經營論集, 第50卷 統合號 (2016年 12月)

Plausibility of the Intuitive Criterion of Cho and Kreps*

Cheol Park**

This paper points out a problem associated with the popular intuitive criterion of Cho and Kreps (1987). We use the Spence job market signaling model to illustrate a rather counter−intuitive prediction of the intuitive criterion and show that the criterion is based on the assumption that low type workers either do not realize possible deviation by high type workers, or do not respond to such deviations.

I. Introduction

Signaling games typically have many Nash equilibria, and some equilibria are unreasonable in the sense that players in those equilibria are to entertain unwarranted belief off the equilibrium path. To rule out these unreasonable equilibria, we usually resort to various types of refinements of the Nash equilibrium. Among these refinements, the intuitive criterion of Cho and Kreps seems to be the most powerful and is the most widely used. In this short paper, we plan to point out inherent assumptions behind the intuitive criterion and present another analysis to illustrate the problem that could arise as a result.

*Financial support from the Institute of Management Research and Institute of Finance and Banking of the Seoul National University is gratefully acknowledged.

**Professor, College of Business Administration, Seoul National University

II. The Spence Job Market Signaling Model

We first specify the model and show the usual Perfect Bayesian Equilibrium of the model. There is nothing new in this part of the paper. We simply summarize textbook discussions. One can find similar exposition in any advanced textbook on microeconomics.

We consider a simplified job market signaling model of Spence (1973) with a single firm hiring workers and a single worker looking for a job. Workers are willing to work for any positive wage $w>0$, and firms are willing to hire them as long as they pay less than workers' productivity. A worker's productivity, denoted by t, can be either high(t=t_H) or low(t=t_L) with $t_H > t_L > 0$. A worker of type t is worth \$$t$ to the firm. Productivity is the worker's private information. The firm's prior belief is given by $\mu = \Pr(t{=}t_H)$ with $0 < \mu < 1$.

Before entering the job market, a worker of type i=H,L can acquire education $e \geq 0$ at a cost of $c(e)$=$\theta_i e$. It is assumed that education does not affect productivity; productivity is completely determined by the type of the worker alone. The key assumption is that it costs more for a low type worker to acquire education: $\theta_H < \theta_L$. This is the famous single crossing condition.

For simplicity, we assume workers' preferences are represented by a linear function in $(w.\theta)$: $u_i(w, e) = w - \theta_i e$. This makes things easier to analyze because indifference curves are straight lines with slopes θ_i: $dw/de|_{\bar{u}} = \theta_i$. The following Figure 1 shows indifference curves of the workers.

Let $p_i(e)$ be the probability that a worker of type i=H, L chooses an education level e, and let $\mu(i|e)$ denote the probability that the worker is type i(i=H, L) conditional on the observation of acquired education e. We will assume that the worker has all the bargaining power so that they get paid:

$$w(e) = \mu(H|e)t_H + [1 - \mu(H|e)]t_L$$

Each worker will choose such e that maximizes $w(e) - \theta_i e$. We will use the perfect Bayesian equilibrium (PBE) as our equilibrium concept. The following gives us the exact definition of PBE for our game.

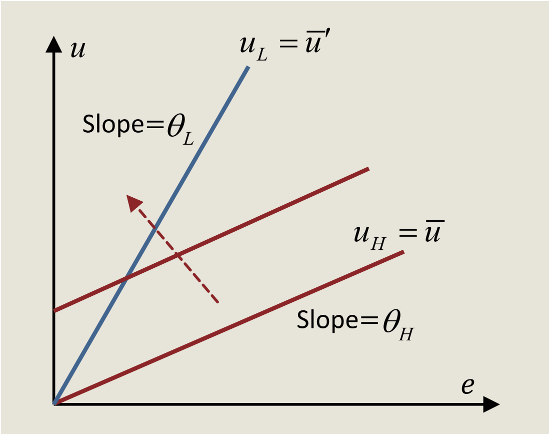

Figure 1. Indifference Curves in (e, u) Space

Definition: A PBE is a strategy $p_i(e)$ and conditional beliefs $\mu(i\,|\,e)$ such that

1. If $p_i(e^*)>0$, then $e^* \in \arg\max_e \left[\mu(H|e)t_H + \mu(L|e)t_L - \theta_i e\right]$
2. $\mu(i|e) = \dfrac{\mu p_i(e)}{\mu p_H(e) + (1-\mu)p_L(e)} (i = H, L)$ whenever $p_i(e) > 0$ for at least one type
3. If $p_H(e) = p_L(e) = 0$, there is no restriction on $\mu(i|e)$
4. Workers are paid $w(e) = \mu(H|e)t_H + \mu(L|e)t_L$

Since education does not affect productivity, the full information equilibrium is $e_H = e_L = 0$ and $(w_H, w_L) = (t_H, t_L)$. However, this is *not incentive compatible* because the low type wants to get paid the high type's wage.

