Distributed Ledgers

Design and Regulation of Financial Infrastructure and Payment Systems

OA Funding Provided By:

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

The title-level DOI for this work is:

doi:10.7551/mitpress/13382.001.0001

10

Regulation and the Use of Distributed Ledger Technology

In this chapter we group together various insights from theory that pertain to optimal regulation of financial systems. In several cases, distributed ledger technology can play a central role in mitigation of bank and market runs and in coordinating payment devices related to the issue of digital assets, for instance. More generally, free and open competition as in the EvryNet platform can suffer from coordination problems, especially under traditional regulation. Market structure and its regulation needs to be part of the overall ex ante design.

10.1 Mitigating Runs on Banks and Markets

DLT can improve on current technology used by banks and markets in order to mitigate resulting runs. Diamond and Dybvig (1983) wrote a seminal paper on bank runs. The idea is that investors deposit their funds in a bank. The bank can invest in a short-term asset or tie up funds with a longer-term investment. Yet depositors retain the option of withdrawing funds early, which they would like to do if they are hit with a preference or other shock that makes it more urgent to have the funds. Ideally, the bank could plan on the fractions of agents in the population who will be urgent (or the opposite, more patient) and invest accordingly. However, if more than that fraction of urgent households gets the idea there is a run on the bank, then patient investors, fearing a disproportionate number will withdraw early, leaving little for the longer term, will run also. The point is that this can become a self-fulling prophesy, a valid if horrible equilibrium. There is a more recent literature on market runs rather than bank runs that draws an analogy (Martin, Skeie, and von Thadden 2013, 2014) so that the logic of runs applies beyond banks.

Though mitigation of runs is an important rationale for intervention, as in deposit insurance, that can bring other distortions, no market discipline on the bank means regulators must take on even more responsibility for what the bank might be doing. Indeed, the monetary authority may have to inject liquidity if banks have taken on too much risk and investments go bad, and banks' anticipating this action exacerbates the problem.

Yet the problem of runs has a partial if not complete solution that can be found within the mechanism design literature itself. A simple version is suspended convertibility, thus reassuring investors that some of their money would remain, regardless. More sophisticated, sequential-service models treat customers differently depending on when they arrive at the trading window—that is, on the history of traders and trades before them, determining their incentives to announce privately observed shocks (Green and Lin 2003).

This is where distributed ledgers can play a key role. Messages in distributed ledgers are essentially time-stamped and immutable, so it is quite natural to think about history based on previous reported transactions in the blockchain as being used to determine consequences for contemporary actions and messages. With independent private values, there is a unique Bayesian Nash equilibrium that eliminates runs entirely and is a dominant strategy equilibrium. Thus multiple equilibria, as in the original Diamond-Dybvig model, is no longer a problem. With correlated private values, some run-like phenomena remain, as the history of traders at any moment in time is, on the one hand, self-reported and yet, on the other, desired as a key statistic, as it is revealing of aggregates. Nevertheless, the risk of runs can be substantially mitigated.

These mechanisms could be used in practice as a private-sector alternative to public liquidity. This could dramatically mitigate the central bank moral-hazard problem, that bailouts ex post create perverse private-sector incentives ex ante. This could also alleviate the concern that private platforms might somehow pose a systemic risk. They can be designed to guard against that. Optimal regulation would then amount to verifying that the platforms that implement these insights are designed properly.

Here, of course, there is a political economy issue. Getting diverse parties to agree to these changes is nontrivial, especially if banks benefit from government largess. Another qualification stems from the latency in networks, which confounds the chronological ordering. A modification of the model to treat identically a set of orders that are received over small intervals of time will need to be designed. Budish, Cramton, and Shim (2015) propose such a mechanism to deal with high-frequency traders. Likewise, a related finance literature outlines the trade-offs between frequency of adjustment and thick markets (Du and Zhu 2017).

