Model of Effects of Variable Staking Rewards

In these notes we will consider a simple model of varying staking rewards and how this causes the pool of stakers to evolve over time.

1 Notation and assumptions

Suppose that we have discrete periods and in each period i some set of users who are staked receive a reward with a rate of return of $r_i \in [0,1]$. Namely, if a participant j stakes $x_{j,i}$ PNK for the entirety of the period i, at the end of the period she receives $r_i \cdot x_{j,i}$ PNK so that $x_{j,i+1} = x_{j,i} \cdot (1 + r_i)$.

We denote $S_i = \{j : \text{ staked at period } i\}$. Furthermore, we denote by $s_i \in [0,1]$ the percentage of PNK staked in that period. Namely, if N_i is the total amount of PNK that exists during that period, then

$$s_i = \frac{\sum_{j \in S_j} x_{j,i}}{N_i}.$$

Suppose that $t \in [0,1]$ is a "target" for the percentage of the community that one hopes to encourage to stake. Then suppose that r_i evolves from one round to another as follows:

$$r_{i+1} = r_i \left(1 + t - s_i\right).$$

The idea here is that when the percentage of the community that is staked s_i approaches the target, the reward rate should stabilize with $r_{i+1} \approx r_i$.

We assume that the set of potential stakers is finite and that each potential staker j has some rate $r^{(j)} \geq 0$ at which she would be willing to stake. For simplicity we assume that for each participant the value $r^{(j)}$ is constant from one period to another. We suppose participant j stakes all of her $x_{j,i}$ PNK in period i if $r_i \geq r^{(j)}$ and stakes nothing if $r_i < r^{(j)}$. Particularly, note that our formula for $x_{j,i+1}$ above is such that we are assuming that at the end of each period users will add any reward they receive to their stake for the next period if they choose to stake in that period.

In this setup, a participant j who is adequately incentivized by the opportunity to obtain fees by arbitrating disputes without such reward/dilution mechanisms is one such that $r^{(j)} = 0$ and who remains staked for any value of r_i. Note, however, that the number of dispute resolution tasks per staker will vary as the staking pool changes so a more sophisticated approach should take these effects into account when modeling for which values of r_i a participant j chooses to stake, see Section 3.

2 Initial observations

The following result rules out the existence of fixed values of r_i other than zero that give equilibria in this model.

Proposition 1. *Let $t \in (0,1)$. Then there does not exist $i \in \mathbb{N}$ such that $r_i = r_{i+1} = r_{i+2} > 0$.*

Proof. Suppose we had some $i \in \mathbb{N}$ such that $r_i = r_{i+1} = r_{i+2} > 0$. Then

$$1 + t - s_i = 1 + t - s_{i+1} = 1 \Rightarrow s_i = s_{i+1} = t.$$

However, note that $r_i = r_{i+1}$ implies that $S_i = S_{i+1}$ as the values of $r^{(j)}$ are fixed from one round to another. Thus, we can calculate:

$$s_i = \frac{\sum_{j \in S_i} x_{j,i}}{N_i} = \frac{\sum_{j \in S_i} x_{j,i}}{\sum_{j \in S_i} x_{j,i} + \sum_{S_i^c} x_{j,i}}$$

and

$$s_{i+1} = \frac{\sum_{j \in S_i} x_{j,i+1}}{N_{i+1}} = \frac{\sum_{j \in S_i} x_{j,i} \cdot (1 + r_i)}{\sum_{j \in S_i} x_{j,i} \cdot (1 + r_i) + \sum_{S_i^c} x_j}.$$

So $s_i = s_{i+1}$ implies

$$r_i \left(\sum_j x_{j,i} \right) \cdot \left(\sum_{j \in S_i} x_{j,i} \right) = r_i \left(\sum_{j \in S_i} x_{j,i} \right) \cdot \left(\sum_{j \in S_i} x_{j,i} \right).$$

We have assumed $r_i > 0$ so this factor can be canceled. Moreover, if $\sum_{j \in S_i} x_{j,i} = 0$, then $t = s_i = 0$ which violates our assumptions. Thus, we can cancel this factor as well to obtain

$$\sum_j x_{j,i} = \sum_{j \in S_i} x_j \Rightarrow x_{j,i} = 0 \forall j \notin S_i.$$

Then $\sum_{j \in S_i} x_{j,i} = N_i$ so we must have $t = 1$ contradicting our assumptions. □

While the previous result ruled out "static" equilibria, it is also interesting to ask if this system allows for situations where reward rates converge to non-zero values, creating a sort of "pratical" equilibrium where the reward rate is only changing slightly even though it is not exactly the same from round to round.

