LEDGER
ledgerjournal.org

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

Governance in Blockchain Technologies & Social Contract Theories: Open Review

Authors: Wessel Reijers,*[†] Fiachra O'Brolcháin, [‡] Paul Haynes[§]

Reviewers: Reviewer A, Reviewer B, Reviewer C

Abstract. The final version of the paper "Governance in Blockchain Technologies & Social Contract Theories" can be found in Ledger Vol. 1 (2016) 134-151, DOI 10.5915/LEDGER.2016.62. There were three reviewers, none of whom have requested to waive their anonymity at present, and are thus listed as A, B, and C. After initial review (1A), the authors submitted a revised submission and response (1B). Reviewer C was asked to affirm that the revisions adequately and substantively addressed the criticisms. Reviewer C responded in the affirmative and supplied further notes (2A). The authors responded (2B). Authors' responses are in bullet form.

1A. Review, Initial Round

Reviewer A:

The topic is very important and relevant to Ledger, and this well-written manuscript certainly makes a meaningful contribution to the field.

I have only detected five minor shortcomings, which I believe can easily be revised by the authors:

1. Consistency in the narrative style in terms of tenses. For example, one can read "Hobbes says" (simple present) and "Hobbes viewed" (simple past).

2. Rephrase and simplify some sentences to improve the readability of the paper, such as the following: "This outlook ties in with anarchist and libertarian critiques of authority: which claim that..."; "Another aspect that early social contract theories have in common is that they relied on direct and voluntary consent between freely choosing rational individuals to justify the system inaugurated by the contract: the state"; and "In the same way that Hobbes' and Rousseau's conceptions of the social contract are derived from very different conceptions of authority, power and legitimacy, but their visions in many ways converge, such that at the

[†]W. Reijers (wreijers@adaptcentre.ie) is a PhD researcher at the School of Computing, Dublin City University
[‡]F. O'Brolcháin (fiachra.obrolchain@dcu.ie) is a postdoctoral researcher at the Institute of Ethics, Dublin City University
[§]P. Haynes (paul.haynes@rhul.ac.uk) is lecturer at the School of Management, Royal Holloway, University of London
* 3HrFGw5nuBup39tzvQT5reEF5gdtx8fDGw

structural level are in many ways 'deeply and decisively similar' so too the conception of the blockchain as a technological solution to politically derived problems, not merely converges with social contract principles but draws on and indeed manifests essential features of social contract theory".

3. The opening sentence of the conclusion section is the following: "Technologies structure the social and the political as society and politics structure technology". I find it a bit ambiguous. Does "as" imply the same degree, i.e., that technology structures society/economy/politics in the same way/level/degree that society/economy/politics structure technology? Mild techno-determinism, which, for example, can be found in Carlota Perez's work, suggests that technologies structure the social and the political, society and politics structure technology, but not in the same degree. Since the authors have not discussed in depth this issue (i.e., how technology shapes society and how it is shaped by it), I would propose that they rephrase this very first sentence by making it more generic.

4. The authors might consider to use a table which would summarise the comparison among social contracts eminent theorists and the use of the blockchain in Ethereum or blockchain in general. This table could improve the readability of the paper, especially for that part of the audience which is not used to scholarly texts.

5. I would suggest that the authors use this citation "Kostakis, V., & Giotitsas, C. (2014). The (A)Political Economy of Bitcoin. tripleC: Journal for a Global Sustainable Information Society, 12(2), 431-440" instead of "Kostakis, V., & Giotitsas, C. "The (A)political economy of bitcoin." In P2P & inov. Rio de Janeiro 2. (2015)." since the latter is a reprint of the former (original).

In sum, my recommendation is acceptance with minor changes.

Reviewer B:

This is a very interesting paper on an important and central topic to the development of blockchain technology and to the blockchain community. I do think, in general, that it should be published in a journal like *Ledger*. But to get there I think it needs to do several things better. In general the problems I see stem in no small part from its length: the piece is about 5500 words long, including notes and bibliography, yet it takes on one of the central issues in Western political theory, three of the most prominent figures in that theory, and an emerging and not entirely clear aspect of a very new technology. It is simply too short to do this adequately: more detail on all of the questions it raises strikes me as necessary for a scholarly article on any topic, let alone such deep, interesting, and important ones. The paper needs to go into much more detail about Hobbes, Locke, Rousseau, and (I would suggest) Rawls and to demonstrate specifically how the blockchain community (and possibly the technology) realizes these ideas. (though see points 8 and 9 below on these two suggestions.)

Several major issues need to be addressed more carefully in a revision. Most importantly, the thesis of the paper needs to be clarified. On page 1, the authors pose the thesis as a question:

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

"whether the blockchain can be conceptualized as the technological manifestation of the social contract, as theorized by classical theorists Thomas Hobbes and Jean-Jacques Rousseau" (1), which they appear to answer in the affirmative. On page 2, they write that their paper "[explores] philosophical ideas common to both the blockchain and classical social contract theory." But toward the end of the paper the thesis mutates somewhat drastically: "though it seems that many essential aspects of social contract theories are present in the notion of social contract as it is incorporated in the Ethereum platform, they are underpinned by game theoretical principles and a firm believe [sic; should obviously be "belief"] in efficiency through market competition" (8) and "Ethereum, as conceptualized by blockchain proponents, is in some essential ways a modern manifestation of social contract theory" (8).

The versions of the thesis offered at the beginning of the paper are too vague. Ascribing "philosophical ideas" to "the blockchain" actually begs the question: why should readers believe that a technology has ideas in it? That seems to be what the authors want to show, but they do this by examining writings about the blockchain, often by developers—this shows that people working on the blockchain share these ideas, but it does not quite show that the ideas are in the blockchain. That is, it does not show that systems actually built with the blockchain would instantiate those ideas.

When in the final pages of the paper the authors introduce the idea of a game-theoretic notion of the social contract, they fail to note that this is a highly controversial and very specific reinterpretation of social contract theory that is not recognized by most current philosophers on the topic. The most famous contemporary social contract theorist is John Rawls, whose works go unmentioned here: it is very hard to reconcile Rawls's ideas with the game-theoretic model briefly mentioned at the end. If the thesis is that the blockchain instances a game-theoretic model of the social contract, that needs to be stated at the beginning, and the relationships of that model to the models familiar from Hobbes, Locke, Rousseau, and Rawls need to be delineated.