III. Separating Equilibrium

In a separating equilibrium, the type of each worker will be known. Thus a low type worker will obtain no education, so $e_L = 0$ whereas a high type obtains some education: $e_H > 0$. For the low type to choose $e_L = 0$ and to be paid t_L, we should have $t_L \geq t_H - \theta_L e_H$. Thus, we get:

$$e_H \geq \frac{t_H - t_L}{\theta_L} \equiv \underline{e}$$

For the high type, we should have $t_H - \theta_H e_H \geq t_L$, Thus, we should have:

$$e_H \leq \frac{t_H - t_L}{\theta_H} \equiv \overline{e}$$

The two critical values $\underline{e}$ and $\overline{e}$ are depicted in the following Figure 2.

The belief supporting this as an equilibrium is

$$\mu(H|e) = \begin{cases} 0 & if\ 0 \leq e < e_H \\ 1 & if\ e \geq e_H \end{cases}$$

The least costly separating equilibrium is such that the high type worker chooses $\underline{e}$. While the requirements for the PBE do not rule out other equilibria with $e > \overline{e}$, we can eliminate these equilibria using an argument of dominance.

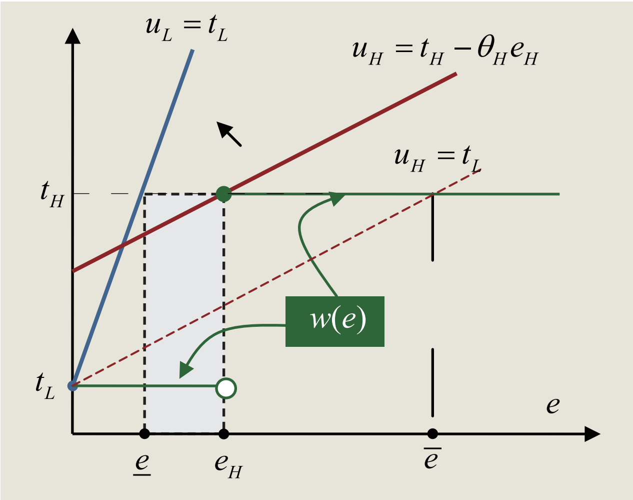

Figure 2. Separating Equilibria

VI. Elimination of All but One Separating Equilibria Using Dominance

Consider a separating equilibrium e_H where $\underline{e} < e_H \leq \overline{e}$. The belief supporting this as an equilibrium satisfies $\mu(H|\underline{e} < e < e_H) = 0$.

A low type worker has no incentive to choose $e \geq \underline{e}$ because even if the firm believes $e \geq \underline{e}$ indicates a high type and pays t_H, a low type worker's expected payoff cannot be higher than t_L (see the Figure 2). In other words, $e \geq \underline{e}$ is a dominated strategy for a low type worker. Thus the correct belief after observing $e \geq \underline{e}$ must be $\mu(H|\underline{e} \leq e < e_H) = 1$. If that is the case, a high type worker choosing $e_H > \underline{e}$ is NOT optimal because $e_H = \underline{e}$ is better. This argument of dominance eliminates all the separating equilibria except for the least-cost equilibrium $\underline{e}$.

1. Pooling Equilibrium

We can now describe the pooling equilibrium. In a pooling equilibrium, all workers acquire the same level of education: $e_H = e_L \equiv e_P$. Since the level of education does not provide any new information, (a) the posterior will be the same as the prior as long as education at the level of e_p is acquired: $\mu(H|e_P) = \mu$, and (b) at e_p, both types will be paid the same wage: $\overline{\omega} \equiv \mu t_H + (1-\mu) t_L$.

