GRAZSCHUMPETERCENTRE

GSC Discussion Paper Series

Paper No. 21

Substituting Trust by Technology: A Comparative Study

Johanna Rath[1]

[1] Institute for Comprehensive Analysis of the Economy (ICAE)

Johannes-Kepler-University Linz, Austria

Johanna.rath@jku.at

Abstract

This study contrasts different effects of applying blockchain technology on a social norm of trust and individual behaviour. The advanced technological features of blockchain could either complete contractual information and prevent coordination failures by substituting the need for trust or allow for some degree of incompleteness in information and favour a reciprocal mechanism of trust to solve for inefficiencies arising out of it. Either way, incomplete information is a necessary condition for the emergence of social norms of trust and reciprocity; hence a change in the completion of contractual information influences the institutional setting that market mechanisms are embedded in. One evolutionary process drives both, the degree of information available and behavioural traits within the society. Technology is neutral, but the way it is applied has different consequences on the institutional setting and thus favours different individual behavioural traits. Blockchain technology might either substitute or complement the need for trust.

Keywords: trust, incomplete contracts, social norms, coordination failure

The Graz Schumpeter Centre is named after the famous Austrian social scientist Joseph A. Schumpeter, who taught and did his research in Graz from 1911 to 1921. It provides a platform of international cooperation and reach for young scientists, working on innovative and socially relevant topics.

Introduction

Trust is a ubiquitous component in human interactions and a natural way of people to solve for situations of missing information. This implies incomplete or asymmetric information about some aspects of the good to be traded (such as quality in the market of lemons; Akerlof, 1970), bounded rationality (e.g. Williamson, 1979) or other circumstances regarding the transaction such ex ante uncertainty (e.g. Knight, 1921) about the future state of the world to be realized. Conversely, in settings of complete information, individuals are able to bargain over the pareto-optimal allocation without the need of governance or third-party intermediaries providing credible commitment or trust in unknown trading partners. To illustrate this, consider a simple market transaction, given complete and symmetric information: If both trading partner know about their own as well as the others valuation of the good, the good's quality, its costs of production for every state of the complete set of future state possible, the contracting partners can rationally bargain over an efficient outcome (Coase, 1960). There would be little reason to set up a contract in written form, let aside a need for (costly) intermediaries to interfere and govern the transaction. Assuming the complete information assumption holds and the set of individual preferences is given the problem of allocation is solved by pure rational logic (e.g. Hayek, 1945).

Still, the complete information assumption is likely to be violated in most real world settings. This led to the emergence of a separated field of research dealing with the question of different modes of governing transactions in the presence of uncertainty and incomplete information. Among its most famous representatives are the Coasian theory of the firm and the transaction costs approach (Coase, 1937; 1960), Williamson's theory of opportunistic behaviour in contracting situations (Williamson, 1979) as well as theories of vertical integration (Williamson, 1985) and property rights (Grossmann & Hart, 1986; Hart & Moore, 1990). All of them are to some extend concerned with the question of how to organize transactions efficiently either by solving for ex post inefficiencies arising out of incomplete information or by incentivizing individuals to choose their behaviour such that an efficient allocation is achieved even in the absence of complete information.

Most recently, recent advances in technological capabilities have brought about a new type of governance. Since blockchain's distributed ledger technology entered the spectrum of governance modes for transactions (Davidson et al., 2018), a new possibility of information provision is available: Besides securely storing and processing information about previous transactions, certain conditions required for the exchange can be set and verified via the

blockchain. If these conditions are met[1], smart contracts automatically perform pre-specified transactions. Several authors (Holden & Malani, 2018; Meier & Sannajust, 2018) have argued for an application of these innovative features to support the functioning of self-regulating market price mechanisms by increasing the degree of verifiable information. This is done by applying revelation and renegotiation mechanisms that incentivize individuals to reveal their private information and prevent opportunistic behaviour in the presence of uncertainty. By providing proper commitments to the mechanisms via blockchain-based default options, the parties achieve an efficient allocation. However, as a consequence of providing and verifying additional information, the degree of contractibility is increased. This again is equivalent to saying that a higher number of complete contracts is available and a higher number of allocations solved by the logic of individual utility maximization without the need to trust in third-party intermediaries or the contracting partner. Such a shift in the available information represents a fundamental change in the institutional setting in which markets are embedded and might ultimately bring about further, unintended consequences.

The objective of this paper is to show how an increased number of complete contracts available might involve an evolutionary change in individual behaviour and the underlying system of norms and institutions, especially on a social norm of trust. Increasing contractibility ultimately promotes different behavioural traits within a society resulting from a path-dependent evolutionary process. Thus, an evolutionary game theoretical model is chosen to study the scope of transformation in the behavioural traits within a population as a (best) response to new circumstances associated with a change in the number of complete contracts. Building on a model by Bowles (2005) I argue, that different social norms of behaviour emerge in the context of complete and incomplete information. Whereby in the first case market-like allocations of rational self-interested individual behaviour are favoured (the mechanism design approach), in the second case conditions for the emergence of social norms of trust and cooperation apply (the social norms approach). I further discuss some problems and limits associated with each approach.

Section 1 deals with the interrelatedness of incomplete information and the need for a social norm of trust. Section 2 presents two approaches of decentral organization to deal with incompleteness in contracting and discusses the role of blockchain technology to support their functionality. In Section 3 an evolutionary model is presented to show how different technological and institutional settings provoke different forms of behaviour. Section 4 concludes.

[1] Think about a level of required wealth on the buyer's digital wallet to execute a certain purchase transaction.

Incomplete Information and a Social Norm of Trust

In 2008 trust in financial institutions was shaken tremendously by one of the deepest financial crisis of previous decades. Consequently, some people felt a strong need to replace hierarchical financial institutions (such as banks, notaries, etc.) and organize the economy in a more decentralised way. It was about this time, that cryptocurrencies like Bitcoin, based on the innovative blockchain technology first entered the public discussion of decentral governance and information provision. The blockchain's distributed ledger ´technology was first developed with the intention to solve the double-spending problem by providing secure and verified information on a digital network and therefore eliminating the need for trust in third-party intermediaries executing institutional tasks of information verification so far. The backbone of this innovative technological feature is to guarantee for a single source of truth by storing and verifying data in a decentralized manner. The blockchain network via consent mechanisms verifies transactions and pre-specified smart contracts are automatically executed via algorithmic solutions (on platforms such as Ethereum[2]). The process of transaction execution is thus decentral (rather than involving centralized, hierarchically-organized institutions) it is secure (it can only be changed by a majority voting on the network and is robust against individual manipulation) and it is public to everyone with access to the public ledger,[3] that keeps track of all past transactions (e.g. Buterin, 2014).