Some general issues:

1. The notions of "social contract," "contract" (which I'll call "ordinary contracts" in the rest of this review), and "smart contract" need to be carefully defined and distinguished. These concepts are not equivalent, and the paper is not at all careful enough to make this clear. The "social contract," with which the paper is primarily concerned, is an abstract, theoretical, *implicit* construct developed by philosophers (as the authors correctly note, historically most familiar from Hobbes, Locke, and Rousseau) to explain why human beings live in societies—but it is not something most citizens are or even should necessarily be familiar with; "contracts" are voluntary and *explicit* agreements between two parties to achieve some mutually-agreed upon goal; and "smart contracts" are software programs that appear in theory to mimic some of the characteristics of "ordinary contracts."

2. "The social contract," in so far as we can synthesize a single argument from Hobbes, Locke and Rousseau, is an abstract agreement among citizens to share political power amongst themselves. What would a "technological manifestation" of such an abstract, implicit agreement be? This is never stated clearly enough.

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

3. It's odd to me that the paper relies so much on Dupont and Maurer (2015), since I find that excellent essay to point much more clearly at even "smart contracts" as being more similar to ledgers or records of contracts rather than the contracts themselves, and the piece points out several differences between ordinary contracts and smart contracts, which the authors here fail to follow.

4. "For Hobbes, a core feature of this state of nature is that it is 'trustless', implying that individuals are unable to come to agreement on certain issues because they cannot trust that all parties involved will honor the agreement." (4): this sort of imposition of blockchain terminology on classical philosophical theorists strikes me as loading the argument. It is by no means clear that the concept of "trust" in blockchain technology (which stemmed from the metaphorical idea of machines being able to "trust" each other) is at all the same thing as the binding concepts in Hobbes. The paper needs to show this, not take it as given. Some of the most prominent blockchain theorists (especially Rachel O'Dwyer) explicitly reject the idea that blockchain "trust" and social "trust" are at all the same idea.

5. "Another aspect that early social contract theories have in common is that they relied on direct and voluntary consent between freely choosing rational individuals to justify the system inaugurated by the contract: the state." (4) This is just wrong. There is no "direct…consent" between the members of society in social contract theory: it is instead a theoretical reconstruction to answer the question "why do people collect in societies at all?" Further, while it is arguable that in Hobbes "the state" is what results from the social contract, in the works of Locke, Rousseau, and more recently Rawls, what the social contract creates is not "the state" but *society*itself. Gestures at "the state" as a single, easily-identifiable entity mark this paper as explicitly anarchist in nature, since few non-anarchist theorists accept the idea that there such an entity is a useful unit for social analysis—see particularly the work of Anthony Giddens on this question. Indeed, it is particularly odd to refer to "the state" in a work that discusses Locke and Rousseau, since both of these theorists very clearly thought that there are *different kinds* of states, and argued vociferously for the superiority of some kinds of states over others.

6. There is a similar failure to distinguish between "state" and "government": anarchists believe these forms are identical, but few other theorists, including most social contract theorists, do. While Hobbes was certainly talking about what he understood as "the state" (though writers like Giddens would argue that "the state" to Hobbes is very different from a contemporary nation-state), Locke and Rousseau were far more interested in "government." Further, both Locke and Rousseau were working hard to create what can at least be provisionally thought of as "decentralized" systems of government. That being the case, it's not clear why we need blockchain politics—on at least some versions of the Locke/Rousseau account, we already have decentralized political power. Again, it is only from an explicitly anarchist perspective that democratic governments are seen as centralized power.

7. "While the Ethereum Wiki page claims that 'ultimately, Ethereum could be used to run countries', much of the discussion concerns a variety of strategies and solutions to replace state monopoly including services such as issuing currency, online voting, decentralized

governance applications and exchange systems without third parties. These discussions show that there remain overt elements of the social contract in Ethereum's blockchain applications." (7) Statements like this one are really puzzling. Since the social contract is nothing but an abstract agreement between individuals to govern each other, there are no overt elements to it other than that basic fact. Further, if the social contract really is the abstract structure suggested by Locke and Rousseau, any social formation whatsoever would retain "overt elements" of it: only a complete Hobbesian "state of nature" would not.

8. The paper is really hampered by the fact that there are so few existing blockchain applications: it talks about several theoretical possibilities for applications that might be built on the Ethereum platform, but since there are so few running applications so far, it almost seems premature to make these arguments, at least about the technology. The arguments are much more cogently about the beliefs of the blockchain/Ethereum development community than they are about a technology that is barely out of beta.

9. The authors never note that the social contract is one of the single most disliked and rejected idea among libertarians and anarcho-capitalists, whom the authors admit make up the majority of the blockchain development community. Is the covert point of this paper to "rescue" the blockchain technology for non-libertarians who still value the social contract? Why doesn't the paper ever discuss the vast amount of anti-social-contract theory among the most prominent political writings in the blockchain community? Given that so little overt support for social contract theory is uncovered in the paper, at times it reads like an apology before the fact, trying to convince readers that even though many in the blockchain community subscribe to philosophies that overtly reject the social contract, bits of it still remain. I don't find this particularly convincing, especially not without much more evidence in support of it.

Reviewer C:

The core thesis the authors defend is that some "essential elements" of the social contract theory are manifested in blockchain technology. According to social contract theory, the consent (in some loose sense of that word) of individuals to be members of a political community (what the authors call "political organizations") is what justifies the enforcement of the rules of that community. Similarly, blockchain technologies create communities to which its members consent (in some sense), and this justifies the enforcement of contracts and other agreements made through blockchain. (I ignore here some points about how this enforcement is decentralized.)

