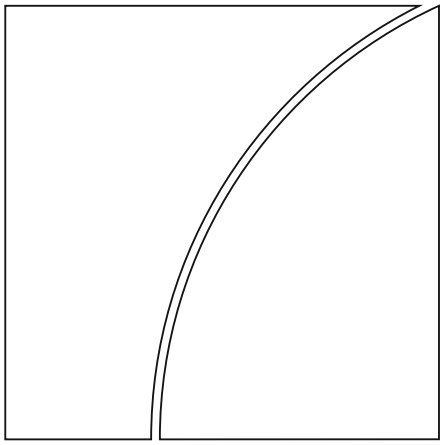
BANK FOR INTERNATIONAL SETTLEMENTS



BIS Working Papers

No 924

Distributed ledgers and the governance of money

by Raphael Auer, Cyril Monnet and Hyun Song Shin

Monetary and Economic Department

January 2021 (revised October 2023)

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

Keywords: market design, money, distributed ledger technology, DLT, blockchain, decentralized finance, global game, consensus.

BIS Working Papers are written by members of the Monetary and Economic Department of the Bank for International Settlements, and from time to time by other economists, and are published by the Bank. The papers are on subjects of topical interest and are technical in character. The views expressed in them are those of their authors and not necessarily the views of the BIS.

This publication is available on the BIS website (www.bis.org).

ISSN 1020-0959 (print)
ISSN 1682-7678 (online)

Distributed Ledgers and the Governance of Money*

Raphael Auer	Cyril Monnet	Hyun Song Shin
BIS	University of Bern and SZ Gerzensee	BIS

October 2023

Blockchain technology makes plausible the classical vision of money as a substitute for a ledger of all past transactions. While it involves updating the ledger through a decentralized consensus on the unique truth, the robustness of the equilibrium that supports this consensus depends on who has access to the ledger and how it can be updated. Using a global game analysis of an exchange economy with credit, we solve for the optimal ledger design that balances security, scalability and decentralization. When intertemporal incentives are strong, a centralized ledger is always optimal. Otherwise, decentralization may be optimal.

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

Keywords: market design, money, distributed ledger technology, DLT, blockchain, cryptocurrencies, decentralized finance, global game, consensus.

*We thank Joseph Abadi, Aleksander Berentsen, Aldar Chan, Francesca Carapella, Jon Frost, Rod Garratt, Piero Gottardi, Hans Gersbach, Hanna Halaburda, Ricardo Lagos, Jacob Leshno, Dirk Niepelt, Jean-Charles Rochet, Harald Uhlig, participants of the 2021 CBER Forum, CB&DC Virtual Seminar Series, 2020 CEBRA annual meeting, the 2020 Summer Workshop on Money and Payments, the 2020 ETH-Zurich Workshop on Future Money and seminar participants at the BIS, Yale University and the University of Zurich for comments. The views presented in this paper are those of the authors and not necessarily those of the Bank for International Settlements.

1 Introduction

Money, whether it be in the form of clay pots, precious coins, or banknotes, is a social convention that serves as a record of goods sold or services rendered in the past. In seminal contributions to monetary theory, Kocherlakota (1998) and Kocherlakota and Wallace (1998) show that the social convention of money can achieve the same allocations as when agents have free access to a universal ledger of all previous transactions in the economy.

In practice, a trusted intermediary, such as a central bank, facilitates the use of money as a social convention by maintaining and transferring balances among depositors. Then the central bank acts as the ultimate guarantor of credit and monetary exchanges. This centralized record-keeping system has been effective. However, the advent of blockchain technology and the concept of a free universal ledger of all past transactions, offers an alternative monetary system that eliminates the need for intermediaries like the central bank.

The universal ledger was given concrete form in the Bitcoin white paper (Nakamoto (2008)), which proposed a digital payment system "without the need for a trusted third party." The Bitcoin protocol guarantees the legitimacy of Bitcoin transactions by controlling the entries in a universal ledger of all transactions, exactly as envisaged in Kocherlakota (1998).[1]

Our paper explores the limits of universal ledgers and considers generalizations of the arrangements for its governance. The ledger could be centrally managed by a single authority, by committees formed of a limited number of privileged agents, or it could be managed by every participant of a decentralized economy.

[1]Recently, a sizeable literature has examined the economics of Bitcoin and other anonymous cryptocurrencies, see e.g. Abadi and Brunnermeier (2021), Biais et al. (2019), and Schilling and Uhlig (2018)). These are "permissionless", ie available for inspection by all and can possibly be updated by anybody. However, they are inherently limited in their economic efficiency, energy intensive, and slow (see i.e. Budish (2018), Chiu and Koeppl (2018), Abadi and Brunnermeier (2021)).

A broader research question is whether and under what circumstances it is economically efficient to implement market designs in which the updating of the ledger is to some extent decentralized (i.e., there is more than one intermediary). In this paper, we analyze the economic forces determining the optimal governance of money – where we understand money broadly as a ledger of past transactions – and examine the conditions under which a decentralized system is more efficient than one based on a single intermediary managing the ledger. To be precise, we do not seek to model the technical subtleties of the ledger and remain general on these aspects. Nor do we analyze whether a universal ledger does better than fiat money. Rather, we revisit Juvenal's analysis of who – or what mechanism – should guard the guardians of the ledger (see also Hurwicz (2008) and Rahman (2012)).

The agents in charge of managing the ledger – the "validators" – should accurately verify transactions and correctly update the ledger. Several versions of so-called "permissioned distributed ledgers" have been proposed as a way to implement decentralized consensus mechanisms for practical applications in setting with known users.[2] We examine both identified and anonymous variants of the technology, focussing on the economic principles underlying the validation protocol and how the incentives to validators can alleviate the two frictions that technology alone cannot solve. First, there is no technical way to force a validator to sign any given transaction. Validation may be costly – especially if this involves costly monitoring of off-chain events.[3] Validators hence need incentives to actively verify and vote on transactions. Second, nothing can technically prevent a validator from validating multiple

[2]Three permissioned distributed ledger systems currently in use are Corda, Hyperledger and Quorum. 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.

[3]The underlying difficulties of writing real-world information into the ledger is widely known as the "oracle problem" in the cryptocurrency industry. See Caldarelli (2020) and Xu et al. (2016) for an introduction to oracles in blockchain.

ledgers with conflicting histories.[4]

We examine theoretically how the optimal validation protocol deals with the frictions embedded in the consensus mechanism underlying the monetary system 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 We start from a credit economy, in which some agents produce early for some other agents and expect these beneficiaries to reciprocate at a later date. However, when the latter "late" producers cannot commit to reciprocate, trade becomes impossible without an external enforcement mechanism.

In similar contexts, Kocherlakota (1998) shows that trustless exchange can be sustained as an equilibrium when agents can freely consult a record-keeping device tracking the default history of the beneficiaries, as long as the device automatically updates itself according to the behavior of the beneficiaries (see also e.g. Rocheteau and Nosal, 2017). Our contribution is to endogenize the validation of records on the ledger and how consensus on actual transactions is reached as an equilibrium outcome in a repeated game setting. As in Abadi and Brunnermeier (2021), Amoussou-Guénou et al. (2019), and Halaburda et al (2021), ensuring the integrity and consistency of the ledger is a quintessential design issue because our economy cannot generate value without it. We assume validators are in charge of reading and updating the ledger of trade histories,[5] and the consensus mechanism relies on supermajority voting. Since verification and communication are costly activities, validators must be compensated for their efforts. And, since validators cannot be trusted to work, or to refuse bribes to falsify an entry, reaching consensus as an equilibrium actually requires that

[4]This opens up the possibility of a "history-reversion attack" (see i.e. Shanaev et al. 2020).

[5]One could interpret that validators act as the-so called "oracles" in decentralized finance applications: Oracles are reference points for external information – such as asset prices or interest rate benchmarks – that are used as an input to calculate the pay-outs of self-executing (i.e. smart) financial contracts.

validators are compensated in excess of costs; i.e. they are paid rents.

We analyze the optimal design of the trade and validation mechanisms, where optimality is defined in terms of maximizing the surplus from the trade net of the validation costs. 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. We consider *internal* validation, when validators are also users of the system, as well as *external* validation, when they are not.

Results We find that decentralizing record keeping (ie having more than a single validator) can be more efficient than relying on a single intermediary. However, such improved governance does not come for free; i.e. ensuring incentives of the validators is costly and requires giving up unanimity and therefore possibly uses only a weak consensus.

Intertemporal incentives are key to characterize the optimal solution. Validators' incentives are sharpened with the reward they obtain for validating transactions and that reward should be high enough to deter them from ever falsifying the ledger. When intertemporal incentives are strong in the sense that the present values of future rewards are high, validators would have much to lose from misbehaving. In this case, a single validator who earns a large rent can be entrusted with managing the ledger. This validator can be drawn from the participants of the credit system, and the size of each transaction is at the first best level.

Our surprising main finding is that precisely when intertemporal incentives are weak, it becomes too costly to prevent a single validator from misbehaving. The optimal design increases the number of validators and reduces the size of each transaction to lower the incentives for bribing. The many validators play a consensus game that has attributes of a public good provision game – consensus is reached if and only if a supermajority of validator agree – which we proceed to solve using global game methods (see Carlson and van Damme,

1993 and Morris and Shin, 1998, 2003). Validators reach a given level of consensus as a unique, dominance solvable equilibrium if and only if they earn a large enough reward that is above some threshold.[6]

We also show that decentralization naturally leads to a "stakeholder economy" in which the participants of the system are in charge of the record keeping: only internal validation can support trade as an equilibrium with more than one validator. There are hence 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.

We consider both permissionless and permissioned variants of the technology, allow for identification and pseudonymity, and model the coexistence of multiple ledgers. In the model presented in the main section, validators pay a fixed one-time cost set by the mechanism. In Online Appendix A, we consider the case where the identity of validators is known and there is no cost to be a validator.[7] In both cases. the mechanism decides whether validation should be permissioned or permissionless. Also on the side of the users, there can either be complete identification or complete anonymity. In all cases, users can switch between multiple ledgers (or, as shown in Online Appendix F, between the ledger and an outside option offering an exogenously fixed payoff.)

Finally, in Online Appendix C we analyze an extension of our model which incorporates a free rider problem in verification: in this extension, validators have an incentive to pretend to verify without exerting the actual monitoring effort, which jeopardizes the legitimacy of the whole ledger. 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'

[6]Our model has the feature that consensus cannot be reached without validators earning rents (see also Abadi and Brunnermeier (2018)).

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

participation constraint can be implemented in our strategic set-up.

In addition to studying the general question of the optimality of decentralized record-keeping, our approach is also useful to understand the so-called "Oracle problem," which relates to the challenge of putting off-chain information onto the chain. Blockchains are dependent on "Oracles" because blockchains are blind to the outside world, as emphasized in Caldarelli (2020). For example, a smart contract may allocate payouts depending on the prevailing inflation rate, but the rate itself needs to be proved by an external data source. Such oracles are centralized, dependable third parties that act as a link between blockchain technology and the outside world. The adoption of oracles is frequently viewed as a "problem" by blockchain enthusiasts as they reintroduce the ideas of centralization and trustworthy third parties.[8] In the model we lay down, trade and production can be seen as off-chain events and validators serve as the oracles who confirm the validity of these off-chain events. These oracles will need to be paid for onboarding information as long as the process of communicating and verifying information is expensive, and incentives imply that in order to ensure the accuracy of the information they onboard, oracles must earn a rent.

Relation to the literature The academic literature on distributed ledgers is dominated by studies of the protocol underlying the Bitcoin blockchain. A sizable literature analyzes the incentives of miners in Bitcoin and similar cryptocurrencies to follow the proof-of-work protocol.[9] Kroll et al (2013) and Prat and Walters (2020) examine free entry and the dynamics of the "mining" market,[10] while Easley et al (2019) and Hubermann et al. (2021)

[8]Garratt and Monnet (2023) show that the use of Oracles is the solution to a fundamental problem of decentralized systems.

[9]The variant with staking one's cryptocurrency holding 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. 2021). However, proof-of-stake can also be attacked via so called "long-run attacks" (see Deirmentzoglou et al. 2019 for a survey). Therefore, proof-of-stake implicitly assumes the existence of some overarching social coordination (see Buterin, 2014).

