Topics in Game Theory: Learning in Economic Examples

Juuso Välimäki

Aalto BIZ and Helsinki GSE

March 2022

1 Introduction

In this third lecture, I embed a model of learning and experimentation in some simple economic models. The first part takes a first look of the interaction between buyers and sellers when the quality of the products sold is uncertain. The key economic question relates to the intertemporal division of the costs and benefits of exploration between buyers and sellers. The presentation is based on Bergemann and Välimäki (2000)

The second main topic is a model of contracting in the principal agent setting when the success probability of the agent's effort is learned over time. We go over the simplest model without informational asymmetries between the contracting parties, but since I have the notes written for the case where the agent's decisions are unobservable to the principal, I include material on that case too. Free disposal applies to that part (last part of Bergemann and Hege (2005), Hörner and Samuelson (2013), and Halac, Kartik, and Liu (2016)).

2 Experimentation in Markets

2.1 The Model

- Two sellers provide quality differentiated products to a unit mass of identical buyers with unit demand in each period.
- The incumbent supplies a product with known quality, while the quality supplied by the entrant is initially unknown.
- The value of the established product is s per period and the new product has a value of either r_L or r_H with $r_L < s < r_H$.
- Let μ be the common prior probability that the new product has value r_H and denote the expected quality by $r(\mu) := \mu r_H + (1-\mu) r_L$.
- Marginal costs of production are identical and normalized to zero.
- Firm j chooses price p_j^t in period $t \in \{1, 2\}$, where $j = 0$ indexes the entrant, and $j = 1$ the incumbent.
- The net utility of a purchase to the buyer is the (expected) quality of the product minus its current price.
- Buyers and sellers maximize the sum of their per period payoffs.
- The revelation of uncertainty takes an extremely simple form. If a fraction x of the buyers experiment with the new product, then its true quality is revealed to all agents in the second period with probability x.
- With the complementary probability, no new information arrives.

2.2 Analysis: second period

- If full experimentation occurs in the first period, i.e., $x = 1$, then with probability μ, the new product is worth r_H in the second period.

- The second period prices are given by Bertrand competition: $p_0^2 = r_H - s$ and $p_1^2 = 0$, and all buyers purchase from the entrant.

- With probability $1-\mu$, the quality is low and the second period prices are $p_0^2 = 0$ and $p_1^2 = s - r_L$, and all buyers purchase from the incumbent.

- Conditional on full experimentation in the first period, the expected second period profits for the two firms are $\pi_0^2 = \mu(r_H - s)$ and $\pi_1^2 = (1-\mu)(s - r_L)$.

- If there is no experimentation in the first period, then second period prices are given by

  $$p_0^2 = \max\{r(\mu) - s, 0\} \text{ and } p_1^2 = \max\{s - r(\mu), 0\},$$

  and the firm with positive price sells to the entire market.

- We assume for the rest of this section that $r(\mu) < s$. This implies that from a myopic point of view, experiments are costly.

2.3 Analysis: first period

- The first period equilibrium prices, p_0^1 and p_1^1, are found by backward induction.
- Since each consumer is of measure zero, the future payoff of an individual buyer is independent of her current product choice.
- The equilibrium condition under Bertrand pricing requires then that the buyer be indifferent between the two offers:

$$r(\mu) - p_0^1 = s - p_1^1 \tag{1}$$

  and hence the price differential has to be equal to the (expected) quality difference.
- Moreover, we require that the non selling firm be indifferent between selling and not selling at equilibrium prices.
- Prices satisfying this requirement are called *cautious*.
- With the linearity of the payoffs in x, either all buyers or none buy from the new firm in equilibrium.
- The values of μ at which experimentation occurs in equilibrium are characterized by two conditions.
  1. The incumbent must prefer to concede the market in the first period and to make sales in the second period if the new good fails in the first period:

$$p_1^1 + s - r(\mu) \leq (1-\mu)(s - r_L).$$

     With cautious pricing, this holds as an equality and

$$p_1^1 = \mu(r_H - s). \tag{2}$$

2. Second, the entrant has to make nonnegative expected profits by selling today and betting on a favorable resolution of uncertainty tomorrow:
$$p_0^1 + \mu(r_H - s) \geq 0. \tag{3}$$

- The values of μ that satisfy (1)-(3) induce experimentation in the first period.

