INSTITUTO DE ECONOMÍA

TESIS de DOCTORADO

2006

Perfectly Competitive Education with Reputational Concerns

Bernardita Vial.

www.economia.puc.cl

Perfectly Competitive Education with Reputational Concerns

por

Bernardita Vial

Licenciado en Ciencias Económicas y Administrativas, Pontificia Universidad Católica de Chile, 1996
Magíster en Economía, Pontificia Universidad Católica de Chile, 2000

Esta tesis se presenta como requerimiento parcial
para optar al grado de

Doctor en Economía

Instituto de Economía
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Comité:
Felipe Zurita (profesor guía)
David K. Levine
Rodrigo Harrison
José Miguel Sánchez
Federico Weinschelbaum

2006

La tesis de Bernardita Vial fue aprobada con distinción:

Felipe Zurita (profesor guía)

Rodrigo Harrison

David K. Levine

José Miguel Sánchez

Federico Weinschelbaum

Pontificia Universidad Católica de Chile

14 de diciembre de 2006

Abstract

Perfectly Competitive Education with Reputational Concerns

by

Bernardita Vial

Doctor en Economía, Pontificia Universidad Católica de Chile

Felipe Zurita (Chair)

This dissertation analyzes how schools choose and signal their quality in the presence of asymmetric information. Student beliefs about the quality offered by schools are updated in a Bayesian fashion according to test results. The present modeling of reputation departs from the existing literature in two ways: (1) the cost of signaling is related to students' types (ability), and (2) there is not a single producer, but many private schools that differ in their reputation and compete for better students. This results in an equilibrium with personalized prices, where schools with a better score history serve the more affluent and more able students. Within this equilibrium, the distribution of schools' reputations converges to a unique long run distribution, the fixed point of the dynamical system that describes the evolution of the distribution of schools' reputations.

Felipe Zurita
Chair

Resumen

Perfectly Competitive Education with Reputational Concerns

por

Bernardita Vial

Doctor en Economía, Pontificia Universidad Católica de Chile

Felipe Zurita (profesor guía)

En esta tesis se estudia cómo los colegios escogen y dan a conocer su calidad en presencia de asimetrías en la información. Los alumnos actualizan sus creencias en forma bayesiana, de acuerdo a los resultados de los colegios en una prueba estandarizada. La modelación de la reputación se diferencia de la literatura existente en dos aspectos: (1) el costo de señalizar depende de las características de los estudiantes, y (2) no hay sólo un productor, sino muchos colegios que compiten por los mejores alumnos. El resultado es un equilibrio en que los alumnos pagan precios personalizados, y en que los colegios de mejor reputación atienden a los alumnos de mayor ingreso y habilidad. En este equilibrio, la distribución de reputaciones de los colegios converge a una distribución única de largo plazo, el punto fijo del sistema dinámico que describe la evolución de la distribución de reputaciones.

Felipe Zurita
Profesor Guía

To my family.

Contents

List of Figures

Acknowledgments

I am very grateful to my advisor, Felipe Zurita, for his exceptional guidance and support. He was very generous with his time and knowledge. I would also like to thank the members of my committee, David K. Levine, Rodrigo Harrison, José Miguel Sánchez and Federico Weinschelbaum for providing very helpful suggestions and comments.

Chapter 1

Introduction

Although widely used by politicians, analysts and parents, the concept of "school quality" is difficult to describe and measure. When scholars ask for "school quality", they often think of the inputs and resources of the school, the curricula contents, and more frequently, the output of the schooling process. Some commonly used measures of output include cognitive achievement (e.g. as measured by standardized tests), completion rates or entrance rates to higher levels of education, and acquisition of some skills and forms of behavior. What all those measures of output have in common, is that they are affected not only by school attributes, but also by the characteristics of the students themselves.

