

Improvement of academic performance by studying in pairs: an analysis under incomplete information about academic skills

Nataly Hernández-Grageda*

(Recibido: noviembre 2017/Aceptado: abril 2018)

Abstract

This paper focuses on analysing the process of partnership formation and its effects on academic performance. We model the formation of studying partnerships as a Bayesian game. Students utility functions are in the spirit of firms' profit functions where the time they devote to study is the input to produce human capital. Academic skills are uniformly distributed, and every student is assumed to know her academic skills. We find that a student decides to study with a classmate because she believes that the classmate has better academics skills. We also find that, in equilibrium, the time a student is willing to spend studying with a classmate increases with the mean and the variance of the distribution of the academic skills. Therefore, under incomplete information, we expect students to devote more time studying with a classmate regardless whether they are studying in a high standard school.

Keywords: academic performance, incomplete information, study partnership formation, human capital, applied Bayesian games.

JEL classification: C72, D71, I21.

* SEPI-ESE, Instituto Politécnico Nacional, Plan de Agua Prieta 66, Miguel Hidalgo, Mexico City 11340, Mexico. E-mail address: nhg812@gmail.com, tel. +52 1 (55) 2719 3309.

Mejora del rendimiento académico mediante el estudio en parejas: un análisis con información incompleta sobre las habilidades académicas

Resumen

Este artículo se enfoca en analizar el proceso de formación de grupos de estudio y sus efectos en el desempeño académico. Modelamos la formación de parejas de estudio como un juego Bayesiano. Las funciones de utilidad están inspiradas en las funciones de beneficios de las empresas, donde el tiempo dedicado a estudiar es el insumo para producir capital humano. Las habilidades académicas están uniformemente distribuidas, y se asume que cada estudiante conoce sus habilidades académicas. Encontramos que un estudiante estudia con algún compañero si considera que este compañero posee habilidades superiores a las suyas. También encontramos que, en equilibrio, el tiempo que un estudiante está dispuesto a estudiar con un compañero aumenta con la media y la varianza de la distribución de habilidades académicas. Por lo tanto, bajo información incompleta, se espera que los estudiantes dediquen más tiempo a estudiar con sus compañeros sin importar si se encuentran inscritos en una escuela de alto rendimiento.

Palabras clave: desempeño académico, información incompleta, formación de parejas de estudio, capital humano, juegos bayesianos aplicados.

Calsificación JEL: C72, D71, I21.

1. Introduction

Some studies have concluded that economic growth is positively influenced by education (Islam *et al.* (2015), Barro (2013) or Hanushek and Woessmann (2008)). As a consequence, improving academic performance is a widely pursued goal.

Among the many factors that might affect academic performance, studying with studying-partners has proven, in empirical and experimental studies, to have a positive effect. Eisenkopf (2010) confirms the positive peer effects in a learning context. He finds that a peer has a motivational influence before and afterwards the actual cooperation takes place.

He also provides some insights in partnership interaction. Another example is Carman and Zhang (2012) who evidence, with data from a Chinese middle school, that peers have positive influence on math test scores. They also find that working in pairs shows positive effects even in contexts different than school. In addition, Falk and Ichino (2006) find that productivity increases when individuals work in pairs. They perform an experiment where people are remunerated (independently of the production level) for doing an activity. There are two groups: in one group there are single workers while in the other group there are pairs of workers. The second group turned out to be more productive in average, even when matched workers did their jobs separately.

Despite the findings of the empirical and experimental studies, to the best of our knowledge, there is no formal model that focuses on analysing the process of partnership formation and its effects on academic performance. This paper aims to fill that gap.

We now describe the model in greater detail. We analyze group formation as a non-cooperative process under incomplete information. Following Lazear (2001), we assume that the acquiring knowledge process is analogous to the production process of a firm. There are two representative individuals in a classroom. As in Robertson and Symons (2003), we assume that time, expressed in hours, devoted to study is the main input to produce human capital. We also assume that the students are aware that each hour of study might improve their human capital. However, each hour devoted to study implies a cost as it means less time for leisure. Then, the net utility of study looks pretty much as a profit function: benefit minus costs. There is incomplete information about the academic skills of the students. Every student is assumed to know its academic skills which are represented by a random variable uniformly distributed. Then, every student must decide, under incomplete information, whether to study along with a classmate and for how long.

We find three main results:

- 1) A student decides to study with a classmate because she believes that the classmate has better academics skills;
- 2) A high-skilled student will spend longer periods of time studying alone, and;
- 3) The time a student is willing to spend studying with a classmate increases with the mean and the variance of the distribution of the academic skills.

