Blockchains, coordination and forks

By Bruno Biais, Christophe Bisière, Matthieu Bouvard and Catherine Casamatta*

I. Introduction

A ledger is a collection of records, e.g., regarding ownership of an asset. In a distributed ledger, network participants update the ledger when transactions occur. In the Bitcoin distributed ledger, the network participants implementing these updates are called "miners:" End users broadcast their transactions to the network. Miners collect transactions in blocks, which they chain to previous blocks. This gives rise to a chain of blocks, recording consecutive states of the ledger. Two key issues are: Which miner gets to propose each update? And how do miners reach a consensus about a unique ledger?

To address these issues, Nakamoto (2008) proposed the "proof of work" protocol: Miners spend computational resources to solve a numerical problem. The first miner who succeeds is selected to propose an update of the ledger, by chaining his block to previous blocks. Since the numerical problem can only be solved by random trials, the miner selected to update the ledger is randomly drawn. If, as suggested by Nakamoto (2008), the miner chains his block to the longest chain of previous blocks, there is a single chain, embodying consensus on a single ledger.

Biais, Bisière, Bouvard and Casamatta (2019) show that this longest chain rule is a Markov perfect equilibrium of the game played by miners. However, they also show that there exist other equilibria, in which forks arise, i.e., in which the chain of blocks splits in two branches, offering two different versions of the ledger. This result stems from two economic forces: i) coordination effects, reflecting strategic complementarities between miners, and ii) vested interests, reflecting some miners' gains from the persistence of a given branch.

Major forks occurred since 2017 on the Bitcoin blockchain, as illustrated in Figure 1. Miners had to choose between adopting an update in the blockchain protocol or not adopting it. In each of these forks, some miners adopted the updates, but others did not, and this disagreement gave rise to competing branches, operating with competing protocols. In this paper, we rely on the model of Biais, Bisière, Bouvard and Casamatta (2019), to offer an economic analysis of these forks, emphasizing the role of coordination effects and vested interests.

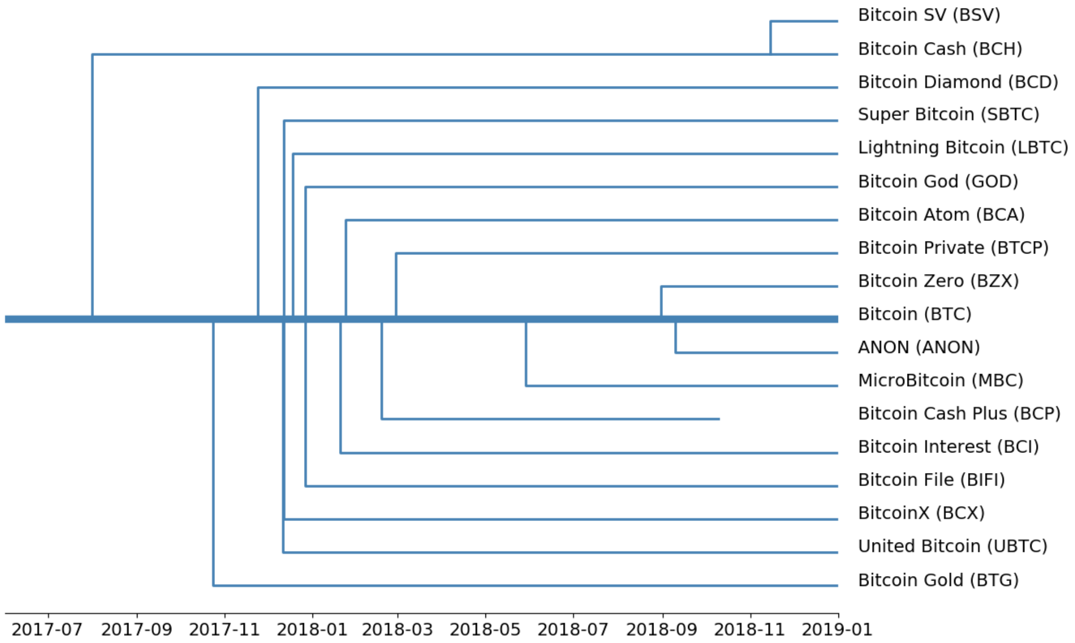
II. Model

Time is continuous, indexed by t, and, for simplicity, we assume there are six risk-neutral miners, $m \in \{1, 2, 3, 4, 5, 6\}$. All miners observe an exogenous flow of transactions that they concatenate into blocks. As explained above, to be selected to append his block to the chain, a miner must solve a numerical problem. In line with practice we assume it takes a random time for miner m to solve his block and that this random time is exponentially distributed, with parameter θ_m. Thus, the probability to solve the numerical problem associated with a block does not depend on the transactions in the block, nor on the time already spent on the problem.

