Information Disclosure in Dynamic Innovation Contests

Jussi Keppo Zhuang Linsheng

Abstract

...

1 Introduction

- Kaggle[1] ...
- Meta-kaggle Risdal and Bozsolik (2022).

1.1 Literature Review

This paper focuses on the two players innovation contest with a continuous time where the players' relative position is public information throughout the game. This is closely related to tug-or-war contest, which, to our knowledge, was first formally given by Harris and Vickers (1987) as a one-dimensional simplification of the multi-stage R&D race. The output processes are model by Brownian motions drifted with effort inputs, which is followed by Budd et al. (1993) who model the state of a dynamic competition of two innovative duopoly firms by a Brownian motion drifted by the effort gap, and solve the equilibrium approximately. Furthermore, Moscarini and Smith (2007) model the tug-of-war state as the gap of the two outputs directly, and draw an analytical equilibrium of the pure strategies.

...

Information disclosure in contest - Bimpikis et al. (2019).

...

Closest paper - Ryvkin (2022).

...

2 The Model

We assume two players, i and j, compete for a prize $\theta > 0$ in a contest. Winner gets the prize $\theta > 0$ and loser gets nothing. The contest starts at time zero. At every time $t \geq 0$, the representative player i chooses an effort level $q_{i,t}$ and burdens a quadratic cost $C_i(q_{i,t}) = c_i q_{i,t}^2/2$, with a lower c_i corresponding to higher ability. The output of player i denoted by $x_{i,t}$ follows

$$dx_{i,t} = q_{i,t}dt + \sigma_i dW_{i,t} \tag{1}$$

[1] https://www.kaggle.com

Here $W_{i,t}$ is a standard Brownian motion and $\sigma_i > 0$ measures the production risk. Denoted by $y_t \equiv x_{i,t} - x_{j,t}$ the output gap of two players at time t. The dynamic of y_t is given by

$$dy_t = (q_{i,t} - q_{j,t})dt + \sigma dW_t \tag{2}$$

where $\sigma^2 = \sigma_i^2 + \sigma_j^2$ and $\sigma dW_t = \sigma_i dW_{i,t} - \sigma_j dW_{j,t}$.

We assume that the contest is equipped with a submission system that allows participants to upload their algorithms at any time and receive immediate feedback. For simplicity, we further assume that agents submit their intermediate results whenever they make progress. This setup enables the contest organizers to monitor all players' progress $x_{i,t}$ and $x_{j,t}$ in real time. This true output level, evaluated by the system, is only known by the game designer but not the two players.

At any time $t > 0$, the contest designer emits a *public* signal of the real output gap y_t. The signal is ambiguous and the game holder controls the ambiguity. The dynamic of signal is

$$dZ_t = y_t dt + \frac{dB_t}{\sqrt{\lambda}} \tag{3}$$

where B_t is standard Brownian motion independent with $(W_{i,t})$ and $(W_{j,t})$, and the parameter λ is set by the game holder to control the precision of signal. The larger the λ, the more accurate the signal would be.

The information set of both players at time $t \geq 0$ is $I_t \equiv \{Z_s : 0 \leq s \leq t\}$. Player i estimates the unknown output gap y_t based on the information set I_t. Let $\tilde{y}_t \equiv E(y_t|I_t)$ be the estimated output gap and $S_t \equiv E[(\tilde{y}_t - y_t)^2|I_t]$ be the estimation variance. According to Chapter 1.2 of Bensoussan (1992), *Kalman-Bucy filter* gives the dynamics of $\tilde{y}_t$ and S_t,

$$d\tilde{y}_t = (q_{i,t} - q_{j,t})dt + \lambda S_t(dZ_t - \tilde{y}_t dt) \tag{4}$$

