Putting 'Good Citizens' in 'The Good Place'?

Jiahong Chen

21.06.2019 doi: 10.17176/20190621-122918-0

Debates over the Chinese Social Credit System frequently end up with either utopian or dystopian scenarios. While the Western media often present the image of an Orwellian totalitarian future, the system is sometimes considered an innovative way of social governance. In this contribution, I will aim to answer the question as to whether a Social Credit System will be more likely to lead a society to a 'digital republic' or a 'digital dictatorship'. After analysing how the Chinese Social Credit System exhibits an enormous gap between policy-making and policy-execution, I argue that instead of a utopia or dystopia, such a system is more likely to lead us to a future of 'digital bureaucracy'.

List-Based Measures vs Score-Based Measures

Other contributors are aware that, despite the important uniqueness of *the* Chinese Social Credit System, the discussion should be kept more generally on the desirability of *a* Social Credit System. As Siems and Mac Dithigh have pointed out, the Chinese Social Credit System takes three different forms in practice: Nationwide blacklists, compliance scores in pilot cities, and social credit scores by financial institutions. Both Reijers and Van't Klooster have chosen a more 'score-based' definition of a Social Credit System. Such a score-based image of a Social Credit System is probably also fuelled by the popularity of the *Black Mirror* episode 'Nosedive', where everyone is assigned a social rating.

It is important that certain practices under the umbrella notion of "Social Credit System" be distinguished, as they might have a significantly different impact on individuals and be subject to very different practical or theoretical justifications. For example, China's most oft-cited social credit policy documents – the 2014 Outline and the 2016 Opinion – do not include measures involving a 'score' or 'rating'. Rather, both documents are overwhelmingly dominated by list-based measures, i.e., blacklists for trust breakers, and 'red'-lists for trust keepers. This is also the case for Shanghai's Social Credit Ordinance 2017.

The prevalence of list-based approaches results largely from the fact it was the Supreme Court who introduced some of the first initiatives to deal with disobedience of court decisions. In the absence of individual insolvency law (and contempt of court law) in China, it has always been a challenge for courts to enforce judgments in civil cases effectively. In response, the Supreme Court decided to publish a list of those who wilfully refuse to obey court decisions (as a reputational punishment) and to restrict their access to certain services (akin to bankruptcy restrictions), which turns out to be rather effective. This partly explains why most of these initiatives did not meet much resistance – or were even welcome by the public – in the first place.

This is not to say, however, that these measures are intrinsically desirable, or that they are essentially different from the score-based measures. Yet, the relatively higher acceptability of these measures mirrors certain important features among different approaches to a Social Credit System. A blacklist released by courts may arguably be less worrisome in that it is subject to judicial control, applies only to serious and wilful disobedience, and does not involve opaque or arbitrary algorithms. These claims are open to challenges, but the point here is that certain practices falling within the scope of a Social Credit System are not necessarily unjustifiable, or at least not far from legal mechanisms that are commonly accepted across various jurisdictions (such as the US's 'no-fly lists'. For more examples, see the contribution by Siems and Mac Sitigh). The less discussed downside of a Social Credit System, as will be explained below, is its natural propensity to be exploited for bureaucratic purposes.

Policy on Paper vs Policy in Reality

One of the reasons why a Social Credit System is susceptible to bureaucratic abuses lies in its potential to be repurposed, which is what has happened in China. While the two key policy documents have not mandated score-based measures, such approaches are embraced by local authorities. In fact, it is well-documented that a large number of Chinese cities are already experimenting with a variety of schemes whereby local residents are treated differently based on their social credit scores. These schemes concern matters ranging from the right to relocate to a megacity to prioritised access to public services. Points can be gained or lost through a wide spectrum of behaviour, including volunteering or blood donation, something certainly not covered by the 2014 Outline or the 2016 Opinion, but implemented in the name of 'contributing to the construction of the Social Credit System'.

Therefore, a crucial point to raise is that, regardless of the arguably good intention and reasonable design of the overarching system, the mere fact that local or sectoral enforcers may easily distort such schemes will on its own warrant serious precaution. The 'function creep' of the Chinese Social Credit System is a case in point: The list-based approach targeting punishable behaviour as advanced by the central policies has ended up as a score-based, local implementation that encourages 'desirable' behaviour.

One might argue such abuse has less to do with the system itself, but more with who controls the system and what constraints are put in place – the 'regulatory environment' in Siems and Mac Sitigh's words. Indeed, under a totalitarian regime, a Social Credit System will be readily available for suppressing dissidents. However, even in a democratic society, there remains a high risk of such a system being repurposed for a fully legitimate policy goal. The criteria of what should be rewarded and what should be punished under a Social Credit System may be exploited to serve a new policy objective that is not necessarily related to the 'social creditworthiness' of an individual. For example, a local government may decide to encourage the consumption of local products by incentivising such purchase with credit points. While it is debatable whether such practice is justifiable, it shows how a Social Credit System, irrespective of the government it serves, is bound to create a strong motivation for bureaucrats to expand the scope of the system to areas without democratic debates or scrutiny.