10.2 The Limits of Competition for Contracts: Coordination and Re-optimized Regulation

It was argued earlier in chapter 8 that when smart contracts are needed, one can envision competition among intermediaries in offering such contracts, consistent with a Walrasian equilibrium, with its Pareto optimality properties. There are, however, caveats and advice to regulators that come from general equilibrium theory. Specifically, following Pesendorfer (1995) and Makowski (1980), one can begin in an economy with incomplete markets and contracts and then allow financial service providers to innovate. One might hope this would complete the markets, but it may not, because of complementarities in uncoordinated innovation. More coordination would be needed. A related point: Ill-informed or outdated regulation can segment an industry and make the needed coordination impossible.

In addition, markets and access should not be entirely open all the time to anyone. When there is private information and minimal scales of operation, then some forms of competition can undercut incentives and optimal diversification. Equity markets should not free-ride on existing infrastructure, for example. Certain kinds of exclusivity are needed. Competition should be ex ante for the right to provide services, not ex post to draw off customers from contracts. As in Acemoglu and Zilibotti (1997), minimum scales of operation across sectors means that risk-sharing is incomplete. Stretching the extensive margin to high minimum-scale projects is good, but this means less funding for each active project; on the margin some high-scale projects receive at the optimum relatively more funding than other zero or lower-end projects, as the resources must come from somewhere. The point: Portfolios are not balanced in funding. Unrestricted trade into equities issued by firms would undercut the optimum; investors will want to diversify their eggs into the various baskets equally. Instead, one intermediary should do all the packaging. There can be ex ante competition for this right. Townsend and Xandri (2018) provide blueprints for market design and regulation in this context.

Retrading among potential deposits in the Diamond-Dybvig model (1983), as in Jacklin (1987), is another negative factor. Retrading undercuts the implicit provision of insurance for the impatient. Distributed ledgers with ex ante contracts and commitment to market structure can prevent this from occurring.

Blueprints for the ultimate design are needed, as otherwise the system experiencing innovations may develop piecemeal and not achieve a constrained-efficient outcome. Regulators need to see the blueprints and understand the big picture.

10.3 Lessons from Monetary Theory for the Regulation of Payment Systems: The Need for Coordination

Lack of key common information may make it difficult to achieve optimal targets and, further, can lead to market crashes. Thus there are clear implications for micro/macroprudential regulation. Ironically, DLT, rather than being a regulatory concern, offers an obvious solution in the context of these problems.

10.3.1 The Impossibility of Decentralized Exchange

A natural objective is to try to achieve the Pareto optimal allocation associated with a Walrasian competitive equilibrium. Ostroy and Starr (1974) ask whether this can be done under a decentralized exchange when information is limited, or rather, whether centralization of some information is necessary.

These problems emerge despite the fact that in the Ostroy and Starr (1974) model, many important items are simply taken as given, so that there are as few obstacles as possible. The target Walrasian allocation is given; the prices as common marginal rates of substitution in the Walrasian allocation are given and fixed; and paths over time of matched agents are known in advance. Furthermore, agents simply act as computers implementing code. Here is the question: How do you write the code to implement this targeted social optimum? The answer: It is impossible to always do this when information is decentralized. The key insight: Information on the distributed ledger should be public in some clearly delineated instances.

In the Ostroy-Starr model, money plays the role of unit of account. There are potentially many agents and many underlying commodities. Agents start out in the model with underlying endowments of commodities. But the model would apply equally well to endowments of securities, various possible fiat monies, or combinations of all these. Actual payment systems handle retail and wholesale trade, securities settlement, and cross-border currency flows.

Agents in the model meet pairwise and then trade. The point is that not everyone is together all the time in one spot—they are matched, for example, in an over-the-counter (OTC) market. Trade in the model is monetized as a payment order in a *quid pro quo* condition: When goods are supplied, the supplier is given unit of account credits, in monetary terms. Likewise, when goods are purchased, the purchaser is assigned debits in the unit of account, in monetary terms. Under a natural *quid pro quo* spot-trade condition, the value of the purchases must match the value of sales in each and every contemporary bilateral transaction. Here, then, there is no credit. The prices used to value commodities and securities come from the target Walrasian allocation (known). Indeed, the entire point is to try to achieve the Walrasian allocation in a decentralized way.

