Application: Job Market Signaling Game

The game we will explore now is attributable to Michael Spence (Quarterly Journal of Economics, 1973). The basic idea is that employees have different innate abilities that affect their productivity. Employers would like to reward employees with abilities that lead to higher productivity, but cannot directly observe these abilities. Alternatively, employees do know if they have the ability to be high productivity workers.

Now suppose that the innate abilities required for high productivity in the work place are the same as those required for high productivity in school. Can employers offer wage contracts that will encourage high productivity workers to distinguish themselves in school?

Suppose we have two players, an employer and employee. There are two types of employee: a high productivity employee denoted by H with probability p and a low productivity employee denoted by L with probability $1 - p$. The employee knows its type, but the employer does not. After nature chooses the employee's type, the employee must choose how much education to receive, $e_t \geq 0$. After the employee chooses how much education to receive, the employer must choose how much to pay the employee given the amount of education the employee has received. To keep things relatively simple, we will assume

(i) the employee's cost of education is $c(H, e) = e^2/2$ for a high type and $c(L, e) = e^2$ for a low type (we really only need to assume the marginal cost of education is higher for the low type: $\frac{\partial c(L,e)}{\partial e} > \frac{\partial c(H,e)}{\partial e}$ for all e, which is referred to as the single crossing condition);

(ii) the employee's output in terms of revenue $y(H, e) = 4 + 4e$ for a high type and is $y(L, e) = 2 + 2e$ for a low type (the important thing here is that the high type is more productive: $y(H, e) > y(L, e)$ for all e);

(iii) the employer earns zero economic profit by setting the wage rate equal to the expected revenue: $w(e) = p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e)$ where $p(H|e)$ is the probability of a high type given education level e, and

(iv) the probability of a high productivity worker is $p = 0.5$.

Figure 1 illustrates these revenues and costs.

This specification yields the payoffs

$$\begin{aligned} p_H(e) &= p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e) - c(H, e) \\ &= 2 + 2p(H|e) + 2e(1 + p(H|e)) - e^2/2 \end{aligned}$$

and

$$\begin{aligned} p_L(e) &= p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e) - c(L, e) \\ &= 2 + 2p(H|e) + 2e(1 + p(H|e)) - e^2 \end{aligned}$$

for the high and low type employees.

Now to answer our question we need to determine if there is a separating Perfect Bayesian equilibrium (PBE) strategy for this game. But before we do that suppose the employer knew the employee's productivity, how much effort would the high and low type employees choose?

The high type employee's optimization problem is

$$\max_{e} \; p_H(e) = 4 + 4e - \frac{e^2}{2},$$

which yields the optimal level of effort $e_H^{CI} = 4$ with the payoff $p_H(e_H^{CI}) = 12$.

The low type employee's optimization problem is

$$\max_{e} \; p_L(e) = 2 + 2e - e^2,$$

which yields the optimal level of effort $e_L^{CI} = 1$ with the payoff $p_L(e_L^{CI}) = 3$. Figure 2 illustrates these equilibrium levels of education and wages.

Note however that $4 + 4 \times 4 - 4^2 = 4 > p_L(e_L^{CI})$. But what does this mean? It means that if the low type were to choose the high type's optimal education and the employer took this to mean that the low type was actually the high type, then the low type would earn more. That is, if the employer does not know the employees' types, then the low type can have an incentive to try to mimic a high type's level of education.

Now let us return to the question of whether or not there is a separating PBE in this game when the employer does not know the employee's productivity. Suppose such an equilibrium exists where e_L^S is the low type employee's equilibrium effort, e_H^S is the high type employee's equilibrium effort, $w_H^S = w(e_H^S) = y(H, e_H^S) = 4 + 4e_H^S$ is the high type employee's equilibrium wage, and $w_L = w(e_L^S) = y(H, e_L^S) = 2 + 2e_L^S$ is the low type employee's equilibrium wage. Assuming $e_H^S \neq e_L^S$, Bayes rule implies $p(H|e_H^S) = 1$ and $p(H|e_L^S) = 0$.

What else must be true for these to be separating PBE strategies and beliefs?

Sequential rationality implies (also known as "Incentive Compatibility Constraints")

(1) $y(H, e_H^S) - c(H, e_H^S) \geq p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e) - c(H, e)$ or

$$4 + 4e_H^S - e_H^{S\,2}/2 \geq 2 + 2p(H|e) + 2e(1 + p(H|e)) - e^2/2$$

for all $e \neq e_H^S$

and

(2) $y(L, e_L^S) - c(L, e_L^S) \geq p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e) - c(L, e)$ or

$2 + 2e_L^S - e_L^{S\,2} \geq 2 + 2p(H|e) + 2e(1 + p(H|e)) - e^2$

for all $e \neq e_L^S$.

We know $p(H|e_L^S) = 0$ and $p(H|e_H^S) = 1$ by Bayes rule, but what about $p(H|e)$ when $e \neq e_L^S$ and $e \neq e_H^S$? Since Bayes rule does not apply, we can choose any $p(H|e)$ that jointly satisfies equations (1) and (2). Let us try $p(H|e) = 0$. That is, we will always assume a low type unless $e = e_H^S$. Equation (1) then implies

$y(H, e_H^S) - c(H, e_H^S) \geq y(L, e) - c(H, e)$ or

$4 + 4e_H^S - e_H^{S\,2}/2 \geq 2 + 2e - e^2/2$

for $e \neq e_H^S$. The maximum of $y(L, e) - c(H, e)$ occurs where $\frac{\partial y(L,e)}{\partial e} = \frac{\partial c(H,e)}{\partial e}$ or for our specific case where $e = 2$ such that $4 + 4e_H^S - e_H^{S\,2}/2 \geq 4$ or $8 \geq e_H^S$.