The authors convincingly defend this core thesis, which is not very controversial. Any voluntary transaction at all manifests these "essential elements". But I don't think their argument as it is supports more controversial and interesting claims about the relationship between blockchain and social contract theory. For example, on page 8 when the authors talk about blockchain "facilitating" certain social interactions that have game-theoretic properties described by social contract theory, I take their claim to be the following. Blockchain is a device that can*create* political communities that have various special properties according to

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

social contract theory: they are in equilibrium, they are justified, legitimate, etc. I don't think the argument as it stands supports this more controversial thesis about the creation of political communities. My concerns all focus on their discussion of the enforcement of rules in blockchain.
What is a political community? Presumably, the authors have in mind something like this: in a political community members agree to and expect others to abide by a particular system of property rights and contracts (among other things), and to be punished for violations. The social contract is a device for justifying the existence of political communities. Now, can blockchain create a political community? According to the "case study" of Ethereum, it can. But I don't see why social contract theorists would agree. For Hobbes, we only have a system of property rights after establishing a sovereign. This sovereign keeps us all "in awe" with the threat of punishment if we violate its rules. If I steal your property, I face banishment from the community, imprisonment, or death. Clearly, it is not in my interest to steal from you given this threat of punishment.

Now, blockchain punishes offenders through banishment from the network. Is the threat of banishment severe enough to make it not in my interest to violate the rules of blockchain? Maybe it is if the blockchain keeps track of my holdings in a currency like bitcoin. Bitcoin only has value insofar as the members of the blockchain agree that it has value. If I am excluded from the network, my bitcoins become worthless to me. So, I have no reason to violate the rules of the blockchain if doing so means losing my bitcoins. But if a blockchain supports a private property regime, the blockchain would be keeping track of many more goods than just digital currency: stuff like houses and cars, say. Some of these goods, unlike bitcoin, would have value for me independently of whether or not I belong to the blockchain network. If my goods have value whether or not I belong to the network, why should the threat of banishment from the network be sufficient to make me abide by the blockchain's rules? For blockchain to "facilitate" political communities where members agree on property rights in tangible items, the threat of banishment from the network would need to be severe enough to make it in no one's interest to violate the rules. Otherwise, members of the network have no reason to trust that their peers will abide by the rules. And without this trust, we have no functioning political community, in the Hobbesian sense.

There are two things I'd like to hear the authors explain in more detail. First, how could blockchain actually create private property regimes? The authors cite a paper that discusses this in footnote 12, but I would find it useful to give an example in the body of the paper. I have some sense of how blockchain creates a ledger that tracks bitcoin holdings, but I think the paper could use more discussion of how blockchain does this for other forms of property besides digital currencies. (I say this as someone who hasn't read any of the cited literature of blockchain.) Second, why think that the enforcement of the rules of blockchain through banishing violators is sufficient to create and sustain stable property regimes? This second question surely can't be adequately addressed in such a short paper. But some discussion of the question would help clarify which claims the authors take themselves to establish and what further work needs to be done.

Other comments:

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

1. Typo—"overs" page 5, in the Hobbes quote, should be "others"

2. Typo—"this late" page 5, in the Hobbes quote, should be "this latter"

3. On page 7 the authors write that "everyone should be able to for instance define her own property right regime or create her own cryptocurrency." Why is this? You can't have a private property regime if no one but yourself agrees to it: if I think I own everything, and you think you own everything, we're just in Hobbes' state of nature, where there is no private property at all.

4. I didn't follow on page 8 how different cryptocurrencies could "compete" with each other. What are the shared "default" rules the authors refer to? Are these rules that are common to all cryptocurrencies? I found it very hard to follow the thread from the game-theoretic properties of social contract theories to why blockchain technologies should lead to a market equilibrium.

5. Typo—"believe" page 8, penultimate paragraph of section 4, should be "belief"

6. Typo—"a" should be deleted in "implies a essential shifts" bottom of page 8

1B. Authors' Response

We would like to thank the reviewers for their very valuable input, which allowed us to revise the manuscript quite significantly. Hopefully, we have managed to address most of the criticisms. Below, we tried to address each comment separately, indicating where and how the comments have been dealt with in the text.

Response to Reviewer A:

The topic is very important and relevant to Ledger, and this well-written manuscript certainly makes a meaningful contribution to the field.

I have only detected five minor shortcomings, which I believe can easily be revised by the authors:

1. Consistency in the narrative style in terms of tenses. For example, one can read "Hobbes says" (simple present) and "Hobbes viewed" (simple past).

- We made sure the narrative style is consistent throughout the document (using present tense).

2. Rephrase and simplify some sentences to improve the readability of the paper, such as the following: "This outlook ties in with anarchist and libertarian critiques of authority: which

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

claim that...”; “Another aspect that early social contract theories have in common is that they relied on direct and voluntary consent between freely choosing rational individuals to justify the system inaugurated by the contract: the state”; and “In the same way that Hobbes’ and Rousseau’s conceptions of the social contract are derived from very different conceptions of authority, power and legitimacy, but their visions in many ways converge, such that at the structural level are in many ways 'deeply and decisively similar' so too the conception of the blockchain as a technological solution to politically derived problems, not merely converges with social contract principles but draws on and indeed manifests essential features of social contract theory”.

- These and a number of other sentences were revised to increase readability and clarity. Also in the additional text, we have tried to prevent ambiguous sentences as much as possible, aiming at analytic clarity.

3. The opening sentence of the conclusion section is the following: “Technologies structure the social and the political as society and politics structure technology”. I find it a bit ambiguous. Does “as” imply the same degree, i.e., that technology structures society/economy/politics in the same way/level/degree that society/economy/politics structure technology? Mild techno-determinism, which, for example, can be found in Carlota Perez's work, suggests that technologies structure the social and the political, society and politics structure technology, but not in the same degree. Since the authors have not discussed in depth this issue (i.e., how technology shapes society and how it is shaped by it), I would propose that they rephrase this very first sentence by making it more generic.

- We have totally rewritten the conclusion. Nevertheless, we clarify our position on this issue in the end of section two (mentioning writings in philosophy of technology).

4. The authors might consider to use a table which would summarise the comparison among social contracts eminent theorists and the use of the blockchain in Ethereum or blockchain in general. This table could improve the readability of the paper, especially for that part of the audience which is not used to scholarly texts.

- We have considered using a table, but because of the multi-dimensional nature of the comparison we provide we didn’t conceive of a feasible way to do this in a compact and coherent manner.

