Queen's Economics Department Working Paper No. 1397

Blockchain-based Settlement for Asset Trading

Jonathan Chiu
Bank of Canada

Thorsten Koeppl
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

1-2018

Blockchain-based Settlement for Asset Trading*

Jonathan Chiu† Bank of Canada

Thorsten V. Koeppl‡ Queen's University

January, 2018

Abstract

Can securities be settled on a blockchain and, if so, what are the gains relative to existing settlement systems? We consider a blockchain that ensures delivery-vs-payment by linking transfers of assets with payments and operates via a Proof-of-Work protocol. The main problem is to overcome settlement fails where participants fork the chain to get rid of trading losses. To deter forking, the blockchain needs to restrict block size and block time in order to generate sufficient transaction fees which finance costly mining. We show that large enough trading volume, sufficiently strong preferences for fast settlement and limited trade size and risk are necessary conditions for blockchain-based settlement to be feasible. Despite mining being a deadweight cost, our estimates based on the market for US corporate debt show that gains from moving to faster and more flexible settlement are in the range of 1-4 bps relative to existing legacy settlement systems.

Keywords: Securities Settlement, Blockchain, Block Size, Block Time, Transaction Fees, Club Good

JEL Classification: G2, H4, P43

*The views expressed in this paper are not necessarily the views of the Bank of Canada. We thank our discussant, Larry Glosten, and the audience of the RFS FinTech conference for their comments. This research was supported by SSHRC Insight Grant 435-2014-1416. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

†Bank of Canada, Ottawa, K1A 0G9, Canada (e-mail: jchiu@bankofcanada.ca).

‡Queen's University, Department of Economics, Kingston, K7L 3N6, Canada (e-mail: thor@econ.queensu.ca).

1 Introduction

A principal risk in securities markets is settlement risk where the seller of a security fails to deliver the security while receiving payment or where the buyer of a security fails to deliver payment while receiving the security. To deal with such risk, securities settlement systems have been put in place in many markets to ensure a *delivery versus payment* (DvP) mechanism where the settlement of the cash and the securities leg in a trade are intrinsically linked. These systems are typically organized around a specialized third-party called *Central Securities Depository* (CSD) which transfers legal ownerships of securities against payment. Still, many other intermediaries such as brokers, custodians and payment agents are involved in facilitating the clearing and settlement of a trade.

Having many intermediaries involved in the settlement process is time- and cost-intensive, and current practices are commonly viewed as rather inefficient.[1] Settlement cycles in many fragmented securities markets tend to be fairly long and fixed at particular time intervals such as T+2 or T+3 to coordinate actions among intermediaries.[2] Similarly, intermediaries operate back office systems that are incompatible with other financial market infrastructure. This leads to duplication of costs for record-keeping and often involves substantial costs for reconciling such records.[3]

Many practitioners believe that *blockchain* or *distributed ledger technology* (DLT) has the potential to radically transform securities settlement. The key innovation is to have a shared database of securities ownership that can be updated without relying on multiple, specialized intermediaries or a third-party infrastructure.[4] Cutting out intermediaries could substantially reduce post-trade processing costs.[5] Some infrastructure could be made redundant altogether by employing *smart contracts* to contain settlement risk. Such contracts are built to automatically provide DvP in the

[1] For example, see the discussion in Benos et al. (2017).

[2] Some transactions can take even longer to settle. For example, the post-trade cycle for syndicated loans can be unpredictable and frequently stretch to three weeks due to legal complications.

[3] As summarized by Benos et al. (2017), "revenue from settlement, custody and collateral management amounted to $40-45 billion in 2013, which represented approximately 13% of the total trade value chain (from execution to settlement)." Broadridge (2015) estimates that the financial industry spends roughly $17bn to $24bn per year in core post-trade processing, reference data, reconciliations, trade expense management, client life-cycle management, corporate actions, tax and regulatory reporting.

[4] In Section 2 below we provide a more detailed discussion of how blockchain technology applies to securities settlement systems.

[5] Mainelli and Milne (2016) estimate a cost saving of at least 50% for security transactions. Santander (2015) estimates a potential annual cost reduction of $20 billion in the banking industry.

absence of a central authority (see Wall and Malm (2016)). Blockchain technology could speed up settlement, both by getting rid of a fragmented post-trade infrastructure and by implementing a faster settlement cycle. Finally, a blockchain could offer the opportunity for market participants to choose how fast a transaction settles. According to FINRA (2017), "... the adoption of DLT may not necessarily lead to implementation of real-time settlement, [but] it has the potential to make settlement time more a feature of the actual market needs of the parties instead of being based on operational constraints." Consequently, using a blockchain for settling securities could allow for flexible settlement times creating value beyond what can be offered by traditional settlement systems especially when transactions are time critical.

Under what condition is it feasible to settle securities on a blockchain and what are the gains relative to using the settlement infrastructure that is currently in place? To answer these questions, we build a model of a blockchain for securities settlement and study its design. The blockchain we study has two distinctive features. First, we assume that the blockchain handles ownership transfers of both securities and payment. This enables a DvP mechanism and, thus, the blockchain has the potential to directly rule out settlement risk. Second, we assume that the *blockchain is permissionless*. There are no designated, third parties that are in charge of updating the information stored on the blockchain.[6]

The blockchain in our paper is thus a record-keeping system that keeps track of securities ownership as well as payments related to securities trades. For securities trades to be settled, the transaction information (transfer of ownership and payment) needs to be recorded on the blockchain. To this end, the investors involved in the trade communicate this information to a peer-to-peer network that is charged with updating the blockchain.

The updating of records is based on a proof of work (PoW) protocol. A group of transaction validators – called miners – is tasked with solving a computationally intensive problem. Whoever solves the problem first obtains a reward and is allowed to update the blockchain. Such competition helps to protect the integrity of the blockchain. After a securities transaction has taken place and been communicated to the network, an investor can still undo it by creating a *fork* on the chain

[6] Alternatively, a blockchain could be *permissioned* where some trusted parties have been designated to update and manage the information stored. Not surprisingly, this model has been favoured by existing financial intermediaries. Other start-ups have worked on a permissionless blockchain such as CoinSpark and Colu based on colored coins technology in the bitcoin network and Lykke who is working on an integrated, secure global marketplace for the exchange of different financial assets.

which is an alternative history of records. If the investor trying to fork in fact wins the competition, he can convince the entire network that the transaction has not been conducted. To avoid such settlement fails, the blockchain system needs to offer large rewards to make the mining competition sufficiently hard.

The key issue here is that mining is a public good. Once there is a sufficient amount of mining activities, settlement fails can be prevented independent of the total number of transactions, making settlement a free resource. Hence, blockchain-based settlement systems need to make settlement a scarce resource in order for investors to pledge transaction fees that generate rewards for mining.

A blockchain can generate fees by limiting the speed at which it is updated. This can be achieved by restricting the block size (how many transactions can be included in each new record) and the block time (how frequently new records are incorporated). When investors have a desire to settle early, congestion on the blockchain can exploit their willingness to pay for early settlement. In essence, congestion creates competition for fast settlement, making the blockchain a club good.

Congestion, however, creates two types of costs for investors. First, investors have to pledge transaction fees that have to be sufficiently high to discourage incentives to fork the chain. Second, settlement lags arise as settlement becomes a scarce resource. The system, thus, needs to balance transaction fees and settlement speed while ensuring that the blockchain is tamper-proof. When settlement is too costly or too slow, investors will choose not to conduct a transaction. Hence, the feasibility of a trustless blockchain depends on whether the system can rule out settlement fails in a cost-effective manner.

A trustless blockchain tends to be more viable for an asset market with a large volume of small transactions. This insight is again related to the public good character of settling trades on a blockchain. The benefit for revoking a trade is related to the individual transaction size, while the cost of doing so depends on the mining reward which is related to the aggregate transaction volume. Larger trades tend to have larger incentives to cause settlement fails; larger volume raises the potential for the blockchain to generate mining rewards. In addition, the incentive to fork the chain increases with the trade exposure to post-trade price movements. Thus, assets with lower price risk are more conducive to blockchain-based settlement. Finally, a blockchain is more viable for time-critical transactions because investors are willing to pay a higher fee for timely settlement. This enables the blockchain to raise more mining rewards.

The optimal design of a permissionless blockchain chooses a block time and block size to maximize the expected net trade surplus. We first derive a constraint that summarizes the incentives for any investors not to fork and cause a settlement fail. This constraint becomes tighter as congestion for settlement decreases. Consequently, one cannot set arbitrarily large block sizes to speed up settlement. Interestingly, one cannot set an arbitrarily low block time either. A shorter block time implies that the total number of blocks needed to be support settlement over a time interval increases. Total rewards have to be split over more blocks reducing the reward per block and, hence, mining competition. This introduces a trade-off between faster block time and smaller block size. Overall, we show that it is optimal to choose the block time and block size that jointly minimize the time to settle all transactions over a trade period, while still generating sufficient fees to rule out settlement fails.

We then calibrate our model to the US corporate debt market to provide an estimate of the gains from blockchain-based settlement. Assuming that intentional forking incurs a small fixed cost, we find that trades can be settled quite cost-effectively on a permissionless blockchain. For a block time of 5 minutes, the optimal block size would optimally lead to a throughput rate of 2.6 transactions per second. This implies an average settlement time of 148 minutes and fees of roughly 0.34 bps per trade.

Interestingly, these results could be improved further by lengthening block time and, simultaneously, increasing block size to satisfy the condition that there are no settlement fails. We find that a block time equal to about 27 minutes is optimal together with a very large block size. This would be a vast improvement relative to the existing settlement regime which has a settlement cycle of $T+2$. Using our calibration, we find that the gains from moving to faster settlement fall in the neighborhood of about 1-4 bps: investors would still prefer a permissionless blockchain even if one subsidized a legacy settlement system by this amount.[7]

To the best of our knowledge, our work is the first paper that explicitly models the distinctive technological features of a blockchain for asset settlement and investigates its feasibility and optimal design both qualitatively and quantitatively. It is still uncertain whether and in what form this technology will be adopted to reform securities settlement systems. For instance, Pinna and Ruttenberg (2016) envisage different potential future scenarios. At one extreme, DLT could be fully implemented via a permissionless blockchain; at the other extreme, the existing core players could

[7]Our result, of course, depends on the calibrated utility gains from faster settlement of trades.

simply adopt the database features of a distributed ledger technology to improve internal efficiency, while still relying on known intermediaries to update the database. It is reasonable to conclude that different systems could be adopted for different environments depending on specific factors such as market structure, characteristics of participants and assets and the regulatory framework.

We have chosen to look at a permissionless blockchain based on PoW as it is currently the most well understood implementation of the technology. Notwithstanding, we extend our analysis to point out that neither the trade-offs nor our conclusion change radically under different assumptions. Using alternative protocols that do not rely on sunk resources still give rise to the problem of forking while incurring a range of other potential issues for trusted record-keeping. Similarly, designating trusted parties such as brokers to maintain and update a blockchain will still, at least qualitatively, feature the same trade-offs.

The academic literature on blockchain while growing rapidly is fairly small and mostly focused on computer science.[8] Research in this area has concentrated on the incentives of miners in a blockchain (see for example Eyal and Sirer (2014) or Sapirshtein et al. (2016)) and technological aspects such as scalability (see for example Croman et al. (2016) or Eyal et al. (2016)). A main limitation of this literature is that it does not model the underlying transactions which determine both the value of the system and the users' incentives to tamper with the system. Modeling the value of trades (and their settlement) for investors is necessary to derive their willingness to pay fees for settlement and their incentives to intentionally fork the chain to cause a settlement fail. Using actual data from a specific asset market makes it then possible to estimate the gains for investors to move to blockchain-based settlement.

The economics and finance literature on blockchain is also very thin. Most contributions are of empirical nature and focus on cryptocurrencies such as Bitcoin.[9] For example, Huberman et al. (2017) and Easley et al. (2017) study mining activities and transaction fees in the Bitcoin network. Two exceptions are Cong et al. (2017) who model the formation of decentralized consensus on a blockchain and Biais et al. (2017) who provide a game-theoretical analysis of strategic mining and how it influences consensus about blockchain information.