Proposition 2. *Let $t \in (0,1)$ and suppose that there exists some user j_0 such that $r^{(j_0)} = 0$ and $x_{j_0,1} > 0$. Then either the sequence r_i converges to 0 or it does not converge.*

Proof. Note that we can express r_i as

$$r_i = r_0 \prod_{k<i} r_k = r_0 \prod_{k<i} (1 + t - s_k).$$

By standard results on infinite products, if this quantity converges to a non-zero value as $i \to \infty$, then s_i must also converge as a sequence to t.

Suppose r_i converges to $R > 0$. As we have supposed that the set of users is finite, we can calculate the minimum non-zero difference between the thresholds at which any user would stake and R as: $m = \min\left\{|r^{(j)} - R| : r^{(j)} \neq R\right\} > 0$. Note that, if there were not at least one user j for which $r^{(j)} \neq R$ then all users would have the same threshold and so they would either all stake or none of them would stake for any given value of r_i. Then the sequence s_i could only take values in $\{0, 1\}$ which is incompatible with having a limit of $t \in (0, 1)$.

We consider two cases:

- For all $\epsilon_0 > 0$ there exists some $i \in \mathbb{N}$ such that $r_i \in [R, R + \epsilon_0)$.
- There exists some ϵ_0 such that there exists no $i \in \mathbb{N}$ such that $r_i \in [R, R + \epsilon_0)$.

We consider the first case where for all $\epsilon_0 > 0$, there exists some $i \in \mathbb{N}$ such that $r_i \in [R, R + \epsilon_0)$. Let $\epsilon_1 = \frac{1}{2}\min\left\{r^{(j)} - R : r^{(j)} > R\right\}$ which is positive as the set of participants is finite. Then there exists $K_1 \in \mathbb{N}$ such that if $i > K_1$, $|r_i - R| < \epsilon_1/4$. As the participant j_0 stakes in every round regardless of the value of r_i, we can say that

$$N_i \geq x_{j_0,i} = x_{j_0,K_1} \prod_{k=K_1}^{i-1} (1 + r_k) \geq x_{j_0,K_1} \cdot (1 + R - \epsilon_1/4)^{i-K_1-1}.$$

Then, as our choice of ϵ_1 guarantees that no participants j such that $r^{(j)} > R$ stake at any period after period K_1, we have

$$\sum_{j:r^{(j)}>R} \frac{x_{j,i}}{N_i} \leq \frac{\sum_{j:r^{(j)}>R} x_{j,K_1}}{x_{j_0,K_1} \cdot (1 + R - \epsilon_1/4)^{i-K_1-1}}.$$

Take $\epsilon = 2\left[1 - \left(t + \frac{1-t}{2}\right)\right]$ which is positive by our assumptions on t. As the set of participants is finite, the value $\frac{\sum_{j:r^{(j)}>R} x_{j,K_1}}{x_{j_0},K_1}$ is fixed, so by the continuity of the function $\frac{1}{(1+R-\epsilon_1/4)^{i-K_1-1}}$ in i, there exists some $K_2 \in \mathbb{N}$ such that if $i > \max\{K_1, K_2\}$, then $\sum_{j:r^{(j)}>R} \frac{x_{j,i}}{N_i} < \frac{\epsilon}{2}$. By our assumption using the values ϵ_0, there in particular exists an infinite sequence $i_1, i_2, \ldots$ such that $i_k \geq \max\{K_1, K_2\}$ and $r_{i_k} > R$. However, then $\left\{j : r^{(j)} \leq R\right\} \subseteq S_{i_k}$, so

$$s_{i_k} \geq 1 - \sum_{j:r^{(j)}>R} \frac{x_{j,i}}{N_i} \geq 1 - \frac{\epsilon}{2} = t + \frac{1-t}{2}.$$

However, as $t \in (0, 1)$ this is incompatible with the sequence s_i having a limit of t.

