Auer, Raphael A.; Monnet, Cyril; Shin, Hyun Song

Working Paper

Permissioned distributed ledgers and the governance of money

Discussion Papers, No. 21-01

Provided in Cooperation with:
Department of Economics, University of Bern

Suggested Citation: Auer, Raphael A.; Monnet, Cyril; Shin, Hyun Song (2021) : Permissioned distributed ledgers and the governance of money, Discussion Papers, No. 21-01, University of Bern, Department of Economics, Bern

This Version is available at:
https://hdl.handle.net/10419/242852

UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Permissioned Distributed Ledgers and the Governance of Money

Raphael Auer, Cyril Monnet, Hyun Song Shin

21-01

January, 2021

DISCUSSION PAPERS

Schanzeneckstrasse 1
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

Permissioned Distributed Ledgers and the Governance of Money*

Raphael Auer
BIS

Cyril Monnet
Gerzensee and University of Bern

Hyun Song Shin
BIS

January 26, 2021

We explore the economics and optimal design of "permissioned" distributed ledger technology (DLT) in a credit economy. Designated validators verify transactions and update the ledger at a cost that is derived from a supermajority voting rule, thus giving rise to a public good provision game. Without giving proper incentives to validators, however, their records cannot be trusted because they cannot commit to verifying trades and they can accept bribes to incorrectly validate histories. Both frictions challenge the integrity of the ledger on which credit transactions rely. In this context, we examine the conditions under which the process of permissioned validation supports decentralized exchange as an equilibrium, and analyze the optimal design of the trade and validation mechanisms. We solve for the optimal fees, number of validators, supermajority threshold and transaction size. A stronger consensus mechanism requires higher rents be paid to validators. Our results suggest that a centralized ledger is likely to be superior, unless weaknesses in the rule of law and contract enforcement necessitate a decentralized ledger.

JEL Codes: C72, C73, D4, E42, G2, L86.

Keywords: digital currencies, money, distributed ledger, blockchain, coordination game, global game, consensus, market design.

*We thank an anonymous referee of the BIS Working Paper Series, Aldar Chan, Francesca Carapella, Jon Frost, Rod Garratt, Piero Gottardi, Hans Gersbach, Dirk Niepelt, Jean-Charles Rochet, participants of the 2020 annual meeting of the Central Bank Research Association, the 2020 Summer Workshop on Money and Payments, the 2020 ETH-Zurich Workshop on Future Money and seminar participants at the BIS and the University of Zurich for useful comments. The bulk of this research was conducted when Cyril Monnet was a visiting economist at the BIS. The views presented in this paper are those of the authors and not necessarily those of the Bank for International Settlements.

1 Introduction

Money is a social convention. People accept money in payment in the expectation that others will do so in the future. Within an equilibrium with monetary exchange, holding money is a record of goods sold or services rendered in exchange for money. In this sense, money is a record-keeping device.

Kocherlakota (1998) and Kocherlakota and Wallace (1998) showed that money as a record-keeping device was capable of doing the job of a complete ledger of all past transactions in the economy. The motto is that "money is memory". In this spirit, monetary theorists have regarded money as performing the role of a publicly available and freely accessible record-keeping device. While the concept of money as memory has been well-known in theoretical circles, the advance of cryptography and digital technology has opened the possibility of taking the idea of a complete digital ledger more literally, and building a monetary system around such a ledger. However, with a public ledger the issues that loom large are who should have the authority to update the ledger and how. This is all the more so given the incentive problems that arise to misrepresent ownership of funds. Under traditional account-based money overseen by an intermediary, for instance a bank, this authority is delegated to intermediaries. The bank updates the ledger by debiting the account of the payer and crediting the account of the receiver.

However, in monetary systems without a central intermediary, the ledger must be updated by other means, such as decentralized ledger technology (DLT), as exemplified by Bitcoin. DLT is a record keeping device in the spirit of Kocherlakota's analysis of money as memory. In permissioned DLT, a known network of validators can update the distributed ledger via the agreement of a supermajority of the validators. The aim is to achieve agreement on the state of some data in a network of nodes that constantly exchange the underlying data file, the ledger.[1] Applications of permissioned DLT are being explored for securities settlement

[1]For example, in a trade finance application, the state might be the delivery status of a set of shipments and respective payments, and the network includes anyone authorized to access the system. Users can transact in this system, for example by initiating a new shipping order. The purchase instruction needs to be written into the ledger, which means that validators need to read the past ledger and verify that the new transaction is indeed genuine. Once this has happened, they vote for the transaction to be included in the ledger. The set of rules by which agreement is achieved is called the consensus algorithm. This is a computer

systems, trade finance solutions, "stablecoins", and central bank digital currencies.[2]

In this paper, we present what we believe is a first economic analysis of permissioned DLT in a monetary economy. We examine the economic opportunities and challenges of this technology focusing on the strategic elements underlying its optimal design.[3] We abstract from the details of the computing or cryptographic implementation and focus on the incentives of the validators that are needed to sustain mutually beneficial exchange as an equilibrium of a game. The validation protocol is constrained by two technological limitations. First, there is no technical way of forcing a validator to verify and sign any given transaction. Second, nothing can technically prevent a validator from validating multiple ledgers with conflicting histories. We examine theoretically how the optimal validation protocol deals with these constraints and derive the optimal number of validators, their compensation, and the optimal voting rule. In turn we can determine how the optimal validation protocol impacts the level of trade in the economy.

A model of credit Our model has three building blocks. The first modeling block consists of an intertemporal model of exchange involving credit. Our economy has two types of infinitely lived agent, early and late producers. In each period, an early producer is randomly matched with a late producer and the pair engages in two subsequent production stages. In the early stage, the early producer produces goods for the late producer. In the late stage, the late producer should reciprocate and produce some goods for the early producer. We impose two main frictions on producers. First, there is private information: late producers can be faulty and cannot produce but early producers cannot tell the difference between faulty and other producers. Second, late producers cannot commit to reciprocating. Therefore there is no trade unless a record-keeping device – the memory of our economy – tracks the actions of late producers and, in particular, whether the late producer has ever defaulted in the past.

protocol that specifies the conditions under which a ledger is considered as valid. Importantly, the consensus mechanism also guides how to choose between multiple versions of a ledger if conflicts should emerge.

[2]See e.g. Townsend, 2020, Baudet et al (2020), Arner et al (2020), Auer et al (2020), and for how to use blockchain technology to settle assets, Chiu and Koeppl (2019). A recent survey indicates that 86% of central banks are conducting research or development in the area of CBDC (Boar and Wehrli, 2021).

[3]Of course, a decentralized payment system too will rely on central bank money for underpinning of a stable value and an elastic supply, see Frost et al (2020). The decentralization concerns how this money circulates in the economy once it is issued by the central bank.

Just as in Kocherlakota (1998), rather than users owning and paying with monetary tokens, the ledger's memory of the production history suffices to allow for trustless exchange: it is well known (see e.g. Rocheteau and Nosal, 2017) that there is an equilibrium with trade when the trading history of a late producer is publicly and freely observable and automatically updates itself according to the behavior of the late producer. In turn, since the ledger's memory is the essential value underpinning of the economy, ensuring its integrity is the quintessential design issue to be solved.

The second modeling block endogenizes the process of updating the history of trades, that is the validation of records on the ledger. We assume that a number of agents known as "validators" are in charge of reading and updating the ledger of trade histories. For each trade involving a late and early producer, some validators have to verify the history of the late producer and communicate the result to the early producer. The history is understood to be "good" (that is, without default) whenever a supermajority of validators say it is so. We assume that verifying histories has a known common cost, while the cost of communicating the result is idiosyncratic, reflecting, e.g. the possibilities of operational failures for some validators. Validators privately learn their cost of communicating histories, which consists of a common component and an idiosyncratic one.

Since verification and communication are both costly activities, validators must be compensated for their efforts and they will expect a payment from the pair of early and late producers whose history they have to validate. Since validation requires a supermajority of validators, this structure gives rise to a global game that we analyze using the approach of Morris and Shin (1998, 2003). As we explained above, validators cannot be trusted, because (1) they cannot commit to verifying histories, so while messages sent by validators are observable, their checking is a costly non-observable action which raises a moral hazard problem and (2) they can accept side payments to record a false entry.

Our third and final building block is the analysis of the optimal design of the trade and validation mechanisms. The optimal mechanism chooses the number of validators, the supermajority threshold, the compensation of validators, as well as the trade allocation that maximize the gains from trade subject to incentive compatibility conditions.

Results We first show that reaching decentralized consensus – a unique equilibrium – among validators entails paying higher rents to validators in order to sustain a higher level of decentralized consensus as an equilibrium outcome. Given their private costs, validators play a game that has attributes of a public good provision game – the public good is provided if and only if a supermajority provides it – which we proceed to solve using global game methods (see Carlson and van Damme, 1993 and Morris and Shin, 1998, 2003). We show that there is a unique, dominance solvable equilibrium if and only if the rewards to being a validator are higher than a threshold that increases in the average cost of validation.

We show that a decentralized consensus fails to be sustained when the rents accruing to validators falls below the threshold. This is so, even when validation would be a possible equilibrium in a complete information game. The reason is that the uncertainty surrounding the fundamental communication cost can reverberate throughout the validation process: validators may choose to abstain from validating a trade when, given their private cost, they believe that other validators will also abstain. However, there is an equilibrium where validators validate a trade as long as their private costs is below some level. Driving the idiosyncratic component of the cost to zero, we find the validation process "works" whenever the fundamental communication cost is below that level. In turn, this gives the probability that the validation process will be successful. That success probability falls with the supermajority threshold, but increases with the payments validators obtain. Therefore a higher supermajority requires higher payments to validators in order to guarantee the same success probability. In other words, reaching a higher level of consensus among validators requires higher rents to be paid to validators.

Our second result is that, despite strategic uncertainty, a trading equilibrium can arise where the ledger truthfully reflects the history of trades. We characterize the optimal trade size, supermajority, and number of validators; where optimality is defined as the surplus from the trade net of the validation costs. Naturally, the optimal solution is constrained by the incentives of late producers and validators. We find that intertemporal incentives are key to characterizing the optimal solution. When intertemporal incentives are strong in the sense that the present values of future rewards are high, validators can be trusted as they would

have much to lose from accepting a bribe. In this case, the supermajority threshold should be high, there should be few validators and they should earn high rents. The trading allocation is high in this case. In the limit, the supermajority tends to unanimity with few (measure zero) validators, trade is efficient, but the rent to the few validators is arbitrarily large. On the contrary, when intertemporal incentives are weak, the future high rents are not enough to deter validators from accepting a bribe. Since the bribe size falls with the number of validators, there should be many. However the supermajority should be relatively low so that consensus is weak, and as a result the validators' rents should be relatively small. In this case, the trade size will also be small.

Our findings therefore suggest a number of initial conclusions. While it is costly to duplicate verification and communication across many validators, we find conditions under which many validators are better than one. Therefore, to use Aymanns et al's (2020) terminology, we find conditions under which a (trading) platform should be vertically disintegrated – a group of agents should handle the interaction between users – rather than vertically integrated, when a single intermediary has the monopoly over managing the interaction of the platform users. Also there are economies of scope in trading and validation: achieving good governance and honest record-keeping is made easier by having validators who also participate in the market themselves and thus have an intrinsic interest in keeping it going smoothly. This also implies that validators should be selected from the market participants.

Our results on the supermajority are naturally dependent on the communication cost being stochastic and unknown. When the communication cost is common knowledge, unanimity is optimal and, to reduce the incentive to bribe validators, they should be sufficiently many. Hence it is typically sub-optimal to have a central validator whenever they have to satisfy incentive constraints and the communication cost is common knowledge.

However, while it is optimal to have many validators absent any other frictions, it gives rise to a free-rider problem in solving information frictions similar to the one in the seminal paper of Grossman and Stiglitz (1976). Since verifying a label is costly, we show that, under some conditions, validators have an incentive to skip verification while still communicating a good label, which jeopardizes the whole legitimacy of the ledger. To resolve this free-rider problem

and maintain the integrity of the ledger, we show that (absent a unanimity rule where all validators are pivotal) the allocation of validators should be dependent on which label they communicate to the ledger and how their communication compares with the supermajority. When the label they communicate differs from the supermajority (which is observable and verifiable), validators should be excluded from trading and validating in the future. In this context, we derive a folk's theorem of sort for validators; as validators become more patient, the free-rider problem has no bite and any allocation satisfying the validators' participation constraint can be implemented in our strategic set-up.

Literature A sizable literature analyzes the incentives of miners in Bitcoin and similar cryptocurrencies to follow the proof-of work protocol.[4] Kroll et al (2013) and Prat and Walters (2020) examine free entry and the dynamics of the "mining" market,[5] while Easley et al (2019) and Hubermann et al (2021) examine the economics of the transaction market. Budish (2018) and Chiu and Koeppl (2019), and show that ensuring the finality of transactions in Bitcoin is very costly as so-called "majority" or "history reversion" attacks are inherently profitable, while Auer (2019) examines whether the transaction market can generate sufficient miner income to ensure the finality.[6] Leshno and Strack (2020) present a generalization of such analysis, demonstrating that no other anonymous and proof-of-work-based cyrptocurrencies can improve upon the performance of Bitcoin. This can serves as a benchmark for the analysis at hand to compare permissioned and permissionless market designs. Further to this, even in the absence of incentives to reverse history, sunspot equilibria

[4]Also the variant with betting on the truth instead of costly computation, i.e. proof-of-stake, is attracting increased attention (see Abadi and Brunnermeier 2018, Saleh 2021, and Fanti et al 2020). However, proof-of-stake can also be attacked via so called "long-run attacks" (see Deirmentzoglou et al 2018 for a survey). Therefore, proof-of-stake implicitly assumes the existence of some overarching social coordination to (see Buterin, 2014).

[5]See also Cong et al. (2019) for an analysis of the concentration of mining and efficiency.

[6]Such attacks are outlined in Nakamoto (2008). In these, the majority of computing power is used to undo a transaction in the blockchain by creating an alternative transaction history that does not contain the transaction. It is noteworthy that other attacks on cryptocurrencies are possible, including the possibility of "selfish" mining analyzed in Eyal and Sirer (2014). Gervais et al. (2016) present a dynamic analysis of the costs and benefits of various attack vectors. Garatt and van Oordt (2020) examine the role of fixed cost of capital formation for the security of Proof-of-Work Based Cryptocurrencies, Böhme et al. (2015) and Schilling and Uhlig (2019) present discussions of broader economic implications and governance issues, respectively.

can arise in proof-of work based blockchains (Biais et al 2019).[7]

The literature on validator incentives and design of permissioned versions of distributed ledgers is sparser. Most closely related to our analysis is Amoussou-Guénou et al (2019), who first modeled the interaction between validators as a game entailing non-observable effort to check transactions and costly voting. They also analyzed that game in terms of moral hazard and public good provision. Relative to their analysis, our contribution is to link the ledger validation game to monetary exchange, establish the uniqueness of the equilibrium via a global game approach, and characterize the optimal mechanism design, in particular in terms of number of validators, size of transactions, and optimal supermajority voting threshold. In our work, all validators are profit-seeking, and the issue at heart is how the market can be designed so that profit-seeking validators actually verify the ledger and validate only correct histories.[8] The focus on dealing with free-riding and coordination relates to several classical strands of papers on the coordination with many actors. Reminiscent of Grossman and Stiglitz (1976), free riding can prevail in the case of multiple validators. Consistent with Biais et al (2019) and Amoussou-Guenou et al (2019) we also derive a folk theorem. Last, we note that we do not model monopoly power on the market for transactions.