We find the last result very interesting as it ensures conditions for students, regardless of the kind of school in which they are studying, to be willing to devote time to study with a classmate. On one hand, high standard schools are expected to have higher mean. However, the academic skills of their students might have a low variance. Then the students in this kind of schools will devote more time studying with a classmate due to the high mean of the academic skills. On the other hand, schools with many students in every classroom, as public schools, are expected to have higher variance. However, as public schools are mostly believed as low standard schools, the academic skills of their students might have a low mean. Then the students in this kind of schools, especially students whose academic skills are above the mean, will devote more time studying with a classmate due to the high variance of the academic skills. Then, in both kinds of schools, there might be very talented students who maximize their transformation function by devoting a high amount of time studying alone and/or with a classmate.

Our results relay on the assumption that the academic performance, taken as acquisition of knowledge, is similar to the production process of a firm. This assumption strong as it is, allows analysing the different factors that enhance group formation in schools. However, as grading could be an imperfect way of measuring the acquisition of knowledge, this model is limited. Another limitation comes from the difficulty of knowing the cost of studying for every student. Despite of the limitations, the model gives useful insights about the peer positive effects on academic performance.

This paper is related to the vast literature on the factor that positively affects academic performance. In this regard, Nonis and Hudson (2006) show that personal ability positively influences academic performance when it is in tandem with longer time for studying. Another example is Crede and Kuncel (2008) who find that study habits, skills, and motivation improve academic performance more than any other non-cognitive variable in college students.

Closer to our work are the non-cooperative models in Corgnet (2010) and Hollard (2000).¹ Corgnet (2010) models the effect that bias on own and peers' abilities has over group formation in workplaces. He assumes that total outcome of working in group can be shared according to an allocation

¹ There are also some studies that characterize group formation by cooperative game models (See for example Banerjee *et al.* (2001)).

rule and an additive way in which the abilities of workers interact. On the other hand, Hollard (2000) provides conditions to guarantee the existence of pure-strategy Nash equilibrium in a class of group formation games with finite sets of actions. In contrast, our model considers a game with infinite sets of actions (intervals). We provide conditions to determine the amount of time studying with a classmate in a Bayes-Nash equilibrium.

The paper is organised as follows: In Section 2, we present our model. In Section 3, we characterise the decision rules. In Section 4, we provide conditions to study with a classmate. In Section 5, we present an example. Section 6 concludes. The appendix contains most of the formal proofs.

2. The model

Our model bases on the theory of the firm and follows the approach of Lazear (2001). There is a classroom full of students who are starting a new course. Consider two representative individuals; say students 1 and 2. Following Osborne and Rubinstein (1994), we assume that student 1 is a female while student 2 is a male. A students' academic skills are represented by a random variable, θ , and each individual knows her/his skills. We also assume that they are aware that each hour of study might improve their human capital. However, each hour devoted to study implies a cost as it means less time for leisure (time with family and friends, in cultural and recreational activities, etcetera).

2.1. *Transforming hours of study into knowledge*

We start by defining the net transformation function of a student as follows:

$$r(x) = \theta f(x) - c(x), \quad (1)$$

where x refers to hours devoted to study alone. The term $\theta f(x)$ represents the accumulated knowledge while the term $c(x)$ captures the cost or disutility of studying. Thus, $r(x)$ refers to the net utility obtained by studying alone. As in Foster and Frijters (2010), the net transformation function $r(x)$ is an ordinal function.

As in the theory of the firm, we assume that f and c are standard production and cost functions, respectively. We assume:

Assumption 1. f is a concave function, twice differentiable: $f' > 0$, $f'' \leq 0$, and $f(0) = 0$; c is a convex function, twice differentiable: $c' > 0$, $c'' \geq 0$, and $c(0) = 0$.

The parameter θ represents the individual academic skills. Then, as a technology parameter does in firm theory, the parameter captures the *type* of every student according to her/his academic skills.