- Conditions (1)-(3) imply that:

$$p_0^1 = \mu\left(r_H - s\right) + r\left(\mu\right) - s \geq -\mu(r_H - s).$$

Hence experimentation occurs in equilibrium whenever

$$\mu \geq \mu^* = \frac{s - r_L}{\left(r_H - r_L\right) + 2\left(r_H - s\right)}.$$

- On the other hand, the socially efficient policy requires experimentation whenever current costs of experimentation are outweighed by future gains:

$$\mu(r_H - s) \geq s - r\left(\mu\right), \text{ or}$$
$$\mu \geq \hat{\alpha} = \frac{s - r_L}{\left(r_H - r_L\right) + \left(r_H - s\right)}.$$

- As $\mu^* < \hat{\mu}$, we conclude that the cautious equilibrium exhibits excessive experimentation.

2.4 Discussion

- This inefficiency can be traced to the divergence of the private cost from the social cost of experiments in equilibrium.

- The social benefit, $\mu(r_H - s)$, coincides with the entrant's private benefit.

- The social cost is given by the myopic losses, $s - r\left(\mu\right)$.

- The private cost of supporting the experiment, i.e. the negative price that the entrant has to quote, is

$$p_0^1 = r\left(\mu\right) - s + \mu\left(r_H - s\right).$$

- The additional term $\mu\left(r_H - s\right)$ is the price of the incumbent, and thus reflects his informational gain through cautious pricing.

- The failure of the buyers to take the future surplus extraction into account reduces the private cost to finance experimentation.

- In contrast to the duopoly, where the identity of the benefiting seller depends on the outcome of the experiment, a monopoly would extract the social surplus at every stage and the equilibrium would be efficient.

- More insight into the discrepancy between the efficient and the equilibrium allocation may be obtained by considering the case where all buyers act collectively and make purchases as a cooperative as in Bergemann and Välimäki (1996).

- In equilibrium, the cooperative is indifferent between the two products at current prices.

- Hence the price differential is equal to the sum of the quality differential *and* the change in the continuation payoff resulting from experimentation,

$$p_0^1 - p_1^1 = r(\mu) - s + \mu s + (1-\mu) r_L - r(\mu) = r(\mu) - s - \mu(r_H - s). \quad (4)$$

- By cautious pricing, p_1^1 equals the expected gain from experimentation for the incumbent when the entrant is selling in the first period.
- Notice that with the cooperative, the expected losses from experimentation for the buyer equal exactly the incumbent's expected gain.
- The equilibrium condition (4) then shows that the private cost of experimentation for the entrant coincides with the social cost $p_0^1 = r(\mu) - s$, and efficiency follows.

References

BERGEMANN, D., AND U. HEGE (2005): "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36, 719–752.

BERGEMANN, D., AND J. VÄLIMÄKI (1996): "Learning and Strategic Pricing," *Econometrica*, 64, 1125–49.

——— (2000): "Experimentation in Markets," *Review of Economic Studies*, 67, 213–234.

HALAC, M., N. KARTIK, AND Q. LIU (2016): "Optimal Contracts for Experimentation," *The Review of Economic Studies*, 83(3), 1040–1091.

HÖRNER, J., AND L. SAMUELSON (2013): "Incentives for experimenting agents," *The RAND Journal of Economics*, 44(4), 632–663.

3 Incentives for Experimentation

The starting point is the simplest contracting model following the work of Holmström and Tirole. The agent is risk-neutral, but limited liability constraints prevent her from buying the franchise outright.

We take a look at how dynamics on learning the profitability of a joint project affect efficient and second-best contracting. The real benefit of this modeling comes in dynamic models, where unobservable actions by the agent allow for possibly misco-ordinated beliefs between the two parties. It is a major surprise that the model remains tractable. Unfortunately, we have to leave those models for further courses, but I include some material on them here as a teaser.

3.1 Bergemann and Hege

3.1.1 Setup

- An entrepreneur, the agent, owns a project.
- The project requires funding that must be obtained from a investor, the principal.
- Cost of funds proportional to amount given.
- Short term contracting, risk neutrality, limited liability.
- The returns from the project are contractible but the agent is subject to moral hazard.
- She may invest or divert the money she gets from the principal.
- Discrete-time, infinite horizon, common discount factor, $\delta < 1$.
- Project is of unknown quality.

- Good project is completed in each period with a probability that is proportional to the amount invested. Bad project is never completed.
- Ex ante symmetric information about the project.
- q_t is the posterior probability that an uncompleted project in t is good.

Timing

- At the beginning of period t, agent proposes investment amount a_t and a share s_t of the returns for herself and $(1 - s_t)$ for the principal.
- Principal accepts or rejects the offer, $d_t \in \{0, 1\}$. If accepts, then pays ca_t upfront to the agent.
- Agent decides whether to invest or not $i_t \in \{0, a_t\}$.
- Project completion observed.