Now we consider the second case where there exists an $\epsilon_0 > 0$ such that there exists no $i \in \mathbb{N}$ such that $r_i \in [R, R + \epsilon_0)$. Note by properties of composites of limits the sequence $z_i = \frac{r_i \cdot s_i \cdot (1-s_i)}{1+r_i \cdot s_i}$ must converge to $\frac{R \cdot t \cdot (1-t)}{1+R \cdot t}$.

Let $\epsilon = \min\left\{\frac{R\cdot t\cdot(1-t)}{1+R\cdot t}, \frac{m}{2}, \frac{\epsilon_0}{2}\right\}$. Note that $\epsilon > 0$ by our assumptions on t, m, and ϵ_0. Let $K_1 \in \mathbb{N}$ be such that if $i > K_1$, $|r_i - R| < \epsilon/4$, let $K_2 \in \mathbb{N}$ be such that if $i > K_2$, $|s_i - t| < \epsilon/4$, and let $K_3 \in \mathbb{N}$ be such that if $i > K_3$, $\left|z_i - \frac{R\cdot t\cdot(1-t)}{1+R\cdot t}\right| < \epsilon/4$. Then take $K \in \mathbb{N}$ to be any value larger than $\max\{K_1, K_2, K_3\}$ and let $i > K$ such that $r_i \neq r_{i+1}$ which must exist by Proposition 1. Note that r_i and $r_{i+1} < R$ as $|r_i - R| < \epsilon/4 \Rightarrow r_i < R + \epsilon/4$ and we have chosen $\epsilon < \epsilon_0$.

Suppose there exists some j for which $r^{(j)}$ is in the closed interval between r_i and r_{i+1}, $r^{(j)} \in [\min\{r_i, r_{i+1}\}, \max\{r_i, r_{i+1}\}]$. Then,

$$0 \leq \max\{r_i, r_{i+1}\} - r^{(j)} < R - r^{(j)} \leq R - \min\{r_i, r_{i+1}\} < \epsilon/4 < m,$$

which contradicts the definition of m. Thus, we know there exists no such j and hence $S_i = S_{i+1}$.

Now, by the triangle inequality $|s_{i+1} - s_i| \leq |s_{i+1} - t + t - s_i| \leq |s_{i+1} - t| + |s_i - t| < \epsilon/2$ for $i > K_2$. However,

$$\begin{aligned}
|s_{i+1} - s_i| &= \left|\frac{\sum_{j\in S_i} x_{j,i}\cdot(1+r_i)}{\sum_{j\in S_i} x_{j,i}\cdot(1+r_i) + \sum_{S_i^c} x_j} - \frac{\sum_{j\in S_i} x_{j,i}}{\sum_{j\in S_i} x_{j,i} + \sum_{S_i^c} x_{j,i}}\right| \\
&= \frac{r_i\left[\left(\sum_{j\in S_i} x_{j,i}\right)\left(\sum_j x_{j,i}\right) - \left(\sum_{j\in S_i} x_{j,i}\right)^2\right]}{\left(\sum_{j\in S_i} x_{j,i} + \sum_{S_i^c} x_{j,i}\right)\cdot\left(\sum_{j\in S_i} x_{j,i}\cdot(1+r_i) + \sum_{S_i^c} x_j\right)} \\
&= \frac{r_i\left(\sum_{j\in S_i} x_{j,i}\right)\left(\sum_{j\in S^c} x_{j,i}\right)}{\left(\sum_{j\in S_i} x_{j,i} + \sum_{S_i^c} x_{j,i}\right)\cdot\left(\sum_{j\in S_i} x_{j,i}\cdot(1+r_i) + \sum_{S_i^c} x_j\right)} \\
&= \frac{r_i\cdot s_i(1-s_i)}{1+r_i\cdot s_i} = z_i > \frac{R\cdot t\cdot(1-t)}{1+R\cdot t} - \epsilon/4 > \frac{3\epsilon}{4},
\end{aligned}$$

which is contradictory as $\epsilon > 0$.