Our own work on cryptocurrency (Chiu and Koeppl (2017)) is also related to this paper. It shares

[8]The Bitcoin system popularizing the idea of DLT in a financial context was first proposed by Nakamoto (2008).

[9]Harvey (2016) provides an overview of different issues related to Bitcoin and cryptofinance. Aune et al. (2017) discuss information leakage when trading in distributed ledgers.

the idea that any blockchain analysis needs to consider the incentives of participants to alter the ledger to their advantage. Indeed, many essential features of the blockchain (e.g. its consensus protocol, mining, reward scheme) are precisely introduced to ensure the immutability of the ledger. Hence, both papers develop incentive constraints – albeit different ones –that the optimal design of a blockchain has to respect and that, therefore, limit the benefits of using the technology.

The key difference of this paper to Chiu and Koeppl (2017) is that the latter studies double-spending in a blockchain for goods trading. A blockchain for securities trading is fundamentally different from a cryptocurrency system for goods trading for several reasons. First, in a securities settlement system, both securities and cash are digital assets recorded in digital ledgers. Hence DvP can be ensured automatically by a smart contract as discussed in Section 2. This is not possible for goods trading as the ownership of goods that are traded is typically not recorded on the blockchain. To ensure DvP there, settlement needs to be delayed until a transaction has been confirmed sufficiently often in the blockchain. This introduces a trade-off between a verification lag and higher mining rewards that is absent with blockchain-based securities settlement. Second, the incentive problem in a goods trading environment is often asymmetric. Buyers have an incentive to cheat by "double spending", while merchants are typically more trustworthy. In contrast, an asset transaction is usually subject to a two-sided incentive problem as both the buyer and the seller can have an incentive to default via a settlement fail. Third, the price of financial assets tends to fluctuate more than that of goods, with larger price movement over shorter horizons. As a result, the incentive to revoke a financial transaction are higher and, hence, settlement speed is more critical. All these factors play important roles in our model causing the analysis to be fundamentally different from our earlier work.

The paper proceeds as follows. Section 2 briefly reviews blockchain technology. In Section 3, we introduce the trading environment, while in Section 4 we model how financial trades are settled on a permissionless blockchain without trust. Section 5 and Section 6 study the optimal design of a permissionless blockchain qualitatively and quantitatively. Our paper concludes by presenting some extensions and showing that our insights are more broadly applicable (Section 7 and 8).

2 A Brief Review of Blockchain Technology

It is useful to first review the basic idea of how to use blockchain technology for settling security trades. A securities settlement system facilitates the transfer of legal ownership of financial assets among investors. Traditionally, this function has been performed by a trusted third-party. This party maintains a centralized ledger which records the ownership of securities and the transfer thereof by crediting and debiting buyers and sellers' accounts after every transaction.

Distributed ledger technology (DLT) – often referred to simply as blockchain – allows for the verification, updating and storage of the record of transaction histories without the use of a designated third-party. It relies on a single ledger that is distributed among many different parties, but that is updated without having a dedicated central administrator. There are two basic versions of DLT. The first one is *trustless* where anyone can access and potentially update the ledger. Consequently, it is often referred to as a *permissionless blockchain*. In the other version, some institutions or individuals are entrusted with direct access to the blockchain and with updating it. Hence, the expression used commonly is *blockchain with trust* or *permissioned blockchain*.

What is a blockchain and what problem does it solve?

The key problem for keeping digital records such as ownership is *time stamping* that ensures that a transaction has been conducted at a particular time in the past. This is achieved via a blockchain. Transactions are grouped into *blocks* at particular times and consecutively chained together over time to form a *blockchain*. Hence, the blockchain contains the entire history of past transactions that can be used to create a ledger to verify asset ownership.

To ensure that this record of ownership is accurate, each block is built upon the previous one. Hence, to change a past block, one needs to change all blocks that have been created since that particular block. If this becomes more costly as the chain increases, older blocks are more secure and one can trust the information stored in the block. Distributing the blockchain among participants provides a decentralized dataset architecture that permits the storage and sharing of transaction records without the need of a third, central party.

Traditional payment and settlement systems rely on a trusted third party, such as a central bank or other specialized entities such as central security depositories or custodians to manage a single, centralized ledger. Such systems have to rely on settlement lags to verify and consolidate informa-

tion from many parties that are involved in an asset trade. With DLT, this information can be shared directly making costly third-party intermediation unnecessary and allowing to consolidate trade reporting and trade reconciliation that is often duplicated and incompatible across different participants in a securities transaction.

How can a permissionless blockchain create a trusted record of ownership?

A permissionless blockchain allows a peer-to-peer network to collectively manage a digital ledger even when there is no central authority and when participants potentially have conflicting interests. To do so, the blockchain relies on a decentralized network of validators to maintain and update copies of the digital ledger. But since the system is permissionless, anyone can be a validator updating the blockchain. Having no a priori trust or even anonymous validators, this raises the key issue that all new information incorporated in the blockchain is accurate and that all participants agree with it. How does a permissionless blockchain then achieve consensus among its participants?

First of all, to reach a consensus in a network, validators need to compete for the right to append a new block to the chain. This competition can take different forms. In the most common consensus protocol, Proof-of-Work (PoW) this process is called mining and involves solving a computationally difficult problem (aka the proof of work). The winner of this mining competition has the right to update the chain with a new block. This ensures that there is agreement within the peers of the network about new blocks being added to the blockchain.

The second issue is that the system needs to prevent users from tampering with the blockchain by proposing fake transactions in the mining process. Each owner of the digital balances relies on a private-public key pair to protect the ownership. Cryptography ensures that only private key holders can move the balances recorded on the ledger, using their public key to prove ownership. Hence, a dishonest user cannot spend other users' balances without compromising the associated private keys.

Still, a dishonest user can potentially eliminate transactions that have been initiated either by himself or by other users. A user might have an incentive to do so when he wants to revoke his own past transaction that has already been agreed upon and broadcasted to the network. To do so, the dishonest user needs to create an alternative history of transactions which involves winning the mining competition against honest miners. If such an attack succeeds, the cheater can undo the transaction by convincing the entire network of an alternative history where the transaction

has never occurred.

Hence, the third issue is to defend the system against such attacks. With a PoW protocol, this is automatic since it is difficult and costly to win competition. Since the probability of winning is proportional to the fraction of computational power owned by a miner, sufficient mining activities help safeguard the blockchain against these attacks. Of course, mining is costly, so that honest mining activities need to be properly incentivized through rewards. These rewards can either be seignorage in a cryptocurrency system such as Bitcoin or, more generally, with transaction fees being paid to the winner of the mining competition. While there are many other consensus protocols, PoW is the only really tested one shown to be successful in the form of the original Bitcoin blockchain in the context of permissionless blockchains.

What is different in a permissioned blockchain?

In a permissioned blockchain with trust, only a set of trusted validators with known identities has access to the blockchain and can update it. As discussed above, in a permissionless blockchain, the consensus protocol and the reward scheme need to be designed properly to prevent users from tampering with the blockchain. In contrast, in a permissioned system, ordinary users cannot tamper with the blockchain directly, but validators still need to behave honestly which can either be achieved through economic incentives or legal enforcement.

The comparison between the two systems thus depends on the degree of commitment and enforcement. If users' incentives were not an issue, then there would be little need to consider a permission-based system or to find a tamper-proof consensus protocol. Similarly, with trusted validators, the optimal design of a blockchain still needs to consider the validators' incentives to tamper with the blockchain. Hence, whether DLT can reap significant advantages relative to existing, centralized securities settlement systems will depend on the costs of providing such incentives or a tamper-proof consensus protocol.

How could blockchain technology improve current settlement arrangements?

Current post-trade settlement arrangements that rely on a designated third party tend to be slow and inefficient. This is mainly related to the nature of dispersed information in the trading process and the costs of reconciling this information.[10] Traditional settlement systems are typically forced

[10]For a detailed overview of these costs see CPSS (2017).

to impose relatively long settlement cycle (typically T+2 or T+3) and fixed fees for settlement. In contrast, blockchain technology can reduce information costs by providing a common, public ledger which can be accessed and shared among all participants. This allows a blockchain-based system to also offer flexibility in settlement times and costs. It can introduce time-varying settlement times that depend on actual needs of markets and participants instead of being based on technological constraints. As participants choose how fast to settle, they can either save costs by accepting longer lags or ensure additional benefits by settling faster for a higher fee.

A second, technological advantage is that – unlike traditional settlement systems – a blockchain does not require intermediaries to ensure Delivery-vs-Payment. In a blockchain-based system, this can be ensured by a *smart contract* which is a self-enforcing, autonomous program without the support of any intermediaries.[11] Trades involving multiple legs such as a transfer of security and a cash payment can then be executed either in their entirety or not at all. In database systems, this is referred to as an *atomic transaction* which is an indivisible and irreducible series of database operations such that either all occur, or none occur.

Furthermore, blockchain-based securities settlement could improve the functioning of markets that are too thin to warrant formal settlement arrangements. Examples are markets like private equity where transactions are few and infrequent or cross-border trading where there are no common record-keeping systems in place. These markets are often dominated by financial institutions that provide expensive services that substitute for settlement systems. A viable threat from settling on a permissionless blockchain can reduce market power and thus can provide indirect costs savings for market participants.[12]

3 Trading Environment

We model a single trading period where investors meet bilaterally and negotiate the terms of trading an asset. The trading period is the interval $[0, 1]$, with trades being negotiated at time $t = 0$. Trades need to be confirmed and settled by transferring ownership of the asset and making a payment at some later time $T \in [0, 1]$.

[11] See Wall and Malm (2016).

[12] Similarly, new infrastructures could be built for assets that so far were not tracked or for markets that were too fragmented to warrant a settlement infrastructure such as diamonds or art work.

There is a measure $\mathcal{M}$ of risk-neutral *sellers* who are endowed with one unit of an indivisible asset. Each asset delivers a dividend denoted by δ. At $t = 0$, sellers have a marginal valuation of the asset given by u_ℓ. There is also a measure $\mathcal{M}$ of *buyers* who do not own an asset, but have a higher valuation $u_h > u_\ell$ than sellers at $t = 0$. This gives an incentive to trade. We assume that these $\mathcal{M}$ sellers and buyers are matched bilaterally to trade at $t = 0$.[13] In addition, these investors value a numeraire good linearly that is used for payments. Once a trade has been settled at $T \in [0, 1]$, the dividends of the asset and the payment are consumed. We assume that there is no discounting.

After transaction is agreed, the investors' valuation of the asset is subject to a random shock. Specifically, the valuation of the buyer and the seller can reverse according to an exponentially distributed random shock with an arrival rate λ. When this shock materializes, the trade surplus turns from positive to negative. This gives rise to a preference for early settlement as a transaction is more time-critical when λ is high. Furthermore, the dividend is subject to a random shock that is also exponentially distributed with an arrival rate ν. Conditional on receiving a shock, the shock is distributed symmetrically across $\mathbb{E}(\delta)$ with extreme values given by $\bar{\delta} = \mathbb{E}(\delta) + \varepsilon_\delta$ and $\underline{\delta} = \mathbb{E}(\delta) - \varepsilon_\delta$. This dividend shock will give investors an incentive to strategically default on trades negotiated at $t = 0$.

In a bilateral trade, the buyer agrees to pay p units of the numeraire good in exchange for the asset. Suppose the transaction is settled at time T and is subject to a transaction cost τ that is shared equally among the buyer and seller. The expected surplus from trading for the buyer is given by

$$S_b = \left[e^{-\lambda T} u_h + (1 - e^{-\lambda T}) u_\ell\right] \mathbb{E}(\delta) - p - \frac{\tau}{2}. \tag{1}$$

The buyer has a high (low) valuation if the trade is settled before (after) the valuation shock hits. Similarly, the seller's surplus is given by

$$S_s = p - \left[e^{-\lambda T} u_\ell + (1 - e^{-\lambda T}) u_h\right] \mathbb{E}(\delta) - \frac{\tau}{2}. \tag{2}$$

Note that the costs of a delayed settlement arise from the valuation shock and that – due to risk neutrality – only the expected dividend $\mathbb{E}(\delta)$ matters. Assuming that the surplus is split equally between the buyer and the seller, we have the following result.