Stage Payoffs

- A completed project yields R and the agent and principal get $s_t R$ and $(1 - s_t) R$ respectively.
- Expected payoffs if (s_t, a_t) is proposed and accepted:
  - if $i_t = a_t$ agent: $a_t q_t s_t R$, principal: $a_t q_t (1 - s_t) R - ca_t$
  - if $i_t = 0$, agent gets ca_t and principal gets $-ca_t$.

Strategies

- Histories are defined as usual: h^0 is an arbitrary fixed history.
- $h^t = (h^{t-1}, a_{t-1}, s_{t-1}, d_{t-1}, i_{t-1}) \in H^t$ and $H := \cup_{t=0}^{\infty} H^t$.
- $s : H \to [0, 1]$ and we write often $s(h^t)$.

- $a : H \to [0, \overline{a}]$ with $\overline{a} < 1$, and we write $a\left(h^t\right)$.
- $d : H \times [0,1] \times [0, \overline{a}] \to \{0, 1\}$.
- $i : H \times [0,1] \times [0, \overline{a}] \to \{0, a_t\}$. Note that there is no investment decision if $d_t = 0$.
- Is there any loss in not considering mixed strategies?

Beliefs on Project Quality

- Discrete-time exponential bandit (good news case).
- Bad project never completed, good completed with probability proportional to investment.
- $q(h^t) := q_t$ is the probability that the project is good.
- $$q_{t+1} = \frac{q_t\left(1 - i_t\right)}{1 - q_t i_t},$$
  where we require that $i_t \in \{0, a_t\}, a_t < 1$.
- Hence conditional on positive investment and no completion, project quality is updated downwards, $q_{t+1} < q_t$.

textbfSolution Concept Histories determine posterior beliefs for project quality $q_t = q\left(h^t\right)$.

Definition 1. A variable $x\left(h^t\right)$ is a *state variable* for the model if after any two histories $h^t, \widehat{h}^t$ such that $x\left(h^t\right) = x\left(\widehat{h}^t\right)$ the continuation games following h^t and $\widehat{h}^t$ are identical.

In this game, $q\left(h^t\right)$ is clearly a state variable.

Definition 2. A strategy $\sigma\left(h^t\right)$ is *Markovian* if $\sigma\left(h^t\right) = \sigma\left(x\left(h^t\right)\right)$.

We write $s\left(q_t\right), a\left(q_t\right), d\left(q_t, s_t, a_t\right), i\left(q_t, s_t, a_t\right)$ for Markovian strategies in this game.

Definition 3. A *Markov Perfect Equilibrium* (MPE) is a subgame perfect equilibrium where all players use Markovian strategies.

Exercise 1. Show by an example that this is different from defining MPE as any equilibrium in the game where the players have only Markovian strategies. Hint: use a repeated extensive form game where the game ends with a probability determined by stage actions (really a stochastic game).

3.1.2 Observable Investments

Benchmark: Certain Project

- In this case we have $q_t = 1$.
- If we insist on Markovian equilibria, $s_t = s, a_t = a$ for all t.
- Principal breaks (at least) even if agent invests and

$$(1-s)\,aR - ca \geq 0 \text{ or } s \geq \frac{R-c}{R}.$$

- Agent invests if

$$V(a) = asR + (1-a)\,\delta V\left(a\right) \geq ca + \delta V\left(a\right).$$

- The equality is the promise keeping constraint and the inequality is the incentive compatibility condition for investing.
- $V\left(a\right)$ can be solved from the equality with break-even contract for principal (i.e. $s = \frac{R-c}{R}$):

$$V\left(a\right) = \frac{a\left(R-c\right)}{1-\delta\left(1-a\right)}.$$

- Hence IC can be written as:
$$R - 2c \geq \frac{\delta a (R - c)}{1 - \delta (1 - a)}$$
or
$$a \leq \frac{1 - \delta}{\delta c} (R - 2c)$$
or
$$R \geq 2c + \frac{\delta a c}{1 - \delta}. \tag{5}$$

- Hence it is optimal to set $a = \overline{a}$ if $R \geq 2c + \frac{\delta \overline{a} c}{1 - \delta}$ and $a = a^* = \frac{1-\delta}{\delta c} (R - 2c)$ otherwise.

- If inequality (5) holds at $a = \overline{a}$, we say that we are in the high return case.

Uncertain Project Quality

- Assume now that $q_0 < 1$.

- After any period with investments $i_t = a_t > 0$ $q_{t+1} < q_t$.

- If $i_t = 0$, then $q_{t+1} = q_t$.

- Under Markovian strategies, $a_{t+1} = a_t$ and $s_{t+1} = s_t$ if $i_t = 0$.

