Notes on Dynamic Screening and Reputation Games

February 22, 2017

The purpose of these notes is to give a big-picture overview of lecture 21, on dynamic screening and commitment, and lecture 22, on incomplete information and reputation. It's easy to get lost in the details of the lectures, so these notes try to refocus on the most important logic and the structure of the problems.

Dynamic Screening

In the dynamic screening game, a seller (with value 0) has two periods to sell an item to a single buyer. The buyer's value is either v_H, with probability λ, or v_L, with probability $1-\lambda$. The seller sets a single price in each period. If the item is purchased in the second period, then payoffs are discounted by δ. There are three strategies:

- Caving: $p_1 = v_L$. Both types of buyers purchase in period 1.
- Bargaining: $p_1 = (1-\delta)v_H + \delta v_L$, $p_2 = v_L$. Under these prices, high-type buyers receive the same payoff in both periods. Low-type buyers are willing to buy in the second.
- Stonewalling: $p_1 = p_2 = v_H$.

Intuitively, caving is preferred to bargaining and stonewalling for low values of λ. If the probability that the buyer is a high type is small, then it makes sense to get revenue from low types as soon as possible.

Define the critical belief $\lambda^* = \frac{v_L}{v_H}$. λ^* is the probability of a high type that makes the seller indifferent between caving and bargaining, following from the payoffs

$$\underbrace{v_L}_{\text{Caving revenue}} = \underbrace{\lambda^* \left((1-\delta)v_H + \delta v_L\right) + (1-\lambda^*)\delta v_L}_{\text{Bargaining revenue}}.$$

Then the seller's strategy is

$$\begin{cases} \lambda < \lambda^* & \text{Cave} \\ \lambda = \lambda^* & \text{Indifferent between caving and bargaining} \\ \lambda > \lambda^* & \text{Bargain or stonewall} \end{cases}$$

The lecture notes do not directly compare stonewalling with caving, so note that if $\lambda < \lambda^*$, caving is strictly better than stonewalling. Even if the high type always purchased at $p_1 = v_H$ under stonewalling, revenue would still be $v_H\lambda < v_H\lambda^* = v_L$.

Also note that the seller will not stonewall at λ^*. When the seller stonewalls, the high type mixes between buying in the first and second periods (as explained in the next section). The crucial point is that for $\lambda = \lambda^*$, the high type always waits to buy until the second period (by the definition of α^* in the notes), so the seller only receives $\delta v_H \lambda^* = \delta v_L < v_L$. The seller would rather cave or bargain.

Stonewalling

Assume that $\lambda > \lambda^*$. The logic for caving and bargaining are familiar, so I focus on stonewalling. The key behavior in stonewalling is that the high-type buyer, v_H, must mix.

To see why, suppose that v_H does not mix and that p_1 is strictly between the bargaining price and v_H, $p_1 \in (p_1^*, v_H)$. (For now, ignore that $p_1 = v_H$ under stonewalling.) Neither of v_H's pure strategies are a PBE:

- If v_H always buys in period 1, then the seller believes all buyers in period 2 are low types, $\mu(v_L \mid \text{Reject at } 1) = 1$, and sets $p_2 = v_L$. This is not a PBE because v_H has a profitable deviation: since $p_1 > p_1^*$, v_H would rather deviate to reject in period 1, be recognized as v_L in period 2, and get $\delta v_L > v_H - p_1$.

- If v_H never buys in period 1, then the seller believes $\mu(v_H|\text{Period } 2) = \lambda > \lambda^*$, making $p_2 = v_H$ optimal. This is not a PBE because v_H could profitably deviate to accept in period 1 and get surplus $v_H - p_1 > 0$.

Since neither pure strategy is a PBE, v_H must mix. There is one last point to this logic: if v_H is willing to mix, then his payoffs in both periods must be the same. For $v_L < p_1 < v_H$, this requires that the seller mix between v_L and v_H in period 2. If the seller does not mix, then v_H could profitably deviate to a pure strategy.

In the end, we only consider the stonewalling equilibrium with $p_1 = p_2 = v_H$, so the seller does not mix. When $p_1 = v_H$, our explanation for never buying does not apply, but Curt's notes suggest that it still is not a PBE because of an "open set problem." I am not sure what he means by this, but the result is: v_H must mix in the stonewalling equilibrium.

The seller decides whether to stonewall or bargain based on which gives the higher payout, which reduces to a condition on δ.

Reputation Games

We've seen a handful of reputation games: an incumbent firm against entrants, rational and possibly irrational players in centipede, repeated ultimatum games, etc. I'll use the incumbent/entrant game as an example.

Consider an incumbent, I, who is believed to be *tough* (rational, cooperative, etc.) at time t with probability μ_t. Each entrant E_t has two options at t: enter, e, or not, n. If E_t enters, I can fight, f, or accommodate, a. A tough I prefers to fight if E_t enters, but a weak I does not want to fight in a one-shot game.

In a general setup, E_t decides whether to do something that a weak I does not want, and a weak I must decide whether to take an immediate loss (i.e. play f) to deter future entrants from taking that action. Once I accommodates, he is known to be the weak type, and all future entrants will play e.

The critical belief μ_t^* is the probability of I being tough such that E_t is indifferent between entering and staying out at t,

$$\underbrace{u_E(e,a)(1-\Pr(f|\mu_t^*,t)) + u_E(e,f)\Pr(f|\mu_t^*,t)}_{E \text{ plays } e} = \underbrace{u_E(n)}_{E \text{ plays } n} . \tag{1}$$

There is one subtle point here: the probability of fighting is different from the belief that I is the strong type. In general, at t,

$$\Pr(f|\mu_t^*,t) = \underbrace{\Pr(\text{Strong }|\mu_t^*)}_{\mu_t^*} + \underbrace{\Pr(\text{Weak }|\mu_t^*)}_{1-\mu_t^*}\Pr(f|\text{ Weak, } t^*). \tag{2}$$

The reasoning is: for $t < T$, a weak I might fight with some probability to deter future entry. The weak I's willingness to do so depends on $T - t$.

