



INSTITUTE FOR RESEARCH IN ECONOMIC AND FISCAL ISSUES

IREF Working Paper Series

A Matter of Trust: The Lockean Minimal State and its New Foe

Silviya Lechner

IREF WORKING PAPER NO. 202304

JULY 2023

IN ENGLISH: [EN.IREFEUROPE.ORG](http://en.irefeurope.org)
IN FRENCH: [FR.IREFEUROPE.ORG](http://fr.irefeurope.org)



INSTITUTE FOR RESEARCH IN ECONOMIC AND FISCAL ISSUES

A Matter of Trust: The Lockean Minimal State and its New Foe

Silviya Lechner
Chair of Political Science

Anglo-American University - Prague
silviya.lechner@aauni.edu

July 2023

Acknowledgements. I wish to thank the Institute for Research in Economic and Fiscal Issues (IREF) for the generous support of my research.

Abstract

Among libertarians, it has been conventional to define the minimal state as a state that enforces contracts and employs coercion to prevent force and fraud only. Its foundations are the basic rights of individuals to liberty, life, and estate. The usual competitor to the minimal state has been the redistributive state which allows individual rights to be restricted for the purposes of social and economic welfare. In this paper, I return to Locke's classical liberal theses and make a case for a Lockean minimal state that fleshes out the importance of trust. The core claim is threefold: 1) that a Lockean minimal state should be understood both as a rights-protecting and as a trust-based political authority; that 2) trust is fundamental to a free society; and that 3) the principal "foe" of the Lockean minimal state is not the redistributive state, but a novel type of social organisation called a social credit system. Social credit systems, as the one introduced in China since 2020, undermine both individual rights and trust or the moral glue that connects individuals inside a free society. The paper concludes with a warning about the possible spread of elements of social credit systems to Western liberal states.

1. Introduction

This paper makes a libertarian case for a Lockean minimal state, and identifies its new "foe": a recently emergent form of socio-political organisation known as a social credit system.¹ The main thesis is that their antagonism turns on the question of trust.² A minimal state can be defined as a political authority that does no more than it is necessary to protect the liberties of its subjects—it enforces rules against violence, theft, and non-observance of contracts—without engaging in redistributive policies (Nozick 1974: ix; Mack 2018: 2). Traditionally, the major competitor to the minimal state has been the redistributive or welfare state, which transfers wealth, income, resources or opportunities from the well-off to the less fortunate members of society. From a libertarian standpoint such redistribution violates the basic rights of individuals. This basic set of rights typically includes rights over our "lives, liberties, and estates" which Locke branded "property" (Locke 1980, Chap. 9, §123; see also Chap. 2, §6).³ In the twentieth century, the most famous defence of the minimal state remains Nozick's *Anarchy, State and Utopia* (1974). Since Nozick's theses have been extensively studied (Paul

¹ Barry (1989) has pointed out that term "libertarianism" covers a range of positions which treat the premise of individualism as central. In the present discussion libertarianism is linked to the idea of a minimal state, and specifically, a Lockean minimal state

² The concept of trust has been studied across various fields such as sociology (Luhmann 1979, 1988; Gambetta 1988a; Barbalet 2009), epistemology (Faulkner 2011), social theory (Hardin 1992, 1999, 2002), ethics (Horsburgh 1960; Baier 1986; Karen Jones 1996) and philosophy (Hollis 1998; Uslander 2002; Hawley 2014; Faulkner 2017; Faulkner and Simpson 2017). In this discussion, the emphasis is on trust in relation to government from a philosophical perspective. The relationship of trust between citizens and their government is modelled on the relationship between individual persons. Hardin (1999: 24) notes that such an approach is needed to ensure that trust in government is compatible with a general (philosophical) account of trust.

³ At §123 Locke uses "property" in a broad sense referring to the whole basket of rights to life, liberty, and estate. In the *Second Treatise* he also uses "property" in a narrow sense to refer only to the right to *estate*.

1981; Wolff 1991; Bader and Meadowcroft 2012), the proposal here is to give a fresh direction to the debate by returning to Locke's classical liberalism and to articulate the principles of a *Lockean* minimal state. The analysis offers three insights centred on the category of trust. First, it claims that a Lockean minimal state is not just a state that protects basic rights (analogously to the Nozickian state is), but that it is a state founded on the morality of trust. Second, it claims that trust serves as a moral glue connecting individuals to one another and to their government in a free society. Third, it identifies a principal antagonism between the Lockean minimal state and the so-called social credit system (construed as ideal models). In 2020 China officially introduced a social credit system which assigns trustworthiness scores to its citizens (State Council Notice 2014). The problem Chinese citizens face, apart from the fact that they enjoy a relatively limited set of rights (measured by the standards of Western liberal states), is that their actions in both the public and private domains are regarded with increasing suspicion by their own government. This distrust further deflates their rights and erodes the moral glue that integrates society.

The discussion is developed in four sections (in addition to this Introduction). Section 2 analyses the concept of trust and distinguishes two alternative conceptions of it. Section 3 elucidates the Lockean minimal state as a bastion of moral trust and basic rights. Section 4 outlines the features of the Chinese social credit system, followed by a more abstract, philosophical analysis of social credit systems in Section 5. The analysis concludes by raising concerns about a possible spread of elements of social credit systems to Western liberal states.

2. Two conceptions of trust

The Lockean minimal state, as presented here, has two core features – it protects basic rights and it is grounded in the trust of the governed. Before we turn to the defining principles of the Lockean state in the next section, it is requisite to outline the concept of trust. This interlude section will distinguish between its two core conceptions – *moral trust* (Oakeshott 1975a: 71; 1975b: 137; Uslaner 2002, 22, 24; Faulkner 2017: 119) and *strategic trust* (Hardin 1992; 1999; cf. Uslaner 2002: 17, 21-22).⁴

Trust is a relationship between agents. It is manifest in various forms of social interaction with a risky outcome whereby one agent makes oneself dependent on another (Luhmann 1979; Baier 1986: 235, 240; Barbalet 2009: 368-369; cf. Hardin 1992: 164).⁵ For example, while I am away on holiday I trust you to water my plants, or, laying ill in bed, I trust you to file an important tax declaration for me before the deadline. The *problem of trust*, as it has been termed more technically (Faulkner 2017), involves reliance on others over whom the commissioning agent has no perfect control.⁶ One part of the problem has to do

⁴ Uslaner (2002: 17) defines strategic trust as a category of *knowledge-based trust* that involves risk. The argument is that the trustor's belief in the trustworthiness of the trustee depends on knowledge or past experience concerning the trustee (or people in general, see also Hardin 1992).