Our paper also has ramifications in the banking literature, starting with Diamond (1984) or Williamson (1986, 1987) where banks are modeled as a way to save on monitoring costs. Another approach, pioneered by Leland and Pyle (1977) and developed by Boyd and Prescott (1986) models banks as information-sharing coalitions. Gu et al. (2016) show that higher rents can discipline intermediaries, while Huang (2019) uses that model to study the optimal number of intermediaries when they have an incentive to divert deposits. A related analysis that study the optimal composition of the money stock between inside and outside money can be found in Monnet (2006), Cavalcanti and Wallace (1999a, b), and Wallace (2005). Global games techniques have also been introduced in the banking literature to study the probability a bank run occurs by Rochet and Vives (2004) and Goldstein and Pauszner

[7]See Carlstens et al (2016) for a related argument based on simulations and Pagnotta (2021) for an examination of multiple equilibria in the presence of a feedback loop between blockchain security and cryptocurrency valuation.

[8]Note that Amoussou-Guénou et al (2019) do not examine history reversion attacks; rather, byzantine attackers are assumed to attempt bringing the system to a halt for exogenous reasons.

(2005).

In game theory, the literature on incentives with public and private monitoring is large and it is beyond the scope of this paper to summarize it all (see Kandori, 2001 for an early survey). However, we would like to mention Rahman (2012), who studies a problem of private monitoring where the observation of the monitor(s) is not verifiable. He shows that sending a false positive to test the "attention" of the monitor can be optimal, or in his own words "the principal allocates private information to provide incentives." When considering the free-rider problem we also find that there must be enough faulty producers to induce the correct behavior from validators. However, our planner does not know if the match involves a faulty producer when it does, while the principal in Rahman (2012) knows when the false positive is sent.

The paper describes in Section 2 the main features of permissioned DLT that we think any model of permissioned DLT should capture. Section 3 then lays down the basic set-up and characterizes benchmark allocations absent a record-keeping device and a freely accessible one. Section 4 defines incentive feasible allocation with DLT, and characterizes the optimal allocation including the optimal number of validators. We analyze the free-rider problem in Section 5.

2 The technology and economics of permissioned distributed ledgers

How does the model streamlined above and developed below relate to the actual underlying technologies? In all DLT applications, the aim is to achieve agreement on the state of some data in a network of nodes that constantly exchange the underlying data file, the ledger. For example, in the case of the permissionless cryptocurrency Bitcoin, the state is the current ownership of bitcoins and the network includes anyone who downloads the blockchain.[9] In a permissioned application in the field of trade finance, the state might be the delivery status

[9] A blockchain is a ledger is composed of files (called blocks, each containing a number of transactions) that are sequentially chained (thus resulting in a chain of blocks).

of a set of shipments and respective payments, and the network includes anyone authorized to access the system.

The set of rules by which agreement is achieved is called the consensus algorithm (or mechanism). This is a computer protocol, specifying the conditions under which a ledger is considered as valid and, importantly, it also guides how to choose between multiple versions of a ledger should conflicts ever emerge. The consensus mechanism sustains decentralized consensus whenever it creates incentives for everyone to follow the rules of its protocol.

DLT applications generally feature two kinds of actors, "users" and "validators". On the one side, users want to transact, whereas validators are those who can include (either alone or as a group, depending on the precise consensus algorithm) new transactions in ledger updates. We note that there is nothing that bars a validator from also being a user. In fact, as we show below, it can be optimal to draw validators from the set of users because, as users of the DLT, validators are interested in preserving its integrity.

The validation process involves both the verification of transactions and the casting of a vote to the ledger. Verification means to read the ledger and analyze whether it is consistent with the current state (ie the version of the ledger that resulted from a previous instance of running the consensus mechanism). Voting is the process by which one or many validator(s) agree to write an authorized transaction into the ledger to update it. Below we will model these two steps of the validation processes in detail.

To exemplify these actors and actions, consider a simple transfer of a token from user L to E in a DLT application such as in Corda's R3, Hyperledger Fabric, or Quorum, or alternatively in the Libra Blockchain of the Libra Association's global stablecoin proposal. [10] User L initiates the transaction via its digital signature. The set of validators V (called "notaries" in Corda) can verify that L is indeed the current owner of the token and that the transfer

[10]To be precise we focus on the details of the validation steps in Corda. In Hyperledger Fabric and Quorum, the relevant logical steps of the transaction are the same, although the technical implementation differs somewhat. Technically, whereas Corda like Bitcoin follows a UTXO model where verification of a transaction involves tracing a token all the way back to its origin, in Hyperledger Fabric and Quorum – as in the cryptocurrency Ethereum – a transaction resembles the account-based system where transaction include an update on the balance of accounts. Tokens are not native units of Hyperledger Fabric or Quorum, but can be constructed or emulated on them.

authorization is authentic. The verification process in Corda includes verifying whether L is authorized to transfer the token and verifying whether L has already spent the token. For the former verification step, each validator needs to check the digital signature of the token owner L, as well as, the entire chain of transactions involving this token. If all signatures from the point of issuance to that of the transfer to L are valid, the token is authentic. For the latter verification step, each validator needs to check the ledger to see if L still owns the token, i.e. for the absence of an older transaction transferring the token to somebody else.

Each validator can then broadcast its vote to include the transaction in the next update of the ledger. In this process, the key issue concerns the conditions under which the votes can be considered as being sufficient for E to assume that the transaction has indeed occurred, i.e. that the transfer is and will continue to be included in the consensus version of the ledger. This is achieved via a voting-based consensus protocol, for example Practical Byzantine Fault Tolerance.[11] It typically involves the following steps: after verification, validators communicate their votes to each other. Once a share of votes exceeds the pre-set threshold – for example, 50% plus 1 – of votes have been communicated, the transaction is understood to be valid under the consensus rule. After this has happened E can, with some but not full certainty, assume that the transaction is final.

The key underlying economic problems – which we analyze below – concern not the technical implementation, but the underlying incentives of the validators that are needed to sustain honest exchange. The first aspect is that there is no technical way to force a validator to sign any given transaction. Validators need to be incentivized to actively verify and vote on transactions. The second aspect is that nothing can technically prevent a validator from voting on multiple ledgers with conflicting histories. The latter opens up the possibility of a "history-reversion attack", in which a first consensus ledger emerges that includes a given transaction, but later a second consensus ledger emerges that does not include it.

A history-reversion attack is one in which the payment for a merchandise first enters the

[11] Indeed, many of today's sandbox-style applications of Corda R3 feature only one notary, and once this notary has signed a transaction into the ledger, it becomes final. As we show below, this is the exact equivalent of today's model of centralised exchange. Looking ahead, Corda (2020) argues that while Byzantine Fault Tolerance-type of consensus mechanism is most commonly used, the platform is pluggable with different consensus mechanisms.

consensus ledger to be sub-sequentially excluded in a future consensus ledger. In the above-stated transfer from L to E, such a fraudulent attack would take place in the following way. L and E agree on an exchange of merchandise from E to L in exchange for the payment of a token from L to E. L initiates the token transfer and after E observes that the transfer has become part of the consensus ledger, it releases the merchandise. Once this has happened, L bribes a sufficient number of validators to vote on a conflicting ledger that does not contain said transfer, thus effectively undoing the transaction.[12]

Such history-reversion attacks are becoming increasingly frequent for permissionless cryptocurrency (see i.e. Shanaevet el. 2020), and, going forward, a mechanism must be found to ensure that it does not happen in permissioned DLT. For this mechanism, importantly, although validator action is not enforceable, it is observable to the participants of the system, and the protocol can thus reward and punish certain validator actions. On the one side, it can reward active participation (also to reimburse for the cost incurred during verification). On the other side, it can punish malfeasance via exclusion from the set of future validators. If validators earn fee income in excess of operation costs, supporting a history-reversion attack thus has the cost of forgoing the net present value of validator profits.

In line with the above considerations, we next model the two-step verification process and the underlying incentives by assuming that 1) some users are faulty, and 2) users are not trustworthy. To model step 1) we will abstract from the idea of users owning a token. As in Kocherlakota (1998), we will rather consider users having a label related to their history of trade. In our model, a simple label turns out to be a set of statistics sufficient to summarize the history of trades. A good label stands for the fact that L is authorized to trade, very much as a token was acquired in a legitimate way. So in our model, validators will verify the entire history of trade as summarized by the label, and communicate the result by sending a signal to the ledger. Our model takes step 2) into consideration by letting L choose to "double spend" by obtaining goods in exchange for the promise to produce later (akin to the

[12]As multiple versions of the ledger can exist, a consensus protocol also must specify a rule to distinguish among them. This tends to be the one version with the most votes. If the consensus rule was 50% plus 1, 50% plus 2 votes are needed for a successful attack. Axiomatically, a history reversion attack thus can only succeed if at least some validators – the exact number depending on the supermajority rule – validate conflicting histories.

promise to deliver the coin) but then not delivering on that promise in due time.

3 The basic model

Our model builds on Gu et al (2013).[13] Time is discrete and infinite. $\beta \in (0,1)$ is the discount factor. Each period is divided in two distinct production/consumption stages, early and late. There is one good per stage, the "early good" and the "late good". Goods are non-storable across stage or across periods. There is a continuum of agents. Agents can be of three types, which are permanent. There is a measure one of early producers, a measure $1-f>0$ of late producers, and a measure f of faulty producers. Early producers cannot distinguish between late and faulty producers. Early producers can produce the early good that late and faulty producers like to consume. Late producers can produce the late good that early producers like to consume. Faulty producers do not produce. Therefore, there can only be gains from trade between early and late producers.

Preferences of early producers are represented by the following utility function[14]

$$U_e(x^e, y^e) = x^e - y^e$$

where x^e is the consumption of the late good by early producers, and y^e is the production of the early good by early producers. Preferences of productive late producers are represented by the following utility function

$$U_\ell(x^\ell, y^\ell) = u(x^\ell) - y^\ell$$

where x^ℓ is the consumption of the early good, and y^ℓ is the production of the late good. The function $u(.)$ is continuous, concave, increasing and $u(0)=0$. We assume there are gains from trade between early producers and productive late producers. That is, there is x

[13] Gu et al (2013) borrows methodological elements from Lagos and Wright (2005). See also Williamson and Wright (2011), and Lagos et al (2017).

[14] Linear utility function for one of the agent (here early producers) allows us to get clean comparative statics, as would do quasilinear utility functions like $x^e - v(y^e)$.

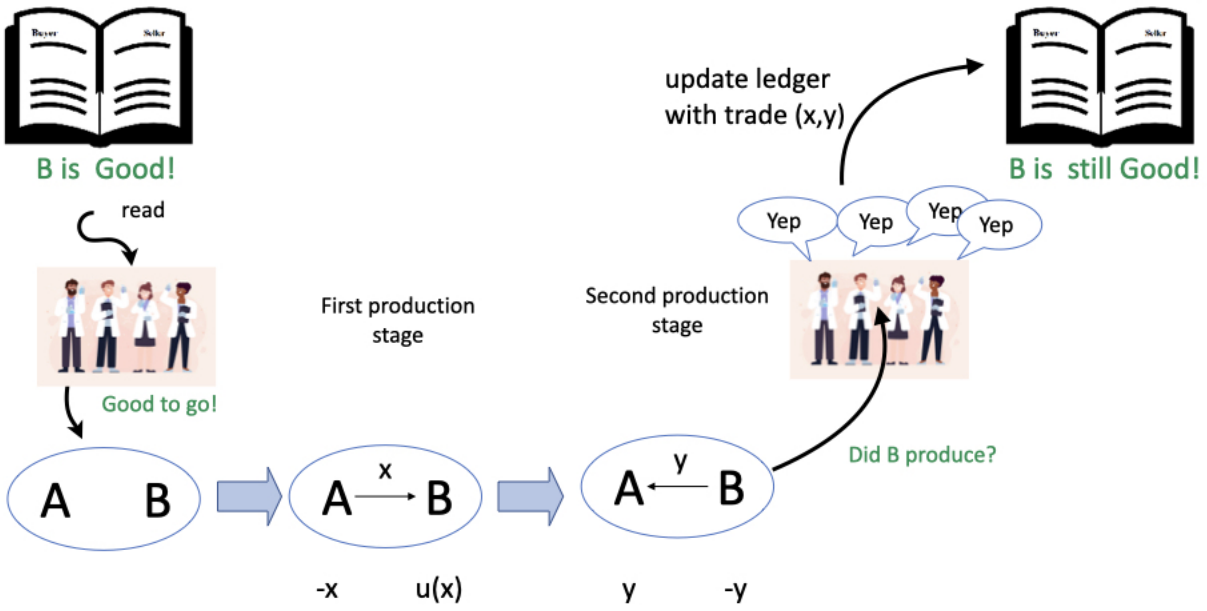

Figure 1: Timeline

such that $u(x) > x$. We denote by x^* the efficient allocation that solves $u'(x^*) = 1$. Finally, preferences of faulty producers are represented by

$$V_\ell(x^\ell) = \rho x^\ell$$

We assume $\rho < 1$ so there are no gains from trade between early producers and faulty producers.

Early and late producers meet pairwise at the start of the early production stage. The matching technology is such that nature selects a measure α of early producers, a measure $\alpha(1-f)$ of late producers and a measure αf of faulty producers. Nature also matches one of the early producer with either one of the late or faulty producers. All other producers remain unmatched for the period. Therefore, the probability of a match for any producer is α. The probability that a match involves a faulty producer is f and with complementary probability, the match involves a late producer. The match is maintained across both stages but it dissolves at the end of the later stage. Since being faulty is private information to

that producer, a faulty producer will behave as if he was productive. As a consequence, we can concentrate on allocations in matches that involve early and late producers only.

Feasibility and efficiency require that $x^e = y^\ell$ and $x^\ell = y^e$. Therefore, we can conveniently drop indices and use $x = x^\ell = y^e$ and $y = x^e = y^\ell$. Hence, an allocation is (x, y) where x denotes the production (consumption) of early (late) producers, and y denotes the production (consumption) of late (early) producers. In this paper we will concentrate on symmetric and stationary allocations. Figure 1 sketches the timeline of our economy.

We can now define a trading mechanism. We restrict trading mechanisms to be in the class of coordination games: In each match the two agents choose a pair $(x, y) \in \mathbb{R}^2_+$. If both choices coincide, the early producer produces x in the early stage and the late producer produces y in the late stage (faulty producers never produce). In this case we say that the allocation (x, y) can be implemented.

If agents can commit, the set of allocations (x, y) that can be implemented is solely defined by the participation constraints of early and late producers. Late producers participate if and only if $u(x) \geq y$, while early producers participate if and only if $(1 - f)y \geq x$ since they can only consume if the producer is not faulty. Therefore all allocations (x, y) such that $(1 - f)u(x) \geq x$ can be implemented. If there are too many faulty producers and f is too large, then the efficient allocation $x = y = x^*$ is not implementable.

In addition to hidden state, we will assume late producers only have a limited ability to commit.[15] Absent commitment and any record-keeping technology, it is routine to show that late producers will never repay early producers and the only implementable allocation is autarky $(x, y) = (0, 0)$.

A ledger technology

To discipline late producers who cannot commit and to tell late producers apart from faulty producers, it is necessary to record the history of trades of these producers in a ledger (see

[15] For the sake of symmetry we can also assume that early producers are unable to commit, but their incentive problem is straightforward: if they do not produce, the late producer will not produce either.

