12-3-2019

Give us a little social credit: to design or to discover personal ratings in the era of Big Data

Abigail Devereaux
New York University

Linan Peng
DePauw University, linanpeng@depauw.edu

Give us a little social credit: to design or to discover personal ratings in the era of Big Data

Abigail Devereaux[1]* and Linan Peng[2]

[1]New York University, Economics, New York, NY, USA and [2]Texas Tech University, Free Market Institute, Lubbock, TX, USA
*Corresponding author. Email: abigail.devereaux@gmail.com

Abstract

In 2014, the State Council of the Chinese Communist Party announced the institution of a social credit system by 2020, a follow-up to a similar statement on the creation of a social credit system issued by the State Council in 2007. Social credit ratings of the type being developed by the State Council in partnership with Chinese companies go beyond existing financial credit ratings in an attempt to project less-tangible personal characteristics like trustworthiness, criminal tendencies, and group loyalty onto a single scale. The emergence of personal credit ratings is enabled by Big Data, automated decision-making processes, machine learning, and facial recognition technology. It is quite likely that various kinds of personal and social credit ratings shall become reality in the near future. We explore China's version of its social credit system so far, compare the welfare and epistemological qualities of an ecology of personal ratings emanating from polycentric sources *versus* a social credit rating, and discuss whether a social credit system in an ideologically driven state is less a tool to maximize social welfare through trustworthiness provision and more a method of preventing and punishing deviance from a set of party-held ideological values.

Keywords: Institutions: design; formation, and operations; Institutions and the macroeconomy

1. Introduction

Personal ratings systems (PRSs) are scoring systems that attempt to project salient characteristics of an individual onto some scale or grade indicative of trustworthiness and other qualities. Financial credit ratings, a form of PRS, are hegemonic in the USA and the EU. The sharing economy relies heavily on mutually rating both sides of a transaction using some variation of a five-point star rating. Matching apps for romantic and professional partnerships use proprietary algorithms utilizing salient user-provided information considered representative of important relationship variables. Good ratings drive traffic to products and services and open more opportunities to the highly rated consumer. Bad ratings, if bad enough, may ban both providers and users from a given platform.

In 2014, as part of its Twelfth Five-Year Plan, the State Council of the Chinese Communist Party (CCP) issued a document titled, 'Planning Outline for the Construction of a Social Credit System'. 'The social credit system is an important part of the socialist market economic system and social governance system', the document begins, tying the fate of Chinese socialism-with-capitalistic-elements to the social credit system (*shehui xinyong tixi* – SCS). According to the State Council of the Chinese Communist Party (2014), the requirements of an SCS are 'establishing the idea of a sincerity (*chengxin*) culture, and promoting honesty and traditional virtues, it uses encouragement for trustworthiness (*shouxin*) and constraints against untrustworthiness (*shixin*) as incentive mechanisms, and its objective is raising the sincerity consciousness and credit levels of the entire society'. The SCS is designed to

go much further than foster trustworthiness in economic interactions. It is also intended to cover moral and political actions.

The 2014 document follows up on a 2007 document called the 'The General Office of the State Council's Opinions Concerning the Construction of a Social Credit System' (State Council of the Chinese Communist Party, 2007), where the State Council outlines the concept of a social credit system whereby 'the construction of a social credit system involves all aspects of economic and social life'. Especially in combatting commercial and industrial fraud, the State Council claims that '[i]t is necessary to play the role of chambers of commerce and associations, promote industry credit construction and industry self-discipline'.

The idea of social credit as a dualistic financial and moralistic trustworthiness mechanism has a deeper history in China: the *dang'an* system of personal files of grades, employment records, and criminal records required for activities like promotion, applying for a baby permit, and receiving a retirement pension and the *hukou* system of household registration (Creemers, 2018). However, we do not believe the SCS in practice is an extension of the personnel dossier system to non-Party members, or an extension of earlier traditions of controlling and shaping the morality of Chinese citizens. The SCS is an unprecedented country-wide standardization of the normative values of Beijing based on new technologies like algorithmic techniques on Big Data and automated decision-making processes.

Part of the motivation for the SCS is to encourage more Chinese to adopt credit, as only about 20% of the Chinese population had a credit report in 2012 (Zhou, 2012). But it seems unlikely that financial credit scoring would not have eventually existed in the absence of a CCP-designed credit score. Unlike most other countries with established credit reporting, China was still largely a cash-based society in the first decade of the 21st century. Chinese household debt to GDP in 2006 was very low, around 10%, though it was to rise steadily to about 30% in 2014 and to over 50% by the end of 2018. In 2006, the People's Bank of China (PBoC) created a 'Credit Reference Center' intended to be a credit-scoring bureau utilizing information supplied by banks (Creemers, 2018). But few people had bank accounts, so the bureau didn't have enough information to develop reliable scoring.

Unexpectedly, the key to modernizing the cash-only Chinese society resided in the advent of smartphones. By 2013, millions of Chinese people had eschewed cash for the digital payment app AliPay, whose transactional data provide a basis for credit scoring. Though executives at Ant Financial (an affiliate of Alibaba) initiated the development of credit scoring using that data, the CCP launched its own social credit score initiative the following year. Ant Financial's Sesame (Zhima) Credit remained one of the top contenders for the SCS until early 2018, when PBoC revoked the credit scoring licenses of the eight commercial credit scoring companies and issued a 3-year license to Baihang Credit, of which the eight companies are 8% shareholders and the National Interest Finance Association of China (NIFA) is a 36% shareholder (PBoC, 2018; Reuters staff, 2018a; Xie 2018).

Whether China's SCS is ultimately realized as a single numerical score or not, it is intended to reflect the social desirability and undesirability of an individual, company, or regional government's behaviors as defined within the constraints of CCP ideology. High enough social credit scores earn agents a place on a 'redlist', and low enough scores on a 'blacklist', a public database of 'discredited individuals' (*laolai*) maintained by the Supreme People's Court of China. Several provinces have devised ways to publicly shame *laolai*, including placing their pictures and personal information on giant screens in parks and squares and replacing ringtones with voice-overs proclaiming: 'the person you are calling is a dishonest debtor' (Zeng, 2018). Moreover, in recent years, different levels of courts (county, city, and provincial) have amplified the enforcement on dishonest debtors, which has been facilitated by Skynet, the state-sponsored Big Data program (Xinhuanet, 2019). For example, the execution bureaus of courts not only have increased the frequency of forcible execution of orders of courts on dishonest debtors, but also been able to track down dishonest debtors who fled, which was very difficult to achieve previously without Big Data.

The SCS is just one manifestation of the use of Big Data to intentionally alter or control individual behavior. Behavioral Big Data, what's being called 'surveillance capital', has become big business in the 21st century (McCahill and Finn, 2014; Zuboff, 2015, 2019). For every individual every day, private

companies and public entities collect huge volumes of personal data: GPS coordinates, satellite imagery, online gaming behavior, how long employees use which apps on their work computers, search results on Google and Baidu, products viewed on Amazon and Alibaba, customer loyalty card usage, PayPal and AliPay transactions, 'liked' posts on Facebook and Twitter, comments made on Reddit and news sites, steps taken and hours spent sitting or exercising, emails and SMS sent and received.

The private mining and use of surveillance capital by large companies to alter behavior, dubbed 'surveillance capitalism', has been cause for recent alarm (Zuboff, 2019). Of particular alarm is the monopolistic quality of some of the major companies, like Google and Facebook, who collect and analyze data using black-box proprietary algorithms to determine how best to modify user behavior to generate returns for their commercial partners. Surveillance capitalism poses a threat to user choice and liberty in the absence of the pressure of competition to soften the predatory edge of decision-altering algorithms. In robust and agile markets, the use of surveillance capital will, like other products, tend to benefit the user as much as the developer. In monopolized and inagile markets, the use of surveillance capital becomes more worrisome from a human rights perspective, particularly when the algorithmic use of data is arrogated by a single entity with the power to write law and conduct military operations. The corpus of one's personal surveillance capital represents a relatively effortless digital footprint of what was once too difficult to trace or legally untraceable. Such a digital footprint could be used to further the agenda of the data owner, particularly if the algorithms and objectives of data use are obscured for reasons of intellectual property or national security.

Users seem to understand that algorithms on their data can be highly imperfect and have the potential to be misused. Of particular concern are ratings systems developed out of surveillance capital. In 2016, the personal ratings app Peeple went live and was almost universally panned by its users and the media.[1] Despite the attitude that rating people on a single scale is 'horrible, cruel, and disgusting' as one Peeple user wrote in his iTunes app store review (see footnote 2), privacy laws in the West have not kept up with Big Data advances (Hartzog and Selinger, 2013). In 2016, the European Union approved the General Data Protection Regulation (GDPR), granting a 'right of explanation' of all decisions made algorithmically or using automated decision-making processes. However, some have noted ambiguity in the regulation that obviates the feasibility of defining or protecting such a right (Wachter *et al.*, 2017). In 2017, China passed a cybersecurity law, updated in 2018, which allows China's Ministry of Public Security (MPS) to conduct penetration tests and record data from any computer system for the purpose of cybersecurity (MPS, 2018).