⁵ Hardin (1992:164) argues that trust involves dependency on another, but not risk; rather, it involves an *expected probability* in a Bayesian context where I base my trust in you on an average of my past experiences of others. Specifically, this Bayesian context includes an optimum average of such past experiences (ibid, 155).

⁶ But even though trust involves reliance, it is not reducible to it (Baier 1986: 234). See also the discussion of strategic trust in the paragraphs that follow.

with the delegation of discretionary power. In trusting others, we make ourselves dependent on those others to care for things that matter to us, and this “caring for” consists in the exercise of discretionary power (Baier 1986: 240). The other part of the problem has to do with risk. In the above-mentioned example, the outcome is risky since it is possible that you might leave the plants to wilt, or you may not deliver the tax declaration on time. Trust would not be a problem in social life if the trustor (the one who invests trust in another) can fully control the trustee (the one who is trusted; see Dunn 1990), or if a guarantee or evidence can be produced to the effect that the trustee will perform as expected (Gambetta 1998b: 233). Indeed, trust represents ‘a gamble, a risky investment’ (Luhmann, 1979: 24).

The agent’s attitude towards the factor of risk generates the two contrasting conceptions of trust: moral trust versus strategic trust. Let us consider trust as a moral category first. In (morally) trusting you to do something, I consciously make myself vulnerable *to you* within the horizon of a future directed exchange (Oakeshott 1975a: 71; see also Barbalet 2009).⁷ Vulnerability is a risk-prone attitude. I know that you might betray my trust (by not performing the entrusted task) but I still decide to go ahead and trust you. Notice that the relationship of trust is not exhausted by the trustee’s discharging the entrusted task, which can be done more or less reliably (i.e., as expected). If you fail to perform what I have entrusted you to do (e.g., to deliver my tax declaration), you are showing a lack of respect for *me*, and not merely harming *my interest* (in having the task performed reliably), although this may in fact be true. Trust is a relationship between persons that concerns a performance but is not contained in the performance itself – trust, if you will, is the ‘spirit’ in which the performance is conducted. In this, trust resembles the notion of good will (Baier 1986).

What makes trust into a moral category? The factors are risk and the preparedness of the trustor (call her “A”) to make herself vulnerable to the trustee (call him “B”). When B reneges on a contract (or promise) based on trust, B not only fails to perform a task, but to keep his word to A (Oakeshott 1975b: 91, 138). A legitimately feels betrayed. In this case, A’s act of trusting B does not make B into a means for transmitting A’s interests, but elevates B to a moral agent with a status, a person, who is entrusted to act on A’s behalf. A feels betrayed and not merely slighted or disappointed when B fails to perform his part of the contract since this failure shows disrespect for both A’s and B’s personhood. Viewed as a moral category, trust fleshes out the idea of persons as rational agents who have value as ends in themselves and not as throughputs for wished-for outcomes (Kant 1996: IV: 429). In sum, moral trust is a relationship between persons, and not a tool for generating reliable outcomes.⁸

There is a parallel modality of trust: strategic trust. Strategic trust downgrades the basic element of personhood inherent in a trust relationship and upgrades the elements of reliance and expected outcomes. Russell Hardin discusses strategic trust under the rubric of “trust as encapsulated interest” (1992: 152-153; 1999; 2002: 3-10). On this construal, “To say that I trust you means I have reason to expect you to act, for your own reasons, *as my agent* with respect to the relevant matter. Your interest encapsulates my interest” (Hardin 1999: 26; see

⁷ Oakeshott considers trust as an element in the undertaking of a moral obligation via promising. On moral promises see David Jones (1996).

⁸ I follow A.I. Melden who conceives rights as relations between persons, in the technical sense of person, as a unit of basic moral concern in moral theory. Violating someone’s rights, Melden argues (1977: 46), is equivalent to committing a wrong against a person. My point is that betraying someone’s trust too is a mode of disrespect for persons.

also Hardin 1992: 154; 2002: 4). Strategic trust supposes a three-part relation, “A trusts B to do x (or with respect to x)” (Hardin 1992: 154; 1999: 26; 2002: 9). Here the emphasis is not on the relationship between A and B, trustor and trustee, as persons (as was the case with moral trust), but on the expectation of an outcome x (Uslaner 2002: 22). A *expects* that B will procure an outcome (which A has an interest in obtaining). In this case, B becomes a means for channelling A’s preferences and interests. This can be seen from the fact that it would not matter to A if someone else other than B, say, C would secure the same outcome x . For A does not care about *B’s act of performing x* , but only about that a state of affairs x obtains.

This interest-based view of trust (strategic trust) is a consequentialist standpoint which prioritises the consequences of actions over the value of persons. The failure to perform an entrusted task counts as a frustration of expectations and not as a moral defect. While such expectations are not purely inductive – such as the expectation that the Sun will rise tomorrow (Hardin 1999: 26; 2002: 7) – they are not moral expectations either. On the conception of strategic trust, if the trustee, B, fails to act in a way that meets A’s (the trustor’s) expectations, the typical complaint is not that B disrespects A (or has broken one’s word to A) but, rather, that B cannot be relied upon to deliver outcome x . Hence, B is no longer considered *trustworthy* by A.⁹ The penalty for untrustworthy behaviour, Hardin explains (1992: 153; 2002: 6), is severance of the continuation of any future exchange with the trustor. Strategic trust, therefore, represents an idiom of self-interest both on the trustor’s side (who relies on the trustee for an intended outcome for self-interested reasons) and on the trustee’s side (who is assumed to have an interest in maintaining connection to the trustor).