[13] We later extend the framework to incorporate trading frictions that give a role for liquidity provision and broker intermediation. For now only the total number of trades $\mathcal{M}$ matters.

Lemma 1. *The transaction price is*

$$p = \frac{u_h + u_\ell}{2}\mathbb{E}(\delta), \tag{3}$$

and, given a settlement time T, the expected surplus from trade is

$$S = S_b + S_s = \left(2e^{-\lambda T} - 1\right) V_0 - \tau, \tag{4}$$

where $V_0 = (u_h - u_\ell)\mathbb{E}(\delta)$.

The price at which the asset is traded is thus independent of the settlement time T, while the total surplus decreases when settlement is delayed. This introduces a trade-off between faster settlement and transaction costs τ. Investors are willing to pay a positive fee τ to shorten the settlement delay T, since transactions are time critical.

4 Settlement on a Permissionless Blockchain without Trust

Suppose there is no centralized or intermediated arrangement to settle trades. Instead, assets and payments are recorded on a public ledger in the form of a *blockchain*. The blockchain is updated over time in a distributed fashion by competing miners who record the investors' instructions to transfer assets and payments to one another. The authenticity of these transfer instructions is ensured by using cryptography and delivery-vs-payment (DvP) can be guaranteed by executing an *atomic transaction*. The main threat to the security of the blockchain is then what we call a *forking attack*: after a transaction has been agreed upon and sent to miners, either the seller or the buyer attempts to alter the blockchain so that the original transaction is inconsistent with the public ledger. Being successful amounts to a default on the original transaction which we call a *settlement fail*.[14]

4.1 Blockchain for Payments and Asset Holdings

The trading period is divided into $\bar{N}$ consecutive discrete subperiods, $n = 1, ..., \bar{N}$, so that each subperiod has length $\Delta = 1/\bar{N}$. There are publicly observable balances of assets and numeraire

[14]For example, an investor can unilaterally default on a trade by giving ownership of the asset or the payment back to himself at a different account in the ledger. This can make it impossible to reconcile the original trade instructions with the ledger and is similar to a *double spending* attack in a cryptocurrency system.

goods denoted by $a_n(i) \in \{0,1\}$ and $m_n(i) \in I\!R_+$ at the start of subperiod n owned by investor i. Due to anonymity, an investor is allowed to hold multiple balances. We use $\mathcal{S}_n = \{a_n(i), m_n(i)\}$ to denote the entire public record of these balances, called the *public state*. We take as given the initial public state $\mathcal{S}_0$ and assume that the initial balances of the numeraire good are large enough to finance asset trading and transaction fees.

In a distributed network, trades are settled by validating that they are feasible and by updating the public state through a process called *mining*. After a trade at $t = 0$, the seller owning entry i and the buyer owning entry j jointly broadcast a transaction message about the terms of trade. We denote this message by $(\Omega^a(i,j), \Omega^m(i,j), \tau) \in \{0,1\} \times I\!R_+ \times I\!R_+$ which specifies that an asset from entry i is to be transferred to entry j against a payment Ω^m from entry j to i involving a transaction fee τ.

In every subperiod, miners compete to update the public state. Updates are in the form of a *block* where miners group transaction messages together, verify that the transfers specified in messages are feasible and earn the right to propose a block upon winning a competition. We denote the block for subperiod n by $\mathcal{B}_n = \{(\Omega^a(i,j), \Omega^m(i,j), \tau)\}$. A message can be incorporated into block n if it is *feasible*,

$$0 \leq \Omega^a(i,j) \leq a_n(i) \tag{5}$$

and

$$0 \leq \Omega^m(i,j) \leq m_n(j). \tag{6}$$

These restrictions make sure that entry i (j) has sufficient assets (numeraire goods) to transfer.

A trade is *settled in subperiod* n if the associated message is incorporated in block n and, consequently, when the public state has been updated according to

$$a_{n+1}(i) = a_n(i) + \sum_j \left[\Omega^a(j,i) - \Omega^a(i,j)\right] \tag{7}$$

$$m_{n+1}(i) = m_n(i) + \sum_j \left[\Omega^m(i,j) - \Omega^m(j,i)\right] \tag{8}$$

to reflect the change in ownership of the asset and the numeraire good. The sequence of blocks $\mathcal{B} = \{\mathcal{B}_n\}_{n=1}^{\bar{N}}$ is publicly observable and generates a sequence of public states $\mathcal{S} = \{\mathcal{S}_n\}_{n=1}^{\bar{N}}$ given $\mathcal{S}_0$ by the updating rule specified above. We call this sequence of blocks a *blockchain*.[15]

[15]The Bitcoin blockchain stores only the sequence of transactions and not the state, $\mathcal{S}$. However, one can easily generate the entire state from the history of transactions. The Ethereum blockchain includes both the state and the transactions in its blocks.

4.2 Mining

There are M miners who compete solving a *Proof-of-Work* problem for each of the $\bar{N}$ subperiods. The miner who wins the competition in a subperiod can propose the block to update the blockchain. Investing computing power q, the probability that a miner solves the problem within a time interval t is given by an exponential distribution with parameter μ

$$F(t) = 1 - e^{-\mu t} \tag{9}$$

where $1/\mu = D/q$ is the expected time to solve the problem. Aggregating over all M miners, the first solution among all miners, is also an exponential random variable with parameter $\sum_{i=1}^{M} \mu_i$. The expected time needed to complete the proof-of-work is thus given by

$$\frac{D}{\sum_{i=1}^{M} q_i}. \tag{10}$$

The parameter D captures the difficulty of the PoW and can be adjusted appropriately so that the expected time for a solution is equal to the frequency at which a block is generated. For simplicity, we assume for the rest of the analysis that exactly one block will be mined per subperiod.[16]

Any particular miner j will be the first one to solve the PoW and propose a new block with probability

$$\phi_j = \frac{q_j}{\sum_{i=1}^{M} q_i}. \tag{11}$$

The winner of the block receives R units of the numeraire good. An individual miner's maximization problem is then given by

$$\max_{q_j} \phi_j R - q_j \tag{12}$$

where we have normalized the cost of computing power to be 1. The FOC is given by

$$\frac{\sum_{i=1}^{M} q_i - q_j}{\left(\sum_{i \neq j} q_i + q_j\right)^2} R = 1. \tag{13}$$

Imposing symmetry across the M miners, $q_i = Q$, we obtain the following result.

Lemma 2. *The expected profit for a miner is*

$$\Pi = \frac{1}{M^2} R \tag{14}$$

[16]Since investors and miners are both risk neutral, the distribution of arrival times of a solution to the PoW is irrelevant for the remainder of our analysis.

and the total computing cost incurred by the mining community is

$$MQ = \frac{M-1}{M}R. \tag{15}$$

Note that, as $M \to \infty$, the expected profit converges to zero and the total computing cost converges to R.[17] Competition dissipates the rent from mining. We assume that the reward R is financed by transaction fees collected from trades and that the winners of the $\bar{N}$ blocks share the total reward equally across the trade period so that the reward per block remains constant across all blocks.

4.3 Settlement Fails through Forking

In a securities settlement system, the transfer of the asset and the payment are linked and jointly recorded on the blockchain. In an atomic transaction, it is infeasible for one side to undo one of the transfers unilaterally. This means that a trade automatically involves *delivery-vs-payment (DvP)* so that one cannot steal the security or the payment. Notwithstanding, the buyer or seller can still eliminate the transaction by creating a fork. In practice, this involves reassigning balances or assets to a different entry so as to invalidate the original transaction. If such forking is successful before the original transaction has been included in the blockchain, the transaction will be declared infeasible, which effectively amounts to an investor strategically defaulting on the trade. In other words, one needs to *avoid intended forking altogether* to ensure that there are *no settlement fails* on the blockchain.

In our framework, a buyer has an incentive to default if the price is higher than the value of the security, while a seller has an incentive to do so when his value of the security is larger than the price received in the trade. To ensure that the blockchain has no settlement fails, we need to focus on the worst possible scenario where the incentives to default are highest. The buyer's maximum payoff from a default is

$$V_b = p - u_\ell \underline{\delta} \geq 0 \tag{16}$$

while a seller's maximum payoff from a default is

$$V_s = u_h \bar{\delta} - p \geq 0, \tag{17}$$

with $V = \max\{V_b, V_s\}$. For simplicity, we assume here that all transaction fees are still being paid with forking.

[17] Later on, we show that setting $M \to \infty$ is indeed optimal.

4.4 Secret Mining

To create a fork, an investor needs to invalidate a transaction by including a message into a block that causes the original message to be infeasible given the public state.[18] Hence, a dishonest investor needs to win the mining game against honest miners just once. We call such an attempt *secret mining*. Depending on the precise details of the blockchain and the features of the securities market, secret mining may be less efficient or more costly than honest mining. We introduce a parameter $\alpha \geq 1$ to capture that the variable costs of secret mining are potentially higher than honest mining.[19] Furthermore, we allow a secret miner to be subject to a fixed cost $\Gamma < V$ which can represent the short-run cost of installing and adjusting hardware, the expected punishment if detected, or the potential loss of relationship with trading partners.

Since we want to exclude settlement fails, we look at a dishonest investor that has the largest incentives to fork the blockchain. He solves

$$\max_{q_0} \left(\frac{q_0}{MQ + q_0} \right) (V + R) - \alpha q_0 - \Gamma. \tag{18}$$

There is an extra return V from winning a block and creating a fork. The investor still has to compete against all the honest miners taking their computational power of MQ as given and facing additional costs of mining (α, Γ). His optimal investment in secret mining is given by

$$q_0 = MQ \left(\sqrt{\frac{(V + R)}{\alpha MQ}} - 1 \right). \tag{19}$$

which implies the following result.

Proposition 3. *As $M \to \infty$, there are no settlement fails on any trade if*

$$V \leq \Gamma + 2\sqrt{R\alpha\Gamma} + R(\alpha - 1). \tag{20}$$

The exposure V is the private reward for an investor to fork and, thereby, default on the trade. If the private reward becomes too big, secret mining is profitable and there will be settlement

[18] For example, the investor can change his asset holdings or his holdings for the numeraire good at the relevant entry by sending holdings to another entry he controls.

[19] One can show that this assumption is equivalent to assuming that the efficiency of a secret miner is lower than that of honest miners. Specifically, assume that a dishonest investor can win the mining game only with probability

$$\phi(q_0) = \frac{\theta q_0}{MQ + \theta q_0}.$$

Hence, the efficiency of secret mining increases with θ. For $M \to \infty$, we obtain the same results with $\theta = 1/\alpha$.

fails with blockchain-based settlement. Similarly, when secret mining becomes cheaper, it is more difficult to avoid default. Importantly, increasing the block reward R can counter secret mining. A larger reward fosters competition among miners and, thus, makes forking less attractive as miners increase their investment. This will be the key channel for designing the blockchain to rule out settlement fails.

4.5 Block Size and Block Time

We call the time interval $\Delta = 1/\bar{N}$ between two consecutive updates on the blockchain *block time*. As Δ increases, the blocks are updated less frequently. The *block size* B limits the total number of transactions included in any block. Given the block size, transaction messages sent to miners can be ranked in terms of the transaction fees τ and partitioned into blocks that contain at most B transactions. The first block $\mathcal{B}_1$ consists of the B transactions that pay the highest fees and is incorporated as the first update on the blockchain at time Δ. The second block $\mathcal{B}_2$ includes the next B transactions and is added at time 2Δ, and so on.