- Hence we can write incentive compatibility for the agent (with break-even for the principal):
$$V (q_t) = a_t (q_t R - c) + (1 - a_t q_t) \delta V (q_{t+1}) \geq c a_t + \delta V (q_t).$$

- Suppose that incentive compatibility with $a_{t+1} = \overline{a}$ binds at q_{t+1}. Then
$$V (q_{t+1}) = \frac{c \overline{a}}{1 - \delta}.$$

- A necessary and sufficient condition for incentive compatibility in t is then:
$$\overline{a} (q_t R - c) + (1 - \overline{a} q_t) \delta \frac{c \overline{a}}{1 - \delta} \geq c \overline{a} + \delta \frac{c \overline{a}}{1 - \delta}.$$

- But this is just the condition for the high return case at return $q_t R$. and scale $q_t \overline{a}$ and hence high return projects satisfy the condition for high enough q_t.

- The critical q above which the project gets full funding is

$$q^* = \frac{2c\,(1-\delta)}{R\,(1-\delta) - \delta\overline{a}c}.$$

- If the project does not get full funding, then

$$a_t\,(q_t R - c) + (1 - a_t q_t)\,\delta\frac{ca_{t+1}}{1-\delta} = ca_t + \delta\frac{ca_t}{1-\delta}.$$

- From this one sees that since q_t is decreasing, the maximal level of funding consistent with incentive compatibility is decreasing in q_t.

Exercise 2. Will the project be abandoned in finite time? We know that q_t drifts downwards but so does a_t. It is easy to see that a_t converges to 0, but will the boundary $q = \frac{2c}{R}$ be reached?

Theorem 1. The game has a unique MPE where high return projects are funded at full speed in the beginning. Low return projects receive limited and restricted funding over time.

Definition 4. A subgame perfect equilibrium of a repeated game is called weakly renegotiation proof if none of its continuation equilibrium payoff vectors is Pareto dominated by any other continuation equilibrium payoff vector.

Exercise 3. Prove or disprove that all weakly renegotiation proof equilibria induce the same path as the unique MPE.

Exercise 4. Can you find subgame perfect equilibria that leave the principal a positive payoff?

3.1.3 Unobservable Investment

- Suppose next that the investment decision is unobservable to the principal.
- Question: How to define Markovian strategies? Not straightforward since the players observe different histories.
- The belief of the agent about project quality can always be required to depend on her own investment decisions alone.
- The principal has really a belief over the agent's beliefs..
- Public history now is h^0 and $h_P^t = (h_P^{t-1}, a_{t-1}, s_{t-1}, d_{t-1})$. Notice that i_{t-1} is missing from the public history.
- After a deviation, the agent has a belief that differs from the belief (about her belief) based on public history.
- Markovian strategies of the principal are unaffected by deviations in investments.
- Let q^A denote the agent's belief computed based on the full experimentation history.
- Let q^P be the belief over agent's belief given public history.
- On equilibrium path $q^P\left(h_P^t\right) = q^A\left(h^t\right) = q\left(h_P^t\right) =: q_t$.

Definition 5. (a^*, s^*, d^*, i^*) is a Markov Perfect equilibrium if the players use Markovian strategies:

$$\begin{aligned}
a^*\left(h^t\right) &= a^*\left(q^A\left(h^t\right), q^P\left(h_P^t\right)\right), \\
s^*\left(h^t\right) &= s^*\left(q^A\left(h^t\right), q^P\left(h_P^t\right)\right), \\
i^*\left(h^t, a_t, s_t, d_t\right) &= i^*\left(q^A\left(h^t\right), q^P\left(h_P^t\right), a_t, s_t, d_t\right), \\
d^*\left(h_P^t\right) &= d^*\left(q^P\left(h_P^t\right), a_t, s_t, \right),
\end{aligned}$$

and $q^A(h^t)$ is computed using Bayes' rule and i^*.

- Note that on equilibrium path, both players choose actions based on the common public belief on the quality of the project q_t.
- Notice also that we must allow the agent to condition her decisions on both beliefs.
- Issue: How to deal with beliefs following observable deviations by the agent? Can these be used to generate additional sequential equilibria? In Horner and Samuelson, this issue does not arise since the principal makes the offers.

Analysis

- Break-even for the investor:

$$q_t^P(1-s_t)\,a_t R \geq c a_t.$$

- Assuming this binds, we have incentive compatibility for agent:

$$a_t(q_t^A R - c) + \left(1 - a_t q_t^A\right)\delta V\left(q_{t+1}^P, q_{t+1}^A\right) \geq c a_t + \delta V\left(q_{t+1}^P, \widehat{q}_{t+1}^A\right).$$

- The value functions are calculated for the pair of public and private beliefs. On path, these coincide and we will write q_t and $V\left(q_{t+1}^P, q_{t+1}^A\right) = V\left(q_{t+1}\right)$ from now on for equilibrium path beliefs. Off path beliefs after not investing are given by $widehat q_{t+1}^A = q_t$.
- We look for a Markov equilibrium where the agent conditions her observable actions on the public information only.
- By this assumption, all future equilibrium offers are unaffected by the current decision to not invest.