Time T

We start reputation games at time T because, by unravelling, $\Pr(f|\text{ Weak, } T) = 0$ (there is no future behavior to affect, so I plays a static best-response). Using $\mu_T = \Pr(f)$, equation (1) gives us the critical belief μ_T^*.

Then, E plays f if and only if $\mu_T < \mu_T^*$. A weak I does not fight if E enters.

Bayes' Rule, Briefly

The use of Bayes' rule to find μ_{t+1} from μ_t and strategies is initially counterintuitive. It's easiest to understand as a counting argument: at $t+1$, what fraction of incumbents who have always fought are tough types? Tough types always fight, but some weak types also fight. The probability is

$$\mu_{t+1} = \frac{\text{Num. Strong Types}}{(\text{Num. Strong Types}) + (\text{Num. Weak Types})\Pr(f|\text{ Weak, } t)} = \frac{\mu_t}{\mu_t + (1-\mu_t)\sigma_t(f)}$$

where $\sigma_t(f)$ is the probability that a weak I fights at t. Instead of the counting argument, we get the same result from applying Bayes' rule:

$$\begin{aligned}
\Pr(\text{Tough }|f) &= \frac{\Pr(f|\text{ Tough})\Pr(\text{Tough})}{\Pr(f|\text{ Tough})\Pr(\text{Tough}) + \Pr(f|\text{ Weak})\Pr(\text{Weak})} \\
&= \frac{(1)\Pr(\text{Tough})}{(1)\Pr(\text{Tough}) + \sigma_t(f)\Pr(\text{Weak})} \\
&= \frac{\mu_t}{\mu_t + \sigma_t(f)(1-\mu_t)}
\end{aligned}$$

Time $T - 1$

At $T-1$, I's behavior can affect whether E_T plays e at T. Assume that E_{T-1} plays e at $T-1$ and consider our familiar cases.

$\mu_{T-1} > \mu_T^*$: This should feel easy. By Bayes' rule, $\mu_T = \mu_{T-1}$ if I fights, so I fights and E_T plays n. (This requires I to be willing to fight today for a certain gain tomorrow, but this should always be true in these problems.)

$\mu_{T-1} < \mu_T^*$: This is the hard case requiring both players to mix simultaneously. Why is simultaneous mixing necessary? Consider I's pure strategies.

- If I plays f with certainty, then $\mu_T = \mu_{T-1} < \mu_T^*$ and E_T will play e. Since I secured no future benefit by fighting, he would rather deviate and play a at $T-1$.
- If I plays a with certainty, then $\mu_T = 1 > \mu_T^*$. But then a weak I would deviate to f at $T-1$ because E_T would believe he is strong and play n at T.

Then I must mix, but I is only willing to mix if f and a have the same expected payoff at $T-1$. Therefore, to support I's mixing, E_{T-1} must also mix. This is why simultaneous mixing is necessary.

In a bit more depth, I must mix so that $\mu_T = \mu_T^*$ (or else, as in the pure strategy case, E_T plays a pure strategy and I wants to deviate). From Bayes' rule, this requires that the weak I play $\sigma_{T-1}(f)$ such that

$$\mu_T^* = \frac{\mu_{T-1}}{\mu_{T-1} + (1-\mu_{T-1})\sigma_{T-1}(f)}$$

To support I's mixing, E mixes with probability $\sigma_{T-1}(e)$ such that

$$\underbrace{\sigma_{T-1}(e)u_I(e,f) + (1-\sigma_{T-1}(e))u_I(n)}_{T-1 \text{ payoff if } f} + \underbrace{u_I(n,f)}_{T \text{ payoff}} = \underbrace{\sigma_{T-1}(e)u_I(e,a) + (1-\sigma_{T-1}(e))u_I(n,a)}_{T-1 \text{ payoff if } a} + \underbrace{u_I(f,a)}_{T \text{ payoff}}$$

(Note that, when making I indifferent, we need to consider future payoffs.) Finally, using I's mixing strategy, we can calculate the total probability of fighting, $\Pr(f \mid \mu_{T-1},\ T-1)$, and use it to calculate μ_{T-1}^*. Once we have the critical belief μ_{T-1}^*, we can consider $\mu_{T-2} < \mu_{T-1}^*$ and repeat.

The Algorithm

The discussion suggests a general algorithm for solving reputation game problems.

1. At T, E_T and I play a one-shot game. E_T's action is determined by whether $\mu_T < \mu_T^*$.
2. At $T-1$, if $\mu_{T-1} > \mu_T^*$ and E_{T-1} enters, then I fights, $\mu_T = \mu_{T-1} > \mu_T^*$, and E_T does not enter, n.
3. At $T-1$, if $\mu_{T-1} < \mu_T^*$, I mixes so that $\mu_T = \mu_T^*$. E_{T-1} mixes over f and a to make I's payoffs for *the rest of the game* equal.
4. Under I's mixing strategy $\sigma_{T-1}(f)$, calculate the total probability that I fights (considering both strong and weak I) and use it to solve for μ_{T-1}^*.
5. Using μ_{T-1}^*, repeat for $T-2, \ldots$, taking care to adjust for changes in continuation payoffs.
6. Find a pattern in the formula for μ_t^* and generalize.