Besides self-interest, the second key ingredient in strategic trust is the concept of reliance. This concept belongs to consequentialist theories of action such as rational choice where the key problem of delegating agency is ultimately one of reliance. In its strategic modality, trust itself is presented as a rational belief concerning the putative reliability of the trustee (Hardin 1992: 153, 157; 1999: 25). A party who has proven reliable over past transactions can be considered trustworthy by the rational trustor. In effect, trust as trustworthiness is a type of reliance whose aim is to minimise the risk of transactions. But the willingness to undertake risk, as we have seen, is a defining element in the moral conception of trust. The proposal in what follows is that we should read the Lockean minimal state as a state animated by the moral conception of trust. Conversely, social credit systems aim to generate trustworthiness.

3. Trust and the Lockean minimal state

This part of the discussion will adduce analytical evidence for the claim that, properly understood, the Lockean minimal state is constituted by two principles: rights (liberties) and moral trust (henceforth simply ‘trust’). Even though the emphasis throughout is on trust, it is requisite to consider both components: trust *and* rights.

⁹ The reader should excuse the infelicity of the term ‘trustworthiness’. In the specialised literature, it has been connected to arguments from reliance and self-interest (see, especially, Hardin 2002) rather than to the (Kantian) conception of moral worth. In this sense, “to be trustworthy” refers to the idea of *being reliable* rather than to the idea of *being worthy* of trust in a moral sense.

The notion of trust is the reason for choosing to develop the present argument via a Lockean minimal state rather than a Nozickian minimal state (from which this notion is absent). There is no need to belabour Nozick's theses except to briefly mention his idea of the "night-watchman state". Nozick's starting point is the dictum, inspired by Kant and Locke, that "individuals have rights, and there are things no person or group may do to them (without violating their rights)" (Nozick 1974: ix). Any policy of redistribution, be it in the name of socialist, utilitarian, or egalitarian projects such as John Rawls's liberal egalitarianism propounded in *A Theory of Justice* (1971), is unjustified because it violates rights. Hence, the only state that can be justified is the night-watchman state, a minimal state that protects its citizens from theft, fraud and physical violence without attempting to promote their welfare beyond this minimal commitment (Nozick 1974: ix). For Nozick, rights are the entitlements of individuals and individuals are separate moral units – they cannot be justifiably fused into a social whole in the name of the common good.¹⁰ This vigorous individualism has made Nozick a prominent figure in libertarianism, as a doctrine asserting the value of individualism in the broadest possible sense.¹¹

Nozick's minimal state is established to protect a set of rights, loosely understood as Lockean rights to life, liberty and estate (or "property" in the broad sense). But while Locke took the state to be a guardian of this same set of rights, he further grounded its authority in the trust of the governed. "Political power," Locke wrote, "is that power, which every man having in the state of nature, has given up into the hands of society, and therein to the governors, whom the society hath set over itself, with this express or tacit trust, that it shall be employed for their good and the preservation of their property" (Locke 1980: Chap. 15, §171). The Lockean minimal state therefore should be seen as constituted by both rights and trust. This hypothesis has a procedural as well as a normative aspect.

Procedurally, the constitution of Lockean government proceeds in two steps. First individuals form "political society" unanimously, via consent, and then political society entrusts the office of rulership to government via majority rule (Locke 1980: Chap. 8, §§ 96,99). Scholars have been preoccupied with the problem of consent in Locke's political theory (Simmons 1993, Chap. 3; Dunn 1967; more generally, Beran 2020). But by Locke's own lights, the constitution of government requires trust beyond consent (Dunn 1984). Locke holds (1980: Chap. 13, §149; Chap. 19, §221) that when government abuses the trust invested into it by the people, the people has a right to depose it and institute a new government. What kind of trust is that? And in what ways (if any) does trust differ from consent, as a normative concept?

As John Simmons argues, Locke uses the term consent to cover promises, contracts, and acts authorising the actions of others (Simmons 1976: 275). It is helpful to distinguish promises (where one undertakes an obligation oneself) from acts of authorisation (where one gives rights to others). Simmons notes that a core feature of consent in the strict sense is that it

¹⁰ Trade-offs can be made between rights and other values (interests, well-being) relative to the domain surrounding the individual person, but not *across* domains assigned to distinct persons (Nozick 1974: 33). Persons ought to remain separate moral units.

¹¹ See note 1 above.

is always given to the actions of others.¹² “Thus, I may consent to my daughter's marriage, to be governed by the decisions of the majority, to my friend's handling my financial affairs. Promises, on the other hand, cannot be made except in special circumstances, concerning how another person will act” (Simmons 1976: 275). Promises put the agent (the promisor) in the spotlight.

Acts of moral trust resemble promises in that they equally presuppose a first-person standpoint. When A promises B to do *x*, A is thereby autonomously undertaking an obligation towards B to do *x*. Recall that when A decides to trust B, A makes oneself vulnerable to B (who in this case is the promisor who may betray A).¹³ A, that is, willingly accepts B's promise on trust and without seeking assurance. If the argument so far bears scrutiny, this willingly assumed vulnerability (moral trust) by the trustor is expressive of moral autonomy.

How does Locke understand the role of trust in the constitution of government? Locke makes trust the fundament of government, and maintains that failure to uphold the trust of the governed justifies the dissolution of government (Locke 1980: Chap. 15, §171; Chap. 19, §221). This pronouncement can be interpreted in two ways. The first interpretation says that the citizens entrust their government with a task on instrumental grounds. Here, the motive of both actors is reciprocal self-interest – obedience is traded for law and order. Hardin's argument for trust as encapsulated interest serves as a model for this interpretation. It suggests that Locke was employing trust as a strategic concept. On this reading, political society (the trustor) would have an interest in law and order, and correlatively, government (the trustee) would have an interest in serving the interest of political society (in law and order). Presumably, government would have a further interest in not being deposed by the people. While it is not incoherent to read Locke's argument in this way, it leads to an interpretive difficulty.