$$\frac{dS_t}{dt} = \sigma^2 - \lambda S_t^2 \tag{5}$$

Hence, the conditional distribution $y_t|I_t \sim \mathcal{N}(\tilde{y}_t, S_t|I_t)$ is fully captured by the mean $\tilde{y}_t$ and variance S_t. If $\lambda = 0$, we have $S_t = S_0 + \sigma^2 t$, i.e., the estimation variance is increasing in time linearly. If $\lambda > 0$, the solution of (5) is

$$S_t = \begin{cases} \bar{S} \cdot \tanh\left\{t \cdot \sigma\sqrt{\lambda} + \tanh^{-1}\left(S_0/\bar{S}\right)\right\} & \text{if } S_0 < \bar{S} \\ \bar{S} & \text{if } S_0 = \bar{S} \\ \bar{S} \cdot \coth\left\{t \cdot \sigma\sqrt{\lambda} + \coth^{-1}\left(S_0/\bar{S}\right)\right\} & \text{if } S_0 > \bar{S} \end{cases} \tag{6}$$

Specifically, $\bar{S} = \sigma/\sqrt{\lambda}$ when $\lambda > 0$ and $\bar{S} = \infty$ when $\lambda = 0$. Please refer to Appendix A for the derivations. Figure 1 shows the evolution of S_t in time: estimation variance S_t converges to *steady state* $\bar{S}$ as time goes by regardless of the starting estimation variance. For simplicity, we henceforth assume that $S_0 = \bar{S}$, hence $S_t \equiv \bar{S}$.

Following Ryvkin (2022), let's consider a dynamic contest with a given fixed deadline. Sup-

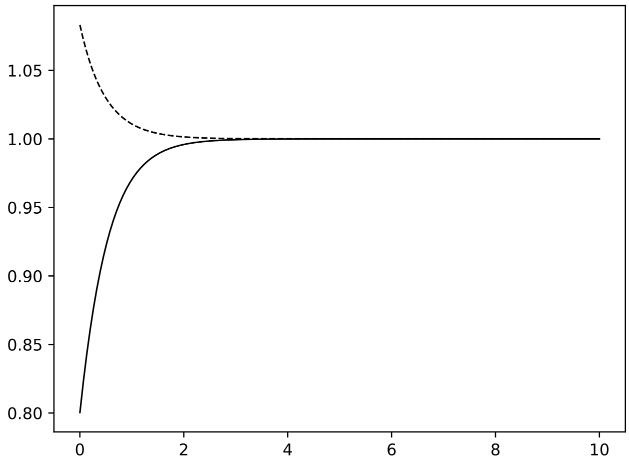

Figure 1: The evolution of S_t in time t, given that $\lambda = 1$ and $\sigma = 1$.

pose the contest is terminated when time $t = T > 0$. Since the variance $\bar{S}$ is fixed as assumed above, the state of the game is fully characterized by a tuple $(\tilde{y}_t, t)$. At any time $0 \leq t < T$, player i optimizes her effort level $q_{i,\tau}$ in the remaining contest period $\tau \in [t, T)$ according to the following optimization problem,

$$V^i(\tilde{y}_t, t; q_{j,t}, \Theta_i) = \max_{\{q_{i,\tau}\}_{\tau=t}^T} \mathbb{E}\left(\theta \cdot 1_{\tilde{y}_T > 0} - \int_t^T C_i(q_{i,\tau}) d\tau \middle| I_t\right) \tag{7}$$

where $\Theta \equiv \{\theta, \lambda, \sigma, c_i, c_j\}$, subject to constraints (4), (5) and $q_{i,\tau} \geq 0$ for all $\tau \in [t, T)$. The optimization problem for player j is just symmetric to that of player i as $V^j(\tilde{y}_t, t) = V^i(-\tilde{y}_t, t)$. The corresponding Hamilton-Jacobi-Bellman (HJB) equation for player i is

$$0 = \max_{q_{i,t} \geq 0} \left[-\frac{c_i q_i^2}{2} + V_y^i \cdot (q_{i,t} - q_{j,t}) + V_t^i + \frac{V_{yy}^i}{2} \lambda \bar{S}^2\right]$$