Kocherlakota, 1998). Such a record helps to discipline late producers by threatening the loss of future consumption in case they do not produce today. The ledger works as follows. For the pair of early and late producers in period t, the ledger records the result of the coordination game $(\tilde{x}_t, \tilde{y}_t)$ as well as the actions of both producers (x_t, y_t). In the coordination, game both producers *agree* that the early producer should produce $\tilde{x}_t$ for the late producer in the early stage and the late producer should produce $\tilde{y}_t$ for the early producer in the late stage. Their respective *actions* x_t and y_t might however differ from their *words* $\tilde{x}_t$ and $\tilde{y}_t$. The ledger always records the transaction truthfully and with no latency. Let $h_t(i)$ be the history for late producer $i \in [0,1]$ up to period t, so $h_t(i) = \left\{(\tilde{x}_0^i, \tilde{y}_0^i, x_0^i, y_0^i), ..., (\tilde{x}_{t-1}^i, \tilde{y}_{t-1}^i, x_{t-1}^i, y_{t-1}^i)\right\}$ where $\tilde{x}_s = x_s = \tilde{y}_s = y_s = 0$ whenever the late producer is not matched in period s. Let $\mathcal{B}$ be the set of history such that the late producer cheated at least once

$$\mathcal{B}_t(i) = \left\{h_t(i) : \exists s \leq t-1 \text{ with } y_s^i \neq \tilde{y}_s^i\right\}.$$

The history of actions translates into a label for late and faulty producers. They can be assigned a good (G) or a bad (B) label. We will also say that late producers can be in a good or bad standing. The mapping between history up to stage t and labels is a function $M : \mathcal{H}_t \to \{G, B\}$ where

$$M(h_t) = \begin{cases} G & \text{if } h_t \notin \mathcal{B}_t(i) \\ B & \text{if } h_t \in \mathcal{B}_t(i) \end{cases}$$

Notice that label B is an absorbing label. The ledger implies that a late producer can be in good (G) or bad (B) standing, while a faulty producer will always be in bad standing. When an agent has label B, early producers rationally expect the agent is a faulty producer or the agent has defaulted in the past, and therefore announces $(\tilde{x}, \tilde{y}) = (0, 0)$ in the coordination mechanism. Therefore an agent with label B will never consume or produce. In other words, if a late producer does not repay an early producer and is assigned label B, he will be in autarky forever.

Then, an allocation (x, y) is implementable if it is incentive feasible (IF). That is, the allocation satisfies participation constraints and late producers have the incentive to repay. These

two participation constraints are

$$\begin{aligned} u(x) - y &\geq 0, \\ y - x &\geq 0, \end{aligned}$$

while late producers will have the incentive to repay early producers if their "repayment constraint" holds,

$$-y + \beta\alpha \frac{u(x) - y}{1 - \beta} \geq 0.$$

If late producers do not produce, they are excluded from the economy, so the right-hand side of the constraint is zero. If they produce, they incur the production cost $-y$, then they are assigned a good label and can participate in trade in the future. The expected value of having a good label is equal to the lifetime gains from trade $(u(x) - y)/(1 - \beta)$ times their probability of trading α. Setting $y = x$, the set of IF allocations is summarized by x such that

$$\beta\alpha u(x) \geq (1 - \beta(1 - \alpha))x.$$

It should be clear that the efficient allocation x^* is implementable if β and α are large enough.

So far, our analysis has been routine because we have taken the functioning of the ledger as given. The objective of our paper is to endogenize the ledger updating process and explain the incentives problem arising from the mechanism used to update the ledger.

4 Permissioned Ledgers

We now endogenize the updating process of the ledger. We assume the ledger is managed by a measure V of "validators." Validators are agents entrusted with validating transactions and updating the ledger, but they are rational agents who need to be incentivized to verify transactions and update the ledger honestly. The resulting history on record should be trusted by all producers. We assume that validators are selected among the pool of late producers in period 0. As a consequence, each validator can also trade with early producers.

However, they cannot validate their own trades. Each period, nature randomly selects αV validators and assigns each of them to work on validating one and only one of the α match. Therefore a measure V of (different) validators works on validating each and every match.[16]

There are two validation stages: in the early stage and in the late stage. In the early stage, the validation process consists of validating the label of producers (late and faulty). To validate the label, validators have to 1) verify the label of producers, and 2) vote, by sending a message to the ledger to communicate the verified label to the ledger. Validator i sends message $m_i^1 \in \{\emptyset, 0, 1\}$, where $\emptyset$ means the validator does not send a message, 0 means the validator sends message B and 1 means the validator sends message G. A consensus is reached and is communicated to the early producer whenever more than a fraction $\tau \in [0, 1]$ of validators cast the same vote. So the label is understood to be G if $\int_{i=0}^{V} m_i^1 \mathbb{I}_{m_i^1 \neq \emptyset} di \geq \tau V$.

In the late stage, the validation process consists of confirming that late producers have produced according to plan, i.e. $\tilde{y} = y$.[17] To confirm production, validators have to 1) verify whether production took place according to plan, and 2) vote, by sending a message to the ledger to communicate the resulting label – if production took place according to plan and the late producers had label G validators will communicate G, but B otherwise. We denote the message of validator i in the late stage as $m_i^2 \in \{\emptyset, 0, 1\}$. Again, a consensus is reached on the new label which is recorded on the ledger whenever more than τV validators cast the same vote; for example, the recorded label is G if $\int_{i=0}^{V} m_i^2 \mathbb{I}_{m_i \neq \emptyset} di \geq \tau V$. We emphasize that τ is a choice variable when designing the consensus algorithm.[18]

Validators incur verification and communication costs. We assume in the early stage that validators incur a linear utility cost $c_v \geq 0$ to verify the label of a producer and an idiosyncratic linear utility cost $c_{s,i} \geq 0$ to send a message to the ledger (or to enter their information on the ledger). Since it is costly to send messages, we assume validators only send a message

[16] Since validators are selected from late producers, $V \leq 1 - f$. The analysis is easily extended when validators are also selected from early producers. It will become clear that a planner would select validators from the set of producers who have the most at stake, because it can use the future value of trades as an incentive device for validators, thus minimizing the rents validators obtain from being able to manage the ledger. Also, the random selection of validators simplifies the argument somewhat.

[17] In practice, this is when double-spending can happen.

[18] We assume the threshold τ is the same in the first and the second stage, but our analysis extends to cases where it differs across the stages.

when they want to communicate that the late producer's label is G.[19]

In order to model the possibility of computer glitches and operational failures, we assume that the private cost of communicating a label $c_{s,i}$ that takes the form

$$c_{s,i} = c_s + \mu_i,$$

where c_s is a common component to all validators, while μ_i is the idiosyncratic element for validator i. The idiosyncratic element μ_i is uniformly distributed over the interval $[-\varepsilon, \varepsilon]$, where ε is a small positive number. For any two distinct validators $i \neq j$, μ_i is independent of μ_j. Finally, we suppose that c_s itself has a uniform ex ante distribution over $[\underline{c}_s, \bar{c}_s]$. Validators learn their cost ahead of the verification game and so ahead of verifying the label.

As Morris and Shin (2003) show, the key to the analysis is the characterization of the strategic uncertainty faced by players. Even if the idiosyncratic component is small relative to the other payoff parameters in the frame, the relative ranking of the costs injects strategic uncertainty in the coordination game. Even if remote, the possibility of computer glitches will imply that validation should not rely on unanimous agreement when there are many validators. In the sequel, the reader should think of $\varepsilon \to 0$.

Late producers can refuse to become validators, so that validators must make a positive expected profit from the validation process. We assume that sending a message is verifiable. So validators who sent a good message are entitled to z^s units of the good in stage $s = 1, 2$ whenever $\int_{i=0}^{V} m_i^s \mathbb{I}_{m_i \neq \emptyset} di \geq \tau V$ and indeed the late producer has label G. Validators receive nothing if they do not send a message to the ledger (or if they send a bad signal) and they cannot work in the late stage if they have not sent a message to the ledger in the early stage. We assume that validators receive these transfers at the end of stage 2 – once the dust settles – and validators value these transfers according to $v(z^1 + z^2)$ – and, to simplify matters further, we will assume $v(\cdot)$ is a linear function.[20] Hence when a fraction $w^s \geq \tau$

[19] We assume validators do not incur costs in the second stage. It is straightforward to extend the model to also analyze this case.

[20] A linear utility function allows us to abstract from possible insurance mechanism among validators. Also, we could assume that validators only consume the late good. In this case, the late producer produce $y + Vz^2$, and the early producer consumes $y - Vz^1$ and each validator collects $z^1 + z^2$. Since early producers

of validators have sent a good message in stage s, early producers in the early stage have to pay $w^1 V z^1$ for each of the $w^1 V$ validators to get z^1, and symmetrically, in stage 2 a late producer has to pay $w^2 V z^2$ for each of the $w^2 V$ validators to get z^2.

Late producers can bribe validators to send a false message: A late producer who has label G when starting the period may get away with not repaying the early producer and may keep its label by making a side payment to τV validators in the late stage (after consuming in the first stage).[21] Validators who accept such a bribe are caught with probability π.

4.1 Payoffs and incentive feasible allocations

Given a threshold τ and a measure of validators V assigned to validate each match, a stationary allocation is a list (x, y, z^1, z^2). An allocation is incentive feasible if it is feasible, it satisfies the incentive constraints of early and late consumers, given τ, the label of late producers is correctly communicated to the ledger, and validators have no incentive to tamper with the record of labels.

Given a stationary incentive feasible allocation (x, y, z^1, z^2), U_i is the expected discounted lifetime utility of late producer i satisfying

$$(1-\beta)U_i = \mathbb{E}_i \left\{ \mathbb{I}_{w_i > \tau} \alpha \left[u(x) - y - w_i V z^2 \right] \right\},$$

where $\mathbb{E}_i$ is the expectation operator of late producer i over w_i, the share of working validators. The late producer only trades if more than τV validators work and validate the trade. In this case the late producer gains $u(x) - y$ from trading but pays $w_i V z^2$ to the working validators. The expected discounted lifetime utility of a validator i with private

are risk-neutral, it turns out this is equivalent to validators consuming both the early and the late goods. Also, at the cost of simplicity, we could assume that the utility of validators is $u(x + z^1 + z^2)$. Also, we could assume that only a share τ of validators receive a reward in stage $s = 1, 2$ since only this number is necessary to reach an agreement.

[21] A late producer who misbehaved some time in the past and enters a period with a label B can bribe validators so as to obtain label G to get to consume. However, validators will not agree to a bribe in the early stage because (1) there is a possibility that the briber is a faulty late producer, but also (2) they would have to trust the late producer to pay the bribe in the late stage. However, if validators accept the bribe, the late producer has no incentive to make good on it.

communication cost $c_{s,i}$ is $U_V(c_{s,i})$ and satisfies

$$\begin{aligned} U_V(c_{s,i}) &= \mathbb{E}_i\left\{\mathbb{I}_{w_i>\tau}\alpha\left[u(x)-y-w_iVz^2\right] \mid c_{s,i}\right\} && (1)\\ &\quad +\alpha\max\left\{0;\mathbb{E}_i\left[-c_v+(1-f)\left(\mathbb{I}_{w>\tau}\left(z^1+z^2\right)-c_{s,i}\right) \mid c_{s,i}\right]\right\}+\beta\mathbb{E}U_V && (2)\end{aligned}$$

where $\mathbb{E}(\cdot \mid c_{s,i})$ is the expectation operator over the common communication cost c_s of validator i conditional on receiving signal $c_{s,i}$. Since validators are selected from the set of late producers, they obtain the expected payoff of late producers. However, they have more information concerning the fundamental communication cost c_s, which they use to compute the probability that the trade be validated. In addition they also get the expected payoff from validating a trade: Given their signal $c_{s,i}$, validators can choose to work or not. If they do not work, they get nothing. If they work, validators incur the verification cost c_v. If they verify and the producer has a good label (which happens with probability $1-f$), they incur the communication cost $c_{s,i}$, and get the reward z^1+z^2, but only when the trade is validated and the indicator function $\mathbb{I}_{w>\tau}=1$. Otherwise they do not get a reward.

Participation constraints. An allocation (x,y,z^1,z^2) satisfies the participation constraints of validators, early and late producers whenever

$$\begin{aligned} \mathbb{E}\left[u(x)-y-wVz^2\right] &\geq 0 && (3)\\ \mathbb{E}\left[y-x-wVz^1\right] &\geq 0 && (4)\\ \mathbb{E}\left[-c_v+(1-f)\left(\mathbb{I}_{w>\tau}\left(z^1+z^2\right)-c_s\right)\right] &\geq 0 && (5)\end{aligned}$$

The expectation operator in (3) and (4) is again on the share of working validators w. Since late producers can refuse to become validators, the last constraint requires that validators expect to make a positive expected profit from the validation process and their expectation operator is over w and c_s.

Repayment constraints. Using a ledger, late producers who do not produce the announced amount y are detected by validators and thus assigned a label B, and are permanently excluded from future trades.[22] Therefore, given the share of working validators is $w \geq \tau$,

[22]If bad payers are also validators, we assume they also loose the right to validate future trades.

the repayment constraint of late producers and validators is respectively

$$-\left(y + wVz^2\right) + \beta U \geq 0. \tag{6}$$

$$-\left(y + wVz^2\right) + \beta \mathbb{E}U_V \geq 0. \tag{7}$$

No bribe. If a validator accepts a bribe, we assume it is caught with probability $\pi \in [0, 1]$. In this case it looses its right to validate future transactions and to consume as a late producer. A validator prefers recording the truth to a false record when the late producer offers its $\bar{z}$ iff

$$\underbrace{z^1}_{\text{producer does not pay } z^2} + \beta \mathbb{E}U_V \geq \underbrace{z^1 + \bar{z}}_{\text{producer bribes } \bar{z}} + (1-\pi)\beta \mathbb{E}U_V$$

When a share w of validators are working on a match, the late producer in this match is willing to pay at most a total of $y + wVz^2$ to get away with production. Given that the ledger requires the agreement of at least τV validators to validate a transaction, the cheating late producer will pay $\bar{z} = (y + wVz^2)/(\tau V)$ to τV validators. Using $\bar{z} = (y + wVz^2)/(\tau V)$, a validator rejects the bribe whenever[23]

$$\pi \beta \mathbb{E}U_V \geq \frac{1}{\tau V}\left(y + wVz^2\right). \tag{8}$$

It remains to find a condition such that, given a threshold τ, the allocation allows for a correct record of the ledger.

Validation threshold. Suppose there is a positive measure of validators in charge of verifying one match. Then the decision of an arbitrary validator to work or to shirk depends on the subjective probability this validator assigns to other validators working or shirking. Therefore, there are many possible equilibria depending on the original beliefs of validators. For example, if one validator believes that other validators will shirk then his optimal strategy is also to shirk. In other words, the uncertainty about the cost of other validators of communicating a label to the ledger may reverberate throughout the system and may

[23] We assume validators act in unison: they either all accept the bribe, or they all reject it.

jeopardize the validation process. We assume the continuation payoff of validators is independent of the current validation results – after all, with operational failures, it is difficult to distinguish whether a validator shirked or experienced a failure.[24] Here we also assume that if validators do not work, they do not send any messages. We relax these assumptions in Section 5, where we also specify the voting game that validators play in detail, and we state conditions under which validators do not send a message without working. Now suppose the validation process requires unanimity. As soon as validators expect a positive measure to abstain from validating, they will also abstain even though they may have received a low signal on the communication cost. As a consequence, we show that validation only occurs correctly when the validation rule is based on supermajority unless payments to validators are arbitrarily large. The higher the supermajority threshold, the more rents should accrue to validators in order to guarantee the integrity of the ledger. We show the following result in the Appendix,

Proposition 1. *Given τ, in the limit as $\varepsilon \to 0$, there is a unique dominance-solvable equilibrium where validators work if and only if the allocation (z^1, z^2) satisfies*

$$z^1 + z^2 \geq \frac{c_s + c_v/(1-f)}{1-\tau}. \tag{9}$$

The proof follows two steps. In the first step, we characterize equilibria when, for some common threshold c_s^*, validators employ switching strategies whereby they work if their cost $c_{s,i}$ below the threshold c_s^* and they shirk otherwise. A validator receiving a cost at threshold level c_s^* is indifferent between working and shirking.