Proponents of this innovative way of information verification argue that blockchain enables a new form of P2P economy: Interactions are fully defined by contractual relationships and complete and symmetric information. Trust in unknown trading partners is redundant and the dependence on third party intermediaries is set to a minimum, because smart contracts can execute contractual content automatically and in a decentralized manner at overall lower transaction costs (e.g. Catalini & Gans, 2016; Swan, 2015; Tapscott & Tapscott, 2016; Voshmgir, 2019). Others raise critical voices, saying that blockchain technology might turn out to be manifestation of current capitalistic structures and associated power relations (Garrod, 2016) or arguing for the need of a new regulatory framework for blockchain application in order to address and prevent illegal activities (e.g. De Filippi & Wright, 2018). Some research focusses on the technology's limits with respect to their ability to substitute political institutions (Reijers et al., 2016; Racsko, 2019; Atzori, 2017), financial services (Böhme et al., 2015) or fiat money (Senner & Sornette, 2019) as well as on blockchain's ability as a system of self-governance (Spithoven, 2019). Others have highlighted the blockchain's ability to offer completely new ways of governance, labelling blockchain as an institutional innovation with the potential to change

[2] https://ethereum.org

[3] The ledger is public to every member of the respective blockchain network. This again, must not necessarily be openly accessible to everyone but the access can be restricted to certain requirements. In general, public and private blockchain networks are distinguished (e.g. Swan, 2015).

the overall institutional setting and to result in an evolution of capitalist institutions (Davidson et al., 2018). A comprehensive literature review on the topic is given by Garrod (2019), who further emphasizes the importance to study the technological transformation process along societal consequences arising out of it. I intend to follow propositions of the latter two by contrasting two possible scenarios in which different features of blockchain technology apply and lead to different institutional frameworks, which again evoke different consequences on social norms of trust and individual behaviour.

The Variety of Governance: Markets, Networks, Hierarchies

The discussion of organizing economic transactions in institutional settings apart from markets dates back to Ronald Coase and his influential work on the theory of the firm (Coase, 1937). In the presence on positive transaction costs, he argues, markets might not provide the most efficient allocation but other forms of organizations, such as firms, might be appropriate to govern economic transactions (Coase, 1960). This approach was further developed by Williamson (1971), who concluded that transactions are most likely organized hierarchically whenever market transactions are costly due to overall uncertainty about the transactional outcome. If costs arising out of bounded rationality and opportunistic behaviour are greater than bureaucratic costs of hierarchical organization, vertical integration is preferable (Williamson, 1971). He elaborated on this point by stating that market transactions are particularly costly in situations of relationship-specific investments (Williamson, 1979, 1985), that is an investment that is most valuable in exchange with one specific contracting party, rather than any other: it is specific to this trading relationship.

Consider the following situation of two contracting parties negotiating over a future trade. For the trade to be successful, one party needs to make a relationship-specific investment. The investing party now faces the potential risk to not recoup the value of their investment, given that subsequently their contracting partner insists on renegotiating the contract or refuses to trade otherwise. The latter might do so, to enforce prices that better meet their preferences, knowing that a relationship-specific investment would imply sunk costs for their contracting partner in case they refuse the trade. Once the relationship-specific investment is made, the contractor's power to renegotiate the contract is affected by the possibility to hold-up the investing party. The resulting uncertainty about the future outcome due to potential opportunistic behaviour is the reason for insufficient amount of investment, and thus, insufficient amount of exchange (Williamson 1979, 1985).

The hold-up problem is one specific form of coordination failure and one example of ex-post inefficiencies that arise out of situations with incomplete, asymmetric information. Both contracting partners could be better off if the potential harm of hold-up could be avoided. One way to achieve this is to vertically integrate the trading partners into a firm.

Grossmann and Hart (1986) contributed to the discussion by analysing the costs and benefits of vertical integration. They argue that contracts by their very nature are incomplete, due to bounded rationality about the future contingencies possible. Some future states of nature and characteristics about the trade are always non-verifiable, thus cannot be written into a contract. Nevertheless, this problem of incomplete contracts can still be solved by assigning residual rights of control to the firm owner, whereby the contracting parties achieve the second best through negotiation in situations of potential coordination failures. Still, the property rights approach faces the problem of how to make observable information verifiable (e.g. the sufficient amount of investment in case of potential hold-up) in order to achieve pareto-efficiency. Several solutions have been proposed in order to tackle this problem of incomplete information in contracts. These include approaches that either secure a sufficient amount of investment, even in the presence of potential hold-up via robust renegotiation mechanisms (Moore & Repello, 1988; Noldeke & Schmidt, 1995; Aghion et al., 1994) or revelation mechanism that incentivize individuals to truthfully reveal their private information about costs and values according to the state of the world revealed (Tirole, 1999; Maskin, 1999). What both mechanism design approaches have in common, is that they require a high level of commitment to the mechanisms proposed, such that the contracting parties do not deviate from the pre-specified path of renegotiation and pareto-efficiency is achieved. One common solution to guarantee commitment is to implement default option in case of deviation. Otherwise, the outcome might require a reassessment and further governance by third-party intermediaries such as courts, to solve any ex post resulting disputes or inefficiencies. If the incentive compatibility assumption holds, incompleteness in contracts is reduced by completing the missing information via renegotiation and revelation mechanisms in a decentralized way[4] (Aghion & Holden, 2011).

The benefits of decentralized governance of individual actions are, among others, emphasised by Hayek (1945) who argues for the market mechanism to provide a determination of values (such as prices) that incorporates a variety of very specific individual information of the members within a society. This set of information is dispersed among many members and consists of specific knowledge inherent to any individual and is only displayed by the individual's active choice to cooperate and reveal information. The most efficient organization of economic transactions, Hayek concludes, is the one that makes the best use of the existing knowledge within a society, implying a certain degree of flexibility to respond promptly to external changes. The prices determined on self-regulating markets reflect a sum of information that cannot be observed nor processed by any single planner. Therefore, any decentral spontaneous order allocation is superior to central institutions, since they make the fullest use of the existing knowledge dispersed in society (Hayek, 1945).