Alternatively, equation (2) implies

$y(L, e_L^S) - c(L, e_L^S) \geq y(L, e) - c(L, e)$ or

$2 + 2e_L^S - e_L^{S2} \geq 2 + 2e - e^2$

for $e \neq e_L^S$ and $e \neq e_H^S$. The maximum of $y(L, e) - c(L, e)$ occurs where $\frac{\partial y(L,e)}{\partial e} = \frac{\partial c(L,e)}{\partial e}$ or for our specific case where $e = 1$, but this is precisely where $y(L, e_L^S) - c(L, e_L^S)$ and $2 + 2e_L^S - e_L^{S2}$ are optimized which implies e_L^S must set $\frac{\partial y(L,e)}{\partial e} = \frac{\partial c(L,e)}{\partial e}$ or for our specific example $e_L^S = 1$. Note that we must also consider the case where $e = e_H^S$ because $p(H|e_H^S) = 1$, not 0. With $p(H|e_H^S) = 1$, equation (2) implies

$y(L, e_L^S) - c(L, e_L^S) \geq y(H, e_H^S) - c(L, e_H^S)$ or

$3 \geq 4 + 4\, e_H^S - e_H^{S\,2}$,

such that $e_H^S \leq -0.236$ or $e_H^S \geq 4.235$. Therefore, $e_L^S = 1$, $8 \geq e_H^S \geq 4.235$, $p(H|e_H^S) = 1$, and $p(H|e) = 0$ for $e \neq e_H^S$ is a separating PBE. Figure 3 illustrates.

So it is possible to have a separating PBE where low and high types reveal their productivity through their choice of effort. An interesting characteristic of this equilibrium is that the high type has to invest more in education than it would if the employer knew everyone's type. It has to or the low type will have an incentive to mimic it.

We should be careful however before declaring victory because what if there is a pooling equilibrium also? Suppose $e = e^P$ for both types such that $p(H|e^P) = p = 0.5$. Are there values of

e^P that can result in a PBE? Let us check. For e^P to be an equilibrium, sequential rationality implies

(3) $$p\, y(H, e^P) + (1 - p)\, y(L, e^P) - c(H, e^P) \geq p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e) - c(H, e) \text{ or}$$

$$3 + 3\, e^P - e^{P\,2}/2 \geq 2 + 2p(H|e) + 2e(1 + p(H|e)) - e^2/2$$

and

(4) $$p\, y(H, e^P) + (1 - p)\, y(L, e^P) - c(L, e) \geq p(H|e)\, y(H, e) + (1 - p(H|e))\, y(L, e) - c(L, e) \text{ or}$$

$$3 + 3\, e^P - e^{P\,2} \geq 2 + 2p(H|e) + 2e(1 + p(H|e)) - e^2$$

for all $e \neq e^P$. Again, before we can proceed further, we need to specify what happens with $p(H|e)$ off the equilibrium path (e.g. when $e \neq e^P$). Let us continue with the assumption that $p(H|e) = 0$ when $e \neq e^P$. Now for a low type, equations (3) and (4) imply

$$p\, y(H, e^P) + (1 - p)\, y(L, e^P) - c(H, e^P) \geq y(L, e) - c(H, e) \text{ or}$$

$$3 + 3\, e^P - e^{P\,2}/2 \geq 2 + 2e - e^2/2$$

and

$$p\, y(H, e^P) + (1 - p)\, y(L, e^P) - c(L, e^P) \geq y(L, e) - c(L, e) \text{ or}$$

$$3 + 3\, e^P - e^{P\,2} \geq 2 + 2e - e^2.$$

Note that which of these constraint is relevant depends on whether $y(L, e^*) + c(L, e^P) - c(L, e^*)$ is greater than or less than $y(L, e^{**}) + c(H, e^P) - c(H, e^{**})$ where $\frac{\partial y(L,e^*)}{\partial e} = \frac{\partial c(L,e^*)}{\partial e}$ and $\frac{\partial y(L,e^{**})}{\partial e} = \frac{\partial c(L,e^{**})}{\partial e}$. For our specific case, $e^* = 1$ such that $3 + 3\, e^P - e^{P\,2} \geq 3$ or $3 \geq e^P$. Alternatively, $e^{**} = 2$ such that $3 + 3\, e^P - e^{P\,2}/2 \geq 4$, which implies $-1 + 3\, e^P - e^{P\,2}/2 \geq 0$, or $5.65 \geq e^P \geq 0.35$. Therefore, $3 \geq e^P \geq 0.35$, $p(H|e^P) = 0.5$, and $p(H|e^P) = 0$ for $e \neq e^P$ is a pooling PBE. Figure 4 illustrates.

This is not such good news if we were hoping for a unique equilibrium or just hoping that what ever equilibrium we found was a separating equilibrium. But things get even worse. Suppose a high productivity employee chose e_H^{SP} and a low productivity employee chose e_H^{SP} with probability q and e_L^{SP} with probability $1 - q$. Bayes rule then implies that $p(H|e_H^{SP}) = \frac{p(e_H^{SP} \mid H)p}{p(e_H^{SP} \mid H)p + (1 - p(e_H^{SP} \mid H))p} = \frac{p}{p + (1-p)q}$ and $p(H|e_L^{SP}) = 0$. Finally, we will again assume $p(H|e) = 0$ for $e \neq e_H^{SP}$. For what e_H^{SP} and e_L^{SP} is this a PBE? For a high and low type, sequential rationality implies

(5) $\frac{p}{p+(1-p)q}y(H, e_H^{SP}) + \frac{(1-p)q}{p+(1-p)q}y(L, e_H^{SP}) - c(H, e_H^{SP}) \geq y(L, e) - c(H, e)$ or

$$2 + 2\frac{1}{1+q} + 2e_H^{SP}\frac{2+q}{1+q} - e_H^{SP\,2}/2 \geq 2 + 2e - e^2/2$$

and

(6) $q(\frac{p}{p+(1-p)q}y(H, e_H^{SP}) + \frac{(1-p)q}{p+(1-p)q}y(L, e_H^{SP}) - c(L, e_H^{SP}))$
$+ (1 - q)(y(L, e_L^{SP}) - c(L, e_L^{SP})) \geq y(L, e) - c(H, e)$ or

$$q(2 + 2\frac{1}{1+q} + 2e_H^{SP}\frac{2+q}{1+q} - e_H^{SP\,2}) + (1 - q)(2 + 2e_L^{SP} - e_L^{SP\,2}) \geq 2 + 2e - e^2$$

for $e \neq e_H^{SP}$. However, we know something more. For a low type to be willing to mix $p_L(e_H^{SP}) = p_L(e_L^{SP})$ implying for our specific case that

(7) $2 + 2\frac{1}{1+q} + 2e_H^{SP}\frac{2+q}{1+q} - e_H^{SP\,2} = 2 + 2e_L^{SP} - e_L^{SP\,2}$.