There is a large empirical literature that tries to identify the school inputs and practices that enhance students' achievement, using standardized test scores as a measure of educational quality. But it is proving difficult to find any robust relationship between school resources or inputs and tests results (see Hanushek (1989)). Two recent studies that use panel data for the US identify a significant "teacher effect": following teachers across time and different groups of students, they find that some of them get systematically better results that others (Rivkin, Hanushek and Kain (2005); Rockoff (2004)). However, this teacher effect is difficult to relate to observable characteristics of the teachers, such as experience or education. Therefore, it seems that teachers differ in their quality, but that the determinants of those differences are difficult to identify from a statistical point of view.

There is also a large empirical literature that tries to identify what students' characteristics affect their achievement, including individual and group or peer characteristics. Regarding individual effects, there is a strong consensus about the importance of family background since the Coleman Report in 1966. Regarding peer effects, the literature is more recent, and there is no consensus about how strong peer effects are, and how do they work (see for example Angrist and Lang (2002), Hanushek, Kain, Markman and Rivkin (2003)).

In this dissertation we contribute to the existing theoretical literature in the economics of education by modeling the school market in the presence of incomplete information: students do not directly observe school quality or school type in advance. Instead, students observe past results in standardized tests, which act as imperfect (public) signals of school quality.

Two different models that capture the incomplete information assumption are considered. The first is a two-period signaling model, where the school type refers to its quality: type A schools always provide high quality education, whereas type B schools always provide low quality. As a result, students who attend type A schools obtain higher educational achievement than students who attend type B schools (for any given ability). The second is a reputation model with imperfect public monitoring. It considers an infinitely repeated game, where the school type refers to the *possibility* of providing high educational achievement to their students: type C schools are competent and can provide high educational achievement, whereas type I schools cannot. Providing high educational achievement (or high quality of education) is costly, and therefore type C schools will do so only if they expect a reward for their effort.

The possibility of signaling the school type in the first model, and of building a good reputation in the second model, arises from the assumption that the probability of obtaining a high test score increases with the student's educational achievement. Therefore, the cost associated to increasing the probability of obtaining a high score is lower for schools that provide higher educational achievement. But the cost of educational achievement also depends on student's ability, being lower for schools that serve higher ability students.[1] As a consequence, in both models the competition among private schools leads to tuition discounts to the more able students, and high ability students attend schools with higher results in equilibrium.

In Chapter 2 we describe the signaling model, and we analyze the assignment of students and the prices charged in a separating equilibrium. In this equilibrium, all schools charge the same price and admit the same kind of students in the first period, but high quality schools exert costly effort in order to signal their type. In the second period, those schools that chose high effort charge a higher price. The main result is that high ability students receive tuition discounts and attend schools with higher results on average in the first period in equilibrium. The consequences of imposing regulations that either restrict tuition discounts or students selection are also analyzed. We find that the only difference between the resultant separating equilibrium in this scenario and the separating equilibrium described earlier, is that schools are better off and high ability students are worse off with the regulation. That is, in this new scenario high ability students attend schools with

[1]To isolate the analysis from possible peer effects, we assume that each school serves only one student each period.

higher results on average, but they lose tuition discounts. The economic intuition of this result is that a regulation of this sort limits the degree of competition among schools: a school that serves a high ability student faces lower costs, but the other schools cannot attract this student by offering lower prices.

In Chapter 3 we develop a model of long run reputation with imperfect public monitoring. We focus on the "high quality equilibrium", an equilibrium where all competent schools make costly effort, and their students obtain high educational achievement.

There are two important results concerning the high quality equilibrium. The first refers to the characterization of equilibrium: students with higher income and/or ability go to better schools, where "better" schools are those who have a better reputation. In this model, those schools that are more selective (i.e. schools that receive higher ability students on average) are the ones with a better reputation of making effort to produce better educational achievement. But in the long run, inept schools have a worse reputation on average.[2] Therefore, more selective schools are also "more competent schools" in our model.

