# A Mathematical Appendix

## A.1 Proof of Lemma 1

In the proof, we will proceed as follows: First, we will study each equilibrium outcome by both identifying the region in the $C_H$-$C_L$ space in which this equilibrium can be realized and analyzing the corresponding payoffs for the firm and agents.

**Opaque Case 1.** We first look at the case where neither $H$ type nor $L$ agents improve on education. In this case, $\gamma_S^E = \gamma_S^B$. The firm's strategy is to use: $P_A = P_B = P_D = 1, P_C = 0$. To guarantee that this is a Nash equilibrium, we need the following conditions to be hold:

$$
\begin{aligned}
&C_H \geq (1-\lambda)R && \text{($H$ type agents will not deviate)}\\
&C_L \geq \lambda R && \text{($L$ type agents will not deviate)}\\
&\gamma_A^E \geq \gamma_{th0} && \text{(the firm will not deviate on $P_A$)}\\
&\gamma_C^E \leq \gamma_{th0} && \text{(the firm will not deviate on $P_C$).}
\end{aligned}
$$

(The first two conditions say that $H$ type and $L$ type agents are better off (in terms of utility) by not switching to improving on education, and the last two conditions say that the firm is better off (in terms of total payoffs) by not switching its strategies on $P_A$ and $P_C$.) The first two conditions specify the regions in $C_H - C_L$ space that can induce this equilibrium (the graph below shows this region). The last two conditions are the direct consequences of Assumption (2):

$$
\frac{\lambda\theta}{\lambda\theta + (1-\lambda)(1-\theta)} \geq \frac{R}{\alpha} = \gamma_{th0} \qquad \frac{(1-\lambda)\theta}{(1-\lambda)\theta + \lambda(1-\theta)} \leq \theta \leq \frac{R}{\alpha} = \gamma_{th0}.
$$

Total payoffs to each side are given by:

$$
\begin{aligned}
\Pi_{firm_O 1} &= \lambda\theta\alpha - (\lambda\theta + (1-\lambda)(1-\theta))R\\
\Pi_{H_O 1} &= \theta\lambda R\\
\Pi_{L_O 1} &= (1-\theta)(1-\lambda)R.
\end{aligned}
$$

where we use $\Pi_{H_O 1}$ and $\Pi_{L_O 1}$ to denote the total payoff of $H$ type and $L$ type agents, respectively. (In the remaining of the proof, we will use $\Pi_{H_O i}$ and $\Pi_{L_O i}$ to denote total payoff of $H$ type and $L$ type agents under case $i$, respectively.)

**Opaque Case 2.** In this case, only $H$ type agents improve on education and we have: $\gamma_B^E = \gamma_D^E = 1$ and $\gamma_A^E = \gamma_C^E = 0$. We find another equilibrium where the firm's strategy is $P_A = P_C = 0$, $P_B = P_D = 1$. The conditions on the parameters are:

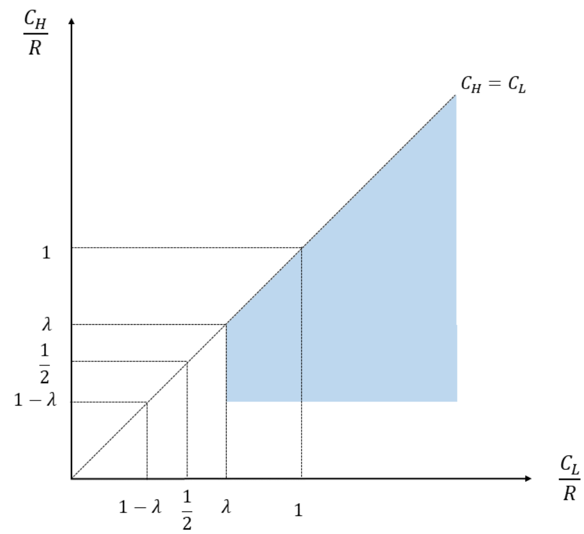


Figure 6: opaque case 1

$C_H \leq R$ ($H$ type agents will not deviate)

$C_L \geq R$ ($L$ type agents will not deviate)

$\gamma_B^E \geq \gamma_{th1}, \gamma_D^E \geq \gamma_{th1}$ (the firm will not deviate on $P_B$ and $P_D$)

$\gamma_A^E \leq \gamma_{th0}, \gamma_C^E \leq \gamma_{th0}$ (the firm will not deviate on $P_A$ and $P_C$).

