The Economics of Optimal Transaction Fee Mechanisms

David Lancashire
david.lancashire@gmail.com

July 7, 2024

Introduction

A socially optimal transaction fee mechanism must be *incentive compatible* for users and block producers and *pareto optimal* for all participants.[1] A school of papers in computer science claims this is impossible, but formal economic analysis confirms it is.[2]

Start with the equation for the *utility possibilities frontier* for a single individual i choosing between two goods.

$$\frac{u_a^i}{u_b^i} = \frac{F_a}{F_b}$$

The LHS shows the *marginal rate of substitution* between two goods. The RHS shows the relative costs of producing each. All combinations which maximize individual utility fall on this curve.[3] We sum the *marginal rates of substitution* for all individuals to find the *utility possibilities frontier* for society as a whole.

$$\sum_{i=1}^{s} \frac{u_a^i}{u_b^i} = \frac{F_a}{F_b}$$

These points are *pareto optimal* as we cannot improve outcomes for anyone without worsening outcomes for someone else. Incentive compatibility is also only possible at these points as otherwise some coalition of participants must be able to improve their utility through the simple byzantine strategy of paying more or less in transaction fees than the honest strategy recommends.

It is well-known that voluntary-trade pulls this equation into equilibrium for any two private goods in a free market.[4] That implies equilibrium for any N-dimensional basket of private goods, which allows us to set good a as our blockchain and good b as the basket of all other private goods.

Our first observation is that if the blockchain were a simple private good then *pareto optimality* is reached at whatever fees-levels users voluntary agree to pay block producers. Collusion is not socially suboptimal in such a case, although this analysis is tangential since the blockchain does not consist of a private good.

Modelling Blockchain Provision as a Public Good

It is better to model blockchains as public goods because they offer diffuse and non-excludable benefits that emerge when block producers compete to include publicly-circulating transactions in blocks:

- **higher security** - competition induces nodes to spend a greater percentage of their income on the security function than would otherwise be rational.
- **faster inclusion** - confirmation times are more sensitive to changes in the transaction fee as multiple nodes now compete to prioritize the highest-paying transactions for inclusion in blocks.
- **censorship resistance** - open competition makes it more profitable for non-censoring nodes to join the network if existing producers refuse to include fee-paying transactions.

We update our utility equation to account for both these public and private benefits:

$$u_{priv}^i$$

[1] Incentive compatibility at a point of pareto optimal production is the technical way to describe our desired properties, since at these points resources are used most efficiently not only within the mechanism to optimize the utility created from the fees paid, but also in relation to all other goods that could theoretically be purchased for the same fee. Unless overall production is *pareto optimal* then *incentive compatibility* is impossible since any global coalition can costlessly increase its collective welfare by adjusting the way resources are allocated in the production of goods and services.

[2] See the papers referenced in the bibliography by Tim Roughgarden, Elaine Shi, Hao Chung and others. The non-standard terms like "side-contract proofing" stem from the apparent belief that collusion must always be a socially suboptimal strategy such that *pareto optimality* and *incentive compatibility* can never be achieved at price levels negotiated directly between users and block producers for private inclusion. It is noteworthy that none of the authors working on this problem seem familiar with public choice theory or the standard approaches in economics used to address production optimization questions.

[3] If this equation is not in equilibrium this individual can increase their utility by shifting their allocation of resources to whichever good produces utility more efficiently. This translates directly into paying a higher or lower fee for on-chain transactions if one of the goods in question is a blockchain, which means that a byzantine strategy of paying a suboptimal fee is preferable to an honest strategy for users in this case and the mechanism cannot be incentive compatible.

[4] We assume a classical model with standard assumptions such as the existence of rational production schedules. Readers without a background in economics can find treatment of this subject in any introductory economics textbook. We recommend Paul Samuelson's "Economics: An Introductory Analysis" (1948) as Samuelson also pioneered the method of analyzing the optimality of public goods provision in the discussion that follows.

becomes

$$u^i_{pub+priv}$$

and

$$F_a$$

becomes

$$F_{pub+priv}$$

Our *utility possibilities frontier* is now:

$$\sum\nolimits_{i=1}^{s} \frac{u^i_{pub+priv}}{u^i_b} = \frac{F_{pub+priv}}{F_b}$$

As in our model for private goods, unless a solution to this curve exists the amount of resources being allocated to our blockchain is not *pareto optimal.* And that means we cannot have an *incentive compatibility* by definition as some subset of users must necessarily exist who can costlessly improve outcomes through the simple byzantine strategy of paying a different transaction fee.