The students should decide how long to study. This time could be spent studying alone and/or with a studying partner. Such a decision is taken under incomplete information as student 1 knows her own type, θ_1 , while she does not know the corresponding type of student 2, given by θ_2 . The opposite holds as well. The probability distributions of θ_1 and θ_2 are common knowledge as described in the following assumption:

Assumption 2. The random variable Θ refers to academic skills and is uniformly distributed on $[\mu - \varepsilon, \mu + \varepsilon]$. Where $\mu > \varepsilon > 0$ and $\mu - \varepsilon > 1$.²

2.2. Transforming hours of study into knowledge when studying along with a classmate

Student 1 has to decide on the amount of time, in hours, she is studying alone, x_1 , and the amount of time, s_1 , she will study along with student 2. Then, student 1 faces the following maximization problem:

$$\max_{x_1, s_1} r_1(x_1, s_1, s_2) = \theta_1 f(x_1) - c(x_1) + \alpha \theta_1 \theta_2 f(s_1 \wedge s_2) - \beta c(s_1 \wedge s_2) \quad (2)$$

subject to $x_1 + s_1 \leq T$. Where T refers to the maximum time allowed to study per day. The term $x_1 \wedge s_2$ captures the amount of time spent studying with a studying-partner. As the students study together if and only if both agree on doing so, the term $x_1 \wedge s_2$ denotes the minimum value of s_1 and s_2 .

The parameters $\alpha \geq 0$ and $\beta \geq 0$ represent weights on the benefits and costs derived from studying with a studying-partner. They could be interpreted respectively as student 1's beliefs about the benefit and cost from

² The variance of the distribution depends completely on ε as it is given by $\varepsilon^2/3$.

studying with student 2. Following the conclusion of Foster and Frijters (2010) about the complementarity of academic skills, we take the product of both academic skills θ_1 and θ_2 . Then, student 1 has benefits $\alpha\theta_1\theta_2f(s_1 \wedge s_2)$ from studying with a classmate. Which depends on the academic skills of both students $\theta_1\theta_2$ and the parameter α .

Our approach of addressing the students' problem is inspired on the problem of duopoly of Cournot.³ However, to capture the particular features of the students' problems, we need to add some technical complications. First, student 1 affects student 2's utility directly. Second, we assume that the students study together if and only if both agree on doing so. Then, the amount of time they study together is given by the minimum value of s_1 and s_2 .

In the following section we provide conditions for a student to spend the optimum amount of time studying alone and along with a studying-partner.

3. Decision rules

Given the incomplete information on the academic skills of the students, we need to ensure that a student will be willing to study with a classmate. A student of type θ_1 will study with a studying-partner of type Θ , if there exists s such that the following inequality holds:

$$E [\alpha \theta_1 \Theta f(s) - \beta c(s) | \theta_1] > 0 \quad (3)$$

As μ is the mean of Θ , an equivalent statement for expression (3) is as follows: A student of type θ_1 *does not* want to study with a classmate if every s satisfies the following inequality:

$$\alpha \theta_1 \mu f(s) \leq \beta c(s) \quad (4)$$

Inequality (4) holds when the expected utility is lower than the corresponding disutility. Particularly, inequality (4) holds in two cases:

- 1) when the benefit from studying with a classmate is low while the cost is high, this is, for $\alpha \approx 0$ and $\beta \gg 0$ (big enough);

³ See, for instance, Chapter 3 in Gibbons (1992) for further details.

- 2) when the expected skills of the potential studying partner, given by μ , is small enough.

3.1. Equilibrium amount of time devoted to study

Once two students are willing to study together, they decide strategically the amount of time studying alone and/ or along with a studying partner. This fact is captured by the Bayes-Nash *symmetric* equilibrium, $(x_1(\theta_1), s_1(\theta_1), x_2(\theta_2), s_2(\theta_2))$, that solves the following optimization problem:⁴

$$\max_{x_i, s_i} E [r_i(x_i, s_i, s_j(\Theta_j); \theta_i)] \quad i, j = 1, 2, \quad (5)$$

with $x_i(\theta_i) + s_i(\theta_i) \leq T$. Where Θ_i has the same distribution of Θ , that is, Θ_1 and Θ_2 are identical and independent random variables uniformly distributed. The decision rule $s : [\mu - \varepsilon, \mu + \varepsilon] \rightarrow \mathbb{R}$ about the time to study with a studying partner is called a *strategy*.

The following proposition states necessary conditions that a Bayes-Nash symmetric equilibrium of the game must satisfy.

Proposition 1. Suppose that Assumptions 1 and 2 hold. Suppose $(x(\theta_1), s(\theta_1), x(\theta_2), s(\theta_2))$ is a symmetric Bayes-Nash equilibrium such that the functions x and s are differentiable. Then, for each player $i = 1, 2$, $x(\theta_i)$, satisfies the following expression:

$$c'(x(\theta_i)) = \theta_i f'(x(\theta_i)) \quad (6)$$

In addition, assume that the function s is strictly increasing, then

$$\frac{c'(s(\theta_i))}{f'(s(\theta_i))} = \frac{\alpha \theta_i}{2\beta} (\theta_i + \mu + \varepsilon). \quad (7)$$

Proof. See the Appendix.