Together, equations (6) and (7) imply that $2 + 2e_L^{SP} - e_L^{SP\,2} \geq 2 + 2e - e^2$ for $e \neq e_H^{SP}$. Recall that $2 + 2e - e^2$ is maximized where $e = 1$, which implies $e_L^{SP} = 1$ if equation (6) is going to hold. Equation (7) then implies

$$2 + 2\frac{1}{1+q} + 2e_H^{SP}\frac{2+q}{1+q} - e_H^{SP\,2} = 3 \text{ or}$$

$$q = (1 + 4\, e_H^{SP} - e_H^{SP\,2})/(1 - e_H^{SP})^2$$

for $e_H^{SP} \neq 1$. Since q is a probability $1 \geq q = (1 + 4\, e_H^{SP} - e_H^{SP\,2})/(1 - e_H^{SP})^2 \geq 0$, implying $e_H^{SP} \geq 3$ and $4.235 \geq e_H^S \geq -0.236$ or $4.235 \geq e_H^S \geq 3$.

Finally, we need to verify that equation (5) for $4.235 \geq e_H^{SP} \geq 3$:

(5') $2 + 2\frac{1}{1+q} + 2e_H^{SP}\frac{2+q}{1+q} - e_H^{SP\,2}/2 \geq 2 + 2e - e^2/2$ or
$2 + 2\frac{1}{1+q} + 2e_H^{SP}\frac{2+q}{1+q} - e_H^{SP\,2}/2 \geq 4$

when $e = 2$, the maximum of $2 + 2e - e^2/2$. Substituting q into equation (5') and simplifying yields shows that it will hold provided $e_H^{SP} \geq \sqrt{2}$. Therefore, we also have a partially separating equilibrium where $4.235 \geq e_H^{SP} \geq 3$, $e_L^{SP} = 1$, $q = (1 + 4\, e_H^{SP} - e_H^{SP\,2})/(1 - e_H^{SP})^2$, $p(H|e_H^{SP}) = (1 + q)^{-1}$, and $p(H|e) = 0$ for $e \neq e_H^{SP}$. Figure 5 illustrates.

In terms of predicting behavior, all this is rather disturbing and we might ask ourselves if we can do better. The fact is we can do better if we are willing to place more assumptions on what constitutes a reasonable off the equilibrium path belief. Here we can get what we need by appealing to the ***Intuitive Criterion***. To define the restrictions of the Intuitive Criterion in the context of our game, we need to start by defining what it means to be ***Equilibrium-Dominated*** **(Gibbons, page 239)**:

Given a PBE in a signaling game, the message m_j from M is ***Equilibrium-Dominated*** for type t_i from T if t_i's equilibrium payoff, denoted by $U^*(t_i)$, is greater than t_i's highest possible payoff from m_j: $U^*(t_i) > \max_{a_k \in A} U_S(t_i, m_j, a_k)$.

What does this mean for our game? The m_j's above correspond to e off the equilibrium path and the a_k corresponds to the $w(e)$ chosen by the employer. Suppose the high type plays $e_H^S = 7$ in a separating PBE with the resulting payoff $4 + 4e_H^S - e_H^{S\,2}/2 = 7.5$. If he played $e = 8$, the firm will think he is a low type and he will earn $2 + 2e - e^2/2 = -14$. However, the highest possible payoff he could earn from choosing $e = 8$ is $4 + 4e - e^2/2 = 4$ if it just so happens that the firm chose to respond to $e = 8$ with $w(e) = 4 + 4e$. Of course, this isn't what happens in equilibrium, but out of equilibrium it could happen. Note that $7.5 > 4$, so even if the best possible world emerged from choosing 8 instead of 7, the high type would still have a lower payoff. Therefore, $e = 8$ is equilibrium-dominated.

Now suppose the high type chose $e = 6$ out of equilibrium instead of 7. His equilibrium payoff from this deviation will be $2 + 2e - e^2/2 = -4$. But the highest possible payoff he could earn if the employer chose a wage other than the equilibrium wage is $4 + 4e - e^2/2 = 10$. Now 10 is greater than 7.5, so it is possible for the high type to improve its payoffs by choosing 6 instead of 7 if the employer beliefs were out of equilibrium. In this case $e = 6$, is not equilibrium-dominated.

Now we are ready to state the restrictions on beliefs specified by the ***Intuitive Criterion***:

If the information set following m_j is off the equilibrium path and m_j is equilibrium-dominated for type t_i then (if possible) the receiver's belief $\mu(t_i| m_j)$ should place zero probability on type t_i (This is possible provided m_j is not equilibrium-dominated for all types in T).

In the context of our game, the receiver is the employer and $\mu(t_i| m_j)$ corresponds to $p(H|e)$.

This restriction is very useful for our game. Figure 6 helps illustrate why. Figure 6 shows $y(H, e) - c(H, e)$ and $y(H, e) - c(L, e)$, which are the highest possible payoffs for the high and low types given e. It also shows $y(H, e_H^S) - c(H, e_H^S)$ and $y(L, e_L^S) - c(L, e_L^S)$, which are the separating PBE payoffs for $e_L^S = 1$, $e_H^S = 8$, $p(H|e_H^S) = 1$, and $p(H|e) = 0$ for $e \neq e_H^S$. Note that there are three important regions in this figure with respect to equilibrium-domination. The education levels in Region I correspond to messages that are not equilibrium-dominated for either type. Therefore, we cannot rule out $p(H|e) = 0$ by the intuitive criterion in this region. The education levels in Region II correspond to messages that are equilibrium-dominated for a low type, but not for a high type, which by the intuitive criterion means that $p(H|e) = 1$, not $p(H|e) = 0$. If we set $p(H|e) = 1$ instead of $p(H|e) = 0$ in this region, then $e_H^S = 8$ is no longer a best response and $e_L^S = 1$, $e_H^S = 8$, $p(H|e_H^S) = 1$, and $p(H|e) = 0$ for $e \neq e_H^S$ is not an equilibrium. The education levels in Region III correspond to messages that are equilibrium-dominated for both types. Therefore, we cannot rule out $p(H|e) = 0$ by the intuitive criterion in this region. Similar arguments can be made to rule out all PBE where $e_L^S = 1$, $8 \geq e_H^S > 4.235$, $p(H|e_H^S) = 1$, and $p(H|e) = 0$ for $e \neq e_H^S$.