[10]See also Cong et al. (2021) for an analysis of the concentration of mining and efficiency.

examine the economics of the transaction market. Budish (2018) and later Chiu and Koeppl (2022) 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 finality.[11] Further to this, even in the absence of incentives to reverse history, sunspot equilibria can arise in proof-of work based blockchains (Biais et al 2019).[12]

The literature on validator incentives and design of permissioned versions of distributed ledgers is sparser.[13] Townsend (2020) focuses on an economics-based approach to the issue of distributed ledgers, exploring novel contracting possibilities enabled by DLT. Most closely related to our analysis are Abadi and Brunnermeier (2021), Halaburda et al. (2021), and Amoussou-Guénou et al. (2019). The latter authors first modeled the interaction between validators as a game entailing non-observable effort to check transactions and costly voting. They also analyzed this game in terms of moral hazard and public good provision. Closely related, Abadi and Brunnermeier (2021) and Halaburda et al (2021), examine the incentives to reach consensus using communication games and their robustness if rational nodes can freely send messages to selected recipients only.

Relative to their analysis, our focus is not on the specific steps of the communication game needed to reach consensus. Rather, our contribution is to link the ledger validation game to monetary exchange in a repeated game setting. We establish the uniqueness of the equilibrium via a global game approach, and characterize the optimal mechanism design, in

[11]Such attacks are outlined in Nakamoto (2008). See Eyal and Sirer (2014) and Gervais et al. (2016) for other attacks. Other references on the economic analysis of Bitcoin are Böhme et al. (2015), Schilling and Uhlig (2019), Garatt and van Oordt (2020) and Leshno and Strack (2020).

[12]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. Halaburda et al. (2021) recently examined possible equilibria and their robustness if rational validators can send messages to selected recipients only or even send conflicting messages to different recipients.

[13]Applications of permissioned DLT are being explored for securities settlement systems, trade finance solutions, "stablecoins", and central bank digital currencies, see also Baudet et al. (2020), Arner et al. (2020), Auer et al. (2020), and Chiu and Koeppl (2019).

particular in terms of the number of validators, size of transactions, and optimal super-majority 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.[14] 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-Guénou et al. (2019) we also derive a folk theorem.

More narrowly, in the context of existing applications in decentralized finance, our model shows how the so-called "Oracle Problem" can be solved via incentive design. On platforms such as Ethereum, oracles serve as reference points for external information – such as asset prices, interest rate benchmarks, or other relevant variables such as the official inflation rate (See Xu et al. (2016) for an introduction to oracles). Since such information is used as an input to calculate the pay-outs of self-executing (i.e. smart) financial contracts, oracles are easy targets for manipulation.[15]

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.[16] 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

[14]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. In a related context, Halaburda et al. (2021), examine the incentives to reach consensus using communication games and their robustness if rational nodes can freely send messages to selected recipients only. Consensus can be reached but not necessarily on the true outcome.

[15]See for example Luu et al. (2016) and Froewis and Boehme (2017).

[16]While it is costly to duplicate verification and communication across many validators, we find conditions under which many validators are better than one. 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.

number of intermediaries when they have an incentive to divert deposits. A related analysis that focuses on 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 of a bank run occurring, eg by Rochet and Vives (2004) and Goldstein and Pauzner (2005).

Section 2 lays down the basic set-up and characterizes benchmark allocations absent a record-keeping device and a freely accessible one. Section 3 defines incentive feasible allocation with DLT, and characterizes the optimal allocation including the optimal number of validators.

2 The model

Our model builds on Gu et al. (2013).[17] Time is discrete and infinite.The discount factor is $\tilde{\beta} \in (0, 1)$. Each period is divided in two distinct production/consumption stages, early and late with one good per stage, the "early good" and the "late good." Goods are non-storable across stages or across periods. There is a continuum of agents with two permanent types. There is a unit mass of early producers and another unit mass of late producers.[18] Early producers can produce the early good that late producers like to consume. Late producers can produce the late good that early producers like to consume. Early and late producers have a survival probability σ at the end of a period. They learn whether they survive at the very end of the period so that their effective discount factor is $\beta = \sigma\tilde{\beta}$. Exiting agents are replaced by new agents of the same type so the distribution of types is stationary.

[17] Gu et al. (2013) borrows methodological elements from Lagos and Wright (2005). See also Williamson and Wright (2011), and Lagos et al. (2017). Berentsen and Schaer (2020) is a very clear exposition of the link between monetary theory and ledgers.

[18] In Online Appendix C we also consider the case where a mass f of late producers have a faulty production technology and can never produce while they can mimic the productive late producers.

Preferences of early and late producers are represented by the following utility function, respectively[19]

$$\begin{aligned} U_e(x^e, y^e) &= x^e - y^e \\ U_\ell(x^\ell, y^\ell) &= u(x^\ell) - y^\ell \end{aligned}$$

where x^e (resp. x^ℓ) is the consumption of early (resp. late) producers, and y^e (resp. y^ℓ) is the production of early (resp. late) producers. The function $u(.)$ is continuous, increasing, concave, and $u(0) = 0$. We assume that there are gains from trade between early and late producers: there exists x such that $u(x) > x$. We denote by x^* the efficient allocation that solves $u'(x^*) = 1$.

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 and late producers and matches one with another pairwise. All other producers remain unmatched for the period. Therefore, the probability of a match for any producer is α. The match is maintained across both stages but it dissolves at the end of the later stage.

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

We restrict trading mechanisms to be in the class of coordination games: In each match the two agents announce a pair $(\tilde{x}, \tilde{y}) \in \mathbb{R}^2_+$. If both announcements coincide, the early producer

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

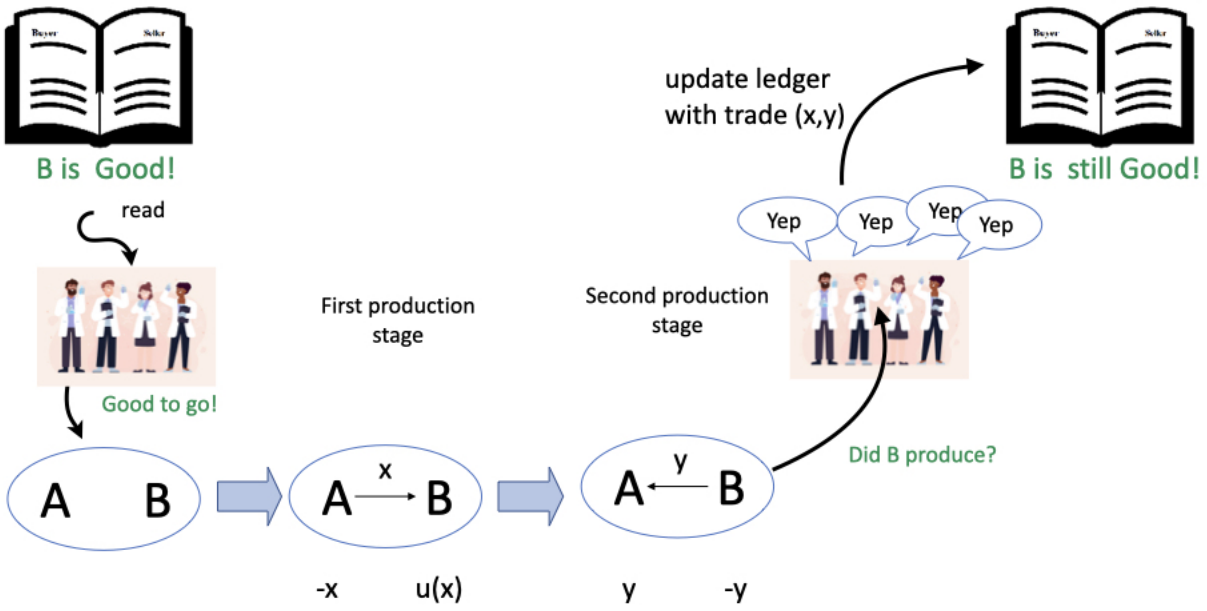


Figure 1: Timeline

produces $x = \tilde{x}$ in the early stage and the late producer produces $y = \tilde{y}$ in the late stage. In this case we say that the allocation (x, y) can be implemented.

We assume late producers only have a limited ability to commit.[20] Absent commitment and any record-keeping technology, it is routine to show that late producers will never produce for early producers and the only implementable allocation is autarky $(x, y) = (0, 0)$.

A ledger technology

To discipline late producers it is necessary to record their history of trades in a ledger. Such a truthful record helps to discipline late producers by threatening the loss of future consumption in case they do not produce today (see Kocherlakota, 1998). The ledger records

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

$\emptyset$ in case producers have no match. Histories of producers can be conveniently summarized using two labels: good (G) or bad (B). We will also say that late producers can be in good or bad standing. A producer will be assigned label B whenever his actions differed from his announcements some time in the past, irrespective of how long ago it was. Otherwise, a producer will be assigned label G. Notice that label B is an absorbing label and an agent carrying label B will never consume or produce.[21]

Then, an allocation (x, y) is implementable if it is incentive feasible (IF): it satisfies participation constraints and late producers produce for the early producer. The two participation constraints are $u(x) \geq y$ and $y \geq x$, while late producers will have the incentive to produce for 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 $-y$ and are assigned a good label so they can trade in the future. The expected value of having a good label is equal to the discounted lifetime gains from trade times the probability of trading α. Setting $y = x$, the set of IF allocations is characterized by all x that satisfy

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

The efficient allocation x^* is implementable if β and α are large enough.

[21] When a late producer has label B, early producers rationally expect that the late producer has defaulted in the past, and therefore announces $(\tilde{x}, \tilde{y}) = (0, 0)$ in the coordination game. Anticipating that all agents in the future will announce $(\tilde{x}, \tilde{y}) = (0, 0)$, a late producer with label B will not produce.

3 Trustless Ledgers

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, without the help of a trusted central authority, and explain the incentives problem arising from the mechanism used to update the ledger.

We next endogenize the updating process of the ledger. We assume the ledger is managed by a measure V of "validators." Validators are rational agents who need incentives to verify transactions and update the ledger honestly.[22] The resulting history should be trusted by all producers. Validators should meet eligibility requirements to gain access to the validation system, and their interactions are invisible to outsiders.

We consider two types of validator: In period 0, a measure V of validators can be selected from the set of late producers (internal validation) or from a set of agents who only consume both goods indifferently (external validation).[23] Internal validators can trade, but they cannot validate their own trades. Each period, all V validators work on validating each of the α match.[24] We assume a validator uses the same strategy across all matches, as described below. Therefore, a validator either works on validating all matches or none. Exiting validators are replaced by new validators in a way we specify below. For symmetry we assume that external validators are also subject to the survival shock.

There is a private cost to start trading on the ledger (this we will later show is optimal). There is a cost γ to late producers to open an account,[25] while the cost to open an account

[22] These validators can be thought of as the notaries in permissioned ledgers such as Corda.

[23] The latter could include early producers.

[24] When validators are selected from late producers, $V \leq 1$. Since there is a continuum of validators, there is also a measure V that will work on validating the trade involving a validator. In this sense, all statements should be qualified with "almost surely."

[25] Only the account and its activities are recorded on the ledger, but the underlying identity of the person controlling the account is not known.

eligible for validation is γ_v, where $v = IV$ when there is internal validation and $v = EV$ when there is external validation. Once an account is open, validators only verify the standing of that account, that is whether the holder of an account defaulted in the past on the liabilities generated with that account.[26] If late producers defaulted on an account, that account is assigned label B and trade is no longer possible with it. Since the full identity of the holder of the account is not known, he or she can open another account. Opening a new account does not give information as to the reasons why that account is opened because there are always new agents who seek to use the ledger.

Each period, the ledger collects T_ℓ from new late producers who are opening accounts and T_v from new validators who are opening accounts. These collections are redistributed lump-sum to respectively (continuing) late producers and (continuing) validators.[27] Since there is no reason for a producer to open more than one account at a time, and for brevity, we will sometimes refer to the "standing of producer i" when referring to the "standing of the account held by producer i."

To simplify the analysis, we assume early producers can read the account label of late producers in the ledger, but they cannot write the outcome of the match on the ledger. Therefore the validation process takes place in the late stage only.[28]

The validation process consists in validators 1) verifying that late producers have produced according to plan, i.e. $\tilde{y} = y$,[29] and 2) sending a message to the ledger – if production took place according to plan and the late producer's account has label G, validators will communicate G, but B otherwise. Consensus is reached on the new label which is recorded

[26] See Schneider and Taudien (2023) for a model of reputation when agents use pseudonyms.