Expression (6) says that, in equilibrium, the optimal time spent studying alone guarantees that the marginal cost equals the marginal product. A similar interpretation could be given to expression (7).

⁴ See Jehle and Reny (2001) and Mas-Colell *et al* (1995) for further details on Bayes-Nash equilibrium.

For student $i = 1$, expression (6) could be written as

$$\beta c'(s(\theta_1)) = \alpha \theta_1 \left(\frac{\theta_1 + \mu + \varepsilon}{2} \right) f'(s(\theta_1)), \quad (8)$$

Recall that the actual value of θ_2 is unknown. Then, from expression (8), in equilibrium student 1 believes that the type of student 2 lays in the midpoint of the interval $[\theta_1, \mu + \varepsilon]$. In other words, in equilibrium, student 1 believes that student 2 has higher academic skills, given by $\theta_2 = (\theta_1 + \mu + \varepsilon)/2$. It seems that the belief of the academic skills of the potential studying-partner is positively related to the variance of the distribution, defined by ε . If ε approaches to zero, then the skills of student 1 will be close to the mean $\theta_1 \approx \mu$ while student 1 believes that student 2's academic skills approaches to the mean, $\theta_2 \rightarrow \mu$. On the other hand, if ε increases, student 1 believes that student 2 has greater academic skills. Then, the present analysis suggests that, under incomplete information, a student decides to study with a classmate because she believes that the classmate has better academics skills.

This result gives support to the empirical findings of Foster and Frijters (2010) who argue that students believe that benefits of working with a peer come from the skills of the peer. This result also suggests, as stated by De Fraja and Landeras (2006), that the potential benefits from working with a classmate can be lost due to segregation by skills.

Nevertheless, this fact not always holds. There are talented students who are keen to collaborate in study groups with less skilled students. Although this behaviour could be explained by altruism, which is not the topic of the present paper, it could be judged as it may benefit low-skilled students rather than high-skilled ones (Vardardottir (2013)).

Worth noting that every student has also to decide the amount of time spent in studying alone. For a student of type θ , the following proposition states such conditions.

Proposition 2. If Assumptions 1 and 2 hold, then, the function x is strictly increasing in θ .

Proof. See the Appendix.

Previous result implies that a high-skilled student will spend longer periods of time studying alone. This could be due to the fact that a high-skilled student faces a more efficient process to accumulate human capital than a less-skilled student. Then, the best strategy for a high-skilled student would be to devote many hours studying alone.

Nevertheless, this fact not always holds. There are some brilliant students who spend just few hours studying (alone and/or along with a classmate). This could be due to two facts:

- 1) these students value leisure more than acquiring knowledge at school. Thus, their cost of studying for one hour is higher.
- 2) Grading is the method mostly used to evaluate the acquisition on knowledge. As a consequence, some brilliant students could obtain good grades just by studying for few hours. Then, in such a context, the best strategy for a high-skilled student would be to devote just the minimum time of study to obtain decent grades. This effect might be seen more often in schools with low standards.

The results from propositions 1 and 2 suggest that a high-skilled student will spend longer periods of time studying alone. In addition, she believes that her studying-partner has greater academic skills. Then, this behaviour leads to the fact that a high-skilled student will match another high-skilled student to study with. This fact guarantees the best way of transforming hours of study into knowledge. On the other hand, this result also suggests that low-skilled student will study (alone and/or with a classmate) for a lower amount of time.

Previous analysis implies that high-skilled students have to devote most of their time (or even their whole life) to study, alone and along with classmates. On the other hand, low-skilled students must study for little time, just to obtain some basic academic skills to allow them to perform a job. This distribution of academic skills seems to be the best for a society.

This analysis implies two problems of asymmetric information:

- 1) identifying clearly the high-skilled students to encourage them to pursue higher education. This identification should be independent of income, race, skin colour or other aspects.
- 2) as grading is an imperfect way of measuring academic performance, there is the need of designing better measures of the acquisition of knowledge.

In the following subsection we analyse the impact of the different parameters on the amount of time studied with a student-partner.

4. Studying with a classmate

In this section we analyse the way the time a student devotes to study with a classmate is affected by some parameters.

4.1 Amount of time devoted to study with a classmate

Recall that the time devoted to study with a classmate, s , depends on the distribution of the academic skills, Θ , and on the weights on the benefits and costs. That is, s depends on the parameters a, β, μ , and ε . The following proposition describes the way s is affected by the corresponding parameters.

