January 30, 2020

Blockchain Enforcement

Hitoshi Matsushima

University of Tokyo

My Recent Works

Blockchain Influences Real-World Business.
Good Influence or Bad Influence?
How do we use blockchain for business?

"Mechanism Design with Blockchain Enforcement," (joint with Shunya Noda, UBC), in preparation.

"Blockchain Disables Real-World Governance," DP CARF-F-459, U-Tokyo, 2019.

"Information Design in Blockchain: A Role of Trusted Intermediaries," DP CARF-F-462, U-Tokyo, 2019.

Basic Question:
How can we enforce real-world business agreement?

1. **Players communicate with each other.**
   Each player $i \in \{1,2,3,...,n\}$ **announces a message** s_i, **where** $n \geq 3$.
   $s = (s_i)_{i \in N}$

2. **Players make an agreement on actions and payments.**
   $a = \alpha(s)$, **where** $a = (a_i)_{i \in N}$ **and** $\alpha(s) = (\alpha_i(s))_{i \in N}$.
   Each player i **is required to select** $a_i = \alpha_i(s)$.

3. **Each player** i **actually selects** $\tilde{a}_i$.
   Each player i **is regarded as a deviator if** $\tilde{a}_i \neq \alpha_i(s)$.

How can we incentivize players to select $\alpha(s)$?

Three Approaches

Legal Enforcement

Implicit Collusion

Blockchain Enforcement (New)

Legal Enforcement

0. Players write explicit contract in advance.
(S,α)

$\Downarrow$

4. Players verify $(S,\alpha,s,\tilde{a})$ to the court.

We need a trusted third party (court). Costly.

Implicit Collusion

1. **Players build trust in advance.**

⇓

2. **If a player deviated, the other players punish her by stopping future business.**

We need a long-term relationship (trust building). Restrictive.

Blockchain Enforcement

No trusted third party
No court
No trust building
Can we still enforce business agreement?

Yes, we can. Use blockchain!

What is Blockchain?

Distributed ledger managed in a decentralized, tamper-proof manner without relying on trust in any parties.

Cryptocurrency (Bitcoin, Ethereum, etc.) is a leading application managing the data about account balance (ownership) of cryptocurrencies

cf. Digital Assets managed by Blockchain, quite limited
Non-digital goods, impossible.

Programmability and Smart Contract

Blockchain allows users to write a computer protocol
termed Smart Contract,
which directly accesses account balance of cryptocurrencies and
makes automated transfers.

Users can make a commitment for contingent payments
by writing a smart contract.

Limitations of Blockchain and Smart Contract

Limited Automation: **Transformation from Real to Digital is not automated.**

Manipulation: **Dishonest Inputs**
Oracle Problem (Limited Automation + Manipulation)

Privacy Invasion: **Blockchain discloses the content to the public.**
Secrecy technology is very costly.

High Commission: **High-tech censoring and too much inputs are very costly.**
Hence, very simple smart contract is preferable.

We must design simple incentive contract:

never use detail of business content,
never use high-tech censoring,
incentivize users to tell the truth.

Even under these limitations, can we have very powerful blockchain enforcement?

Yes, we can!

Blockchain Enforcement

Smart Contract as Judgement Mechanism

Players write a smart contract in advance, termed Judgement Mechanism.
(M,T,γ)

$M = \underset{i\in N}{\times} M_i$	**message (input) space**
$T = (T_i)_{i\in N} \in R_+^n$	**deposit of cryptocurrencies**
$\gamma = (\gamma_i)_{i\in N}$	**side-payment rule**
$\gamma_i : M \rightarrow [0,T_i]$	

Each player i deposits $T_i + \varepsilon$ in cryptocurrencies in advance.
After players select real-world actions a, each player i inputs message m_i to blockchain,
Each player i receives $T_i + \varepsilon - \gamma_i(m)$, that is, burns $\gamma_i(m)$, in cryptocurrencies.

Blockchain Enforcement

Payers do not need to write explicit contract such as (S,α).

Instead players just write a simple computer protocol (M,T,γ).

Players do not need to verify either $(S,\alpha,s,\tilde{a})$ or (M,T,γ,m) to the court.

Assumption: Players know who deviated in real world.

State Space $\Omega=\{0,1\}^n$, $\omega=(\omega_i)_{i\in N}$, $\omega_i\in\{0,1\}$
0, non-deviator
1, deviator
Each player knows the true state. (She can verify what she did to the other players.)

How does judgement mechanism function?

“Digital (Automated) Courts”

Each player i announces about who deviated through input m_i.

Mechanism judges who are deviators according to Majority Rule.

Payment $\gamma_i(m)$ works as a punishment.

Design judgement mechanism to incentivize players to tell the truth.

Three designs of Judgement Mechanisms:

Design I: **Weak Implementation**

Truth telling is a Nash equilibrium, but not unique.