[27] In equilibrium, $T_\ell = (1-\sigma)(1-V)\gamma$ and $T_v = (1-\sigma)V\gamma_v$. Therefore the $\sigma(1-V)$ late producers who do not exit the ledger will receive $T_\ell/(\sigma(1-V)) = (1-\sigma)\gamma/\sigma$, while the σV validators who do not exit will receive $(1-\sigma)\gamma_v/\sigma$.

[28] We also worked out a version of the model in which early producers cannot read the ledger. In this case, validators inform the early producers about the late producer's label. The changes are only cosmetic.

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

on the ledger whenever at least a fraction $\tau \in [0,1]$ of the V validators communicate the same label, i.e. cast the same vote. We emphasize that τ is a choice variable when designing the consensus protocol.[30]

Validators incur verification and communication costs. Validators incur an additive utility cost $c_v \geq 0$ to verify a late producer's action and an idiosyncratic additive utility cost $c_{s,i}$ 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 when they want to communicate that the late producer's label is G.[31]

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

$$c_{s,i} = c_s + \mu_i, \tag{1}$$

where c_s is a common component to all validators uniformly distributed over $[\underline{c}_s, \bar{c}_s]$, while μ_i is the idiosyncratic element for validator i that 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. . Validators learn their cost ahead of the verification game and so ahead of verifying the label.

To bring agents to become validators, they must make a positive expected profit from the validation process. Sending a message is verifiable, so validators who correctly sent message G are entitled to z units of the late good whenever enough validators agree that the label is G. Validators receive nothing if they do not cast a vote. Validators value these transfers in

[30] 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.

[31] For tractability, we allow the communication costs to be negative so that we can use symmetric distributions with no mass point. The results would also obtain when we use distributions with a mass point at zero. We assume validators do not incur costs in the second stage. It is straightforward to extend the model to analyze this case too.

an additive and linear way.[32] When a fraction $w \geq \tau$ of validators send label G at the end of stage 2, the late producer has to produce wVz to compensate these validators. Then these validators update the ledger.[33]

Notice that validators play a game that is similar to a public good contribution game: They get a reward only if a sufficient number of validators confirm a trade. Given the structure of the validation costs, we solve this game using tools from the literature on global games. 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. Although remote, the possibility of computer glitches will imply that validation should not rely on unanimous agreement when there are many validators. In particular, the equilibrium in the limiting case as $\varepsilon \to 0$ gives rise to some degree of coordination failure and associated inefficiencies in contrast to the (common knowledge) case when $\varepsilon = 0$, which is typically associated with multiple equilibria. In the sequel, we limit our attention to the limiting case of $\varepsilon \to 0$.

Finally, we assume late producers can bribe validators to send a false message: A late producer whose account has label G when starting the period may get away with not repaying the early producer while keeping its label by "bribing", i.e. making a side payment to τV validators in the late stage (after consuming in the first stage).[34] Validators who accept a

[32] A linear utility function allows us to abstract from possible insurance mechanism among validators. Also, at the cost of simplicity, we could assume that the utility of validators is $u(x + z)$. Finally, we could assume that only a share τ of validators receive a reward since only this number is necessary to reach an agreement.

[33] We do not explicitly model the updating process. An intuitive narrative is that each validator updates his own copy of the ledger and sends it to all other validators. Validators compare all copies and coordinate on adopting the copy that has been sent the most times and at least a number τ of times.

[34] A late producer who misbehaved at 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 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. Bribing could be interpreted as the attempt to coordinate a sufficient mass of users on a fork or on restarting the ledger.

bribe are caught with probability π in which case their current account loses the privilege to trade and validate.[35] However, those validators can open a new account allowing them to trade and/or validate in the future.

Finally, we assume late producers, including validators, have the option to switch to another platform. We assume the benefit from trading on the platform is a fraction $\eta \leq \beta$ of their current expected payoff. This is similar to assuming that it takes time to switch platform, and that the role of the agent on the new platform is the same as the one on the current platform.[36] In Online Appendix F we consider an extension where the outside option is not related to the current expected payoff and agents can choose between being validators or just late producers. These changes do not affect the qualitative results.

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 of account fees and production/consumption $(\gamma, \gamma_v, x, y, z)$. 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' account is correctly communicated to the ledger, and validators have no incentive to tamper with the record of labels.

Given a stationary incentive-feasible allocation $(\gamma, \gamma_v, x, y, z)$, U_i is the expected discounted lifetime utility of late producer i with label G, satisfying

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

where $\mathbb{E}_i$ is the expectation operator of late producer i over w_i, the share of working valida-

[35] We assume it is costless for late producers to bribe validators because if the latter reject it, late producers can always revert to paying early producers as planned.

[36] A validator remains a validator, maybe thanks to a prior technological investment.

tors. The late producer is entitled to the transfer T_ℓ. In addition he gains $u(x) - y$ from trading with probability α. The late producer's account retains label G only if $w_i \geq \tau$ validators validate the trade. In this case, the late producer pays $w_i V z$ to the working validators and can trade in the future. With probability $1 - \alpha$, the late producer does not trade but retains his label into next period. The expected discounted lifetime utility of an internal validator i with private communication cost $c_{s,i}$ is $U_{IV}(c_{s,i})$ and satisfies

$$\begin{aligned} U_{IV}(c_{s,i}) &= T_{IV} + \mathbb{E}_i \left\{ \alpha \left[u(x) - y + \mathbb{I}_{w_i \geq \tau} \left\{ -w_i V z + \beta \mathbb{E} U_{IV} \right\} \right] + (1-\alpha) \beta \mathbb{E} U_{IV} \mid c_{s,i} \right\} \quad (2) \\ &\quad + \alpha \max \left\{ 0; \mathbb{E}_i \left[-c_v + (\mathbb{I}_{w \geq \tau} z - c_{s,i}) \mid c_{s,i} \right] \right\} \quad (3) \end{aligned}$$

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 $c_{s,i}$.[37] Internal validators receive transfer T_{IV}. Since internal validators are selected from the set of late producers, they also obtain the expected payoff of late producers. In addition they get the expected payoff from validating trades: 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 communicate that the late producer has produced according to his announcement they incur the communication cost $c_{s,i}$, and get the reward z, but only when the trade is validated ($\mathbb{I}_{w \geq \tau} = 1$). Otherwise they do not get the reward z.

Similarly, the expected discounted lifetime utility of an external validator i with private communication cost $c_{s,i}$ is $U_{EV}(c_{s,i})$ and satisfies,

$$U_{EV}(c_{s,i}) = T_{EV} + \alpha \max \left\{ 0; \mathbb{E}_i \left[-c_v + (\mathbb{I}_{w \geq \tau} z - c_{s,i}) \mid c_{s,i} \right] \right\} + \beta \mathbb{E} U_{EV}, \quad (4)$$

where T_{EV} is the transfer to external validators.

[37] Validators have more information concerning the fundamental communication cost c_s, which they use to compute the probability that the trade be validated.

Participation constraints. An allocation $(\gamma, \gamma_v, x, y, z)$ satisfies the participation constraints of validators, early and late producers whenever,

$$\begin{aligned} -\gamma + U &\geq \eta U && (5) \\ -\gamma_v + U_v &\geq \eta U_v && (6) \\ y - x &\geq 0 && (7) \\ -c_v + \mathbb{E}_{w \geq \tau | c_{s,i}} z - c_{s,i} &\geq 0 && (8) \end{aligned}$$

Constraints (5) and (6) require that late producers and validators are better off paying the account fee to use that ledger rather than using the competing ledger offering them ηU. (7) is the (simplified) participation constraint of the early producers. Finally, since late producers can shirk from validation, (8) requires that validators expect to make a positive expected profit from the validation process. The expectation operator in (8) applies to the share of working validators w.

Repayment constraints. Using a ledger, in equilibrium late producers who do not produce the announced amount y are detected by validators and their account is thus assigned a label B. As a consequence, this account is permanently blocked from all economic activities. Similarly, an internal validator whose account is in default loses its future right to validate and their account is also blocked. Agents whose account has been blocked can open a new late producer's account at cost γ, a validator's account at cost γ_v, or obtain the payoff ηU by using the competing ledger. Therefore, given the share of working validators is $w \geq \tau$, the repayment constraint of late producers and internal validators is respectively

$$\begin{aligned} -(y + wVz) + \beta U &\geq \beta \max\{-\gamma + U; -\gamma_v + \mathbb{E}U_{IV}; \eta U\} && (9) \\ -(y + wVz) + \beta \mathbb{E}U_{IV} &\geq \beta \max\{-\gamma + U; -\gamma_v + \mathbb{E}U_{IV}; \eta U\}. && (10) \end{aligned}$$

No bribe. If a validator accepts a bribe, we assume she is caught with probability $\pi \in [0,1]$. In this case she loses her right to validate future transactions and to consume as a late producer using her old account. Yet again, she can open a new account. A validator prefers recording the truth to a false record when the late producer offers $\bar{z}$ iff

$$\beta \mathbb{E}U_s \geq \underbrace{\bar{z}}_{bribe} + (1-\pi)\beta \mathbb{E}U_s + \pi\beta \max\{\mathbb{I}_{s=IV}\left(-\gamma + U\right); -\gamma_s + U_s; \eta U_s\}$$

where $s = IV, EV$ to denote internal of external validation, respectively. 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 $y + wVz$ to get away with production. Given that the ledger requires the agreement of at least τV validators to validate a (fake) transaction, the cheating late producer will pay $\bar{z} = (y + wVz)/(\tau V)$ to τV validators. Therefore a validator rejects the bribe whenever[38]

$$\pi\beta \mathbb{E}U_s \geq \frac{1}{\tau V}\left(y + wVz\right) + \pi\beta \max\{\mathbb{I}_{s=IV}\left(-\gamma + U\right); -\gamma_s + U_s; \eta U_s\}. \tag{11}$$

Finally, we need to identify a condition such that, given a threshold τ, the allocation allows for a truthful record of the ledger.

Validation threshold. Suppose there is a positive measure of validators in charge of verifying each 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. 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 jeopardize the validation process.

We assume the continuation payoff of validators is independent of the current validation

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

results.[39] Also if validators do not work, they do not send any messages thus eliminating a possible free rider problem. We relax these assumptions in Online Appendix C, where we examine an extension that also specifies the voting game that validators play, 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 signal that the communication cost is small. 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 satisfies*

$$z \geq \frac{c_s + c_v}{1 - \tau}. \tag{12}$$

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 $c_{s,i}$ is 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

[39] For example, it is difficult to distinguish whether a validator shirked or his message technically failed to reach other validators. However, see Green and Porter (1991) and Monnet and Quintin (2021) show how punishments followed by forgiveness may discipline agents who can hide behind the veil of "bad luck."

the event that a fraction q of the V validators will work. Since this validator is indifferent between working or not, and his subjective beliefs of the share of working validators is uniform, i.e. $g(\tau \mid c_s^*) = 1$, c_s^* solves

$$-c_v + [(1-\tau)(z - c_s^*) + \tau(-c_s^*)] = 0.$$

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 logically affects the threshold value c_s^*: By increasing the validator's rents z, 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.[40]

As a corollary, notice that the payment z to validators 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. Finally, given τ, the probability that the trade will go through when c_s is uniformly distributed is the probability that

$$c_s \leq (1-\tau)z - c_v \equiv c_s^*.$$

If the ledger is required to allow *all* legitimate trades involving a producer with label G will

[40]Morris and Shin (2003) show 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.

always go through, then z should be set to

$$z \geq Z(\tau) \equiv \frac{1}{1-\tau}\left(\bar{c}_s + c_v\right) \equiv \frac{C}{1-\tau} \tag{13}$$

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 Online Appendix D, we relax this assumption and we analyze the optimal validation protocol for any thresholds c_s^* 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 $(\gamma, \gamma_v, x, y, z)$ that satisfies (5)-(11) and (13).

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

Limiting distribution of communication costs

From now we consider the global game limiting case where $\varepsilon \to 0$, and (13) holds. In this limit, all validators face the same communication cost equal to $\bar{c}_s$. So unless they are bribed, all validators 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 is equal to one and we can write (5)-(12) with $w_i = 1$. We can further simplify the set of IF allocations by setting the participation constraint of early producers (7) at equality.