[4] For a comprehensive literature review on incomplete contracts and the theory of the firm see also Aghion & Holden (2011).

Still, in the broad discussion of institutional governance the markets vs. central planner contiuum is not sufficient to explain the wide varierty of organizational forms presented until today (e.g. Powell, 1990). Among decentral mechanism of governance the price mechanims active in competitive markets is only one example, networkbased cooperation demonstrates another way. Nevertheless, Hayek's argument to choose a mechanism according to his ability to make the best use of the existing knowledge and react to external changes is a helpful tool in the following comparative analysis of market- versus networkbased governance on the blockchain.

Contrasting Decentral Approaches to Incomplete Contracts

Williamson (1979, 1985), who first identified the hold-up problem offered a solution of long-term trading relationships in which possible future trades and associated pay-offs would be taken into account by the players, thus preventing opportunistic behaviour ex ante. This cooperative solution builds on reputation to escape the potential coordination failure. Similarly, Powell (1990) proposed a network governance structure that is organised in a decentralized way but still not guided by a market price mechanism. One main difference of networks compared to market allocation is the role of trust: In settings of clearly specified property rights and complete contractual information, there is no need to trust one another (apart to some trust and commitment to the existing legal framework). Whereby complete and symmetric information is a pre-condition for efficient market allocations, incomplete information is a necessary condition for a situation of cooperation and voluntarily sharing information within a network. The latter is characterized by promoting the complementary strengths of the exchanging partners, as well as a long-term pattern of interactions in which sanctions are of normative kind rather than legally enforced. Any form of conflict is not resorted ex post by courts or third-party intermediaries but ex ante via a norm of reciprocity and concerns about one's future reputation. This implies that actors' preferences are not given and independent of the institutional setting but are interrelated to the actor's environment as well as the actions of others. Coordination failures are prevented as the parties in a network agree to forego their own interest for the sake of the mutual beneficial outcome (Powell, 1990; Bowles, 2005). This was first pointed out by Axelrod & Hamilton (1894) who found that a social norm of trust and reciprocity achieve the pareto-optimal outcome in prisoner's dilemma situation by favouring cooperation. Based on the assumption of long-term interactions and repeated play, cooperative individual behaviour is promoted as well as punishing those who refuse to cooperate. The emerge of a social norm of reciprocity and trust within a network (tit-for-that strategy) further promotes cooperation and mutual learning, thus favours the voluntary exchange and use of information and leads to mutual beneficial outcomes.

Quite contrary to that, the mechanism design approach aims at completing the missing information by providing compatible incentives to reveal one's private information, thus make

missing information observable and verifiable. Coupled with renegotiation mechanisms that limit the number of possible future outcomes, mechanism design is said to overcome coordination failures (such as hold-up) in a slightly different way. I contrast this in the following with ex post solving for inefficiencies arising out of coordination failures via a social norm of trust and reciprocity, thus a social norms approach. Table 1 summarizes some main differences of mechanism-design and social-norm-approaches.

	Mechanism-Design-Approach	**Social-Norm-Approach**
Precondition:	complete information (common knowledge)	incomplete information
Mechanism:	designed mechanisms, compatible incentives	social norms of trust and reciprocity (decreasing conflict of interest)
Behaviour, Preferences:	Self-interested, rational	adaptive, endogenous
Duration:	single, one-shot interaction	long-term relationsships, repeated play
Outcome:	market-based, spontaneous order allocations	cooperation (mutually beneficial outcomes), networks
Information:	revealed, verified and stored on the blockchain	used and exchanged voluntarily (without permanently storing it on the blockchain)

Table 1: two contrasting approaches of solving coordination failures (own representation)

1.) Mechanism Design Approach via the Blockchain

Several authors have dealt with the question of how to implement findings from mechanism design into blockchain's computational codes (e.g. Holden & Malani, 2018; Meier & Sannajust 2018), mainly with the focus to solve for asymmetric information and the hold-up problem described by Williamson (1979, 1985). By guiding individual behaviour via mechanisms of constraints and incentives, different types of coordination failures are prevented ex ante or inefficiencies are resolved ex post. Necessary conditions to implement such mechanisms are

provided by blockchain technology: The 'rules of the game' could be published in a so-called Decentralized Autonomous Organisation (DAO), a platform, on which pre-specified rules can be set up for all members and single smart contracts can be written and automatically executed (Buterin, 2014). This implies coded default options or penalties, which the blockchain's algorithms automatically trigger, once individuals deviate from the predefined path of renegotiation. Thus, a neutral third party (the blockchain network) supervises the trade without any counterparty or interpretational risks involved. A combination of renegotiation and revelation mechanisms substitutes the need for trust in unknown contracting parties, as well as central third party intermediaries (Holden & Malani, 2018).

Renegotiation mechanisms, as proposed for example by Aghion, Dewatripont and Rey (1994), either ban renegotiation or pre-specify the path of renegotiation in a way such that the social optimum is achieved and hold-up is prevented. Each party is offered the possibility to request a default trade, which gives them potentially full return on their (relationship-specific) investment and guarantees for proper incentives for rational individual behaviour. Thus, blockchain and smart contracts can prevent deviation from the mechanism by including penalties for making a second offer (leading to a different outcome). This penalty is triggered automatically and spread in the blockchain network. Thus, the penalty flows to a neutral third party as required for compatible incentives[5]. Therefore, blockchain technology can offer sufficient amount of commitment to the renegotiation mechanism, the incentive compatibility assumption holds (Holden & Malani, 2018).