It is well-known from the global game literature (e.g. Morris and Shin, 2003) that, for the marginal player whose cost is exactly equal to the threshold value c_s^*, the density over the share of working agents is uniform over $[0, 1]$. Hence, the validator assigns a probability q to the event that a fraction q of the V validators will work. Since this validator is indifferent

[24] However, see Green and Porter (1991) and Monnet and Quintin (forthcoming).

between working or not, c_s^* solves

$$-c_v + (1-f)\left[\int_\tau^1 (z_1 + z_2 - c_s^*)\, g(\tau \mid c_s^*) d\tau + \int_0^\tau (-c_s^*)\, g(\tau \mid \tilde{c}_s^*) d\tau\right] \;=\; 0$$

When this agent does not work, it gets nothing. If it works, it will find that the label of the producer is B with probability f. In this case, it does not cast a vote to the ledger and does not get any compensation. With probability $1-f$ the label is G and it will cast the vote to the ledger by paying his cost c_s^*. However it only gets compensated if more than τV validators are working. Since its subjective beliefs of the share of working validators is uniform, i.e. $g(\tau \mid c_s^*) = 1$, c_s^* solves

$$-c_v + (1-f)\left[(1-\tau)\,(z_1 + z_2 - c_s^*) + \tau\,(-c_s^*)\right] = 0.$$

When the noise vanishes, all individual costs necessarily converge to the common value c_s. Therefore, when $c_s \leq c_s^*$, all validators will work to validate a trade and the ledger will record labels correctly, while when $c_s > c_s^*$ none of them will. The allocation $z_1 + z_2$ logically affects the threshold value c_s^* defined by (9) holding with equality: By increasing the validator's rents $z_1 + z_2$, the validation protocol can ensure that validation happens for higher levels of the communication cost.

Hence, the first step of the proof establishes that the validation game has a unique equilibrium in switching strategies. The second step of the proof establishes that this unique equilibrium in switching strategies is also the only strategy profile of the players that survives the iterated deletion of strictly dominated strategies. In other words, the game is dominance solvable. Morris and Shin (2003) shows that a sufficient condition for dominance solvability in our setting is that the payoffs satisfy strategic complementarity – that is, the payoff to working is weakly increasing in the proportion of other validators who work. Since this condition is satisfied in our game, we can apply the global game results in Morris and Shin (2003) to conclude that our game is dominance solvable.

As a corollary, notice that the payment to validators, as measured by $z^1 + z^2$ is positively linked to the supermajority level τ when there are validation costs. Therefore, the ledger

can only retain integrity when unanimity is required if payments to validators are arbitrarily large. The reason is that the "seed of doubts" reverberates throughout the validation game, so that validators can never be certain that *all* validators will work unless payments are arbitrarily large. With finite payments only a supermajority can sustain the ledger's integrity. Finally, given τ, the probability that the trade will go through when c_s is uniformly distributed is the probability that

$$c_s \leq (1-\tau)(z_1+z_2) - \frac{c_v}{1-f} \equiv c_s^*.$$

If the ledger is required to allow *all* legitimate trades involving a producer with label G will *always* go through, then $z_1 + z_2$ should be set to

$$z_1 + z_2 = \frac{1}{1-\tau}\left(\bar{c}_s + \frac{c_v}{1-f}\right) \tag{10}$$

where $\bar{c}_s$ is the maximum possible communication cost. In this case, and given τ, validators will always work. Below, we assume this is the case. In the Appendix, we relax this assumption and we analyze the optimal validation protocol for any thresholds c_s^* defined by (9) holding with equality and we let the designer choose what c_s^* should be. Also, we derive sufficient conditions under which $c_s^* = \bar{c}_s$ is optimal.

We can now define incentive feasible allocations.

Definition 1. Given τ and V, an incentive feasible allocation is a list (x, y, z^1, z^2) that satisfies (3)-(8) and (10).

In the sequel we solve for the optimal design of the validation protocol, and how it affects incentive feasible allocations.[25] We simplify matters further by assuming that the distribution for the communication cost c_s converges to one that gives all the mass to just one point $\bar{c}_s$.

[25] In a different context, see e.g. Gersbach et al (2020) for an analysis of the optimal committee size under majority rule as well as the references therein.

4.2 Degenerate distribution of communication costs

From now we consider the limiting case where $\varepsilon \to 0$, (10) holds, and the distribution of the communication costs converges to a degenerate distribution at $\bar{c}_s$. So unless they are bribed, all validators always verify the label, always verify that production took place according to plans, and always cast the right vote to the ledger. Therefore the share of working validators w converges to one and we can write (3)-(9) with $w_i \to 1$. In this section, we consider the incentives of a late producer to make side payments to validators so that they record a false trade, and we analyze the incentives of validators to accept that bribe.

We can further simplify the set of IF allocations by setting the participation constraint of early producers at equality.

Since the payment to validators should be minimized, (10) binds so validators take home $z^1 + z^2 = Z(\tau)$, where

$$Z(\tau) \equiv \frac{\bar{c}_s + c_v/(1-f)}{1-\tau}$$

When $z^1 + z^2 = Z(\tau)$, the participation constraint of validators (5) is always satisfied. Let $R(\tau)$ be the expected rent of validators,

$$\begin{aligned} R(\tau) &\equiv (1-f)\left[Z(\tau) - (\bar{c}_s + c_v)\right] - f c_v \\ &= \frac{\tau(1-f)\bar{c}_s + c_v}{1-\tau} - c_v \end{aligned}$$

We can set the participation constraint for early producers (4) at equality and replace $y = x + Vz^1$. Since validators earn a rent, $U_V \geq U$ and (7) is satisfied whenever (6) is. Then the set of IF allocations is characterized by

$$\frac{\beta\alpha}{1-\beta}\left[u(x) - (x + VZ(\tau))\right] \geq (x + VZ(\tau)) \tag{11}$$

$$\pi\frac{\beta\alpha}{1-\beta}\left[u(x) - (x + VZ(\tau)) + R(\tau)\right] \geq \frac{1}{\tau V}(x + VZ(\tau)) \tag{12}$$

Constraint (11) says that late producers are better off retaining their good label by repaying x to early producers and $VZ(\tau)$ to validators, than getting a bad label and lose the expected

lifetime discounted payoff from trading net of the payment to validators. Constraint (12) compares the payoff a validator would obtain from accepting the maximum bribe a late producer would offer $(x + VZ(\tau))/\tau V$, to the expected loss of accepting such a bribe, given they are caught with probability π. In addition to losing the expected lifetime discounted payoff from trading net of compensating validators, validators would also use the validation rents they earn $R(\tau)$. Notice that since early producers have linear utility, their production $x+z^1$ is compensated by late producers through their production $y = x+z^1$. In this context, it does not matter who bears the cost to compensate validators $Z(\tau)$ and only the total cost $Z(\tau)$ matters for the allocation.[26]

It will be convenient to define the default factor as

$$\delta \equiv \frac{\pi\beta\alpha}{1-\beta}.$$

The higher δ is, the lesser are the incentives of validators to accept a bribe, either because they would lose a lot of trading opportunities (as captured by a large α) or because they would very likely be caught cheating (as captured by a high π) or because they care a lot about the future (as captured by a high β). We now focus on the number of validators V. The more validators there are, the higher the total production must be to make the payment to validators – this is the term $-VZ(\tau)$ in the first inequality. However, given τ more validators also means that the bribe per validator diminishes, which relaxes the constraint of validators – this is the term $\frac{1}{\tau V}(x + VZ) = x/(\tau V) + Z(\tau)/\tau$ on the RHS of the validator's incentive constraint. The optimal number of validators will trade-off both effects. When the measure of validators is large, notice that if $\pi \to 1$, the repayment constraint of late producers becomes the binding one so that validators never accept a bribe. This is intuitive: when V is large, no validator gets a very large bribe, but if the mechanism almost surely observes when they accept a bribe, they almost certainly loose the expected lifetime payoff from trading. Then, only the incentive to repay of the late producer matters. However, note that if (11) is binding while (12) is slack, V can be reduced up to the point where (12) binds.

[26]This is obviously not robust to a generalization of the preferences of early producers. In the Appendix, we solve the case when early producers have preferences for consumption represented by a strictly concave utility function.

We summarize the discussion above in the following result.

Lemma 1. *The distribution of the cost to remunerate validators $Z(\tau)$ among early or late producers is indeterminate. If π and V are large enough, validators will never accept a bribe.*

4.3 Optimal design

Since the probability of a productive match is $1-f$ in each period, the objective function of a planner is the sum of the early and late producers' utility when they trade, and the rent of validators from operating the ledger for a measure α of trades,

$$\alpha(1-f)\left[u(x)-y-Vz^2+\left(y-x-Vz^1\right)\right]+\alpha VR(\tau),$$

or simplifying and replacing $z^1+z^2=Z(\tau)$,

$$\alpha(1-f)\left\{u(x)-x-V\left(Z(\tau)-\frac{R(\tau)}{1-f}\right)\right\}$$

Replacing the expression for $R(\tau)$ it is natural to find that the net payments to validators $(1-f)Z(\tau)-R(\tau)$ is a positive constant, equal to the average verification cost $(1-f)\bar{c}_s+c_v$. Therefore, the size of the payment to validators $Z(\tau)$ only matters for incentives but has no impact on the objective function: it is a mere transfer between late producers. A planner chooses the trading size x, the number of validators and the threshold τ to solve

$$\alpha(1-f)\max_{x,V\geq 0,\tau\in[0,1]}\left\{u(x)-x-V\left(\bar{c}_s+c_v+\frac{f}{1-f}c_v\right)\right\}$$

subject to (11) and (12). Hence the objective function is decreasing in V and as a result we can show in the Appendix:

Lemma 2. *The incentive constraint of validators (12) always binds while the incentive constraint of producers (11) never binds.*

Replacing the expression for the validation rent in (12), and re-arranging, we obtain:

$$\delta\left[u(x)-x\right] \geq \frac{1}{\tau}\left[\frac{x}{V}+Z(\tau)\right]-\delta\left[\left(1-f-V\right)Z(\tau)-(1-f)\bar{c}_s-c_v\right] \tag{13}$$

Since the objective function is independent of τ, the planner will choose τ to minimize the right hand side of (13). Increasing τ reduces the maximum bribe size per validator. However, it necessarily increases the payment to validators $Z(\tau)$ to ensure the validation process works. When intertemporal incentives are strong, so that $1 \leq \delta\left(1-f-V\right)$, this second effect also comes into effect to reduce the RHS of (13): validators have much to lose by accepting a bribe. In this case, it is optimal to set $\tau = 1$, even if $Z(\tau) \to \infty$. Alternatively, when intertemporal incentives are weak, $1 > \delta\left(1-f-V\right)$, the optimal $\hat{\tau}$ trades-off the lower bribe size with the increase payment to validator and it solves

$$\frac{1-\hat{\tau}}{\hat{\tau}} = \sqrt{\frac{\left[1-\delta\left(1-f-V\right)\right]\left(\bar{c}_s+\frac{c_v}{1-f}\right)}{\left[\frac{x}{V}+\bar{c}_s+\frac{c_v}{1-f}\right]}} \tag{14}$$

It is easy to verify that when $1 > \delta\left(1-f-V\right)$, the second order condition for a minimum is satisfied. The optimal threshold $\hat{\tau}$ is decreasing in V but increasing in x. $\hat{\tau}$ is also increasing in π, β or α. When π, β or α increase, the net rent $R(\tau)-VZ(\tau) > 0$ of validators becomes more important in determining the incentives of validators relative to the bribe size $(x+VZ)/\tau V$, either because they have a higher chance of losing them – when π increases – or because they have a higher lifetime discounted value – when β or α increases. Since the net rent is increasing in τ, the planner will choose to increase $\hat{\tau}$ following a rise in π, β or α. Also, notice that raising τ implies higher rents to validators. Therefore, unlike in traditional models of limited commitment, higher trustworthiness as captured by higher values for π, β, or α imply more rents to validators.

It is useful to look at the first order conditions in detail. Denoting the Lagrange multiplier on the validators' IC constraint by λ , the first order conditions with respect to x and V

respectively are

$$[u'(x)-1](1+\delta\lambda)-\lambda\frac{1}{\hat{\tau}V} = 0 \quad (15)$$

$$\left[\frac{R(\hat{\tau})}{1-f}-Z(\hat{\tau})\right]+\lambda\frac{x}{\hat{\tau}V^2}-\delta\lambda Z(\hat{\tau})-\lambda_{1-f} \leq 0 \quad (16)$$

(15) says that a marginal increase in x will increase the gains from trade by $u'(x)-1$. Also the lifetime discounted payoff of validators will increase by $\delta[u'(x)-1]$, which relaxes their IC constraint, but the bribe they can be offered increases by $1/V$, which tightens their IC constraint. Importantly, equation (15) makes clear that $V>0$ unless $\lambda=0$. (16) says that a marginal increase in the number of validators will increase the total payoff by the rent they receive $R(\hat{\tau})$, but it will also increase the total expected validation cost born by producers $Z(\hat{\tau})$. This also reduces the lifetime discounted payoff of validators and so tightens their IC constraint, but it implies the bribe they can each receive will decrease by x/V^2, which relaxes their constraint. (16) holds with equality whenever $V>0$ and $V=1-f$ whenever $\lambda_{1-f}>0$. The optimal x and V trade off these effects. We can now state our main result.

Proposition 2. *The constrained optimal solution $(\hat{x},\hat{V},\hat{\tau})$ solves (13) at equality and (14)–(16) and is characterized by four regions:*

1. [central system] *If $\delta>1/(1-f)$ then the allocation is arbitrarily close to the one with a single validator, $\hat{V}\to 0$, $\hat{\tau}\to 1$ and $\hat{x}\to x^*$, but it requires arbitrarily large payments, i.e. $Z(\hat{\tau})\to\infty$, while $\lim_{\hat{\tau}\to 1}V(\hat{\tau})Z(\hat{\tau})=\frac{x^*}{\delta(1-f)-1}$.*

2. [partially distributed system] *If $\bar{\delta}<\delta\leq 1/(1-f)$, the constrained optimal number of validators is $\hat{V}>0$, and only a supermajority $\hat{\tau}<1$ is optimal. Each validator receives a finite payment $Z(\hat{\tau})<\infty$. The constrained optimal allocation is $\hat{x}<x^*$.*

3. [fully distributed system] *If $\delta_0<\delta\leq\bar{\delta}$, all late producers are validators $\hat{V}=1-f$, and $\hat{\tau}=\left(1+\sqrt{\frac{(1-f)c_s+c_v}{x+(1-f)c_s+c_v}}\right)^{-1}$. The constrained optimal allocation is $\hat{x}<x^*$.*

4. [no trade] *If $\delta\leq\delta_0$, there is no validation protocol that can decentralize trade.*

Proof. We first consider the solution when $\hat{\tau}\to 1$. Then

$$\frac{R(\hat{\tau})}{1-f}-Z(\hat{\tau}) = -(\bar{c}_s+c_v)-\frac{f}{1-f}c_v$$

while $Z(\hat{\tau}) \to \infty$. We now show that λ and V both converge to zero. From (16) we obtain either $\lambda \to 0$ and/or $V \to 0$. It is clear that if $V > 0$ and $\lambda > 0$ then the LHS of (16) is necessarily negative as $\tau \to 1$ so that $V = 0$, a contradiction. Now rewrite (16) as

$$\frac{(\bar{c}_s + c_v) + \frac{f}{1-f} c_v}{\left[\frac{x}{\hat{\tau} V^2} - \delta Z(\hat{\tau})\right]} = \lambda$$

Since $\lambda \geq 0$ we obtain $x \geq \tau \delta V^2 Z(\hat{\tau})$. Since x is bounded from above but $Z(\hat{\tau}) \to \infty$, we must have $V \to 0$ as $\hat{\tau} \to 1$. Further, since (11) never binds, it must be that $VZ(\tau)$ converges to a positive constant.[27] Therefore,

$$\frac{\lambda}{V} = \frac{(\bar{c}_s + c_v) + \frac{f}{1-f} c_v}{\left[\frac{x}{\hat{\tau}} - \delta V^2 Z(\hat{\tau})\right]} V \overset{\hat{\tau} \to 1}{\to} 0.$$

