the welfare changes of digital services. The shift from offline to online commerce leads to product variety, quality and convenience gains. Therefore, such research has to collect detailed information on every online/offline transaction each household has conducted, and on the products offered at the local physical stores. This leads to a longitudinal survey of two rounds from 2,800/3,800 rural households and 11,500 price quotes from physical village stores. Measuring the welfare of other digital services will be even more complicated if it involves free services and new products. And this is why the International Monetary Fund and national governments have called for new methods to measure the digital economy.[60]

How Data Valuation Problems Matter Politically

As data plays an increasingly important role as an economic input, GDP (per capita) can generate greater measurement error when used to proxy for economic strength. But why does this statistical problem matter politically?

[60] https://www.imf.org/en/News/Seminars/Conferences/2018/04/06/6th-statistics-forum

A sole focus on GDP makes it difficult to accurately assess the relative strengths of world economies. It may distort Chinese perceptions of the outside world, and foreign perceptions of Chinese economic resilience. For instance, Brynjolfsson et al. (2019) designed a new measure of economic power by incorporating the welfare contribution of free digital products, which they called GDP-B. Adding Facebook alone causes GDP-B to grow by 1.91% per year; by comparison, average real GDP growth over the same period was 1.83% per year. In other words, Facebook alone added 0.05–0.11% to GDP-B annual growth. The US growth rate has therefore been underestimated.

A look beyond GDP also helps examine the resilience of the Chinese regime. Many have taken the recent decline of GDP growth as a sign of weakness. It is puzzling that – despite the overwhelming consensus that the Party relies on economic performance for legitimacy – the slowdown has not led to any major political unrest. One explanation might be the rise of the digital economy that has unleashed consumer welfare.

Conclusion

As digital technologies transform how individuals work, trade and live, they have also begun to influence how states interact. At the core of this change is the proliferation of data as a strategic asset. Many have equated data to oil to highlight the value of data. But as this paper indicates, data differs from oil in nature, and data politics operates on unique foundations that are worth researching.

At a deeper level, data is affecting international politics by changing the basis of power. A state's strength now lies not only in its military power or trade, but its capacity to collect, refine, and utilize data. Data has also become a new frontier of geopolitical conflict. As indicated by the US–China feud over 5G, the battle for digital supremacy will ignite looming tensions between states. And the Russian meddling in social media around 2016 U.S. election shows that personal data can be used as a powerful weapon to influence foreign states.

Data politics is at play, and it will only increase in salience over time.

References

Acemoglu D, Makhdoumi A, Malekian A, Ozdaglar A. Privacy-constrained network formation. Games and Economic Behavior 2017;105:255-275.

Acquisti A, Taylor C, Wagman L. The economics of privacy. Journal of economic Literature 2016;54(2):442-92.

Agrawal A, Gans J, Goldfarb A. Prediction machines: the simple economics of artificial intelligence. Harvard Business Press; 2018.

Aitken A. Measuring Welfare Beyond GDP. National Institute Economic Review 2019;249(1):R3-R16.

Allen, MMC. "Comparative capitalisms and the institutional embeddedness of innovative capabilities." *Socio-Economic Review* 11, no. 4 (2013): 771-794.

Brynjolfsson E, Collis A. How Should We Measure the Digital Economy?. Harvard Business Review; 2019.

Brynjolfsson E, Collis A, Diewert WE, Eggers F, Fox KJ. GDP-B: Accounting for the value of new and free goods in the digital economy (No. w25695). National Bureau of Economic Research; 2019.

Bughin J, Lund S. The ascendancy of international data flows. Vox EU 2017;9.

Chander A, Lê UP. Data nationalism. Emory LJ 2014;64:677.

Choi JP, Jeon DS, Kim BC. Privacy and personal data collection with information externalities. Journal of Public Economics 2019;173:113-124.

Carriere-Swallow MY, Haksar MV. The economics and implications of data: an integrated perspective. International Monetary Fund 2019.

Couture V, Faber B, Gu Y, Liu L. Connecting the Countryside via E-Commerce: Evidence from China. AER: Insights 2020.

Dickson BJ. Red capitalists in China: The party, private entrepreneurs, and prospects for political change. Cambridge University Press; 2003.

Erlich Y, Shor T, Pe'er I, Carmi S. Identity inference of genomic data using long-range familial searches. Science 2018;362(6415):690-694.