Still, some uncertainty about the future remains. The blockchain code would have to be written ex ante just like any conventional paper contract. Possible future states of the world that cannot be anticipated ex ante are missing in the smart contract on blockchain networks, just as they do in conventional contracting. Any unexpected event might lead to the need of adapting the former conditions of the contract to new circumstances, thus to renegotiate the contract. At the same time, this is also the reason why any ex post requested modification of a contract represents a potential hold-up: The challenge is to verify if either the change in circumstances is real, thus the renegotiation is appropriate and desired by both parties or the intention to renegotiation is driven by opportunistic behaviour of one party holding-up the other. Differentiating one from the other and verifying the real state revealed has been the task of institutions (like courts) until now, but could be taken over by the blockchain in the future by applying so-called revelation mechanisms (e.g. Maskin, 1999; Morre & Repullo, 1988). Those mechanisms applied via the blockchain shape incentives to truth-fully reveal private information about real world conditions and solve the implementation problem. In a nutshell, the mechanism goes as follows: The agents

[5] The penalty actually flows to a randomly chosen and unanimous set of nodes on the blockchain network. This is metaphorically, blockchain putting money on the street and people rushing to take it (Holden & Malani, 2018).

make announcement about their valuations (signalling) and receive pay-offs dependent on their announcement. The highest pay-offs are generated when telling the truth. This makes private information of individuals observable and thus verifiable to outsiders, thus to the blockchain network. Again, smart contracts can increase commitment to a certain path of interaction that individuals have to follow in order to reveal missing information and complete contracts. Incentive compatibility is secured by automatically triggering a default option in case the agents leave the pre-specified path, which in this case means to not reveal information truthfully (Holden & Malani, 2018).

Meier and Sannajust (2018) deal with the hold-up problem from a similar perspective, but focus more on the ability of blockchain technology to lower the overall transaction costs (c.f. also Catalini & Gans, 2016) and to avoid opportunistic behaviour by establishing symmetric information on the network. This is provided as any information sent and shared via the blockchain is stored and accessible by every member of the network. Hence a greater degree of transparency is achieved. Thus, information in smart contracts on the blockchain network is complete and symmetric. In case of ex post disputes and renegotiation required by one or both parties due to uncertainty, the authors propose the use of artificial intelligence taking over the task of courts and evaluating the situation based on the data available (Meier & Sannajust, 2018). Summarizing, features of blockchain technology as proposed by these authors have the potential to increase the number of complete contracts, providing a mode of governance unknown until today. It eliminates the need to trust in one's contracting partner and facilitates market-like, spontaneous order allocations by providing credible commitment to pre-specified negotiation mechanisms without interventions from central institutions.

However, there are some limits to this approach: Aghion et al. (2012) argued that complex, designed (revelation) mechanisms are not robust against small information perturbations about the contractual content. The mechanism requires that the contracting partners can agree on and both have knowledge about the precise values of costs and valuation of the exchanging good. If this is not the case, truth-telling is no longer the exclusive equilibrium of the game – in some cases, the information perturbations even cause truth-telling to no longer be an equilibrium at all. This means that a partial completion of the information can lead to a breakdown of the mechanism, with unpredictable outcome. Contracting in this situation again might require other forms of governance, like courts resolving the conflict (Aghion et al., 2012; Holden & Malani, 2018).

The degree of completion of contractual information possible, and hence the efficient functioning of the mechanisms presented above further depends on the ability of technical components to verify real world data. Consider a smart employment contract, in which worker's effort would need to be verified by the blockchain network in order to provide complete

information. In the near future, Internet of Things technologies might reveal great potentials for interactions of analogue and digital spheres, but this will also touch sensitive topics of data protection and individual privacy.

Additionally the degree of complete information depends on blockchain's ability to communicate and interact with other spheres of contracting. For example, a renegotiation mechanism, based on a take-it-or-leave-it-offer, requires a ban of renegotiation, thus a ban of second offers to guarantee for credible commitment. To prevent parties from renegotiating outside the blockchain sphere, the technology would need to be able to capture and detect any second offer agreements of renegotiation outside the blockchain network to actually ban any kind of renegotiation possible (that is e.g. to trigger the default option in case of a second verbal agreement). Otherwise, the conditions for efficient market exchanges on the blockchain require additional interference and regulations of legal institutions (e.g. a ban of second offer agreements outside the blockchain network) to resolve inefficiencies arising.

Moreover, there are types of knowledge, which, even if they were to be shared voluntarily, cannot be verified. One example to illustrate this is tacit knowledge (Nelson & Winter, 1982) that is best transferred learning-by-doing and can hardly be made verifiable in a contract or similar. Thus, tacit knowledge and other forms of intangible assets are also said to be best transferred and utilized in network settings. More generally, agreements on specific assets and completely verified information are best organized in markets, whereas exchanges under uncertainty about the outcome are best organized in networks, in which voluntarily sharing of information is more likely because trust is built through repeated interaction (Powell, 1990).

The arguments stated above hint to certain problems that might arise in completing contractual information implying that the degree of contractibility increases by blockchain application, but might still not provide complete information unless further technical and legal options are provided. Nevertheless, in order for rational, self-interested individuals to arrive at an efficient solution via bargaining, the common knowledge assumption must hold. If this cannot be provided by blockchain technology, one might reconsider the request for complete information and ask whether there are certain circumstances in which incomplete contracts are preferable. This might imply to maintain incompleteness even if partly completing information is possible.

2.) Social Norms Approach via the Blockchain

Whenever information is non-verifiable, other modes of governance might apply to use the existing knowledge without verifying it. For example, as proposed by Powell a notion of "[t]rust reduces complex realities far more quickly and economically than prediction, authority, or bargaining." (Powell, 1990, p. 305) Previously, Williamson (1985) pointed to the possibility of forming long-term trading relationships based on mutual trust in the trading partner to solve for

consequences of incompleteness. Assuming repeated interactions, a reputation mechanism promotes a tit-for-that strategy of individual behaviour leading to the superior cooperative equilibrium and preventing hold-up. In other words, a social norm of trust and reputation is built to secure the cooperative outcome by decreasing the conflict of interest between the trading partners.

The hold-up problem can be extended to any situation in which cooperation would give both contracting partners higher pay-offs, but still on an individual level is not utility-maximizing and thus not rationally chosen by self-interested single players[6]. The individual pay-offs for both contracting partners playing the self-interested rational strategy is lower than the outcome of both partners choosing to cooperate. Thus, mutual cooperation denotes a pareto-improvement compared to individually rational behaviour in the presence of complete information. Whenever information is non-verifiable, and thus incomplete, social norms and conventions potentially help to resolve the coordination failures. Furthermore, the level of trust is closely related to the level of completeness in information: The greater the degree of verified information, and thus the degree of completeness in contracts, the lower the need to trust the contracting partner. The same is true conversely: If contractual content is incomplete, the required condition to establish a social norm of trust and reciprocity is given.