Then (15) implies $x \to x^*$. When the solution for $\hat{V}$ is interior, we can simplify (15) and (16) to obtain

$$\frac{\left[Z(\hat{\tau}) - \frac{R(\hat{\tau})}{1-f}\right]}{\left[\frac{x}{\hat{\tau} V^2} - \delta Z(\hat{\tau})\right]} = \lambda$$

and

$$u'(x) - 1 = \frac{\lambda}{1 + \delta\lambda} \frac{1}{\hat{\tau} V}$$

[27] From (13) holding at equality,

$$\begin{aligned} \delta \left[u(x) - x\right] &= \left(\frac{1}{\tau V}\right) \left[x + VZ(\tau)\right] - \delta \left[R(\tau) - VZ(\tau)\right] \\ \delta V \left[u(x) - x\right] &= \left(\frac{1}{\tau}\right) x + VZ(\tau)/\tau - \delta \left[1 - f - V\right] VZ(\tau) - V\delta \left[(1-f)\bar{c}_s + c_v\right] \end{aligned}$$

and taking the limit as $\tau \to 1$ (which implies $V(\tau) \to 0$), since $u(x) - x$ is bounded,

$$\begin{aligned} 0 &= x^* + \lim_{\tau \to 1} V(\tau) Z(\tau) - \delta \left[1 - f - V(\tau)\right] V(\tau) Z(\tau) \\ 0 &= x^* + \left[1 - \delta (1-f)\right] \lim_{\tau \to 1} V(\tau) Z(\tau) + \delta \lim_{\tau \to 1} V(\tau)^2 Z(\tau) \end{aligned}$$

Suppose $V(\tau)^2 Z(\tau)$ would converge to a strictly positive constant. Then $\lim_{\tau \to 1} V(\tau) Z(\tau)$ would converge to $+\infty$ which would violate the equality above. Hence,

$$\lim_{\tau \to 1} V(\tau) Z(\tau) = \frac{x^*}{\delta(1-f) - 1}.$$

or

$$u'(x) \;=\; 1 + \frac{VZ(\hat{\tau})}{x} - \frac{R(\hat{\tau})V}{x(1-f)\left(1+\delta\lambda\right)}$$

The right hand side of the above equation is always higher than 1, so that generically $x < x^*$. The constrained optimal solution is $(\hat{x}, \hat{V})$ that solves (15), (16) and (13) holds with equality. $\hat{\tau}$ is given by (14).

Consider the case where $V = 1 - f$. Using (14), it is easy to check that $\tau = \left(1 + \sqrt{\frac{C}{x+C}}\right)^{-1}$. The rest of the proof for this case is in the Appendix as it involves some tedious algebra. But in this region of the parameter, x is increasing δ. So as δ goes to zero, (13) can never be satisfied (even setting $x = 0$), and for any validation protocol (τ, V) the ledger would not reflect the label truthfully so that trade cannot take place. □

Proposition 2 states that centralized validation by a single validator is not necessarily optimal. Only when validators are sufficiently trustworthy and there are relatively few faulty producers is it optimal to have a single validators. However, in our context, a single validator requires an arbitrarily large payment, as otherwise the single validator can be easily bribed. In reality, feasible payments may be bounded and in such a case, a single validator will never be optimal. Still, the single validator case is a useful benchmark to illustrate the forces driving the solution towards a single validator.

When π is so small that it is hard to detect wrong behavior by validators, it will be optimal to require many validators. This is even more surprising as we abstracted from any industrial organization issues but focused only on incentives. To understand why this is the case, it is useful to look at the comparative statics. If the relative risk aversion is greater than one, $V'(x) \leq 0$: increasing the allocation x makes it more expensive for validators to be caught cheating, which implies the planner can optimally reduce the number of validators. Then we obtain the following comparative statics (details of the calculation are in the Appendix),

- x and τ are (weakly) increasing with β, π, and α but decreasing with $c_v + (1-f)\bar{c}_s$.
- V is (weakly) decreasing with β, π, and α but increasing with $c_v + (1-f)\bar{c}_s$.

It is intuitive that the optimal level of V decreases, while that of x increases with β, π and α: Everything else constant, when validators are more patient, or when they trade more often, they give up more when they are caught accepting a bribe to enter a false record on the ledger. Since validators are less likely to accept a bribe, the amount of the bribe per validator can increase. This is achieved by both increasing the size of each trade x and reducing the number of validators.

It may be surprising that the optimal number of validators increases with the total cost of validation. However, the intuition is as above: when the cost to validate each trade increases, the size of each trade x drops. This lowers the overall lifetime surplus of validators, who, as a result, have less to lose from accepting a bribe. It is therefore easier to bribe them. To limit this effect, it is optimal to increase the number of validators so that (given x) each of them can only be offered a smaller bribe, which they will then reject. The number of validators acts as the variable of adjustment to limit their incentives to accept bribe. The higher the number of validators , the less they can each receive given x, and the less likely they can be bribed.

For the validation threshold $\hat{\tau}$, as the total validation cost increases validators are more likely to believe that fewer other validators will work. Therefore, they are less likely to work whenever $\hat{\tau}$ is high. To maintain the legitimacy of the ledger, the threshold $\hat{\tau}$ should fall in order to induce all validators to work. Similarly, we have explained why the direct effect of increasing α, β, π is to increase $\hat{\tau}$. Since $\hat{\tau}$ is also decreasing in V and increasing in x, both direct and indirect effects are reinforcing each other so that the total effect of increasing α, β, π is also to increase $\hat{\tau}$.

5 The validation game

In this section, we specify the details of the validation game and we analyze a free-rider problem inherent to the validation protocol: validators have the incentives to abstain from verifying a label, but still send the message that the label of the producer is G. The severity of this free rider problem could undermine the existence of an equilibrium with trade, as the

ledger would loose integrity.[28]

We keep some of the features of the optimal allocation. In particular, as payments are indeterminate, we simplify notation by using $z = z^1 + z^2$. Also, recall that as $\varepsilon \to 0$ and absent the free-rider problem, all validators should be expected to work as long as the communication cost is lower than some threshold (that we set at $\bar{c}_s$).

5.1 The free-rider problem

In this section, we describe the validation game that validators play in detail. In the first stage, a strategy for validator i consists of a verification strategy, $\nu_i \in [0,1]$, a voting strategy $\sigma_i \in [0,1]$ which is the probability to send a message and the validator i choice of message $m_i \in \{\emptyset, 0, 1\}$ to send. We call *shirkers* those validators who do not verify labels, and we call *workers* those validators who do. Define $\mathbf{m} = (m_1, m_2, ..., m_V)$ and $\mathcal{I}(\mathbf{m}) = 1$ if $\int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di > \tau V$ and $\mathcal{I}(\mathbf{m}) = 0$ otherwise.[29]

The public history at the end of period t consists of the public history h_t at the end of period $t-1$, as well as the result of the validation process $\mathcal{I}(\mathbf{m})$ and the production of the late producer, that we can summarize with the label of the producer $\ell \in \{G, B\}$.[30] We focus on strategies for validators that depend only on the public history and the information acquired during the current period. The equilibrium concept is Bayesian perfection: strategies are Nash equilibrium given the information validators have and validators are Bayesian so that they update their belief using Bayes' rule.

Recall that validators are dealing with legitimate late producers with probability $1-f$ and the probability that the producer is faulty is f. Also, recall that validators who do not send a message are not entitled to a payment. Given the allocation (x, y, z) is incentive feasible – so that a late producer who is found to have label G will produce for the early producer –

[28] See also Amoussou-Guenou, et al (2019).

[29] Again we assume that $m_{i,k} = \emptyset$ counts as $m_{i,k} = 0$.

[30] Since we concentrate on incentive feasible allocations (x, y, z) the producer's label is a sufficient statistics for the outcome in a match because a late producer with label G will produce so that the early producer will also produce, while a late producer with label B is not expected to produce so that the early producer will not produce.

the expected payoff of a working validator from validating the transaction is

$$-c_v + (1-f)\left\{\begin{array}{c}\sigma_i(G)\left(-c_s + E_i\left[\mathcal{I}(\mathbf{m})z + \beta U_V(h_t,(\mathcal{I}(\mathbf{m}),G)) \mid m_i = 1\right]\right)\\ +(1-\sigma_i(G))E_i\left[\mathcal{I}(\mathbf{m})\times 0 + \beta U_V(h_t,(\mathcal{I}(\mathbf{m}),G)) \mid m_i = \emptyset\right]\end{array}\right\}$$
$$+f\left\{\begin{array}{c}\sigma_i(B)\left(-c_s + E_i\left[\mathcal{I}(\mathbf{m})\times 0 + \beta U_V(h_t,(\mathcal{I}(\mathbf{m}),B)) \mid m_i = 1\right]\right)\\ +\left(1-\sigma_i(B)\right)E_i\left[\mathcal{I}(\mathbf{m})\times 0 + \beta U_V(h_t,(\mathcal{I}(\mathbf{m}),B)) \mid m_i = \emptyset\right]\end{array}\right\} \quad (17)$$

where $U_V(h_t,(\mathcal{I}(\mathbf{m}),\ell))$ is the continuation payoff of the validator given the new history $(h_t,(\mathcal{I}(\mathbf{m}),\ell))$ and $\ell \in \{G,B\}$. $U_V(\cdot) = U_V$ whenever the continuation payoff does not depend on the validator's actions. Also, we assume that if the majority agrees that the producer has label G, then validation and communication happens in the second stage.

We now explain the different elements in (17). In the early stage, the working validator incurs cost c_v to verify the label. If the label is G (which happens with probability $1-f$) the validator sends message $m_i = 1$ with probability $\sigma_i(G)$ and nothing otherwise (again, we anticipate that not sending messages $m_i = \emptyset$ is better than sending $m_i = 0$, since it communicates the same information at a lower cost). If the index $\mathcal{I}(\mathbf{m}) = 1$,[31] the match is validated and trade can take place. Then validators who sent a message get z from the late producer, and they verify production takes place in stage 2 and communicate the result to the ledger. If $\mathcal{I}(\mathbf{m}) = 0$, the transaction is not validated and working validators get nothing. E_i is the expectation of validator i over the index function $\mathcal{I}(\mathbf{m})$ given the validator's information summarized by message m_i. With probability f, the working validator learns the producer's label is B. Then with probability $\sigma_i(B)$ the validator sends message $m_i = 1$ but he expects to receive zero, even if (at least) τ validators send $m = 1$ because he knows the buyer has a bad label and will not produce. With probability $1-\sigma_i(B)$, validator i sends no message (or message 0), and does not expect any payments. In any case, the working validator knows production will not take place in stage 2 and so he does not verify or communicate anything in stage 2. Notice that in this section, the expected future payoff of validators $U_V(h_t,(\mathcal{I}(\mathbf{m}),\ell))$ where $\ell \in \{G,B\}$ only depend on public history and so can vary depending on the outcome of the validation process. It should be obvious that $\sigma_i(B) = 0$: a worker will not send

[31] $\mathcal{I}(\mathbf{m}) = 1$ if at least $\tau - 1$ other validators send $m = 1$ if $m_i = 1$, and at least τ other validators send $m = 1$ if $m_i = \emptyset$.

message $m_i = 1$ for a producer with label B.

The expected payoff of a shirker is

$$\bar{\sigma}_i \left\{ \begin{array}{c} (1-f)\left(-c_s^1 + E_i\left[\mathcal{I}(\mathbf{m})z + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), G)) \mid \bar{m}_i = 1\right]\right) \\ +f\left(-c_s^1 + E_i\left[\mathcal{I}(\mathbf{m}) \times 0 + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), B)) \mid \bar{m}_i = 1\right]\right) \end{array} \right\}$$
$$+ (1-\bar{\sigma}_i) \left\{ \begin{array}{c} (1-f)E_i\left[\mathcal{I}(\mathbf{m}) \times 0 + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), G)) \mid \bar{m}_i = \emptyset\right] \\ +fE_i\left[\mathcal{I}(\mathbf{m}) \times 0 + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), B)) \mid \bar{m}_i = \emptyset\right] \end{array} \right\}$$

A shirker does not know the buyer's label before sending her message $\bar{m}_i$. Again, given index $\mathcal{I}(\mathbf{m})$, the future payoff of validators is the same for all validators irrespective of their action. Note that a shirker only shirks in period 1, here. Since it observes the index $\mathcal{I}(\mathbf{m})$, she can learn the label of the producer and she can verify production and send a message relative to that production in period 2 only if the label is G.[32]

Now suppose $\tau < 1$. We can show that in equilibrium, not all validators will be working/cooperating.

Lemma 3. *Suppose validation does not require unanimity $\tau < 1$. There is an equilibrium where **all** validators work whenever*

$$f \geq \frac{c_v}{c_s}.$$

Proof. Suppose all validators are working and send message $m = 1$ if the label is G and do not send a message otherwise. Since $\tau < V$, any validator i is not pivotal, because changing the value of one message will not change the overall index value. Then the value of working and sending $m = 1$ if the label is G and $m = \emptyset$ otherwise is

$$-c_v + (1-f)\left(-c_s + z + \beta U_V(h_t, (1, G))\right) + f\beta U_V(h_t, (0, B)) \tag{18}$$

[32]It does not matter here as the verification and communication costs are zero in the second stage. We showed that this also holds when those costs are positive.

while the expected value of shirking is

$$\bar{\sigma}_i \left\{(1-f)\left(-c_s + z + \beta U_V(h_t,(1,G))\right) + f\left(-c_s + \beta U_V(h_t,(0,G))\right)\right\}$$
$$+ (1-\bar{\sigma}_i)\left\{(1-f)\beta U_V(h_t,(1,G)) + f\beta U_V(h_t,(0,B))\right\}$$

So a shirker sends $m = 1$ whenever the expected payment is greater than the cost of always sending a message:

$$(1-f)z \geq c_s.$$

Using (9) at equality to replace for z, a shirker sends $m = 1$ whenever

$$\tau \geq f - \frac{c_v}{c_s}.$$

So when $\tau < f - \frac{c_v}{c_s}$, shirkers prefer to send no message and they never get a payment. So in this case, working always give a higher payoff to validators than shirking (and not sending a message). If $\tau \geq f - \frac{c_v}{c_s}$, shirkers are better off sending a message. Then working gives a higher payoff than shirking (and sending a message) when $f \geq \frac{c_v}{c_s}$. □

Stated slightly differently, Lemma 3 says that there is a free-rider problem whenever $f < c_v/c_s$. This is intuitive: when $fc_s < c_v$, free-riders who expect at least τ validators to work save the verification cost c_v but incur the cost of sending a message c_s when they should not send it (when the producer is faulty). As a corollary, we deduce that incentives to free-ride are high whenever $c_s \to 0$, because the cost of sending a message when one should not is negligible.

Notice that the only punishment that free-rider incurs is the cost of sending a message when the producer is faulty in which case they will not receive a payment. We now look at other forms of punishments. First, the worst punishment when payoffs can only depend on public history, is that the system shuts down if, collectively, validators make a mistake. This means that $U_V(h_t,(1,B)) = 0$. A late producer with label B will not produce and so the system will detect that the validation process was flawed. However, a late producer with label G who received the wrong validation will not produce (because the early producer will not

produce) and so will not be distinguished from a late producer with label B. In this case the system cannot detect the flawed validation. So we must have $U_V(h_t,(0,G)) = U_V(h_t,(0,B))$. We define a *uniform* mechanism as one that gives validators the same continuation payoff to "observationally equivalent" outcome and when the validation process gave a correct outcome, that is

$$U_V(h_t,(0,G)) = U_V(h_t,(0,B)) = U_V(h_t,(1,G)) \equiv U_{Vt}.$$

Following the steps in the proof of Lemma 3, we can conclude that uniform mechanisms do not relax the free-rider problem.