The strategies of the miners map the history of the blockchain into i) which block to mine, and ii) to which previous block to chain it. Whenever a miner does not solve his block, he can either decide to continue mining the same block, or abandon it and

* Biais: HEC Paris, 1 Rue de la Libération, 78350 Jouy-en-Josas, France, biaisb@hec.fr. Bisière: Toulouse School of Economics, Université Toulouse Capitole (TSM-Research), 21 Allée de Brienne, 31000 Toulouse, France, christophe.bisiere@tse-fr.eu. Bouvard: Desautels Faculty of Management, McGill University, 1001 Rue Sherbrooke Ouest, Montréal, QC H3A 1G5, Canada, matthieu.bouvard@mcgill.ca. Casamatta: Toulouse School of Economics, Université Toulouse Capitole (TSM-Research), 21 Allée de Brienne, 31000 Toulouse, France, catherine.casamatta@tse-fr.eu.

Figure 1. Some Bitcoin forks since 2017.



Note: The figure only depicts forks for which market data are available. Each branch starts at the time of the fork, and lasts as long as market data are available.

start mining a new one.

Each block includes a "coinbase" transaction stating that a given number of cryptocurrency units are created and allocated to the miner mining the block. If the miner solves the block, this reward is registered on the branch of the chain to which he appended his block. If the vast majority of the miners are active on that branch, the corresponding version the ledger enjoys consensus, and the units of cryptocurrency it records are valuable. In contrast, if no miner or only one miner is active on that branch, the units of cryptocurrencies are valueless. Thus, we assume the value of the reward earned by a miner, for a block on a given branch of the blockchain, is increasing with the number of miners on that branch. Correspondingly we denote it by $G(K)$, where $K \leq M$ is the number of miners mining on the branch (or branches) that contains the block. For simplicity, we assume $G(6) = G(5) = 1$, $G(4) = \frac{3}{4}$, $G(3) = \frac{1}{2}$, $G(2) = \frac{1}{4}$, and $G(1) = G(0) = 0$.

Finally, we assume that miner m exits the game at $t = z_m$, exponentially distributed with parameter λ_m. At that time, he sells all the rewards collected throughout the game to an otherwise identical miner who replaces him.[1]

III. Hard forks

Suppose that, after the n^{th} block has been solved, an upgrade in the blockchain protocol is proposed, and miners must choose between adopting it and continuing to use the incumbent version. The upgrade triggers a "hard fork:" a block mined with the old protocol cannot be chained to a block mined with the new protocol.[2] This incompatibility precludes the existence of a single chain where both protocols are used. Building on Biais, Bisière, Bouvard and Casamatta (2019), we now offer an equilibrium analysis of participants' choices to adopt the upgrade or not.

[1] This assumption maintains stationarity while keeping the expected profit of a miner bounded, see Biais, Bisière, Bouvard and Casamatta (2019).

[2] This is different from a soft fork, in which blocks mined with the old protocol can be chained to blocks mined with the new one.

Since 2010, the Bitcoin protocol set the maximum size of a block to one megabyte. As the number of transactions on the Bitcoin blockchain grew, it became necessary to upgrade the protocol to increase throughput. Developers and miners exchanged their views on how to conduct this upgrade, and reached the so called "New York Agreement" in May 2017 over the SegWit2x project. They agreed to introduce the SegWit upgrade (a soft fork), and to implement in November 2017 a hard fork that would increase block size to two megabytes (the "2x" part of the project). However, an alternative hard fork, Bitcoin Cash, occurred on August 1st, 2017. A prominent supporter of Bitcoin Cash was Bitmain, a major manufacturer of ASICs (specialized mining equipment) and a large mining pool operator. Bitmain owned a patent on a mining-enhancing technology called AsicBoost, which use was limited with SegWit, but not with Bitcoin Cash. Thus, Bitmain derived large private benefits from the adoption of Bitcoin Cash instead of SegWit.