Fairness vs Efficiency

Even if the Social Credit Systems are tightly restricted to certain areas, they would nevertheless be subject to another more fundamental criticism in relation to the theoretical and practical limitations of rational control over individual behaviour. Along with the rise of modern bureaucracy, this challenge arises from the long-standing debate in political philosophy over the extent to which collective efficiency should be prioritised over individuality. The ratings of individuals may be employed as a 'convenient' or 'efficient' way of allocating scarce resources by the government. In cities where social housing is in shortage, for example, local councils may decide to give priority to those with higher scores as it would be easier for them than assessing the degree of individual needs on a case-by-case basis. While this is still better than price mechanisms (Van't Klooster), it is not necessarily fairer than the 'old-fashioned' model that involves qualitative human judgment.

The functioning of a Social Credit System thus carries certain characteristics of a bureaucratic regime: 1) Rationalisation and optimisation of efficiency (quantitative evaluation and influence of behaviour of citizen); 2) specialisation and division of labour (credit recorders, calculators, enforcers and auditors from different sectors); 3) clear formal rules and limited discretion (criteria of 'positive' and 'negative' behaviour); and 4) impersonality (insensitivity to individual circumstances). The term 'bureaucracy' is used in a neutral sense and does not necessarily carry any desirable or undesirable connotations. That said, bureaucracy enabled by a Social Credit System would very likely have some unintended consequences, such as the scenario of 'computer says no', where citizens are denied services or resources when personal situations

are not taken into account. The mundanity of the activities covered by such a system and the potential diffusion of responsibilities would also subject a Social Credit System to further criticisms.

The next question would be, are these characteristics essential to the concept of a Social Credit System, or are they simply the results of the (ab)uses of such a system? While I agree with other contributors that the actual effectiveness and harm of a Social Credit System would depend on the regulatory environment (Siems and Mac Sitigh), a Social Credit System, as a social engineering project, brings forth changes in technological configurations and power dynamics by altering the regulability[1] of individual behaviour. Such potentials are independent of the conditions of the regulatory environment and should not be dismissed. A Social Credit System may serve a totalitarian regime or a democratic one, but either way, it will bring existing bureaucracy to another level.

Conclusion: 'The Good Place' and the Future of Bureaucracy

A Social Credit System faces two primary challenges regardless of the regulatory environment: First, it is prone to unchecked expansion of scope by bureaucrats; second, it strengthens the bureaucratic propensity pre-existing in a society. In answering the question of the debate – whether Social Credit Systems will lead to a 'digital republic' or a 'digital dictatorship' – one would have first to answer the question whether an ultimate bureaucratic society could be considered a utopia or a dystopia. To me, it could go either way, depending on the role of bureaucracy in that society. For a hypothetical society sitting perfectly between a republic and a dictatorship, a Social Credit System is most likely to reinforce bureaucracy – amplifying its merits, and also its flaws. The characteristics of a Social Credit Systems raised above – the tendency to be transferred for an incompatible policy objective or simply employed for administrative efficiency – are unlikely to drive the society to a utopian future but, at the same time, they will not necessarily lead the society closer to a totalitarian future, either.

While the idea of the Orwellian 'Big Brother' certainly captures dangers of a Social Credit System, a better metaphor seems to be the world of *The Good Place*, where one's afterlife will end up in 'The Good Place' or 'The Bad Place' based on the accumulated 'absolute moral worth' of their every single action, and the value of an 'undefined action' to be decided by a relevant department. Rather than a digital republic or a digital dictatorship, a Social Credit System is more likely to lead to a 'digital bureaucracy', a future where social credits are produced and consumed as a way of social indifference (rather than social suppression) imposed on individualism. On the macro-level, Social Credit Systems might be intended to make 'good citizens' (Reijers) – whether for the right or wrong reasons – but on the micro-level, they might be executed to make it easier for bureaucrats to make mundane decisions of who should go to 'The Good Place', and who should go to 'The Bad Place'.

Jiahong Chen is currently supported by the EPSRC grant "From Human Data to Personal Experience" (EP/M02315X/1), and was previously funded by the China Scholarship Council/University of Edinburgh Joint Scholarship Programme during his PhD.

[1] For further discussion of the notion of 'regulability', see Lessig, L. (1998). 'The laws of cyberspace.' *Readings in cyberethics,* 134, 136.

ABOUT THE AUTHOR

Jiahong Chen

Dr Jiahong Chen is a Researcher Fellow in IT Law of Horizon Digital Economy Research, University of Nottingham.Dr Jiahong Chen is a Researcher Fellow in IT Law of Horizon Digital Economy Research, University of Nottingham.

SUGGESTED CITATION

Chen, Jiahong: Putting 'Good Citizens' in 'The Good Place'?, Verfassungsblog, 21.06.2019, doi: 10.17176/20190621-122918-0.

Licensed under CC BY-SA 4.0.

Langzeitarchiviert bei IntRechtDok unter: doi.org/10.17176/20190621-122918-0.