5.2 Individual mechanisms

We now define an individual mechanism as one where both the current payoff and the continuation value depend on the publicly observable action of validators. To be precise, we consider that the ledger assigns label B to a validator who is caught sending a message that differ from the "supermajority" τ of validators. As a result, this validator loses the ability to validate but also the opportunity to trade in the future. Such mechanisms specify individual continuation values $U_V(h_t,(\mathcal{I}(\mathbf{m}),\ell);m)$ as a function of the result of the validation process $\mathcal{I}(\mathbf{m})$, whether the producer produced or not $\ell \in \{G,B\}$ and the validators' message m. A validator goes against the majority whenever $\mathcal{I}(\mathbf{m}) \neq m$. In this case, the worse punishment is the level of utility the validator would obtain in permanent autarky. So we set

$$U_V(h_t,(\mathcal{I}(\mathbf{m}),\ell);m) = \begin{cases} 0 & \text{if } \mathcal{I}(\mathbf{m}) \neq m, \\ U_{Vt} & \text{otherwise.} \end{cases}$$

Then we show

Lemma 4. *Using an individual mechanism and $\tau < 1$, there is an equilibrium where all validators work whenever*

$$f \geq \frac{c_v}{c_s + \beta U_{Vt}}$$

When validators work, they lose c_v, but they send the right signal. When they don't work,

either they prefer to never send a signal, or they send a signal. When (uninformed) shirkers prefer not to send a signal, they often get it wrong when there are many good producers. In this case, validators prefer working than shirking whenever their loss c_v is less than the expected net loss of not sending a signal when they should $(1-f)(z-c_s+\beta U_{Vt})$. But if (uninformed) shirkers prefer to send a signal, they will get it wrong with probability f in which case they lose $c_s^1+\beta U_{Vt}$. So they prefer to work whenever the expected loss of sending a wrong signal $f(c_s+\beta U_{Vt})$ is higher than the verification cost.

Notice that validators working can now be an equilibrium even if $c_s = 0$. So we have the following Folk's theorem

Lemma 5. *[**"Folk" Theorem**] Let $\beta \to 1$. Using an individual mechanism and $\tau < 1$, there is an equilibrium where all validators work whenever the late producer's participation constraint (11) is satisfied,*

$$u(x) > x + VZ(\tau)$$

Proof. The existence of the equilibrium requires

$$f \geq \frac{c_v}{c_s + \beta U_{Vt}}$$

In equilibrium, $\beta \to 1$ implies that $\beta U_V \to \infty$ as long as $u(x) > x + VZ(\tau)$. Then $\frac{c_v}{c_s+\beta U_{Vt}} \to 0$. Therefore, validators work whenever $f \geq 0$. So validators always work as long as $u(x) > x + VZ(\tau)$. □

6 Conclusion

In this paper, we have presented an economic analysis of permissioned decentralized ledger technology in an economy where money is essential. To our knowledge, our analysis is the first economic analysis of permissioned DLT in such a context. It links a ledger validation game to monetary exchange, establishes the uniqueness of the equilibrium via a global game approach, and characterizes the optimal mechanism design, examining the optimal supermajority voting rule, number of validators, and size of transactions.

We believe our analysis is a timely one, as permissioned DLT is rapidly becoming an industry standard for digital currencies and in other applications. In particular, our results can shed light on the burgeoning literature on central bank digital currency insofar as it gives conditions under which a central authority should manage the ledger of transactions.[33] The economic discussion of technology and the economics of central bank digital money has thus far centered on the balance sheet effects and related systemic implications.[34] Here, we focus not on balance sheets and the issue of how the value of a currency can be guaranteed (central backing is of the essence for a CBDC irrespective of our analysis), but on the governance of money when it is used in exchange as the record-keeping device of society.

Of course, we have made simplifying assumptions in order to better grasp the basic economics of memory. Future work should relax some of these. For instance, we have assumed that one individual can only have one account, so that the reputation of the individual and his/her account are intertwined. However, ledgers only record transactions for one account and it is usually difficult to trace the identity of the owner of the account. However, our analysis would extend directly to reputation on accounts rather than on individuals.[35]

Also and to simplify the analysis we have assumed that validators all agree to accept bribes in unison. It would be interesting to also study the cooperative games between validators in more detail. We have also taken as given that agents use a private permissioned ledger as they want to preserve their anonymity in trades. Tirole (2020) and Chiu and Koeppl (2020) make progress on this front. In order to better compare the different types of ledger, future work should also include the benefit from preserving anonymity. Also our mechanism design approach implies that we have ignored every industrial organization aspect of DLT, which might be significant if this technology were to be widely adopted in the future.

[33]See the stock-taking exercise of pursued technological designs in Auer et al (2020).

[34]See among others Andalfatto (2018), Brunnermeier and Niepelt (2019), Fernández-Villaverde et al (2020), Keister and Monnet (2020).

[35]For an analysis of why it is important to distinguish individuals from accounts, see Li and Wang (2019).

Appendix

A. Proof of Proposition 1

Proposition 3. *Given τ, in the limit as $\varepsilon \to 0$, there is a unique dominance-solvable equilibrium where validators work if and only if the allocation (z^1, z^2) satisfies*

$$1 - \tau \geq \frac{c_s + c_v/(1-f)}{z^1 + z^2} \tag{19}$$

Let $z = z^1 + z^2$. Assume that each validators receive private cost

$$c_{s,i} = \gamma_s + \varepsilon_i$$

where ε_i is uniformly distributed over $[-\varepsilon, \varepsilon]$. Given a validation threshold τ, the expected payoff of validator i is

$$-c_v + \begin{cases} (1-f)\left(-c_{s,i} + z + \beta U_V\right) + f\beta U_V & \text{if } \int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di > \tau V \\ (1-f)\left(-c_{s,i}\right) + \beta U_V & \text{otherwise} \end{cases}$$

The expected payoff of a validator is given by (18) when the measure of validators sending a message is higher than τV. We can rewrite the expected payoff as

$$\beta U_V + (1-f)z \times \begin{cases} -\frac{\tilde{c}_v^1}{1-f} - \tilde{c}_{s,i} + 1 & \text{if } \int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di > \tau V \\ -\frac{\tilde{c}_v^1}{1-f} - \tilde{c}_{s,i} & \text{otherwise} \end{cases} \tag{20}$$

where we normalized the cost by the validator's rent, as $\tilde{c}_{\cdot,i} = c_{\cdot,i}/z$. This normalized cost is necessarily lower than 1 and it is uniformly distributed since γ_s is uniformly distributed. Notice that if a validator expects $\int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di < \tau V$, this validator will not even verify the label in the first place. The structure of the above payoff is the same as the one in the public good game analyzed in Morris and Shin (2002) and their results extend almost directly. We repeat their argument here for completeness.

Let w be the random variable $\int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di/V$ measuring the fraction of working validators. This is a random variable because the communication strategy m_i of each validator i depends on their communication cost. Also this random variable belongs to the interval $[0, 1]$. The distribution of w is important because it gives the probability that a trader with label G is able to trade and compensate validators, and in fine whether it is worth it in expected terms for a validator to verify and communicate the label to the ledger. Let $g(w \mid \tilde{c}_{s,i})$ be the subjective density over w for a validator with private cost $\tilde{c}_{s,i}$ and total verification and communication cost $-\frac{\tilde{c}_v}{1-f} - \tilde{c}_{s,i}$. We conjecture that validators adopt a switching strategy whereby they work whenever their total cost is lower than some level $C^* \equiv \frac{\tilde{c}_v}{1-f} + \tilde{c}^*_{s,i}$. Since the normalized cost is uniformly distributed over the interval $\left[\frac{\gamma_s - \varepsilon}{z}, \frac{\gamma_s + \varepsilon}{z}\right]$, the total cost is also uniformly distributed over $\left[\frac{\frac{c^1_v}{1-f} + \gamma_s - \varepsilon}{z}, \frac{\frac{c^1_v}{1-f} + \gamma_s + \varepsilon}{z}\right]$, and validators working are those with $C_i < C^*$. Therefore, the measure of validators working is

$$w = \frac{\frac{\tilde{c}^1_v}{1-f} + \tilde{c}^*_s - \frac{\frac{c^1_v}{(1-f)} + \gamma_s - \varepsilon}{z}}{2\frac{\varepsilon}{z}} = \frac{c^*_s - (\gamma_s - \varepsilon)}{2\varepsilon}$$

So for some $q \in [0, 1]$, there is a value for the **common** communication cost $\gamma_s(q)$ such that $w = q$. This is

$$\gamma_s(q) = c^*_s + \varepsilon - 2\varepsilon q$$

Hence, $w < q$ iff $\gamma_s > \gamma_s(q)$. We now need to find the probability that $\gamma_s > \gamma_s(q)$. Considering the validator with total cost C^*, the posterior density over γ_s conditional on his communication cost being c^*_s is uniform over the interval $[c^*_s - \varepsilon, c^*_s + \varepsilon]$. Hence, the probability that $\gamma_s > \gamma_s(q)$ is

$$\frac{c^*_s + \varepsilon - \gamma_s(q)}{2\varepsilon} = \frac{c^*_s + \varepsilon - (c^*_s + \varepsilon - 2\varepsilon q)}{2\varepsilon} = q.$$

Therefore

$$G(w < q \mid c^*_s) = q,$$

so that by differentiation, for all w

$$g(w \mid c^*_s) = 1$$

and the density over w is uniform at the switching point $\tilde{c}^*_s$. Hence, the probability that the

validation process will fail is

$$G(\tau) = \int_0^{\tau} g(w \mid \tilde{c}_s^*) dw = \tau.$$

The validator with private cost c_s^* is indifferent between working and shirking. Therefore the switching point $\tilde{c}_s^*$ solves

$$\begin{aligned} -\frac{c_v}{(1-f)z} + \int_{\tau}^{1} (1 - \tilde{c}_s^*)\, g(\tau \mid \tilde{c}_s^*) d\tau + \int_0^{\tau} (-\tilde{c}_s^*)\, g(\tau \mid \tilde{c}_s^*) d\tau &= 0 \\ 1 - G(\tau \mid \tilde{c}_s^*) - \tilde{c}_s^* &= \frac{c_v}{(1-f)z} \\ 1 - \tilde{c}_s^* - \frac{c_v}{(1-f)\left(z - c_v^2 - c_s^2\right)} &= G(\tau \mid \tilde{c}_s^*) \\ 1 - \tau - \frac{c_v}{(1-f)z} &= \tilde{c}_s^* = \frac{c_s^*}{z} \end{aligned}$$

Hence, as $\varepsilon \to 0$, all validators with signal $c_{s,i} \leq c_s^*$ will work while all other validators will shirk. The probability that the validation process succeeds is $1 - \tau$. Then, as $\varepsilon \to 0$, all validators will work whenever $c_s \leq c_s^*$ and they will all shirk whenever $c_s > c_s^*$. The argument to show uniqueness is standard from Morris and Shin (2003) given the payoff of any validators (20) to communicating with the ledger is increasing in the measure of validators who also communicate with the ledger.

Therefore, the probability that a trade validation process goes through is the probability that $c_s \leq c_s^*$, or

$$\int_{\underline{c}_s}^{c_s^*} \frac{dc_s}{\bar{c}_s - \underline{c}_s} = \frac{(1-\tau)z - \frac{c_v}{(1-f)} - \underline{c}_s}{\bar{c}_s - \underline{c}_s}.$$

B. Proof of Lemma 2

Proof. We first show that (12) must bind. Suppose it does not. The objective function is decreasing in V and (11) is relaxed as V is lowered. So if (12) does not bind, it is optimal to set $V = 0$, and the solution is $\tilde{x} > 0$ which is the solution to $\max [u(x) - x]$ subject to $\frac{\beta\alpha}{1-\beta} [u(x) - x] \geq x$. However, since $\tilde{x} > 0$ and $\tau \leq 1$, it is clear that (12) cannot be satisfied when $V = 0$. Therefore (12) must bind. Suppose now both (11) and (12) bind. Then, given

τ, the solution is given by

$$\begin{aligned}
\frac{\beta\alpha}{1-\beta}\left[u(x)-(x+VZ(\tau))\right] &= (x+VZ(\tau)) \\
\tau V\pi\frac{\beta\alpha}{1-\beta}\left[u(x)-(x+VZ(\tau))+R(\tau)\right] &= (x+VZ(\tau))
\end{aligned}$$

Hence from the first equation,

$$x+VZ(\tau)=\frac{\beta\alpha}{1-\beta+\beta\alpha}u(x)$$

and from the second,

$$\begin{aligned}
\tau V\pi\frac{\beta\alpha}{1-\beta}\left[u(x)-\frac{\beta\alpha}{1-\beta+\beta\alpha}u(x)+R(\tau)\right] &= \frac{\beta\alpha}{1-\beta+\beta\alpha}u(x) \\
\tau V\pi\frac{\beta\alpha}{1-\beta}\left[\frac{1-\beta}{1-\beta+\beta\alpha}u(x)+R(\tau)\right] &= \frac{\beta\alpha}{1-\beta+\beta\alpha}u(x) \\
(1-\tau V\pi)\,\beta\alpha u(x) &= \tau V\pi\frac{\beta\alpha}{1-\beta}\left(1-\beta+\beta\alpha\right)R(\tau) \\
u(x) &= \frac{\tau V\pi}{1-\tau V\pi}\left(1+\frac{\beta\alpha}{1-\beta}\right)R(\tau)
\end{aligned}$$

Hence

$$\begin{aligned}
x+VZ(\tau) &= \frac{\beta\alpha}{1-\beta+\beta\alpha}u(x) \\
x+VZ(\tau) &= \frac{\beta\alpha}{1-\beta+\beta\alpha}\frac{\tau V\pi}{1-\tau V\pi}\left(\frac{1-\beta+\beta\alpha}{1-\beta}\right)R(\tau) \\
x+VZ(\tau) &= \frac{\tau V\pi}{1-\tau V\pi}\frac{\beta\alpha}{1-\beta}R(\tau)
\end{aligned}$$

The problem of the planner then becomes

$$\begin{aligned}
\alpha(1-f)\left\{u(x)-x-V\left(Z(\tau)-\frac{R(\tau)}{1-f}\right)\right\} &= \\
\alpha(1-f)\left\{\frac{\tau V\pi}{1-\tau V\pi}\left(1+\frac{\beta\alpha}{1-\beta}\right)R(\tau)-\frac{\tau V\pi}{1-\tau V\pi}\frac{\beta\alpha}{1-\beta}R(\tau)+V\frac{R(\tau)}{1-f}\right\} &= \\
\alpha(1-f)\left\{\frac{\tau\pi}{1-\tau V\pi}+\frac{1}{1-f}\right\}VR(\tau) &= \\
\alpha(1-f)\left\{\pi\frac{\tau}{1-\tau V}+\frac{1}{1-f}\right\}V\left(\frac{\tau(1-f)\bar{c}_s+\tau c_v}{1-\tau}\right)
\end{aligned}$$

This is strictly increasing in both V and τ. Hence the solution is $V=1-f$ and $\tau=1$. However, this implies $u(x)\to\infty$ and $x\to\pm\infty$. If $x\to-\infty$, we get a contradiction. If $x\to+\infty$, the planner's objective function is

$$\alpha(1-f)\left\{u(x)-x-V\left(\bar{c}_s+c_v+\frac{f}{1-f}c_v\right)\right\}\to-\infty$$

which cannot be optimal. Therefore, (11) and (12) cannot both be binding. This shows that only (12) binds. □

C. Proof of Proposition 2

Proof. Here we complete the proof of our main result when $V=1-f$ so that $\lambda_{1-f}>0$. The first order condition gives

$$\begin{aligned}
[u'(x)-1](1+\delta\lambda)-\lambda\frac{1}{\hat{\tau}V} &= 0 \qquad (21)\\
\left[\frac{R(\hat{\tau})}{1-f}-Z(\hat{\tau})\right]+\lambda\frac{x}{\hat{\tau}V^2}-\delta\lambda Z(\hat{\tau}) &> 0 \qquad (22)
\end{aligned}$$