5. I would suggest that the authors use this citation “Kostakis, V., & Giotitsas, C. (2014). The (A)Political Economy of Bitcoin. tripleC: Journal for a Global Sustainable Information Society, 12(2), 431-440” instead of “Kostakis, V., & Giotitsas, C. “The (A)political economy of bitcoin.” In P2P & inov. Rio de Janeiro 2. (2015).” since the latter is a reprint of the former (original).

- We have changed this reference.

In sum, my recommendation is acceptance with minor changes.

Response to Reviewer B:

This is a very interesting paper on an important and central topic to the development of blockchain technology and to the blockchain community. I do think, in general, that it should be published in a journal like Ledger. But to get there I think it needs to do several things better. In general the problems I see stem in no small part from its length: the piece is about 5500 words long, including notes and bibliography, yet it takes on one of the central issues in Western political theory, three of the most prominent figures in that theory, and an emerging and not entirely clear aspect of a very new technology. It is simply too short to do this adequately: more detail on all of the questions it raises strikes me as necessary for a scholarly article on any topic, let alone such deep, interesting, and important ones. The paper needs to go into much more detail about Hobbes, Locke, Rousseau, and (I would suggest) Rawls and to demonstrate specifically how the blockchain community (and possibly the technology) realizes these ideas. (though see points 8 and 9 below on these two suggestions.)

- We have elaborated more in detail on our interpretations of social contract theories and included Rawls as well in our analysis. Moreover, we have narrowed the scope of the paper by deleting the entire last section of the paper (on the social contract in Ethereum) and focusing only on the justification of governance and the modeling of governance in blockchain technologies as compared with social contract theories.

Several major issues need to be addressed more carefully in a revision. Most importantly, the thesis of the paper needs to be clarified. On page 1, the authors pose the thesis as a question: "whether the blockchain can be conceptualized as the technological manifestation of the social contract, as theorized by classical theorists Thomas Hobbes and Jean-Jacques Rousseau" (1), which they appear to answer in the affirmative. On page 2, they write that their paper "[explores] philosophical ideas common to both the blockchain and classical social contract theory." But toward the end of the paper the thesis mutates somewhat drastically: "though it seems that many essential aspects of social contract theories are present in the notion of social contract as it is incorporated in the Ethereum platform, they are underpinned by game theoretical principles and a firm believe [sic; should obviously be "belief"] in efficiency through market competition" (8) and "Ethereum, as conceptualized by blockchain proponents, is in some essential ways a modern manifestation of social contract theory" (8).

- We have changed the thesis of our paper, improving its clarity and using it in a consistent way throughout the paper. It consists of two parts now: (1) to what extent the justification of blockchain governance reflects justifications of governance offered by social contract theories. This question mostly looks at what the blockchain community says *about* the technology. The second part is (2) to what extent the models of governance offered by social contract theories reflect the model of blockchain governance. Here, we do look at the design features of the technology. In order to justify our stance, we clarify our stance regarding the way in which technologies can embody political ideas in the end of section 2; citing works in philosophy of technology.

The versions of the thesis offered at the beginning of the paper are too vague. Ascribing

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

"philosophical ideas" to "the blockchain" actually begs the question: why should readers believe that a technology has ideas in it? That seems to be what the authors want to show, but they do this by examining writings about the blockchain, often by developers—this shows that people working on the blockchain share these ideas, but it does not quite show that the ideas are in the blockchain. That is, it does not show that systems actually built with the blockchain would instantiate those ideas.

- This issue is discussed now in the end of section 2 (where we explicate how to understand that a technology embodies political ideas) and in section 4, where we explicitly look at blockchain governance in terms of design features of the blockchain. We explicitly do not want to offer a technological determinist account, but instead discuss to what extent the relevant design features of the blockchain can be said to reflect aspects of the models of governance offered by social contract theories.

When in the final pages of the paper the authors introduce the idea of a game-theoretic notion of the social contract, they fail to note that this is a highly controversial and very specific reinterpretation of social contract theory that is not recognized by most current philosophers on the topic. The most famous contemporary social contract theorist is John Rawls, whose works go unmentioned here: it is very hard to reconcile Rawls's ideas with the game-theoretic model briefly mentioned at the end. If the thesis is that the blockchain instances a game-theoretic model of the social contract, that needs to be stated at the beginning, and the relationships of that model to the models familiar from Hobbes, Locke, Rousseau, and Rawls need to be delineated.

- We have embedded the game-theoretical nature of the model of governance offered by blockchain technologies in the discussion of the social contract theories; showing how it can be most closely aligned with Hobbes but also how it eventually leads to some important inconsistencies with the theories of Rousseau and Rawls.

Some general issues:

1. The notions of "social contract," "contract" (which I'll call "ordinary contracts" in the rest of this review), and "smart contract" need to be carefully defined and distinguished. These concepts are not equivalent, and the paper is not at all careful enough to make this clear. The "social contract," with which the paper is primarily concerned, is an abstract, theoretical, implicit construct developed by philosophers (as the authors correctly note, historically most familiar from Hobbes, Locke, and Rousseau) to explain why human beings live in societies—but it is not something most citizens are or even should necessarily be familiar with; "contracts" are voluntary and explicit agreements between two parties to achieve some mutually-agreed upon goal; and "smart contracts" are software programs that appear in theory to mimic some of the characteristics of "ordinary contracts."

- We have clarified these issues in the paper, offering definitions of "contract", "smart contract" and "social contract" – notably on p.4

2. "The social contract," in so far as we can synthesize a single argument from Hobbes, Locke

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

and Rousseau, is an abstract agreement among citizens to share political power amongst themselves. What would a "technological manifestation" of such an abstract, implicit agreement be? This is never stated clearly enough.

- We don't refer to the notion of "technological manifestation" any more in the current manuscript. Rather, we have nuanced the argument, in line with the two questions discussed above.

3. It's odd to me that the paper relies so much on Dupont and Maurer (2015), since I find that excellent essay to point much more clearly at even "smart contracts" as being more similar to ledgers or records of contracts rather than the contracts themselves, and the piece points out several differences between ordinary contracts and smart contracts, which the authors here fail to follow.

- We have tried to account for the interesting differences outlined by the paper of Dupont and Maurer on p.4.