Another important hard fork, Bitcoin Gold, which occurred in October 2017, also featured private benefits. Bitcoin Gold's core developers justified the need for a hard fork by arguing that the network was threatened by the dominance of a small set of ASICs manufacturers who could dictate their terms to miners.[3] By preventing the use of ASICs, Bitcoin Gold would restore the Bitcoin blockchain to its original decentralized structure. Beyond these stated motives, the developers of Bitcoin Gold premined $100,000$ coins before the blockchain was open to other miners, and included a hidden fee of 0.5% of all block rewards in the code proposed to set up Bitcoin Gold mining pools. It took miners some time before they could remove this fee from the code.

To model the above discussed private benefits, we assume miner m's reward from solving a block with the upgrade is multiplied by a factor $1 + b_m$ where $b_m \in (-1, +\infty)$. Our first proposition shows that, in this context, a permanent fork can arise, in which some miners adopt the upgrade while the others do not (see Biais, Bisière, Bouvard and Casamatta, 2019 for the proof).

PROPOSITION 1: *If for* $m \in \{1, 2\}$, $b_m \geq 2$ *while for* $m \in \{3, 4, 5, 6\}$, $b_m \leq 2$, *there exists an equilibrium in which,*

- *until the* n^{th} *block is solved, all miners mine the same chain.*
- *after the* n^{th} *block is solved, miners* $m \in \{1, 2\}$ *mine one chain using the upgraded protocol, while miners* $m \in \{3, 4, 5, 6\}$ *mine a different chain using the old protocol.*

To understand the economic forces underlying Proposition 1, note first that miners' actions are strategic complements. Because block rewards are worth more when they belong to chains where more miners are active ($G(.)$ is increasing), miners have an incentive to coordinate on the same chain. Therefore, sustaining an equilibrium with a fork, as in Proposition 1, requires a countervailing force. In the short run, this role is played by private benefits associated with one version of the protocol.

Consider miner 1 who strongly favours the upgrade ($b_1 \geq 2$). To show that his equilibrium action is to adopt the upgrade, we need to show he cannot benefit from a one-shot deviation, in which he keeps mining with the old protocol until the next event. If that next event is that he solves the next block, the reward he receives for this block when leaving the game is $G(4)$ under the deviation. This is lower than the reward obtained on the equilibrium path, $G(2)(1 + b_1)$. For all other possible realizations of the next event, miner 1's action is irrelevant to his payoff.

While it is straightforward that private benefits tilt the decisions of miners 1 and 2 towards adopting the upgrade, the above analysis shows that this choice entails an opportunity cost: the upgrade confines

[3] *"Manufacturers can produce ASICs at a tiny cost, but miners have to buy at a high price. This violates the one-CPU-one-vote ethos as described in the Bitcoin white paper, because while everyone can buy CPU at the same price, the same is not true for ASIC hardware."* Robert Kuhne, Bitcoin Gold contributor in Bitcoinmagazine.com, October 11th, 2017.

miners 1 and 2 to a minority chain with low rewards. This endogenous cost makes the upgrade unattractive to miners 3 to 6, who are better off mining on the majority chain.

Thus in the equilibrium of Proposition 1, a minority of miners are active on the chain that gives them private benefits while a majority of miners are active on the chain without private benefits. The latter effect emphasises that strategic complementarities play a major role in generating persistent forks.

The examples of Bitcoin Cash and Bitcoin Gold raise the question of the long-term viability of private benefits: patented technologies can become obsolete, hidden fees can be removed. The equilibrium in Proposition 1 also relies on an endogenous mechanism that tends to perpetuate forks once they have been triggered. To see this, consider miners who played their equilibrium strategies, on their respective chains, until block B_{n+k}. Let N_m denote the number of blocks solved by miner m on his chain since the fork (i.e., after block B_n). We refer to N_m as miner m's vested interest in his chain. Now, consider a one-shot deviation where miner 1 switches to the old protocol right after B_{n+k} was solved, and until the next event. Miner 1's deviation can now affect the value of all the vested interests he has accumulated since the fork. Suppose indeed the next event is that miner 1 has to exit the game. In that case, the value of his vested interests is $G(1)(1+b_1)N_m = 0$ under the deviation, while it is $G(2)(1+b_1)N_m = \frac{1+b_1}{4}N_m$ under the equilibrium strategy. That is, as time goes by, miners have increasing incentives to keep developing the chain where they already mine, in order to defend the value of their vested interests (even if private benefits decrease). This entrenchment suggests that while private benefits can be key to trigger a fork, vested interests can play an important role in the long run, and potentially substitute for declining private benefits.