Using (21) to eliminate λ, the definition $R(\tau)$ and $Z(\tau)$, as well as the expression for τ, (22) becomes

$$\left(\sqrt{\frac{C}{x+C}}\right)\left\{\frac{1+\sqrt{\frac{C}{x+C}}}{(1-f)}\right\}\left[\frac{x}{C}-\frac{1}{[u'(x)-1]}\right]>\delta \qquad (23)$$

where x is given by the incentive constraint of validators holding at equality, which we can write as

$$\delta\left[u(x)-x-C\right]=\frac{C}{1-f}\left(1+\sqrt{\frac{C}{x+C}}\right)\left[\frac{x}{C}+\frac{(1+\sqrt{\frac{C}{x+C}})}{\sqrt{\frac{C}{x+C}}}\right]\equiv H(x)$$

It is tedious to check that $H'(x)>0$ and using the implicit function theorem that $dx/d\delta>0$. The left hand side of (23) is decreasing in x if

$$\frac{1}{u'(x)}+\sqrt{\frac{\rho}{xu'(x)}C}<1,$$

where $\rho=-u''(x)x/u'(x)$. This will hold if C and x are small enough and $\rho\geq 1$ so that $xu'(x)$ is decreasing in x. Assuming this is the case, since $dx/d\delta>0$, x declines as δ decreases. So the LHS of (23) increases when δ decreases and the condition will be satisfied for δ low enough and below some $\bar{\delta}$. □

D. Proof Lemma 4

Proof. Suppose at least τV validators work. If one of the remaining validators shirks, he obtains expected payoff

$$\begin{aligned}\bar{\sigma}_i\left\{(1-f)\left(-c_s+z+\beta U_V(h_t,(1,G),m=1)\right)+f\left(-c_s+\beta U_V(h_t,(0,B),m=1)\right)\right\}\\ +(1-\bar{\sigma}_i)\left\{(1-f)\beta U_V(h_t,(1,G),m=0)+f\beta U_V(h_t,(0,B),m=0)\right\}\quad=\\ \bar{\sigma}_i\left\{-c_s+(1-f)z\right\}+\bar{\sigma}_i(1-f)\beta U_{Vt}+(1-\bar{\sigma}_i)f\beta U_{Vt}\end{aligned}$$

If the buyer has a good label, all other validators send $m_i=1$, so $\mathcal{I}(\mathbf{m})=1$ irrespective of the decision of the shirker. However, the shirker only gets the reward if he also sends $m=1$. If he sends a message when the producer has a bad label, he does not receive a reward and he

gets the autarkic payoff in the future. Therefore, a shirker sends a signal ($\sigma_i = 1$) whenever

$$-c_s + (1-f)(z + \beta U_{Vt}) > f\beta U_{Vt}$$

and does not otherwise.

The expected utility of a working validator (who sends signal G if the producer has label G and nothing if the producer has label B) is as before,

$$\begin{aligned} -c_v + (1-f)(-c_s + z + \beta U_V(h_t,(1,G),m=1)) + f\beta U_V(h_t,(0,B),m=0) &= \\ -c_v + (1-f)(-c_s + z) + \beta U_{Vt} & \end{aligned}$$

The remaining validator will work whenever

$$\begin{aligned} -c_v + (1-f)(-c_s + z) &> -c_s + (1-f)z - f\beta U_{Vt} \\ -c_v + (1-f)(-c_s + \hat{z}) &> -c_s + (1-f)\hat{z} + (1-f)\beta U_{Vt} - \beta U_{Vt} \\ c_s + \beta U_{Vt} - c_v &> (1-f)(c_s + \beta U_{Vt}) \\ 1 - \frac{c_v}{c_s + \beta U_{Vt}} &> (1-f) \end{aligned}$$

So if $(1-f) \geq \frac{f(c_s+\beta U_{Vt})}{(z-c_s+\beta U_{Vt})}$, the remaining validator works whenever

$$1 - \frac{c_v}{c_s + \beta U_{Vt}} > (1-f) \geq \frac{f(c_s + \beta U_{Vt})}{(z - c_s + \beta U_{Vt})},$$

and he would send a signal if he were to shirk.

However, if $-c_s + (1-f)(z + \beta U_{Vt}) < f\beta U_{Vt}$ we have $\bar{\sigma}_i = 0$ and in this case the remaining validator decides to work whenever

$$\begin{aligned} -c_v + (1-f)(-c_s + z) &> -(1-f)\beta U_{Vt} \\ (1-f)(-c_s + z + \beta U_{Vt}) &> c_v \\ 1 - f &> \frac{c_v}{z - c_s + \beta U_{Vt}}. \end{aligned}$$

So if

$$\frac{f\left(c_s+\beta U_{Vt}\right)}{z-c_s+\beta U_{Vt}}>(1-f)>\frac{c_v}{z-c_s+\beta U_{Vt}}$$

the remaining validator works (and he would not send a signal if he were to shirk). Notice that this case is only possible if $f\left(c_s+\beta U_{Vt}\right)>c_v$, or

$$1-\frac{c_v}{c_s+\beta U_{Vt}}>1-f$$

Therefore, combining both condition, validators prefer to work whenever

$$1-\frac{c_v}{c_s+\beta U_{Vt}}>1-f>\frac{c_v}{z-c_s+\beta U_{Vt}}$$

Replacing z using (10), we obtain

$$\begin{aligned}(1-f)(z-c_s+\beta U_{Vt}) &> c_v\\ (1-f)\left[\frac{1}{1-\tau}\left(c_s+\frac{c_v}{1-f}\right)-c_s+\beta U_{Vt}\right] &> c_v\\ \frac{\tau}{1-\tau}\left((1-f)c_s+c_v\right)+(1-f)\beta U_{Vt} &> 0\end{aligned}$$

which always true. Hence, validators prefer to work whenever

$$f>\frac{c_v}{c_s+\beta U_{Vt}}$$

This concludes the proof. □

E. Case with non-degenerate cost distribution

In this Appendix, we consider the case where the distribution of the common cost c_s is non degenerate. We show that a weak sufficient condition for the planner to choose $c_s^*=\bar{c}_s$ defined as $c_s^*\equiv(1-\tau)(z^1+z^2)-\frac{c_v}{1-f}$ is

$$\frac{\bar{c}_s+\frac{c_v}{1-f}}{(1-f)\left(\bar{c}_s-Ec_s\right)}\geq\delta.$$

In this case, all validators will work, irrespective of their private communication costs.

Let c_s^* be defined as above. So validators only verify a trade whenever $c_s \leq c_s^*$. When c_s is uniformly distributed over $[0, \bar{c}_s]$ the probability that validators verify a trade is simply the probability that $c_s \leq c_s^*$, that is $c_s^*/\bar{c}_s$. Validators verify whenever $c_s \leq c_s^*$ and these validators take home

$$Z(\tau) \equiv \frac{c_s^* + c_v/(1-f)}{1-\tau}$$

When $z^1 = Z$, the participation constraint of validators (37) is always satisfied

$$\mathbb{E}_{c_s \leq c_s^*} \left[-c_v + (1-f)\left(Z - c_s\right) \right] \geq 0$$

Let $R(\tau, c_s)$ be the expected rent of validators when the fundamental communication cost is $c_s \leq c_s^*$,

$$\begin{aligned} R(\tau, c_s) &\equiv (1-f)\left[Z(\tau) - c_s\right] - c_v \\ &\equiv (1-f)\left[\frac{c_s^*}{1-\tau} - c_s\right] + \frac{\tau c_v}{1-\tau} \end{aligned}$$

We can set the participation constraint for early producers (34) at equality and replace $y = x + wVZ(\tau)$. Then the set of IF allocations is characterized by

$$\delta \frac{c_s^*}{\bar{c}_s} \left[u(x) - (x + VZ(\tau))\right] \geq (x + VZ(\tau)) \tag{24}$$

$$\delta \int_0^{c_s^*} \left[u(x) - (x + VZ(\tau)) + R(\tau, c_s)\right] \frac{dc_s}{\bar{c}_s} \geq \frac{1}{\tau V}\left(x + VZ(\tau)\right) \tag{25}$$

Notice that validators only work with probability $\frac{c_s^*}{\bar{c}_s}$. Therefore, producers (including validators) can trade only with probability $\frac{c_s^*}{\bar{c}_s}$.

E..1 Optimal design

We need to adapt the objective function to the new setup. Since the probability of a productive match is $1-f$ in each period, the objective function of a planner is the sum of the early and late producers' utility when they trade, and the expected rent of validators from

operating the ledger for a measure α of trades,

$$\int_0^{c_s^*} \left\{ \alpha(1-f) \left[u(x) - y + \left(y - x - Vz^1 \right) \right] + \alpha V R(\tau, c_s) \right\} \frac{dc_s}{\bar{c}_s},$$

or replacing y and z^1, as well as $Z(\tau) - R(\tau, c_s)/(1-f) = c_s + \frac{c_v}{1-f} > 0$, a planner chooses the trading size x, the number of validators V, the threshold τ, and the threshold c_s^* to solve

$$\alpha(1-f) \max_{x,V \geq 0, c_s^* \leq \bar{c}_s, \tau \in [0,1]} \int_0^{c_s^*} \left\{ u(x) - x - V \left(c_s + \frac{1}{1-f} c_v \right) \right\} \frac{dc_s}{\bar{c}_s}$$

subject to (43) and (44). Using the same steps as he simpler case, we can show that (43) is always slack when (44) holds. Re-arranging the constraint,

$$\pi \frac{\beta\alpha}{1-\beta} \frac{c_s^*}{\bar{c}_s} \left[u(x) - x \right] \geq \left(\frac{1}{\tau V} \right) \left(x + VZ(\tau) \right) - \frac{\pi\beta\alpha}{1-\beta} \int_0^{c_s^*} \left(R(\tau, c_s) - VZ(\tau) \right) \frac{dc_s}{\bar{c}_s} \tag{26}$$

Since the objective function is independent of τ, the planner will choose τ to minimize the right hand side of (26). The first order condition for the optimal threshold $\hat{\tau}$ gives

$$\frac{1-\hat{\tau}}{\hat{\tau}} = \sqrt{\frac{\left[1 - \delta \frac{c_s^*}{\bar{c}_s} (1 - f - V) \right] \left(c_s^* + \frac{c_v}{(1-f)} \right)}{\frac{x}{V} + c_s^* + \frac{c_v}{(1-f)}}} \tag{27}$$

This threshold is well defined if

$$1 > \delta \frac{c_s^*}{\bar{c}_s} (1 - f - V)$$

and otherwise $\hat{\tau} = 1$.

Then it is useful to look at the first order conditions in detail. When λ is the Lagrange multiplier on the validators' IC constraint, the first order conditions with respect to x, V

and c_s^* respectively are -

$$
\begin{aligned}
[u'(x)-1]\,(1+\delta\lambda)\,\frac{c_s^*}{\bar{c}_s}-\lambda\frac{1}{\hat{\tau}V} &= 0 && (28)\\
\int_0^{c_s^*}\left[\left(c_s+\frac{1}{1-f}c_v\right)\right]\frac{dc_s}{\bar{c}_s}+\lambda\frac{x}{\hat{\tau}V^2}-\delta\lambda\frac{c_s^*}{\bar{c}_s}Z(\hat{\tau}) &= 0 && (29)\\
\left\{u(x)-x-V\left(Z(\tau)-\frac{R(\tau,c_s^*)}{1-f}\right)\right\}\frac{1}{\bar{c}_s} && && (30)\\
+\lambda\left\{\begin{array}{c}\delta\left[\{u(x)-x-VZ(\tau)\}+R(\tau,c_s^*)\right]\frac{1}{\bar{c}_s}\\ +\delta\int_0^{c_s^*}\left[-V\frac{\partial Z(\tau)}{\partial c_s^*}+\frac{\partial R(\tau,c_s)}{\partial c_s^*}\right]\frac{dc_s}{\bar{c}_s}-\frac{1}{\tau V}V\frac{\partial Z(\tau)}{\partial c_s^*}\end{array}\right\} &\geq 0 && (31)
\end{aligned}
$$

We now determine conditions so that the solution is $c_s^*=\bar{c}_s$. Suppose this is the case. Then the expression in $\{.\}$ in (31), the second part of the FOC which is multiplied by λ (which pertains to the behavior of the IC when the planner increases c_s^*) is

$$\delta\frac{1}{\tau V}x\frac{1}{\bar{c}_s}+\frac{1}{\tau(1-\tau)}\left\{\delta\left[\bar{c}_s+\frac{c_v}{1-f}\right]\frac{1}{\bar{c}_s}-1\right\}+\delta\int_0^{\bar{c}_s}\left[(1-f-V)\frac{1}{1-\tau}\right]\frac{dc_s}{\bar{c}_s}$$

Hence, if

$$\delta\left[1+\frac{c_v}{\bar{c}_s(1-f)}\right]\geq 1$$

then the LHS of the IC is increasing with c_s^* and it is optimal to set $c_s^*=\bar{c}_s$, as long as the objective function is also increasing in c_s^* when evaluated at $\bar{c}_s$, that is

$$u(x)-x-V\left(Z(\tau)-\frac{R(\tau,\bar{c}_s)}{1-f}\right)\geq 0$$

From (26)

$$\delta\left[u(x)-x+\underbrace{\int_0^{\bar{c}_s}R(\tau,c_s)\frac{dc_s}{\bar{c}_s}}_{=ER(\tau,c_s)}-VZ(\tau)\right]=\frac{1}{\tau V}\left(x+VZ(\tau)\right)$$

Hence,

$$\begin{aligned} u(x) - x - V\left(Z(\tau) - \frac{R(\tau, \bar{c}_s)}{1-f}\right) &= \\ \frac{1}{\delta\tau V}\left(x + VZ(\tau)\right) + V\frac{R(\tau, \bar{c}_s)}{1-f} - ER(\tau, c_s) &= \\ \frac{1}{\delta\tau V}x + V\frac{R(\tau, \bar{c}_s)}{1-f} + \left[\frac{1}{\delta\tau} - (1-f)\right]\left(\bar{c}_s + \frac{c_v}{1-f}\right) + (1-f)Ec_s + c_v & \end{aligned}$$

Since $\tau \leq 1$, and a sufficient condition for the RHS to be positive is

$$\begin{aligned} \frac{1}{\delta}\left(\bar{c}_s + \frac{c_v}{1-f}\right) - (1-f)\left(\bar{c}_s + \frac{c_v}{1-f}\right) + (1-f)\left(Ec_s + \frac{c_v}{1-f}\right) &\geq 0 \\ \frac{1}{\delta}\left(\bar{c}_s + \frac{c_v}{1-f}\right) - (1-f)\left(\bar{c}_s - Ec_s\right) &\geq 0 \\ \frac{\bar{c}_s + \frac{c_v}{1-f}}{(1-f)\left(\bar{c}_s - Ec_s\right)} &\geq \delta. \end{aligned}$$

Therefore, f large enough (for example) would allow the inequality to be satisfied. Also, if Ec_s is close enough to $\bar{c}_s$. In those cases, $c_s^* = \bar{c}_s$, and the solution if given by the FOC of the planner's problem,

$$\begin{aligned} \left[u'(x) - 1\right](1 + \delta\lambda) - \lambda\frac{1}{\hat{\tau}V} &= 0 \\ Ec_s + \frac{1}{1-f}c_v + \lambda\frac{x}{\hat{\tau}V^2} - \delta\lambda Z(\hat{\tau}) &= 0 \end{aligned}$$

together with the binding IC,

$$\delta\left[u(x) - x\right] = \left(\frac{1}{\tau V}\right)\left(x + VZ(\tau)\right) - \delta\int_0^{\bar{c}_s}\left(R(\tau, c_s) - VZ(\tau)\right)\frac{dc_s}{\bar{c}_s} \tag{32}$$

F. Concave utility for early producers

In this Appendix, we lay down the analysis when early producers also have a concave utility function. We analyze the case when the distribution of the communication cost is degenerate at $\bar{c}_s$. Participation constraints are

$$u(x) - y - Vz^2 \geq 0 \tag{33}$$

$$v(y) - x - Vz^1 \geq 0 \tag{34}$$

$$-c_v + (1-f)\left(z^1 + z^2 - c_s\right) \geq 0 \tag{35}$$

From the PC of early producers holding at equality,

$$y = v^{-1}(x + Vz^1) \equiv \Phi(x + Vz^1)$$

where Φ is increasing and convex, and the PCs become

$$u(x) - \Phi(x + Vz^1) - Vz^2 \geq 0 \tag{36}$$

$$-c_v + (1-f)\left(z^1 + z^2 - c_s\right) \geq 0 \tag{37}$$

Repayment constraints:

$$-\left(\Phi(x + Vz^1) + Vz^2\right) + \beta U \geq 0. \tag{38}$$

$$-\left(\Phi(x + Vz^1) + Vz^2\right) + \beta \mathbb{E}U_V \geq 0. \tag{39}$$

No bribe.

$$\pi\beta U_V \geq \bar{z}$$

When a share w of validators are working on a match, the late producer in this match is willing to pay at most a total of $y + wVz^2$ to get away with production. Given the ledger requires the agreement of at least τV validators to validate a transaction, the cheating late producer will pay $\bar{z} = (y + Vz^2)/(\tau V)$ to τV validators. Using $\bar{z} = (y + wVz^2)/(\tau V)$, a

validator rejects the bribe whenever[36]

$$\pi\beta U_V \geq \frac{1}{\tau V}\left(y + Vz^2\right). \tag{40}$$

$$\pi\beta U_V \geq \frac{1}{\tau V}\left[\Phi(x + Vz^1) + Vz^2\right]. \tag{41}$$

Validation threshold.

$$z^1 + z^2 \geq \frac{c_s + c_v/(1-f)}{1-\tau}. \tag{42}$$

Since the payment to validators should be minimized, (10) binds so validators take home

$$z^1 + z^2 = Z(\tau) \equiv \frac{\bar{c}_s + c_v/(1-f)}{1-\tau}$$

Cheapest to deliver.