By definition, we have $\lambda \bar{S}^2 = \sigma^2$. Under the assumption of inner solution, we plug into the first order conditions $q_{i,t} = V_y^i / c_i$ and $q_{j,t} = -V_y^j / c_j$, we have the system of equations

$$\frac{1}{2c_i}(V_y^i)^2 + \frac{1}{c_j} V_y^i V_y^j + V_t^i + V_{yy}^i \frac{\sigma^2}{2} = 0$$
$$\frac{1}{2c_j}(V_y^j)^2 + \frac{1}{c_i} V_y^j V_y^i + V_t^j + V_{yy}^j \frac{\sigma^2}{2} = 0$$

subject to boundary conditions $V^i(-\infty, t) = 0$, $V^i(+\infty, t) = \theta$, $V^j(-\infty, t) = \theta$, $V^j(+\infty, t) = 0$, $V^i(\tilde{y}_T, T) = \theta \cdot 1_{\tilde{y}_T > 0}$ and $V^j(\tilde{y}_T, T) = \theta \cdot 1_{\tilde{y}_T < 0}$.

The Nash equilibrium is summarized in the following lemma:

Lemma 1 (Ryvkin 2022)**.** *In the Markov perfect equilibrium, the players' efforts in state* $(y, t) \in$

$\mathbb{R} \times [0, T)$ *are given by*

$$m_i(y,t) = \frac{e^{-z^2/2}}{\sqrt{2\pi\sigma^2(T-t)}} \cdot \frac{\sigma^2}{2} \left[\gamma(\rho_i) + \gamma(\rho_j)\right] \left[1 - \rho(z)^2\right] \left[1 + \rho(z)\right]$$

$$m_j(y,t) = \frac{e^{-z^2/2}}{\sqrt{2\pi\sigma^2(T-t)}} \cdot \frac{\sigma^2}{2} \left[\gamma(\rho_i) + \gamma(\rho_j)\right] \left[1 - \rho(z)^2\right] \left[1 - \rho(z)\right]$$

where $z = y/(\sigma\sqrt{T-t})$, $\rho(z) = \gamma^{-1}\left(\Phi(z)\left[\gamma(\rho_i) + \gamma(\rho_j)\right] - \gamma(\rho_j)\right)$ *and*

$$\gamma(u) = \frac{u}{1-u^2} + \frac{1}{2}\ln\frac{1+u}{1-u}, \quad u \in (-1,1)$$

$$\rho_i = \frac{e^{w_i} + e^{-w_j} - 2}{e^{w_i} - e^{-w_j}}, \quad \rho_j = \frac{e^{w_j} + e^{-w_i} - 2}{e^{w_j} - e^{-w_i}}, \quad w_{i(j)} = \frac{\theta}{\sigma^2 c_{i(j)}}.$$

Remark 1. *By definition, it is not hard to see that* $\gamma(\cdot)$ *is strictly increasing on* $(-1,1)$*, ranging from* $-\infty$ *to* ∞*. Moreover, both* ρ_i *and* ρ_j *lie in the interval* $(-1,1)$*.*

We include a simplified version of the proof in the appendix.

...

3 Model Estimation

In this section, we ....

Unknown parameters to be estimated:

- σ (Total innovation risk): prior
- c_i and c_j (Capacities)

Data: Data are submission times $\{t_k^i\}_{k=1}^{N_i}$ and $\{t_k^j\}_{k=1}^{N_i}$.

Likelihood: We assume the submission times follow inhomogeneous Poisson process, controlled by the effort functions $\tau_i(t)$ and $\tau_j(t)$. Then, during any time interval of the contest $\mathcal{S}$, the Poisson arrival rate of submissions of the representative player i is given by $\int_{s\in\mathcal{S}} \tau_i(s)ds$. The likelihood function of any realization of this point process $\{t_k^i\}_{k=1}^{N_i}$ is given by

$$p\left(\{t_k^i\}_{k=1}^{N_i}|\tau_i\right) = \exp\left\{-\int_{s\in\mathcal{D}} \tau_i(s)ds\right\}\prod_{k=1}^{N_i}\tau_i(t_k^i)$$

We assume the submission intensity $\tau_i(t)$ if proportional to the effort level $m_i(\tilde{y}_t, t)$. More specifically, we assume

$$\tau_i(t) = \bar{\tau}_i \cdot m_i(\tilde{y}_t, t)$$

4 Application

...