4. "For Hobbes, a core feature of this state of nature is that it is 'trustless', implying that individuals are unable to come to agreement on certain issues because they cannot trust that all parties involved will honor the agreement." (4): this sort of imposition of blockchain terminology on classical philosophical theorists strikes me as loading the argument. It is by no means clear that the concept of "trust" in blockchain technology (which stemmed from the metaphorical idea of machines being able to "trust" each other) is at all the same thing as the binding concepts in Hobbes. The paper needs to show this, not take it as given. Some of the most prominent blockchain theorists (especially Rachel O'Dwyer) explicitly reject the idea that blockchain "trust" and social "trust" are at all the same idea.

- We acknowledge that we mistakenly used the notion of trust wrongly and created the impression that human trust is similar to trust in the blockchain. We now make explicit that different kinds of trust are implied, also referring to the excellent paper of O'Dwyer, on pages 7 and 8)

5. "Another aspect that early social contract theories have in common is that they relied on direct and voluntary consent between freely choosing rational individuals to justify the system inaugurated by the contract: the state." (4) This is just wrong. There is no "direct…consent" between the members of society in social contract theory: it is instead a theoretical reconstruction to answer the question "why do people collect in societies at all?" Further, while it is arguable that in Hobbes "the state" is what results from the social contract, in the works of Locke, Rousseau, and more recently Rawls, what the social contract creates is not "the state" but society itself. Gestures at "the state" as a single, easily-identifiable entity mark this paper as explicitly anarchist in nature, since few non-anarchist theorists accept the idea that there such an entity is a useful unit for social analysis—see particularly the work of Anthony Giddens on this question. Indeed, it is particularly odd to refer to "the state" in a work that discusses Locke and Rousseau, since both of these theorists very clearly thought that there are different kinds of states, and argued vociferously for the superiority of some kinds of states over others.

- We have deleted this claim and reworded the overall discussion in such a way that no reference is made to the "state" that could be in any sense be understood as being a "nation-state". Instead, we chose to refer to "governance" as the process leading to "government" as the actual focus of our interest.

6. There is a similar failure to distinguish between "state" and "government": anarchists believe these forms are identical, but few other theorists, including most social contract theorists, do. While Hobbes was certainly talking about what he understood as "the state" (though writers like Giddens would argue that "the state" to Hobbes is very different from a contemporary nation-state), Locke and Rousseau were far more interested in "government." Further, both Locke and Rousseau were working hard to create what can at least be provisionally thought of as "decentralized" systems of government. That being the case, it's not clear why we need blockchain politics—on at least some versions of the Locke/Rousseau account, we already have decentralized political power. Again, it is only from an explicitly anarchist perspective that democratic governments are seen as centralized power.

- We acknowledge this inconsistency. Therefore, we have tried to drop the references to the "state" understood as the nation state and redirected the attention to the process of governance. We also don't intend that we *need* blockchain technologies to instantiate decentralized governance (as in: without these technologies it would be impossible to do so). Rather, we try to show to what extent the conception of decentralized governance in Rousseau can be said to be similar to the one in blockchain governance.

7. "While the Ethereum Wiki page claims that 'ultimately, Ethereum could be used to run countries', much of the discussion concerns a variety of strategies and solutions to replace state monopoly including services such as issuing currency, online voting, decentralized governance applications and exchange systems without third parties. These discussions show that there remain overt elements of the social contract in Ethereum's blockchain applications." (7) Statements like this one are really puzzling. Since the social contract is nothing but an abstract agreement between individuals to govern each other, there are no overt elements to it other than that basic fact. Further, if the social contract really is the abstract structure suggested by Locke and Rousseau, any social formation whatsoever would retain "overt elements" of it: only a complete Hobbesian "state of nature" would not.

- We agree that this statement and some related statements were very confusing. We therefore re-phrased our theses throughout the paper and deleted these confusing sentences.

8. The paper is really hampered by the fact that there are so few existing blockchain applications: it talks about several theoretical possibilities for applications that might be built on the Ethereum platform, but since there are so few running applications so far, it almost seems premature to make these arguments, at least about the technology. The arguments are much more cogently about the beliefs of the blockchain/Ethereum development community than they are about a technology that is barely out of beta.

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

- Yes, this is necessarily a drawback of any critical reflection on emerging technologies (not just blockchain technologies, but also AI, forms of robotics, human enhancement etc.). Nevertheless, with the proper nuance (we have tried to extensively nuance our claims wherever appropriate, throughout the paper), we think that these questions can be at least asked and (partially) be answered. Especially since these technologies, perhaps more than other emerging technologies, are already tested out and implemented on a large scale (notably Bitcoin of course).

9. The authors never note that the social contract is one of the single most disliked and rejected idea among libertarians and anarcho-capitalists, whom the authors admit make up the majority of the blockchain development community. Is the covert point of this paper to "rescue" the blockchain technology for non-libertarians who still value the social contract? Why doesn't the paper ever discuss the vast amount of anti-social-contract theory among the most prominent political writings in the blockchain community? Given that so little overt support for social contract theory is uncovered in the paper, at times it reads like an apology before the fact, trying to convince readers that even though many in the blockchain community subscribe to philosophies that overtly reject the social contract, bits of it still remain. I don't find this particularly convincing, especially not without much more evidence in support of it.

- One of the motivations of writing this paper has been to engage with the notion of the "social contract" as it is used in many writings of proponents of blockchain technologies, by comparing it with the academic literature on the topic. Thereby, we did not want to suggest that these are in any way in agreement with one-another. We are aware that the "ideology behind" blockchain technologies is more of an anarchist/libertarian kind; though it does not extensively engage with the academic roots of these ideologies (there are some interesting Reddit threads in which people of the blockchain developers community discuss their interest in actual Anarchist and Libertarian works and most of them said they had no knowledge of these or exclusively had read Proudhon's work. Even though we did not have time not the sufficient research done to include Proudhon in our discussion, he actually proposes a type of social contract theory that would be highly compatible with the one that could be derived from "blockchain governance"). In any case, we have tried to refer more extensively to the mismatch between the actual ideological backing of the blockchain developers community and the social contract traditions we discuss.