Proposition 3. Suppose that Assumptions 1 and 2 hold. In addition, suppose that the strategy $s(\theta, a, \beta, \mu, \varepsilon)$ is differentiable with respect to each of the parameters, then

$$\frac{\partial s}{\partial a} > 0 \quad (9)$$

$$\frac{\partial s}{\partial \beta} < 0 \quad (10)$$

$$\frac{\partial s}{\partial \mu} > 0 \quad (11)$$

$$\frac{\partial s}{\partial \varepsilon} > 0 \quad (12)$$

Proof. See the Appendix.

Expression (9) says that, keeping everything else constant, in equilibrium, the time a student is willing to spend studying with a classmate increases with her belief about the benefits obtained from doing it, a . The more a student expects to obtain from studying along with a classmate, the longer time she will devote to do it.

Expression (10) says that, keeping everything else constant, in equilibrium, the time a student is willing to spend studying with a classmate decreases with her belief about the costs from doing it, β . The more a student expects to waste from studying along with a classmate, the lower time she will devote to do it.

Expression (11) implies that, keeping everything else constant, in equilibrium, the time a student is willing to spend studying with a classmate increases with the mean of the distribution of the academic skills. Then, in schools with high standards, where the academic skills of the average student are expected to be higher, students will devote longer amounts of time to study with classmates.

Finally, *expression (12)* ensures that, keeping everything else constant, in equilibrium, the time a student is willing to spend studying with a classmate

increases with the variance of the distribution of the academic skills. At a first sight this result seems not to be very intuitive. However, in a classroom where the variance of the academic skills is high, there is a chance that at least one very talented student exists. The apposite holds as well: there is a chance to find there a very bad student. In such environment, a student might be motivated to try to match that very talented student. Then, she will be willing to devote more time studying with a class mate expecting to pick the most talented student as her studying-partner. This behaviour might be seen more often in students whose academic skills lay above the mean, μ , as they believe that the academic skills of their studying-partner is higher. A high variance could be seen in schools, as public schools (that are mostly believed as low standard schools) that have many students in every classroom. Then, as a consequence, in public schools, students with academic skills above the mean could be motivated to devote more time studying with a classmate.

Expressions (11) and (12) together ensure conditions for students, regardless of the kind of school in which they are studying, to be willing to devote time to study with a classmate. On one hand, high standard schools are expected to have higher mean, μ . However, the academic skills of their students might have a low variance. Then the students in this kind of schools will devote more time studying with a classmate due to the high mean of the academic skills. On the other hand, public schools (with many students) are expected to have higher variance, ε . However, the academic skills of their students might have a low mean. Then the students in this kind of schools, especially students whose academic skills are above the mean, will devote more time studying with a classmate due to the high variance of the academic skills. Then, in both kinds of schools, there might be very talented students who maximize their transformation function by devoting a high amount of time studying alone and/or with a classmate.

4.2. Expected amount of time studying with a classmate

In general, every student wants to devote a different amount of time study with a classmate. Then, the actual time a couple of students devote to study together is given by the minimum value between $s(\theta_1)$ and $s(\theta_2)$. The following proposition ensures that the comparative statics given in Proposition 3 is also valid for the amount of time two students expect to study together.

Proposition 4. Suppose that two students are randomly chosen from a classroom in which the academic skills are uniformly distributed. If both students are willing to study together, then, the expected time they study in pairs

$$E[\{s(\Theta_1, \alpha, \beta, \mu, \varepsilon) \wedge s(\Theta_2, \alpha, \beta, \mu, \varepsilon)\}] \quad (13)$$

increases with a , μ , and ε , whereas it decreases with β .

Proof. See the Appendix.

The intuition for this proposition follows closely the intuition of proposition 3. As in there, students who are in classrooms where the academic skill are uniformly distributed with a high mean and high variance will be willing to devote longer time to study with a classmate.

5. An example

In order to illustrate our findings, consider a framework where students have a net transformation function given by: $r(x) = \theta x - x^2$. that this, $f(x) = x$ and $c(x) = x^2$. It is direct to verify that if a student decides to study alone, she will devote $x = \theta/2$ hours to study in order to maximize her transformation function.