Figure 7 shows why we cannot also rule out the PBE with $e_L^S = 1$, $e_H^S = 4.235$, $p(H|e_H^S) = 1$, and $p(H|e) = 0$ for $e \neq e_H^S$ using the intuitive criterion. In this instance, we only have two regions. In Region I, all education levels/messages off the equilibrium path for a high type are equilibrium-dominated, but none are for the low type. Therefore, $p(H|e) = 0$ in this region satisfies the intuitive criterion. In Region II, all education levels/messages off the equilibrium path are equilibrium-dominated for both types. Therefore, the intuitive criterion does not rule out $p(H|e) = 0$ in this region.

Figure 8 shows how the intuitive criterion can be used to rule out pooling PBE. The figure explicitly considers the pooling PBE where $e^P = 2$, $p(H|e^P) = 0.5$, and $p(H|e^P) = 0$ for $e \neq e^P$. In Regions I and V, all education levels/messages off the equilibrium path are equilibrium-dominated for both types. Therefore, the intuitive criterion does not rule out $p(H|e) = 0$ in this region. In Region II, the education levels/messages off the equilibrium path are equilibrium-dominated for a high type, but not for a low type. Therefore, the intuitive criterion does not rule out $p(H|e) = 0$ in this region. In Region III, none of the education levels/messages off the equilibrium path are equilibrium-dominated for either type, so the intuitive criterion again does not rule out $p(H|e) = 0$ in this region. In Region IV, the education levels/messages off the equilibrium path are equilibrium-dominated for a low type, but not for a high type, so the intuitive criterion does rule out $p(H|e) = 0$ in this region and instead implies $p(H|e) = 1$. But with $p(H|e) = 1$ in this region, $e = 2$ is no longer a best response for a high type.

Similar arguments can be made to rule out all pooling PBE where $3 \geq e^P \geq 0.35$, $p(H|e^P) = 0.5$, and $p(H|e^P) = 0$ for $e \neq e^P$ and all partially separating PBE where $4.235 > e_H^{SP} \geq 3$, $e_L^{SP} = 1$, $q = (1 + 4\, e_H^{SP} - e_H^{SP\,2})/(1 - e_H^{SP})^2$, $p(H|e_H^{SP}) = (1 + q)^{-1}$, and $p(H|e) = 0$ for $e \neq e_H^{SP}$. The end result is that the intuitive criterion leaves us with a unique separating PBE: $e_L^S = 1$, $e_H^S = 4.235$, $p(H|e_H^S) = 1$, and $p(H|e) = 0$ for $e \neq e_H^S$. What more can we ask for?

Figure 1: Revenues and costs for high and low type employees.

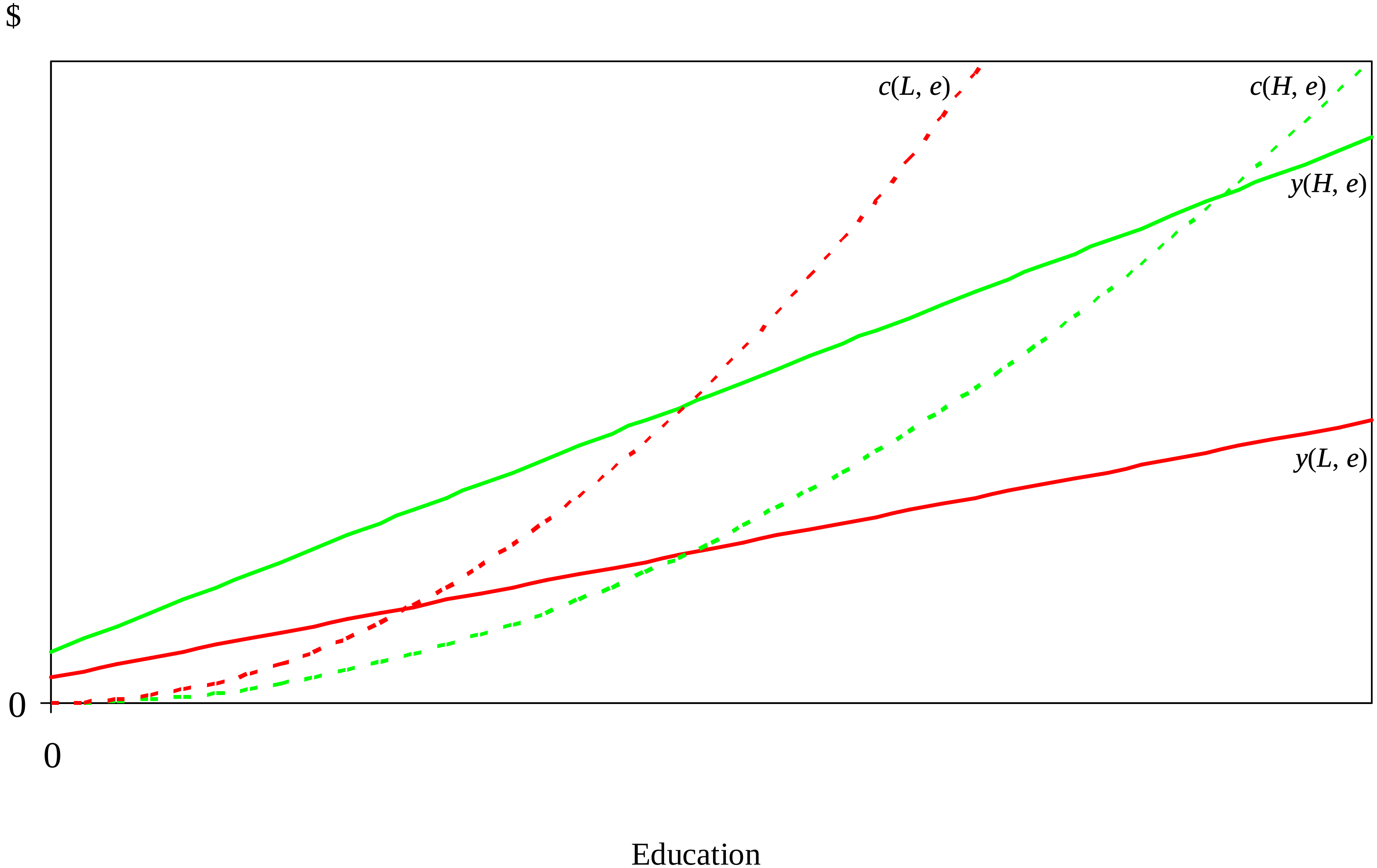

Figure 2: Complete information equilibrium education and wages.