In this paper, we assume data democratization is inevitable, and will greatly reduce the transaction costs of both collecting surveillance capital and developing personal ratings. We therefore focus not on the cost of ratings systems or whether or not they'd emerge in a polycentric fashion, but on the epistemological content of personal ratings that arise in a polycentric fashion *versus* the epistemological content of centrally designed personal ratings. We also consider the more expedient, adaptive uses of ratings systems that arise from polycentric *versus* centralized sources of surveillance capital, particularly as methods of tracking and controlling individual behavior.

The utility, coordinativity, and potential danger of personal ratings depend strongly on the mix of political and economic institutions in which they and their surveillance capital are developed, much in the same way that a decentralized price discovery process has different informational qualities than a centralized price setting process (Hayek, 1945). We place our case-study, China, in a framework of exploration *versus* exploitation, then consider how centralized *versus* polycentric ratings processes effect the epistemological and coordinative qualities of ratings. The exploitation *versus* exploration framework, as opposed to traditional equilibrium analysis, is essential for understanding the different effects of public *versus* polycentric innovation, as exploration and exploitation in combination are analogous to creative destruction (Gupta et al., 2006; March, 1991; Ostrom, 2005, 2010; Simon, 1996).

---

[1]As of 15 March 2018, the Peeple app has a 1.7/5 rating on the iTunes app store. The top 1-star review pans the app on principle, calling the idea of rating people 'horrible, cruel, and disgusting' (user JackTripodi, posted on 4 December 2016).

2. A study of centralized social ratings: China's SCS

Chinese planning takes the form of a 'champagne pyramid'. Officials identify some dominant contradiction at the top of the pyramid that impedes social progress, then identify the subproblems and sub-subproblems derivable from the dominant contradiction that must be resolved so society can progress to its next stage of development[2] (Creemers, 2017). Official resolution of dominant contradictions requires institutional design, the political form of which is expert rule (Koppl, 2018). The dominant contradiction in the SCS design document is a lack of trust and trustworthiness in financial transactions, commercial exchanges, personal interactions, judicial actions, and in government (State Council of the Chinese Communist Party, 2014).

Social credit ratings are a type of PRS geared toward social ends that do not necessarily reflect the subjective ends of rated individuals. Social credit ratings define trustworthiness as a hegemonous concept subject to design of experts, in contrast to fragmented personal ratings emergent from polycentric sources subject to competition and network effects present in regular markets and thus reflective of a subjective, context-dependent form of trustworthiness. It is noteworthy that the catch-all characteristic of the SCS, in that it reflects economic, moral, and political actions, is not a bug of the CCP's scoring system, but a feature. It is this mix of economic, moral, and political characteristics in contrast to an ecology of fragmented and specialized ratings that is the subject of our modeling Section 3. First, we take a thorough look at how the SCS has been designed, rolled out, and received thus far.

Proposed measures of the SCS

There are three basic requirements for the workability of China's SCS, according to the State Council of the Chinese Communist Party (2014): (1) a credit system covering the entire country and its supporting database, (2) a 'trustworthiness incentive mechanism', and (3) an 'untrustworthiness disciplinary mechanism'. The language validating the expedient necessity of an SCS argues that the SCS is, essentially, coordinative of the plans of individuals and businesses in the Chinese market socialist system. The document stipulates that efforts will be made to improve the government image of honesty and transparency, through standards the government sets for itself. The disciplinary mechanism is designed to be harsh, with one former party official stating that the social credit system is crafted so that 'discredited people become bankrupt' (Liu, 2018).

Businesses are a special focus of the report. Chinese corporations rely largely on state-controlled credit, as there is a missing market in rewarding good investments and punishing bad investments. The banking sector is highly regulated, which has had the unintended emergence of a $9 trillion shadow banking sector to enable firm-bank-firm loans, trust loans, and peer-to-peer loans, often to higher-risk debtors who do not have access to legal credit (Weinland, 2019). The SCS is also intended as a way to track compliance with safety and quality standards, blacklist companies that do not comply, and use the blacklist to target market bans and force withdrawals of some businesses. Banks are brought into the credit rating system as a way of reducing fraud. Tax evasion is a particular focus by using the social credit system's databases to better track property ownership and income. Tax violators go on a blacklist (Credit China, 2019). Pricing comes under even stricter control to 'standardize market price order'. The document urges the use of social credit scores as determinants of who contracts with whom in any given business deal, effectively shutting people out of commercial life if their credit scores are too low.

Despite the many stipulations about citizen behavior and blacklisting, there is no system in place for preventing abuse by officials, ensuring data quality and score accuracy, allowing citizens to dispute scores, or protecting citizen privacy. By the end of 2018, there were over 14 million individuals on the blacklist, with over 4 million added in 2018 alone, and only 2.2 million taken off the blacklist as a result of having engaged in some forms of mitigating behavior or reparations (Credit China,

---

[2]For example, see discussions of Dobson and Kashyap (2006) on the banking system reform and Coase and Wang (2012) on the Great Leap Outward in the late 1970s.

2019). As Chen and Cheung (2017) put it, '[i]ndividuals risk being reduced to transparent selves before the state in this uneven battle' (357).

Furthermore, it is unclear the Chinese suffer such a paucity of trust services to justify the creation of a single-scaled centralized social credit system. The 21st century Chinese business practices demonstrate elaborate trust-gaining practices, where social gatherings with new employees and clients are essential to establishing professional trust (Chua, 2012). Though the Chinese had relatively low credit usage as of 2014 (36.3% of people aged 15+ compared to 51.4% in the United States), they had high levels of formal bank account ownership (79%) (Fungáčová and Weill, 2014). Furthermore, Chinese credit usage as of 2014 was a good deal *higher* than China's wealthy protectorate Hong Kong (25.8%), and higher even than long-developed France (28.3%) (WorldBank Global Findex, 2014). While China's household debt to GDP was only 10% in 2006, it increased to 52.6% by the end of 2018, just below Germany (Trading Economics, 2019).

Furthermore, the apparent credit-shyness of Chinese citizens may not be due to a lack of general trustworthiness but to the state's heavy-handedness in banking, and the ways in which credit and collection in China differs from credit and collection in developed democracies. Chinese banks are nationalized; credit is issued from and controlled by the Chinese government. Credit is semi-collectivized at the familial level, such that Chinese banks can force family-members, spouses, and even ex-spouses to settle unpaid debts (Wang and Buckley, 2019). In some circumstances, like for a 'malicious overdraft', the CCP considers unpaid debts to be criminal offenses. The debtor must then prove the overdraft was non-malicious (Liu, 2011). Even if a Chinese citizen avoids a visit from the police after defaulting, they may still get a visit from debt collectors. Chinese debt collectors employ harsher tactics than in developed democracies and have been known to kidnap people and hold them until their debts are repaid (Feng, 2017).

Technical structure of the SCS

During the first period of the SCS rollout (2014–17), China gave eight companies the opportunity to come up with a credit scoring system. The two biggest contenders were WeChat developer Tencent's partner, China Rapid Finance, and Ant Financial's Sesame (Zhima) Credit. Zhima Credit was by far the most popular, used by 58% of respondents in a study by Kostka (2018), followed by Tencent Credit at 31%. Scores are linked to a Chinese citizen's official ID (*shenfenzheng*). A different app for Chinese youths, called China Youth Credit Management or CY Credit, is being developed by Tsinghua's Unigroup in partnership with the Communist Youth League Central Committee (CYLCC) and the National Development and Reform Commission (NDRC) (Jing, 2019).

By the beginning of 2018, the PBoC approved Baihang Credit Scoring to collect information, store data, and use that data to create credit scores for individuals and companies (Reuters staff, 2018b). The eight companies who worked on developing their own version of social credit scoring are shareholders of Baihang Credit. It is not yet clear whether or not Baihang Credit will rate citizens on Zhima Credit's scale, or if it will use a different scale. It is clear in the planning documentation (in particular the State Council of the Chinese Communist Party, 2014, 2016) that officials intend for the SCS to be a single, comprehensive score that encompasses economic, moral, and political actions. The final version of the SCS may result in a sorting mechanism for redlists and blacklists, finer control in the form of CCP-approved behavioral nudges, or both.

To get a sense of how the rollout has worked so far, credit scores administered through Zhima Credit range between 350 and 950 points. Some credit scores, like very early versions of 'mass credit' (*dazhong xinyong*) tested in 2010 in Suining, Jiangsu province, begin at 1000 points and result in a graded A–D classification (Creemers, 2018). Regardless of the ratings range, a feature of most of the credit scores tested so far is that users earn or lose points based on things they do (Botsman, 2017a). In the case of Zhima Credit, its scoring takes into account five categories of action-types that are weighted and subject to a proprietary algorithm: (1) financial credit history (35%), (2)

purchase-related behavior (25%), (3) having a steady income and possessing assets (20%), (4) personal information (15%), and (5) social behavior on the platform (5%) (Creemers, 2018).