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. Since the payment to validators should be minimized, (13) binds so validators earn $Z(\tau)$ and the participation constraint of validators (8) is always satisfied. Also, the incentive constraints of late producers and validators are relaxed when the cost of opening an account is set as high as possible, that is $\gamma = (1-\eta)U$ and $\gamma_v = (1-\eta)U_v$. Then denote the expected rent of validators as $R(\tau)$,

$$R(\tau) \quad = \quad \tau Z(\tau). \tag{14}$$

Define the default factor as (recall that the effective discount factor is $\beta = \sigma\tilde{\beta}$)

$$\delta \quad \equiv \quad \frac{\pi\sigma\tilde{\beta}(1-\eta)\alpha}{1-\sigma\tilde{\beta}-(1-\eta)(1-\sigma)/\sigma},$$

and replace the expressions for the transfers, to find the set of IF allocations characterized by[41]

$$\frac{\delta}{\pi}\left[u(x)-x-VZ(\tau)\right] \quad \geq \quad x+VZ(\tau) \tag{15}$$

$$\delta\left[u(x)-x-VZ(\tau)+R(\tau)\right] \quad \geq \quad \frac{1}{\tau V}\left(x+VZ(\tau)\right) \tag{16}$$

The higher δ is, the weaker 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 future income (as captured by a high $\tilde{\beta}$), or the outside option is really poor (η is low). Also, the default factor can be increasing or decreasing in the survival rate σ. There

[41] Since we set the participation constraint for early producers (7) at equality, we can use $y = x$. Also, since validators earn a rent, $U_{IV} \geq U$ and (10) is satisfied whenever (9) is.

are two counteracting effects from a higher survival rate: On one hand, agents effectively become more patient, which improves incentives, but on the other hand a higher survival rate decreases the transfer that surviving agents get from the ledger, which weakens discipline. Which effect dominates depends on the strength of the outside option: if η is low enough the default factor falls with a higher σ.

According to constraint (15), 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 losing the expected lifetime discounted payoff from trading net of the payment to validators. Constraint (16) states the condition for validators to reject a bribe: the payoff from accepting the maximum bribe a late producer offers, $(x + VZ(\tau))/\tau V$, must be lower than 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 lose the expected rents they earn $R(\tau)$. With external validation, (16) simplifies to

$$\delta R(\tau) \geq \frac{1}{\tau V}\left(x + VZ(\tau)\right) \tag{17}$$

Inequality (17) requires that losing future validation rents with probability π should be a larger cost to external validators than the benefit of accepting the largest bribe a late producer would offer.[42]

We next focus on the number of validators V. The more validators there are, the higher the overall payment for validation – this is the term $VZ(\tau)$ on the left-hand side of (15). However, given τ, more validators also means a lower bribe for each validator, thus relaxing their incentive constraint as shown on the RHS of (16) or (17). The optimal number of

[42] Our mechanism that requires a supermajority τ of *all* validators implies that producers can offer a smaller bribe relative to a mechanism that would base consensus on a supermajority of *voting* validators (instead of all validators). Also, using a mechanism relying only on voting validators may be problematic because validators cannot be induced to vote when some producers are faulty as we consider in Online Appendix C, as then they do not expect any rewards.

validators will trade off both effects. When the measure of validators is large, the repayment constraint of late producers becomes the binding one when $\pi \to 1$, 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 lose the expected lifetime payoff from trading. Then, only the incentives of the late producer matter. However, note that if (15) is binding while (16) is slack, V can be reduced up to the point where (16) binds. We summarize the discussion above in the following result.

Lemma 1. *If π and V are large enough, validators will never accept a bribe.*

Optimal design

The objective of a planner is to maximize the gains from trade net of the validation costs. A planner chooses the costs to open accounts γ and γ_v, the trading size x, the type of validation (internal or external), the number of validators and the threshold τ to solve

$$\alpha \max_{\gamma,x,V\geq 0,\tau\in[0,1]} \{u(x) - x - VC\} \tag{18}$$

subject to (15) and (16) for internal validation or (17) for external validation. These constraints have been simplified by using $\gamma = (1-\eta)U$ and $\gamma_v = (1-\eta)U_v$. But this is without loss of generality: the cost to open the account is redistributed lump sum to all agents in the same period it is levied on late producers. Since this is a mere redistribution, it does not affect the objective of the planner. Also the higher the lump sum transfers to agents, the more relaxed their incentive constraints. Therefore, as we guessed above, the cost to open an account can be set as high as possible, that is as high as the lifetime benefit of trading on the ledger net of the available outside option. Conveniently, the size of the payment to validators $Z(\tau)$ only matters for incentives but has no impact on the objective function because it is

a transfer between producers and validators. Also note that whether validation is internal or external does not affect the social planner's objective function (18). However, comparing (16) and (17) shows that as long as trade is sufficiently beneficial, or $u(x) > (x + VZ)$, then (16) is always satisfied whenever (17) is. In other words, the set of incentive feasible allocations is larger with internal validation, and internal validation weakly dominates external one. But not only this. As we show below, whenever trade can be supported, centralization is optimal with external validation. Next, we start by considering the optimal design with external validation.

Optimal design with external validation. With external validation the only relevant constraints are (15) and (17), which we can write respectively, as

$$\delta u(x) \geq (\delta + \pi)\left[x + V\frac{C}{1-\tau}\right] \tag{19}$$

$$(1-\tau)x \leq \left(\delta\tau^2 - 1\right)VC \tag{20}$$

It is easy to show that the validators' incentive constraint (20) always binds,[43] so we can use it to replace for the total cost of validation, VC in the late producer's repayment constraint (19) and the objective function. The problem of the planner then becomes

$$\alpha \max_{x,\tau\in[\bar{\tau},1]} \left\{u(x) - \frac{(\delta\tau^2 - \tau)}{(\delta\tau^2 - 1)}x\right\}$$

subject to

$$\delta u(x) \geq (\delta + \pi)\frac{\delta\tau^2}{(\delta\tau^2 - 1)}x$$

[43]Suppose it does not at the solution $(\tilde{x}, \tilde{\tau}, \tilde{V})$. Then reduce V until it does. This increases the objective function (18), while relaxing the participation constraint of late producers. So $(\tilde{x}, \tilde{\tau}, \tilde{V})$ could not be the solution, a contradiction.

Inspecting (20), notice that a necessary and sufficient condition for the existence of external validation is that intertemporal incentives are strong enough, in the sense that $\delta \geq 1$. Otherwise, the only incentive feasible allocation is autarky with $V = 0$. With $\delta > 1$, it is straightforward to verify that the constraint is relaxed when τ is highest. Also, the objective function is maximized when $\tau = 1$. Therefore, with external validation, the solution is $\tau = 1$ and $V \to 0$, and the optimal trading size $\tilde{x}(\delta) \leq x^*$ where $\tilde{x}(\delta) = x^*$ if $\delta \geq \delta^* > 1$ so that the constraint is not binding or it is defined as the solution to $u(\tilde{x}) = \tilde{x}(\delta + \pi)/(\delta - 1)$ otherwise, with $\tilde{x}(\delta) \to 0$ as δ decreases to 1. The payment to the validator is $VZ \to \tilde{x}(\delta)/(\delta - 1)$.

Optimal design with internal validation. Next, consider the optimal design with internal validation. As we show in the appendix, the social planner's objective function (18) is decreasing in V, and so we obtain

Lemma 2. *With internal validation, the incentive constraint of validators (16) always binds while the incentive constraint of late producers (15) never binds.*

Replacing the expression for the validation rent (14) in (16), and re-arranging, we obtain:

$$\delta \left[u(x) - x\right] \geq \frac{x}{\tau V} + \frac{C}{1-\tau}\left[1 - \delta(1-V)\tau\right] + \delta C \tag{21}$$

Since the objective function (18) is independent of τ, the planner chooses τ to minimize the right hand side of (21). Two forces are at play. On the one hand, increasing τ reduces the maximum bribe size per validator. On the other hand, increasing τ increases the payment to validators $Z(\tau)$ to ensure that validators indeed verify and validate. When intertemporal incentives are strong, so that $1 \leq \delta\,(1 - V)$, this second effect reduces the right-hand side of (21): Validators have much to lose by accepting a bribe. In this case, as with external validation, it is optimal to set $\tau = 1$, even if $Z(\tau) \to \infty$. Alternatively, suppose intertemporal

incentives are weak, $1 > \delta(1-V)$, then the optimal $\hat{\tau}$ trades-off the lower bribe size with the higher payment to validator and it solves[44]

$$\frac{1-\hat{\tau}}{\hat{\tau}} = \sqrt{\frac{[1-\delta(1-V)]VC}{x+VC}} \tag{22}$$

The optimal threshold $\hat{\tau}$ is decreasing in V but increasing in x. $\hat{\tau}$ is also increasing in π, β or α, as captured by δ. The intuition is apparent from (15): When π, β or α increase, the net rent $R(\tau)-VZ(\tau) > 0$ of validators becomes more important for the incentives of validators relative to the bribe size $(x+VZ)/\tau V$, either because they have a higher chance of losing it – when π increases – or because they have a higher lifetime discounted value – when β or α increases. So following an increase in δ, the planner gives more net rent to validators by increasing τ. So unlike in traditional models of limited commitment, higher trustworthiness as captured by a higher δ implies more rents to validators.

Importantly, while external validation only implements autarky when intertemporal incentives are weak, internal validation can do much more. This is intuitive: With internal validation, the planner can use the value of trading in the future to discipline validators, thus allowing the planner to choose lower rents (and therefore a lower feasible bribe) validators obtain from being able to manage the ledger.

The optimal choice of x and V trades off several effects. First, increasing x toward x^* brings additional gains from trade, but at the cost of increasing the bribe size per validator which tightens their incentive constraint. Second, increasing V relaxes the incentive constraint of validators, but increases the cost of validation. We can now summarize these considerations in our main result.

Proposition 2. *The constrained optimal solution $(\hat{x}, \hat{V}, \hat{\tau})$ solves (21) at equality and (22)–*

[44]It is easy to verify that when $1 > \delta(1-V)$, the second order condition for a minimum is satisfied.

(26) and is characterized by four regions:

1a. [centralized system - efficient trade size] *If $\delta \geq \delta^* > 1$, external and internal validations are characterized by $\hat{V} \to 0$, $\hat{\tau} \to 1$ and $\hat{x} \to x^*$. Validation requires arbitrarily large payments, i.e. $Z(\hat{\tau}) \to \infty$, while $\lim_{\hat{\tau}\to 1} V(\hat{\tau})Z(\hat{\tau}) = \frac{x^*}{\delta - 1}$. Welfare is higher with internal validation.*

1b. [centralized system - inefficient trade size] *If $1 \leq \delta \leq \delta^*$, external validation is characterized by $\hat{V} \to 0$, $\hat{\tau} \to 1$ and $\hat{x} \to \tilde{x}(\delta) \leq x^*$. Internal validation is characterized by $\hat{V} > 0$, $\hat{\tau} < 1$ and $\hat{x} < x^*$. Welfare is higher with internal validation.*

2. [partially distributed system] *If $\bar{\delta} < \delta \leq 1$, only internal validation can decentralize trade. 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 $\delta_0 < \delta \leq \bar{\delta}$, only internal validation can decentralize trade. All late producers are validators $\hat{V} = 1$, and $\hat{\tau} = \left(1 + \sqrt{\frac{C}{x+C}}\right)^{-1}$. 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.*

Proposition 2 states that the optimal design of the ledger requires centralized validation by a single validator only when validators are sufficiently trustworthy. However, moving toward centrality necessarily imposes a move toward unanimity. In the limit, this requires an arbitrarily large payment to the single authority managing the ledger. In reality, feasible payments may be bounded and in such a case, a single validator will never be optimal. Still, the single validator case offers a useful benchmark that illustrates the forces driving the solution toward centrality. When intertemporal incentives become weak, the single authority can be more easily convinced to do wrong and the optimal validation protocol moves away from centrality. The centrifugal forces manifest themselves also in a reduction in trade size,

and a departure from unanimity or a high supermajority rule. Both margins reduces the feasible bribe to validators and therefore their incentives to do wrong.[45]

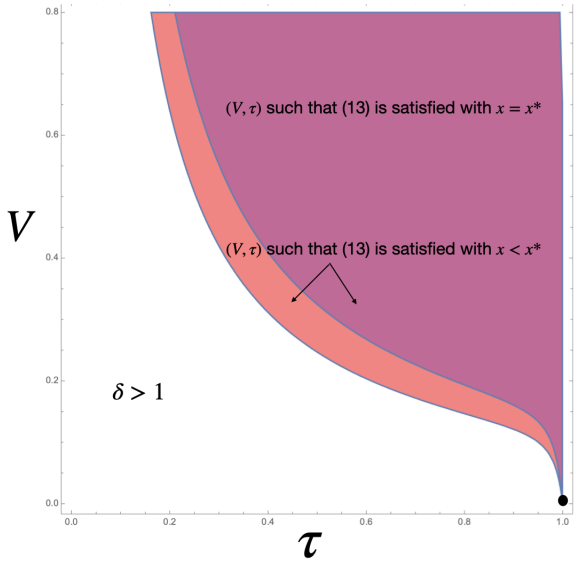
Underlying the results in Proposition 2 are the following comparative statics (details of the calculations are in the appendix): x and τ are (weakly) increasing with δ (i.e. β, π, and α) but decreasing with validation costs C. V is (weakly) decreasing with δ but increasing with validation costs C.