- The payoff to the agent from any belief q_t from the equilibrium path offers is:

$$q_t[a_t\left(1-s_t\right)R+(1-a_t)a_{t+1}\delta\left(1-s_{t+1}\right)R+(1-a_t)\left(1-a_{t+1}\right)a_{t+2}\delta^2\left(1-s_{t+2}\right)R+...]$$

- Therefore the payoff is linear in q_t for a fixed sequence of offers and we have:
$$V\left(q_{t+1}^P,\widehat{q}_{t+1}^A\right)=\frac{q_t}{q_{t+1}}V\left(q_{t+1}\right).$$

- Hence the incentive compatibility condition is much simplified:
$$q_tR-2c\geq\frac{q_t}{q_{t+1}}\delta V\left(q_{t+1}\right),$$
where we have used Bayes' rule:
$$\left(1-a_tq_t\right)=\left(1-a_t\right)\frac{q_t}{q_{t+1}}.$$

- Again when funding is restricted, incentive constraint binds and we have a nice difference equation.

- Together with the Bellman equation
$$V\left(q_t\right)=a_t(q_tR-c)+\left(1-a_tq_t\right)\delta V\left(q_{t+1}\right),$$
we can write a differential equation in (a_t,q_t)
$$a_t=\frac{1}{1-q_{t+1}}\left(\frac{1}{2}a_{t+1}\delta-(1-\delta)\left(\frac{q_{t+1}R}{2c}-1\right)\right).$$

- This can be solved for the unique $\overline{q}$ consistent with full funding at $a_t=a_{t+1}=\overline{a}$.

- It can be shown that $\frac{2c}{R}\leq\overline{q}\leq 1$ if either
$$\begin{aligned} i)\ R &\geq 2c+\frac{\delta ac}{1-\delta}\text{ and }\delta\leq\frac{2-2\overline{a}}{2-\overline{a}}\text{ or}\\ ii)\ R &\leq 2c+\frac{\delta ac}{1-\delta}\text{ and }\delta\geq\frac{2-2\overline{a}}{2-\overline{a}}.\end{aligned}$$

- In the first case, a_t is decreasing over time, in the second case, it is increasing over time.

3.2 Hörner and Samuelson

- Still short term contracting: consider sequential equilibria of the game.
- Unobservable investment case.
- Make the principal the proposer.
- Consider continuous time limit. Notice that the game is defined in discrete time, but its properties for small Δ are analyzed using differential equations.
- Must show (and H&S do show) that equilibrium behavior in the discrete models converges to the behavior given by the differential equations.
- Size of the project fixed.
- Do a really careful analysis with the fine details related to Markovian equilibria.
- If you are interested in the area of dynamic incentives, this is very educational reading.

3.2.1 Setup

- Principal proposes, agent accepts or rejects and decided whether to invest if accept. Otherwise timing as before.
- Time interval between periods $\Delta \to 0$, $\delta = (1 - r\Delta)$.
- Probability of success from a good project is $\lambda\Delta$.
- Public belief conditional on experimentation evolves according to

$$\dot{q}_t = -\lambda q_t (1 - q_t).$$

- Principals payoff $v\left(q\right)$ on equilibrium path must satisfy:

$$v\left(q_t\right) = \left(\lambda q_t R\left(1 - s\left(q_t\right)\right) - c\right)\Delta + \left(1 - r\Delta\right)\left(1 - \lambda q_t \Delta\right) v\left(q_{t+\Delta}\right).$$

- Taking limits:

$$\left(r + \lambda q\right) v\left(q\right) = \lambda q R\left(1 - s\left(q\right)\right) - c - \lambda q\left(1 - q\right) v'\left(q\right).$$

- Agent's payoff:

$$\begin{aligned} w\left(q_t\right) &= \lambda q_t R s\left(q_t\right)\Delta + \left(1 - r\Delta\right)\left(1 - \lambda q_t \Delta\right) w\left(q_{t+\Delta}\right) \\ &= c\Delta + \left(1 - r\Delta\right)\left(w\left(q_{t+\Delta}\right) + \left(\frac{q_t}{q_{t+\Delta}} - 1\right) w\left(q_{t+\Delta}\right)\right) \end{aligned}$$

- The first equality is the promise keeping constraint and the second is the binding incentive compatibility constraint.