It is well-known that the most efficient way for any government to impose law and order is by suppressing the rights of the subjects. Against this, Locke's intent in the *Second Treatise* was to defend the liberty of the subjects by limiting the scope of political authority. These liberties comprise the basic rights to life, liberty and estate, which exist independently from political authority (as moral rights) and may be termed “first order rights” (Mack 2013: 75-78). Since they can be violated in the state of nature by transgressors, Locke allows for a parallel set of “second order rights” of enforcement, punishment, and restitution (Mack 2013: 75-78). Each individual is permitted to use these second-order rights to prevent or redress violations of *anyone's* basic rights (Locke 1980: Chap. 2, §7). The normative regime which regulates the distribution of rights of both kinds in the Lockean state of nature is the law of nature. The law of nature holds that each agent is free just in case others do not interfere with the legitimate exercise of her basic rights (Locke 1980: Chap.2, §6). Freedom, on this construal, implies that A's basic rights impose corresponding obligations on other agents to respect A's rights.

For Locke, government has no special mandate on its own other than protecting basic rights. This stems from Locke's premise that rights have the same liberty preserving function

¹² In noting that consent involves an authorisation of other persons' actions, Simmons (1976: 275) draws on an earlier argument by Plamenatz (1968:3).

¹³ In this sense, the logical structure of promises and acts of moral trust (where the trustor accepts a promise) is reversed. Their *moral* structure (in terms of autonomy) however is similar.

in both the private and the public domains. No principled distinction is drawn between public morality and private morality on this view (Mack 2018: 2). A (private) relationship between two individuals (as moral agents) is not *per se* distinct from a (public) relationship between a group of freely associated individuals (political society) and its government. Reason of state, on this view, cannot justify *government* in curtailing citizens' basic rights any more than one person can be justified in curtailing the rights of another. This is the ethos of the Lockean minimal state. But if so, the first interpretation above where law and order are traded for obedience – on instrumental or welfare promoting grounds – cannot be supported in a way consistent with the rights-protecting premises of a Lockean minimal state and indeed with any liberal rights-based framework for justifying political authority.

How is Locke's rights-based account relevant to the question of trust and government? If, as Locke thinks, the business of government is to protect basic rights, then the mandate of government cannot rest on purely strategic trust. The would-be citizens cannot properly be said to entrust government on the expectation that it will prove itself trustworthy in the task of securing their rights, if this leaves government the option to rewrite such rights, for example, by treating rights as interests or interchangeable units of well-being. To enact this option would be to collapse rights, as rigid freedom-protecting perimeters surrounding the person¹⁴, into interests, as shifting indexes of wants and desires. Once trust is rendered a function of interest (as suggested by the doctrine of strategic trust) it cannot be coherently grafted onto Locke's view of basic *rights* as protected domains of personal freedom.

It is thus pertinent to attempt a different reading of Locke's statement that the grant of political authority can be revoked whenever popular trust has been abused. This statement is both about trust in government, and about the nature and scope of political authority. This brings us to the second interpretation of Locke's argument for trust and government in the *Second Treatise*.

Locke can be interpreted as saying that government is revocable if it has violated its *comprehensive mandate* entrusted to it as political authority.¹⁵ Government, it should be noted, is not an agency for discharging particular tasks or seeking particular outcomes—its mandate (to make law) is itself constituted by law, as a broad, or comprehensive, domain for regulating human conduct. According to GEM Anscombe (1990), government constitutes a domain of juridically coercive political authority that is *pro tanto* justified. That is, government is distinct from purely voluntary associations in that its claims to authority are coercively underwritten, and from organised mafias in that such claims are normatively grounded – they presuppose a *right* to rule tied to a corresponding *obligation* to obey on part of the citizens (Anscombe 1990: 147-151). For Locke, legal coercion does not sacrifice individual liberty since the mandate of government is limited by the rule of law. The principles of the rule of law are summed up in the injunction that the “*legislative*, or supreme authority, cannot assume to it self [sic] a power to rule by extemporary arbitrary decrees, but is *bound to dispense justice*, and decide the rights of the subject by *promulgated standing*

¹⁴ Nozick has explicated the idea in terms of rights as “side constraints” which block interferences into the domain of individual freedom (Nozick 1974: 29, 33).

¹⁵ By “authority” in this discussion I mean *practical authority*, roughly legislative authority over persons, to be contrasted with theoretical authority based on expert knowledge (Hart 1990: 107).

laws and known authorized judges” (Locke 1980: Chap. 11, §136: original emphasis).¹⁶ The rule of law is appealed to in order to protect the basic *rights* of the citizens via a stable juridical framework whose maintenance does not itself slide into rights violations.

If this second interpretation is plausible, Locke’s social contract portrays legitimate government as one entrusted with a *broad* form of authority. This type of authority, in lieu of its juridical mandate that has no predetermined expiration date, does not allow strategic calculations of trustworthiness based on knowledge or past experience. The decision to trust government cannot be compared to a decision to hire a roof maker that appears trustworthy.¹⁷ Under conditions of uncertainty where the trustee is to be granted an open-ended mandate, the concept of strategic trust is non-operative. Thus, the citizens in Locke’s account should be viewed as trustors who take the risk of making themselves vulnerable to their rulers in the broad domain of *political authority*. But, and here the idea of mutual trust is key, the government that is trusted should, in its turn, be prepared to trust its citizens in the equally broad domain of *personal liberty*.

In summary, then, the Lockean minimal state rests on rights and on mutual trust. Within this type of state, the governed are prepared to take the risk of trusting their government, and government equally is risking the option of less than total state control. It refrains from seeking to oversee all spheres of citizen conduct, thereby offering juridical protection of the sphere of individual liberty for each citizen. The trust involved in each case is moral trust and can be described under the schema “A trusts B”, or “B trusts A” (Faulkner 2017: 119; see also Uslaner 2002: 22, 24). So construed, moral trust is a two-way relationship that omits any reference to a third element such as a particular outcome *x*. (Recall that on the strategic model of trust, trust is necessarily a three-way relationship under the schema “A trusts B *to do x*”). Trust as a moral category, not the practice of governmental control and verification, is the root of the Lockean minimal state as a custodian of liberty.