The second important result is more general, and is presented in Chapter 4. It refers to the evolution of reputations in games with imperfect monitoring in a competitive setting. A firm's reputation (i.e. the short run players' belief about the firm's type) is updated according to the realization of some public signal whose distribution function depends on the firm's action. Since different firms obtain different realizations of the public signal across time, firms generally differ in their reputation. An exogenous probability

[2] More precisely, the reputation for competent schools first-order stochastically dominates the reputation of inept schools. See section 3.

of replacement is introduced as in Mailath and Samuelson (2001) to sustain a long run reputation effect. Therefore, there is no long run convergence of beliefs: a bad realization of the public signal always has an important effect on a firm's reputation. As a consequence, schools' types are not even eventually (asymptotically) revealed. Our main result is that the distribution of firms' reputations converges to a unique long run distribution. This long run distribution is the fixed point of the dynamical system that describes the evolution of the distribution of schools' reputation. That is, although each school's reputation changes every period even in the long run, the distribution of reputations has a steady state.

Much of the recent theoretical literature in the economics of education has focused on equilibrium models for local public goods (see for example Nechyba (2000) and Hoxby (1999)), and on the effect of competition in the school market in the presence of peer effects. In Epple and Romano (1998), the only school attribute that affect students' results is the peer composition: students' educational achievement is positively related to the mean ability of the student body in the school attended, which is observable. Since students are willing to pay for good peers, lower ability/higher income students cross subsidize higher ability/lower income students, and they attend the same schools in equilibrium. All private schools are ex-ante identical, but ex-post there is a strict hierarchy of school qualities. Moreover, those schools that serve more affluent and more able students obtain better results. The define "stratification by income" as the prediction that more affluent students attend schools with better results in equilibrium (holding ability fixed), and "stratification by ability" analogously.

Epple, Figlio and Romano (2004) test the central predictions of the Epple and

Romano model. They find evidence of stratification by income and ability, both between the public and private sector, and within the private sector. That is, the probability of attending a private school instead of a public one increases with income, and also the probability of attending a higher-quality school within the private sector (and similarly with ability). They also find evidence of discounting to ability in the private sector: higher ability students pay lower tuitions.

In the Epple and Romano model the stratification by income and ability and the strict hierarchy of school qualities arise solely from the "peer effect". In contrast, in our model the stratification arises from a "reputational effect". This distinction is important, because in Epple and Romano's model a selective school has no merit in itself, while in ours more selective schools are ex-ante different: they are better, in the sense of having a higher probability of being competent. Therefore, the consequence of a reform that harms or favors the operation of more selective schools in the scenario described by Epple and Romano differs markedly from ours. Furthermore, some policy recommendations that follow from Epple and Romano's model would have negative consequences in the presence of "reputational effects". For instance, Epple and Romano (2002) propose a voucher that decreases with student ability and with no extra charges allowed, the idea being achieving the benefits of a voucher program without "cream skimming". But to obtain a high quality equilibrium, a necessary condition is that better reputation schools are allowed to charge higher fees.

Epple and Romano's model predicts that the equilibrium assignment of students produces stratification by income and ability, but it is not clear how this equilibrium assignment is reached. This is because there is no element in the model that predicts which schools will receive better students and obtain better results in equilibrium, since the reason why those schools obtain better results is just because the ability of their students is higher (it all boils down to a coordination problem among students).

De Fraja and Landeras (2006) assume that the students' effort and the quality of the schools' teaching are determinants of student educational achievement, in addition to peer composition. Therefore, schools differ not only in their peer composition as in Epple and Romano, but also in their choice of teaching quality. They also assume that school past performance affects the school's reputation and school peer composition, since better test results attract more able students. A prediction of their model is that competition produces stratification by ability. Therefore, they find a similar prediction to Epple and Romano's, but in their model there is a reputation-based explanation for the equilibrium assignment of students. However, they don't address the way in which school reputation is created and maintained.