Farboodi M, Mihet R, Philippon T, Veldkamp L. Big data and firm dynamics. In AEA papers and proceedings 2019;109:38-42.

Farrell, H., & Newman, A. L. (2019). Of privacy and power: The transatlantic struggle over freedom and security. Princeton University Press.

Gallagher M, Miller B. Legitimation and control: Social media governance in china. Working Paper 2018.

Goldfarb A, Trefler D. AI and international trade (No. w24254). National Bureau of Economic Research; 2018.

International Monetary Fund (IMF). Measuring the Digital Economy; 2018.

Jernigan C, Mistree BF. Gaydar: Facebook friendships expose sexual orientation. First Monday 2009;14(10).

Jin GZ. Artificial intelligence and consumer privacy (No. w24253). National Bureau of Economic Research; 2018.

Jones CI, Tonetti C. Nonrivalry and the Economics of Data (No. w26260). National Bureau of Economic Research Working Paper 2019.

King G, Pan J, Roberts ME. How censorship in China allows government criticism but silences collective expression. American Political Science Review, 2013;107(2):326-343.

King G, Pan J, Roberts ME. Reverse-engineering censorship in China: Randomized experimentation and participant observation. Science 2014;345(6199):1251722.

Landry PF, Lü X, Duan H. Does performance matter? Evaluating political selection along the Chinese administrative ladder. Comparative Political Studies 2018;51(8):1074-1105.

Landry PF. (2008). Decentralized Authoritarianism in China: the Communist Party's control of local elites in the post-Mao era. New York: Cambridge University Press; 2008. p. 42-48

Lee KF. Tech companies should stop pretending AI won't destroy jobs. MIT Technology Review 2018.

Li WC, Nirei M, Yamana K. Value of data: there's no such thing as a free lunch in the digital economy. US Bureau of Economic Analysis Working Paper, Washington, DC 2019.

Liu J. China's data localization. Chinese Journal of Communication, 2020;13(1):84-103.

Liu L. From Click to Boom: The Political Economy of E-Commerce in China (Doctoral dissertation, Stanford University) 2018.

Liu L, Weingast B. Law, Chinese Style: Solving the Authoritarian's Legal Dilemma through the Private Provision of Law. Working paper 2019.

Liu L, Weingast B. Taobao, Federalism, and the Emergence of Law, Chinese Style. Minnesota Law Review 2018;102(4).

MacCarthy M. New directions in privacy: Disclosure, unfairness and externalities. ISJLP 2010;6:425.

Milhaupt CJ, Zheng W. Beyond ownership: state capitalism and the Chinese firm. Geo. LJ 2014;103:665.

Oi JC. Rural China takes off: Institutional foundations of economic reform. Univ of California Press; 1999.

Ong LH. Thugs and outsourcing of state repression in China. The China Journal 2018;80(1):94-110.

Polishchuk, L. "Institutional outsourcing." VOPROSY ECONOMIKI 9 (2013).

Pearson MM. China's New Business Elite: The Political Consequences of Economic Reform. Univ of California Press; 1997.

Roberts ME. Censored: distraction and diversion inside China's Great Firewall. Princeton University Press; 2018.

Shih V, Adolph C, Liu M. Getting ahead in the communist party: explaining the advancement of central committee members in China. American Political Science Review 2012;106(1):166-187.

Shirk SL. The political logic of economic reform in China (Vol. 24). Univ of California Press; 1993.

Steinfeld, ES. Playing our game: Why China's rise doesn't threaten the West. Oxford University Press, 2010.

Tsai KS. Adaptive informal institutions and endogenous institutional change in China. World Politics 2006;59(1):116-141.

Wallace JL. Juking the stats? Authoritarian information problems in China. British Journal of Political Science 2016;46(1):11-29.

Woetzel J, Seong J, Wang KW, Manyika J, Chui M, Wong W. China's digital economy: A leading global force. New York: McKinsey Global Institute 2017.

Varian Hal. Artificial Intelligence, Economics, and Industrial Organization in Ajay K. Agrawal, Joshua Gans, and Avi Goldfarb, eds., The Economics of Artificial Intelligence: An Agenda, University of Chicago Press, 2018.

Xu X. To Repress or to Co-opt? Authoritarian Control in the Age of Digital Surveillance. American Journal of Political Science 2020.

Huang Y. Capitalism with Chinese characteristics: Entrepreneurship and the state. Cambridge University Press; 2008.