4. The Chinese social credit system: a sketch

At this juncture, the focus of the argument shifts from the Lockean minimal state towards its novel contender or “foe” – a social credit system. Both the Lockean minimal state and its social credit counterpart are analysed as paradigms or models. This, however, is not a purely theoretical enterprise since an actual social credit system – *shehui xinyong tixi* – was officially introduced in China in 2020 after a decade of planning (State Council Notice 2014). In this section, the empirical features of the Chinese social credit experiment will be outlined before proceeding to develop a more abstract, philosophical account of social credit systems in

¹⁶ The principles of rule of law, more specifically, include “established, settled, known law”, “known and indifferent judges” (Locke 1980: Chap. 2, §126), separation of powers, and the injunction that the rulers are subject to the laws they make (Locke 1980: Chap. 12, §143).

¹⁷ Uslaner (2002: 24) distinguishes between the disposition to have faith in people in general, and the decision to trust specific persons (roof contractors, car dealers, etc). In the latter case, it makes sense to search for information in order to assess someone’s trustworthiness, but not in the former case. This faith in people is faith in a “generalized other” (ibid.) It represents an *attitude* of trusting (which is not a *belief* based on information or assessment of trustworthiness). Faulkner (2017:119) draws on Uslaner’s argument.

Section 5. Perforce the ensuing analysis is speculative considering the novelty of the problem discussed.

The Chinese social credit system is unprecedented in that it singles out the trustworthiness of individuals within its borders as the target of social and political control. According to its official governmental definition, it is “a socioeconomic system that attempts to evaluate the trustworthiness of individuals, companies and social organizations, and state agencies in China with mechanisms of rewards and punishments” (State Council, 2014 in Jiang 2020, 93). At its heart, the SCS is a ranking system. It utilises quantitative indicators similar to the ones used by financial rating companies in the West such as FICO (in the US) and SCHUFA (in Germany) to assess the creditworthiness of individuals seeking bank loans. There are two major differences though. The first is that the criterion of *social credit* – *chengxin* or *xinyong* – translated as “trustworthiness” is distinct from financial creditworthiness (Zhang 2020, 566, 570). It has ethical connotations, as one of the main purposes of the system is to produce “good citizens” (Reijers 2019; Orgad and Reijers 2021; Liang and Chen, 2022).

The second difference is that the social credit system is not designed to evaluate the conduct of individuals as private persons or free market participants. Rather, its aim is to evaluate their conduct as *citizens* (Orgad 2019; Orgad and Reijers 2021). The state assigns trustworthiness scores to its citizens – often by relying on private ranking agencies as mediators— and uses these scores as inputs within an administrative system of rewards and punishments (Creemers 2018).¹⁸ Individuals who are considered untrustworthy are debarred from boarding high-speed trains or airplanes preventing them from leaving the county, or are denied access to governmental jobs, real estate purchases, and higher education opportunities for their children (Ziang 2020, 94; Liang and Chen 2022, 126).¹⁹

At present, the Chinese social credit system operates with three nodes of information aggregation. These includes China-wide blacklists (and redlists) which record good (and respectively, bad) citizens behaviour, SCS pilots at municipal and city level (currently over 70), and the financial data overseen by the People’s Bank of China (Zhang 2020, 570; Ziang 2020, 93). Local authorities monitor citizen behaviour in diverse areas (internet use, travel, payment of utility bills, compliance with court orders, default on bank loans and so on) (Lee 2019, 956). Whenever the system receives negative feedback from more than three separate authorities about an untrustworthy individual, it aggregates and uploads their data to a website *CreditChina* established in 2015 (Lee 2019, 956).

It is easy to see that the Chinese social credit system claims a sweeping discretion over domains of personal choice and conduct that, in lieu of principles of individual autonomy and rights, are normally excluded from political oversight within liberal states. What is the significance of this for liberal politics and liberal political theory? Commentators worry that the social credit experiment conjures up a surveillance state (Botsman 2017), that it exacerbates authoritarianism (Creemers 2018; Ford 2019), or that it promotes the

¹⁸ Private financial organizations such as Sesame Credit (run by a subsidiary of Alibaba Group) used to provide social credit ratings (combining financial *and* socio-ethical data) of individuals before the People’s Republic of China suspended their license in 2019.

¹⁹ Ziang (2020) mentions all three areas of restricted control without the factor access to higher education services for one’s children.

“datafication” of society (Lee 2019). The claim here is that social credit systems pose a deeper challenge to the Western liberal state. They challenge the idea of government as limited by independent moral constraints such as basic rights, the distinction between “man and citizen”, and the separation between the private and the public sphere, distinctions which have been formative for the liberal state since the dawn of European modernity.

5. Social credit systems

This section generalises the contingent features of the Chinese social credit system into a more abstract model of social credit systems. It puts forward two hypotheses holding that: (1) the introduction of a social credit system transforms the state into a market for services, rendering the role of law perfunctory; and that (2) trust becomes a commodity which circulates in the social and political market place. In essence, by turning trust into a commodity, social credit systems threaten to erode the notion of trust as a moral glue that underpins the relationship between individuals as autonomous agents, and between citizens and their government (central to the model of a Lockean minimal state).

5.1. *The marketisation of the state*

The first hypothesis suggests that a social credit system discloses a “switch” towards the marketisation of the state. This new form of state resembles a firm insofar as it is *exclusively* concerned with the efficiency of service provision. Efficiency is defined in terms of cost (Buchanan 1999, Chap. 1), and the purpose is to provide social and political control at the lowest possible cost. A marketised state becomes inverted into an administrative branch positioned *within* a broader system of production and distribution of social credit. As such it differs principally from the juridically constituted authority of the liberal state, of which the Lockean minimal state is an exemplar. The marketised state is not concerned to be a source of known, settled law, to be an impartial arbitrator, or to treat its citizens as equals under the law—on the contrary, the social credit system generates and manages different *hierarchies* of citizens—“good” vs. “bad”, “trustworthy vs. untrustworthy”. While judicial offices and courts continue to exist formally after the introduction of the system, legal judgements are not primarily enforced by traditional juridical means such as police, nor adjudicated by higher courts of appeal. The practice of legal enforcement is *de facto* replaced by a rewards/punishment scoring mechanism. The rationale, it seems, is that it is less costly to rely on a scoring mechanism than to invest in burdensome and time-consuming court procedures or in the education of lawyers, judges, and police officers. A failure by an individual to comply with a legal judgment no longer counts as a legal offence, but results in the lowering of one’s social credit score. Legal judgments are diluted into entries within the administrative apparatus of the social credit organisation whose ultimate goal is efficiency.