The process of substituting trust by completing contracts via blockchain technology has the potential to outperform decentralized social norms of trust and reciprocal behavioural traits that depend on a certain degree of incompleteness in contracts. Both presenting a decentralized solution to incompleteness in contracting, mechanism design via blockchain might compete with existing institutional solutions of social norms mechanisms of trust and reciprocity: The former solving prisoner's dilemma by completing missing information in contracts and the latter by establishing behavioural traits besides rational self-interestedness.

The amount of data and information shared in the blockchain network is a question of technological feasibility, but also a question of choice. As argued above, there might be reasons to not complete contracts via mechanism design, even if feasible. The lack of complete information gives rise to certain endogenous enforcement strategies (Bowles, 2005) that outperform market mechanisms in some aspects. The functioning of social norms mechanisms of trust and fairness yields the most efficient solution to coordination failures or principal agent relations, not by incentivizing individual behaviour but by decreasing the conflict of interests.

[6] One main difference between hold-up and situations of prisoners dilemma lies in the fact that in the latter the game is played but individually rational utility-maximizing behaviour does not lead to the pareto-optimum (simultaneous play), while in presence of the former playing the game (or trading the good) is prevented all together by trading partners playing their individually rational utility-maximizing strategies (sequential play).

An evolutionary model is chosen to cope with the dynamic process of differential replication, and continuative an evolutionary distribution of behavioural traits. In this view, individuals are not the most basic element of social dynamics but rather adapt their behaviour to others and change their behavioural traits accordingly. In this process of cultural adaption some chance event might arise. Such a change, like the invention of blockchain technology and a following change in preferences due to completion of contractual information can be illustrated as an endogenous process in these evolutionary models. It thus goes beyond rational calculation of future (rationally) expected payoffs. Individual's behaviour in Bowles' evolutionary game theory model is based on a simply but convincing learning rule: a "copy-those-who-do-well-strategy" (Bowles, 2005). Following a replicator dynamic function, individuals constantly face opportunities to update their strategies. These models are helpful in presenting how cooperative behaviour arises in presence of incompleteness in contracting situations (Bowles, 2005).

The benefits of both trading partners arise by embracing the norm and generating trust, such that higher levels of cooperation are sustained. Further, a deeper investigation of the reputational model gives hints to other forms of blockchain applications, beside increasing contractibility: The large information base of blockchain's distributed ledger, accessible to every member on the network, can impact the way of detecting the potential trading partner's behaviour in past transactions. If blockchain is able to reduce such detection costs it can support the functioning of reputational mechanisms, thus complementing social norms mechanisms. Again, this implies to resist the completion of contractual contents, since behavioural traits of reputational mechanisms only emerge if certain incompleteness is maintained.

One basic assumption in reputational mechanisms is that partners have some information about the other. In the following I replicate a model by Bowles (2005) in which individuals pay a cost of detection δ to find out about past behavioural traits of their future trading partner. Now, one can either choose to condition their own behaviour on their partner's behaviour in previous transactions, that is inspect the partners past transactions and then adapt their own strategy (cooperate, if cooperation of the partner was detected in past transactions and defect otherwise). Call those inspectors (I): people choosing an equilibrium strategy to build up a reputation of being conditionally cooperative. Diversely, individuals can choose the strategy of unconditionally-defect (D) that is to not cooperate in any case but play a (rational) individual's own pay-off maximizing strategy.

Define α to be the fraction of inspectors in a population. The game is as an extended version of the prisoner's dilemma (with pay-off structure $a>b>c>d$ and $a + d > 2b$), with strategies "inspect" and "unconditionally defect". The former receives a pay-off b if they inspect a cooperative partner and c if an unconditionally defector is faced, reduced by the cost of detection δ. Similarly, pay-offs for unconditional defectors are c.

The relevance of the amount of detection costs δ in this setting is high:

- $\delta > 0$ allows for the equilibrium of D/D, implying $\alpha = 0$
- $\delta < b - c$ renders possible the cooperative equilibrium of I/I, leading to $\alpha = 1$

The model provides three equilibria, two stable corner solutions, as well as a possible third interior equilibrium of $\alpha \in (0, 1)$.

In equilibrium, strategy's pay-offs must be equal:

$$\pi^{I}_{(\alpha)} = \pi^{D}_{(\alpha)}$$

$$\alpha(b - \delta) + (1 - \alpha)(c - \delta) = c$$

- In case of $\alpha \in (0, 1)$, equality in payoffs gives $\alpha^* = \frac{\delta}{b-c}$

The optimal fraction of inspectors in the population depends on the amount of detection costs δ, as well as payoffs b and c. Further the stationary level of α^* is unstable, because $\frac{d(\pi^{I}(\alpha)-\pi^{D}(\alpha)}{d\alpha} = b - \delta > 0$. The equilibrium is not self-correcting for deviations. Put differently, small deviations from α^* do not result in a return to α^* but increase the expected payoff of one strategy relative to the other, and hence leads to the dynamic process that results in the corner solutions $\alpha = 1$ or $\alpha = 0$, depending on the historical composition of defectors and inspectors in the population. The result of the copy-those-who-do-well-strategy is a combination of the individuals own choice and the product of the choices of the others. Social norms and their evolution over time depend on how individuals adopt and hence, which behavioural traits are copied (or reproduced) and which are abandoned. As the interior equilibrium is unstable (as described above), the only evolutionary stable states of the system are the corner solutions $\alpha = 1$ and $\alpha = 0$. Any exogenous event of chance places the population at the evolutionary stable equilibria (the corner solutions) as a result of a path-dependent process (Bowles, 2005).

The key takeaway from this is that the cooperative, pareto-optimal equilibrium can be reached by building social norms of reciprocity and trust as for example via a mechanism of reputation assuming sufficiently low detection costs. This again can be provided by blockchain's large information base, hence increasing the probability of the cooperative equilibrium ($\alpha = 1$). If the level of detection costs δ is interrupted by applying blockchain technology it impacts the inspector-pay-off-function as well as the interior, unstable equilibrium. Decreasing detection cost leads to an upwards shift of the inspector-pay-off-function (leading to $\alpha^*_{BC} < \alpha^*$) and increases the basin of attraction of the cooperative equilibrium $\alpha \in (\alpha^*_{BC}, 1]$. If blockchain technology could achieve $\alpha = 0$, the non-cooperative equilibrium in strategies (D/D) would be eliminated in a sense that the strategies payoffs would be reduced to b and c (with $b > c$), leading

to 100 % inspectors in the population ($\alpha = 1$), thus strongly supporting the cooperative equilibrium.