Response to Reviewer C:

The core thesis the authors defend is that some "essential elements" of the social contract theory are manifested in blockchain technology. According to social contract theory, the consent (in some loose sense of that word) of individuals to be members of a political community (what the authors call "political organizations") is what justifies the enforcement of the rules of that community. Similarly, blockchain technologies create communities to which its members consent (in some sense), and this justifies the enforcement of contracts and other agreements made through blockchain. (I ignore here some points about how this enforcement is decentralized.)

The authors convincingly defend this core thesis, which is not very controversial. Any voluntary transaction at all manifests these “essential elements”. But I don’t think their argument as it is supports more controversial and interesting claims about the relationship between blockchain and social contract theory. For example, on page 8 when the authors talk about blockchain “facilitating” certain social interactions that have game-theoretic properties described by social contract theory, I take their claim to be the following. Blockchain is a device that can create political communities that have various special properties according to social contract theory: they are in equilibrium, they are justified, legitimate, etc. I don’t think the argument as it stands supports this more controversial thesis about the creation of political communities. My concerns all focus on their discussion of the enforcement of rules in blockchain.

- This comment is in line with the main concern of the second reviewer, namely that different theses are presented in the manuscript that are either too general or not sufficiently defended. We have tried to tackle this issue by proposing a more consistent thesis that consists of two questions that correspond to the two main sections of the paper. Now, the issues of for instance the game-theoretical properties of blockchain governance are connected to these sub-questions (see e.g. p.8).

What is a political community? Presumably, the authors have in mind something like this: in a political community members agree to and expect others to abide by a particular system of property rights and contracts (among other things), and to be punished for violations. The social contract is a device for justifying the existence of political communities. Now, can blockchain create a political community? According to the “case study” of Ethereum, it can. But I don’t see why social contract theorists would agree. For Hobbes, we only have a system of property rights after establishing a sovereign. This sovereign keeps us all “in awe” with the threat of punishment if we violate its rules. If I steal your property, I face banishment from the community, imprisonment, or death. Clearly, it is not in my interest to steal from you given this threat of punishment.

- To shortly reply to this comment: because of the difficulty of focusing on the idea of “political community” (referring to the comment of the second reviewer: should we understand this as a state, a nation state?), we re-directed the focus to governance, as the process that leads to the forming of a political community.

Now, blockchain punishes offenders through banishment from the network. Is the threat of banishment severe enough to make it not in my interest to violate the rules of blockchain? Maybe it is if the blockchain keeps track of my holdings in a currency like bitcoin. Bitcoin only has value insofar as the members of the blockchain agree that it has value. If I am excluded from the network, my bitcoins become worthless to me. So, I have no reason to violate the rules of the blockchain if doing so means losing my bitcoins. But if a blockchain supports a private property regime, the blockchain would be keeping track of

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

many more goods than just digital currency: stuff like houses and cars, say. Some of these goods, unlike bitcoin, would have value for me independently of whether or not I belong to the blockchain network. If my goods have value whether or not I belong to the network, why should the threat of banishment from the network be sufficient to make me abide by the blockchain's rules? For blockchain to "facilitate" political communities where members agree on property rights in tangible items, the threat of banishment from the network would need to be severe enough to make it in no one's interest to violate the rules. Otherwise, members of the network have no reason to trust that their peers will abide by the rules. And without this trust, we have no functioning political community, in the Hobbesian sense.

- This is a very valid point. Currently, only for *digital assets,* smart contracts can in fact be enforced by the blockchain. However, in the context of IoT (see p.3 and p.11), in which the world of physical object is linked with the world of digital assets, enforcement *can* happen through the blockchain (for instance by disabling a physical device whenever the contract is violated). We explicitly deal with the question of punishment by exclusion on p.11, by arguing that in order for it to be a deterrent, a particular blockchain that is linked to the physical world would need to be dominant enough (i.e. the cost of switching to another blockchain would need to be great enough for exclusion to be a deterring force). Such situations are very well possible, considering that specific blockchains (think of e.g. Ripple becoming the standard for inter-bank settlements) can become dominant, or standards in particular domains.

There are two things I'd like to hear the authors explain in more detail. First, how could blockchain actually create private property regimes? The authors cite a paper that discusses this in footnote 12, but I would find it useful to give an example in the body of the paper. I have some sense of how blockchain creates a ledger that tracks bitcoin holdings, but I think the paper could use more discussion of how blockchain does this for other forms of property besides digital currencies. (I say this as someone who hasn't read any of the cited literature of blockchain.) Second, why think that the enforcement of the rules of blockchain through banishing violators is sufficient to create and sustain stable property regimes? This second question surely can't be adequately addressed in such a short paper. But some discussion of the question would help clarify which claims the authors take themselves to establish and what further work needs to be done.

- We now discuss how property rights could be organized on the blockchain on p. 3. The question of sufficiency of banishment for rule-compliance is dealt with on p.11.

Other comments:
1. Typo—"overs" page 5, in the Hobbes quote, should be "others"
2. Typo—"this late" page 5, in the Hobbes quote, should be "this latter"
3. On page 7 the authors write that "everyone should be able to for instance define her own property right regime or create her own cryptocurrency." Why is this? You can't have a private property regime if no one but yourself agrees to it: if I think I own everything, and you think you own everything, we're just in Hobbes' state of nature, where there is no private property at all.

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

- Yes, this is a very valid point. This is the idea though behind the blockchain infrastructure: that people can promote their blockchain application (e.g. their own currency) in a digital "market space". This could lead to an interesting debate of whether market behavior in such a context would actually lead to a situation in which a substantial number of heterogeneous blockchain apps (e.g. many different cryptocurrencies) would actually co-exist, or whether it would lead to monopolies or oligopolies (which perhaps seems a more plausible situation at the moment). However, such a discussion would fall outside of the scope of our paper we think.

4. I didn't follow on page 8 how different cryptocurrencies could "compete" with each other. What are the shared "default" rules the authors refer to? Are these rules that are common to all cryptocurrencies? I found it very hard to follow the thread from the game-theoretic properties of social contract theories to why blockchain technologies should lead to a market equilibrium.