The first two conditions specify the regions (see the graph below) that can induce this equilibrium. The last two constraints are trivially satisfied.

Total payoffs to each side:

$$\Pi_{firm_O 2} = \theta(\alpha + \beta) - \theta R$$
$$\Pi_{H_O 2} = \theta(R - C_H)$$
$$\Pi_{L_O 2} = 0.$$

**Opaque Case 3.** In the case, both $H$ type and $L$ type agents improve on education. The values of $\gamma^E$'s are given below:

$$\gamma_A^E = \gamma_C^E = 0$$
$$\gamma_B^E = \frac{\lambda\theta}{\lambda\theta + (1-\lambda)(1-\theta)}$$
$$\gamma_D^E = \frac{(1-\lambda)\theta}{(1-\lambda)\theta + \lambda(1-\theta)}.$$

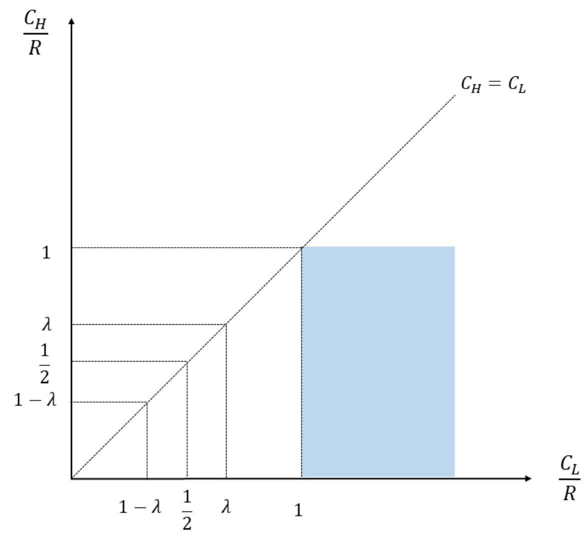


Figure 7: opaque case 2

The firm's strategy is to use $P_A = P_C = P_D = 0, P_B = 1$. To guarantee that this is a Nash equilibrium, we need the following conditions to hold:

$$\begin{aligned}
&C_H \leq \lambda R && \text{($H$ type agents will not deviate)}\\
&C_L \leq (1-\lambda)R && \text{($H$ type agents will not deviate)}\\
&\gamma_B^E \geq \gamma_{th1} && \text{(the firm will not deviate on $P_B$)}\\
&\gamma_D^E \leq \gamma_{th1} && \text{(the firm will not deviate on $P_D$).}
\end{aligned}$$

The first two conditions specify the regions (see graph below) that can induce this equilibrium. The last two conditions are direct consequences of Equation 1 and 2.

Total payoffs to each side:

$$\begin{aligned}
&\Pi_{firm_O 3} = \lambda\theta(\alpha+\beta) + (1-\lambda)(1-\theta)\beta - (\lambda\theta + (1-\lambda)(1-\theta))R\\
&\Pi_{H_O 3} = \theta\lambda R - \theta C_H\\
&\Pi_{L_O 3} = (1-\theta)(1-\lambda)R - (1-\theta)C_L.
\end{aligned}$$