As the level of patience or frequency of trading among validators increases (represented by a higher δ), the optimal number of validators (V) decreases while the optimal trade size (x) increases. This is intuitive because when validators are more patient or trade more frequently, they have more to lose from any wrongdoings. Therefore, they are less likely to engage in such behavior, as δ increases. Then the planner can increase the size of each trade and reduce the number of validators needed. This choice of V and x indirectly affects τ and reinforces the planner's decision to increase the net rent for validators by choosing a larger threshold, as we already explained above. All in all, these effects work together to increase the threshold τ as δ increases. It is also worth noting that as the system becomes fully distributed with only one validator ($V = 1$), the supermajority threshold will converge to simple majority (τ approaching $1/2$) as the trade size approaches zero, even though there is no constraint requiring τ to be greater than or equal to $1/2$.

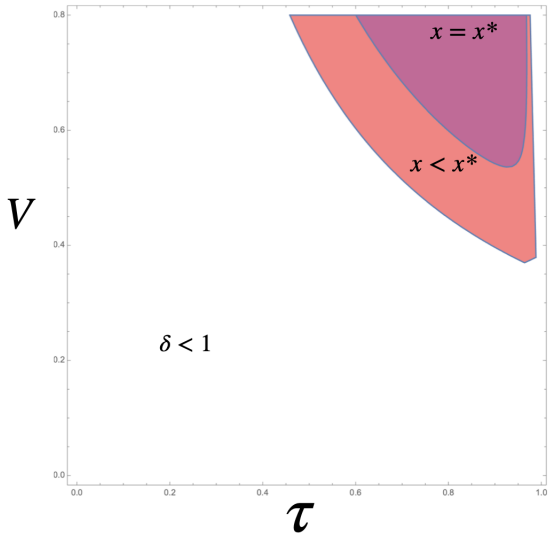
Centrifugal forces also include the cost of validation C. This may be surprising as increasing the number of validators also increases duplication costs to the detriment of social welfare. However, validation costs are reducing the overall lifetime discounted surplus of validators, who, as a result are more easily convinced to do wrong. Then it is optimal to increase the number of validators so that (given x) each of them can only be offered a smaller bribe.

[45]Put another way, the safeguards necessary for the ledger's integrity makes decentralized consensus expensive and reduces trade size. In short, the ledger is hard to scale. This conundrum introduces trade-offs sometimes known as the ledger's "scalability trilemma" (see ie Buterin (2021)). The trilemma is posed in terms of the challenge of attaining a ledger that is simultaneously decentralized, secure, and scalable.

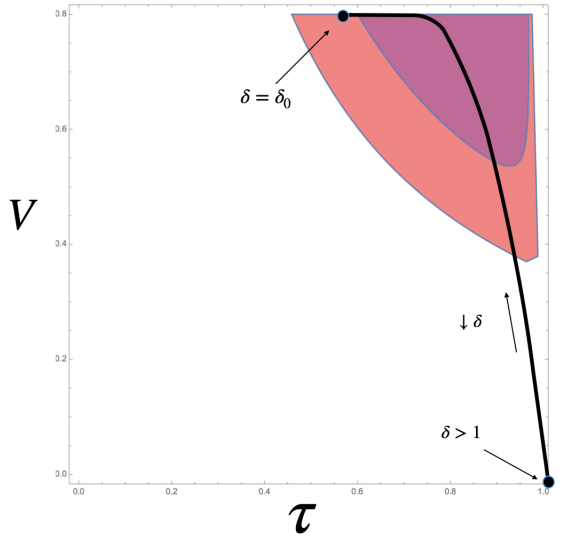
Reaching consensus with a higher number of validators however does not come for free: As validation cost C is higher, validators are more likely to believe that fewer other validators will work. Hence maintaining the same level of consensus requires a larger rent be paid to each validator as Proposition 1 shows. To maintain the legitimacy of the ledger while keeping costs in check, the threshold τ should fall.



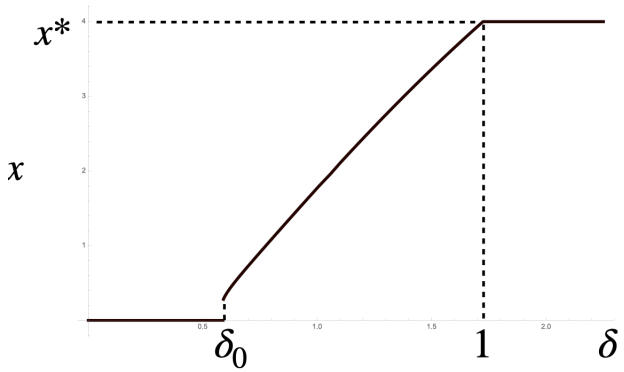
(a) $\delta > 1$, IF choice of (V, τ) for $x = x^*$, $x < x^*$.



(b) $\bar{\delta} < \delta < 1$ IF choice of (V, τ) for $x < x^*$, $x = x^*$



(c) Optimal solution for (V, τ) as δ falls



(d) Optimal solution for x

Figure 2: Incentive feasible choice of (V, τ)

Figure 2 illustrates Proposition 2 setting $\eta = 0$ so that the value of the outside option is zero. Given some $x \leq x^*$ and the default factor δ, Figure 2 shows the set of (V, τ) for which the IC of validators (21) is satisfied. When δ is relatively high, $x = x^*$ and $(V, \tau) = (0, 1)$ satisfy (21) and this is the best design for the system. In this case, the solution is given by the black dot in the lower right corner of Figure 2a. In contrast, Figure 2b shows incentive feasible allocations when δ is relatively smaller: Then the allocation $x = x^*$ and $(V, \tau) = (0, 1)$ is no longer incentive feasible. The figure shows that implementing $x = x^*$ is feasible for some (V, τ), but only for relatively large V. This is costly and the planner does better by choosing a lower x which decreases the bribe size and allows it to select fewer validators (lower the number of validators V), as shown by the red area in Figure 2b. Reducing x slightly below x^*, the planner only makes a second order loss, but realizes a first order gain as V decreases, thus reducing the overall validation costs. Therefore, as δ decreases, x declines and (V, τ) moves up along the black curve, as shown in Figure 2c. As δ falls below the threshold $\bar{\delta}$ toward δ_0, $V = 1$ and τ moves westward on the black line until it reaches $1/2$ and $x = 0$. For levels of δ below δ_0, there is no equilibrium with trade. Finally, as we move to the left along the black curve, the trade size x decreases toward zero.

4 Conclusion

In this paper, we have presented an economic analysis of decentralized ledger technology in an economy where money is essential. To our knowledge, our analysis is the first economic analysis of 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 supermajority voting rule, number of validators, and size of transactions.

We believe our analysis is a timely one, as 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 and stablecoins insofar as it gives conditions under which a central authority should manage the ledger of transactions.[46] 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.[47] 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 a societal record-keeping device used as a substitute for money.

Of course, we have made simplifying assumptions in order to better grasp the basic economics of money as memory. Future work should relax some of these.

For instance - 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.

References

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

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

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

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.

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].

Berentsen, Aleksander and Fabian Schaer (2020) "Bitcoin, Blockchain, and Cryptoassets: A Comprehensive Introduction," MIT-Press, Cambridge, Massachusetts.

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.

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

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology,

and governance. Journal of Economic Perspectives, 29(2), 213-38.

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.

Buterin, V (2014a): "Proof of stake: how I learned to love weak subjectivity", blog.ethereum.org, 25 November.

Buterin, Vitalik (2021) "Why sharding is great: demystifying the technical properties", available at https://vitalik.ca/general/2021/04/07/sharding.html

Caldarelli, Giulio (2020) Understanding the Blockchain Oracle Problem: A Call for Action. Information, 11, 509; doi:10.3390/info11110509

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

Carlsten, M., Kalodner, H., Weinberg, S. M., & Narayanan, A. (2016, October). On the instability of bitcoin without the block reward. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (pp. 154-167).

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.

Chiu, Jonathan and Thorsten Koeppl (2022) "The Economics of Cryptocurrencies - Bitcoin

and Beyond," Canadian Journal of Economics 55(4), 1762-1798, https://doi.org/10.1111/caje.12625.

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

Cong Lin William, Zhiguo He and Jiasun Li (2021) "Decentralized Mining in Centralized Pools," *The Review of Financial Studies* 34(3), 1191–1235, hhaa040, https://doi.org/10.1093/rfs/hhaa040.

Deirmentzoglou, E., Papakyriakopoulos, G., & Patsakis, C. (2019). A survey on long-range attacks for proof of stake protocols. IEEE Access, 7, 28712-28725.

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

Easley, D., O'Hara, M., & Basu, S. (2019). From mining to markets: The evolution of bitcoin transaction fees. Journal of Financial Economics, 134(1), 91-109.

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

Fanti, Giulia, Leonid Kogan, and Pramod Viswanath (2021) "Economics of Proof-of-Stake Payment Systems," MIT Sloan Research Paper No. 5845-19.

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

Froewis, Michael and Rainer Boehme (2017) "In code we trust? Measuring the control flow immutability of all smart contracts deployed in Ethereum," in J Garcia-Alfaro, G. Navarro-Arribas, H Hartenstein, and J Herrera-Joancomarti (eds), Data privacy management, cryp-

tocurrency and blockchain technology, Springer, pp. 357-72.

Garratt, Rodney and Cyril Monnet (2023) "An impossibility theorem on truth-telling in fully decentralised systems," Bank for International Settlements Working Paper 1117.

Garratt, 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.

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 Pauzner (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.

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

Halaburda, Hanna, Zhiguo He, and Jiasun Li (2021) "An Economic Model of Consensus on Distributed Ledgers," mimeo University of Chicago.

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*,

forthcoming.

Hurwicz, Leonid (2008) "But Who Will Guard the Guardians?" American Economic Review 98, 577-85.

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.

Luu, Loi, Duc-Hiep Chu, Hrishi Olickel, Prateek Saxena and Aquinas Hobor (2016) "Making smart contracts smarter," in Proceedings of the 2016 ACM SIGSAC Conference on Computer

and Communications Security, pp. 254-69.

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

Monnet, Cyril and Erwan Quintin (2021) "Optimal Financial Exclusion," *American Economic Journal: Microeconomics 13, 101-34.*

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.

Nakamoto, Satoshi (2008) "Bitcoin: A Peer-to-Peer Electronic Cash System," https://bitcoin.org/bitcoin.pd

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.

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 (2020), ISBN electronic: 9780262361194.

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.

Xu, Xiwei, Cesare Pautasso, Liming Zhu, Vincent Gramoli, Alexander Ponomarev, An Binh Tran, and Shiping Chen (2016) "The blockchain as a software connector," in Proceedings of the 13th Working IEEE/IFIP Conference on Software Architecture (WICSA), pp. 182-191.

Appendix

In this Appendix containing all proofs, we assume that there is a fraction f of late producers who do not have the ability to produce any goods. We refer to these producers as "faulty" producers. In turns they are only necessary to eliminate the free riding problem facing validators that we consider in section C. Otherwise, all other results go through with $f = 0$. For completeness we have written the proof with $f > 0$.