- Cryptocurrencies can compete by having different properties (e.g. proof of stake vs. proof of work & offering different degrees of pseudonymity). A cryptocurrency can advertise itself in the digital market place of cryptocurrencies and compete with other cryptocurrencies (for instance: Dogecoin and Zerocoin competing with Bitcoin). However, it's a big question whether we can speak of fair competition in this context, or whether Bitcoin and Ether can for instance already be said to create monopolies or oligopolies.

5. Typo—"believe" page 8, penultimate paragraph of section 4, should be "belief"
6. Typo—"a" should be deleted in "implies a essential shifts" bottom of page 8

2A. Review, Second Round

Reviewer B:

Overall, this is a significantly improved essay that is even more clearly about a vital topic in the blockchain community than was the first version, and I am happy to recommend publication with only minor revisions. I particularly admire the depth with which the paper engages the theories of Hobbes, Rousseau, and Rawls, that last of whom I think the paper benefits greatly from having added.

In general, the paper does need a careful round of copy-editing.

There are three parts of the argument where I see some of the ambiguity that I found in the earlier version cropping up, and where some effort might be made to distinguish between two different ideas.

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

The first is ambiguity about “social contracts” & “smart contracts,” best exemplified in this passage from page 10. when the authors write “creating different social contracts and ‘voting’ for them,” I’m not sure that is consistent with any of the models that have been discussed earlier. Almost by definition, people cannot choose different social contracts: the whole point is that all members of a given society have agreed (tacitly and abstractly) to the terms of the social contract. As the authors suggest elsewhere, it is more the blockchain as a whole—or particular instances of the blockchain, such as Ethereum or The DAO—that resemble the social contract, rather than individual contracts within a given blockchain:

The model of governance of the Ethereum platform is perhaps best described by Binmore, who states that “a social contract is”…“an equilibrium profile of strategies, one for each citizen. When the social contract operates, each citizen will therefore be optimizing when he follows the rules of behavior prescribed by his strategy” (1998: 355). This conception of the social contract, which intertwines it with game theoretical foundations, is used as the basis for designing the social interactions through blockchain technologies. Players, miners or eventually “citizens” compete with each other by creating different social contracts and “voting” for them (which simultaneously means investing in them). At the same time, participants are consenting by default with the agreed upon rules in a particular smart contract. Thanks to the power of the default, or market equilibrium, every outcome of the “game” will be most effective for the collective of participants. Thus, the aspects of social contract theories in blockchain governance on Ethereum are underpinned by game theoretical principles and a firm belief in efficiency through market competition.

The second is some lingering ambiguity about the location of rules, exemplified by this passage from page 11. The authors write that “disobeying the rules is made impossible.” But that most clearly refers to the blockchain software itself, not to the human beings who may be using the blockchain for whatever purpose. The farther the blockchain gets away from pure exchange of currency-like tokens, the clearer this becomes. At any rate, the locus of the “social contract” in this formulation would seem much more clearly to be instances of blockchain software rather than the individuals running the software:

Within a single blockchain, disobeying the rules is made impossible and will lead to exclusion from the system – i.e. the blockchain is totalitarian in terms of rule-enforcement, which makes it comparable to Hobbes’ Leviathan. Moreover, no blockchain can be altered or manipulated by the individuals who use it to contract with one-another. Because fraud and counterfeit are rendered structurally impossible, once a person has contracted with someone else through the blockchain she has no other choice but to abide by its rules.

This second point leads to the third, which might also be addressed somewhere, even if briefly. The main thrust of this paper strikes me as being about something like the models of governance and society itself that we see manifested in the blockchain *communities* (and, of course, that might proceed from actual instances of running blockchains). But at times like the last passage just quoted, the subject slips slightly to something like how blockchain governance actually works, which the authors don’t really spend enough time addressing—nor should they—in part because we have very few examples of actually-running blockchain smart contract systems. Such systems remain highly speculative. The discussion of The DAO

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

is good in this regard, but prior to that, I'd suggest reading through the paper for places, such as the last quoted paragraph, where the subject seems to be what *will* happen in actual blockchain governance systems, as that strikes me as highly speculative in nature. Note that this is partly what makes the paper so welcome: these issues should be discussed, in depth, prior to the creation of actual blockchain governance systems, lest we build a political system whose nature we really don't understand (which is my own personal take on what at least some blockchain proponents are actually doing).

A few minor notes:

1. Page 1: "Blockchain technology...is often regarded as the brainchild of a movement of anarchists, computer scientists and crypto-enthusiasts who saw Bitcoin as a long-awaited realization of an old "cypherpunk" dream of money that is free from control of the state and commercial banks."

"often regarded" suggests that what follows isn't entirely true and will be disagreed with in what follows. But it really isn't. I think the authors mean to say that *although* blockchain tech was the brainchild of anarchists etc. (which is true in a historical sense), it *may* realize governance models that don't fit those and may be attractive to people of other political persuasions, which is in fact what the article attempts to demonstrate.
2. in general, I'd try to avoid the genitive construction for Hobbes (ie., Hobbes') and rewrite the sentences that use it. In my lexicon the correct genitive is Hobbes's, but that is just as ugly & I'd still try to get around it.

3. page 14: missing closing quotation mark here: "veil of ignorance, being non-discriminatory, though it negates this idea because power-relations are predefined in the public ledger.

2B. Authors' Response to Second Round

Response to Reviewer B:

- We would like to thank the reviewer for the insightful additional remarks.
  We believe that the current revision has addressed the ambiguities identified by the reviewer and that we have undertaken a copy-editing of the article that has improved its consistency and eliminated the remaining errors.

Review of " Governance in Blockchain Technologies & Social Contract Theories" for Ledger

Overall, this is a significantly improved essay that is even more clearly about a vital topic in the blockchain community than was the first version, and I am happy to recommend publication with only minor revisions. I particularly admire the depth with which the paper

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

engages the theories of Hobbes, Rousseau, and Rawls, that last of whom I think the paper benefits greatly from having added.

In general, the paper does need a careful round of copy-editing.

- We have edited the entire document to provide the copy editing suggested, adding the corrections (the use of Hobbes' and the missing quotation marks) and a wide variety of small proof reading alternations to improve consistency and set a more consistently scholarly/academic tone.