- By shirking, the agent can consume $c\Delta$ and remain more optimistic and hence $w\left(q_{t+\Delta}\right)$ is adjusted by $\frac{q_t}{q_{t+\Delta}}$ as already discussed in Bergemann and Hege.

- Taking limits gives:

$$\begin{aligned} 0 &= \lambda q R s\left(q\right) - \lambda q\left(1 - q\right) w'\left(q\right) - \left(r + \lambda q\right) w(q) \\ &= c - \lambda q\left(1 - q\right) w'\left(q\right) - \left(r + \lambda q\right) w\left(q\right) + \lambda w\left(q\right) \end{aligned}$$

since by Bayes' rule:

$$\lim_{\Delta \to 0}\left(\frac{q_t}{q_{t+\Delta}} - 1\right) w\left(q_{t+\Delta}\right) = \lambda\left(1 - q\right) w\left(q\right).$$

3.2.2 Results

- Analysis can be done using these three differential equations for $v\left(q\right), w\left(q\right)$ and $s\left(q\right)$. Experimentation must stop at

$$\underline{q} = \frac{2c}{\lambda R}$$

in complete analogy with Bergemann and Hege.

- $s(q)$ can be easily eliminated by summing the principal's and the agent's Bellman equation.
- Boundary conditions for $v(q)$ and $w(q)$:

$$v(\underline{q}) = w(\underline{q}) = 0.$$

- Solving these equations gives a limiting solution for some parameter values. The constraint is that $v(q)$ and $w(q)$ must be nonnegative for all $q \in [\underline{q}, 1]$.
- This does not always happen. The cases where this is violated correspond to the cases where dynamic agency cost is high.
- The remedy to this: Principal delays, i.e. makes offers that are not accepted on equilibrium path.
- Horner and Samuelson continue from here to analyze the entire set of Markovian and non-Markovian strategies.
- A rich set of dynamic possibilities emerges.

3.3 Halac, Kartik and Liu

- Back to the full contracting setting: long-term contracts.
- Add an element of adverse selection: agent knows her own success probability.
- Discrete-time and potentially infinite horizon.
- Risk neutrality, no limited liability.
- Hence the physical environment is similar, but the economic modeling is very different.
- This is really screening over time where moral hazard by the agent shapes the dynamic screening contracts.

3.3.1 Model

- In each stage, agent chooses to work or shirk $a_t \in \{0, 1\}$ Working is at cost $c > 0$ per period.
- The project is good with prior probability q_0.
- The agent has a type $\theta \in \{L, H\}$ determining her success probability λ^θ with $\lambda^H > \lambda^L$.
- Agent knows her own type but the principal assigns prior π_0 to the type being good.
- Good project succeeds with probability $a_t \lambda^\theta$ if the agent is of type θ.
- Notice that as before, the project must be good and the agent must work for the project to succeed with positive probability.
- The reward to the principal from a completion is R.

3.3.2 Contracting

- At the beginning of the game, the principal offers a contract $C = (T, W_0, b, l)$, where T is the effective length of the contract, W_0 is the initial payment to/from agent, $b = (b_1, ..., b_T)$ is a sequence of bonuses contingent on success and $l = (l_1, ..., l_T)$ is a sequence of penalties imposed on the agent for not completing the project.
- Notice that there is a lot of redundancy in the specification of the contracts: Initial payment together with a sequence of bonuses (bonus contract) or initial payment with only penalties imposed (clawback contract) would be sufficient. This is an easy exercise and Proposition 1 in the paper.

3.3.3 Payoffs

- Conditional on a contract C, agent type θ, and a sequence of actions $a = (a_t)_{t=0}^{T}$,the payoff to the principal can be written as

$$\begin{aligned}\Pi_0^\theta (C,a) &= -W_0 - (1-q_0)\sum_{t=1}^{T}\delta^t l_t \\ &\quad + q_0 \sum_{t=1}^{T}\delta^t \left[\prod_{s<t}\left(1-a_s\lambda^\theta\right)\right]\left[a_t\lambda^\theta (R-b_t) - \left(1-a_t\lambda^\theta\right) l_t\right].\end{aligned}$$

- Similarly, the payoff to agent of type θ from contract C and action sequence a is

$$\begin{aligned}U_0^\theta (C,a) &= W_0 + (1-q_0)\sum_{t=1}^{T}\delta^t (l_t - ca_t) \\ &\quad + q_0 \sum_{t=1}^{T}\delta^t \left[\prod_{s<t}\left(1-a_s\lambda^\theta\right)\right]\left[a_t\left(\lambda^\theta b_t - c\right) + \left(1-a_t\lambda^\theta\right) l_t\right].\end{aligned}$$