**Opaque Case 4.** In this case, $H$ type agents improve on education with probability $p_H$

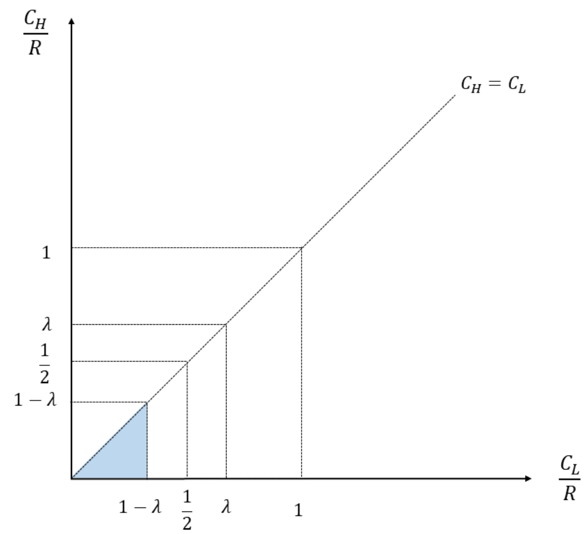


Figure 8: opaque case 3

and $L$ type agents do not improve on education. The values of $\gamma^E$'s are given below:

$$\gamma_A^E = \frac{\theta(1-p_H)\lambda}{\theta(1-p_H)\lambda + (1-\theta)(1-\lambda)}$$
$$\gamma_C^E = \frac{\theta(1-p_H)(1-\lambda)}{\theta(1-p_H)(1-\lambda) + (1-\theta)\lambda}$$
$$\gamma_B^E = \gamma_D^E = 1.$$

The firm's strategy is to use $P_A = p_4, P_C = 0, P_B = P_D = 1$. To guarantee that this is a Nash equilibrium, the following conditions should hold:

$$C_H = (1-\lambda p_4)R \qquad (H \text{ type agents are indifferent})$$
$$C_L \geq (1-(1-\lambda)p_4)R \qquad (L \text{ type agents will not deviate})$$
$$\gamma_A^E = \gamma_{th0} \qquad (\text{the firm is indifferent on } P_A)$$
$$\gamma_C^E \leq \gamma_{th0} \qquad (\text{the firm will not deviate on } P_C).$$

The last condition is satisfied following Assumption 2:

$$\gamma_C^E = \frac{\theta(1-p_H)(1-\lambda)}{\theta(1-p_H)(1-\lambda) + (1-\theta)\lambda} \leq \frac{(1-\lambda)\theta}{(1-\lambda)\theta + \lambda(1-\theta)} \leq \frac{R}{\alpha} = \gamma_{th0}.$$

The first condition could be used to represent $p$ in terms of the other parameters and, similarly, the third condition can be used to represent $p_H$ in terms of the other parameters.

Specifically, we have:

$$p_4 = \frac{1}{\lambda}\left(1 - \frac{C_H}{R}\right) \tag{A.1}$$

$$p_H = 1 - \frac{R(1-\theta)(1-\lambda)}{(\alpha - R)\theta\lambda}. \tag{A.2}$$

Given the fact that $p$ and $p_H$ are values between 0 and 1, we can calculate the range for $C_H$ and $C_L$ that lead to this equilibrium (shown in the graph below):

$$(1-\lambda)R \leq C_H \leq R$$
$$\frac{R - C_H}{R - C_L} \leq \frac{\lambda}{1-\lambda}.$$

where the first inequality follows from the first condition and the second inequality follows from the second condition.

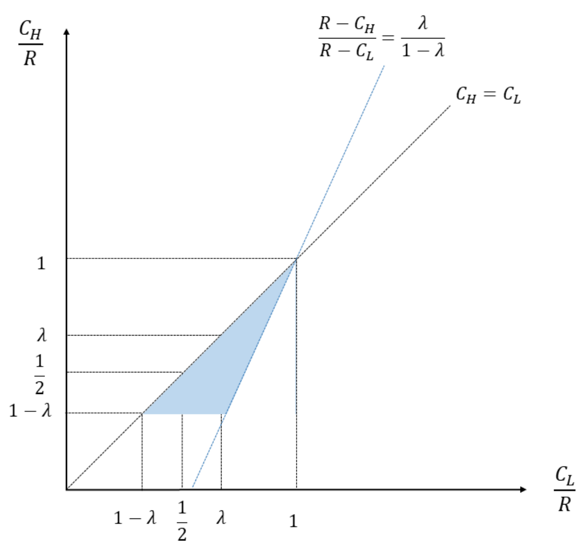


Figure 9: opaque case 4

Total payoffs to each side:

$$\Pi_{firm_O 4} = \theta p_H(\alpha + \beta) - \theta p_H R$$
$$\Pi_{H_O 4} = \theta(R - C_H)$$
$$\Pi_{L_O 4} = p_4(1-\theta)(1-\lambda)R.$$