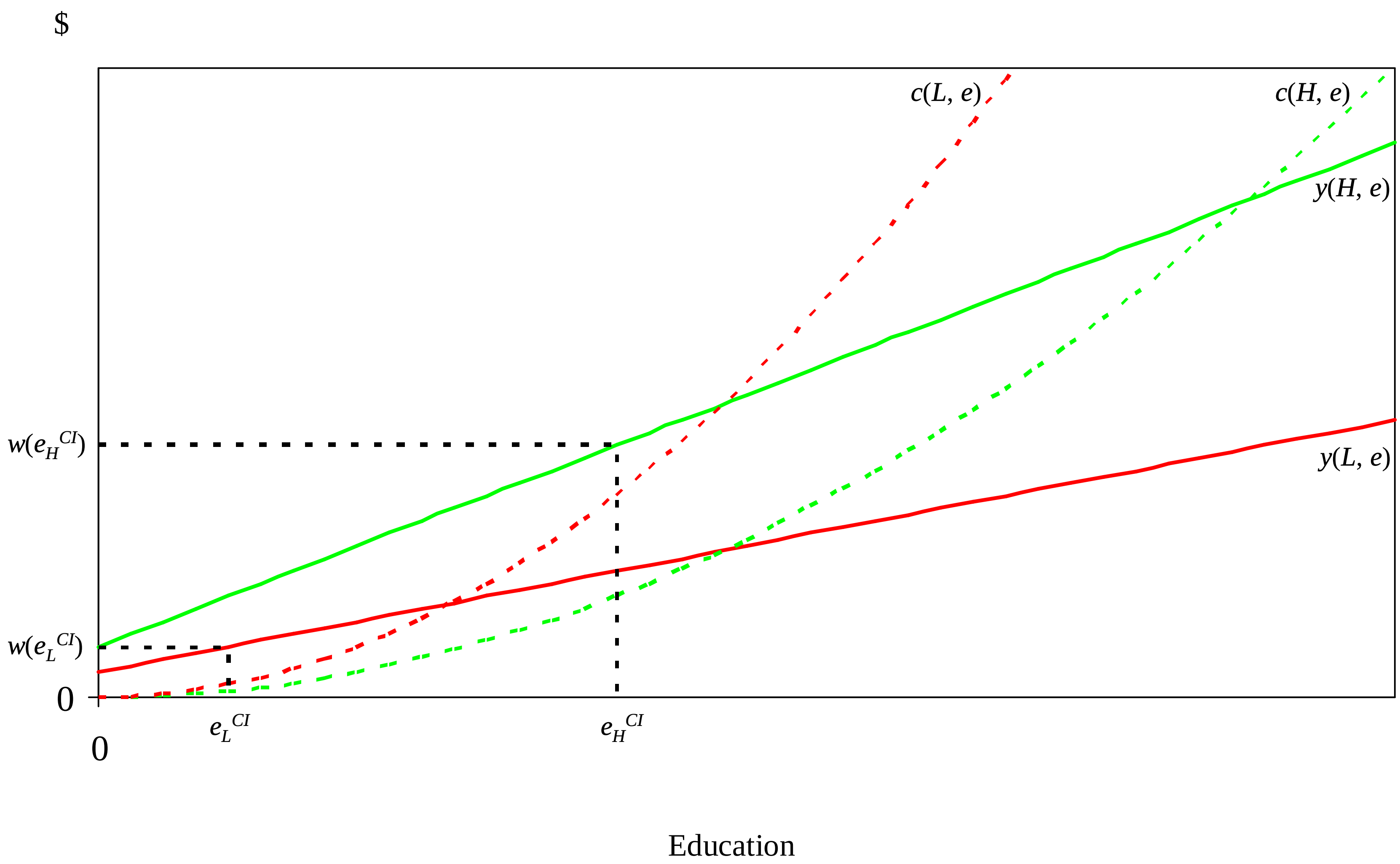

Figure 3: Separating PBE equilibrium education and wages.