A Proof of Proposition 1

Proposition. *[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 satisfies*

$$1 - \tau \geq \frac{c_s + c_v/(1-f)}{z} \tag{23}$$

Assume that each validator receives private cost

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

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

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

where $m_i = 1$ if validator i sends label G (recall that validators do not send a message when the account label is B). The expected payoff of a validator is given by (35) when the measure

of validators sending a message is higher than τV. We can rewrite the expected payoff as

$$\beta U_V + (1-f)z \times \begin{cases} -\frac{\tilde{c}_v^1}{1-f} - \tilde{c}_{s,i} + 1 & \text{if } \int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di > \tau V \\ -\frac{\tilde{c}_v^1}{1-f} - \tilde{c}_{s,i} & \text{otherwise} \end{cases} \tag{24}$$

where we have 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 the common cost component c_s is uniformly distributed. Note 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 gives the probability that a trader with label G is able to trade and compensate validators, and 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 information $\tilde{c}_{s,i}$ 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 $C^* \equiv \frac{\tilde{c}_v}{1-f} + \tilde{c}_s^*$. Since the normalized cost $\tilde{c}_s$ 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 $\left[\frac{\frac{c_v^1}{1-f} + \gamma_s - \varepsilon}{z}, \frac{\frac{c_v^1}{1-f} + \gamma_s + \varepsilon}{z}\right]$. Since all validators with $C_i < C^*$ are working, the measure of working validators is

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

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

$$c_s(q) = c_s^* + \varepsilon - 2\varepsilon q$$

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

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

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

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

Hence, as $\varepsilon \to 0$, all validators with private 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 (24) 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

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

and validation will always go through whenever

$$1 - \tau \geq \frac{\bar{c}_s + c_v/(1-f)}{z}.$$

B Proof of Lemma 2

Proof. We first show that (16) must bind. Suppose it does not. The objective function is decreasing in V and (15) is relaxed with lower V. So if (16) 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 $\delta [u(x) - x] \geq \pi x + \frac{(1-\beta)}{\alpha}\delta \bar{U}$. However, since $\tilde{x} > 0$ and $\tau \leq 1$, it is clear that (16) cannot be satisfied when $V = 0$. Therefore (16) must bind. Suppose now both (15) and (16) bind. Then, given τ, the solution is given by

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

Using these two equations we can write the objective function of the planner, as

$$\alpha \left\{ u(x) - x - V \left(Z(\tau) - R(\tau) \right) \right\} = \alpha \left\{ \pi \frac{\tau}{1 - \tau V} + 1 \right\} V \tau \left(\frac{\bar{c}_s + c_v}{1 - \tau} \right).$$

This is strictly increasing in both V and τ. Hence the solution is $V = 1$ 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 $\alpha \left\{ u(x) - x - V \left(\bar{c}_s + c_v \right) \right\} \to -\infty$, which cannot be optimal. Therefore, (15) and (16) cannot both be binding. This shows that only (16) binds. □

C Proof of Proposition 2

Proof. First we show that internal validation does at least as well as external validation. Both problems have the same objective function. The constraints for external validation are

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

while the constraint for internal validation is

$$\delta \left[u(x) - (x + Vz) + R(\tau) \right] \geq \frac{1}{\tau V} \left(x + Vz \right)$$

since $u(x) - (x + Vz) \geq 0$, and $\pi < 1/(\tau V)$, the constraint set for internal validation is weaker than the one for internal validation. This shows the claim.

Next we concentrate on the results for internal validation (the one for external validation is in the text). The first order conditions with respect to x (25) and the one with respect to V

(26) are

$$[u'(x)-1](1+\delta\lambda)-\lambda\frac{1}{\hat{\tau}V} = 0 \quad (25)$$

$$\left[\frac{R(\hat{\tau})}{1-f}-Z(\hat{\tau})\right]+\lambda\frac{x}{\hat{\tau}V^2}-\delta\lambda Z(\hat{\tau})+\lambda_0-\lambda_{1-f} = 0 \quad (26)$$

where λ is the Lagrange multiplier on the validators' IC constraint, and λ_{1-f} is the one on $V \leq 1-f$ and λ_0 is the one on $V \geq 0$.

We first consider the solution when $\hat{\tau} \to 1$. Then $Z(\hat{\tau}) \to +\infty$ and

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

We now show that λ and V both converge to zero. To the contrary, suppose $V>0$ ($\lambda_0=0$) and $\lambda>0$. Since $Z(\hat{\tau}) \to +\infty$ as $\hat{\tau} \to 1$, the LHS of (26) is necessarily negative as $\hat{\tau} \to 1$. Hence $\lambda_0>0$ and/or either $\lambda \to 0$ to reestablish the equality, which contradicts $V>0$ and $\lambda>0$. Hence either $V=0$ or $\lambda=0$ or both. . Starting with $V \in (0,1-f)$, we can rewrite (26) 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 (15) never binds, it must be that $VZ(\hat{\tau})$ converges to a positive constant, so that $V^2 Z(\hat{\tau}) \to 0$. Therefore,

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

so that $\lambda \to 0$. Then (25) implies $x \to x^*$. The positive constant VZ is such that the

incentive constraint of late producers is satisfied, that is

$$\begin{aligned}
\frac{\delta}{\pi}\left[u(x^*) - x^* - Vz\right] &\geq x^* + VZ(\tau) \\
\delta\left[u(x^*) - x^*\right] &\geq \pi x^* + (\delta+\pi)VZ(\tau) \\
\delta u(x^*) &\geq (\delta+\pi)x^* + (\delta+\pi)VZ(\tau) \\
\frac{\delta}{\delta+\pi}u(x^*) - x^* &\geq VZ(\tau)
\end{aligned}$$

Also, the positive constant VZ is such that the incentive constraint of validators is satisfied, that is

$$\begin{aligned}
\delta\left[u(x) - x - VZ + \tau Z\right] &\geq \frac{1}{\tau V}\left(x + VZ(\tau)\right) \\
\delta\tau V\left[u(x) - x\right] - \delta\tau V^2 Z + \delta\tau^2 VZ &\geq \left(x + VZ(\tau)\right)
\end{aligned}$$

as $\tau \to 1$, $V \to 0$ and $VZ \to \textit{constant}$. Hence the inequality above tends to

$$\delta VZ \geq x^* + VZ$$

which requires $\delta > 1$, and we obtain in the limit $VZ \geq \frac{x^*}{\delta - 1}$. Combining that last restriction on VZ with the IC of late producers, we find

$$\delta \geq \frac{u(x^*) + \pi x^*}{u(x^*) - x^*} > 1.$$

When the solution for $\hat{V}$ is interior, we can simplify (25) and (26) 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

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

or

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

The right hand side of the above equation is always higher than 1, so that generically $x \leq x^*$. The constrained optimal solution is $(\hat{x}, \hat{V})$ that solves (25), (26) and (21) holds with equality. $\hat{\tau}$ is given by (22).

Consider the case where $V = 1-f$. Using (22), it is easy to check that $\tau = \left(1 + \sqrt{\frac{(1-\delta f)(1-f)C}{x+(1-f)C}}\right)^{-1}$. Also when $V = 1 - f$ so that $\lambda_{1-f} > 0$, the first order condition gives

$$[u'(x) - 1]\,(1 + \delta\lambda) - \lambda \frac{1}{\hat{\tau} V} \;\; = \;\; 0 \tag{27}$$

$$\left[\frac{R(\hat{\tau})}{1-f} - Z(\hat{\tau})\right] + \lambda \frac{x}{\hat{\tau} V^2} - \delta\lambda Z(\hat{\tau}) \;\; > \;\; 0 \tag{28}$$

Using (27) to eliminate λ, the definition of $R(\tau) = (1-f)\tau Z(\tau)$ and $Z(\tau) = C/(1-\tau)$, as well as the expression for τ, (28) becomes

$$\left(\sqrt{\frac{(1-\delta f)(1-f)C}{x + (1-f)C}}\right) \left\{\frac{1 + \sqrt{\frac{(1-\delta f)(1-f)C}{x+(1-f)C}}}{(1-f)}\right\} \left[\frac{x}{(1-f)C} - \frac{1}{[u'(x) - 1]}\right] > \delta \tag{29}$$

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

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

It is tedious to check that $H'(x) > 0$ and using the implicit function theorem that $dx/d\delta > 0$

(when $f \to 0$). The left hand side of (29) is decreasing in x if

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

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 (29) increases when δ decreases and the condition will be satisfied for δ low enough and below some $\bar{\delta}$.

□

Online Appendix

A Known identities

In this Appendix, we show that the equilibrium is almost identical when the identity of producers and validators are known. Then punishment strategies can exclude these agents from all future activities when they are caught cheating.

When individual history is public information, without loss of generality, we can then set $\sigma = 1$ so that agents do not exit the economy and there are no newborn. Hence there are no transfer and $T = 0$ for all agents. Also, we can set the outside option to zero for all late producers and validators. The assumption here is that there is an economy wide agreement to ban these agents from all future activities when they are caught cheating. In this case the participation constraints are $U \geq 0$ and $U_v \geq 0$, while the repayment constraints are

$$\begin{aligned} -(y + wVz) + \beta U &\geq 0, \\ -(y + wVz) + \beta \mathbb{E}U_{IV} &\geq 0. \end{aligned}$$

The condition that validators do not accept bribes simplifies to

$$\pi\beta\mathbb{E}U_v \geq \frac{1}{\tau V}(y + wVz).$$

The rest of the analysis follows by defining the default factor as

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

Then the set of IF allocations is characterized by

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

For the case with external validation, the last constraint is replaced with

$$\delta R(\tau) \geq \frac{1}{\tau V}\left(x+VZ(\tau)\right).$$

Using the proper rescaling of δ, Proposition 2 is unchanged.

B Permissionless validation with free entry

In this Appendix, we analyze the case where any agent can become a validator, although at a cost. The fixed cost of entry is set to ε that agents pay once and for all — for instance, buying the necessary computer equipment and connecting to the network. Since the validation is permissionless, we assume that given there are V validators (determined in equilibrium by free entry), one is selected at random to validate a trade. With α trade to validate, a validator has a probability α/V of being selected to validate a trade. As in the paper, any validator incurs the cost c_v to validate the trade and c_s to write and make the new ledger available to the community of validators. In exchange the validators gets a fee z as a compensation for the work. We set $C = c_s + c_v/(1-f)$.

Given V, the expected payoff of a validator then is $\bar{U}_F$ defined by

$$(1-\beta)\bar{U}_F = \frac{\alpha}{V}(1-f)\left(z-C\right)$$

Participation implies the restriction, $\bar{U}_F \geq 0$. Free entry implies validators will enter as long

as $-\varepsilon + \bar{U}_F \geq 0$. This equation holding with an equality pins down the number of validators in equilibrium,

$$V = \frac{\alpha(1-f)\,(z-C)}{(1-\beta)\varepsilon} \tag{30}$$

We assume that validators can be prevented from downloading the blockchain if they are caught cheating (this is the best case scenario for permissionless validation). Validators do not accept a bribe $\bar{z}$ whenever

$$\beta \bar{U}_F \geq \bar{z} + (1-\pi)\beta \bar{U}_F.$$

Using $\bar{z} = y + z$ as well as $\bar{U}_F = \varepsilon$, we obtain

$$\pi\beta\varepsilon \geq (y+z)\,.$$

Using the PC of early producers $y \geq x$ at equality, an allocation is incentive feasible whenever it satisfies,

$$u(x) - x - z \;\geq\; 0 \tag{31}$$

$$z \;\geq\; C \tag{32}$$

$$\pi\beta\varepsilon \;\geq\; x + z \tag{33}$$

The planner wants to maximize

$$\alpha(1-f) \max_{x,z} \{u(x) - x - V\,(C + (1-\beta)\varepsilon)\}$$

subject to the three constraints above. Using (30), we can easily see that the planner's objective function is decreasing in z, so the planner will set $z = C$ (which implies $V \to 0$)

and being constrained by (33), the allocation will be

$$x = \begin{cases} x^* & \text{if } x^* < \pi\beta\varepsilon - C \\ \pi\beta\varepsilon - C & \text{otherwise} \end{cases}$$

Hence permissionless ledgers will implement the efficient allocation x^* as long as it is small enough relative to the entry cost into validation.

C The validation game

In this Appendix, we specify the details of the validation game for internal validation. We analyze a free-rider problem inherent to the validation protocol: validators have 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 lose integrity.[48]

We keep some of the features of the optimal allocation. 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 $\bar{c}_s$).

C.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

[48] See also Amoussou-Guénou, et al. (2019).

$\int_{i=1}^{V} m_i \mathbb{I}_{m_i \neq \emptyset} di > \tau V$ and $\mathcal{I}(\mathbf{m}) = 0$ otherwise.[49]

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 $\ell \in \{G, B\}$.[50] 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 – the expected payoff of a working validator from validating the transaction is

$$\begin{aligned} -c_v + (1-f) \left\{ \begin{array}{c} \sigma_i(G) \left(-c_s + E_i \left[\mathcal{I}(\mathbf{m}) z + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), G)) \mid m_i = 1\right]\right) \\ +(1-\sigma_i(G)) E_i \left[\mathcal{I}(\mathbf{m}) \times 0 + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), G)) \mid m_i = \emptyset\right] \end{array} \right\} \\ +f \left\{ \begin{array}{c} \sigma_i(B) \left(-c_s + E_i \left[\mathcal{I}(\mathbf{m}) \times 0 + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), B)) \mid m_i = 1\right]\right) \\ + (1-\sigma_i(B)) E_i \left[\mathcal{I}(\mathbf{m}) \times 0 + \beta U_V(h_t, (\mathcal{I}(\mathbf{m}), B)) \mid m_i = \emptyset\right] \end{array} \right\} \end{aligned} \tag{34}$$

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.