Transaction values are placed on the ledgers as flows as they occur, and this results in new commodity/asset positions as stocks. Of course, as emphasized from the outset of this book and from the discussion of accounts and ledgers in chapter 8, the corresponding trade ledgers (flow) and asset ledgers (change in stocks) must be consistent.

In a key example provided by Ostroy and Starr (1974) and a technical correction by Kim (2015), it is shown that the appropriate trades across agents when they meet in subsets can require centralized knowledge of the underlying environment and trade histories. In particular, knowledge of identities of agents, histories of trade, and initial excess demands are needed, not only pairwise of those contemporaneously matched, but also of others with whom the contemporaneous set of matched traders has not been matched previously. The idea is straightforward: Implementation is both forward- and backward-looking. One has to know where the system should be headed, the target, and the remaining options in the future to reach that target, hence what trades need to have been accomplished in the past in order to make this feasible. Sometimes there are multiple choices of trades to make and more than one way to do things within a given contemporaneous pairing. Guidance is needed, from the forward and backward perspective. However, not all past histories are crucial; in the key example, there is only one instance in which information would need to be shared.

In the Ostroy-Starr (1974) and Kim (2015) example, private information about initial excess demand is the source of the potential problem. However, one can think of some initial trading period appended onto the beginning of the Ostroy-Starr model, differing from the Ostroy-Starr initial period—perhaps earlier trading rounds that determine excess demand at the time we tune into the Ostroy-Starr initial period. These initial excess demands are the deviations remaining from the ultimate target. If there were a common-consensus verified ledger to which all traders had access, then this required information would be known.

Ostroy and Starr (1974) do also discuss alternatives that mitigate the need for centralized common information. They describe what could be termed a *monetary solution*, sufficient ex ante liquidity in one of the commodities, termed the *money good*, ample enough so that expenditures of commodities or asset purchases of any agent can be financed out of this liquidity, regardless of who meets whom when and regardless of underlying economy-wide efficient targets for trade. But that amount of liquidity, whether a commodity or fiat money or tokens, is large and potentially costly, as liquidity held for this purpose is not invested in economic activities. This is made more explicit in other models.[1] Real-time gross settlement systems are in fact frequently run as hybrids with liquidity-saving mechanisms made possible by computer algorithms and queues (Martin and McAndrews 2008).

Another Ostroy-Starr alternative is a central warehouse, such as a very large broker-dealer with whom all can trade, though such a large entity could bring other distortions—namely, market power. A third alternative is credit, as if from a Walrasian banker, of the kind witnessed in early trade fairs (Townsend 1990). In the Ostroy-Starr model, however, this requires two more rounds of trade at the beginning and end: to get the credit, then the featured round of trading, and then repayment. In their model, this violates the desired criterion to achieve all trade in one round of pairings only.

It is tempting to think of the Walrasian banker as a central bank or digital reserve bank. In practice, for good reasons, central banks worry about intraday exposure and thus require good collateral. In the spirit of Ostroy-Starr, but going beyond the model, central banks also worry about the ultimate motive for trades, such as interbank borrowing transactions that are passing through their payment system but are not designated as such. Central banks would like to know more about what is driving transactions, given macroprudential concerns.

10.3.2 Information Problem with Private Monies: Circulating Private Debt and Multiple Media of Exchange Equilibria

A somewhat related setting is found in Townsend and Wallace (1987), who focus on payments made via high-velocity privately issued debt. Securities can serve as payment devices and circulate (remember the English tally sticks of chapter 5). The point here is that this is an issue with e-securities or e-assets. E-tokens are designed to facilitate liquidity and trade, but there can be problems.