The models we present in this dissertation have some elements in common with the previous literature in education, but they also incorporate elements of both General Equilibrium and Game theories. We explicitly model the way in which school reputation is created and maintained in a competitive school market with asymmetric information, where students' beliefs about the quality offered by schools are updated in a Bayesian fashion according to test results. The present modeling of reputation departs from the existing literature in two ways: (1) the cost of signaling is related to students' types (ability), and (2) there is not a single producer, but many private schools that differ in their reputation and compete for better students. This results in an equilibrium with personalized prices and stratification by income and ability: schools with a better score history serve the more affluent and more able students.

CHAPTER 2

Test scores and school quality: a signaling model

In this Chapter we present a signaling model in a two-period competitive school market. The model considers students (one -period players) characterized by their ability (b), and schools (two-period players) characterized by their type. To isolate the analysis from possible peer effects each school is assumed to serve only one student each period.

Students' utility function is increasing in the unique consumption good, z, and the student's educational achievement, a. In turn, educational achievement is an increasing function of the student's ability and the school's quality, q. There are two types of schools: high quality schools (A), which provide q_A, and low quality schools (B), which provide q_B, with $q_B < q_A$. School quality is neither observable nor contractible.

In the second period, students observe the results of a standardized test taken by all schools in the previous period. Test results are deterministically related to student's achievement and school effort. Educational achievement and school effort both contribute to increase test results. Therefore, the effort (and cost) required to produce a given result r is decreasing on b, which is observable for the schools. Furthermore, the cost is lower for A than for B schools given the student's ability. Hence, high quality schools can separate themselves from low quality schools by using test scores as a signal of their type.

We analyze the assignment of students and the prices charged in a separating equilibrium. Under this equilibrium, all schools charge the same price and admit the same students in the first period, but high quality schools exert costly effort in order to signal their type. In the second period, those schools that chose high effort charge a higher price. The main result of this Chapter is that high ability students receive tuition discounts and attend schools with better results on average in the first-period equilibrium.

We also analyze the consequences of restricting tuition discounts or selection of students. We find that the only difference between the resultant separating equilibrium in this scenario and the separating equilibrium described earlier, is that schools are better off and high ability students are worse off if tuition discounts are not allowed. That is, we still find that under this new scenario high-ability students attend schools with higher results on average, but they loose tuition discounts.

2.1 The Model

2.1.1 Students

There is a continuum of students characterized by their ability $b \in \left[\underline{b}, \overline{b}\right]$. Each student observes his own ability, but not the ability of other students. The cumulative distribution function (henceforth, cdf) of abilities is denoted F_b, continuous and constant over time.

Students' utility function is increasing in the unique consumption good, z, and the educational achievement, a. In turn, educational achievement depends on b and school quality, $q \in \{q_A, q_B\}$. Students do not observe school quality in advance: educational achievement is an "experience good". In the first period ($t = 0$), they only observe the prices posted and the admission policy of each school. In the second period they also observe how each school did in a standardized test taken at $t = 0$.

Students hold homogeneous beliefs regarding how likely it is that a given school is of type is A, μ. Then, μ is a school's characteristic. We will abuse notation and refer to a school with reputation μ as to a "school μ". Student's expected utility of attending a school μ is therefore:

$$E[u] = \mu u(z, a(b, q_A)) + (1-\mu) u(z, a(b, q_B)), \tag{2.1}$$

where $u(z,a)$ is increasing in z and a, and $a(b,q)$ is increasing in b and q. We assume that u and a are differentiable. We also assume that the expected utility is zero if the student does not attend any (public or private) school.

Students can always choose either a public, low quality school, or a private school μ. The public school is free. Therefore, a student with income y and ability b will choose the public school if and only if:

$$u(y, a(b, q_B)) \geq \mu u(y - p(\mu, b), a(b, q_A)) + (1-\mu) u(y - p(\mu, b), a(b, q_B)) \tag{2.2}$$

for all available schools, where $p(\mu, b)$ is the price charged by the private school μ to a student with ability b.