There are three parts of the argument where I see some of the ambiguity that I found in the earlier version cropping up, and where some effort might be made to distinguish between two different ideas.

The first is ambiguity about "social contracts" & "smart contracts," best exemplified in this passage from page 10. when the authors write "creating different social contracts and 'voting' for them," I'm not sure that is consistent with any of the models that have been discussed earlier. Almost by definition, people cannot choose different social contracts: the whole point is that all members of a given society have agreed (tacitly and abstractly) to the terms of the social contract. As the authors suggest elsewhere, it is more the blockchain as a whole—or particular instances of the blockchain, such as Ethereum or The DAO—that resemble the social contract, rather than individual contracts within a given blockchain:

The model of governance of the Ethereum platform is perhaps best described by Binmore, who states that "a social contract is"…"an equilibrium profile of strategies, one for each citizen. When the social contract operates, each citizen will therefore be optimizing when he follows the rules of behavior prescribed by his strategy" (1998: 355). This conception of the social contract, which intertwines it with game theoretical foundations, is used as the basis for designing the social interactions through blockchain technologies. Players, miners or eventually "citizens" compete with each other by creating different social contracts and "voting" for them (which simultaneously means investing in them). At the same time, participants are consenting by default with the agreed upon rules in a particular smart contract. Thanks to the power of the default, or market equilibrium, every outcome of the "game" will be most effective for the collective of participants. Thus, the aspects of social contract theories in blockchain governance on Ethereum are underpinned by game theoretical principles and a firm belief in efficiency through market competition.

- This ambiguity indeed persisted in the document. We have wrongfully created the impression that smart contracts are somehow separate social contracts that people can vote for or chose between. Therefore, we re-phrased this passage on p.10 – making clear that the particular instances of the blockchain rather than smart contracts can be seen as resembling a social contract. We deleted the earlier sentence that explicitly causes the ambiguity.

The second is some lingering ambiguity about the location of rules, exemplified by this passage from page 11. The authors write that "disobeying the rules is made impossible." But

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

that most clearly refers to the blockchain software itself, not to the human beings who may be using the blockchain for whatever purpose. The farther the blockchain gets away from pure exchange of currency-like tokens, the clearer this becomes. At any rate, the locus of the "social contract" in this formulation would seem much more clearly to be instances of blockchain software rather than the individuals running the software:

Within a single blockchain, disobeying the rules is made impossible and will lead to exclusion from the system – i.e. the blockchain is totalitarian in terms of rule-enforcement, which makes it comparable to Hobbes' Leviathan. Moreover, no blockchain can be altered or manipulated by the individuals who use it to contract with one-another. Because fraud and counterfeit are rendered structurally impossible, once a person has contracted with someone else through the blockchain she has no other choice but to abide by its rules.

- We think that this ambiguity was to some extent dealt with in the original text (for instance by stressing that "within a single blockchain", disobeying the rules is made impossible). However, to make this more clear we added explicitly: "Important to note, however, is that this structural impossibility only exists within the system that runs on the blockchain. Participants running the software can circumvent this structural impossibility by opting out to use a certain blockchain technology or by switching between different blockchain technologies." (p.11). To some extent, this is also dealt with on p.12 where the rule-abiding within the system is juxtaposed with the rule-abiding outside of the system (for instance respecting property rights that are stored on a blockchain). We discuss this with reference to IoT property rights that could be organized on a blockchain.

This second point leads to the third, which might also be addressed somewhere, even if briefly. The main thrust of this paper strikes me as being about something like the models of governance and society itself that we see manifested in the blockchain communities (and, of course, that might proceed from actual instances of running blockchains). But at times like the last passage just quoted, the subject slips slightly to something like *how blockchain governance actually works*, which the authors don't really spend enough time addressing—nor should they—in part because we have very few examples of actually-running blockchain smart contract systems. Such systems remain highly speculative. The discussion of The DAO is good in this regard, but prior to that, I'd suggest reading through the paper for places, such as the last quoted paragraph, where the subject seems to be what will happen in actual blockchain governance systems, as that strikes me as highly speculative in nature. Note that this is partly what makes the paper so welcome: these issues should be discussed, in depth, prior to the creation of actual blockchain governance systems, lest we build a political system whose nature we really don't understand (which is my own personal take on what at least some blockchain proponents are actually doing).

- We believe that this ambiguity is partially addressed in the last paragraph of the conclusion, in which we discuss the limitations of our paper. However, we have explicitly added: "By doing so, we do not intend to provide an account of how blockchain government actually works, for such an account would be highly speculative in the current state of affairs in which no instance of wholly functioning

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62

> blockchain governance exists, but rather of similarities between models of governance as they are being claimed to manifest themselves through the use of blockchain technologies and those discussed by social contract theories" (p.9)

A few minor notes:

page 1: "Blockchain technology...is often regarded as the brainchild of a movement of anarchists, computer scientists and crypto-enthusiasts who saw Bitcoin as a long-awaited realization of an old "cypherpunk" dream of money that is free from control of the state and commercial banks."

"often regarded" suggests that what follows isn't entirely true and will be disagreed with in what follows. But it really isn't. I think the authors mean to say that although blockchain tech was the brainchild of anarchists etc. (which is true in a historical sense), it may realize governance models that don't fit those and may be attractive to people of other political persuasions, which is in fact what the article attempts to demonstrate.

- We re-phrased this sentence to avoid this misunderstanding for the reader; framing it as the historical development of blockchain technologies.

In general, I'd try to avoid the genitive construction for Hobbes (ie., Hobbes') and rewrite the sentences that use it. In my lexicon the correct genitive is Hobbes's, but that is just as ugly & I'd still try to get around it.

- We have tried to avoid this construction as much as possible throughout the paper.

page 14: missing closing quotation mark here:

"veil of ignorance, being non-discriminatory, though it negates this idea because power-relations are predefined in the public ledger.

- We included the quotation mark.

Ledger is published by the University Library System of the University of Pittsburgh as part of its D-Scribe Digital Publishing Program and is cosponsored by the University of Pittsburgh Press.

ISSN 2379-5980 (online)
associated article DOI
10.5915/LEDGER.2016.62