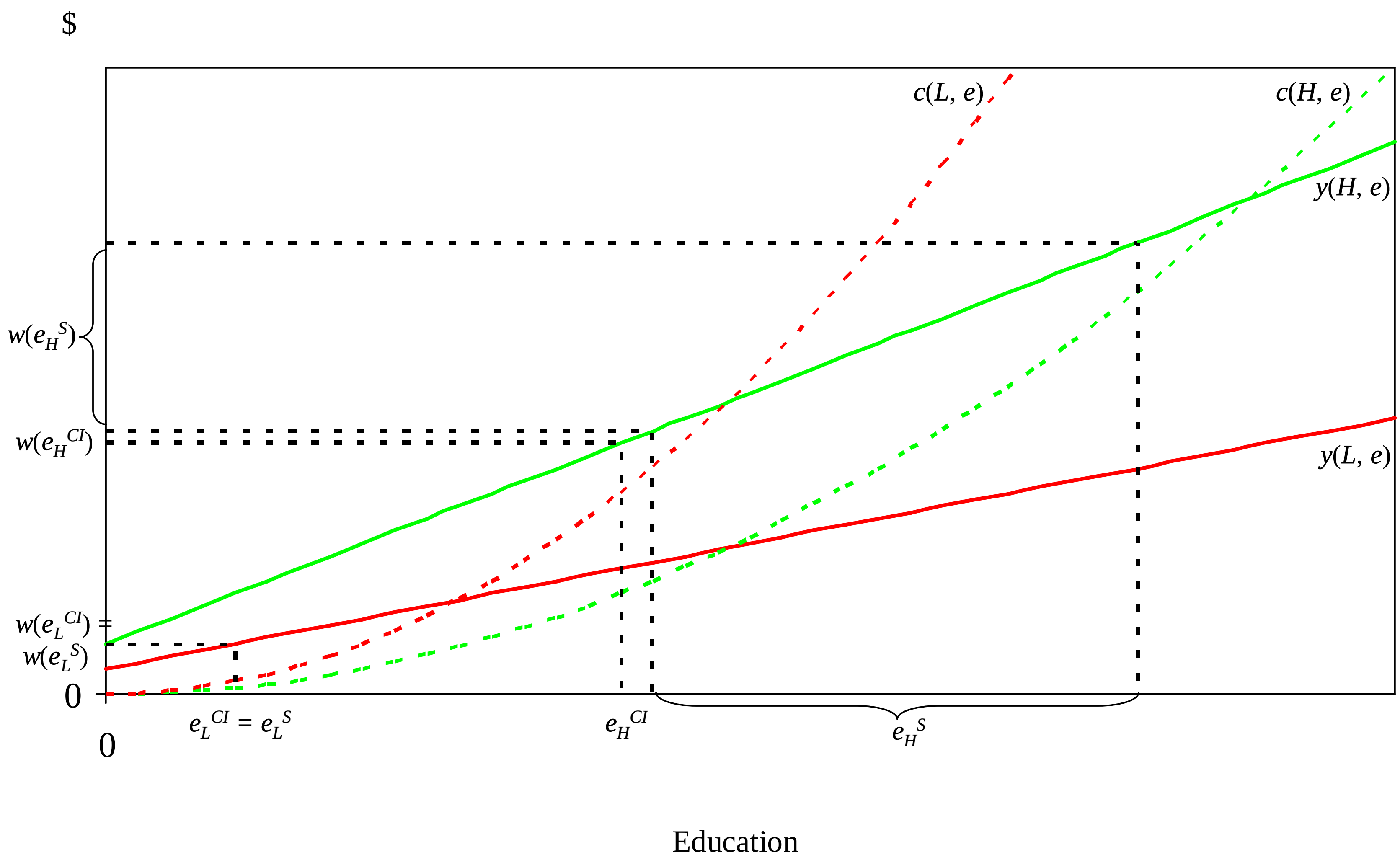

Figure 4: Pooling PBE equilibrium education and wages.

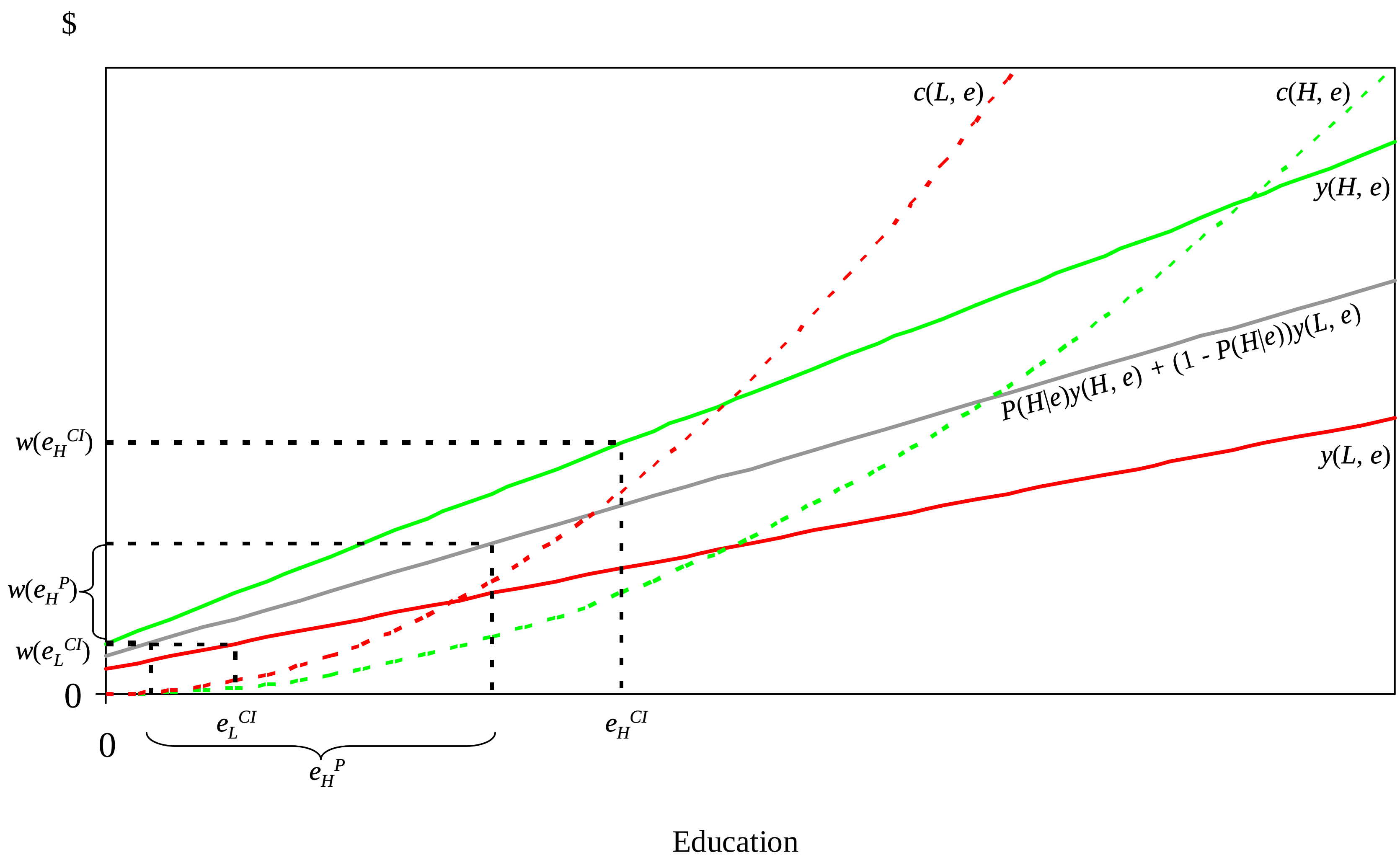

Figure 5: Partially Separating PBE equilibrium education and wages.