Once the student decided to study with a classmate, under asymmetric information, student 1's net utility is.

$$r_1(x_1, s_1, s_2) = \theta_1 x_1 - x_1^2 + \alpha \theta_1 \theta_2 (s_1 \wedge s_2) - \beta (s_1 \wedge s_2)^2 \quad (14)$$

From expression (3) we know that student 1 is willing to study along with a classmate if

$$\alpha \theta_1 \mu > \beta s \text{ for some } s,$$

and from expressions (6) and (7), the amounts of time this student devotes to study alone and with a classmate are given by

$$x(\theta) = \frac{\theta}{2}, \text{ and } s(\theta) = \frac{\alpha \theta}{4\beta} (\theta + \mu + \varepsilon) \quad (15)$$

Note that the amount of time the student study alone, $x = \theta/2$, is independent on whether the student study with a classmate. When a student has the

chance to study with a classmate, she will add this amount of time to the hours she already decided to study alone. Then, having the chance of studying along with someone else increases the amount of time devoted to study. This is due to the fact that $x(\theta)$ and $s(\theta)$ are increasing functions.

Therefore, the symmetric Bayes-Nash equilibrium is

$$(x_1, y_1, x_2, y_2) = \left(\frac{\theta_1}{2}, \frac{\alpha\theta_1(\theta_1 + \mu + \varepsilon)}{4\beta}, \frac{\theta_2}{2}, \frac{\alpha\theta_2(\theta_2 + \mu + \varepsilon)}{4\beta} \right) \quad (16)$$

As was expected, expression (14) ensures that, in equilibrium, more qualified students devote more time to study alone and with a classmate.

6. Concluding remarks

Does a student improve her academic performance by studying with a classmate? Our analysis suggests that the answer may rely on the distribution of the academic skills in every classroom. Among the many results delivered by the analysis, we show that a student decides to study with a classmate because she believes that the classmate has better academics skills. Then, the time a student is willing to spend studying with a classmate increases with the mean and the variance of the distribution of the academic skills. We find interesting the result that a high-skilled student will spend longer periods of time studying alone.

There are several directions for future research. In the paper, the distribution of academic skills is exogenously given. In a more comprehensive model, students would realise the lost from being in a classroom with low mean and/or variance and, consequently, would take actions to modify the distribution of these factors. Such actions may include obtaining information to identify accurately the most talented students.

Appendix

Proof of Proposition 1. Let $(x(\theta), s(\theta))$ be a symmetric Bayes-Nash equilibrium. Given $s(\theta_j)$, the expected payoff for player i is

$$E[r_i(x_i, s_i, s(\theta_j)) | \theta_i] = \theta_i f(x_i) - c(x_i) + E[\alpha\theta_i\Theta_j f(s_i \wedge s(\theta_j)) - \beta c(s_i \wedge s(\theta_j)) | \theta_i].$$

Since s is strictly increasing, and so there exists the inverse s^{-1} , the expectation above splits into two intervals $[\mu - \varepsilon, \theta]$ and $[\theta, \mu + \varepsilon]$ for some θ . Given s_i , pick the point $\theta \in [\mu - \varepsilon, \mu + \varepsilon]$ such that $s_i = s(\theta)$. Thus $s(\theta_j) \leq s_i$ if and only if $\theta_j \leq s^{-1}(s_i)$, equivalently, $\theta_j \in [\mu - \varepsilon, s^{-1}(s_i)]$. That is

$$\begin{aligned} E[\alpha\theta_i\Theta_j f(s_i \wedge s(\Theta_j)) - \beta c(s_i \wedge s(\Theta_j)) | \theta_i] &= \int_{\mu-\varepsilon}^{s^{-1}(s_i)} [\alpha\theta_j f(s(\theta_j)) - \beta c(s(\theta_j))] \frac{d\theta_j}{2\varepsilon} \\ &\quad + \int_{s^{-1}(s_i)}^{\mu+\varepsilon} [\alpha\theta_j f(s_i) - \beta c(s_i)] \frac{d\theta_j}{2\varepsilon} \\ &= \int_{\mu-\varepsilon}^{s^{-1}(s_i)} [\alpha\theta_j f(s(\theta_j)) - \beta c(s(\theta_j))] \frac{d\theta_j}{2\varepsilon} \\ &\quad + \frac{\alpha\theta_i f(s_i)}{2\varepsilon} \left[\frac{(\mu + \varepsilon)^2}{2} - \frac{(s^{-1}(s_i))^2}{2} \right] \\ &\quad - \frac{\beta c(s_i)}{2\varepsilon} [\mu + \varepsilon - s^{-1}(s_i)] \end{aligned}$$