**Opaque Case 5.** In this case, $H$ type agents improve on education and $L$ type agents improve on education with probability $p_L$. We have:

$$\gamma_A^E = \gamma_C^E = 0$$
$$\gamma_B^E = \frac{\theta\lambda}{\theta\lambda + (1-\theta)p_L(1-\lambda)}$$
$$\gamma_D^E = \frac{\theta(1-\lambda)}{\theta(1-\lambda) + (1-\theta)p_L\lambda}.$$

The firm's strategy is to use $P_A = P_C = 0, P_B = 1, P_D = p_5$. To guarantee that this is a Nash equilibrium, the following conditions should hold:

$$C_H \leq (\lambda + p_5(1-\lambda))R \qquad (H \text{ type agents will not deviate})$$
$$C_L = ((1-\lambda) + p_5\lambda)R \qquad (L \text{ type agents are indifferent})$$
$$\gamma_B^E \geq \gamma_{th1} \qquad (\text{the firm will not deviate on } P_B)$$
$$\gamma_D^E = \gamma_{th1} \qquad (\text{the firm will not deviate on } P_D).$$

The second condition can be used to represent $p$ in terms of the other parameters while the last condition can be used to represent $p_L$ in terms of the other parameters. Specifically,

$$p_5 = \frac{C_L - R}{\lambda R} + 1 \tag{A.3}$$
$$p_L = \frac{\alpha\theta(1-\lambda) - (R-\beta)\theta(1-\lambda)}{(1-\theta)\lambda(R-\beta)}. \tag{A.4}$$

The third condition is satisfied following Equation 2 and 3:

$$\gamma_B^E = \frac{\theta\lambda}{\theta\lambda + (1-\theta)p_L(1-\lambda)} \geq \gamma_{th0}.$$

Given the fact that $p$ and $p_L$ are values between 0 and 1, we can calculate the range for $C_H$ and $C_L$ that lead to this equilibrium:

$$(1-\lambda)R \leq C_L \leq R$$
$$\frac{R - C_H}{R - C_L} \geq \frac{1-\lambda}{\lambda}.$$

Total payoffs to each side:

$$\Pi_{firm_O5} = \theta\lambda(\alpha + \beta - R) + (1-\theta)p_L(1-\lambda)(\beta - R)$$
$$= \frac{2\lambda - 1}{\lambda}\theta(\alpha + \beta - R)$$
$$\Pi_{H_O5} = (\theta\lambda + \theta(1-\lambda)p_5)R - \theta C_H$$
$$\Pi_{L_O5} = 0.$$

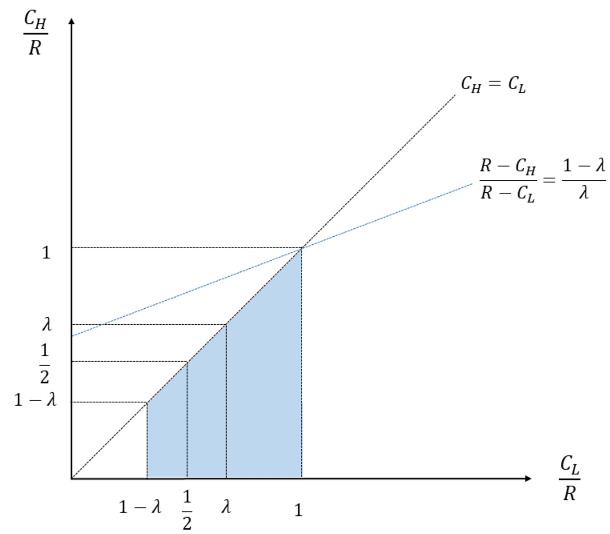


Figure 10: opaque case 5

**Dealing with multiple equilibria.** Per our analysis of the above five cases, there are several regions where multiple equilibria exist. According to the dynamics of the game, in the opaque case, agents move first and the firm moves next. Thus, the actual equilibrium outcome would be the one gives agents the largest total utilities for each agents' type. (In theory, finding such an equilibrium is not always possible; fortunately, it is in our case.)