Two applications for blockchain technology have been contrasted to solve for coordination failures arising from opportunism and uncertainty: (1) completing the contract by programming designed revelation and renegotiation mechanisms into the blockchain code (mechanism design approach) or (2) facilitating the functioning of reputational mechanism to implement the evolutionary stable norm equilibrium of cooperation (social norms approach). Thus, the consequences of blockchain technology on behavioural traits within the society are (to some degree) a question of how blockchain will be applied in the future. Clearly the two approaches differ in their assumptions and derivations. Further, it has already been indicated that they are mutually exclusive: A rise of social norms mechanisms is based on certain behavioural traits that emerge only in the presence of incompleteness. This process of endogenous preferences is shown in the following section.

Coevolution of Preferences and Institutions

The behavioural traits in a population as well as completeness in contracting are both subject to one underlying mechanism that connects the required level of trust to a specific set of preferences (Bowles, 2005). A completion of contractual content via smart contracts on blockchain application (mechanism design approach) might thus jeopardize these existing institutional solutions (social norms approach).

There is also some experimental evidence supporting this claim. Bartling et al. (1012) show, that different forms of contracts (incomplete, complete) provoke different forms of behaviour. In their experiment, people were ask to choose between complete sales contract (with ex ante fixed and clearly specified tasks) and incomplete employment contracts (that specified a principal agent relationship over some period, but no specific task to be executed by the agent). They find that incompleteness in employment contracts gives rise to reciprocity and hence trustful behaviour via a mechanism of reputation. These personal relationships yielded wages and effort levels in the experiment exceeding equilibria levels predicted by rational choice theory in settings of complete contracts (Bartling et al., 2012).

Individuals in the following model, taken from Bowles (2005)[7], are assumed to have some interest in the psychological make-up of their contracting partners. Moreover, one has some means to cause a change in the other's preferences, due to the long-term relationship and the accompanied effective threat of terminating the relationship. This threat of terminating the relationship has impacts on the individual's preferences, and thus on the emergence of certain behavioural traits. Social norms of trust and reciprocity are built to deal with the incentive

[7] This model was first suggested by Peter Skott, as quoted in Bowles (2005, p 261, Footnote).

problems and leading to situations of durable, personal exchange. Hence the degree of incompleteness in contracting is strongly related to the structures of markets and institutions. Using the blockchain to complete contracts increases the probability of the non-cooperative equilibrium, and hence selfish behavioural traits within the population. To see that this is the case, the evolutionary path and interdependencies of the degree of completeness and the associated changes in preferences, and thus the distribution of behavioural traits are closer examined in the following.

Consider a population that consists of buyers and sellers that are randomly paired for a single interaction. They trade over a good of high (H) or low (L) quality, with the level of quality (H or L) chosen by the seller, and costly to verify for the buyer. The buyer makes the first move to offer a contract. She can choose to offer a complete (C) or incomplete (I) contract. Sellers on the other hand can be reciprocators (R), reacting to a complete contract offered with mistrust as well as reacting to an incomplete contract with trust in the exchanging partner. But sellers can also be selfish (S) providing low quality, irrespective of the offered contract. If the buyer offers a complete contract, the seller receives some fixed compensation that is just enough to compensate her for providing the low quality good. The reciprocal seller (R) reacts to the complete contract by providing low quality and pays costs of δ. In the incomplete contract offer it is pre-specified (by assumption), that the buyer has to bear the costs of providing low quality plus half of the profit resulting from the exchange (that is, half of the higher pay-offs π^H for high quality, and π^L for low quality provision). Reciprocators (R) detect the incomplete contract offer, again implying additional detection costs of δ, and react with trust in a cooperative exchange by providing high quality. Further assume that $\pi^H > 2\pi^L$ and $\pi^H - \pi^L > 2\delta$. Pay-offs are given in table 2.

	R	**S**
I	$\frac{\pi^H}{2}, \frac{\pi^H}{2} - \delta$	$\frac{\pi^L}{2}, \frac{\pi^L}{2}$
C	$\pi^L, -\delta$	$\pi^L, 0$

Table 2: Endogenous preferences (Bowles, 2005)

Denote ω the fraction of R-sellers in the population, and ρ the fraction of I-contracts offered by the buyers. Now calculating the equilibrium fraction of ω^* and ρ^* by equating the expected pay-offs for buyers offering complete or incomplete contracts (as well as for sellers behavioural types respectively), gives

$$v^I = \omega \left(\frac{\pi^H}{2}\right) + (1 - \omega)\left(\frac{\pi^L}{2}\right)$$

$$v^C = \omega\,(\pi^L) + (1 - \omega)(\pi^L) = \pi^L$$

$$v^I = v^C, \text{ gives } \omega^* = \frac{\pi^L}{\pi^H - \pi^L}$$

The same procedure for the calculation of fraction of incomplete contracts gives ρ^*:

$$v^R = \rho \left(\frac{\pi^H}{2} - \delta\right) + (1 - \rho)(-\delta)$$

$$v^S = \rho \left(\frac{\pi^L}{2}\right)$$

$$\rho^* = \frac{2\delta}{\pi^H - \pi^L}$$

Now, in every period, players have the opportunity to update their strategies if higher pay-offs can be generated by choosing a different behaviour strategy. The replicator dynamic equations describing the systems dynamics over time are given by

$$\frac{d\rho}{dt} = \rho\,(1 - \rho)(v^I - v^C)$$

$$\frac{d\omega}{dt} = \omega\,(1 - \omega)(v^R - v^S)$$

The three stationary states resulting from this derivations are ρ=ω=0, ρ=ω=1 and the unstable equilibrium (ω^*, ρ^*). The only asymptotically evolutionary stable states are the corner solutions (Bowles, 2005).