We define the reservation price p_R for a school μ as the price that solves:

$$u(y, a(b, q_B)) = \mu u(y - p_R, a(b, q_A)) + (1-\mu) u(y - p_R, a(b, q_B)). \tag{2.3}$$

It follows from (2.3) that $p_R = 0$ if $\mu = 0$, and the reservation price is increasing on μ:

$$\frac{\partial p_R}{\partial \mu} = \frac{u(z, a(b, q_A)) - u(z, a(b, q_B))}{\mu u_z(z, a(b, q_A)) + (1-\mu) u_z(z, a(b, q_B))} > 0, \tag{2.4}$$

where $\frac{\partial u(z,a)}{\partial z} \equiv u_z(z,a)$ and $\frac{\partial u(z,a)}{\partial a} \equiv u_a(z,a)$.

In order to isolate the analysis of the equilibrium assignment of students from possible differences in the willingness to pay among students with different ability level, we assume that preferences satisfy:

$$\frac{\partial p_R}{\partial b} = \frac{\mu u_a\left(y - p_R\left(\mu\right), a\left(b, q_A\right)\right) \frac{\partial a(b,q_A)}{\partial b}}{\mu u_z\left(z, a\left(b, q_A\right)\right) + \left(1-\mu\right) u_z\left(z, a\left(b, q_B\right)\right)} + \frac{\left(\left(1-\mu\right) u_a\left(y - p_R\left(\mu\right), a\left(b, q_B\right)\right) - u_a\left(y, a\left(b, q_B\right)\right)\right) \frac{\partial a(b,q_B)}{\partial b}}{\mu u_z\left(z, a\left(b, q_A\right)\right) + \left(1-\mu\right) u_z\left(z, a\left(b, q_B\right)\right)} = 0. \quad (2.5)$$

Therefore, reservation prices are independent of b. We will denote the reservation price by $p_R\left(\mu\right)$.

Example 2.1 *If $u = za$ and $a = q^{\gamma} b^{\beta}$, the reservation price p_R is:*

$$p_R = \mu y \frac{q_A^{\gamma} - q_B^{\gamma}}{\mu q_A^{\gamma} + \left(1-\mu\right) q_B^{\gamma}}. \quad (2.6)$$

2.1.2 Schools

There is a continuum of schools. Each school serves only one student each period. Schools observe students characteristics, and they can choose which student to serve.

There are two types of schools, A and B, which differ in the in their quality: $q_A > q_B$. The fraction of type A schools is λ. Students do not observe school quality, but at $t = 1$ they observe past results of a standardized test, $r \geq 0$. The test result is an increasing function on the educational achievement and the school effort. Therefore, the effort (and cost) required to produce a given result r at time t is decreasing on the ability level of the student served at t, and is lower for type-A than for type-B schools. We denote the cost associated to a result r and a student with ability b in a type τ school as $c_{\tau}\left(r, b\right)$.

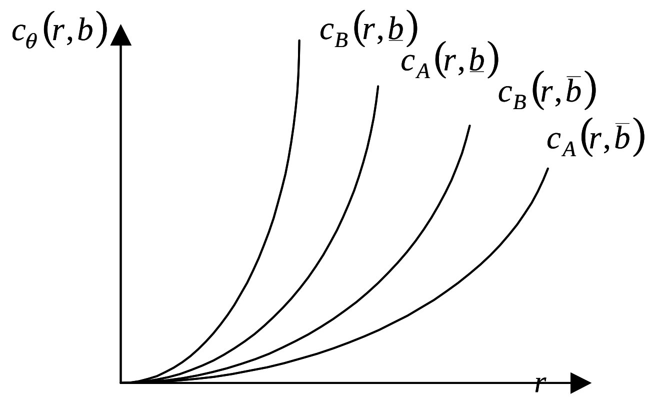

Figure 2.1: The cost function $c_\theta(r,b)$

We assume that c_τ is an unbounded function, continuous and differentiable, and satisfies:

$$c_A(0,b) = c_B(0,b) = 0 \text{ for all } b \in \left[\underline{b},\bar{b}\right], \tag{2.7}$$

$$c_A(r,b) < c_B(r,b) \text{ for all } (r,b) \in (0,\infty) \times \left[\underline{b},\bar{b}\right], \tag{2.8}$$

$$\frac{\partial c_B(r,b)}{\partial r} > \frac{\partial c_A(r,b)}{\partial r} > 0 \text{ for all } b \in \left[\underline{b},\bar{b}\right], \text{ and} \tag{2.9}$$

$$\frac{\partial c_\tau(r,b)}{\partial b} < 0 \text{ for } \tau \in \{A,B\} \tag{2.10}$$

as depicts Figure 2.1 for $b = \underline{b}$ and $b = \bar{b}$.