Charity work, separating and recycling rubbish, donating blood, and sharing links from a state-sponsored news agency praising the CCP are the kinds of deeds that earn citizens points on their credit score (Osborne, 2018; Zeng 2018). Citizens with high social credit scores can be rewarded with free gym facilities, high-speed broadband access, discounts on energy bills, cheaper public transport, foreign travel visas, shorter wait times in hospitals, better insurance premiums, better interest rates at banks, no-deposit rentals, higher quality schooling for their children, access to better restaurants, public approbation, and better dating pools on dating sites like Baihe (Ma, 2018; Margolis, 2017; Zeng, 2018).

Bad deeds that negatively impact one's social credit score include a wide range of behavior, not all forms of which are explicitly known. Most citizens find the social rating system opaque (Chen and Cheung, 2017). Beyond the obvious – holding too much unpaid debt – behavior that hurts one's credit score includes: '…not showing up to a restaurant without having cancelled the reservation, cheating in online games, leaving false product reviews, and jaywalking' (Zeng, 2018); sharing a news article about China's recent stock market collapse (Osborne, 2018); not walking your dog enough, refusing to carry out military service, spending too much time playing video games or posting on social media, spreading 'fake news', and being a bad driver (Ma, 2018); eating, sleeping, or playing loud music on the train, reading the wrong books, not paying utility bills on time, having the wrong friends, saying negative things about one's friends or economy or government, having friends or family that say negative things about their friends or economy or government, and of course, mentioning Tiananmen Square (Botsman, 2017a).

If one's score is low enough, or one's behavior is deemed bad enough by the CCP, she is placed on a publicly accessible blacklist. Citizens with bad scores are 'discredited individuals' (*laolai*). Sometimes, all it takes is a single offensive action to find oneself blacklisted. Early documents establishing the practice and motivation for blacklisting describe it as using restrictions and public shaming to punish non-compliance with court-ordered punishments and fines (Creemers, 2018: 14). According to the Supreme People's Court (2010), which establishes the major framework for the SCS blacklist, being blacklisted can restrict you from: buying plane or train tickets, especially for traveling abroad; sending your children to private schools; eating in certain restaurants and staying in certain hotels; purchasing insurance; purchasing or renovating property; obtaining financial credit; obtaining social security and other benefits; establishing a social organization; obtaining government contracts; becoming a Party member; joining the military; managing a state-owned enterprise (SOE) or acquiring other management or 'responsible' positions; taking jobs in the food sector, the drugs sector, or in construction; using public parks, fishing areas, beaches or forests. Reportedly, the full range of punishments is much larger and continually expanding. For instance, your dog may be taken away from you for not walking it enough (Ma, 2018), and your internet can be throttled (Botsman, 2017a).

Dissent can earn a citizen a blacklisting, as happened in the case of the journalist Liu Hu who vocally criticized the Chinese government. In 2017, Hu attempted to buy a plane ticket and discovered he'd been deemed 'not qualified'. He later discovered that he could not buy property, take out a loan, or travel on the posher Chinese trains (Vanderklippe, 2018). As of September of 2019, there had been over 15 million cases of blacklisted individual, while those blacklisted individuals had been barred from over 31 million flights and 6 million high-speed train trips (Court Order Execution, 2019).

A 'redlist' of highly rated citizens also exists, though it is much less clear how to positively impact one's SCS and get on the redlist than it is how to negatively impact one's score and get on the blacklist (Engelmann *et al.*, 2019).

The SCS rollout

The SCS has so far consisted of a fragmented set of initiatives, undertaken by different companies and counties, and rolled out in different regions at different times, some before the issuance of the SCS

foundational document (State Council of the Chinese Communist Party, 2014). Not all SCSs have been digital; some towns developed manual pen-and-paper SCSs. Gradation differed, as did blacklisting and redlisting, among the different rollout cities (Mistreanu, 2018). Citizen feedback was integrated into some of the experiments. Officials revised their SCS in Suining County over intense pushback over point deductions for making 'unauthorized' petitions to higher authorities (Hao, 2018). While it is reported some prominent blacklisted individuals were not notified in advance of their status nor given an option to appeal (Wang, 2017), General Office of the CCP *et al.* (2016) does include notification and appeal in its blueprint for the blacklist.

It must be stressed that regardless of the low-tech experiments run in some counties, the SCS is intended to utilize large volumes of data and automate decision-making processes in its administration and operation (Hoffman, 2018; Creemers, 2018: 20–2). Facial recognition glasses, facial recognition at ATMs, and facial recognition cameras on Chinese streets have already resulted in thousands of arrests, including high-profile arrests of wanted criminals who had alluded capture for years. China's Skynet and Xue Liang programs are geared toward complete coverage of all public places by surveillance cameras by 2020, including the development of a centralized database of footage available at the local and national levels (Zhang, 2018). Even as public facial recognition is being tested in other places like London (Dearden, 2018), it carries a bevy of concerns about privacy, false positives, and racial discrimination. Recently, Microsoft pulled its largest database of almost 100,000 individuals over ethical concerns (Murgia, 2019). It was the largest facial recognition set in the world, used by several large companies including Alibaba, Sensetime, and Mengvi – the two latter companies have created Chinese equipment used to track Uighur minorities in Xinjiang province. Arrests and the internment of religious minorities like the Uighurs in China have raised concerns about human rights and lack of oversight in a country where government officials both have license to exploit user data at will and where facial recognition is coupled with automated decision-making processes that can immediately and algorithmically impact social credit scores (Vincent, 2018a, 2018b).

Reception of the SCS by Chinese citizens

The SCS has been surprisingly popular. Over 80% of respondents in a study by Kostka (2018) either somewhat or strongly approve of SCSs, 19% remain neutral, and perhaps most strikingly, only 1% either strongly or somewhat disapprove of SCSs. The study author notes that the survey results may be skewed positive by their authoritarian context (ibid.: 9), though he believes the SCS doubters are likely part of the neutral 19% (ibid.: 11). Older people, wealthier people, more educated people, and urbanites were more likely to approve of SCSs than other groups. People generally approved of local government-run credit systems (64%) more than commercial credit systems (55%) (ibid.: 12–15). SCSs are largely perceived, thus far, as useful tools that provide more access to credit, help citizens avoid scams, improve environmental regulatory compliance, and standardize ratings system for, say, the food safety of local restaurants (ibid.: 21–22). In total, 76% of respondents agree there is a problem with mutual trustworthiness in China (ibid.: 22). Some citizens say they have seen the behavior of their fellow citizens respond positively during the implementation of the social credit system in their area (Ma, 2018). The survey and interview data indicate that trustworthiness mechanisms are useful tools; respondents found both local and commercial ratings systems to be useful in various ways, with somewhat higher approval for the government-run solutions.

These results do not imply that a centrally designed and administered SCS is necessary or better than the alternative in the long run. While the lack of interpersonal trust in China has been correlated with marketization in some research (Zhang and Xin, 2019), it seems to contradict the observed high and average levels of interpersonal trustworthiness in marketized places with Chinese language and culture like Hong Kong and Taiwan, respectively (Ward *et al.*, 2014). Additionally, both places have western-style systems to determine financial trustworthiness for determining eligibility for

loans and credit. In total, 61% people in Hong Kong have three or more credit cards in their wallet, and 64% people use mobile wallets (JD Power, 2019).

3. The emergence of personal ratings from polycentric sources

As economists, we are especially interested in the subjective value of personal ratings, as proxied by their epistemological quality. Personal ratings are intended to represent individual characteristics and behaviors are correlated with trustworthiness in commercial, financial, and social interactions. Trust and trustworthiness are central considerations to any system of personal ratings, as noted in China's SCS rollout documents (State Council of the Chinese Communist Party, 2014, 2016). Personal ratings are intended to be trust-enhancing; the more general and less specific the ratings, the more broadly they are intended to be indicative of trustworthiness.

Trust is present in '[v]irtually every commercial transaction' (Arrow, 1982) and can radically change the outcome of any transaction (Ostrom, 2010). Trust modeling in the neoclassical framework has often taken the form of suggesting ways to ameliorate incomplete, asymmetric information, or costly information (Akerlof, 1970). Often, anticipatory regulation is suggested to rectify problems of information (see discussion in Klein, 1997: 3). Anticipatory regulation isn't strictly necessary, as evidenced by the increasing number of organizations (including the Canadian government and Dell) who have begun to adopt guiding principles to help reduce deadweight losses in the event that some abnormally bad outcome during the execution of a contract between two parties (Frydlinger and Hart, 2019).

Communication is an informal institution. Communication between individuals to reduce uncertainty and improve information symmetry is central to the development of trust. Developing a forum for communication requires common rules and norms of the space, typically encapsulated by membership, regional, or ritualistic restrictions. See, for example: the creation of gossip networks (Merry, 1984), chatting (Tullock, 2005), going out for cocktails with potential business partners to build *guanxi* (Botsman, 2017b), congregating in particular establishments with membership restrictions as in the London cafes that hatched modern stock trading (Stringham, 2015), and the creation of tracking mechanisms like those used by Japan's financial clearinghouses (Ryser, 1997). Such networks can utilize information more efficiently than established leadership hierarchies (Banerjee et al., 2014).