4.1 Synthetic Data

Before applying our estimation procedure to real-world contest data, we evaluate its potential on synthetically generated data.

Based on the following assumptions, we will next generate the submission data for two players in a data analysis competition:

- Total uncertainty: $\sigma = 5$
- Capacities: $c_i = 1.5$ and $c_j = 2$
- Contest duration: from 2025-01-01 to 2025-03-31
- Starting point $\tilde{y}_0 = 0$
-

4.2 Case Study

...

5 Conclusion

References

Bensoussan, A. (1992). *Stochastic Control of Partially Observed Systems*. Cambridge University Press, 1 edition.

Bimpikis, K., Ehsani, S., and Mostagir, M. (2019). Designing dynamic contests. *Operations Research*, 67(2):339–356.

Budd, C., Harris, C., and Vickers, J. (1993). A model of the evolution of duopoly: Does the asymmetry between firms tend to increase or decrease? *The Review of Economic Studies*, 60(3):543–573.

Harris, C. and Vickers, J. (1987). Racing with uncertainty. *The Review of Economic Studies*, 54(1):1–21.

Moscarini, G. and Smith, L. (2007). Optimal dynamic contests. *Working Paper*, pages 1–18. Science (June).

Øksendal, B. (1998). *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag Berlin Heidelberg, 5 edition.

Risdal, M. and Bozsolik, T. (2022). Meta kaggle.

Ryvkin, D. (2022). To fight or to give up? dynamic contests with a deadline. *Management Science*, 68(11):8144–8165.

A Solve S_t in Equation (5)

If S is in steady state $dS/dt = 0 \Leftrightarrow S = \bar{S} \equiv \sigma/\sqrt{\lambda}$. If S is not in steady state, i.e. $S \neq \bar{S}$, we first isolate the two variables and get

$$\frac{dS}{\sigma - \lambda S^2} = dt$$

Then, we take the integral on both sides

$$t = \int \frac{dS}{\sigma - \lambda S^2} = \frac{1}{\sigma\sqrt{\lambda}} \int \frac{dS\sqrt{\lambda}/\sigma}{1 - (S\sqrt{\lambda}/\sigma)^2} \equiv \frac{1}{\sigma\sqrt{\lambda}} \int \frac{du}{1 - u^2}$$

where $u = S\sqrt{\lambda}/\sigma = S/\bar{S}$. Hence,

$$\sigma\sqrt{\lambda} \cdot t = \begin{cases} \tanh^{-1}(u) - K_1, & \text{if } |u| < 1 \\ \coth^{-1}(u) - K_2, & \text{if } |u| > 1 \end{cases} = \begin{cases} \tanh^{-1}(S/\bar{S}) - K_1, & \text{if } S < \bar{S} \\ \coth^{-1}(S/\bar{S}) - K_2, & \text{if } S > \bar{S} \end{cases}$$

Thus, we conclude the non-steady state case that

$$S = \begin{cases} \bar{S} \cdot \tanh(\sigma\sqrt{\lambda} \cdot t + K_1), & \text{if } S < \bar{S} \\ \bar{S} \cdot \coth(\sigma\sqrt{\lambda} \cdot t + K_2), & \text{if } S > \bar{S} \end{cases}$$

Finally, we determine the constants K_1, K_2 by the initial condition S_0 and have

$$K_1 = \tanh^{-1}(S_0/\bar{S})$$
$$K_2 = \coth^{-1}(S_0/\bar{S})$$