Given a block size B, the number of blocks needed to settle all transactions is thus given by

$$N = ceiling\left[\frac{\mathcal{M}}{B}\right] \tag{21}$$

where $ceiling[x]$ the smallest integer greater than or equal to x. The total time spent on settlement is $T = \Delta N$. In the analysis below, we treat $\mathcal{M}/B$ as an integer for simplicity.[20] Also, we assume that all transactions can be settled in the interval $[0, 1]$, i.e., that $T \leq 1$ or, equivalently, that $\bar{N} \geq N$.

5 System Design of Permissionless Blockchain

5.1 Transaction fees

With their messages, investors announce transaction fees for validation. There will be a diminishing sequence of threshold fees $\tau_1 \geq \tau_2 \geq ... \geq \tau_N$ which give the minimum fee τ_n required for validation in block n. In other words, a transaction offering a fee τ will be validated in block $n' \leq n$ if and

[20] In the numerical exercise, we consider the general case where $\mathcal{M}/B$ can take non-integer values so that the last block is only partially filled.

only if $\tau \geq \tau_n$. For transaction fees to be optimal, we require that no investor can be better off by changing his transaction fee. Since all transactions are identical, this implies that the surplus to be the same across all transactions. Otherwise, some investors be better off by changing their transaction fees. This implies that transaction fees are described by the first order difference equation

$$\tau(n-1) - \tau(n) = 2\left(e^{-\lambda\Delta(n-1)} - e^{-\lambda\Delta n}\right) V_0 \tag{22}$$

for all $n = 2, \ldots, N-1$ with the boundary condition $\tau(N) = 0$.

Restricting block size makes early settlement a scarce resource for which investors have to compete by posting positive fees. These fees decrease with later blocks as settlement is delayed. In the last block, N, there is no incentive to post any transaction fees anymore as settlement is delayed to a maximum.[21]

Condition (22) then implies that

$$\tau(n) = 2V_0\left(\rho^n - \rho^N\right), \tag{23}$$

where $\rho = e^{-\lambda\Delta} \in (0,1)$ is akin to a discount factor across blocks. Differentiating, we obtain

$$\frac{d\tau(n)}{d\rho} = 2V_0\rho^{n-1}\left(n - N\rho^{N-n}\right). \tag{24}$$

Hence, we have the following result.

Lemma 4. *When ρ is sufficiently close to 1, $\tau(n)$ increases with time criticality λ and block time Δ.*

Investors are willing to pay higher fees when transactions are more urgent and when the time delay between blocks increases. Aggregating transaction fees, we obtain that the total reward financed by transaction fees is given by the expression in the following lemma.

Lemma 5. *The total reward for mining is given by*

$$\bar{N}R = 2V_0\left[B\left(\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}}\right)\left(1 - e^{-\lambda\Delta N}\right) - \mathcal{M}e^{-\lambda\Delta N}\right]. \tag{25}$$

[21] In the numerical exercise, if N is not an integer, the last block N is only partially filled, but still features zero transaction fees.

The total mining reward depends on the block size B according to a function of the form

$$f(B|\rho) = B\left(\frac{\rho}{1-\rho}\right)(1-\rho^{\frac{\mathcal{M}}{B}}) - \mathcal{M}\rho^{\frac{\mathcal{M}}{B}} \tag{26}$$

The function f is positive, bounded, continuous and is equal to 0 for $B = 0$ and $B = \mathcal{M}$. Hence, there are two opposing effects of block size on revenue. Holding fees constant, as the block size increases more revenue is generated since more trades can be settled in earlier blocks at a higher transaction fee. As the block size increases, however, the fee investors are willing to pay decreases. The intuition is that investors only pay fees when early settlement is scarce. Reducing the block size creates a congestion effect so that investors need to compete for early settlement by posting larger fees. For $B \to 0$, the first effect dominates while for $B \to \mathcal{M}$ the second effect does.

5.2 Equilibrium without Settlement Fails

Given the design parameters (B, Δ) of the settlement system, an equilibrium is defined as a fee schedule $\tau(n)$ such that (i) the fee schedule satisfies the first-order difference equation (22), (ii) there are no settlement fails, i.e. equation (20) is satisfied and (iii) the surplus is positive for all trades.

By setting a block time Δ, the settlement system determines the total number of blocks $\bar{N}$ in a trading period. The revenue per block is given by

$$R = 2V_0\frac{f(B|\rho)}{\bar{N}}. \tag{27}$$

Using this result in constraint (20) to ensure that there are no settlement fails, we obtain

$$V \leq \Gamma + 2\sqrt{2V_0\left(\frac{f(B|\rho)}{\bar{N}}\right)\alpha\Gamma} + 2V_0\left(\frac{f(B|\rho)}{\bar{N}}\right)(\alpha - 1). \tag{28}$$

Whether or not this constraint can be satisfied – and, hence, whether an equilibrium without settlement fails exists – depends on two factors. First, the maximum default exposure V cannot be too large relative to the ex-ante surplus V_0. Second, secret mining cannot be too cheap. Ultimately, the requirement that there are no settlement fails puts limits on the minimum block reward that is necessary to avoid settlement fails, which in turn depends on the block size and block time. We will discuss this issue in more depth in Section 6 below.

Proposition 6. *For any parameters (α, Γ) and ex-ante surplus V_0, an equilibrium without settlement fails exists only if the maximum default exposure V is not too large.*

5.3 Optimal Block Size and Block Time

For characterising how to set block size and block time optimally, we use the expected net surplus as our objective function

$$\mathcal{W}(B) = \sum_{n=1}^{N} B\left(2\rho^n - 1\right) V_0 - \sum_{n=1}^{N} B\tau(n). \tag{29}$$

This function consists of the gains from trade less mining costs which are equal to the total transaction fees as we look at the case where $M \to \infty$ so that all mining revenue is spent on computational investments. The optimal design of a permissionless blockchain then maximizes the expected net surplus $\mathcal{W}(B)$ subject to the constraint (28). We can rewrite the objective function as

$$\mathcal{W}(B) = \mathcal{M}\left(2\rho^N - 1\right) V_0 = \mathcal{M}\left(2e^{-\lambda\Delta N} - 1\right) V_0. \tag{30}$$

Hence, the optimal block size and block time minimizes the *total settlement lag* ΔN where the number of blocks $N = \mathcal{M}/B$ is inversely related to block size. Note that it is infeasible to set block time to 0 and block size to $\mathcal{M}$ whenever the fixed costs of secret mining Γ are sufficiently low as this would generate no revenue and, thus, would violate the constraint (28).

This gives us the following trade-offs for setting block size and block time. Reducing the block size delays settlement of time-critical trades, thus reducing welfare. At the same time, however, reducing the block size creates congestion which is necessary for investors to pay transaction fees to compete for scarce early settlement. Lower block sizes help raise rewards to finance mining activities, relaxing the constraint (28).

A shorter block time increases the discount factor ρ for settlement and thus raises the expected net surplus $\mathcal{W}(B)$. However, shortening block time has an ambiguous effect on the constraint (28). First, the function $f(\rho|B)$ that defines total rewards for a given block size is non-monotone in ρ. Second, since $\bar{N} = 1/\Delta$, a shorter block time leads to more blocks $\bar{N}$ over which the total revenue needs to be distributed. There is thus a cost for lowering block time. This implies that one cannot set block time arbitrarily low.[22] Given the discreteness of block size, we cannot characterize analytically how to jointly set the optimal block size and block time and resort to numerical exercises in the next section. However, we have the following partial result on comparative statics.

[22] We interpret here our time interval $[0, 1]$ as fixed. In reality, this corresponds to the operating time of the settlement system with trades arriving throughout the interval. If one could dynamically adjust block time in response to trading demand, we could interpret $B\bar{N} = \frac{B}{\Delta} \geq \mathcal{M}$ as the required capacity of the blockchain in a particular time interval. One can show that a faster block time can then always be supported with lower block sizes.

Proposition 7. *The optimal block size and block time are set to minimize the total settlement lag ΔN that is consistent with no settlement fails.*

Given Δ (B), the optimal block size (block time) increases (decreases) with the number of trades $\mathcal{M}$ and the trade surplus V_0, but decreases (increases) with the incentives to secretely mine (i.e. lower costs of secret mining (α, Γ) and a larger default exposure ε_δ).

The intuition for these results is straightforward. First, a larger trade volume implies higher revenue for the settlement system that can be used to incentivize honest miners with larger rewards. Similarly, a larger trade surplus increases the willingness to pay for earlier settlement, raising rewards for honest mining. Hence, one can increase the block size or have a faster block time to speed up settlement. Second, whenever the threat of secret mining increases, one needs to create more congestion through smaller block sizes and slower block times. This is the case when secret mining becomes less costly and when the maximum gain V from a settlement fail increases.

6 Quantitative Results

6.1 Optimal Block Size

We now continue our analysis with a quantitative illustration of a permissionless blockchain. Our objective is to see (i) whether settlement on a permissioned blockchain is feasible for some financial markets and (ii) how such a system compares with existing legacy systems where transactions are settled only within a few business days. To do so, we calibrate our model to the US market for corporate debt using aggregate statistics from the TRACE reporting system.

In our benchmark calibration, a trading period is set to eight hours and we first keep the block time fixed at five minutes (i.e. $\Delta = (12 \cdot 8)^{-1}$) to focus on the effects of varying the block size.[23] Guided by the statistics from TRACE (see Mizrach (2015)), we set the total number of trades to $\mathcal{M} = 45000$ and the individual trade size to $\mathbb{E}(\delta) = 1$ which we interpret to be in million dollars.[24] The maximum default exposure is assumed to be 3%, i.e. $\varepsilon_\delta = 0.03 \cdot \mathbb{E}(\delta)$, within a trading day.[25]

[23]The average block time is about 10 minutes for Bitcoin and about 12-15 seconds for Ethereum.

[24]The daily average number of trades was about 45,000 in 2015. During the period from 2007 to 2013, the average daily transaction size was $1.2mn for the 1,000 most actively traded and $0.6 mn for less actively traded bonds.

[25]The extreme values of 2-year, 5-year and 10-year treasury bonds in 2012 imply deviations of respectively 0.21%,

Following Duffie, Garleanu and Pedersen (2007), we set $u_h - u_\ell = 0.01$ to capture the surplus from trade given our daily frequency. Furthermore, the preference shock on average arrives once every six months ($\lambda = 0.0056$). This translates into about 250 trades being affected by the valuation shock every day. Finally, we assume that a secret miner faces the same computational costs as a professional miner ($\alpha = 1$), but is subject to a fixed cost $\Gamma = 0.01 \cdot \mathbb{E}(\delta)$. For the benchmark, this gives a value of \$10,000.

Table 6.1: **Benchmark Parameter Values**

$\bar{N}$	Δ	$\mathcal{M}$	$\mathbb{E}(\delta)$	u_h	u_ℓ	λ	ε_δ	α	Γ
96	0.0104	45000	1	1.005	1.995	0.0056	0.03	1	0.01

Figure 6.1 reports the optimal block size (B), along with the average settlement time for a trade, $\left(\sum_{n=1}^{N} Bn\Delta\right)/\mathcal{M}$, and the average transaction fee per trade, $\left(\sum_{n=1}^{N} B\tau(n)\right)/\mathcal{M}$. For the benchmark case, we obtain an optimal block size of 774 transactions. Hence, the blockchain supports about 2.6 transactions per second.[26] The implied average settlement time per trade is 148 minutes, with the average fee per trade being equal to \$34 or roughly 0.34 basis points of the average trade size of \$1mn.

We next look at the effects of increasing transaction volume, time criticality and the maximum default exposure. The first row of Figure 6.1 shows that larger trade volume ($\mathcal{M}$) allows the blockchain to increase block size (see Proposition 7). The reason is that more trades imply a larger total of transaction fees. Hence, less congestion is needed to support sufficient rewards in order to discourage secret mining. In our benchmark, this also leads to a shorter average settlement time and less transaction fees.

This is important as a permissionless blockchain is thus scalable from an economics point of view. Mining is a public good for validating individual transactions: once there is a sufficient amount of mining activities – or, equivalently, transaction fees – secret mining can be prevented independent of the total number of transactions. This is because the benefit for an investor to create a fork is

1. 35% and 3.39%. We assume that the volatility of corporate bond prices is close to the upper bound of this range.