Shearmur and Klein (1997) argue that when societies become more anonymous, we need some other way of signaling trustworthiness, what they call 'seals of approval' (Klein, 1997: 4). Thus, society is a 'a flowing patchwork of reputational communities' (ibid.: 5) where people seek out and aspire to obtain seals of approval. Businessmen in the early 20th century rural North Carolina, for instance, carried proof of membership in the Baptist church as evidence of their trustworthiness (Weber, 2002(1948)). One of the primary signals of entry-level employability in modern developed democracies is some form of university degree or certificate (Caplan, 2018). Systems to produce seals of approval are typically rooted in social norms and small-town practices, and though these services tend to expand in size and scope and become standardized, they also become more expensive the larger the society and more anonymous the transaction (Shearmur and Klein, 1997: 6–7, 37).

The quality of any seal of approval depends on (1) how well the seal of approval correlates with characteristics relevant to the success of any trust-requiring transaction, and (2) the network of people who co-produce and utilize the seal of approval. Businessmen in rural North Carolina relied on the reputation of the Baptist church for its trustworthy membership and on a large network of people who were both members of the Church and who interact with members of the Church. More recently, distributed and decentralized online interactions, which exhibit the informational asymmetry of anonymity, have given rise to many new kinds of seals of approval like product and provider star-ratings, seller feedback, and mutual provider/user ratings in the sharing economy. Distributed and decentralized ledger technology, exemplified by the blockchain and related applications like smart contracts, is greatly reducing the cost of enforcing contracts in large and anonymous markets (Norgaard et al., 2018).

Large volumes of continuously updating data coupled with rapidly improving technologies for harnessing and making sense of that data will undoubtedly continue to provide new means and lower costs for *both* the centralized and decentralized generation of seals of approval. Before they were explicitly tasked and licensed to do so, Ant Financial planned to develop a credit scoring system when the cost of obtaining detailed personal and transactional information drastically reduced thanks to the ubiquity of smartphones and AliPay. As such, for our analytical purposes, we are specifically interested in the comparative advantages and disadvantages of a developed ecology of personal ratings emanating from polycentric sources *versus* a developed centralized system of hegemonous personal ratings.

The emergence and construction of personal ratings

Especially as it is rooted in information asymmetries and free rider problems, trustworthiness has been frequently categorized as a public good (Anomaly, 2017). In traditional economic analysis, trustworthiness is a solution to a commitment problem whereby the gains of deception in the short run are balanced against the gains of long-term cooperation. This calculus dictates the extent to which privately provided trust is effective. Since it never pays to be honest in a society characterized by a paucity of trustworthiness, such a state of affairs is self-reinforcing, and thereby (according to a static equilibrium analysis) virtually impossible to alter without public intervention.

Personal ratings represent another way to mitigate the cost of trustworthiness in large and anonymous societies. Personal ratings are seals of approval indicative of personal characteristics relevant to an interpersonal transaction. When hiring a rideshare driver, riders might care about drivers who demonstrate good driving skills, hygiene, and perhaps, conversation. Similarly, drivers might prefer riders to be quiet, sober, and clean. Whether or not either driver or rider is attractive might not matter to the quality of the ride for some; for others, it might. Specialized personal ratings of this sort are ubiquitous, though typically localized to a certain platform of interaction. Specialized ratings are specialized in that they refer to a small set of characteristics deemed relevant to signaling trustworthiness or quality in a particular interaction, rather than the entire corpus of an individual's characteristics or behaviors as in social credit ratings.

PRSs may also consist of how many followers or 'likes' someone has and how often their content has been shared by others (common on platforms like Twitter and Instagram), the difference between how often their comments have been upvoted or downvoted (common on platforms like Reddit), and whether their reviews have been rated 'helpful' by other users (common on platforms like Yelp and Amazon). Giant databases like Experion specialize in acquiring information about people along as many axes possible, and are often used by legal entities, employers, and the government to score a person's credit or criminal history.[3]

There are three delivery methods for PRSs: *crowd-sourced*, where anonymous strangers determine the rating; *algorithmic*, where the app developers create a black-box algorithm that determines the rating; and *mutual*, where two (or three) people known to each other rate each other from their interaction. There are three types of PRSs: *single-scale*, *single-aspect*, and *multi-aspect*. Single-scale ratings attempt to reduce the salient bits of an individual to a single number (Peeple[4], Zhima Credit). These apps have raised the most ethical concerns among users (see Reagle (2015) on Peeple and similar apps, and Margolis (2017) on Zhima Credit). These apps can be crowd-sourced (Peeple) and algorithmic (Zhima Credit), though the television show *Black Mirror* foreshadowed how single-scale ratings could conceivably be mutual. Single-aspect ratings attempt to rate particular aspects of a person like their professional capabilities (RateMyProfessors[5], Uber, AirBnB)

---

[3]Note especially that we do not include scorable characteristics like IQ, EQ, MBTI, or automated decision-making algorithms associated with parole decisions in what we consider to be personal ratings, as they are not generated by virtue of some ongoing social process.

[4]https://itunes.apple.com/us/app/peeple/id1008896593?mt=8.

[5]http://www.ratemyprofessors.com.

attractiveness (Spontana[6]), intelligence (Best IQ Test[7]), and MBTI-type personality (PersonalityMatch[8]). These apps can be crowd-sourced (Spontana), algorithmic (Best IQ Test), and mutual (Uber). Multi-aspect ratings allow people to rate each other (and themselves) on many axes, most notably a feature of dating apps (OkCupid, Tindr, eHarmony) and employee/employer-rating apps (Dots[9], Pluggd, TaskRabbit). These apps can be crowd-sourced (Dots), algorithmic (OkCupid), and mutual (Pluggd, TaskRabbit).

We summarize and provide examples of PRSs in Table 1 (including scorable characteristics). Unless marked with a single asterisk (*), all personal ratings are commercially offered.

For PRSs to be coordinative of the ends of users, they should advance a user's subjective ends. If we consider traditional credit scores, a credit score is useful insomuch as it signals a debtor's creditworthiness. In the language of risk assessment creation, a risk assessment is useful in that it assigns likelihoods of individuals engaging in some kind of behavior that matches the distribution of future observed behavior. Lenders don't care so much if a potential client is a good driver, or whether she shared an article critical of her country on Twitter. Multi-aspect ratings of this sort provide a better signal if their composition is based on a narrow set of characteristics highly correlated with the desired signal. The algorithmic method of relating characteristics to signal is also important. Certainly, ratings should weight characteristics most relevant to the signal the most. Characteristics might themselves be ratings, or simple states that represent historical behavior.

Suppose a particular PRS is based on k characteristics $C^1, \ldots, C^k$. Characteristics should be rated on or converted to the same numerical scale $[\sigma_1, \sigma_2]$, $\sigma_1 < \sigma_2$. Define a mapping $s{:}C^k \times N \to \gamma \in [\sigma_1, \sigma_2]$, where $N = [1, \ldots, i, \ldots, n]$ and such that there are n agents in the system. Then, if C_i^j is the particular value of that characteristic for agent or household i, $s(C_i^j)$ is the score of that value on the range $[\sigma_1, \sigma_2]$. We can then define weights α_i^j for each characteristic C_i^j, such that $\sum_{j=1}^{k} \alpha_i^j = 1$. A good PRS is one where characteristics are weighted according to their relevance to a particular transaction.

A simple model of a generalized PRS might look something like:

$$\text{PRS}(C^1, \ldots, C^k, i) = \alpha_i^1 s(C_i^1) + \alpha_i^2 s(C_i^2) + \cdots + \alpha_i^k s(C_i^k) \tag{1}$$

By specifying characteristics as 'bins' corresponding to traits considering positively correlated to the desirable behavior or feature represented by the credit score, the coefficients as scores for each bin, and the function *PRS* as a mapping from individual i with a vector of coefficients $\boldsymbol{\alpha}_i$ to a score proxying a single aspect (like trustworthiness), our formulation corresponds with generalized models of risk assessment (Kleinberg *et al.*, 2016).

A particular instance of a personal rating as defined in (1) then adjusts the weights according to the relevance to the desired signal and the demographics of the individual or agent type i. A SCS is a special case of a generalized PRS where the weights are not demographically adjusted, i.e. are the same for everyone in all transactional contexts. We can write an SCS like so:

$$\text{SCS}(C^1, \ldots, C^k, i) = \alpha^1 s(C_i^1) + \alpha^2 s(C_i^2) + \cdots + \alpha^k s(C_i^k) \tag{2}$$

A good indicator is a metric that correlates well with whatever hidden information it presumes to proxy. Perhaps creditworthiness has something to do with the length someone has been at their current job, for instance, but it's likely less important than whether or not and for how long they've paid their bills on time.

Algorithms for data collection and inclusion are also highly important to the quality of personal ratings based on these characteristics. Does a pattern-matching algorithm intended to detect faces

---

[6]https://spontanaapp.com.
[7]https://itunes.apple.com/us/app/best-iq-test-pro/id1047228171?mt=8.
[8]https://play.google.com/store/apps/details?id=com.personalityperfect.app&hl=en.
[9]Childers (2017).