- In the region where Case 4 and Case 5 overlap, Case 4 always gives higher payoff to both $H$ type and $L$ type agents:

$$\Pi_{H_O 4} = \theta(R - C_H) > (\theta\lambda + \theta(1-\lambda)p_5)R - \theta C_H = \Pi_{H_O 5}$$
$$\Pi_{L_O 4} = p_4(1-\theta)(1-\lambda)R > 0 = \Pi_{L_O 5}.$$

  where the inequalities follow since $p_4 = \frac{R - C_H}{\lambda R}$ and $p_5 = \frac{C_L - R}{\lambda R} + 1$ are values between 0 and 1.

- In the region where Case 4 and Case 1 overlap, Case 1 always gives higher payoff to both $H$ type and $L$ type agents:

$$\Pi_{H_O 1} = \theta\lambda R \geq \theta(R - C_H) = \Pi_{H_O 4}$$
$$\Pi_{L_O 1} = (1-\theta)(1-\lambda)R \geq p_4(1-\theta)(1-\lambda)R = \Pi_{L_O 4}.$$

  where the inequalities follow since $\frac{C_H}{R} \geq 1 - \lambda$ in the overlapped region and $p_4 = \frac{R - C_H}{\lambda R}$ is between 0 and 1.

- In the region where Case 1 and Case 2 overlap, Case 1 always gives higher payoff for both $H$ type and $L$ type agents:

$$\Pi_{H_O 1} = \theta\lambda R \geq \theta(R - C_H) = \Pi_{H_O 2}$$
$$\Pi_{L_O 1} = (1-\theta)(1-\lambda)R \geq 0 = \Pi_{L_O 2}.$$

  where the first inequality follows since $\frac{C_H}{R} \geq 1 - \lambda$ in the overlapped region.

- In the region where Case 1 and Case 5 overlap, Case 1 always gives higher payoff for both $H$ type and $L$ type agents:

$$\Pi_{H_O 1} \geq \Pi_{H_O 4} \geq \Pi_{H_O 5}$$
$$\Pi_{L_O 1} \geq \Pi_{L_O 4} \geq \Pi_{L_O 5}.$$

  where the first inequality follows since $\frac{C_H}{R} \geq 1 - \lambda$ in the overlapped region.

## A.2 Proof of Lemma 2

Similar to the proof of the opaque case, we proceed by analyzing the same five cases analyzed in the opaque case. We show that only the equilibrium outcomes corresponding to cases 1 to 3 are sustainable.

**Transparent Case 1.** In this case, neither $H$ type nor $L$ type agents improve on education. We have: $\gamma_E^E = \theta$ and $\gamma_F^E = 0$. The firm's strategy is to use $P_E = 0$ and $P_F = 1$. To guarantee that this is a Nash equilibrium, the following conditions should hold:

$$C_H \geq R \qquad \text{($H$ type agents will not deviate)}$$
$$C_L \geq R \qquad \text{($L$ type agents will not deviate)}$$
$$\gamma_E^E \leq \gamma_{th0} \qquad \text{(the firm will not deviate deviate on $P_E$).}$$

The last condition follows from Equation 2:

$$\gamma_E^E = \theta < \frac{R-\beta}{\alpha} < \frac{R}{\alpha} = \gamma_{th0}.$$

The region described by the first two conditions on $C_H$ and $C_L$ is shown in Figure 11. The payoffs to each side:

$$\Pi_{firm_T 1} = 0$$
$$\Pi_{H_T 1} = 0$$
$$\Pi_{L_T 1} = 0.$$

where we use $\Pi_{H_T i}$ and $\Pi_{L_T i}$ to denote total utilities of $H$ type and $L$ type agents under case $i$, respectively.