In his seminal 1954 paper on The Pure Theory of Public Expenditure, Paul Samuelson noted that all solutions require individuals to contribute resources in proportion to the utility they receive. In decentralized mechanisms where utility is unobservable and participants cannot be compelled to pay for it, rational users can free-ride on the contributions of others by underreporting how much they value the public good. Samuelson hypothesized that this might make it impossible to design any decentralized mechanism to fund public goods at socially optimal levels.

Leonid Hurwicz – Samuelson's student and inventor of the concept of *incentive compatibility* – later generalized Samuelson's critique to all informationally decentralized games, pointing out that in decentralized mechanisms participants must communicate via signaling, such that "byzantine" strategies are always available in any game where users can lie about their utility levels and willingless to spend resources.

We can nonetheless prove a solution exists for public blockchains.

Escaping the Samuelson-Hurwicz Trap: The Proof

In review, our blockchain provides a combination of public and private benefits. Public benefits are diffuse and non-excludable, while private benefits accrue privately and in proportion to the fees paid.

$$u_{pub+priv}$$

Our provision of public goods depends on the public circulation of transactions in ways that induce open competition for collection of their fees. Minimal public goods emerge in mechanisms where private transactions predominate. The maximum amount requires all participants to share all fee-bearing transactions publicly.

So our cost function is not:

$$F_{pub+priv}$$

but rather:

$$F_{priv}$$

We modify our *utility possibilities frontier* to include a new variable x that reflects the probability between 0 and 1 that any transaction is a *public transaction* rather than a *private transaction.*

$$\sum\nolimits_{i=1}^{s} \frac{u^i_{(pub*x)+priv}}{u^i_b} = \frac{F_{priv}}{F_b}$$

Rational users prefer a high value for x as it increases utility. Block producers want to consume transactions shared by their peers, but minimize competition for collection of the transactions they already have in their mempool. Their dominant strategy is to consume *public transactions* and hoard *private transactions.*

It follows that no network which fails to incentivize block producers to share transactions publicly can have a *pareto optimal* transaction fee mechanism. A theoretical proof this problem makes optimality impossible to achieve in all proof-of-work and proof-of-stake networks may be found in the paper *On Bitcoin and Red Balloons.*

At least one class of consensus mechanisms does exist in which all participants prefer to share transactions publicly. In routing work mechanisms income is higher for block producers who share unconfirmed transactions with their peers.

Under these conditions, block producers also prefer to share transactions and public transactions crowd out private transactions as the value of x approaches 1 in equilibrium. The blockchain achieves *pareto optimality* at the point where transaction fees reflect the private utility of inclusion to users. Incentive compatibility is also achieved as users are dividing their resources between goods based solely on the private utility generated, which cannot be further increased by collusion with others.[5]

[5] One of the less obvious insights that comes from this proper solution is recognition of why collusion is suboptimal in

Conclusion

It it possible to design a *pareto optimal* transaction fee mechanism, but all solutions require block producers to prefer sharing rather than hoarding transactions. This finding reinforces the value of traditional economic tools like the study of the *utility possibility frontier* as a lense through which seemingly intractable problems in computer science can be simply and elegantly solved.

proof-of-work and proof-of-stake. Strictly speaking, when users collude with block producers in those networks, the users are not increasing their utility by paying a lower fee, since rational block producers will not add transactions that are not profitable to them regardless of whether those transactions are public or private transactions. What is happening is that block producers can increase their change of collecting public transactions by restricting competition for the inclusion of their own set of private transactions. This effects a lateral wealth transfer from block producers who share to those who do not, which creates zero-sum benefits between block producers which may then be extended to users in exchange for exclusive access to their own transactions.

References

[1] Moshe Babaioff, Shahar Dobzinski, Sigal Oren, and Aviv Zohar. On bitcoin and red balloons. In *SeCO Workshops*, 2011.

[2] Oguzhan Ersoy, Zhijie Ren, Zekeriya Erkin, and Reginald L. Lagendijk. Information propagation on permissionless blockchains. *CoRR*, abs/1712.07564, 2017.

[3] Leonid Hurwicz. On the concept and possibility of informational decentralization. *The American Economic Review*, 59(2):513–524, 1969.

[4] Leonid Hurwicz. But who will guard the guardians. *The American Economic Review*, 98(3):577–585, 2008.

[5] Paul Samuelson. *Economics: An Introductory Analysis.* McGraw-Hill Book Company, 1948.

[6] Paul Samuelson. The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4):387–389, 1954.

[7] Paul Samuelson. Diagrammatic exposition of a theory of public expenditure. *The Review of Economics and Statistics*, 37(4):350–356, 1955.

[8] Youjia Zhang and Pingzhong Tang. Collusion-proof and sybil-proof reward mechanisms for query incentive networks, 2023.