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

[50] 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.

We now explain the different elements in (34). 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 $\mathcal{I}(\mathbf{m}) = 1$,[51] 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 message $m_i = 1$ for a producer with label B.

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

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 their 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. Since she 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.

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

$$-c_v + (1-f)\left(-c_s + z + \beta U_V(h_t,(1,G))\right) + f\beta U_V(h_t,(0,B)) \tag{35}$$

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 (12) 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 was assigned the wrong label 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 following these "observationally equivalent" outcomes and when the validation process resulted in 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.

C.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 differs 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*

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

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

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

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,

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

The remaining validator will work whenever

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

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

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

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

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

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

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 conditions, 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 (13), 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. □

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 effectively vote that the producer has label B. Therefore 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 $(1-f)\left(z - c_s + \beta U_{Vt}\right)$. 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 (15) 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 $u(x) > x + VZ(\tau)$. □

D 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

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

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 obtain

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

When $z^1 = Z$, the participation constraint of validators (49) 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 (46) at equality and replace $y = x + wVZ(\tau)$. Then the set of IF allocations is characterized by

$$\delta \frac{c_s^*}{\bar{c}_s} \left[u(x) - (x + VZ(\tau))\right] \geq (x + VZ(\tau)) \tag{36}$$

$$\delta \int_0^{c_s^*} \left[u(x) - (x + VZ(\tau)) + R(\tau, c_s)\right] \frac{dc_s}{\bar{c}_s} \geq \frac{1}{\tau V}\left(x + VZ(\tau)\right) \tag{37}$$

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

D.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,

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

or replacing y, 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 (55) and (56). Using the same steps as the simpler case, we can show that (55) is always slack when (56) holds. Re-arranging the constraint,

$$\pi\frac{\beta\alpha}{1-\beta}\frac{c_s^*}{\bar{c}_s}\left[u(x)-x\right]\geq\left(\frac{1}{\tau V}\right)\left(x+VZ(\tau)\right)-\frac{\pi\beta\alpha}{1-\beta}\int_0^{c_s^*}\left(R(\tau,c_s)-VZ(\tau)\right)\frac{dc_s}{\bar{c}_s} \tag{38}$$

Since the objective function is independent of τ, the planner will choose τ to minimize the right hand side of (38). The first order condition for the optimal threshold $\hat{\tau}$ gives

$$\frac{1-\hat{\tau}}{\hat{\tau}}=\sqrt{\frac{\left[1-\delta\frac{c_s^*}{\bar{c}_s}\left(1-f-V\right)\right]\left(c_s^*+\frac{c_v}{(1-f)}\right)}{\frac{x}{V}+c_s^*+\frac{c_v}{(1-f)}}} \tag{39}$$

This threshold is well defined if

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

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 -

$$\begin{aligned}
[u'(x) - 1](1 + \delta\lambda)\frac{c_s^*}{\bar{c}_s} - \lambda\frac{1}{\hat{\tau}V} &= 0 && (40)\\
\int_0^{c_s^*}\left[\left(c_s + \frac{1}{1-f}c_v\right)\right]\frac{dc_s}{\bar{c}_s} + \lambda\frac{x}{\hat{\tau}V^2} - \delta\lambda\frac{c_s^*}{\bar{c}_s}Z(\hat{\tau}) &= 0 && (41)\\
\left\{u(x) - x - V\left(Z(\tau) - \frac{R(\tau, c_s^*)}{1-f}\right)\right\}\frac{1}{\bar{c}_s} && && (42)\\
+\lambda\left\{\begin{array}{c}\delta\left[\{u(x) - x - VZ(\tau)\} + R(\tau, c_s^*)\right]\frac{1}{\bar{c}_s}\\ +\delta\int_0^{c_s^*}\left[-V\frac{\partial Z(\tau)}{\partial c_s^*} + \frac{\partial R(\tau, c_s)}{\partial c_s^*}\right]\frac{dc_s}{\bar{c}_s} - \frac{1}{\tau V}V\frac{\partial Z(\tau)}{\partial c_s^*}\end{array}\right\} &\geq 0 && (43)
\end{aligned}$$

We now determine conditions so that the solution is $c_s^* = \bar{c}_s$. Suppose this is the case. Then the expression in $\{.\}$ in (43), 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 (38)

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

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,

$$\delta\left[u(x)-x\right]=\left(\frac{1}{\tau V}\right)\left(x+VZ(\tau)\right)-\delta\int_{0}^{\bar{c}_s}\left(R(\tau,c_s)-VZ(\tau)\right)\frac{dc_s}{\bar{c}_s} \tag{44}$$

E Concave utility for early producers

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

$$\begin{aligned} u(x)-y-Vz &\geq 0 && (45)\\ v(y)-x &\geq 0 && (46)\\ -c_v+(1-f)\left(z-c_s\right) &\geq 0 && (47)\end{aligned}$$

From the PC of early producers holding at equality,

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

where Φ is increasing and convex, and the PCs become

$$\begin{aligned} u(x)-\Phi(x)-Vz &\geq 0 && (48)\\ -c_v+(1-f)\left(z-c_s\right) &\geq 0 && (49)\end{aligned}$$

Repayment constraints:

$$-\left(\Phi(x)+Vz\right)+\beta U \;\geq\; 0. \tag{50}$$

$$-\left(\Phi(x)+Vz\right)+\beta \mathbb{E}U_V \;\geq\; 0. \tag{51}$$

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 $\bar{z}=(y+Vz)/(\tau V)$ to τV validators. Using $\bar{z}=(y+wVz)/(\tau V)$, a validator rejects the bribe whenever[52]

$$\pi\beta U_V \;\geq\; \frac{1}{\tau V}\left(y+Vz\right). \tag{52}$$

$$\pi\beta U_V \;\geq\; \frac{1}{\tau V}\left[\Phi(x)+Vz\right]. \tag{53}$$

Validation threshold.

$$z \geq \frac{c_s + c_v/(1-f)}{1-\tau}. \tag{54}$$

Since the payment to validators should be minimized, (13) binds so validators take home

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

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

Cheapest to deliver.

The participation constraint of validators (49) is always satisfied. Let $R(\tau)$ be the expected rent of validators,

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

We can set the participation constraint for early producers (46) at equality and replace $y = \Phi(x)$. Since validators earn a rent, $U_V \geq U$ and (51) is satisfied whenever (50) is. Then the set of IF allocations is characterized by

$$\frac{\beta\alpha}{1-\beta}\left[u(x) - \Phi(x) - Vz\right] \geq \Phi(x) + Vz \tag{55}$$

$$\pi\frac{\beta\alpha}{1-\beta}\left[u(x) - \Phi(x) - Vz + R(\tau)\right] \geq \frac{1}{\tau V}\left(\Phi(x) + Vz\right) \tag{56}$$

F Outside option

In this Appendix, we assume late producers, including validators, have the option to trade on another platform with a net payoff $\bar{U} \geq 0$ (net of opening an account). The outside option for validators and late producers is the same because if there were two outside options, one for validators and one for non-validators, they would always choose the one giving the highest payoff.

In this setting, the participation constraints of validators, early and late producers become,

$$\begin{aligned} -\gamma_v + U_v &\geq \bar{U} && (57)\\ -\gamma + U &\geq \bar{U} && (58)\\ y - x &\geq 0 && (59)\\ -c_v + \mathbb{E}_{w\geq\tau|c_{s,i}} z - c_{s,i} &\geq 0 && (60) \end{aligned}$$

and given the share of working validators is $w \geq \tau$, the repayment constraint of late producers and internal validators is respectively

$$\begin{aligned} -(y + wVz) + \beta U &\geq \beta \max\{-\gamma + U; -\gamma_v + \mathbb{E}U_{IV}; \bar{U}\} && (61)\\ -(y + wVz) + \beta \mathbb{E}U_{IV} &\geq \beta \max\{-\gamma + U; -\gamma_v + \mathbb{E}U_{IV}; \bar{U}\}. && (62) \end{aligned}$$

as in the main text, agents whose account has been blocked can open a new late producer's account at cost γ, a validator's account at cost γ_v, or obtain the (net) payoff $\bar{U}$ by using the competing ledger. Using $\bar{z} = (y + wVz)/(\tau V)$, a validator rejects the bribe whenever[53]

$$\pi\beta\mathbb{E}U_s \geq \frac{1}{\tau V}(y + wVz) + \pi\beta \max\{\mathbb{I}_{s=IV}(-\gamma + U); -\gamma_s + U_s; \bar{U}\}. \quad (63)$$

The outside option does not affect the analysis of the validation threshold.

The incentive constraints of late producers and validators are relaxed when the cost of opening an account is set as high as possible, that is $\gamma = U - \bar{U}$ and $\gamma_v = U_v - \bar{U}$. Then using $R(\tau)$ to denote the expected rent of validators,

$$R(\tau) = \tau Z(\tau) \quad (64)$$

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

defining the default factor as (recall that the effective discount factor is $\beta = \sigma\tilde{\beta}$)

$$\delta \equiv \frac{\pi\sigma\tilde{\beta}\alpha}{1-\sigma\tilde{\beta}-(1-\sigma)/\sigma}.$$

and replacing the expressions for the transfers, the set of IF allocations is characterized by[54]

$$\frac{\delta}{\pi}\left[u(x)-x-Vz\right] \geq x+VZ(\tau)+\frac{(1-\beta)}{\alpha}\frac{\delta}{\pi}\bar{U} \tag{65}$$

$$\delta\left[u(x)-x-Vz+R(\tau)\right] \geq \frac{1}{\tau V}\left(x+VZ(\tau)\right)+\frac{1-\beta}{\alpha}\delta\bar{U} \tag{66}$$

With external validation, (66) simplifies to

$$\delta R(\tau) \geq \frac{1}{\tau V}\left(x+VZ(\tau)\right)+\frac{1-\beta}{\alpha}\delta\bar{U} \tag{67}$$

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 future income (as captured by a high $\tilde{\beta}$). Also, it is easy to check that the default factor is decreasing in the survival rate σ. There are two counteracting effects from a higher survival rate: On one hand, agents effectively become more patient, which improves incentives, but on the other hand a higher survival rate decreases the transfer that surviving agents get from the ledger, which weakens discipline. The latter effect always dominates.

[54]Since we set the participation constraint for early producers (7) at equality, we can use $y = x$. Also, since validators earn a rent, $U_{IV} \geq U$ and (10) is satisfied whenever (9) is.

Optimal design with external validation. With external validation the only relevant constraints are (65) and (67), which we can write respectively, as

$$\begin{aligned} \delta u(x) &\geq (\delta+\pi)\left[x+V\frac{C}{1-\tau}\right]+\frac{(1-\beta)}{\alpha}\delta\bar{U} \qquad (68)\\ (1-\tau)x &\leq \left(\delta\tau^2-1-(1-\tau)\tau\frac{1-\beta}{\alpha}\delta\frac{\bar{U}}{C}\right)VC \qquad (69)\end{aligned}$$

It is easy to show that the validators' incentive constraint (69) always binds,[55] so we can use it to replace for the total cost of validation, VC in the late producer's repayment constraint (68) and the objective function. The problem of the planner then becomes

$$\alpha \max_{x,\tau\in[\tilde{\tau},1]}\left\{u(x)-\frac{\left(\delta\tau^2-\tau-(1-\tau)\tau\frac{1-\beta}{\alpha}\delta\frac{\bar{U}}{C}\right)}{\left(\delta\tau^2-1-(1-\tau)\tau\frac{1-\beta}{\alpha}\delta\frac{\bar{U}}{C}\right)}x\right\}$$

subject to

$$\delta u(x) \geq (\delta+\pi)\left[\frac{\left(\delta\tau^2-(1-\tau)\tau\frac{1-\beta}{\alpha}\delta\frac{\bar{U}}{C}\right)}{\left(\delta\tau^2-1-(1-\tau)\tau\frac{1-\beta}{\alpha}\delta\frac{\bar{U}}{C}\right)}x\right]+\frac{(1-\beta)}{\alpha}\delta\bar{U}$$

Inspecting (69), notice that a necessary and sufficient condition for the existence of external validation is that intertemporal incentives are strong enough, in the sense that $\delta \geq 1$. Otherwise, the only incentive feasible allocation is autarky. With $\delta > 1$, it is straightforward to verify that the constraint is relaxed when τ is highest. Also, the objective function is maximized when $\tau = 1$. Therefore, with external validation, the solution is $\tau = 1$ and $V \to 0$, and the optimal trading size $\tilde{x}(\delta) \leq x^*$ where $\tilde{x}(\delta) = x^*$ if $\delta \geq \delta^*(\bar{U}) > 1$ so that the constraint is not binding or it is defined as the solution to $u(\tilde{x}) = \tilde{x}(\delta+\pi)/(\delta-1)+(1-\beta)\bar{U}/\alpha$ otherwise, with $\tilde{x} \to 0$ as δ decreases to 1. The payment to the validator is $VZ \to \tilde{x}(\delta)/(\delta-1)$.