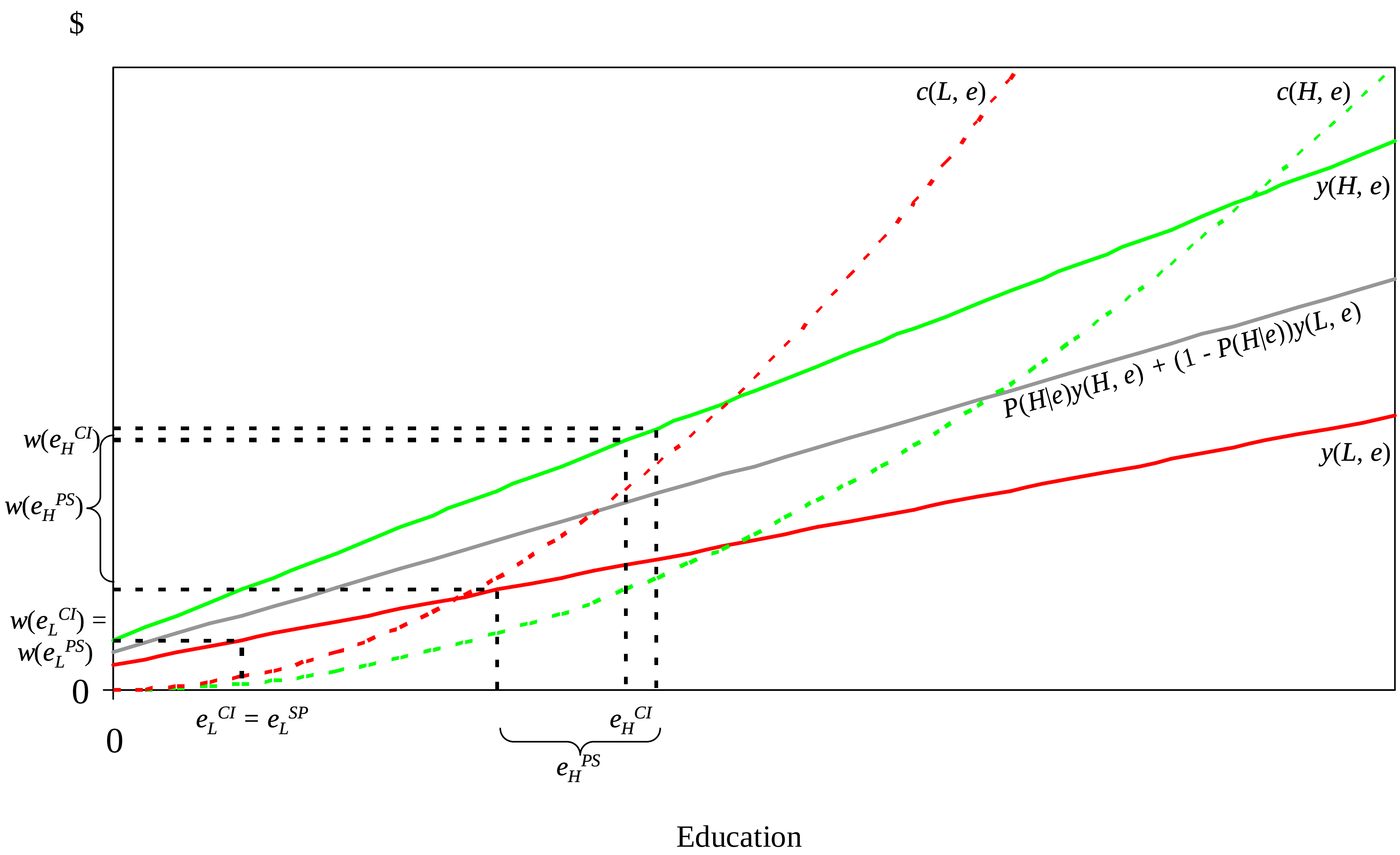

Figure 6: Equilibrium-domination for the separating PBE with $e_H^S = 8$.