The impetus behind social credit systems is buying social and political control on the cheap. Setting up a system of this sort requires a sizable initial investment by the political authorities, especially if a digital infrastructure for monitoring the subjects in real time is introduced, but, once set up, the system perpetuates itself with little official administrative and juridical oversight. Efficiency in governance is achieved by employing the citizens

themselves—and not the state or its apparatuses—to monitor each other’s behaviour and to report one’s fellow citizens to the social credit system for being bad or untrustworthy. This is why an individual’s trustworthiness score cannot be a financial information (“creditworthiness”) confidentially kept inside bank books; it has to be a social credit score (“trustworthiness”) that is publicly visible in one’s immediate community—nowadays, on the internet.

The authority of the Western liberal state from Hobbes onwards has rested on the pillars of law and justified coercion. The scope of this coercive juridical authority has in principle been limited. Even Hobbes, the architect of strong state authority, was adamant that law ought to have limits: that it should not attempt to regulate the inner moral world of individuals (their motives, thoughts, hopes and fears) but only their external interactions with one another in the *foro externo* (Hobbes 1968, Chap. 15, 215 [79]). This is in sharp contrast to the idealised image of a social credit system outlined here. Inside this system, the preferred method for regulating the conduct of citizens is not law, but co-optation. Indeed, a social credit system represents a market for services coupled to a mechanism of social co-optation.

Co-optation seeks to colonise the inner world of the subject. The aim is to reform the individual “from the inside” so that she no longer *wants* to violate the established rules of the game (assuming that such unruly wishes may have been entertained *ab initio*). Generalised social pressure, manifest into publicly visible social credit scores, is the avenue for achieving this goal.

Notice that in a Hobbesian state the private/inner realm is separated from the public/outer realm, and that inside the private realm citizens enjoy certain rights that cannot be infringed by the state without proper justification. The import of individual rights is even more prominent in a Lockean minimal state, which comprises a juridical “penumbra” surrounding a “core” of basic moral rights. As an institutional structure of law and justified legal coercion, the Western liberal state (be it in its Lockean or Hobbesian denomination) is supposed to prevent one individual from violently encroaching the rights of another. Such violence is crime and what counts as crime is stipulated by antecedent law - *Nullum crimen sine lege*.

The logic of social credit systems is fundamentally different. In lieu of the prospective, highly contingent nature of the social credit enterprise, whose digital infrastructure can be updated in real time, the definition of “crime” becomes Bayesian rather than juridical.²⁰ Such definitions are no longer determined by settled, publicly promulgated laws; nor does “crime” refer to the violation of rights. Instead it refers to the latest (freshly updated) ordinance of the State or the Party, so aptly presented by George Orwell in *Animal Farm* (1989). The individual subject of the social credit system is co-opted to comply with its ever-shifting decrees and regulations that are extra judicial at base. The message to the citizens, then, simply is: “Conform to state power whatever the scope and nature of this power may be”. The

²⁰ The purpose of a social credit system is not to wait for rights violations, or crime, to occur in order to then engage the courts and the penal system. Rather, the aim is to deter crime *before* it has occurred. As in the film *Minority Report*, the precrime police unit relies on predictions which, in fact, often have little to do with any actually committed acts of crime. The big data, utilised via digital, algorithmic infrastructures, in some of the largest city pilots running the social credit experiment in China augment this tendency of projecting targets in the future, or prospectively, rather than seeking moral and legal responsibility, on the basis of antecedent laws.

modus operandi of the social credit system resembles that of the culture industry under advanced capitalism, powerfully criticised by Theodor Adorno:

In contrast to the Kantian, the categorical imperative of the culture industry no longer has anything in common with freedom. It proclaims: you shall conform, without instruction as to what; conform to that which exists anyway, and to that which everyone thinks anyway as a reflex of its power and omnipresence. The power of the culture industry's ideology is such that conformity has replaced consciousness (Adorno 1991: 104).

5.2. The commodification of trust

Let us grant for a moment the plausibility of the first hypothesis, stating that a social credit system constitutes a market for services. The second hypothesis suggests that the principal commodity operating in this market is trustworthiness. Recall that on the second interpretation of the role of trust in the Lockean minimal state, trust was seen as a moral category. By hypothesising that a social credit system relies on trustworthiness (the core element of strategic trust) and not on moral trust, this section aims to adduce additional reasons as to why this type of system represents an obverse model to the Lockean minimal state.

It would however be a mistake to equate the notion of trustworthiness presupposed by the social credit system, as presented here, with Hardin's concept of trust as trustworthiness. Hardin argued in a consequentialist fashion (see Section 2) that the reason for B's (the trustee) compliance with A's (the trustor's) investment of trust is not the recognition that A has made oneself vulnerable for the sake of B (as it would be the case if trust were a moral category). To the contrary, B was assumed to have an interest in having A's interest at heart in performing an entrusted task, given an expectation of negative consequences accruing to B in case B neglected to act in A's interest. This conception of strategic trust presupposes a minimum moral content reflected in the concept of interest. At the very least, it presupposes that B cannot have a declared interest in harming A. In an analogous way, a promise (as a matter of exhibiting good will towards the promisee) can be distinguished from a threat (a promise to do something the promisee does not want done; Searle 1969: 58-59). Hardin's conception of trustworthiness exhibits a core difference from the social credit model of trustworthiness, a model which is removed from the ethical world (even in its consequentialist mode) and located inside the market place for services. Trustworthiness becomes a commodity whose circulation depends on information over targets of expected behaviour stochastically generated by the social credit mechanism. Neither party involved in a putative transaction that implies trust (read, trustworthiness) has any volitional control over such commodity.

But, then, who controls the distribution of trustworthiness in social credit systems? Typically, the government, as a third-party external to the transactions between designated individuals, would posit target expectations of what counts as trustworthy behaviour. For example, in the Chinese social credit system people who visit their parents are considered more trustworthy than those who do not (Bertelsmann Stiftung Report 2019). By whom? Certainly not by their parents, but by the system, and ultimately, by the Chinese government.