Table 1. Examples of PRSs, by type of personal rating and delivery mechanism

	Single-aspect	Multi-aspect	Single-scale
Crowd-sourced	Instagram; RateMyProfessors; Spontana	Facebook; Dots; Reddit	Peeple
Algorithmic	IQ[a]; MBTI[a]	eHarmony; OkCupid; Tindr; background checks[b]; automated decision-making[a]	Zhima Credit[b]; insurance premiums[b]
Mutual	Uber; Lyft	Pluggd; TaskRabbit; AirBnB	Black Mirror's *Nosedive*

[a]Scorable characteristics, but not social ratings.
[b]These have elements of public scoring, depending on the system under consideration.

in a crowd actually do so with reasonable accuracy? Does a matching algorithm intended to match romantic partners actually increase the number of successful relationships compared to brute-force search? Like for tangible products, algorithms compete with each other to do a better job satisfying some user's subjective ends, and can suffer a loss of quality absent a robust competitive environment. Since actual market transactions generate crowd-sourced and mutual ratings, they should be more accurate (Page, 2007) and less prone to expert and epistemological bias (Koppl, 2018). However, there can be economies of scale in personal ratings, especially considering that ratings are network goods.

The construction of ratings may favor more centralized or public aggregation methods *if* the requisite technological infrastructure to aggregate, filter, track, and publish ratings isn't advanced enough to make small-scale ratings feasible, as summarized in Table 2. We assume in this paper, however, that technology has penetrated a system enough to allow personal ratings to emerge from polycentric sources.

Can algorithmic, single-scale ratings systems proxy trustworthiness?

In this section, we consider the epistemological difference between two systems: an ecology of personal ratings emanating from polycentric sources, and a centralized PRS. We keep our analysis brief and mostly illustrative, as a full model that includes the emergence and co-production of personal ratings as a network good is out of the scope of this particular paper. To frame the analysis, we rely heavily on the model of creditworthiness developed by Glover and Corbae (2015).

Glover and Corbae set up the problem of issuing debt to households with hidden characteristics as highly correlated with their ability to pay-off the debt in the future. Credit scores are calculated by looking at a household's observable default history, $d^t = (d_0, \ldots, d_{t-1})$, and developing a probability $\pi_{i,t}(b_t, d^t)$ at time t that a household i pays b_t in some future period. The credit score proxies that probability; the better the proxy (signal), the more profitable the lender. Suppose Q_t is the initial payout at time t to household i by the lender. Then we can write the lender's profit function $\Pi_{i,t}$ as

$$\Pi_{i,t}(Q_t, b_t, d^t) = -Q_t + \pi_{i,t}(b_t, d^t)b_t \tag{3}$$

A lender will engage a debtor i iff $\Pi_{i,t} > 0$. If $\pi_{i,t}$ is a good proxy for hidden characteristics that correlate well with i's ability to pay b_t in the future, then utilizing $\pi_{i,t}$ tends to maximize the profit of lenders and offer credit to all borrowers who are indeed able to pay back their loans, thereby maximizing their utility from taking debt.

A social credit rating can have two functions: (a) as a proxy for trustworthiness that is generally coordinative of interactions requiring trust or assurance, or (b) as a method for producing behavioral

Table 2. The tradeoffs between the accuracy and cost of delivering personal ratings given a high volume of market transactions, by type of delivery method

Delivery method	Accuracy (high volume)	Cost (high volume)
Crowd-sourced/mutual	High	High
Algorithmic	Low	Low

compliance in citizens regardless if it is in their best subjective interest. The first function is that of a benevolent designer, who seeks to increase social welfare in a Pareto or Kaldor–Hicks sense. We shall consider whether a SCS is a better tool than an ecology of PRSs in the subjective welfare sense, to start, then consider whether the SCS in China, in its current and probable future form(s), is a tool of welfare enhancement or a tool of social control.

Let's use (3) as a way of investigating how well the SCS can proxy for an indicator like $\Pi_{i,t}$. For the SCS to be a good proxy, we want to see how well it correlates with $\pi_{i,t}$:

$$\pi_{i,t}(b_t, C_i^1, \ldots, C_i^k) = \gamma_{i,t} + \beta_{i,t} SCS(C_i^1, \ldots, C_i^k) + \epsilon_{i,t} \tag{4}$$

where $\gamma_{i,t}$ is some constant at time t, $\beta_{i,t}$ is correlation coefficient, and $\epsilon_{i,t}$ is an error term. Plugging (2) into (4), we can rewrite the above as:

$$\pi_{i,t}(b_t, C_i^1, \ldots, C_i^k) = \beta_{i,t}\alpha^1 s(C_i^1) + \beta_{i,t}\alpha^2 s(C_i^2) + \cdots + \beta_{i,t}\alpha^k s(C_i^k) + \epsilon_{i,t} \tag{5}$$

Suppose now we consider the cluster of correlations between the characteristics and the probability that household i provides the producer a benefit b_t in the future. We only include significant characteristics, non-negative correlations, and non-negative scores:

$$\pi_{i,t}(b_t, C_i^1, \ldots, C_i^k) = \Gamma_{i,t} + \chi_{i,t}^1 s(C_i^1) + \cdots + \chi_{i,t}^k s(C_i^k) + \mu_{i,t}$$

To the extent that the SCS is a good indicator of the probability that household i will provide benefit b_t in the future, we'd expect $\chi_{i,t}^r$ to be close to $\beta_{i,t}\alpha^r$, and $\alpha^r = 0$ for insignificant characteristics. Clearly, the SCS cannot include characteristics for which (1) there is no significant correlation to $\Pi_{i,t}$, and especially for which (2) there is a negative correlation to $\pi_{i,t}$.

Create a distance measure $\delta(\chi_{i,t}^r, \beta_{i,t}\alpha^r)$ between the coefficients $\chi_{i,t}^r$ and $\beta_{i,t}\alpha^r$, perhaps the mean squared distance. When δ is small, the SCS may be a good proxy for $\pi_{i,t}$. But δ will fail to be a good proxy for $\Pi_{i,t}$ when the measured characteristics of people are not good proxies for their ability or willingness to provide benefit b_t in the future. The failure is twofold, producing false negative for some groups of people and a false positive for other groups of people. We believe our results, though not as rigorously proved, align well with similar results regarding the tradeoffs between predictability and fairness as demonstrated for risk assessments like algorithms to help determine criminal recidivism (Kleinberg *et al.*, 2016).

In an ecology of personal ratings emanating from polycentric sources, other ways of measuring trustworthiness could emerge to handle groups that do not fit into some 'standard pattern' of trustworthiness. The SCS would then tend to reduce overall welfare as it would provide a false signal that would expose the producer to more risk than profitable and/or bar households from engaging in utility-increasing interactions based on inaccurate perceptions of their untrustworthiness in that particular domain of interaction. Go back to the example of taking credit: a bad indicator of creditworthiness results in lost producer *and* consumer surplus in the credit market. Furthermore, a SCS subverts the Kleinian format of trust services beginning in community practice and therefore tends to reflect the values of the designers and not the community those services were designed to assist.

What recourse do groups of people (including producers) have when a SCS fails? In an ecology of personal ratings, they can simply opt to use another system, though the question of switching costs and the compatibility between ratings systems must be addressed. Since our thesis is about comparing a mature ecology of personal ratings with a social credit system, it stands to reason that as long as the transaction costs of switching between ratings systems are lower than the opportunity costs of using a subpar ratings system, an ecology of personal ratings should outperform a social credit system in terms of social welfare.

4. The epistemological considerations of exploration

How does a polycentric ecology of PRSs do *versus* a SCS when dynamically utilizing knowledge to produce trust services – that is, does a SCS outperform or underperform a polycentric ecology of PRSs?

Social systems are characterized by creative evolution, the co-production of unpredictable and unprestatable system-level outcomes which then feedback into the system, changing individual behaviors as unpredictably and unprestatedly (Beckage *et al.*, 2013). Exploitation is exemplified by neoclassical search intended to dissipate the advantages of a current institutional framework. Exploitation works by, say, finding more complete contractual arrangements to reduce the uncertainty surrounding the expected behaviors in any given interaction or exchange (Barzel, 1997; Cheung, 1983). Exploration involves evolving the system into entirely new spaces of economic opportunities and available actions by means of an endogenous novelty-generation mechanism. Exploration works via, say, a small set of entrepreneurs expanding the set of potentially trust-enhancing solutions and thereby disrupting stable product spaces (Kirzner, 1997; Koppl *et al.*, 2015).

Exploitation tends to be more tractable to analysis than exploration because it closes the system under consideration. As such, a great deal of ink has been spilled by economists using an exploitative framework of analysis. Under stable conditions where change is constrained to a recognizable pattern of variance, the exploitative framework can take an analyst far. These are not the characteristics of a creatively evolving socioeconomic system, certainly not one the size of China's.

Suppose for the sake of argument that a SCS is welfare-enhancing compared to an existing alternative. That does not imply the SCS is welfare-enhancing in the long run *versus* a long-run alternative. SCSs are brittle to knowledge problems, as demonstrated in the previous section. Even if an SCS is designed to be statistically useful in the absence of some other ratings system, people change, societies change, technology further develops. The algorithmic structure, normative teleologies, and the data on which SCSs depend are past-oriented. Making SCSs predictive, such that scores indicate the probability someone will be trustworthy in the future, does not solve this problem. Predictive algorithms are also past-oriented due to the nature of the data on which they learn. These problems do not differ in substance from those explicated by Hayek (1945) with respect to utilizing knowledge under planning.