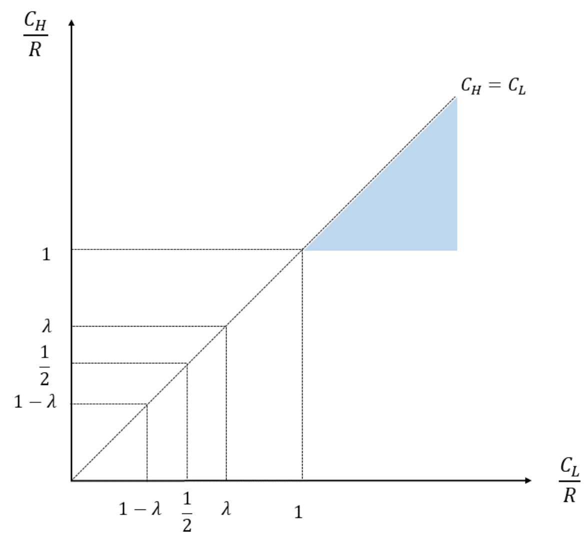


Figure 11: transparent case 1

**Transparent Case 2.** In this case, only $H$ type agents improve on education. We have: $\gamma_E^E = 0$ and $\gamma_F^E = 1$. The firm's strategy is to use $P_E = 0$ and $P_F = 1$. The conditions needed to guarantee that this is a Nash equilibrium are:

$$\begin{aligned} &C_H \leq R && (H \text{ type agents will not deviate}) \\ &C_L \geq R && (L \text{ type agents will not deviate}) \\ &\gamma_E^E \leq \gamma_{th0} && (\text{the firm will not deviate on } P_E) \\ &\gamma_F^E \geq \gamma_{th1} && (\text{the firm will not deviate on } P_F). \end{aligned}$$

The last two conditions are trivially satisfied (by Assumption 1, we have $0 < \gamma_{th1} < \gamma_{th0} < 1$.) The region described by the first two conditions above is shown in Figure 12. The payoffs to each side:

$$\begin{aligned} &\Pi_{firm_T 2} = \theta(\alpha + \beta - R) \\ &\Pi_{H_T 2} = \theta(R - C_H) \\ &\Pi_{L_T 2} = 0. \end{aligned}$$

**Transparent Case 3.** In this case, both $H$ type and $L$ type agents improve on education. We have: $\gamma_E^E = 0$ and $\gamma_F^E = \theta$. The firm's strategy is to use $P_E = 0$ and $P_F = 1$. The

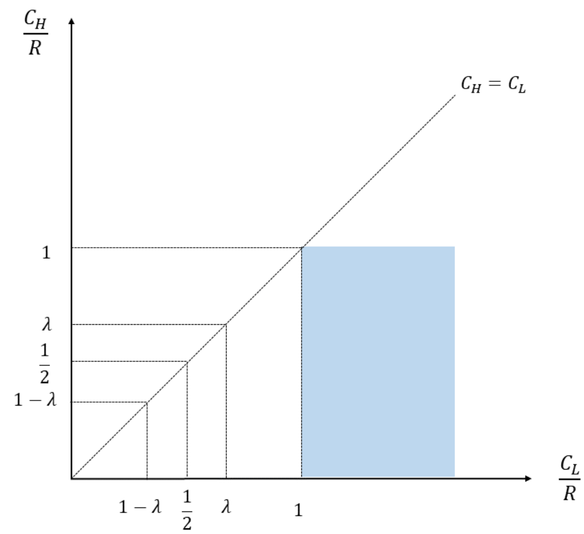


Figure 12: transparent case 2

conditions needed to guarantee that this is a Nash equilibrium are:

$$
\begin{aligned}
&C_H \leq R && (H \text{ type agents will not deviate})\\
&C_L \leq R && (L \text{ type agents will not deviate})\\
&\gamma_F^E \geq \gamma_{th1} && (\text{the firm will not deviate on } P_F).
\end{aligned}
$$