- We write say that IR^θ is satisfied at C^θ for sequence a if

$$U_0^\theta\left(C^\theta, a\right) \geq 0.$$

- We say that $IC^{\theta,\theta'}$ holds if

$$\max_a U_0^\theta\left(C^\theta, a\right) \geq \max_a U_0^\theta\left(C^{\theta'}, a\right).$$

- We say that IC_a^θ holds if

$$a \in \arg\max_{a'} U_0^\theta\left(C^\theta, a'\right)$$

3.3.4 Benchmarks

First-Best

- If the agent operates the project herself, then she continues as long as
$$q_t^\theta R\lambda^\theta \geq c.$$
Using Bayes' rule for the belief after t failures (and assuming agent works in all prior period),
$$q_{t+1}^\theta = \frac{q_0\left(1-\lambda^\theta\right)^t}{q_0\left(1-\lambda^\theta\right)^t + 1 - q_0}.$$

- To get the optimal number of experiments for each type t^θ, take the largest integer below the solution to
$$\frac{q_0\left(1-\lambda^\theta\right)^{t^\theta}}{q_0\left(1-\lambda^\theta\right)^{t^\theta} + 1 - q_0} = \frac{c}{R\lambda^\theta}$$
or
$$t^\theta = \frac{\ln\left(\frac{1-q_0}{q_0}\right) - \ln\left(\frac{R\lambda^\theta - c}{c}\right)}{\ln\left(1-\lambda^\theta\right)}.$$

- The key point to notice here is that t^θ is not monotonic in λ^θ.

- This is natural. Higher λ implies more accurate learning hence fewer trials necessary to become pessimistic.

- On the other hand, lower λ implies a smaller probability of getting successes hence stopping at higher posteriors.

- These two effects cannot be signed in general, i.e. it could be that $t^H > t^L$ or $t^L > t^H$.

- Since the principal can use only time to vary the contractual terms, this makes it relatively hard to find appropriate single crossing conditions for optimal screening contracts.

No Adverse Selection

- In that case, we have pure moral hazard.
- With risk neutral agents and without limited liability, sell the project to the agent at expected value
- Hence first-best is achievable.
- With limited liability, we would be in the long term contracting version of Horner and Samuelson.

No Moral Hazard

- Suppose next that the type of the agent is unknown but the actions are observable.
- Then the principal sees in each period a signal of the agent's type that is correlated with true type.
- This signal (success) has different distribution for the two types.
- Hence can use a Cremer-McLean -type contract to separate in period $t = 0$ and extract all rent.
- Conclusion: to get an interesting problem, we need both adverse selection and moral hazard (or different assumptions on limited liability).

3.3.5 Second-Best

Case where $t^L < t^H$.

- By Proposition 1, it is without loss of generality to use clawback contracts $C^\theta = \left(T^\theta, W_0^\theta, l^\theta\right) := \left(T^\theta, W_0^\theta, b^\theta, l^\theta\right)$ with $b_t^\theta = 0$ for all t.
- It is clear that the principal does not have to leave rent to low-type agent.

- Hence the low type agent gets an initial payment W_0^L that covers exactly the sum of expected penalties and costs of effort during the implemented sequence of actions.

- High type agent gets an information rent since the probability of avoiding future penalties from any period t onwards is always larger for the high type conditional on choosing the same sequence of actions at the low type (this is clearly just a lower bound for the information rent).

- The main theorem in the paper gives the following characterization to the optimal contract in this case.

Theorem 2. Assume $t^H > t^L$. Then there is an optimal clawback contract where $C^H = \left(t^H, W_0^H, l_{t^H}^H\right)$ with $l_{t^H}^H < 0 < W_0^H$, $l_t^H = 0$ for $t < t^H$, and $C^L = \left(\bar{t}^L, W_0^L, l^L\right)$ with $\bar{t}^L \leq t^L$ such that :

1. For all $t \leq \bar{t}^L$,
$$l_t^L = \begin{cases} -\left(1-\delta\right)\frac{c}{q_t^L \lambda^L} \text{ if } t < \bar{t}^L, \\ -\frac{c}{q_t^L \lambda^L} \text{if } t = \bar{t}^L. \end{cases}$$

2. IR^L binds.

3. Both agents work in all periods if they choose their own contract and H works if she chooses the contract of L.

- Notice that stopping is efficient for H (no distortion on the top).

- L stops (weakly) too early (to reduce information rent going to H).

- H pays no penalties unless she fails to complete the process by the efficient stopping time. This backloads the incentive effects of the contract (i.e. gives the best incentives to work at stages where H is the most pessimistic about the project).