The upshot is that within a social credit framework trust is no longer a relationship between self-interested individuals (Hardin 1992, 1999), or between moral persons

(Oakeshott 1975b: 137ff). It has been transformed into a commodity, a trustworthiness score, whose price falls or raises as determined by the contingently changing targets of the administrative state. This state itself is reduced to a branch of the market-like social credit system. Trustworthiness is now the equivalent of money in the political market place, a market place for services which displaces the juridically underwritten liberal state.

6. Conclusion

It is time to take stock of the preceding analysis of trust in the context of a Lockean minimal state. Locke's insight was that trust constitutes the moral bedrock of government, and the argument presented here extended it by showing that citizens not only entrust government to do a particular, narrowly determined task but invest their moral trust into it as part and parcel of a comprehensive mandate of political authority. On this perspective, moral trust is a normative glue: it relates trustor and trustee as persons who have respect for each other as ends in themselves. Trust as a moral category is principally distinct from strategic trust (I trust that you will perform as expected because, presumably, you have an interest in taking into account my interest in fulfilling the expectation).

The examination of the Lockean minimal state revealed that its ability to protect rights depends on the fact that it is a state founded upon moral trust. Moral trust is a "glue" in both the private and public domains—it consolidates the private transactions between individuals as well as the relationship between political society and government. Further, the justification of the Lockean minimal state was shown to depend on its recognition of normative limits, rights, which block public officials from illegitimately using their claim to authority to infringe the sphere of individual liberty. As articulated here, the Lockean minimal state constitutes an institutional edifice grounded in moral trust that contains a core of basic rights, enwrapped in a legal penumbra. Both moral trust and basic rights are absent from social credit systems. Such systems are not relying, as the Lockean minimal state is, on justified legal coercion that is meant to protect liberty; their essence comprises market principles meant to guarantee efficiency.

Within social credit systems, the legal shell of the state is contingently shrinking and expanding in response to information provided by the social credit market. The "rules of the game" are not expressed in known and settled public law; these rules represent directives stochastically updated to meet the efficiency needs of the social system as a whole. On this social credit model, citizens cannot be said to have rights in the proper sense of a protected perimeter of liberties. Individuals and their rights (which are rights in name only) have been turned into resources to be efficiently managed.

The trustworthiness scores of citizens represent the currency of the social credit system. Citizens (if the term is at all appropriate) are induced to comply with the system via a mechanism of co-optation. Instead of guaranteeing their rights by a juridical framework which has moral limits, the mechanism regulates their conduct through market incentives (rewards and punishments). As a result, the moral perimeter guarding each citizen from encroachment by the rest and by the government has been erased. Government is now in a position to control not merely what the citizen does, but *who* the citizen is.

In conclusion, it can be said that social credit systems mark a fundamental transformation away from the Western liberal state, and its Lockean manifestation, and towards a market of services where even trust has been commodified. It is nonetheless important to recognise that, in time, elements of social credit systems may begin to infiltrate Western liberal states. One alarming sign is the recent weakening of the rule of law, in the US and certain EU countries such as Poland and Hungary, producing an asymmetric relationship between state and citizens that entails the diminution of individual rights. Another alarming sign is the widespread use of ratings designed to assess the trustworthiness of agents in online platforms. For the time being, these measure the performance of individuals and firms as private parties and commercial agents but what if liberal governments decided to adopt similar rating practices to evaluate individuals as citizens (i.e., participants in the public sphere)? In light of the preceding discussion, the severe repercussions such developments may have for the domain of individual liberty should give us a pause.

References

Adorno, Theodor W. (1991). The culture industry reconsidered. In: Theodor Adorno, *The Culture Industry*, ed. J. M. Bernstein. London and New York: Routledge, 98-106.

Anscombe, G.E.M. (1990). On the source of the authority of the state. In: Joseph Raz (ed.), *Authority*. New York: New York University Press, 142–173.

Bader, Ralf M., and John Meadowcroft (eds.) (2012). *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*. Cambridge: Cambridge University Press.

Baier, Anette (1986). Trust and antitrust. *Ethics*, 96, 231–60.

Barbalet, Jack (2009). A characterization of trust, and its consequences. *Theory and Society*, 38(4) [Special Issue: Emotion and Rationality in Economic Life], 367-382.

Barry, Norman (1989). Libertarianism: some conceptual problems. *Royal Institute of Philosophy Supplements*, 26, 109-127.

Beran, Harry (2020). *The Consent Theory of Political Obligation*. London and New York: Routledge. (Originally published in 1987).

Bertelsmann Stiftung Report (2019) *China's Social Credit System*.

https://www.bertelsmannstiftung.de/fileadmin/files/aam/Asia-Book_A_03_China_Social_Credit_System.pdf

(Accessed 10 Nov 2019)

- Botsman, Rachel (2017). Big data meets big brother as China moves to rate its Citizens. *Wired UK*. (October 21, 2017) <https://is.gd/VdmJYa>. (Accessed 10 June 2022).
- Buchanan, James (1999). *Cost and Choice: An Inquiry into Economic Theory*. Indianapolis: Liberty Fund. (Originally published in 1969).
- Creemers, R. (2018). China's social credit system: An evolving practice of control. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792 (Accessed 10 June 2022)
- Dunn, John (1967). Consent in the political theory of John Locke. *The Historical Journal*, 10 (2), 153-182.
- Dunn, John (1984). The concept of “trust” in the politics of John Locke. In: Richard Rorty, J.B. Schneewind, and Quentin Skinner (eds.), *Philosophy in History: Essays on the Historiography of Philosophy*. Cambridge: Cambridge University Press, 279–301.
- Dunn, John (1990). Trust and political agency. In: John Dunn, *Interpreting Political Responsibility: Essays 1981-1989*. Princeton: Princeton University Press, 26-44.
- Faulkner, Paul (2011). *Knowledge on Trust*, Oxford: Oxford University Press.
- Faulkner, Paul (2017). The problem of trust. In: Paul Faulkner and Thomas Simpson (eds.), *The Philosophy of Trust*. Oxford: Oxford University Press, 109-128.
- Faulkner, Paul and Thomas Simpson (eds.) (2017). *The Philosophy of Trust*. Oxford: Oxford University Press.
- Ford, Cristie (2019). Seeing like an authoritarian state. In: Orgad, Liav and Wessel Reijers (eds.) *A Dystopian Future? The Rise of Social Credit Systems*. EUI Working Paper RSCAS 2019/94. Florence: European University Institute, 15-17.
- Gambetta, Diego (ed.) (1988a). *Trust: Making and Breaking Cooperative Relations*. Oxford: Blackwell.
- Gambetta, Diego (1988b). Can we trust trust? In: Diego Gambetta (ed.), *Trust: Making and Breaking Cooperative Relations*. Oxford: Blackwell, 213-237.
- Hardin, Russell (1992). The street-level epistemology of trust. *Analyse & Kritik* 14, 152-176.
- Hardin, Russell (1999). Do we want trust in government? In: Mark E. Warren (ed.), *Democracy and Trust*. Cambridge: Cambridge University Press, 22-40.
- Hardin, Russell (2002). *Trust and Trustworthiness*. New York: Russell Sage Foundation.