We now take the partial derivatives with respect to x_i and s_i

$$\begin{aligned} \frac{\partial E[r_i | \theta_i]}{\partial x_i}(x_i, s_i, s(\theta_j)) &= \theta_i f'(x_i) - c'(x_i), \\ \frac{\partial E[r_i | \theta_i]}{\partial s_i}(x_i, s_i, s(\theta_j)) &= \frac{1}{2\varepsilon} [\alpha\theta_i s^{-1}(s_i) f(s_i) - \beta c(s_i)] \frac{ds^{-1}}{ds_i} \\ &\quad + \frac{\alpha\theta_i}{4\varepsilon} \left[-2f(s_i) s^{-1}(s_i) \frac{ds^{-1}}{ds_i} + [(\mu + \varepsilon)^2 - (s^{-1}(s_i))^2] f'(s_i) \right] \\ &\quad + \frac{\beta}{2\varepsilon} \left[c(s_i) \frac{ds^{-1}}{ds_i} + (s^{-1}(s_i) - \mu - \varepsilon) c'(s_i) \right] \\ &= \frac{1}{2\varepsilon} \left[\frac{\alpha\theta_i f'(s_i)}{2} [(\mu + \varepsilon)^2 - (s^{-1}(s_i))^2] + \beta [s^{-1}(s_i) - \mu - \varepsilon] c'(s_i) \right]. \end{aligned}$$

The necessary conditions for $x(\theta_i)$ and $s(\theta_i)$ imply that

$$\frac{c'(x(\theta_i))}{f'(x(\theta_i))} = \theta_i \quad \text{and} \quad \frac{c'(s(\theta_i))}{f'(s(\theta_i))} = \frac{\alpha\theta_i}{2\beta} (\theta_i + \mu + \varepsilon).$$

Proof of Proposition 2. Consider expression (5)

$$c'(x(\theta_i)) = \theta_i f'(x(\theta_i))$$

Taking the derivative with respect to θ_i in both sides

$$\frac{\partial c'(x(\theta_i))}{\partial \theta_i} = \frac{\partial \theta_i f'(x(\theta_i))}{\partial \theta_i}$$

we have

$$c''(x(\theta_i)) x'(\theta_i) = \theta_i f''(x(\theta_i)) x'(\theta_i) + f'(x(\theta_i))$$

then

$$\left[\frac{c''(x(\theta_i))}{c'(x(\theta_i))} - \frac{f''(x(\theta_i))}{f'(x(\theta_i))} \right] x'(\theta_i) = \frac{1}{\theta_i}$$

note that $\frac{1}{\theta_i} > 0$ by Assumption 2. Thus Assumption 1 implies

$$\left[\frac{c''(x(\theta_i))}{c'(x(\theta_i))} - \frac{f''(x(\theta_i))}{f'(x(\theta_i))} \right] > 0$$

therefore

$$x'(\theta_i) > 0.$$

Proof of Proposition 3. Let us consider first s depending on the parameter μ , that is $s(\theta, \mu)$. Consider the necessary condition

$$\frac{c'(s(\theta, \mu))}{f'(s(\theta, \mu))} = \frac{\alpha\theta}{2\beta}(\theta + \mu + \varepsilon).$$

Taking the partial derivative with respect to μ in both sides we have

$$\frac{f'(s(\theta, \mu))c''(s(\theta, \mu))\frac{\partial s}{\partial \mu}(\theta, \mu) - c'(s(\theta, \mu))f''(s(\theta, \mu))\frac{\partial s}{\partial \mu}(\theta, \mu)}{[f']^2} = \frac{\alpha\theta}{2\beta},$$

equivalently

$$\left[\frac{c''(s(\theta, \mu))}{c'(s(\theta, \mu))} - \frac{f''(s(\theta, \mu))}{f'(s(\theta, \mu))} \right] \frac{\partial s}{\partial \mu}(\theta, \mu) = \frac{1}{\theta + \mu + \varepsilon}.$$

Then, assumptions 1 implies

$$\left[\frac{c''(s(\theta, \mu))}{c'(s(\theta, \mu))} - \frac{f''(s(\theta, \mu))}{f'(s(\theta, \mu))} \right] > 0,$$

and since

$$\frac{1}{\theta + \mu + \varepsilon} > 0,$$

we conclude

$$\partial s / \partial \mu > 0.$$

Analogously for the other parameters. \square

Proof of Proposition 4. Let us consider first s depending on the parameter μ , that is $s(\theta, \mu)$. Consider expression

we cannot use derivatives to show the monotonicity since $s(\theta_1, \mu) \wedge s(\theta_2, \mu)$ is not necessary differentiable even when $s(\cdot, \mu)$ is so.