[55]Suppose it does not at the solution $(\tilde{x},\tilde{\tau},\tilde{V})$. Then reduce V until it does. This increases the objective function (18), while relaxing the participation constraint of late producers. So $(\tilde{x},\tilde{\tau},\tilde{V})$ could not be the solution, a contradiction.

Then following the same steps as for Proposition 2 in the main text, we obtain

Proposition 3. *The constrained optimal solution $(\hat{x}, \hat{V}, \hat{\tau})$ solves (21) at equality and (22)–(26) and is characterized by four regions:*

1a. [centralized system - efficient trade size] *If $\delta \geq \delta^*(\bar{U}) > 1$, external and internal validations are characterized by $\hat{V} \to 0$, $\hat{\tau} \to 1$ and $\hat{x} \to x^*$. Validation requires arbitrarily large payments, i.e. $Z(\hat{\tau}) \to \infty$, while $\lim_{\hat{\tau}\to 1} V(\hat{\tau})Z(\hat{\tau}) = \frac{x^*}{\delta - 1}$. Welfare is higher with internal validation. $\delta^*(\bar{U})$ is increasing in the outside option $\bar{U}$.*

1b. [centralized system - inefficient trade size] *If $1 \leq \delta \leq \delta^*(\bar{U})$, external validation is characterized by $\hat{V} \to 0$, $\hat{\tau} \to 1$ and $\hat{x} \to \tilde{x}(\delta) \leq x^*$. Internal validation is characterized by $\hat{V} > 0$, $\hat{\tau} < 1$ and $\hat{x} < x^*$. Welfare is higher with internal validation.*

2. [partially distributed system] *If $\bar{\delta}(\bar{U}) < \delta \leq 1$, only internal validation can decentralize trade. 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 $\hat{x} < x^*$. $\bar{\delta}(\bar{U})$ is increasing in the outside option $\bar{U}$.*

3. [fully distributed system] *If $\delta_0(\bar{U}) < \delta \leq \bar{\delta}(\bar{U})$, only internal validation can decentralize trade. All late producers are validators $\hat{V} = 1$, and $\hat{\tau} = \left(1 + \sqrt{\frac{C}{x+C}}\right)^{-1}$. The constrained optimal allocation is $\hat{x} < x^*$. $\delta_0(\bar{U})$ is increasing in the outside option $\bar{U}$.*

4. [no trade] *If $\delta \leq \delta_0(\bar{U})$, there is no validation protocol that can decentralize trade.*

Proof. First we show that internal validation does at least as well as external validation. Both problems have the same objective function. The constraints for external validation are

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

while the constraint for internal validation is

$$\delta\left[u(x)-(x+Vz)+R(\tau)\right]\geq\frac{1}{\tau V}\left(x+Vz\right)+\frac{1-\beta}{\alpha}\delta\bar{U}$$

since $u(x)-(x+Vz)\geq 0$ the constraint set for internal validation is weaker than the one for internal validation. This shows the claim.

Next we concentrate on the results for internal validation (the one for external validation is in the text). The first order conditions with respect to x (70) and the one with respect to V (71) are

$$\begin{aligned}\left[u'(x)-1\right](1+\delta\lambda)-\lambda\frac{1}{\hat{\tau}V} &= 0 \qquad (70)\\ \left[\frac{R(\hat{\tau})}{1-f}-Z(\hat{\tau})\right]+\lambda\frac{x}{\hat{\tau}V^2}-\delta\lambda Z(\hat{\tau})+\lambda_0-\lambda_1 &= 0 \qquad (71)\end{aligned}$$

where λ is the Lagrange multiplier on the validators' IC constraint, and λ_1 is the one on $V\leq 1$ and λ_0 is the one on $V\geq 0$.

We first consider the solution when $\hat{\tau}\to 1$. Then $Z(\hat{\tau})\to+\infty$ and

$$R(\hat{\tau})-Z(\hat{\tau}) = -\left(\bar{c}_s+c_v\right)$$

We now show that λ and V both converge to zero. To the contrary, suppose $V>0$ ($\lambda_0=0$) and $\lambda>0$. Since $Z(\hat{\tau})\to+\infty$ as $\hat{\tau}\to 1$, the LHS of (71) is necessarily negative as $\hat{\tau}\to 1$. Hence $\lambda_0>0$ and/or either $\lambda\to 0$ to reestablish the equality, which contradicts $V>0$ and $\lambda>0$. Hence either $V=0$ or $\lambda=0$ or both. . Starting with $V\in(0,1)$, we can rewrite (71) as

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

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 (15) never binds, it must be that $VZ(\tau)$ converges to a positive constant, so that $V^2 Z \to 0$. Therefore,

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

so that $\lambda \to 0$. Then (70) implies $x \to x^*$. The positive constant VZ is such that the incentive constraint of late producers is satisfied, that is

$$\frac{\delta}{\pi}\left[u(x^*) - x^* - Vz\right] \;\geq\; x^* + VZ(\tau) + \frac{(1-\beta)}{\pi\alpha}\delta\bar{U}$$

which we can rewrite as

$$\frac{\delta}{\delta+\pi}u(x^*) - x^* - \frac{(1-\beta)}{\alpha}\frac{\delta}{\delta+\pi}\bar{U} \geq VZ(\tau).$$

Also, the positive constant VZ is such that the incentive constraint of validators is satisfied, that is

$$\delta\left[u(x) - x - VZ + \tau Z\right] \;\geq\; \frac{1}{\tau V}\left(x + VZ(\tau)\right) + \frac{1-\beta}{\alpha}\delta\bar{U}.$$

Cross-multiplying by τV and re-arranging, we obtain

$$\delta\tau V\left[u(x) - x\right] - \delta\tau V^2 Z + \delta\tau^2 VZ \geq \left(x + VZ(\tau)\right) + \tau V\frac{1-\beta}{\alpha}\delta\bar{U}.$$

When $\tau \to 1$, $V \to 0$ and as $VZ \to \textit{constant}$ the inequality above tends to

$$\delta VZ \geq x^* + VZ$$

which requires $\delta > 1$, and we obtain in the limit $VZ \geq \frac{x^*}{\delta - 1}$. Combining that last restriction

on VZ with the IC of late producers, we find that $\delta \geq \delta^*(\bar{U}) > 1$ since

$$\frac{\delta}{\delta+\pi}u(x^*) - x^* - \frac{(1-\beta)}{\alpha}\frac{\delta}{\delta+\pi}\bar{U} \geq \frac{x^*}{\delta-1}$$

implies

$$\delta \geq \frac{u(x^*) - \frac{(1-\beta)}{\alpha}\bar{U} + \pi x^*}{u(x^*) - \frac{(1-\beta)}{\alpha}\bar{U} - x^*}.$$

When the solution for $\hat{V}$ is interior, we can simplify (70) and (71) to obtain

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

and

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

or

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

The right hand side of the above equation is always higher than 1, so that generically $x \leq x^*$. The constrained optimal solution is $(\hat{x}, \hat{V})$ that solves (70), (71) and (21) holds with equality. $\hat{\tau}$ is given by (22).

Consider the case where $V = 1$. Using (22), it is easy to check that $\tau = \left(1 + \sqrt{\frac{C}{x+C}}\right)^{-1}$. Also when $V = 1$ so that $\lambda_1 > 0$, the first order condition gives

$$[u'(x) - 1]\,(1+\delta\lambda) - \lambda\frac{1}{\hat{\tau}V} = 0 \tag{72}$$

$$\left[\frac{R(\hat{\tau})}{1-f} - Z(\hat{\tau})\right] + \lambda\frac{x}{\hat{\tau}V^2} - \delta\lambda Z(\hat{\tau}) > 0 \tag{73}$$

Using (72) to eliminate λ, the definition of $R(\tau) = \tau Z(\tau)$ and $Z(\tau) = C/(1-\tau)$, as well as

the expression for τ, (73) becomes

$$\left(\sqrt{\frac{C}{x+C}}\right)\left\{1+\sqrt{\frac{C}{x+C}}\right\}\left[\frac{x}{C}-\frac{1}{[u'(x)-1]}\right]>\delta \tag{74}$$

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

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

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 (74) 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 (74) increases when δ decreases and the condition will be satisfied for δ low enough and below some $\bar{\delta}$. That threshold $\bar{\delta}$ is increasing in $\bar{U}$, and we can show that in the region where $V=1$, $dx/d\bar{U}<0$.

Finally, as δ keeps decreasing below $\bar{\delta}$, $x\to 0$ so that $H(x)\to 4C$, while $\delta\left[u(x)-x-\frac{1-\beta}{\alpha}\bar{U}-C\right]<0$. Therefore, there is a $\delta_0(\bar{U})$ below which there is no $x\in\mathbb{R}$ that satisfies the incentive constraint of validators, and that constraint shows that $\delta_0(\bar{U})$ is increasing with $\bar{U}$. □

The bounds of the four regions are increasing with the outside option $\bar{U}$, which reveals that as it becomes more attractive to switch to a competing arrangement offering net payoff $\bar{U}$, more intertemporal incentives (higher δ) is required to implement the allocation. How the constrained optimal solution $(\hat{x},\hat{V},\hat{\tau})$ changes with $\bar{U}$ depends on parameters. This is

apparent from inspecting (66): Suppose $x < x^*$ but close to x^*, then increasing $\bar{U}$ will not play to increase x because the LHS of (66) will not increase much since x is already close to x^*, while the bribe size would increase. Hence x will decrease. V could increase to make sure that x is not reduced by "too much." But this could backfire if C is large. In that case, it may be best to reduce V as well.

Previous volumes in this series

923 January 2021	Optimal bank leverage and recapitalization in crowded markets	Christoph Bertsch and Mike Mariathasan
922 January 2021	Does regulation only bite the less profitable? Evidence from the too-big-to-fail reforms	Tirupam Goel, Ulf Lewrick and Aakriti Mathur
921 January 2021	Firm-specific risk-neutral distributions with options and CDS	Sirio Aramonte, Mohammad R Jahan-Parvar, Samuel Rosen and John W Schindler
920 January 2021	Investing like conglomerates: is diversification a blessing or curse for China's local governments?	Jianchao Fan, Jing Liu and Yinggang Zhou
919 January 2021	Understanding bank and nonbank credit cycles: A structural exploration	C Bora Durdu and Molin Zhong
918 January 2021	Sovereign credit and exchange rate risks: Evidence from Asia-Pacific local currency bonds	Mikhail Chernov, Drew Creal and Peter Hördahl
917 January 2021	Trade sentiment and the stock market: new evidence based on big data textual analysis of Chinese media	Marlene Amstad, Leonardo Gambacorta, Chao He and Dora Xia
916 January 2021	Firm-level R&D after periods of intense technological innovation: The role of investor sentiment	Sirio Aramonte and Matthew Carl
915 December 2020	The globalisation of inflation in the European emerging countries	Horatiu Lovin
914 December 2020	Demographic shifts, macroprudential policies, and house prices	Jieun Lee and Hosung Jung
913 December 2020	Forecasting expected and unexpected losses	Mikael Juselius and Nikola Tarashev
912 December 2020	Regulatory capital, market capital and risk taking in international bank lending	Stefan Avdjiev and Jose Maria Serena
911 December 2020	Bilateral international investments: The Big Sur?	Fernando Broner, Tatiana Didier, Sergio L Schmukler and Goetz von Peter
910 December 2020	Recessions and mortality: a global perspective	Sebastian Doerr and Boris Hofmann

All volumes are available on our website www.bis.org.