Polycentric systems under market orders allow for public, private, and public–private solutions to problems of trustworthiness provision; more importantly, polycentric systems are context-aware and dynamically adaptive. Examples of successful polycentric provision of public goods include polycentric policing services (Ostrom, 2010, Stringham, 2015), lighthouses and lightships (Candela and Geloso, 2018; Foldvary and Klein, 2003), common-pool resource provision (Ostrom, 2010), and even stock market regulation (Stringham, 2015). The experimentation phase of the SCS in China was itself, to some degree, polycentric; China revoking the credit-scoring licenses of its eight commercial contenders collapses that polycentricity to a single, centralized source, as if the epistemological efficiency of polycentric systems is only important in the first draft of an idea and not in its dynamic realization and re-realization. It is but one indication that the SCS is intended as a tool of social and political stability as much as it is intended to be a trustworthiness mechanism.

Polycentric systems under exploration tend to have their hegemonous orders, equilibria, and institutions disrupted by the discovery of novel solutions to social problems. Disruption isn't a bug of polycentric systems, it's a feature. In order for knowledge to advantage everyone in the

system, processes must be allowed to unfold in unknowable and potentially disruptive ways. Under exploration, entrepreneurs continuously disrupt systemic ratings hegemonies with potentially more coordinative methods of providing assurance. Mutual ratings in the sharing economy and star ratings for traditional products and services alert users to untrustworthy producers, unsafe business practices, and other salient characteristics, albeit imperfectly. Deep Web entrepreneurs of dark e-commerce sites are solving the problem of extreme social extendedness and anonymity by employing blockchain-ledger smart contracts that vet and punish defectors on either side of a sales contract, essentially symmetrizing risk under the conditions of extreme information asymmetry (Norgaard et al., 2018).

Polycentric orders under exploration are characterized by complex and nonlinear systemic behavior such that outcomes cannot be predicted via the aggregation of individual actors engaged in standard Bayesian reasoning (Arthur, 2013; Epstein, 2006; Miller and Page, 2009). Normative and centralized social credit ratings represented by equation (2) have no endogenous learning mechanism to improve the ratings system, unlike the natural winnowing and incentives to innovate present in a marketized polycentric order. The SCS is designed to reflect the normative values of the CCP, whose expert-constructed realization may fail to realize either the CCP's ends or the end of maximizing subjective welfare, especially when values change.

Thus, expert design of and rule over PRSs can fail along a variety of axes (Koppl, 2018). Experts must pick sides. Groups whose subjective values are closest to the normative values promoted by the experts will be winners in expert-constructed systems. Those whose subjective values deviate from the normative values of experts lose. In China, dissidents certainly lose – the 2019 Hong Kong protests against legalizing extradition of alleged criminals from Hong Kong to the Chinese mainland succeeded in part due to the ability of protesters to subvert facial recognition and data collection of their activities (Allen-Ebrahimian, 2019; Cuthbertson, 2019); religious people from the relatively large Uighur Muslim and tiny Kaifeng Jewish communities lose (Mozur, 2019). Not only are these groups not compensated under some Kaldor–Hicks scheme to maximize social welfare, they are punished, often harshly. Unlike polycentric PRSs, SCSs do not ask each person what she believes is best for her, they impose a worldview of what is best for that person onto her.

Furthermore, China's movement toward algorithmic processes involving feedback loops and automated decision-making is troubling from both a scientific and human rights perspective (Brehm and Loubere, 2018). Karen Yeung (2017) discusses many of the problems in automated decision-making and -guiding of what she calls *hypernudge*, a Big Data-enabled, algorithmic form of nudging behaviors. Devins *et al.* (2017) look at the algorithmization of decision-making and -guiding processes from the perspective of law and regulation. From the perspective of subjective welfare maximization, the 'Big Data paradigm' as they call it perpetuates "the belief that theory is no longer necessary because applied mathematical and statistical techniques – based on algorithms – can 'analyze' data and find optimal solutions better than human programmers" (ibid.: 7). The authors note especially that theoryless correlations of the sort generated by machine learning on Big Data tend to overweight averages, recursively augment bias, and definitionally ignore creative evolution.

Calude and Longo (2017) warn against a veritable 'deluge of spurious correlations' in Big Data sets, such that machine learning on these sets may be intractably opaque to objective inference (assuming a benevolent social planner). The other side of the spurious correlations coin is missing data: characteristics that are more salient to the trustworthiness in certain interactions may be ignored because they are costlier to determine than less salient characteristics. Consider the false inference that if a person is good at A, she must be good at B or C or D: if A is 'attractiveness', this reduces to the infamous halo effect (Rosenzweig, 2007). Unfortunately, the problems above aren't very tractable: adding more data come with more spurious correlations, selectively adding data come with bias. Finally, Devereaux (2019) discusses how centrally administered cybernetic hypernudge suffers from the same epistemological problems as top-down central planning, a grave concern for the future of an algorithmized, automatic SCS-behavior feedback loop.

Varian (2014) reminds readers that estimating complex models comes at a high computational cost, requiring variable selection over the set of potential predictors by the modeler. Thus, there is no theoryless data in data modeling, and where there's theory, there's room for the projection of bias onto the system, particularly in systems that are ideologically driven.

5. Conclusion

We have demonstrated that social credit systems, a subset of PRSs, are plausibly welfare-inferior to private ratings that are allowed to emerge from polycentric sources and form an ecology in which ratings complement each other, compete, and adapt to the ever-changing values of users. Fragmenting personal ratings along axes relevant to users *and* closer to user knowledge improves the epistemological content of that information. Algorithmic delivery has the potential to further drain ratings of salience and representativeness, and is subject to bias and spurious correlations.

Polycentric ecologies of personal ratings will not be universally subjectively useful; nothing is. Nor will all possible useful personal ratings emerge within a single institutional ecology. 'Smart' technologies, like data-driven technologies that use algorithms to improve user experiences and outcomes, are emerging in a piecemeal fashion. Furthermore, the emergence of blockchain technologies may be quite disruptive to many current social and political hegemonies (Cowen, 2018) and may require different institutional ecologies to be most effective (MacDonald, 2019).

However, SCSs are plausibly *superior* at attaining the normative goals of a ruling political party. China's SCS is not designed to maximize the subjective value of its users, but to align Chinese citizens with the ideology of the CCP and to punish deviants and other destabilizing elements. In this, it might be effective. China's is an authoritarian system. In a different socioeconomic system with a mix of public and private services, citizens can ostensibly appeal to a public entity when they have been abused by a private entity. In addition, the institution of the market grants users some measure of competitive and financial power to reallocate to themselves welfare lost to participating in unfair ratings systems. In China, deviations and non-conformity merit punishment in the form of blacklisting, which is designed to be severe in comparison to infractions (State Council of the Chinese Communist Party, 2016). Despite a promised process of notification and appeal, the SCS is not opt-out; citizens have no real recourse if they believe the SCS or its system of appeals is abusive or ineffective.

SCSs are, therefore, a useful tool to keep the powerful in power, by identifying and punishing those whose behavior threatens the hegemony of the powerful. China has recovered in many ways from the catastrophe of Maoist economics and the Cultural Revolution. Measures like the SCS threaten to take China back to a time where China was more focused on the ideology of communism than advancing the well-being of its people. The SCS weaponizes Big Data against its citizenry. The CCP's stated reasons for the creation of the SCS sound good on paper, but upon reflection dangle the freedom and wealth of market participation in front of Chinese citizens like a carrot, transforming a force that has greatly improved the lives and trajectories of entire communities into a lever of control.

References

Akerlof, G. A. (1970), 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism', *The Quarterly Journal of Economics*, **84**(3): 488–500.

Allen-Ebrahimian, B. (2019), 'How China's repression playbook backfired in Hong Kong', *Vox*, 29 August, https://www.vox.com/world/2019/8/29/20835183/hong-kong-protests-china-repression-tiananmen-square-beijing-response (accessed 5 September 2019).

Anomaly, J. (2017), 'Trust, Trade, and Moral Progress: How Market Exchange Promotes Trustworthiness', *Social Philosophy and Policy*, **34**(2): 89–107.

Arrow, K. (1982), *Gifts and Exchanges*, Princeton, NJ: Princeton University Press.

Arthur, W. B. (2013), *Complexity Economics*, Oxford: Oxford University Press.

Banerjee, A., A. G. Chandrasekhar, E. Duflo and M. O. Jackson (2014), 'Gossip: Identifying central individuals in a social network', *National Bureau of Economic Research*.

Barzel, Y. (1997), *The Economic Analysis of Property Rights*, Cambridge: Cambridge University Press.

Beckage, B., S. Kauffman, L. J. Gross, A. Zia and C. Koliba (2013), 'More Complex Complexity: Exploring the Nature of Computational Irreducibility Across Physical, Biological, and Human Social Systems', In *Irreducibility and Computational Equivalence*. Berlin, Heidelberg: Springer, pp. 79–88.

Botsman, R. (2017a), 'Big Data meets Big Brother as China moves to rate its citizens', *Wired.co.uk*, 21 October, https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion (accessed 18 November 2018).