[26] This is in the range of current blockchain implementations. Bitcoin – limited by its block size – has a throughput limit of about 7 transaction per second and Ethereum currently seems to be able to handle about 25 transactions per second. Also, note that transactions in our model are not spaced out over the entire trading period. Thus, one can think of the benchmark throughput as peak rates.

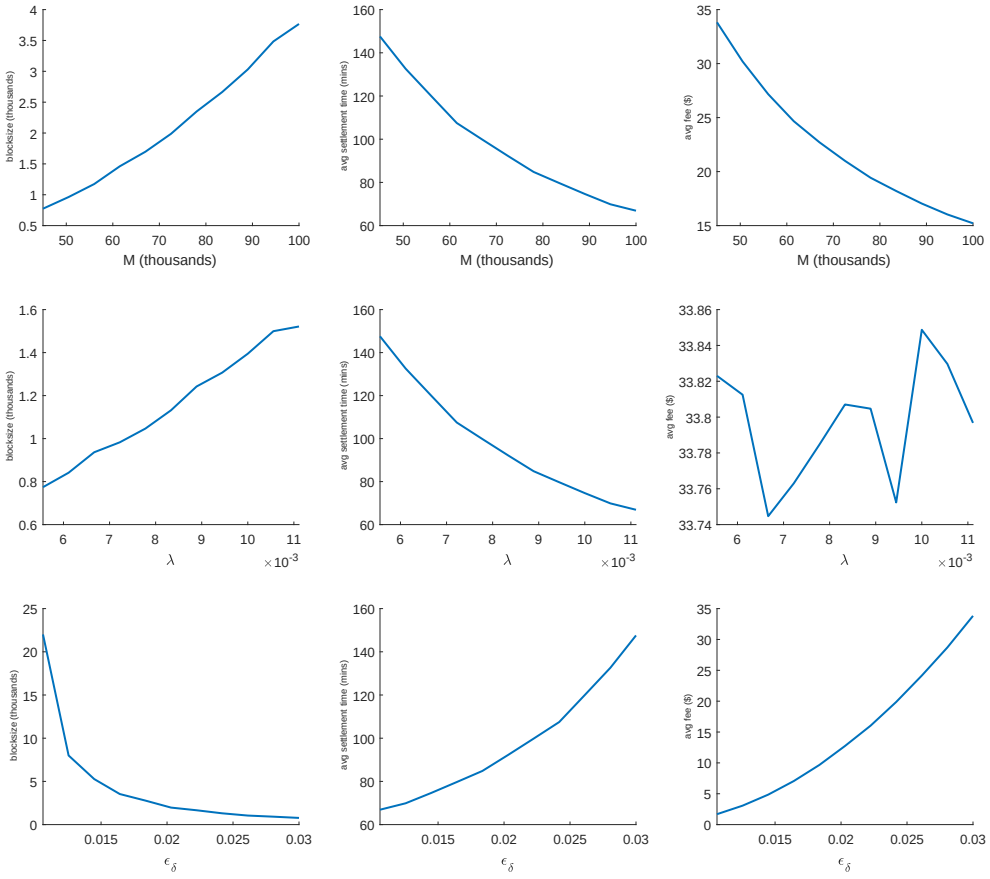

Figure 6.1: Optimal block size, implied average settlement times and fees

related to the *individual* transaction size, while the cost of doing so depends on the mining reward which is related to the *aggregate* transaction volume. Consequently, a permissionless blockchain tends to be more efficient for settling a large volume of small transactions.

The second row in Figure 6.1 shows the effects when trades become more time critical, i.e., when λ increases. Investors then have a larger value for settling trades early. According to Lemma 4, this increases the transaction fees investors are willing to pay for a given block size and, thus, allows the system to increase block size to speed up the average settlement time. There are then two opposing effects on the average transaction fees. On the one hand, given a block size, higher λ increases the investors' willingness to pay high fees to speed up settlement. On the other hand, whenever the block size increases, congestion is lower so that investors need to compete less for fast settlement. This explains why λ has a non-monotonic impact, but increasing trend on the average transaction fee as shown in the right panel of the figure.

Our last comparative statics exercise looks at how the maximum default exposure matters for the design of blockchain-based settlement. As ε_δ increases, higher mining activities are necessary to counter increased incentives for secret mining. As implied by Proposition 7, the system now needs to reduce block size to create more congestion in order to raise the rewards for honest mining.

6.2 Optimal Block Time

Figure 6.2 shows the effects of shortening block time Δ from 5 to 4 minutes, i.e., we now set $\Delta = 0.0083$. This also increases the total number of blocks available within an 8 hour trading period to 120. As discussed above, the total reward raised by transaction fees is now spread out over more blocks. In addition, according to Lemma 4, a shorter block time tends to reduce investors' incentives to pay a high fee to compete for an earlier settlement. The incentives to fork, however, remain the same for every block. To ensure that there are no settlement fails, the optimal system needs to restrict block size to generate extra revenue. The optimal block size given a block time of 4 minutes is once again the largest one that satisfies constraint (28). Even though some trades settle faster, average settlement time actually increases here despite having more frequent blocks.

What is then the optimal block time for our calibration? Table 6.2 reports the optimal block size and time together with the implied transaction fees and settlement times. Surprisingly, for our benchmark it is optimal to choose a fairly long block time which allows the system to increase block

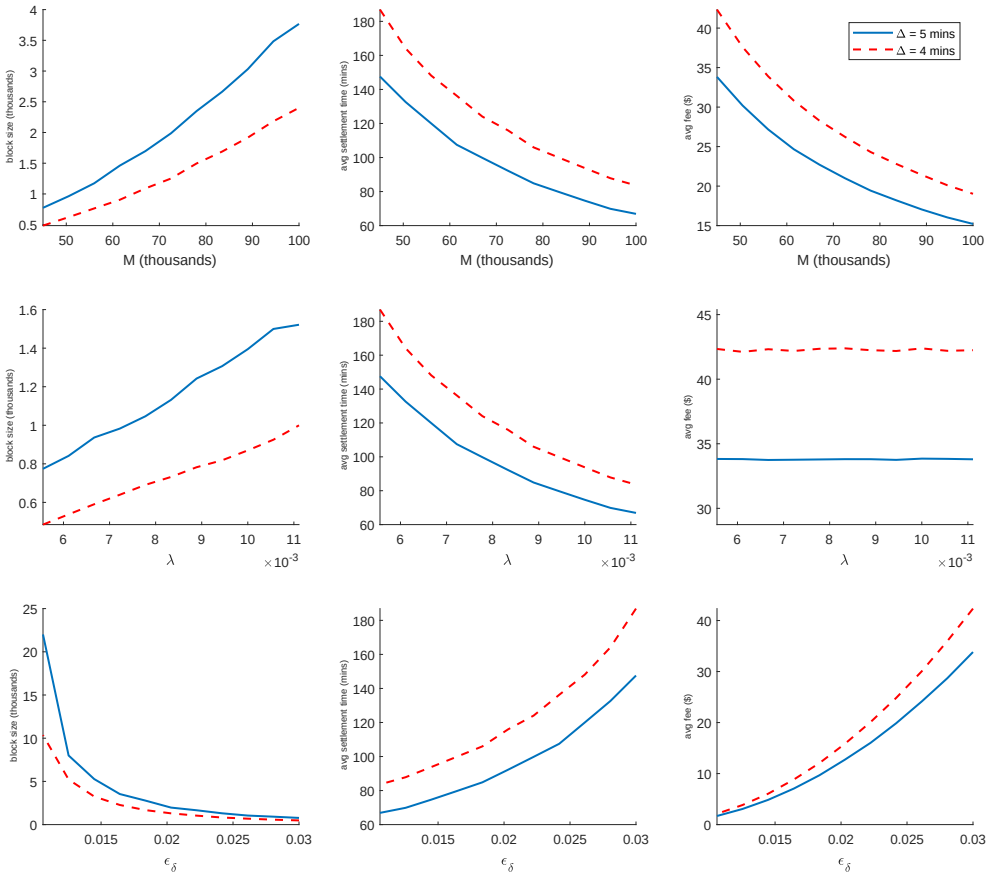

Figure 6.2: Effects of shorter block time Δ

size. This cuts both the average settlement time and the transaction fees by about 1/5 relative to when block time is five minutes. The total number of blocks per trading period also reduces from 96 to about 18. The implied total revenue from settlement fees is about $270,000 per trading period.[27]

When trades become more time critical, investors are willing to pay more in order to settle faster. This allows the system to speed up settlement with faster block times keeping similar block sizes. Finally, as the default exposure becomes less extreme, block time can be shortened further as there are less incentives to fork. Interestingly, however, it is also optimal to restrict block size more.

Table 6.2: Optimal System Design

	Block size (thousand)	**Block time (min)**	**Avg fee ($)**	**Avg settlement time (min)**
Benchmark	44.31	27.25	6.18	26.93
λ =0.0111	44.41	19.25	8.75	19.04
ε_δ =0.01	42.25	6.25	1.35	5.89

6.3 Feasibility of Blockchain

In the benchmark exercise, we calibrate the model to the market for corporate debt and show that an optimally designed permissionless blockchain can provide a viable settlement system for this market. We ask now for what characteristics of assets, markets and investors is a permissionless blockchain a feasible option.

Figure 6.3 illustrates how the feasibility of a permissionless blockchain depends on some key characteristics. Specifically, we plot the minimum trade volume that is required for ruling out settlement fails for a permissionless blockchain using three different values for how time critical trades are: the preference shock arrives on average quarterly ($\lambda = 0.0111$), bi-annually ($\lambda = 0.0056$) or annually

[27]The optimal block time is derived under the assumption that the length of the trading period, $\Delta\bar{N} = 1$, is fixed. If it were adjustable, it would be optimal to set the minimum feasible block time necessary to settle all transactions that exclusively take place at time $t = 0$. To fully study the trade-off in setting the length of the trading period, however, one should modify the model to allow transactions to arrive sequentially over time. This extension is beyond the scope of this paper and thus is left for future research.

($\lambda = 0.0028$). Settlement without fails in a permissionless blockchain is feasible for the region above the curves shown in each of the figures. Note that we once again fix block time at the benchmark value except for the last figure, but vary block size to make settlement possible.

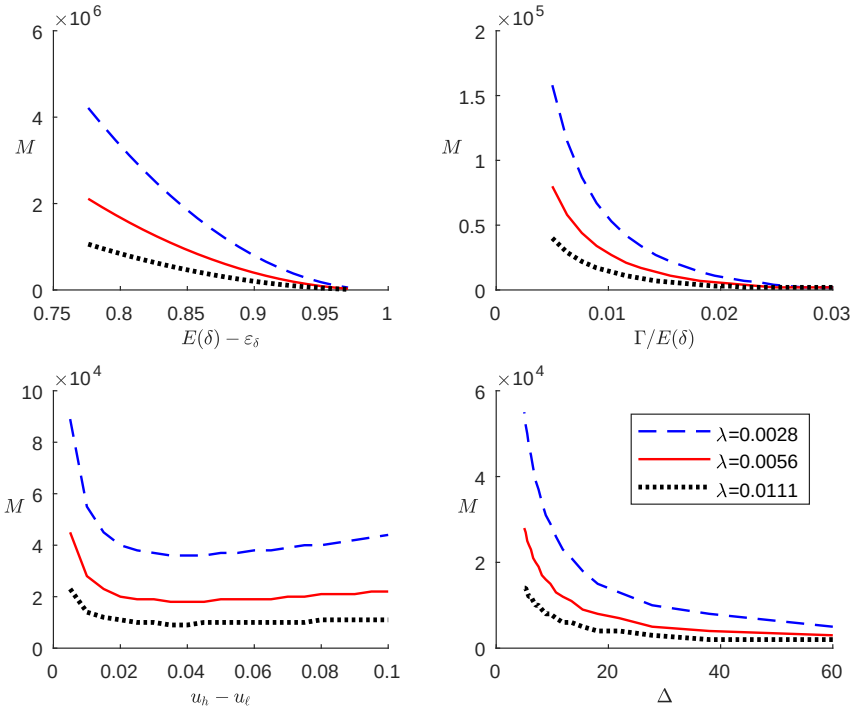

Figure 6.3: Feasibility of blockchain-based settlement

In general, a high enough trade volume $\mathcal{M}$ makes settlement on a permissionless blockchain system feasible. This is a direct consequence of the public good nature of mining with a PoW protocol. More transactions mean more revenue ceteris paribus that can finance rewards for miners. Similarly, as trades become more time critical (higher λ), the investors' willingsness to pay transaction fees for earlier settlement increases, thus raising revenue for mining.