□

Remark 1. *This argument wound up being somewhat more subtle than expected. Intuitively, the lack of equilibria at intermediate rates of staking rewards is largely a result of strong assumptions on the behavior of participants to immediately stake their rewards causing the percentage of PNK staked to change, and hence the reward rate to change, even if the set of participants who are incentivized to remain the same. The complexity of the argument is to show that these different factors don't somehow simulatenously converge to give a stable equilibrium.*

Both

- *Let $t \in (0,1)$. Then either the sequence r_i converges to 0 or it doesn not converge, and*

- *Let $t \in (0,1)$ and suppose that there exists some user j_0 such that $r^{(j_0)} = 0$. Then the sequence r_i converges to 0*

seem likely to be true individually though the current proof only gives this combined statement.

In cases where r_i converges, Proposition 2 essentially shows that the long-run effect of this mechanism is to give more weight to the participants j with $r^{(j)} = 0$, namely those participants who do are already adequately incentivized to stake without a staking reward. Indeed, if $r_i \to 0$, these are the only participants who remain staked for large values of i.

Nonetheless, one can easily construct examples as the following:

Example 1. *Suppose we begin with initial conditions where $t = 1/2$ and $r_1 = .3$, but there is a participant j_0 such that $r^{(j_0)} = 0$ but $\frac{x_{j_0,1}}{N_1} = .6N_1$.*

Note that even though $s_1 \geq .6 > t$ causes the reward rates r_i to decrease, the participant j_0 remains staked and her percentage of the stake does not decrease. Thus we continue to have $s_i \geq .6$.

Nonetheless, note that participants j with $r^{(j)} > 0$ are not totally diluted. We can compute

$$N_i \leq N_1 \prod_{k=1}^{i-1}(1 + r_k) \leq N_1 \prod_{k=1}^{i-1}(1 + .3 \cdot (1 + .5 - .6)^{k-1}) \to 16.3047 \text{ as } i \to \infty.$$

Here the second inequality uses the fact that $s_i \geq \frac{x_{j_0,i}}{N_i} \geq .6$ for all i. Thus, in this example, the supply of tokens stabilizes at, at most, roughly 16 times its initial amount. Then a participant who never staked would be diluted to, at worst, roughly 6.1% of her initial percentage of the token pool.

3 Limitations of model

Note that our model is relatively simplistic, and as we noted in Remark 1, the strong assumptions on participants restaking their rewards contribute to the lack of equilibria at non-zero reward rates. However, one would not necessarily expect real human beings to immediately stake the rewards they receive each period. Moreover, we have assumed that users have a single threshold $r^{(j)}$ at which they stake or destake their PNK, when maybe it could make more sense to model users as "inertial" in that they will make the effort to stake or destake at different thresholds. Then a user that stakes because r_i exceeds her threshold $r^{(j)}$ might only destake if r_i drops below some lower threshold $r^{(j)}_{\text{destake}} < r^j$. One might try to consider a model that combines some of the above ideas with a model for PNK token price to model effects such as whether participants' $r^{(j)}$ change as the price changes.

Finally, as pointed out in Section 1, the above model essentially takes participants j such that $r^{(j)} = 0$ as corresponding to those who are adequately incentivized to stake through fees from dispute resolution. However, as the staking pool varies, individuals may receive more or fewer dispute resolution tasks. Thus, a model that takes into account which jurors are staked in a given period both in terms of the income they can receive from dispute resolution tasks and in terms of the staking rewards may be the following:

Take $r^{(j)}$ and r_i defined as above. Let D be the total amount of income that can be received from dispute resolution tasks in the period i after deducting for gas expenses, juror effort, etc. Then, one's return on staking 1 PNK is

$$r'_i = r_i + \frac{D}{N_i}$$

and one could model a participant j as staking all ofher PNK in period i if $r'_i \geq r^{(j)}$ and stakes nothing if $r'_i < r^{(j)}$.

References