The third condition follows by Equation 3. The first two conditions specify the regions in the $C_H - C_L$ space as shown in Figure 13. The payoffs to each side:

$$
\begin{aligned}
&\Pi_{firm_T 3} = \theta(\alpha+\beta) + (1-\theta)\beta - R\\
&\Pi_{H_T 3} = \theta(R - C_H)\\
&\Pi_{L_T 3} = (1-\theta)(R - C_L).
\end{aligned}
$$

**Transparent Case 4.** In this case, $H$ type agents improve on education with probability $p_H$ and $L$ type agents do not improve on education. We have:

$$
\begin{aligned}
\gamma_E^E &= \frac{(1-p_H)\theta}{(1-p_H)\theta + (1-\theta)}\\
\gamma_F^E &= 1.
\end{aligned}
$$

The firm's strategy is to use $P_E = p$ and $P_F = 1$. The conditions needed to guarantee that

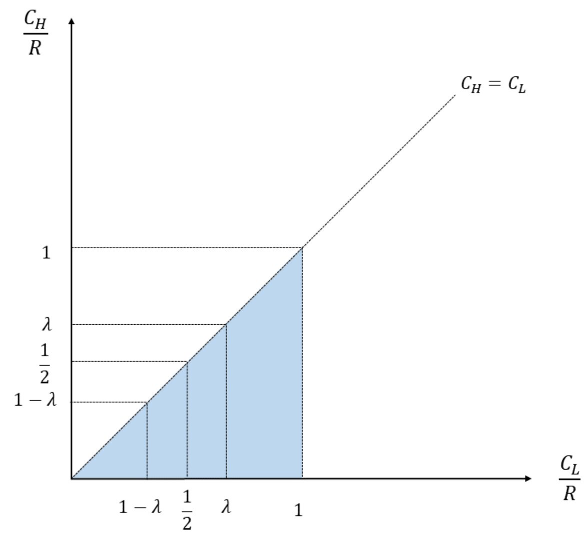


Figure 13: transparent case 3

this is a Nash equilibrium are:

$$
\begin{aligned}
&C_H = (1-p)R && (H \text{ type agents are indifferent})\\
&C_L \geq (1-p)R && (L \text{ type agents will not deviate})\\
&\gamma_E^E = \gamma_{th0} && (\text{the firm is indifferent on } P_E).
\end{aligned}
$$

Since $p_H$ is between 0 and 1, $0 < \gamma_E^E < \theta$. The last condition requires $0 < \frac{R}{\alpha} < \theta$, or equivalently $\alpha > \frac{R}{\theta}$. In the range of $\alpha$ that we are considering (i.e., $\frac{(\theta\lambda+(1-\theta)(1-\lambda))R}{\theta\lambda} < \alpha < \frac{R}{\theta}$, by Assumption 2), this equilibrium cannot be sustained.

**Transparent Case 5.** In this case, $H$ type agents improve on education and $L$ type agents improve on education with probability $p_L$. We have:

$$
\begin{aligned}
\gamma_E^E &= 0\\
\gamma_F^E &= \frac{\theta}{\theta + (1-\theta)p_L}.
\end{aligned}
$$

The firm's strategy is to use $P_E = 0$ and $P_F = p$. The conditions needed to guarantee that this is a Nash equilibrium are:

$$
\begin{aligned}
&C_H \leq pR && (H \text{ type agents will not deviate})\\
&C_L = pR && (L \text{ type agents are indifferent})\\
&\gamma_F^E = \gamma_{th1} && (\text{the firm is indifferent on } P_F).
\end{aligned}
$$

Note that, the second condition implies $p = \frac{C_L}{R}$. Given that $p$ is between 0 and 1, any value of $C_L$ between 0 and $R$ is valid. As for the last condition, $p_L$ is between 0 and 1 implies $\theta < \gamma_F^E < 1$. But, by Assumption 3, $\beta > R - \theta\alpha$, which implies $\gamma_{th1} = \frac{R-\beta}{\alpha} < \theta$. Thus, the last condition cannot be satisfied and, therefore, this equilibrium cannot be sustained.