- The following comparative statics holds for the model:

Proposition 1. Assume $t^H > t^L$. The second-best stopping time for L is weakly increasing in q_0 and weakly decreasing in c and π_0.

Sketch of the Proof

- For easier arguments, let Γ^θ be the set of periods where agent of type θ makes a choice. In periods $t \notin \Gamma^\theta$, the agent is locked out (by a large negative bonus, say). Consider w.l.o.g. only clawback contracts .

- The proof consists of a number of steps:

1. At all $t \in \Gamma^L$, $a_t^L = 1$ is without loss of generality. Proof: assume $a_t^L = 0$ and make the (discounted) penalty in t payable at the last $s \in \Gamma^L$ such that $s < t$. Since $a_t^L = 0$, the utility to L and the profit if $\theta = L$ is unchanged. The utility to H in the new contract is no larger (why?) under the new contract. Hence the new contract satisfies the constraints.

2. Write a relaxed program (RP 1) to maximize profit over contracts where $a_t^L = 1$ and the incentive compatibility requirement is just that
$$U_0^H\left(C^H, a^H\right) \geq U_0^H\left(C^L, a^L\right).$$
In other words, incentive compatibility for H is required to hold only relative to the alternative of choosing C^L and working at all $t \in \Gamma^L$ At the optimal solution to this problem, IR^L and $IC^{H,L}$ must bind. Using these binding constraints, the program becomes one of maximizing social surplus net of information rent for H subject to IC_a^H, IC_a^L and where $a_t^L = 1$ for all $t \in \Gamma^L$. Call this program (RP 2). Note that now the contract for each type can be solved separately.

3. $\Gamma^L = \{1, ..., T^L\}$ for some T^L. If not, then either move earlier the serious periods (to fill any gap between serious periods) or terminate the contract. One of these will increase profit.

4. For each T^L, construct the unique sequence $l_t\left(T^L\right)$ such that IC_a^L binds at all $t \in \{1, ..., T^L\}$. Hint: backward induction.

5. Argue that at any solution (RP 2) must use the penalties from the previous step. Otherwise one can reduce information rent for H. Here the restriction to clawback contracts is of help.

6. Hence the solution to (RP 2) is found by maximizing over T^L. Monotone comparative statics -type argument to show that $T^L \leq t^L$.

7. Using the fact that $T^L \leq t^L$ it can be shown that the optimal choice for H given the contract constructed in 3-6 is such that it is optimal for H to choose $a_t^{H,L} = 1$ (period t action when having chosen contract C^L) for all $t \leq T^L$.Furthermore it is optimal to maximize social surplus on C^H , i.e. $T^H = t^H$. Setting $l_t^H = 0$ for all $t < T^H$ and fixing $l_{T^H}^H$ to give agent H her information rent in expectation makes sure that agent L does not choose C^H.

Specific Comments

- Could be done using bonus contracts by Proposition 1.
- No problem without public observability of successes.

Case where $t^L < t^H$.

- Much harder because steps 6 and 7 fail in this case.
- Paper does not say much about the optimal solution.
- Undiscounted case is analyzed though.

3.3.6 General Comments

- The focus on λ^θ as the source of incomplete information makes the model hard to analyze (the two cases must be handled separately).

- If we had the cost of effort or discounting as the uncertain variables, the analysis would be much easier (to the point of being obvious).
- Maybe one could use these other sources of private information to start the analysis with less restrictive contracting assumptions.
- Limited liability?
- Possibly considering dynamic contracting (one period contracts) as in B&H and H&S.

3.4 Literature for Learning and Contracting

Papers related to repeated or long-term contracting under repeated moral hazard

"Getting it Done: Dynamic Incentives to Complete a Project," Robin Mason and Juuso Välimäki, Journal of the European Economic Association, 2015, 13 (1), 62-97

"Breakthroughs, Deadlines, and Self-Reported Progress: Contracting for Multistage Projects," Brett Green and Curtis R. Taylor, American Economic Review 2016, 106(12): 3660?3699.

"Dynamic Delegation of Experimentation," Yingni Guo, American Economic Review, 2016, 106(8):1969-2008.

"Experimentation in Organizations," Sofia Moroni, unpublished, November, 2017.

"Collaborating," Alessandro Bonatti and Johannes Hörner, American Economic Review, 2011.

"Career Concerns with Exponential Learning," Alessandro Bonatti and Johannes Hörner, Theoretical Economics, 2017.

Games with experimentation

A great survey:

”Learning, Experimentation and Information Design,” Johannes Hörner and Andrzej Skrzypacz, Econometric Society World Congress.