Botsman, R. (2017b), *Who Can You Trust?: How Technology Brought Us Together and Why It Might Drive Us Apart*, London: Penguin UK.

Brehm, S and N. Loubere (2018), 'China's dystopian social credit system is a harbinger of the global age of the algorithm', *Phys.org*, 15 January, https://phys.org/news/2018-01-china-dystopian-social-credit-harbinger.html (accessed 21 March 2018).

Calude, C. S. and G. Longo (2017), 'The Deluge of Spurious Correlations in big data', *Foundations of Science*, **22**(3): 595–612.

Candela, R. A. and V. J. Geloso (2018), 'The Lightship in Economics', *Public Choice*, **176**(3–4): 479–506.

Caplan, B. (2018), *The Case Against Education: Why the Education System is a Waste of Time and Money*, Princeton, NJ: Princeton University Press.

Chen, Y. and A. S. Cheung (2017), 'The Transparent Self Under Big Data Profiling: Privacy and Chinese Legislation on the Social Credit System', *The Journal of Comparative Law*, **12**(2): 356–378.

Cheung, S. (1983), 'The Contractual Nature of the Firm', *The Journal of Law & Economics*, **26**(1): 1–21.

Chua, R. Y. (2012), 'Building Effective Business Relationships in China', *MIT Sloan Management Review*, **53**(4): 1.

Coase, R and N. Wang (2012), *How China Became Capitalist*, New York, NY: Palgrave Macmillan.

Court Order Execution (2019), Available publicly on http://zxgk.court.gov.cn/ (accessed 19 August 2019).

Cowen, T. (2018), 'Why does tech have so many political problems?' *Marginal Revolution*, 5 August, https://marginalrevolution.com/marginalrevolution/2018/08/tech-many-political-problems.html (accessed 18 November 2018).

Credit China (2019), "2018年失信黑名单年度分析报告发布." www.gov.cn, 19 February, http://www.gov.cn/fuwu/2019-02/19/content_5366674.htm (accessed 22 June 2019).

Creemers, R. (2017), 'Part I: Not a Moonshot, but a Legacy of Central Planning', In 'China's Plan to 'Lead' in AI: Purpose, Prospects, and Problems', *New America* online, 1 August, https://www.newamerica.org/cybersecurity-initiative/blog/chinas-plan-lead-ai-purpose-prospects-and-problems/ (accessed 8 November 2019).

Creemers, R. (2018), 'China's Social Credit System: An Evolving Practice of Control', SSRN.

Cuthbertson, A. (2019), 'Hong Kong protesters use lasers to avoid facial recognition cameras and blind police', *The Independent*, 1 August, https://www.independent.co.uk/news/world/asia/hong-kong-protests-lasers-facial-recognition-ai-china-police-a9033046.html (accessed 5 September 2019).

Dearden, L. (2018), 'Facial recognition cameras scanning unwitting tourists and Christmas shoppers in London's West End', *The Independent*, 17 December, https://www.independent.co.uk/news/uk/home-news/facial-recognition-cameras-london-met-police-suspects-arrests-identity-a8687481.html (accessed 19 June 2019).

Devereaux, A. N. (2019), 'The Nudge Wars: A Modern Socialist Calculation Debate', *The Review of Austrian Economics*, **32**(2): 139–158.

Devins, C., T. Felin, S. Kauffman and R. Koppl (2017), 'The Law and Big Data', *Cornell JL & Public Policy*, **27**, 357.

Dobson, W. and A. K. Kashyap (2006), 'The contradiction in China's gradualist banking reforms', *Brookings Papers on Economic Activity*, **2006**(2): 103–162.

Engelmann, S., M. Chen, F. Fischer, C. Y. Kao and J. Grossklags (2019), 'Clear Sanctions, Vague Rewards: How China's Social Credit System Currently Defines Good and Bad Behavior', *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 69–78), ACM.

Epstein, J. M. (2006), *Generative Social Science: Studies in Agent-Based Computational Modeling*, Princeton, NJ: Princeton University Press.

Feng, E. (2017), 'China 'granny Gang' Jailed in Lending Clampdown', *Financial Times*, 11 August, https://www.ft.com/content/34ad1c98-7cd1-11e7-9108-edda0bcbc928 (accessed 11 June 2018).

Foldvary, F. E. and D. B. Klein (Eds.), (2003), *The Half-Life of Policy Rationales: How New Technology Affects Old Policy Issues*, New York: NYU Press.

Frydlinger, D and O. D. Hart (2019), *Overcoming Contractual Incompleteness: The Role of Guiding Principles (No. w26245)*, Cambridge, MA: National Bureau of Economic Research.

Fungáčová, Z. and Weill, L. (2014). A view on financial inclusion in Asian countries. *BOFIT Policy Brief* (August 2014).

General Office of the Communist Party of China, and General Office of the State Council. (2016). 'On the Facilitation of Supervision, Warning, and Punishment of Debtors". Chinese Communist Party, 25 September 2016. http://www.gov.cn/zhengce/2016-09/25/content_5111921.htm (accessed 1 September 2018).

Glover, A. and D. Corbae. (2015), 'A simple dynamic theory of credit scores under adverse selection', in *2015 Meeting Papers* (No. 1265), Society for Economic Dynamics.

Gupta, A. K., K. G. Smith and C. E. Shalley (2006), 'The interplay between exploration and exploitation.', *Academy of management journal*, **49**(4): 693–706.

Hao, K. (2018), 'Is China's social credit system as Orwellian as it sounds?' *MIT Technology Review*, 26 February, https://www.technologyreview.com/f/613027/chinas-social-credit-system-isnt-as-orwellian-as-it-sounds/ (accessed 25 June 2019).

Hartzog, W. and E. Selinger (2013), 'Big Data in Small Hands', *The Stanford Law Review*, **66**, 81.

Hayek, F. A. (1945), 'The Use of Knowledge in Society', *The American Economic Review*, **35**(4): 519–530.

Hoffman, S. (2018), 'Managing the State: Social Credit, Surveillance and the CCP's Plan for China', in N. Wright (ed), *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative*.

JD Power (2019). Hong Kong Retail Banking Satisfaction Study.

Jing, M. (2019), 'This small team is building a social credit system app for China's youth to determine who's naughty or nice', *South China Morning Post*, 25 March, https://www.scmp.com/tech/apps-social/article/3003158/small-team-building-social-credit-system-app-chinas-youth (accessed 22 June 2019).

Kirzner, I. M. (1997), 'Entrepreneurial discovery and the competitive market process: An Austrian approach', *Journal of economic literature*, **35**(1): 60–85.

Klein, D. B. (Ed.) (1997), *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, Ann Arbor, Michigan: University of Michigan Press.

Kleinberg, J., S. Mullainathan and M. Raghavan (2016), 'Inherent trade-offs in the fair determination of risk scores', *arXiv preprint, arXiv:1609.05807*.

Koppl, R. (2018), *Expert Failure*, Cambridge, UK: Cambridge University Press.

Koppl, R., S. Kauffman, T. Felin and G. Longo (2015), 'Economics for a Creative World', *Journal of Institutional Economics*, **11**(1): 1–31.

Kostka, G. (2018), 'China's Social Credit Systems and Public Opinion: Explaining High Levels of Approval', *New Media & Society*, 1461444819826402.

Liu, M. (2011), 'Credit Card Fraud Under Chinese Criminal Law', *Frontiers of Law in China* **6**(3): 369–86.

Liu, X. (2018), 'Social credit system must bankrupt discredited people: former official', *China Global Times*, 20 May, http://www.globaltimes.cn/content/1103262.shtml (accessed 19 August 2018).

Ma, A. (2018), 'China has started ranking citizens with a creepy 'social credit' system – here's what you can do wrong, and the embarrassing, demeaning ways they can punish you', *Business Insider*, 29 October, https://www.businessinsider.com/china-social-credit-system-punishments-and-rewards-explained-2018-4 (accessed 18 November 2018).

MacDonald, T. J. (2019), *The Political Economy of Non-Territorial Exit*, Cheltenham, UK: Edward Elgar Publishing.

March (1991), 'Exploration and exploitation in organizational learning.', *Organization science*.

Margolis, J. (2017), 'A Big Brother approach has qualities that would benefit society', *Financial Times*, 31 October, https://www.ft.com/content/ffe78e52-bd54-11e7-823b-ed31693349d3 (accessed 15 March 2018).

McCahill, M and R. L. Finn (2014), *Surveillance, Capital and Resistance: Theorizing the Surveillance Subject*, New York, NY: Routledge.

Merry, S. E. (1984), 'Rethinking gossip and scandal', in D. Black (ed), *Toward a general theory of social control*, Academic Press, pp.271–302.

Miller, J. H. and S. E. Page (2009), *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, (Vol. **17**), Princeton, NJ: Princeton University Press.

Ministry of Public Security (2018), 'Regulations on Internet Security Supervision and Inspection by Public Security Organs (公安机关互联网安全监督检查规定)', http://www.mps.gov.cn/n2254314/n2254409/n4904353/c6263180/content.html (accessed 23 June 2019).

