The Economics of Optimal Transaction Fee Mechanisms

David Lancashire
david.lancashire@gmail.com

July 9, 2024

Introduction

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

To determine the optimal amount of fees our blockchain must collect, we start with the definition 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 ratio between the costs of producing each. This equation falls into equilibrium when the user is getting the optimal amount of utility from the resources they are spending on goods.

We sum this across all individuals to find the *utility possibilities frontier* for society as a whole.

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

All points defined by this equation are *parteo optimal* – we cannot increase utility for any individual without reducing it for at least one other person. It follows that unless a mechanism delivers production at a point on this curve, *incentive compatibility* is impossible as at least a subset of users can increase their utility through the byzantine strategy of spending their money.

Voluntary-trade pulls this equation into equilibrium for any two private goods in a free market.[3] That implies equilibrium across any N-dimensional basket of private goods and between any one good and a basket of all other private goods.

We therefore set a as our blockchain and good b as the basket of all other private goods.

Modelling Blockchain Provision as a Public Good

We now adjust our equation to include the non-excludable benefits that are created when block producers compete to include *public transactions* in blocks:[4]

- **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 modify the *utility possibilities frontier* to account for these public benefits:

$$u_{priv}^i$$

becomes

$$u_{pub+priv}^i$$

and

[1] pareto optimality is a necessary condition for incentive compatibility, as otherwise a subset of users always exists who costlessly increase utility by spending more or less income on transaction fees.

[2] See the papers referenced in the bibliography by Tim Roughgarden, Elaine Shi, Hao Chung and others. It is noteworthy that none of the authors cite or show familiarity with standard economic approaches to defining and solving utility-optimization problems.

[3] This assumes a classical model with standard assumptions such as rational production schedules. Readers without a background in economics can find treatment of this subject in any introductory economics textbook, but we recommend Paul Samuelson's "Economics: An Introductory Analysis" (1948) as Samuelson is cited below for his method of analyzing the optimality of public goods provision in the discussion that follows.

[4] Public transactions pay their fee on-chain and circulate publicly such that any block producer can include them in the blockchain on equal term. Private transactions either pay their fee off-chain or are hoarded by block producers to minimize competition for collection of their feee.

$$F_a$$

becomes

$$F_{pub+priv}$$

Our *utility possibilities frontier* is now:

The value of x depends on the preferences that users and block producers have.

Rational block producers will not discriminate against public transactions, so rational users always prefer *public transactions* as it strictly increases their utility.

Block producers prefer minimizing competition for collection of transactions fees. Their dominant strategy is to free-ride on *public transactions* and hoard *private transactions.*

It follows that without incentivizing block producers to share transactions publicly we cannot achieve *pareto optimality* or design an incentive compatible mechanism. This formally proves the *transaction fee mechanism design* problem is strictly reducible to the *sybil problem.* 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.*

Given the existence of routing work mechanisms which incentivize block producers to share unconfirmed transactions with peers, we observe that *pareto optimality* is possible. That further implies incentive compatibility is achieved since utility is maximized in these mechanisms by letting users pay whatever fee-level correlates with the utility they receive privately from transaction inclusion.

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

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. In addition to providing a fascinating exception to Paul Samuelson's 70-year old conjecture about public goods, this finding reinforces the value of traditional economic tools like the *utility possibility frontier* as a lense through which seemingly intractable problems in computer science can be simply and elegantly solved.