The feasibility region in each subfigure depends then on how some other variable influences the constraint (28). More extreme default exposures cause settlement fails to become more attractive and, hence, require larger revenues for mining. The opposite is the case when the costs of secret mining are large relative to the trade size (larger $\Gamma/\mathbb{E}(\delta)$). The gains from trade $u_h - u_\ell$ have a U-shaped effect. When $u_h - u_\ell$ is small, investors have low incentives to pay high fees, tightening the constraint. When $u_h - u_\ell$ is very high, the incentives to default eventually increase, once again tightening the constraint. Consequently, using a permission-based blockchain is more viable for

intermediate values of $u_h - u_\ell$.

The most counterintuitive result is that shortening block time potentially interferes with blockchain-based settlement. As shown in the last panel of Figure 6.3, there is a minimum block time so that such settlement is feasible. This is for two reasons. First, the rewards per block are becoming increasingly small as block time falls. And second, investors are less willing to pay to speed up settlement. Hence, shorter block times require either a larger market or more time ciritcal trades.

6.4 Blockchain vs. Legacy Settlement System

We now compare the performance of a permissionless blockchain to a conventional centralized settlement system. In a centralized system, a trusted third party maintains and updates a centralized ledger. Hence, a costly PoW protocol is not required. Furthermore, the centralized system imposes a fixed settlement lag, currently $T + 2$ for corporate bonds, and a fixed transaction fee due to technological and institutional constraints. To the contrary, a blockchain-based settlement is more flexible along both dimensions.

Table 6.3 compares the efficiency of a legacy settlement system relative to that of a blockchain. Specifically, we compute the transaction fees (as a fraction of the transaction value in basis points) that a representative investor is willing to pay at $t = 0$ for choosing the centralized system over a permissionless blockchain. When the fee reported in the table is negative, the investor would switch to a blockchain based system even when settling in a centralized system is free. In that case, a subsidy is needed to induce investors to use the legacy system. Our assessment depends of course on the calibrated gains from faster settlement (λ) and the costs associated with mining in the PoW protocol so that there are no settlement fails ($\mathcal{M}$ and ε_δ).

The three cases shown in the table vary these parameters. For the benchmark cases (in bold), we obtain that investors are willing to pay a fee in the neighbourhood of 2 bps (or about \$200 given the normalized trade size of \$1mn) to move from costless $T+2$ settlement to a blockchain-based system. The advantage of blockchain-based settlement diminishes as settlement time T_{cen} in the centralized system falls and naturally disappears when settlement is close to immediate ($T_{cen} \to 0$).[28] Once again, larger trading volume, higher time criticality, and lower default exposure make settlement

[28] One should interpret these results as upper bounds. Delaying settlement can have also advantages for investors and centralized settlement systems often provide other services.

Table 6.3: Legacy system vs. Blockchain (equivalent fees in bps)

Case I: $\varepsilon_\delta = 3\%$, $\mathcal{M} = 45000$

$\lambda \backslash T_{cen}$	0	0.1	0.5	1	2
0.0111	0.69	0.47	-0.41	-1.51	-3.68
0.0056	**0.68**	**0.57**	**0.13**	**-0.42**	**-1.52**
0.0037	0.68	0.61	0.31	-0.05	-0.79

Case II: $\varepsilon_\delta = 3\%$, $\mathcal{M} = 100000$

$\lambda \backslash T_{cen}$	0	0.1	0.5	1	2
0.0111	0.32	0.10	-0.78	-1.88	-4.05
0.0056	**0.31**	**0.20**	**-0.24**	**-0.79**	**-1.89**
0.0037	0.31	0.23	0.06	-0.43	-1.16

Case III: $\lambda = 0.0056$, $\mathcal{M} = 45000$

$\varepsilon_\delta \backslash T_{cen}$	0	0.1	0.5	1	2
1%	0.03	-0.08	-0.52	-1.07	-2.17
2%	0.25	0.14	-0.30	-0.85	-1.95
3%	**0.68**	**0.57**	**0.13**	**- 0.42**	**- 1.52**

on a blockchain more attractive.

7 Extensions

In this section, we consider several extensions of the basic model. We first look at the optimal degree of mining and then study a slightly altered setup that allows for heterogenous investors or trades. Finally, we briefly look at the case of a permissioned blockchain with trust as well as the role of brokers for blockchain-based settlement.

7.1 Optimal Number of Miners

In the benchmark, we assume $M \to \infty$. Here, we show that competitive mining is indeed optimal in a blockchain that is based on a PoW protocol. For any given number of miners M, the constraint to rule out settlement fails is given by

$$V \leq \Gamma + 2\sqrt{\left(\frac{M-1}{M}\right) R\alpha\Gamma} + \left(\alpha\left(\frac{M-1}{M}\right) - 1\right) R. \tag{31}$$

The total mining cost is $\frac{M-1}{M}R$ which is a deadweight loss financed by transaction fees. Consider now increasing the number of miners, but keeping mining costs constant by lowering the reward R correspondingly. This relaxes the constraint without increasing the deadweight cost associated with the PoW protocol. Hence, one can increase block size and, thus, achieve faster and cheaper settlement. This shows that it is optimal to have $M \to \infty$, or equivalently, for mining to be as competitive as possible.

7.2 Heterogeneous Investors

We now consider a more general case where agents are heterogeneous with respect to their liquidity shock λ. Specifically, there are I types of investors with preference parameters given by $\lambda_1 > \lambda_2 > ... > \lambda_I$. Denote the fraction of type i by π_i so that the number of investors of type i is $\mathcal{M}_i = \pi_i \mathcal{M}$. Note that all investors have the same V and V_0 so that the constraint (28) is the same for all i.

The trade surplus for type i

$$S_i(n) = \left(2e^{-\lambda_i \Delta n} - 1\right) V_0 - \tau \tag{32}$$

is increasing in i. Hence, investors with less time critical trades have higher surplus. Lemma 4 implies that, for ρ sufficiently close to 1, type i chooses to validate earlier than type $j > i$. One can then define thresholds $\underline{n}_i, \bar{n}_i$ such that type i chooses to settle between the $\underline{n}_i$-th and the $\bar{n}_i$-th blocks. Transaction fees are defined iteratively by

$$\tau(n) = 2\left(\rho_i^n - \rho_i^{\underline{n}_{i+1}}\right) V_0 + \tau(\underline{n}_{i+1}) \tag{33}$$

for $n \in \{\underline{n}_i, ..., \bar{n}_i\}$ with $\tau(\underline{n}_{I+1}) = 0$.

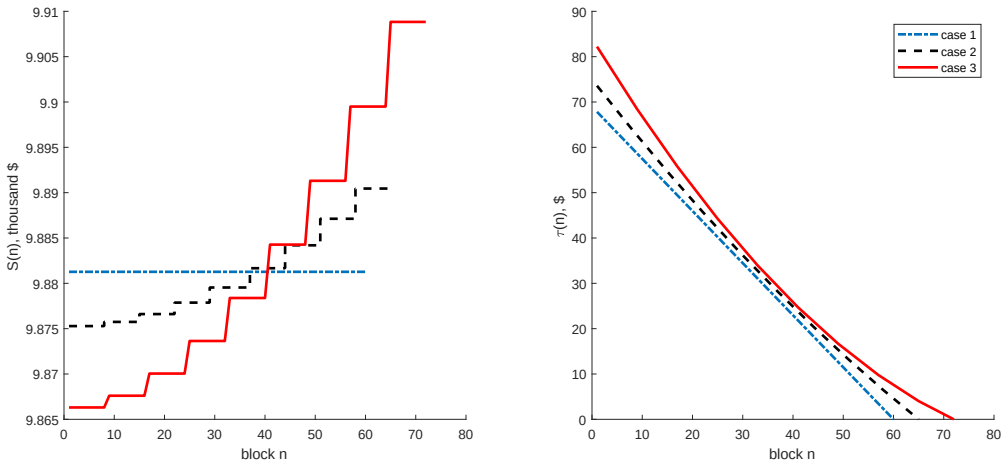

Figure 7.1: Heterogeneous Investors

Table 7.1: Impact of Heterogeneous Investors

	Avg. settlement time (min)	Avg. settlement fee ($)	Median Surplus ($)
Case I	153	33.89	9881
Case II	165	34.11	9880
Case III	180	33.71	9878

We use a numerical example to illustrate the effects of heterogeneous preferences on the optimal block size given the block time is 5 minutes. We set $I = 9$ and look at three different cases. Case 1 is the benchmark case where all investors are homogeneous with $\lambda_i = 0.0056$ for all i. In

case 2, investors have heterogeneous preferences such that $\lambda_i \in [0.0044, 0.0067]$. In case 3, the range of preference heterogeneity increases further to $\lambda_i \in [0.0028, 0.0083]$. Figure 7.1 depicts the equilibrium surplus and transaction fees of investors settling in different blocks. Interestingly, as heterogeneity increases, the optimal block size decreases. Investors with preferences for fast settlement sort into early blocks even though at the expense of larger transaction fees. Investors with less time-critical trades choose to delay settlement. Not surprisingly, as investors' preferences become more heterogeneous, transaction fees become more responsive to settlement time. That is, impatient investors need to bear a larger fraction of the total mining cost than more patient investors. Looking at the median investor who does not change across the three cases, we find that his average settlement time increases, but will small variations in fees and overall surplus.

7.3 Heterogeneous Transaction Sizes

Investors can have different transaction sizes. To model this, we assume that the expected payoff of the assets traded takes the form $\mathbb{E}(\delta_i) = \mathbb{E}(\delta)(1 + \xi_i)$. Note that a larger transaction size leads to higher prices and higher trade surplus V_0. Hence, investors with larger transaction sizes have an incentive to settle earlier and, hence, pay higher transaction fees.

We again look at three cases with different degrees of heterogeneous transaction sizes: (i) the benchmark case where $\xi_i = 0$ for all i, (ii) $\xi_i \in [-0.1, 0.1]$ and (iii) $\xi_i \in [-0.15, 0.15]$. The maximum exposure V increases with ξ_i so that investors with larger transaction sizes have a larger incentive to secretly mine. Hence, holding the block reward R constant, the constraint that rules out more settlement fails tightens with larger transaction sizes. This implies that one needs to increase the block reward R to ensure that there are no settlement fails.

Even though some investors are willing to pay higher transaction fees, Figure 7.2 shows that the block size has to decrease with more heterogeneous transaction sizes in order to generate sufficient rewards for mining. This once again arises from the public good character of mining. Only the largest transaction size drives how much mining is necessary to avoid forking by investors. While all other transactions benefit from mining, they would only require a smaller reward fraction to safeguard against secret mining.

To compare with the homogeneous case, we can look at a median investor having $\xi_i = 0$. This

investor now pays higher transaction fees and the average settlement of his trade is pushed back.[29] Hence, his expected trade surplus declines.