Design II: **Unique Implementation**

Oracle Problem

A variant of truth telling is the unique Nash equilibrium.

Design III: **Hybrid of Designs I and II**

False Charge Problem

Neither non-deviators nor honest reporters are fined a large amount.

＊Blockchain Enforcement is a New Implementation Problem

The court is unavailable.
Signals are verifiable.
Impossibility Theorem

cf. Standard Implementation Problem
The court is available.
Signals are not verifiable.
Possibility Theorem
(Virtual, PNE, Undominated NE)

Design I: Weak Implementation

Truth telling is a Nash equilibrium, but not unique.

$$M_i = \{0,1\}^{n-1},\;\; m_i = (m_i^j)_{j\neq i},\;\; m^j = (m_i^j)_{i\neq j},\;\; m_i^j \in \{0,1\}$$

Player i is fined a large amount T_i according to Majority Rule:

$$f_j(m^j) = 1 \qquad \text{if } \sum_{i\neq j} m_i^j > \frac{n-1}{2}$$

$$f_j(m^j) = 0 \qquad \text{otherwise}$$

Player i is fined a little bit if $m_i^j \neq f_j(m^j)$.

Hence, side payment rule is designed as

$$\gamma_i(m) = T_i f_i(m^i) + \frac{\varepsilon \sum_{j\neq i} \left| m_i^j - f_j(m^j) \right|}{n-1}$$

Design I: Weak Implementation

Purely Self-Interested Players $u_i(\tilde{a}) - \gamma_i(m)$

Strict Nash equilibrium m **defined as** $\gamma_i(m) > \gamma_i(\tilde{m}_i, m_{-i})$ **for all** $\tilde{m}_i \neq m_i$

Honest profile $m = m(\omega)$ $m_i^j(\omega) = \omega_j$ **for all** $i \in N$ **and** $j \neq i$.

Theorem 1: In Design I, $m(\omega)$ **is a strict Nash equilibrium.**

Difficulty in Implementation

It is generally inevitable that
the set of all Nash equilibria is independent of $((S,\alpha),s,\tilde{a})$.
(cf. Standard Implementation Theory)

Oracle Problem

If all players are purely self-interested,
unique (full) implementation is impossible.
(cf. Standard Implementation Theory)

We need a (new) theory of focal point.

This study incorporates behavioral aspects
into blockchain enforcement!

Psychological Preferences

A player belongs to one of three types:

Pure Self-Interest

Honest: **Purely Honest, or**
Compromise between Honest and Pure Self-Interest

Adversarial: **Purely Adversarial, or**
Compromise between Adversarial and Pure Self-Interest

Incomplete Information

Which type does a player belong to? Only she knows.

The other players expect that she is

honest	**with prob.** δ_H
adversarial	**with prob.** δ_A
purely self-interested	**with prob.** $1-\delta_H-\delta_A$

Both δ_H **and** δ_A **are close to zero.**

We still have serious multiplicity in Design I.

We need a new design (II)!

Design II: Unique Implementation

A variant of truth telling is the unique Nash equilibrium.

Player i announces how likely player $j \neq i$ is to be a deviator. $m_i^j \in [0,1]$

Player i is fined by T_i according to Modified Majority Rule:

$$
\begin{aligned}
&f_j(m^j) = 1 && \text{if } \left|\left\{ i \in N \setminus \{j\} \mid m_i^j > \frac{1}{2} \right\}\right| > \frac{n-1}{2} \\
&f_j(m^j) = 0 && \text{otherwise}
\end{aligned}
$$

Proper Scoring Rule is incorporated into side payment rule:

$$(1-m_i^j)^2 m_k^j + (m_i^j)^2 (1-m_k^j)$$

Proper scoring rule incentivizes a purely self-interested player to announce the expected average of the other players' opinions including honest and adversarial.

Side payment rule is designed as

$$\gamma_i(m) = T_i f_i(m^i) + \frac{\varepsilon}{(n-1)(n-2)} \sum_{j \neq i} \sum_{k \neq i,j} \left\{ (1-m_i^j)^2 m_k^j + (m_i^j)^2 (1-m_k^j) \right\}.$$

Design II: Unique Implementation

Which is greater, δ_H and δ_A, is crucial in implementation:

Theorem 2: Consider purely honest, purely adversarial, and purely self-interested.

In Design II, there exists the unique Nash equilibrium (the unique survival of iterative elimination of dominated strategies).

In equilibrium, a purely self-interested player announces

$$m_i^j = \frac{\delta_H}{\delta_H + \delta_A} \qquad \text{if player } j \neq i \text{ is not a deviator,}$$

$$m_i^j = \frac{\delta_A}{\delta_H + \delta_A} \qquad \text{if player } j \neq i \text{ is a deviator.}$$

If $\delta_H > \delta_A$, deviators are correctly detected and punished almost certainly.