We first show that $s(\theta_1, \mu) \wedge s(\theta_2, \mu)$ is increasing in μ for fixed θ_1 and θ_2 . Consider $\mu' \in \mathbb{R}$ such as $\mu' > \mu$. From Proposition 3 we have

$$s(\theta_1, \mu') > s(\theta_1, \mu)$$

$$s(\theta_2, \mu') > s(\theta_2, \mu)$$

then

$$s(\theta_1, \mu') \wedge s(\theta_2, \mu') > s(\theta_1, \mu) \wedge s(\theta_2, \mu).$$

Now multiplying by their joint probability density function $F(\theta_1, \theta_2)$ both sides of expression and integrating with respect to θ_1 and θ_2

$$\iint \{s(\theta_1, \mu') \wedge s(\theta_2, \mu')\} F(\theta_1, \theta_2) d\theta_1 d\theta_2 > \iint \{s(\theta_1, \mu) \wedge s(\theta_2, \mu)\} F(\theta_1, \theta_2) d\theta_1 d\theta_2.$$

We conclude that

$$E[s(\Theta_1, \alpha, \beta, \mu, \varepsilon) \wedge s(\Theta_2, \alpha, \beta, \mu, \varepsilon)] \text{ increases with } \mu.$$

Analogously for the other parameters. \square

References

- Banerjee, S.; H. Konishi & T. Snmez (2001). "Core in a simple coalition formation game". *Social Choice and Welfare*, 18 (1), 135-153.
- Barro, R. J. (2013). "Education and economic growth". *Annals of Economics and Finance*, 14(2), 301-328.
- Carman, K. G. & L. Zhang (2012). "Classroom peer effects and academic achievement: Evidence from a Chinese middle school". *China Economic Review*, 23 (2), 223-237.
- Corgnet, B. (2010). "Team Formation and Self-serving Biases". *Journal of Economics & Management Strategy*, 19 (1), 117-135.
- Crede, M. & N. Kuncel (2008). "Study Habits, Skills, and Attitudes: The Third Pillar Supporting Collegiate Academic Performance". *Perspectives on Psychological Science*, 3 (6), 425-453.
- De Fraja, G. & P. Landeras (2006). "Could do better: the effectiveness of incentives and competition in schools". *Journal of Public Economics*, 90 (1), 189-213.
- Dutta, P. K. (1999). *Strategies and games: theory and practice*. MIT press.
- Eisenkopf, G. (2010). "Peer effects, motivation, and learning". *Economics of Education Review*, 29 (3), 364-374.
- Falk, A. & A. Ichino (2006). "Clean evidence on peer effects". *Journal of Labor Economics*, 24 (1), 39-57.
- Foster, G. & P. Frijters (2010). "Students' beliefs about peer effects". *Economics Letters*, 108 (3), 260-263.
- Gibbons, R. (1992). *A primer in game theory*. Harvester Wheatsheaf.
- Hanushek, E. A. & L. Woessmann (2008). "The role of cognitive skills in economic development". *Journal of economic literature*, 46 (3), 607-668.
- Hollard, G. (2000). "On the existence of a pure strategy Nash equilibrium in group formation games". *Economics Letters*, 66 (3), 283-287.
- Islam, T.M.T. Minier, J. & J. P. Ziliak (2015). "On Persistent Poverty in a Rich Country". *Souther Economic Journal*, 81 (3), 653-678.
- Jehle, G. A. & P. J. Reny (2001) *Advanced Microeconomic Theory*. Addison Wesley Longman.
- Kiss, D. (2013). "The impact of peer achievement and peer heterogeneity on own achievement growth: Evidence from school transitions". *Economics of Education Review*, 37, 58-65.
- Lazear, E. P. (2001). "Educational Production". *Quarterly Journal of Economics*, 116 (3), 777-803.

- Mas-Colell, A.; M. D. Whinston & J. R. Green (1995). *Microeconomic theory*. New York: Oxford university press.
- Nonis, S. A. & G. I. Hudson (2006). "Academic Performance of College Students: Influence of Time Spent Studying and Working". *Journal of Education for Business*, 81 (3), 151-159.
- Osborne, M. J. & A. Rubinstein (1994). *A course in game theory*. MIT press.
- Robertson, D. & J. Symons (2003). "Do peer groups matter? Peer group versus schooling effects on academic attainment". *Economica*, 70, 31-53.
- Vardardottir, A. (2013). "Peer effects and academic achievement: a regression discontinuity approach". *Economics of Education Review*, 36, 108-121.