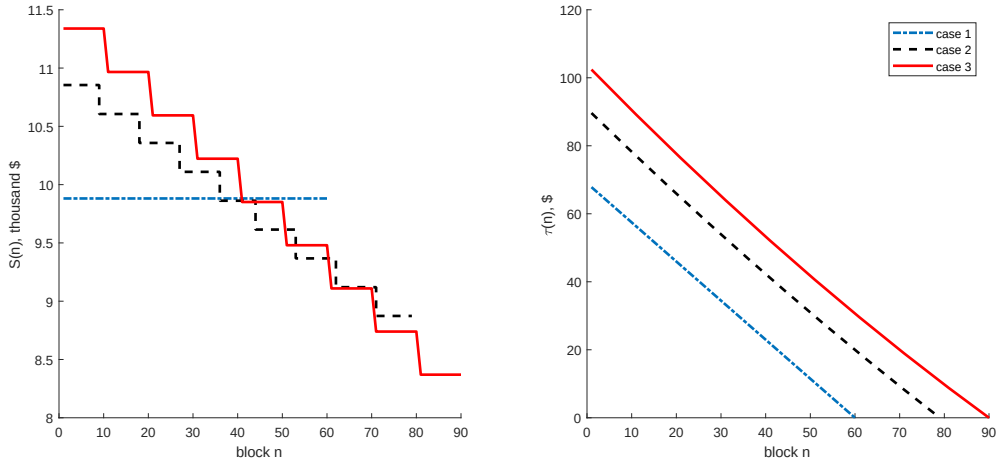

Figure 7.2: Heterogeneous Trades

Table 7.2: Impact of Heterogeneous Trades

	Settlement time (min)	Settlement fee ($)	Median Surplus ($)
Case I	153	33.89	9881
Case II	200	44.03	9862
Case III	228	49.84	9851

7.4 Intermediation by Brokers

Brokers provide not just clearing functions, but also improve clients' access to markets and provide trade financing. Hence, if blockchain-based settlement is not compatible with broker intermediation, then the gain from moving to a blockchain will be reduced. In this section, we consider the extreme

[29] Note that a similar situation arises if assets had different risk characteristics $\varepsilon_{\delta,i}$, but the transaction size would be constant for all investors, i.e., $\mathbb{E}(\delta_i) = 1$. This again tightens constraint (28) with the most risky transaction determining necessary block rewards and block size.

case where brokers can intermediate trades when settlement is centralized, but cannot intermediate trades at all when a permissionless blockchain is used for settlement. Specifically, we introduce brokers that improve liquidity in the market by mitigating search frictions and providing trade financing. Our goal is again to compare a blockchain system without brokers and a legacy system with brokers.

Suppose investors value the dividends with different marginal utility u at time 0: $u = u_h$ with probability κ and $u = u_\ell$ otherwise, where $u_h > u_\ell \geq 0$. Due to preference shocks and different asset holdings, there are then four investor types denoted by ho, lo, hn, ln. Their measures are $s\kappa$, $s(1-\kappa)$, $(1-s)\kappa$ and $(1-s)(1-\kappa)$, respectively. The lo types are potential sellers and the hn types are potential buyers. All other investors are inactive. At time 0, asset sellers and buyers are randomly matched in pairs to negotiate a trade. The number of matches is determined by a matching function

$$\mathcal{M} = \chi \min\{s(1-\kappa), (1-s)\kappa\}, \tag{34}$$

where $\chi < 1$ captures matching efficiency. The unconditional probability of a buyer finding a seller to trade with is then given by

$$\chi \min\{\frac{s(1-\kappa)}{(1-s)\kappa}, 1\}. \tag{35}$$

To pay for a transaction, we assume that buyers need to carry the numeraire good in advance.

When settling on a legacy system, investors can use brokers that provide two services. First, a buyer trading through a broker can have better access to markets. We capture this by assuming the matching efficiency is larger, $\chi_b = 1 > \chi$, when a broker is used. Second, brokers can help buyers save on the costs of carrying the numeraire good by allowing buyers to *trade on margin*. Since only a fraction of a broker's clients will have a trading opportunity, each buyer only needs to carry $m < p$ units of the numeraire good. The rest can be borrowed from the broker and will be repaid at the end of the trading period.

Extending our benchmark calibration, we follow Duffie, Garleanu and Pedersen (2007) to set $s = 0.8$ so that 80% of investors hold a position. We choose $\kappa = 0.9920$ to get an annual turnover of about 50% as reported by Edwards, Harris, and Piwowar (2004). The brokers margin requirement m is set so that all potential trades can be financed, or

$$m(1-s) = s(1-\kappa)p \tag{36}$$

implying that

$$\frac{m}{p} = \frac{s(1-\kappa)}{1-s} = 3.2\%. \tag{37}$$

Borrowing costs are set equal to 5% per annum.

Using our benchmark calibration, Table 7.3 compares a legacy system with brokers and a permissionless blockchain without brokers. In the first row, we look at the case where $\chi = 1$ so that brokers do not improve matching efficiency, but allow for margin trading. Here, investors are still willing to switch to a blockchain-based system even when fees for margin investment and settlement are zero.

Table 7.3: Legacy System w/ Brokers vs. Blockchain w/o Brokers

$\chi \backslash T_{cen}$	0	0.1	0.5	1	2
1.00	2.00	1.88	1.44	0.88	- 0.22
0.99	2.98	2.88	2.43	1.87	0.77
0.90	11.88	11.77	11.32	10.77	9.66
0.50	51.41	51.29	50.85	50.29	49.19

As brokers start to also have an advantage by matching buyers and sellers, staying in the legacy systems becomes quickly much more attractive. This shows that for blockchain-based settlement to be effective, it needs to accommodate broker services with our estimates indicating how much such services are worth to investors.

7.5 Permissioned Blockchain with Trust

We briefly look at a permissioned blockchain where designated third parties ("validators") are put in charge for updating ownership records. In financial markets, traditional intermediaries such as brokers, dealers or banks perform trade related services beyond settlement and, thus, can be seen as the natural set of validators that have the right to update the blockchain. Hence, the incentives of these intermediaries to maintain their reputation substitutes for the PoW protocol of a permissionless blockchain.

To capture these facts, we develop a stylized setting where N_B brokers intermediate all trades, but

are also the validators of the permissioned blockchain. When a broker intermediates for a client who buys an asset for p, the broker's maximum exposure is given by

$$p - p(\underline{\delta}), \tag{38}$$

as he has to send p to the seller, but the securities obtained have only a value of $p(\underline{\delta})$ after a shock. The expression above is thus the maximum short-fall of the broker in case a client defaults on the trade. Equivalently, the extreme exposure on the sell-side is given by

$$p(\bar{\delta}) - p. \tag{39}$$

These exposures can arise due to margin buying by clients or securities lending to clients for their trades. Importantly, since brokers intermediate a large number of transactions, they can have an unbalanced position and their exposure against either the buy or the sell side can aggregate into large positions. Given this exposure, brokers thus may also have an incentive to create a settlement fail in a permission-based blockchain either on their clients' behalf or to get rid of their own exposure against their clients. For our model, we have that $p(\underline{\delta}) = u_\ell \underline{\delta}$ and $p(\bar{\delta}) = u_h \bar{\delta}$ when a broker fully internalizes the incentives of his clients.

Assume now that the broker has a fraction $1/N_B$ of the B transactions in the block and that they are all either buy-side or sell-side transactions. If the broker has the right to propose a block, he can invalidate all transactions of his clients in a block so that the maximum exposure of the permissioned blockchain would be given by

$$\frac{B}{N_B} \max\{V_s, V_b\} = \frac{B}{N_B} V. \tag{40}$$

The cost for a broker of causing a fork is the potential loss of the charter value from being a validator which is proportional to the expected revenue per broker R/N_B. For simplicity, suppose that each broker/validator proposes a block in every subperiod after which it is randomly decided which proposal is used to update the blockchain. Hence, the probability of getting a block in the chain for a broker is $1/N_B$. Similar to equation (28) in the permissionless case, this gives us a constraint to rule out settlement fails

$$\frac{1}{N_B} \frac{B}{N_B} V \leq \gamma \frac{R}{r N_B} \tag{41}$$

where γ is the probability of losing one's charter value after causing a fork and r is an interest rate

to derive the net present value of remaining a validator.[30]

The block size once again plays a role in controlling how binding this constraint will be. A smaller block size decreases the potential gain from forking while increasing the revenue R from settlement and, hence, the charter value of being a validator. Figure 7.3 depicts the effects of the number of brokers on block size, the average settlement time and average transaction fees for our benchmark parameters and an annual interest rate of $r = 5\%$. As the number of brokers N_B increases, the maximum exposure and the probability for creating a settlement fail falls, thereby reducing the expected gain from cheating. As a result, the system can increase block size and reduce transaction fees. The average settlement time is thus shorter. Naturally, as the detection probability γ drops, the block size needs to be restricted further to limit exposure and to raise more transaction fees in order to incentivize validators.

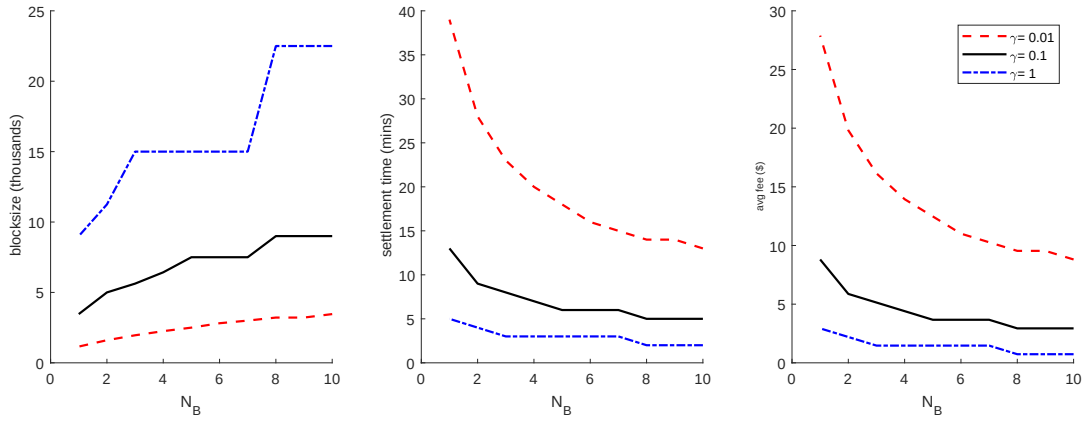

Figure 7.3: Optimal permissioned system

In conclusion, the trade-offs in blockchains with and without trust are similar. Even in a permission-based system – unless the validators are perfectly trustworthy – the optimal design of the system still depends on the validators' incentives to tamper with the blockchain. A key advantage for the permissioned system, however, is that no computational resources are wasted through a socially

[30] More generally, one can think of the constraint taking the form

$$\mathcal{P}(N_B)\frac{B}{N_B}\max\{p - p(\underline{\delta}), p(\bar{\delta}) - p\} \leq L(R, N_B).$$

The function L is a general loss function increasing in block rewards and decreasing in the number of validators. The probability of creating a fork $\mathcal{P}(N_B)$ is likely to depend on the exact specification of the consensus protocol for updating the blockchain, but will be decreasing in the number of validators. For example, in a threshold signature scheme, it might be harder to form a dominating collusion when N_B is large.

costly mining process. Any transaction fees are merely a transfer from investors to validators and do not affect overall welfare. Notwithstanding, how efficient such a system can be depends crucially on how costly it is to rule out settlement fails. This is likely to depend on the precise consensus protocol for updating the blockchain, the identities of validators, as well as the ability of participants to monitor the validation process.[31]

8 Discussion and Conclusion

This paper studies the feasibility and design of a permissioned blockchain for securities settlement. We have focused exclusively on PoW for its consensus protocol. Notwithstanding, the trade-offs involved are not specific to that particular protocol. In any distributed ledger, a set of participants has the delegated authority to process transactions and update the blockchain. These validators, however, sometimes have an incentive to revoke trades by altering the records on the blockchain. To ensure that the blockchain is tamper-proof, the system needs to make dishonest actions sufficiently costly. With PoW, this is based on computational costs. In other protocols like *Proof-of-Stake* (PoS), it is the opportunity cost of holding collateral. Still in other protocols based on voting it would be the cost of compromising validators to achieve a threshold agreement.