- Hart, H.L.A. (1990). Commands and authoritative reasons. In: Joseph Raz (ed.), *Authority*. New York: New York University Press, 92–114.
- Hawley, Katherine (2014). Trust, distrust and commitment. *Noûs*, 48 (1), 1–20.
- Hobbes, Thomas (1968). *Leviathan*, ed. C.B. Macpherson. London: Penguin (Originally published in 1651).
- Hollis, Martin (1998). *Trust within Reason*. Cambridge: Cambridge University Press.
- Horsburgh, H. J. N. (1960). The ethics of trust. *The Philosophical Quarterly*, 10 (41), 343-354.
- Jiang, Min (2020). A brief prehistory of China’s social credit system. *Communication and the Public*, 5(3-4), 93–98.
- Jones, David E. (1966). Making and keeping promises. *Ethics*, 76 (4), 287–296.
- Jones, Karen (1996). Trust as an affective attitude. *Ethics*, 107 (1), 4–25.
- Kant, Immanuel (1997). *Groundwork of the Metaphysics of Morals*, ed. Mary Gregor. Cambridge: Cambridge University Press. (Originally published in 1785).
- Lee, Claire Seungeun (2019). Datafication, dataveillance, and the social credit system as China’s new normal. *Online Information Review*, 43 (6), 952-970.
- Liang, Fan, and Yuchen Chen (2022). The making of “good” citizens: China's social credit systems and infrastructures of social quantification. *Policy & Internet*, 14, 114 – 135.
- Locke, John (1980). *Second Treatise of Government*, ed. C.B. Macpherson. Indianapolis: Hackett. (Originally published in 1690).
- Luhmann, Niklas (1979). *Trust and Power*. New York: Wiley.
- Luhmann, Niklas (1988). Familiarity, confidence, trust: problems and alternatives. In: Diego Gambetta (ed.), *Trust: Making and Breaking Cooperative Relations*. Oxford: Blackwell, 94-107
- Mack, Eric (2012). Locke. In: Gerald F. Gaus and Fred D’Agostino (eds.), *The Routledge Companion to Social and Political Philosophy*. London and New York: Routledge, 71-81.
- Mack, Eric (2013). *Locke*. New York and London: Bloomsbury.

- Mack, Eric (2018). *Libertarianism*. Cambridge: Polity.
- Melden, A.I. (1977). *Rights and Persons*. Berkley and Los Angeles: University of California Press.
- Nozick, Robert (1974). *Anarchy, State, and Utopia*. New York: Basic Books.
- Oakeshott, Michael (1975a). Introduction to Leviathan. In: Michael Oakeshott, *Hobbes on Civil Association*. Indianapolis: Liberty Fund, 1–79.
- Oakeshott, Michael (1975b). The moral life in the writings of Thomas Hobbes. In: Michael Oakeshott, *Hobbes on Civil Association*. Indianapolis: Liberty Fund, 80–140.
- Orgad, Liav (2019). Social credit systems: China and the West. In: Orgad, Liav and Wessel Reijers (eds.) *A Dystopian Future? The Rise of Social Credit Systems*. EUI Working Paper RSCAS 2019/94. Florence: European University Institute, 42-44.
- Orgad, Liav and Wessel Reijers (eds.) (2019). *A dystopian future? The rise of social credit systems*. EUI Working Paper RSCAS 2019/94. Florence: European University Institute, 1-44.
- Orgad, Liav and Wessel Reijers (2021). How to make the perfect citizen? Lessons from China’s social credit system. *Vanderbilt Journal of Transnational Law*, 54 (5), 1087-1121.
- Orwell, George (1989). *Animal Farm*. London: Penguin. (Originally published in 1945).
- Paul, Jeffrey (1981) *Reading Nozick*. Totowa, NJ: Rowman & Littlefield.
- Plamenatz, John P. (1968). *Consent, Freedom, and Political Obligation*, 2nd ed. London: Oxford University Press.
- Rawls, John (1971). *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Reijers, Wessel (2019). How to make “the perfect” citizen? In: Orgad, Liav and Wessel Reijers (eds.), *A Dystopian Future? The Rise of Social Credit Systems*. EUI Working Paper RSCAS 2019/94. Florence: European University Institute, 1-3.
- Searle, John R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. London: Cambridge University Press.
- Simmons, A John (1976). Tacit consent and political obligation. *Philosophy & Public Affairs*, 5 (3), 274-291
- Simmons, A John (1993). *On the Edge of Anarchy*. Princeton, NJ: Princeton University Press.

State Council Notice (2014). Notice concerning issuance of the planning outline for the construction of a social credit system (2014-2020), *State Council of the People's Republic of China* (June 14, 2014), [<https://digichina.stanford.edu/work/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/>] (accessed 10 June 2022)

Uslaner, Eric M. (2002). *The Moral Foundations of Trust*. Cambridge: Cambridge University Press.

Wolff, Jonathan (1991). *Robert Nozick: Property, Justice and the Minimal State*. Cambridge: Polity.

Zhang, Chenchen (2020). Governing (through) trustworthiness: technologies of power and subjectification in China's social credit system. *Critical Asian Studies*, 52 (4), 565-588.