It should be noted that not all planned hard forks succeed. For instance, the SegWit2x technology that was to be adopted through a fork on the original Bitcoin chain in November 2017 never gained traction despite the predictions of many bloggers, developers and mining pool operators. This illustrates how coordination motives may derail forks, as formalized in the following proposition (whose proof is in Biais, Bisière, Bouvard and Casamatta, 2019.)

PROPOSITION 2: *Whatever the private benefits, there exist two equilibria: In one, the upgrade is adopted by all miners, in the other no miner adopts the upgrade.*

Proposition 2 highlights that strategic complementarities are strong enough in our game to override any private benefits. This feature of the model is the product of two assumptions we deem realistic in the context of permissionless blockchains. First, a block reward is worthless when only one miner is active on the chain this block belongs to. Second, the absolute size of private benefits, through their multiplicative form, depends on the value of the block rewards. This captures the idea that a miner cannot derive private benefits from an upgrade if no one else adopts this upgrade. In the equilibrium of Proposition 1, the fact that both miner 1 and miner 2 adopt the upgrade creates enough consensus to make blocks on their chain valuable: the reward for each block mined with the upgraded protocol is $G(2) = \frac{1}{4}$ which is then magnified by a factor $1 + b_m$. By contrast, the value of a block reward would be $G(1)(1 + b_m) = 0$ if only one miner adopted the upgrade. This implies that no miner has an incentive to unilaterally deviate by adopting the upgrade if no other miner adopts it. The same logic underlies the equilibrium where all miners adopt the upgrade.

Confronting Propositions 1 and 2 further clarifies the respective role of strategic complementarities and private benefits in a blockchain. While private benefits associated to one protocol are necessary to trigger the fork in Proposition 1, Proposition 2 clearly shows that they are not sufficient to break consensus. This reinforces the idea that coordination motives are critical not only to equilibria without forks, but also to those with forks. In that sense, the strategic complementarities created by miners being paid in the cryptocurrency of the chain they mine are a double-edged sword. On the one hand, strategic complementarities are necessary to generate consensus, i.e., the coordination of all agents on a single version of the ledger. On the other hand, strategic complementarities may contribute to breaking the consensus when combined with private benefits.

Finally, the welfare implications of forks can be ambiguous for miners. For instance, consider the equilibrium in Proposition 1 and suppose all miners, including 3 to 6, strictly prefer the upgrade ($b_m > 0$, $\forall m$). Then the equilibrium in Proposition 1 is Pareto-dominated by the equilibrium in Proposition 2 where all miners adopt the upgrade. First, the fork reduces the value of each block reward from $G(6)$ to $G(4)$ or $G(2)$, capturing the loss from breaking consensus. In addition, in Proposition 1, miners 3 to 6 forgo the private benefits from adopting the upgrade. The welfare comparison between Proposition 1 and the equilibrium in Proposition 2 where all miners keep mining with the old protocol is ambiguous, however. As in the previous comparison, the loss from breaking consensus puts the the equilibrium with fork at a disadvantage. On the other hand, the equilibrium with fork allows miners 1 and 2 to reap the benefits from the upgrade, which may be large enough to offset the consensus loss. Given that miners' interests are aligned towards coordinating on the upgrade, equilibrium multiplicity and coordination on Pareto-inferior equilibria are another illustration of the role of strategic complementarities in this framework.

IV. Conclusion

This model shows how proof of work, though designed to implement distributed consensus, may fail to do so. Our approach, motivated by recent events on the Bitcoin blockchain, focuses on the implementation of protocol upgrades as a trigger for forks. Hard forks reflect the fragile nature of consensus in a permissionless blockchain where participants need to agree not only on the state of the ledger but also on any change to the blockchain environment. Paradoxically, it is the success of Bitcoin that made the need for technological evolutions more acute. This in turn created opportunities for forks and cast doubts about the long-term viability of this cryptocurrency.[4] Designing upgrades that maintain consensus in a permissionless blockchain is a particularly difficult task as disagreement among participants arises from the structure of rewards induced by proof of work.[5]

REFERENCES

Biais, B., C. Bisière, M. Bouvard and C. Casamatta, 2019, "The Blockchain Folk Theorem," *Review of Financial Studies* forthcoming.

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

[4] See, e.g., *"Down to 'fork' - Bitcoin sinks through $5,000 as sentiment sours,"* in the Financial Times, November 19th, 2018.

[5] For instance, Bitcoin Cash tried to limit the risk of permanent forks by planning regular hard forks to implement upgrades. The recent fork with Bitcoin SV shows the limits of this approach.