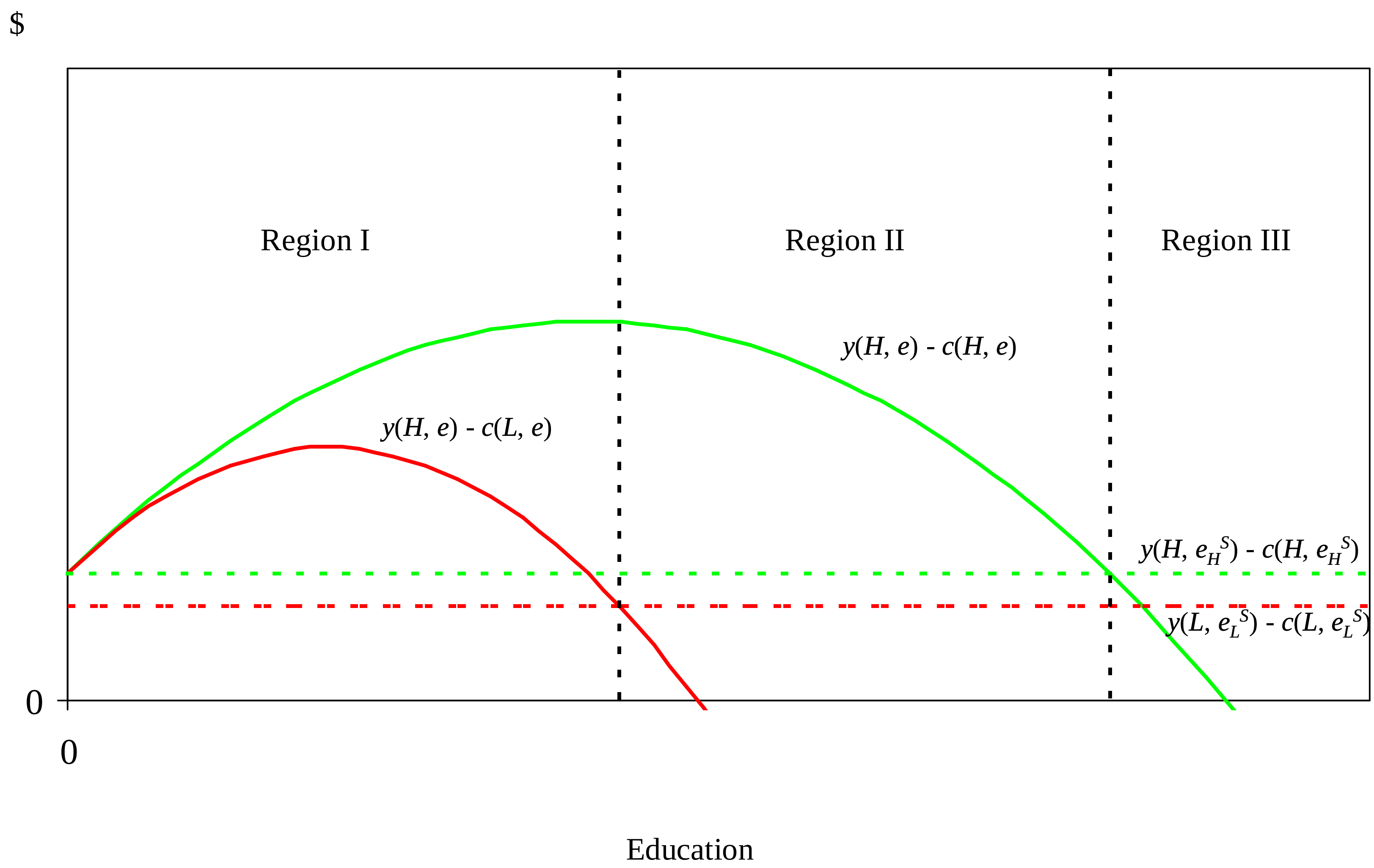

Figure 7: Equilibrium-domination for the separating PBE with $e_H^S = 4.235$.

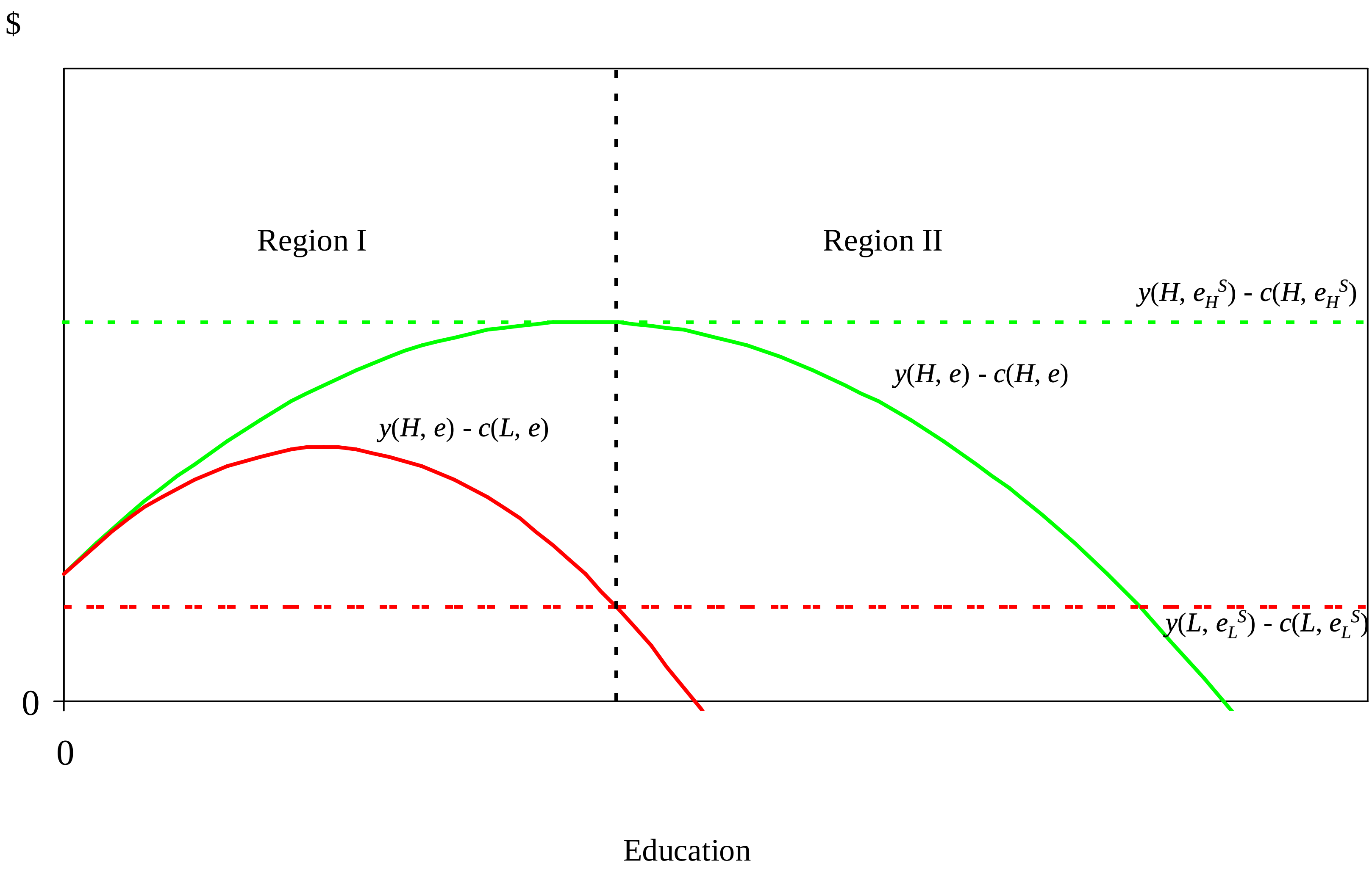

Figure 8: Equilibrium-domination for the pooling PBE where $e^P = 2$.