Now, the question is, what will happen if $e \neq e_P$ is observed. This is the question of the off-the-equilibrium path belief. Suppose that when faced with $e \neq e_P$, the firm believes that the worker is high type with probability μ'. Then, the low type will choose e_p if

$$\mu' t_H + (1-\mu') t_L - \theta_L e \leq \mu t_H + (1-\mu) t_L - \theta_L e_P \equiv \overline{t} - \theta_L e_P$$

If $e < e_P$, it is required $\mu' \leq \mu$. If $e \geq e_P$, μ' could be higher than μ.

Suppose the belief is such that

$$\mu(H|e) = \begin{cases} \mu & if\ e \geq e_P \\ 0 & if\ e < e_P \end{cases}$$

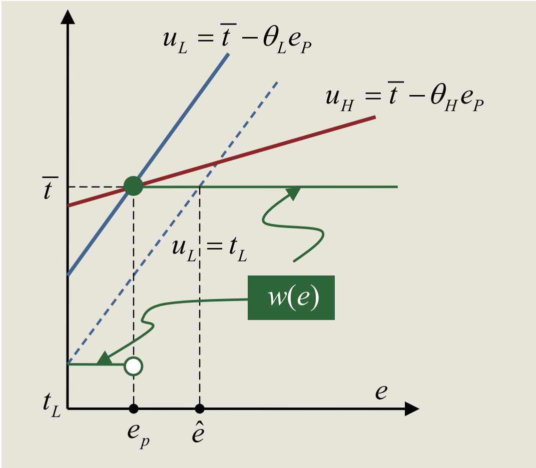

Figure 3. Pooling Equilibria

Then, a low type will choose e_P if and only if $\bar{t} - \theta_L e_P \geq t_L$. Thus, we get the following inequality.

$$0 \leq e_P \leq \frac{\mu(t_H - t_L)}{\theta_L} \equiv \hat{e}$$

Since education does not lead to a higher wage, why would anyone pay for it? In other words, $e_P = 0$ seems to be the only reasonable pooling equilibrium. But, the PBE requirements do not rule out equilibria with $e_P > 0$. Thus there is a continuum of pooling equilibria.

2. Elimination of All Pooling Equilibria Using the Intuitive Criterion

Consider a deviation $e' > e_P$. This deviation cannot be ruled out by the dominance consideration alone because a low type worker can get more than his equilibrium payoff if the firm somehow believes that the deviation has come from the high type.

Now, consider a deviation e''. A low type worker has no incentive to choose this deviation, because he cannot get more than his equilibrium payoff $\bar{t} - \theta_L e_p$. Therefore, the intuitive criterion requires the firm should believe that the deviation e'' comes from a high type worker. Then, as the Figure 4 shows, the high type worker had better choose e'' than e_p. This breaks all

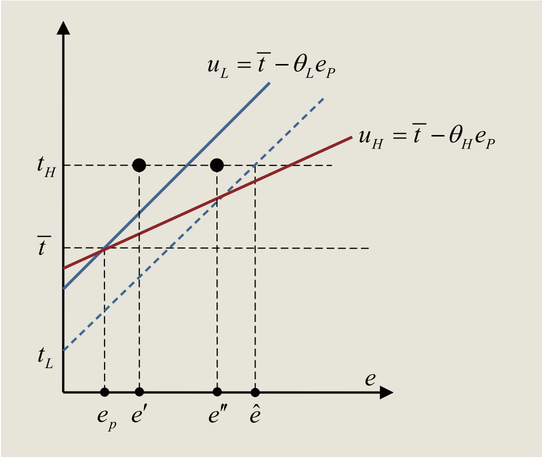

Figure 4. Intuitive Criterion and Pooling Equilibria

the pooling equilibria.

Since the deviation must come from the high type worker $\mu(H|e'')=1$ must hold. Since the pooling equilibrium is supported by belief $\mu(H|e'')=\mu$, this will break all the pooling equilibria.

Cho-Kreps' intuitive criterion may give you unreasonable outcome. Consider the case where $\mu \approx 1$ so that there are very few low type workers. The allocation from the separating equilibrium

$$\left(t_H - \frac{\theta_H (t_H - t_L)}{\theta_L}, t_L\right)$$

is Pareto inferior to that from the pooling equilibrium $(\bar{t}, \bar{t})$. This can be easily seen from the fact that as $\mu \to 1$, we have $(t_H - \bar{t}) \to 0$. Yet, this is the prediction of Cho-Kreps.