## A.3 Derivation of the lower and upper bound of $\alpha$ and $\beta$

In this paper we assume $\alpha$ to be in a certain range to eliminate uninteresting scenarios:

$$\frac{(\theta\lambda + (1-\theta)(1-\lambda))R}{\theta\lambda} < \alpha < \frac{R}{\theta}.$$

In the opaque case, when there is no agent improves on the causal feature, we want the firm to hire some agents based on the information in the correlational feature instead of not hiring anyone. (not hiring anyone in this case is uninteresting because it will trivially drive everyone improving on the causal feature). Thus we want $\alpha$ to be large enough to incentivize the firm hiring agents who have value 1 on the correlational feature. In the transparent case, when there is no agent improves on the causal feature, all the agents are mixed together in the feature space: they all have the same values on both the causal and correlational feature. The firm will either hire everyone or not hire anyone, depending one whether the average productivity of all the agents exceeds the salary or not. We want $\alpha$ to be small enough that the firm will not hire anyone in this case. (hiring everyone in this case is uninteresting because no one will have incentive to improve on the causal feature regardless of the cost of improving). Specifically, rewrite the left inequality as $\theta\lambda\alpha + (1-\theta)(1-\lambda) \times 0 > (\theta\lambda + (1-\theta)(1-\lambda))R$. In the initial distribution of the opaque case, (i.e., where everyone has a value of 0 on the causal feature), there are $\theta\lambda$ $H$ type agents and $(1-\theta)(1-\lambda)$ $L$ type agents who have value 0 on the causal feature and value 1 on the correlational feature. The inequality means that their total productivity (left hand side) should be larger than the total salary paid to them (right hand side). In other words, the firm has an incentive to hire all of these agents. If this is not the case, then the firm will not hire anyone with value 0 on the causal feature even if they have value 1 on the correlational feature which will trivially incentivize individuals to improve on the causal features. The right inequality means in the transparent case where the correlational feature is gamed, if everyone has value 0 on the causal feature, the firm will not hire anyone.

In this paper we also assume $\beta$ to be in a certain range to eliminate uninteresting scenarios:

$$R - \theta\alpha < \beta < R - \frac{\theta(1-\lambda)\alpha}{\theta(1-\lambda) + (1-\theta)\lambda}.$$

$\beta$ is the marginal effect of education on the agent's productivity.

Rewrite the left part in-equation as $\theta(\alpha+\beta)+(1-\theta)\beta > (\theta+(1-\theta))R$. In the transparent case where everyone games on the correlational feature and everyone improves on the causal feature, there are $\theta$ $H$ type agents and $1-\theta$ $L$ type agents who have value 1 on both feature. The inequality means their total productivity (left hand side) is larger than the total salary paid to them (right hand side). In other words, the firm will have incentive to hire all of them. If this is not the case, then in transparent scenario no one will improve on the causal feature and the firm will end up hiring no one. As for the right part in-equation, rewrite it as $\theta(1-\lambda)(\alpha+\beta) + (1-\theta)\lambda\beta < (\theta(1-\lambda) + (1-\theta)\lambda)R$. In the opaque case where no one games on the correlational feature but everyone improves the causal feature, there are $\theta(1-\lambda)$ $H$ type agents and $(1-\theta)\lambda$ $L$ type agents who have value 1 on the causal feature but value 0 on the causal feature. This in-equation means their total productivity (left hand side) is smaller than the total wage paid to them (right hand side). in other words, the firm will have no incentive to hire anyone of them. If this is not the case, then in the opaque case improving on the causal feature will ensure an agent to be hired regardless of his value on the correlational feature, which will again, lead to an uninteresting equilibrium.