If $\delta_A = 0$, deviators are correctly detected and punished certainly.

Robustness: Compromise between Psychology and Monetary Benefit

It is more realistic to assume that even behavioral players consider monetary benefits.

Honest player j selects m_j to minimize

$$\gamma_j(m)+C_H(\alpha(s),\tilde{a},M,T,\gamma,m)$$
$$=\gamma_j(m)+\lambda_H\sum_{i\neq j}\Big[\chi_{\alpha_i(s)}(\tilde{a}_i)(1-m_j^i)^2+\Big\{1-\chi_{\alpha_i(s)}(\tilde{a}_i)\Big\}(m_j^i)^2\Big].$$

Adversarial player j selects m_j to minimize

$$\gamma_j(m)+C_A(\alpha(s),\tilde{a},M,T,\gamma,m)$$
$$=\gamma_j(m)+\lambda_A\sum_{i\neq j}\Big[\chi_{\alpha_i(s)}(\tilde{a}_i)(m_j^i)^2+\Big\{1-\chi_{\alpha_i(s)}(\tilde{a}_i)\Big\}(1-m_j^i)^2\Big].$$

Theorem 2 is robust in compromise:

Theorem 3: **Consider compromise case.**
In Design II, there exists the unique Nash equilibrium (the unique survival of iterative elimination of dominated strategies).
If $\delta_H > \delta_A$, deviators are correctly detected and punished almost certainly.
If $\delta_A = 0$, deviators are correctly detected and punished certainly.

False Charge Problem

Design II has the following drawbacks:

New Issues in Implementation
Specific to Blockchain Enforcement

Consider $\delta_A > 0$.

Non-deviators are fined a large amount with a positive probability.

Honest reporters are fined a large amount with a positive probability.

Can we design judgement mechanism that overcomes false charge problem?

Yes, we can, by considering Design III!

Design III

Neither non-deviators or honest reporters are never fined a large amount.
(Only a tiny punishment permitted.)

We need to incentivize even adversarial players to tell the truth.

We need additional deposits, but that's not all to be consistent with unique implementation.

Assumption: Psychological cost has upper bound $K > 0$:
for every $(\alpha(s), \tilde{a}, M, T, \gamma, m)$,

$$0 \leq C_A(\alpha(s), \tilde{a}, M, T, \gamma, m) \leq K,$$
$$0 \leq C_H(\alpha(s), \tilde{a}, M, T, \gamma, m) \leq K$$

Design III

A player inputs multiple messages to each digital court.

Each player i inputs $(Z+1)(n-1)$ messages at once.
For each $j\neq i$, player i inputs one message about how likely player j deviated.
She inputs multiple Z messages about whether player j deviated.

$$\boldsymbol{m}_j = (\boldsymbol{m}_j^i)_{i\in N\backslash\{j\}}$$

$$\boldsymbol{m}_j^i = (\boldsymbol{m}_j^i(0), \boldsymbol{m}_j^i(1), \ldots, \boldsymbol{m}_j^i(Z))$$

$$\boldsymbol{m}_j^i(0) \in [0,1]$$

$$\boldsymbol{m}_j^i(z) \in \{0,1\} \quad \text{for all} \quad z \in \{1,2,...,Z\}.$$

Design III

Hybrid of Designs I and II

Like Design II, each player's 0-th announcement (from 0 to 1) is incentivized by proper scoring rule.

Like Design I, for every $z \in \{1, 2, ..., Z\}$,
each player's z-th announcement (0 or 1) is incentivized by
matching with the other players' (z-1)-th announcement.

Deviators are detected according to
Majority Rule based on the Z-th (last) announcements.

Theorem 4: **Make a minor symmetry assumption on psychological cost. Consider sufficiently large Z.**
We can construct a judgement mechanism (Design III) that satisfies the uniqueness of Nash equilibrium (unique survival of iterative elimination of dominated strategies).
Deviators are correctly detected and punished with certainty.
Honest reporters are never fined a large amount.

Because of upper bound K and sufficiently large Z, psychological cost is almost unchanged by small change of announcement.
Hence, with the help of just small fine, even adversarial player is willing to announce honestly for later announcement.

Concluding Remark:
Cartelization without Real-World Logistics

"Smart Contract + Blockchain" is a very convenient tool for business enforcement, because even untrusted can create and enforce transactions without relying on the law.
This convenience triggers anti-social behavior.
We should prevent drug crime by observing real-world logistics.

It is hard to crack down anti-social behavior
if it does not involve real-world logistics.

Example: **Cartel in Auction**
Anti-Competitive Commitment Device:
"The winner sends 100 dollars to the loser."

↑ **This is, I believe, the main problem in the use of**
"Smart Contract + Blockchain"