Thus, if blockchain technology has some impact on the degree of contractual completeness possible it further influences the distribution of social norms and thus behavioural traits within a population. At ω=ρ=1 with all reciprocators and incompleteness, the incomplete contracts offered are not due to a non-feasibility of complete contracts, but more so as a best response to the fraction of reciprocators in the population. Hence, applying blockchain technology to complete contracts (mechanism design approach) might have severe consequences on social norms of trust and the frequency of reciprocators in a population. As only two evolutionary stable equilibria appear in this model, a higher contractibility shifts the system along an evolutionary path to universal selfishness as a best response to exclusively complete contracts.

In this setting, the opportunities to solve coordinate failures via social norms mechanisms are reduced due to a substitution of trust. Incomplete parts in contracts are the reason for social norms of trust that enable exchange partners to mutually benefit by generating higher pay-off equilibria than outcomes of self-interested, rational calculations. Further, the dynamical process

indicates that a greater amount of complete contracts might be required in the future to meet the higher frequency of self-interested behavioural traits in the population. This might result in further measures that aim at standardizing goods and increasing verifiability of contractual content, including new technological innovations (e.g. for monitoring worker's effort in smart employment contracts) and an ongoing digitalization and verification of goods and social interactions. Even if this might imply some advantages due to increased cost efficiency, it raises the necessity of a discussion about further institutional intervention to secure individual freedom and protection of privacy.

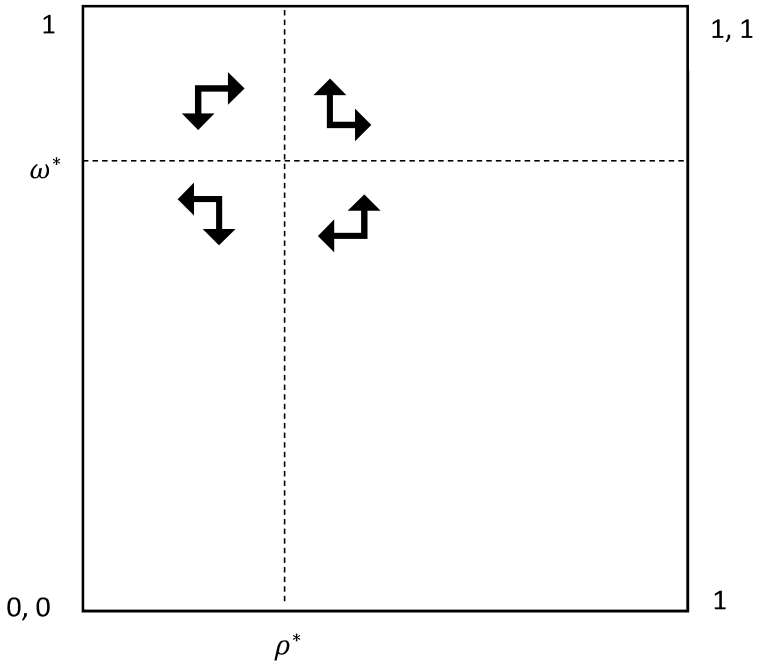

Figure 1: **The coevolution of (complete) contracts and behavioural traits:** The arrows indicate the dynamic movement of the system caused by small deviations from the non-self-correcting, unstable equilibrium (**ω***, **ρ***). A change in contractibility results in a change in preferences and hence drives the evolutionary system towards greater numbers of complete contracts offered (as a best response to a lower fraction of reciprocators in the population) until ρ=ω=0 or vice versa (own representation, as presented in Bowles, 2005).

Discussion: Rules vs. Values

Derivations of mechanism design approaches are based on a fundament of assuming rational, self-interested agents including rational expectations and hence, common knowledge of information of the agents involved. The basic performance figure to evaluate those mechanisms is the efficiency of allocation formed under these assumptions. Nevertheless, an evolutionary view enables several equilibria, some of them with overall higher payoffs than the individual utility maximizing Nash-equilibrium. Social norms in this framework secure the achievement of the overall higher pay-off equilibrium of cooperation.

To argue for social norm mechanisms is to argue for internalized values of individuals that do not require further governance to solve for coordination failures. Every member of a society acting as if someone was looking, that is to listen to their consciousness, is sufficient to avoid or abolish the consequences of market failure. These values might be different according to cultural and sociological circumstances developed throughout a historical process; therefore, they are constantly (re-)negotiated within a society and are modifiable, and thus flexible to uncertain events.

To replace them successfully (that is establish the higher-payoff equilibrium) by complete contracts might yield greater obstacles. Rules in complete contracts are comparably inflexible and need to be able to cope with any revealed state possible. If this is not the case, conventional institutions like courts will have to deal with uncertain and hence unspecified events in any case and will thus most likely not be replaceable by blockchain technology. Technology based on a process of rationalization cannot overcome the overall uncertainty of the future. Further assumptions need to hold for the efficient working of the designed mechanisms presented.

The limits of blockchain application maintaining the immutability-, decentralization- and open-to-everyone-feature is best illustrated with a short example: When the Ethereum platform was launched for the very first time in 2014 there were some flaws in the code, enabling a number of "hackers"[8] to exploit the code to their benefits. This was a turning point in the discussion of blockchain applications: for the very first time some disagreements within the network appeared. Should the "hacker's" transaction be voted invalid by the network to reverse the transaction or is the immutability feature of the network worth keeping with the implication that the "hackers" can keep the money? These ethical and ideological differences about how to deal with unwanted, but immutable transactions lead to a separation of the network in Ethereum, as known today (with a new, adapted code for higher security against these types of exploitation) and Ethereum Classic (the share of people sticking with the original code)[9].Taking this into account, the application of blockchain technology on a broader level will require legal regulations or a strongly segmented blockchain world.

Conclusion

In this paper the consequences of completion of contractual content by blockchain technology on well-established social norms of trust were investigated. Two decentral approaches to deal with incomplete contracts were discussed. Mechanism design approaches aim at revealing private information via revelation mechanisms and guide individuals via renegotiation mechanisms to achieve a pareto-efficient solution, thus substituting the need for trust.

[8] I use quotation marks, because the action of exploiting the code didn't require lots of technical finesse, as a hacker's attack usually does. "Opportunists" might be a valid denomination as well.
[9] See e.g. https://coincentral.com/ethereum-classic-vs-ethereum/ for an extended version of the story.