Mistreanu, S. (2018), 'Life Inside China's Social Credit Laboratory', *Foreign Policy*, 3 April, https://foreignpolicy.com/2018/04/03/life-inside-chinas-social-credit-laboratory/ (accessed 9 October 2018).

Mozur, P. (2019), 'One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority', *The New York Times*, 4 April, https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html (accessed 5 September 2019).

Murgia, M. (2019), 'Microsoft quietly deletes largest public face recognition data set', *Financial Times*, 6 June, https://www.ft.com/content/7d3e0d6a-87a0-11e9-a028-86cea8523dc2 (accessed 19 June 2019).

Norgaard, J. R., H. J. Walbert and R. A. Hardy (2018), 'Shadow markets and hierarchies: comparing and modeling networks in the Dark Net', *Journal of Institutional Economics*, **14**(5): 877–899.

Osborne, S. (2018), 'China has made obedience to the State a game', *The Independent*, 22 December, http://www.independent.co.uk/news/world/asia/china-has-made-obedience-to-the-state-a-game-a6783841.html (accessed 3 March 2018).

Ostrom, E. (2005), *Unlocking Public Entrepreneurship and Public Economies* (No. 2005/01), Wider Discussion Papers/World Institute for Development Economics (Unu-Wider).

Ostrom, E. (2010), 'Beyond Markets and States: Polycentric Governance of complex Economic systems', *American Economic Review*, **100**(3): 641–72.

Page, S. E. (2007), 'Making the Difference: Applying a Logic of Diversity', *Academy of Management Perspectives*, **21**(4): 6–20.

People's Bank of China (PBoC) (2018), '关于百行征信有限公司（筹）相关情况的公示 [Notice on relevant information regarding Baihang Credit Co., Ltd.]', http://www.pbc.gov.cn/rmyh/105208/3456248/index.html (accessed 25 June 2019).

Reagle, J. (2015), 'The problem with rating people on the new app Peeple', *The Conversation*, 2 October, http://theconversation.com/the-problem-with-rating-people-on-the-new-app-peeple-48491 (accessed 26 November 2018).

Reuters staff (2018a), 'China to bar people with bad "social credit" from planes, trains', *Reuters*, 16 March, https://www.reuters.com/article/us-china-credit/china-to-bar-people-with-bad-social-credit-from-planes-trains-idUSKCN1GS10S (accessed 17 March 2018).

Reuters staff (2018b), 'China's first private credit-scoring firm starts operations: Xinhua', *Reuters*, 24 May, https://www.reuters.com/article/us-china-credit/chinas-first-private-credit-scoring-firm-starts-operations-xinhua-idUSKCN1IP02E (accessed 22 June 2019).

Rosenzweig, P. (2007), 'Misunderstanding the Nature of Company Performance: The Halo Effect and Other Business Delusions', *California Management Review*, **49**(4): 6–20.

Ryser, M. (1997), 'Sanctions without Law: The Japanese Financial Clearinghouse Guillotine and its Impact on Default Rate', *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, Ann Arbor: University of Michigan Press, pp. 225–240.

Shearmur, J and D. B. Klein (1997), 'Good Conduct in the Great Society: Adam Smith and the Role of Reputation', in *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, 29–45.

Simon, H. A. (1996), *The Sciences of the Artificial*, Cambridge, MA: MIT Press.

State Council of the Chinese Communist Party (2007), '国务院办公厅关于社会信用体系建设的若干意见 [Several Opinions of the General Office of the State Council on the Construction of Social Credit System]', http://www.gov.cn/zwgk/2007-04/02/content_569314.htm (accessed 22 June 2019).

State Council of the Chinese Communist Party (2014), '社会信用体系建设规划纲要 [Outline of Social Credit System Construction Planning]', http://www.gov.cn/zhengce/content/2014-06/27/content_8913.htm (accessed 13 March 2018).

State Council of the Chinese Communist Party (2016), 'Opinions Concerning Accelerating the Construction of Credit Supervision, Warning and Punishment Mechanisms for Persons Subject to Enforcement for Trust-Breaking', Translation used https://chinacopyrightandmedia.wordpress.com/2016/09/25/opinions-concerning-accelerating-the-construction-of-credit-supervision-warning-and-punishment-mechanisms-for-persons-subject-to-enforcement-for-trust-breaking/ (accessed 19 August 2018).

Stringham, E. (2015), *Private governance: Creating order in economic and social life*, Oxford University Press.

Supreme People's Court (2010), "Supreme People's Court's Provisions on Limiting the Discredited's Purchase", The Supreme People's Court, 20 October, http://www.court.gov.cn/shenpan-xiangqing-1650.html (accessed 8 November 2019).

Trading Economics (2019), 'China Households' Debt to GDP'. *Trading Economics*, https://tradingeconomics.com/china/households-debt-to-gdp (accessed 23 June 2019).

Tullock, G. (2005), in C. K. Rowley (ed.), *The Selected Works of Gordon Tullock. The Organization of Inquiry*, Indianapolis: Liberty Fund.

Vanderklippe, N. (2018), 'Chinese blacklist an early glimpse of sweeping new social-credit control', *The Globe and Mail*, 3 January, https://www.theglobeandmail.com/news/world/chinese-blacklist-an-early-glimpse-of-sweeping-new-social-credit-control/article37493300/ (accessed 8 November 2019).

Varian, H. R. (2014), 'Big data: New tricks for econometrics', *Journal of Economic Perspectives*, **28**(2): 3–28.

Vincent, J. (2018a), 'Artificial Intelligence is Going to Supercharge Surveillance, *The Verge*, 23 January, https://www.theverge.com/2018/1/23/16907238/artificial-intelligence-surveillance-cameras-security (accessed 6 March 2018).

Vincent, J. (2018b), 'Chinese police are using facial recognition sunglasses to track citizens', *The Verge*, 8 February, https://www.theverge.com/2018/2/8/16990030/china-facial-recognition-sunglasses-surveillance (accessed 6 March 2018).

Wachter, S., B. Mittelstadt and L. Floridi (2017), 'Why a Right to Explanation of Automated Decision-Making Does not Exist in the General Data Protection Regulation', *International Data Privacy Law*, **7**(2): 76–99.

Wang, M. (2017), 'China's Chilling 'Social Credit' Blacklist', *The Wall Street Journal*, 11 December, https://www.wsj.com/articles/chinas-chilling-social-credit-blacklist-1513036054 (accessed 19 August 2018).

Wang, A. and C. Buckley (2019), 'Stuck With an Ex-Husband's Debt, a Journalist Fights for Divorced Women', *The New York Times*, 20 December, https://www.nytimes.com/2019/12/20/world/asia/china-divorce-debts.html (accessed on 28 Dec 2019).

Ward, P. R., L. Mamerow and S. B. Meyer (2014), 'Interpersonal Trust Across Six Asia-Pacific Countries: Testing and Extending the 'High Trust Society' and 'Low Trust Society' Theory', *PLoS ONE*, **9**(4): e95555.

Weber, M. (2002 (1948)), *The Protestant ethic and the 'spirit' of capitalism and other writings*. London: Penguin.

Weinland, D. (2019), 'China Shadow Banking Cools for First Time in a Decade', *Financial Times*, 18 March, https://www.ft.com/content/fa026b16-492f-11e9-bbc9-6917dce3dc62 (accessed 8 November 2019).

WorldBank (2014), WorldBank Global Findex study, 2014. *WorldBank.com*. Data from https://databank.worldbank.org/home.aspx (accessed 21 March 2018).

Xie, S. Y. (2018), 'Jack Ma's Giant Financial Startup Is Shaking the Chinese Banking System', *The Wall Street Journal*, July 29, https://www.wsj.com/articles/jack-mas-giant-financial-startup-is-shaking-the-chinese-banking-system-1532885367 (accessed 8 November 2019).

Xinhuanet (2019), 'Answers to Overcoming the Difficulty of Enforcement of People's Courts' Orders', 12 August, http://www.xinhuanet.com/2019-08/12/c_1124863446.htm (accessed 8 November 2019).

Yeung, K. (2017), "Hypernudge': Big Data as a Mode of Regulation by Design', *Information, Communication & Society*, **20**(1): 118–136.

Zeng, M. J. (2018), 'China's Social Credit System puts its people under pressure to be model citizens', *The Conversation*, 23 January, https://theconversation.com/chinas-social-credit-system-puts-its-people-under-pressure-to-be-model-citizens-89963 (accessed 6 March 2018).

Zhang, J. (2018), 'The Xue Liang program, a boost to the Chinese market', in 'Top Video Surveillance Trends for 2018', *IHS Markit video surveillance group*.

Zhang, Y. and Z. Xin (2019), 'Rule Comes First: The Influences of Market Attributes on Interpersonal Trust in the Marketization Process', *Journal of Social Issues*, **75**(1): 286–313.

Zhou, C. (2012), 'Credit Information Database in China', *Presented at 4th Credit Reporting and Risk Management Conference, Kuala Lumpur, 5–9 November 2012*.

Zuboff, S. (2015), 'Big Other: Surveillance Capitalism and the Prospects of an Information Civilization', *Journal of Information Technology*, **30**(1): 75–89.

Zuboff, S. (2019), *The age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Profile Books.