The idea that securities can serve as payment devices should be familiar. In New York markets, for example, brokers experience shortages of various securities. Under rehypothecation of collateral, a lender who gives up cash for securities as collateral becomes a borrower in turn, passing the collateral on along a chain. Singh (2011) finds the velocity of circulation of treasuries is now higher than the standard monetary aggregate (M2). This can compete with fiat currency, as in Muley (2016). Carlson et al. (2016), Greenwood, Hanson, and Stein (2016), Krishnamurthy and Vissing-Jorgensen (2012), and others are all consistent in finding a liquidity premium for those treasuries. On-the-run treasuries have become money-like assets, one might say more like money than money itself. This has policy implications, clearly, though exactly what to do about it depends on the point of view of the analyst.

Table 10.1 from Townsend and Wallace (1987) provides an instructive example environment. There are four agents, four periods, and two locations. Agents have endowments of a consumption good that varies over time, but again, as earlier, one can generalize and imagine these are other objects such as securities. One can trace out chains of named debt from the issuer at the issue date passing through third parties to the issuer at the redemption date (here no reneging is allowed). This circulating private debt is the medium of exchange in contemporaneous transactions, supporting trade in other short-term noncirculating securities and the consumption commodity. One can take the consumption good as the numeraire, but again, as earlier, the idea can be generalized.

The point of the model is that there is a coordination problem. There are many potential equilibria, each of which achieves the same target Pareto optimal allocation, the complete-markets equilibrium real allocation. But these equilibria vary in who is issuing the debt initially and hence what objects are circulating—that is, what objects are serving as payment

Table 10.1
Circulating private debt: who meets whom when.

	Location	
Date	1	2
1	(1,2)	(3,4)
2	(1,3)	(2,4)
3	(1,2)	(3,4)
4	(1,3)	(2,4)

This illustrative table has four traders. Agent 1 is always at location 1 and agent 4 is always at location 2. Agents 2 and 3 alternate locations. There is scope for bilateral borrowing at dates 1,3 and 2,4 as the same agents are paired. There is also scope for circulating debt through chains of pairings—for example, an IOU issued by agent 1 at $t=1$ is passed to agent 2, who in turn presents it to 4 at $t=2$, who passes it to 3 at $t=3$, with redemption by the issuer, agent 1 at $t=4$. There are many other feasible circulating debt chains.
Source: Townsend and Wallace (1987).

devices. In a key example, to be specific, agents 1 and 2 are matched in location one, and agents 3 and 4 are matched at location two. Agents 1 and 4 stay put at their respective locations, but agents 2 and 3 keep switching back and forth across locations from one period to the next. Again, there are many equilibria that achieve the Pareto optimal target: Either all the debts that are allowed to circulate could be issued by initial parties in the first of the two locations, or by the parties in the other, second location, or they could be issued in various particular convex combinations. But by assumption, in the informationally decentralized market environment, there is no way for traders in one location to know what is going on in the other. Too much or too little debt as liquidity could be issued.

This can cause problems later in subsequent markets. Some circulating debts would be "overissued," resulting in a precipitous drop in their prices later on in the trading cycle. Agent-traders would suffer from unnecessary excessive fluctuations in their intertemporal consumption profile. Interestingly, the drop does not happen right away, as trade in short-term assets in early periods after the mismatch is discovered can partially compensate. Eventually some, but not all, parties along the transaction chains suffer drops in consumption—namely, those carrying the circulating debt across locations (Spector and Townsend 2019). Failure to achieve coordination can link up with observed chaotic conditions. Bills of exchange were traded in the London money markets, and these crashed, leading to arguments for the creation of a central bank.

DLT keeping track and verifying initial issues of long-term money-like debt in exchange for consumption or other objects, if public, would achieve in the example environment the necessary coordination. A related point: Not all information need be shared all the time. Here it is only information on initial security issues. Interestingly, and a warning to policymakers, there are no liquidity premia associated with the circulating private debt to be discerned from the data, yet the coordination problem remains. The crash comes as a surprise. The lesson for policymakers: A clear understanding of the environment and consequent tracking of transactions data is needed.