3. Plausibility of the Intuitive Criterion

Let's start with the most efficient allocation $(\bar{t}, \bar{t})$ where no worker acquires education and ask whether this is stable in some sense. The high type worker will try to distinguish himself by acquiring some education. At first he may think that by acquiring education between e^* and $\hat{e}$, he will be able to persuade the firm that he is in fact a high type and receive a high wage t_H.

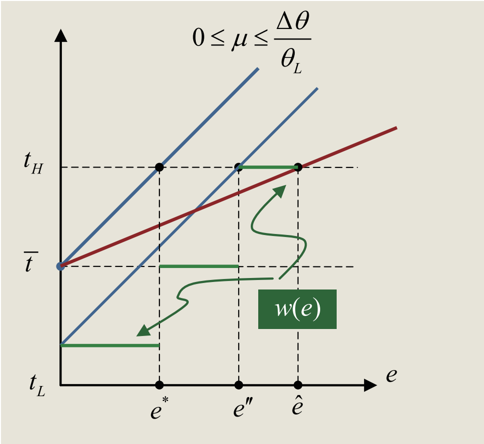

Figure 5. Deviation by High Type Workers

By choosing education between and e^* and $\hat{e}$, the high type worker is telling the firm the following.

> "No low type worker will choose this level of education because it will give him a payoff less than his equilibrium payoff $\bar{t}$. Since only a high type worker can afford to acquire this much education, you must think that I am a high type."

To this, the firm may reply:

> "Your argument is correct only if low type workers do not know what you are doing and keep believing that they will get the equilibrium payoff of $\bar{t}$. If they know what you are trying to do, they will not believe that they will get the payoff of $\bar{t}$. If I take your argument and believe that you are a high type worker, then these low type workers who do not deviate will get t_L, not $\bar{t}$. Therefore, low type workers now have incentives to deviate. Yet they will never choose an education level more than e''. Hence, if you want to persuade me that you are a high type worker, you had better choose education level between e'' and $\hat{e}$."

In other words, the intuitive criterion of Cho and Kreps implicitly assumes that the low

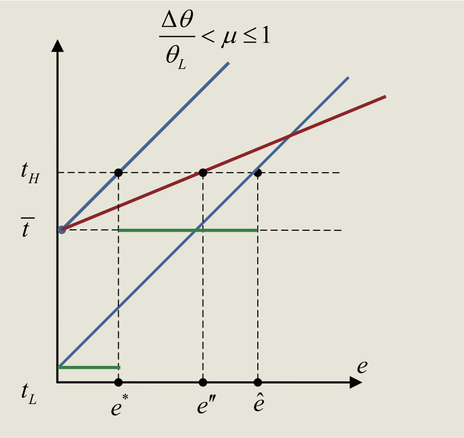

Figure 6. When No Successful Deviation is Possible

type workers will not realize the possibility of deviation by high type workers. If low type workers anticipate such deviation, they will not compare the payoff from deviation with their equilibrium payoff; they will compare it with their payoff when the deviation by the high type worker is successful.

If we accept the argument above, a high type worker can successfully deviate only when $e'' \le \hat{e}$. This is equivalent to:

$$(1-\mu) \ge \frac{\theta_H}{\theta_L} \Leftrightarrow \mu \le \frac{\theta_H - \theta_L}{\theta_L} < 1$$

Thus, if this inequality holds, the high type worker can successfully persuade the firm by deviating and no pooling equilibrium will survive. On the other hand, if the inequality does not hold, or if there are a lot more high type workers, then pooling equilibrium will survive.

The following figure shows the case with $\mu > \frac{\theta_H - \theta_L}{\theta_L}$. In this case, there is no level of education a high type worker can choose that will successfully persuade the firm that he is indeed a high type worker.

V. Conclusion

We illustrate a problem that can arise when we apply the popular intuitive criterion of Cho and Kreps. We believe that a more clear picture will emerge if we model the job market signaling process more properly. For example, in the usual model of job market signaling, the time it takes to get education is not explicitly modeled. If we use some model akin to the overlapping generation model where workers of different generations observe past contract offered for a given level of education, then we may be able to sort out unreasonable equilibria more easily and without a lot of ad-hoc assumptions. But, this work needs time to complete as well.

References

Cho, I.-K. and D. Kreps (1987), "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179-222.

Spence, A. M. (1973), "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355-374.