Blockchain technology in this setting can guarantee for proper commitments to the designed mechanisms by offering automatically triggered default options. Contrary to this, social norms approaches maintain certain degrees of incompleteness and decrease the conflict of interests by mutual cooperation to voluntary exchange the relevant information. Decreasing detection costs to find out about the contracting partner's reputation, blockchain is further able to complement a social norm of trust, supporting their functionality.

It was shown that one underlying evolutionary process drives both, the degree of completeness in information and the behavioural traits within the population: The higher the contractibility of information, the higher the degree of self-interested behavioural traits and thus, the lower the probability that the former achieve to end up at the superior cooperative equilibrium generating overall higher pay-offs. The same is true conversely for certain degrees of incompleteness and cooperative behaviour of individuals. Technology is neutral; the way it is used is not. Applying blockchain technology in order to complete informational content might potentially crowd out social norm mechanisms and substitute the need for trust. Contrary, blockchain might complement social norms of trust in a different way that is to support reputational effects by reducing detection costs.

References

Aghion, P., Dewatripont, M., & Rey, P. (1994). *Renegotiation Design with Unverifiable Information*. Econometrica 62, 257-282.

Aghion, P. & Holden, M. (2011). *Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years*?. Journal of Economic Perspectives 25 (2), 181-97

Aghion, P., Fudenberg, D., Holden, R., Kunimoto, T. & Tercieux, O. (2012). *Subgame-Perfect Implementation under Value Perturbations.* The Quarterly Journal of Economics 127 (4), 1843–1881.

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

Atzori, M. (2017). *Blockchain technology and decentralized governance: Is the state still necessary?*. Journal of Governance and Regulation (print) 6 (1)

Axelrod, R. M., & Hamilton, W. D. (1984). *The evolution of cooperation*. New York: Basic Books.

Bartling, B.; Fehr, E., & Schmidt, K. M. (2012). Use and Abuse of Authority: A Behavioural Foundation of the Employment Relation. Bonn: Forschungsinstitut zur Zukunft der Arbeit (IZA).

Bowles, S. (2005). *Microeconomics – Behaviour, Institutions and Evolution*. Princeton University Press.

Böhme, R.; Christin, N.; Edelman, B. & Moore, T. (2015). *Bitcoin: Economics, technology, and governance*. The Journal of Economic Perspectives 29, 213–238.

Buterin, V. (2014). *Ethereum: A next-generation smart contract and decentralized application platform.* https://github. com/ethereum/wiki/wiki/% 5BEnglish% 5D-WhitePaper.

Catalini, C. & Gans, J. S. (2016). *Some simple economics of the blockchain*. Discussion paper, National Bureau of Economic Research.

Coase, R. (1937). *The Nature of the firm*. Econometrica 4(16), 386-405.

Coase, R. (1960). *The Problem of Social Cost*. Journal of Law and Economics 3, 1-44.

Davidson,S.; de Filippi, P.; Potts, J. (2018). *Blockchains and the economic institutions of capitalism*. Journal of Institutional Economics 14 (4), 639 - 658.

De Filippi, P. & Wright, A. (2018). *Blockchain and the Law - The Rule of Code*. Harvard University Press.

Garrod, J. Z. (2016). *The Real World of the Decentralized Autonomous Society*. Triple C 14(1), 62-77.

Garrod, J. Z. (2019). *On the property of blockchains: comments on an emerging literature*. Economy and Society, 48 (4), 602-623.

Grossman, S. & Hart, O. (1986). *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration.* Journal of Political Economy 94 (4), 691-719.

Hart, O. & Moore, J. (1990). *Property rights and the nature of the firm*. Journal of Political *Economy* 98 (6), 1119-1158.

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

Holden, R. & Malani, A. (2018). *Can Blockchain Solve the Hold-Up Problem in Contracts?*. NBER Working Paper No. w25833. available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3093879 (last accessed 18.March 2020)

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. The Riverside Press

Maskin, E. (1999). Nash equilibrium and welfare optimality. *Review of Economic Studies* 66, 23-38.

Moore, J. & Repullo, R. (1988). Subgame Prfect Implementation, *Econometrica* 56 (5), 1191-1220.

Meier, O. & Sannajust, A. (2018). *Blockchain Revolution: A New Issue for the Hold-up Problem?*. fma conference papers. available at: http://www.fmaconferences.org/NewOrleans/Papers/Blockchainhold_up_problem_SANNAJUST-MEIER.pdf (last accessed 18. March 2020)

Nelson, R.R. and Winter, S.G. (1982). *An Evolutionary Theory of Economic Change.* Harvard University Press.

Noldeke, G. & Schmidt, K. M. (1995). *Option contracts and renegotiation: a solution to the hold-up problem.* RAND Journal of Economics 26 (2), 163-179.

Powell, W. W. (1990). *Neither Markets nor Hierarchy: Network Forms of Organization*. Research in Organizational Behaviour 12. 295-336.

Racsko, P. (2019). *Blockchain and Democracy*. Society and Economy 41 (3), 1-17.

Reijers, W.; O'Brolcháin, F. & Haynes, P. (2016). *Governance in Blockchain Technologies and Social Contract Theories*. Ledger Journal 1, 134-151.

Senner, R. & Sornette, D. (2019). *The Holy Grail of Crypto Currencies: Ready to Replace Fiat Money?*. Journal of Economic Issues 53 (4), 966-1000

Spithoven, A. (2019). *Theory and Reality of Cryptocurrency Governance*. Journal of Economic Issues 53 (2), 385-393.

Swan, M. (2015). *Blockchain - Blueprint for a New Economy*. O'Reilly Media Inc.

Tapscott D. & Tapscott, A. (2016). Blockchain Revolution: How the Technology Behind Bitcoin and Other Cryptocurrencies Is Changing the World. Penguin.

Tirole, J. (1999). Incomplete Contracts: Where Do We Stand?. *Econometrica* 67 (4), 741-781.

Voshmgir, S. (2019). *Token Economy: How Blockchains and Smart Contracts Revolutionize the Economy.* BlockchainHub.

Williamson, O. E. (1971). The Vertical Integration of Production: Market Failure Considerations. *American Economic Review* 61 (2), 112-123.

Williamson, O. E. (1979). Transaction Cost Economics: The Governance of Contractual Relations, *Journal of Law and Economics* 22 (2), 233-261, University of Chicago Press

Williamson, O. E. (1985). The Economic Institutions of Capitalism. New York: Free Press