Consequently, what is common to all protocols is that honest behaviour needs to be properly incentivized through rewards which are financed by transaction fees collected from investors. Therefore, any blockchain-based settlement system – whether permissionless or permissioned – needs to congest settlement so that users are willing to pay transaction fees. This is precisely what our paper has focused on – settlement on a blockchain needs to be a club good.

It is still an open question whether alternative, potentially cheaper protocols can fully replicate the benefits offered by the original idea of having costly mining to ensure consensus on a distributed ledger (see for example BitFury Group (2015)). It is beyond the scope of this paper to investigate this topic, but we briefly discuss the possibility of a PoS protocol in the appendix.

For policy makers and regulators, three key themes emerge form our analysis. First, to ensure DvP, it is important to link digital ledgers for asset ownership and payments together to support atomic

[31] For example, in a system where a central bank works as a validating notary, the probability of brokers manipulating the blockchain can be seen as very small. Similarly, if extreme penalties can be enforced, then validators have no incentives to cheat.

trades. This could give an explicit role for government who could provide a digital currency that could be used with securities settlement systems.[32] Second, the feasibility of using a blockchain for settlement depends on a sufficient volume of transactions, high enough costs for tampering with the blockchain (possibly in the form of fines) and a limited default exposure. Here, regulation and supervision could play a role to ensure such conditions. Finally, in case of a permissionless blockchain, coordination to adjust its design might prove difficult. Here, a government can help to coordinate the different participants to reach agreement.

[32]Currently, such systems tend to rely on large-value payment systems for payment which are often government or central bank run.

References

Aune, R. T., A. Krellenstein, M. O'Hara, and O. Slama. 2017. Footprints on a Blockchain: Trading and Information Leakage in Distributed Ledgers. *The Journal of Trading* 12.3:5–13.

Benos, E., R. Garratt, and P. Gurrola-Perez. 2017. The Economics of Distributed Ledger Technology for Securities Settlement. Bank of England Staff Working Paper No. 670.

Biais, B., C. Bisière, M. Bouvard and C. Casamatta. 2017. The Blockchain Folk Theorem. Mimeo.

BitFury Group. 2015. Proof of Stake versus Proof of Work. White Paper.

Broadridge. 2015. Charting a Path to a Post-Trade Utility. How Mutualized Trade Processing can Reduce Costs and Help Rebuild Global Bank ROE. Broadridge White Paper.

http://www.broadridge.com/broadridge-insights/Charting-a-Path-to-a-Post-Trade-Utility.html

Chiu, J. and T. Koeppl. 2017. The Economics of Cryptocurrency – Bitcoin and Beyond. QED Working Paper 1389.

Cong, L. W., Z. He, and J. Zheng. 2017. Blockchain Disruption and Smart Contracts. Mimeo.

CPSS (2017). Decentralized Ledger Technology in Payment, Clearing and Settlement. Discussion Paper 157.

Croman, K., C. Decker, I. Eyal, A.E. Gencer, A. Juels, A. Kosba, A. Miller, P. Saxena, E. Shi, E.G. Sirer and D. Song. 2016. On Scaling Decentralized Blockchains. *International Conference on Financial Cryptography and Data Security*. Springer Berlin Heidelberg.

Duffie, D., N. Gârleanu, and L. Pedersen. 2007. Valuation in Over-the-Counter Markets. *The Review of Financial Studies*. 20.6: 1865–1900.

Edwards, A. K., L.E. Harris and M. S. Piwowar. 2004. Corporate Bond Market Transparency and Transaction Costs. Working paper, The Securities and Exchange Commission.

Easley, D., M. O'Hara and S. Basu. 2017. From Mining to Markets: The Evolution of Bitcoin Transaction Fees. Mimeo.

Eyal, I., A.E. Gencer, E.G. Sirer and R. Van Renesse. 2016. Bitcoin-ng: A Scalable Blockchain

Protocol. *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pp. 45–59.

Eyal, I. and E.G. Sirer. 2014. Majority is Not Enough: Bitcoin Mining is Vulnerable. *International conference on financial cryptography and data security.* Springer, Berlin, Heidelberg.

FINRA. 2017. Distributed Ledger Technology: Implications of Blockchain for the Securities Industry. https://www.finra.org/sites/default/files/FINRA_Blockchain_Report.pdf

Harvey, C. 2016. Cryptofinance. SSRN, https://ssrn.com/abstract=2438299.

Huberman, G., J.D. Leshno and C. C. Moallemi. 2017. Monopoly without a Monopolist: An Economic Analysis of the Bitcoin Payment System. Mimeo.

Mainelli, M. and A. Milne. 2016. The Impact and Potential of Blockchain on the Securities Transaction Lifecycle. SWIFT Institute Working Paper No. 2015–007, May.

Mizrach, B. 2015. Analysis of Corporate Bond Liquidity. FINRA Research Notes.

Nakamoto, S. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System.

Oliver Wyman and Euroclear. 2016. Blockchain in Capital Markets – The Prize and the Journey. *White Paper.*

Pinna, A., and W. Ruttenberg. 2016. Distributed Ledger Technologies in Securities Post-trading. Revolution or Evolution. European Central Bank (ECB) Occasional Paper Series no. 172, April.

Wall, E., and G. Malm. 2016. Using Blockchain Technology and Smart Contracts to Create a Distributed Securities Depository. Manuscript.

Santander, Oliver Wyman and Anthemis. 2015. The FinTech 2.0 Paper: rebooting financial services. http://santanderinnoventures.com/wp-content/uploads/2015/06/The-Fintech-2-0-Paper.pdf

Sapirshtein, A., Y. Sompolinsky and A. Zohar. 2016. Optimal Selfish Mining Strategies in Bitcoin. *International Conference on Financial Cryptography and Data Security.* Springer, Berlin, Heidelberg.

A Appendix – Proofs and Derivations

A.1 Proof of Proposition 3

Profits for secret mining are given by

$$\left(\frac{\sqrt{\frac{V+R}{\alpha MQ}}-1}{\sqrt{\frac{V+R}{\alpha MQ}}}\right)(V+R)-\alpha MQ\left(\sqrt{\frac{V+R}{\alpha MQ}}-1\right)-\Gamma. \tag{A.1}$$

These profits have to be negative for there to be no incentives to create a fork. Rewriting, we obtain

$$\alpha MQ\left(\sqrt{\frac{V+R}{\alpha MQ}}-1\right)^2\leq\Gamma. \tag{A.2}$$

Using $MQ=R$ as $M\to\infty$, we obtain

$$V\leq\Gamma+2\sqrt{R\alpha\Gamma}+R(\alpha-1), \tag{A.3}$$

which completes the proof.

A.2 Proof of Lemma 5

The total reward is given by

$$\begin{aligned}\sum_{i=1}^{\bar{N}}R&=\sum_{n=1}^{N}B\tau(n) &\text{(A.4)}\\ &=\sum_{n=1}^{N}B(\tau(n)-\tau(N)) &\text{(A.5)}\\ &=2V_0\sum_{n=1}^{N}B\left(e^{-\lambda\Delta n}-e^{-\lambda\Delta N}\right) &\text{(A.6)}\\ &=2V_0\left[B\left(\frac{e^{-\lambda\Delta}}{1-e^{-\lambda\Delta}}\right)\left(1-e^{-\lambda\Delta N}\right)-\mathcal{M}e^{-\lambda\Delta N}\right]. &\text{(A.7)}\end{aligned}$$

A.3 Proof of Proposition 7

Note that V is a constant and that the RHS of constraint (28) is a continuous function of B. Since the expected surplus from trade $\mathcal{W}(B)$ is increasing in B, the optimal block size B^* is the largest one that satisfies the constraint (28). Furthermore, at B^*, the RHS of the constraint must be decreasing in B. Otherwise, a larger B is feasible and increases surplus.

The comparative statics results follow straight from the constraint (28). A change in ε_δ only influences the maximum exposure V. This implies that an increase in V tightens the constraint and requires an increase in the function $f(B|\rho)$ which one can only achieve by reducing the block size B to generate more revenue. Similarly, increasing the surplus from a trade, V_0, allows a further increase in B. Any changes in the parameters for secret mining, (α, Γ), have similar effects.

The same arguments can be applied to derive the effects on the optimal block time.

B Additional Material – For Online Publication Only

B.1 System Comparison with Heterogeneous Types

Table B.1 compares the blockchain with the legacy system when the degree of time-criticality λ is heterogeneous, while Table B.2 reports the case with heterogeneous trade size $\mathbb{E}(\delta)$. Not surprisingly, investors with more time critical trades have a larger benefit from moving to flexible settlement times on a blockchain. When some trades have a larger default exposure, blockchain-based settlement becomes less desirable for the median investor as the costs for avoiding settlement fails increases.

Table B.1: Legacy Systems vs. Blockchain – Heterogeneous Investors

$\lambda_i \backslash T_{cen}$	0	0.1	0.5	1	2
0.0333	1.52	0.86	- 1.77	- 5.00	- 11.31
0.0111	0.88	0.66	- 0.22	- 1.32	- 3.49
0.0056	**0.57**	**0.46**	**0.02**	**- 0.53**	**- 1.63**
0.0028	0.34	0.28	0.06	- 0.21	- 0.76
0.0014	0.19	0.16	0.05	- 0.09	- 0.36

Table B.2: Legacy Systems vs. Blockchain – Heterogeneous Trades

$\xi_i \backslash T_{cen}$	0	0.1	0.5	1	2
10%	0.90	0.78	0.29	-0.32	-1.52
5%	0.89	0.77	0.31	-0.27	-1.42
0%	**0.87**	**0.76**	**0.32**	**-0.23**	**-1.33**
-5%	0.84	0.74	0.32	-0.20	-1.25
-10%	0.81	0.71	0.31	-0.19	-1.17

B.2 Proof of Stake Protocol

We assume that the right for updating the blockchain is allocated randomly across people. The probability that one can update the chain is proportional to the balances pledged as collateral by a user. Suppose the opportunity cost of each unit of balances is r which might capture the interest rate over the pledging period. An individual validator chooses the collateral balance k_j to solve

$$\max_{k_j} \phi_j R - r k_j, \tag{B.1}$$

where

$$\phi_j = \frac{k_j}{\sum_{i=1}^{M} k_i}. \tag{B.2}$$

The total balance pledged is thus

$$MK = \frac{M-1}{rM} R. \tag{B.3}$$

Similarly, a dishonest investor chooses the collateral the collateral balance k_0 to solve

$$\max \left(\frac{k_0}{MK + k_0} \right) (V + R) - r\alpha k_0 - \Gamma. \tag{B.4}$$

Hence the optimal investment in secret mining is given by

$$k_0 = MK \left(\sqrt{\frac{(V+R)}{r\alpha MK}} - 1 \right). \tag{B.5}$$

The decision problems of honest and dishonest investors are thus equivalent to those under the PoW protocol. However, a *Proof-of-Stake* (PoS) protocol as described above can potentially improve the efficiency of the system for two reasons. First, while computational investment MQ by miners under the PoW protocol is a deadweight loss, the cost of collateral rMK might be less costly from a social point of view. For example, if validators need to pledge cryptocurrecies issued by the system as collateral, then the associated interest cost is simply redistributed within the system. Of course, there might be other costs such as a misallocation of cryptocurrency among participants. Second, the system can potentially penalize cheaters by confiscating their collateral. This can be linked to the costs α or Γ and, thus, directly deter dishonest mining.

Even though a system based on PoS may be cheaper to mitigate incentive problems, it may still be subject to other problems. For example, if all balances were owned initially by one issuer, this issuer can easily propose a long fork starting from the initial block to claim all existing balances in any future dates. As pointed out by BitFury Group (2015), in PoW protocols, this so-called

long-range "attack is prevented by the enormous amount of computational power needed to build the blockchain from scratch; however, this task is within the realm of possibility with proof of stake. As the attacker is able to move coins freely in the blockchain he is building, he has a much higher dimensionality of the search space". Finally, PoS protocols still suffer from other, unresolved issues. For example, it could be difficult to reach a consensus if there is a nothing-at-stake problem or if one needs to ensure that any new participants can decide unambiguously between competing chains. Our model is not built to address these issues.