Given $Z(\tau)$ the planner will choose z^1 and z^2 that minimizes the total cost of delivering the amount $Z(\tau)$. That it the planner will choose z^1 and z^2 to solve

$$\min_{z^1 \geq 0} \Phi(x + Vz^1) + V\left(Z(\tau) - z^1\right)$$

with FOC,

$$\Phi'(x + Vz^1) - \lambda - 1 = 0$$

where λ is the Lagrange multiplier in the minimization problem. Hence, $z^1 = 0$ whenever $\Phi'(x) \geq 1$.

When $z^1 + z^2 = Z$, the participation constraint of validators (37) is always satisfied. Let $R(\tau)$ be the expected rent of validators,

$$\begin{aligned} R(\tau) &\equiv (1-f)\left[Z(\tau) - (\bar{c}_s + c_v)\right] - fc_v \\ &= \frac{\tau(1-f)\bar{c}_s + c_v}{1-\tau} - c_v^1 \end{aligned}$$

[36]We assume validators act in unison: they either all accept the bribe, or they all reject it.

We can set the participation constraint for early producers (34) at equality and replace $y = \Phi\left(x + VZ(\tau)\right)$. Since validators earn a rent, $U_V \geq U$ and (39) is satisfied whenever (38) is. Then the set of IF allocations is characterized by

$$\frac{\beta\alpha}{1-\beta}\left[u(x) - \Phi\left(x + Vz^1\right) - Vz^2\right] \geq \Phi(x + Vz^1) + Vz^2 \tag{43}$$

$$\pi\frac{\beta\alpha}{1-\beta}\left[u(x) - \Phi\left(x + Vz^1\right) - Vz^2 + R(\tau)\right] \geq \frac{1}{\tau V}\left(\Phi(x + Vz^1) + Vz^2\right) \tag{44}$$

References

Abadi, Joseph and Markus Brunnermeier (2018) "Blockchain Economics," NBER Working Papers 25407.

Amoussou-Guenou, Yackolley, Bruno Biais, Maria Potop-Butucaru, and Sara Tucci-Piergiovanni (2019) "Rationals vs Byzantines in Concensus-based Blockchains," Research report, HAL ID: hal-02043331.

Andolfatto, David (2020) "Assessing the Impact of Central Bank Digital Currency on Private Banks," *The Economic Journal*, ueaa073.

Arner, Douglas W., Raphael Auer, and Jon Frost (2020) "Stablecoins: risks, potential and regulation," *Financial Stability Review* 39, 95-123.

Auer, Raphael (2019) "Beyond the doomsday economics of 'proof-of-work' in cryptocurrencies", BIS Working Papers, no. 765, January.

Auer, Raphael and R Boehme (2020): "The technology of retail central bank digital currency", BIS Quarterly Review, March, p. 85-97.

Auer, Raphael, Giulio Cornelli, and Jon Frost (2020) "Rise of the central bank digital currencies: drivers, approaches and technologies", BIS Working Papers, no. 880, August.

Aymanns, Christoph, Mathias Dewatripont, and Tarik Roukny (2020) "Vertically Disintegrated Platforms" SSRN, February.

Baudet Mathieu, George Danezis, Alberto Sonnino (2020) "FastPay: High-Performance Byzantine Fault Tolerant Settlement" arXiv:2003.11506v2 [cs.CR].

Biais, Bruno, Christophe Bisière, Matthieu Bouvard, and Catherine Casamatta (2019) "The blockchain folk theorem", *The Review of Financial Studies*, vol. 32, n. 5, May 2019, pp. 1662–1715.

Boar, Codruta and Andreas Wehrli (2021) "Ready, steady, go? - Results of the third BIS survey on central bank digital currency", BIS papers 114.

Bonneau, Joseph (2016) "Why buy when you can rent?" International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg.

Boyd, John and Prescott, Edward (1986) "Financial Intermediary Coalitions," *Journal of Economic Theory*, 38, 211–232.

Brunnermeier, Markus and Dirk Niepelt (2019) "On the Equivalence of Private and Public Money," *Journal of Monetary Economics* 106, 27-41.

Budish, Eric (2018) "The economic limits of bitcoin and the blockchain", NBER Working Papers, no 24717, June.

Cavalcanti, Ricardo and Neil Wallace (1999a), "A Model of Private Bank Note Issue," *Review of Economic Dynamics,* 2, 104–136.

Cavalcanti, Ricardo and Neil Wallace (1999b), "Inside and Outside Money as Alternative Media of Exchange," *Journal of Money, Credit, and Banking*, 31, 443–457.

Calle, George and Daniel Eidan (2020) "Central Bank Digital Currency: an innovation in payments", Whitepaper, R3, April

Carlsson, Hans and Eric E. van Damme (1993) "Global Games and Equilibrium Selection," *Econometrica* 61, 989- 1018.

Chiu, Jonathan and Thorsten Koeppl (2017) "The Economics of Cryptocurrencies - Bitcoin and Beyond," No 1389, Working Papers, Queen's University, Department of Economics.

Chiu, Jonathan and Thorsten Koeppl (2019) "Blockchain-Based Settlement for Asset Trad-

ing," *Review of Financial Studies* 32, 1716-1753.

Chiu, Jonathan and Thorsten Koeppl (2020) "Payments and the D(ata) N(etwork) A(ctivities) of BigTech Platforms," Mimeo Queen's University.

Committee for Payments and Markets Infrastructure (2017) "Distributed ledger technology in payment, clearing and settlement - an analytical framework", February

Cong Lin William, Zhiguo He and Jiasun Li (2020) "Decentralized Mining in Centralized Pools," *The Review of Financial Studies*, hhaa040, https://doi.org/10.1093/rfs/hhaa040.

Corda (2020) "Corda OS 4.6 Developer Documentation", https://docs.corda.net/docs/corda-os/4.6/key-concepts-notaries.html, accessed on 11.11.2020.

Deepesh Patel and Emmanuelle Ganne (2019) "Blockchain & DLT in trade: a reality check," World Trade Organisation, November.

Diamond, Doug (1984), "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51, 393–414.

Garatt, Rodney and Maarten R.C. van Oordt (2020) "Why Fixed Costs Matter for Proof-of-Work Based Cryptocurrencies," Bank of Canada Working Paper 2020-27.

Grossman, Sanford J., and Joseph E. Stiglitz (1976) "Information and competitive price systems," *The American Economic Review* 66, 246-253.

Eyal I., Sirer E.G. (2014) "Majority Is Not Enough: Bitcoin Mining Is Vulnerable," in: Christin N., Safavi-Naini R. (eds) Financial Cryptography and Data Security. FC 2014. Lecture Notes in Computer Science, vol 8437. Springer, Berlin, Heidelberg

Giulia Fanti, Leonid Kogan, and Pramod Viswanath (2019) "Economics of Proof-of-Stake Payment Systems," Mimeo, University of Illinois.

Fernandez-Villaverde, Jesus, Daniel Sanches, Linda Schilling, and Harald Uhlig (2020) "Central Bank Digital Currency: Central Banking For All?," NBER Working Papers 26753.

Frost, Jon, Hyun-Song Shin and Peter Wierts (2020) "An early stablecoin? The Bank of Amsterdam and the governance of money," BIS Working Papers, no. 902.

Gersbarch, Hans, Akaki Mamageishvili, and Oriol Tejada (2020) "Appointed Learning for the Common Good: Optimal Committee Size and Efficient Rewards," CEPR Working Paper DP15311.

Gervais, Arthur, Ghassan O Karame, Karl Wüst, Vasileios Glykantzis, Hubert Ritzdorf, Srdjan Capkun (2016) "On the security and performance of proof of work blockchains," CCS, Proceedings of the 2016 ACM SIGSAC conference on computer and communications security, October

Goldstein, Itay and Ady Pauszner (2005) "Demand–Deposit Contracts and the Probability of Bank Runs," *Journal of Finance* 60, 1293-1327.

Gu Chao, Fabrizio Mattesini, Cyril Monnet, and Randall Wright (2013) "Banking: A New Monetarist Approach," *Review of Economic Studies* 80, 636-662.

Green, Edward, and Robert Porter (1984) "Noncooperative Collusion under Imperfect Price Information," *Econometrica* 52, 87-100.

Huang, Angela (2019) "On the Number and Size of Banks: Efficiency and Equilibrium," Mimeo, National University of Singapore.

Huberman, Gur, Jacob Leshno, and Ciamac C. Moallemi (2021) "Monopoly without a monopolist: An economic analysis of the bitcoin payment system," *Review of Economic Studies*, forthcomming.

Judmayer, Aljosha, Nicholas Stifter, Philipp Schindler, and Edgar Weippl (2018) "Pitchforks in Cryptocurrencies," In Data Privacy Management, Cryptocurrencies and Blockchain Technology, pp. 197-206. Springer, Cham, 2018.

Kandori, Michihiro (2001) "Introduction to Repeated Games with Private Monitoring," *Journal of Economic Theory* 102, 1–15.

Keister, Todd and Cyril Monnet (2020) "Central Bank Digital Currency: Stability and Information," Mimeo, Rutgers University.

Kocherlakota, Narayana (1998) "Money Is Memory," *Journal of Economic Theory* 81, 232-251.

Kocherlakota, Narayana and Wallace, Neil (1998) "Incomplete Record-Keeping and Optimal Payment Arrangements," *Journal of Economic Theory* 81(2), 272-289.

Kroll, Joshua A., Ian C. Davey, and Edward W. Felten (2013) "The economics of Bitcoin mining, or Bitcoin in the presence of adversaries," Proceedings of WEIS. Vol. 2013.

Lagos, Ricardo and Randall Wright (2005) "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy* 113, 463-484.

Lagos, Ricardo, Guillaume Rocheteau, and Randall Wright (2017) "Liquidity: A New Monetarist Perspective," *Journal of Economic Literature* 55, 371-440.

Leland, H. E. and Pyle, D. H. (1977) "Informational Asymmetries, Financial Structure and Financial Intermediation," *Journal of Finance*, 32, 371–387.

Leshno, Jacob, and Philipp Strack (2020) "Bitcoin: An Axiomatic Approach and an Impossibility Theorem," *American Economic Review: Insights* 2, 269-86.

Li, Yiting, and Chien-Chiang Wang (2019) "Cryptocurrency, Imperfect Information, and Fraud," Munich Personal RePEc Archive (MPRA) Paper No. 94309.

Monnet, Cyril (2006) "Private vs Public Money," *International Economic Review* 47(3), 951–960.

Monnet, Cyril and Erwan Quintin (forthcoming) "Optimal Financial Exclusion," *American Economic Journal: Microeconomics.*

Morris, Stephen and Shin, Hyun Song (1998) "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review* 88, 587-97.

Morris, Stephen, and Hyun Song Shin (2002) "Measuring Strategic Uncertainty," manuscript London School of Economics.

Morris, Stephen, and Hyun Song Shin (2003) "Global Games: Theory and Applications," in Advances in Economics and Econometrics: Proceedings of the Eighth World Congress of the Econometric Society, vol. 1, edited by Matthias Dewatripont, Hansen, Lars P. and Stephen J. Turnovsky, chap. 3, pp. 56–114. Cambridge University Press.

Pagnotta, Emiliano (2021) "Decentralizing Money: bitcoin prices and blockchain security" *The Review of Financial Studies* (forthcoming).

Prat, Julien, and Benjamin Walter (2018) "An Equilibrium Model of the Market for Bitcoin Mining," CESifo Group Munich No. 6865.

Rahman, David (2012) "But Who Will Monitor the Monitor?" *American Economic Review* 102, 2767-97.

Rochet, Jean-Charles and Xavier Vives (2004) "Coordination Failures and The Lender of Last Resort," *Journal of the European Economic Association* 2, 1116–1147.

Rocheteau, Guillaume and Ed Nosal (2017) "Money, Payments, and Liquidity," MIT Press.

Saleh, Fahad (2021) "Blockchain Without Waste: Proof-of-Stake" *Review of Financial Studies*, Forthcoming.

Shanaev, Savva Arina Shuraeva, Mikhail Vasenin, Maksim Kuznetsov (2020) "Cryptocurrency Value and 51% Attacks: Evidence from Event Studies" *The Journal of Alternative Investments*, forthcoming.

Schilling, Linda and Harald Uhlig (2019) "Some Simple Bitcoin Economics," *Journal of Monetary Economics* 106, 16-26.

Sveriges Riksbank (2020) "The Riksbank's e-krona pilot", February.

Teutsch, Jason, Sanjay Jain, and Prateek Saxena (2016) "When cryptocurrencies mine their own business," International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg.

Tirole, Jean (2020) "Public and Private Spheres and the Authentic Self.," Mimeo, Toulouse School of Economics.

Townsend, Robert M. (2020) "Distributed Ledgers: Design and Regulation of Financial Infrastructure and Payment Systems", MIT Press, forthcoming.

Wallace, Neil (2005) "From Private Banking To Central Banking: Ingredients of a Welfare Analysis," *International Economic Review*, Vol. 46, No. 2, 619–632.

Williamson, Stephen (1986) "Costly Monitoring, Financial Intermediation and Equilibrium Credit Rationing," *Journal of Monetary Economics*, 18, 159–179.

Williamson, Stephen (1987) "Financial Intermediation, Business Failures, and Real Business Cycles," *Journal of Political Economy*, 95, 1196–1216.

Williamson, Stephen and Randall Wright (2011) "New Monetarist Economics: Models," in Handbook of Monetary Economics, vol. 3A, B. Friedman and M. Woodford, eds, Elsevier, 25-96.