Each schools takes prices as given, and chooses an admission policy and a test result so as to maximize its expected profit, that is, the expected present value of earnings, with discount factor δ. Each period students observe schools' prices and admission policies, as well as their past results in the standardized test at $t = 1$.[1] With this information they choose a school among all the schools that had admitted them. Figure 2.2 shows the timeline of the game. If low quality schools charge different prices or choose a different

[1]The current student generation does not observe the prices charged or the ability of students admitted by each school in the previous period.

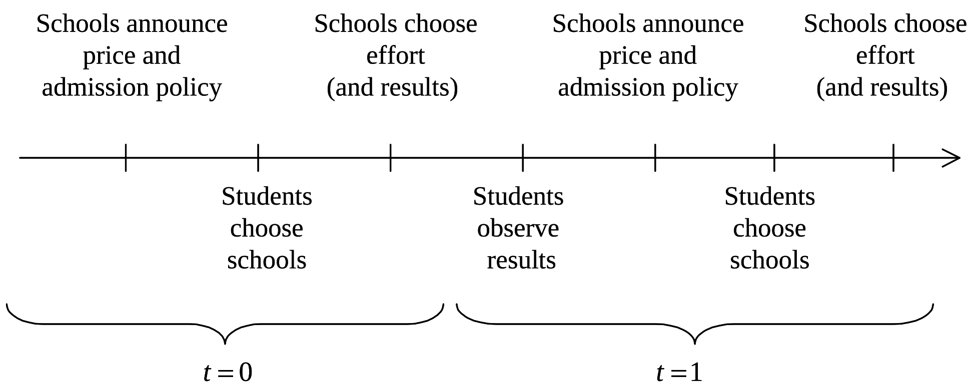

Figure 2.2: The timeline of the game

pool of students, they reveal their type. Hence, low quality schools follow high quality schools in the price charged and the students chosen.

2.1.3 Equilibrium

We focus on perfectly competitive perfect bayesian equilibrium in pure strategies that satisfy Cho and Krep's intuitive criterion.

Since there are only two periods, the result chosen at $t = 1$ by schools A and B is $r = 0$. We denote the result chosen by a type τ school with a student with ability b at $t = 0$ as $r_\tau(b)$. We focus in equilibria with two levels of results: r^* and 0.

We denote the price charged at $t = 0$ to a student with ability b by $p^0(b)$, and likewise by $p^1(r, b)$ the price charged at $t = 1$ by a school that chose a result r to a student with ability b.

A competitive (either fully or partially) separating equilibrium is therefore a belief μ, a signal r^*, a price function and an assignment of students for periods $t = 0$ and $t = 1$

that satisfies the following conditions:

Competitive separating equilibrium 1 (CSE1) $r_\tau(b)$ is optimal for the schools.

Competitive separating equilibrium 2 (CSE2) Correct initial beliefs and Bayesian updating. That is, if $\mu(h)$ denote the posterior belief after a history h, then $\mu(\phi) = \lambda$ and $\mu(r)$ is obtained using the Bayes rule according to the equilibrium behavior of schools.

Competitive separating equilibrium 3 (CSE3) The assignment of students is optimal for the students. That is, no student wants to attend a school different from the one he attends under the equilibrium assignment.

Competitive separating equilibrium 4 (CSE4) The assignment of students is optimal for type A schools. That is, no school wants to attract a student different from those who attend this school under the equilibrium assignment.

Competitive separating equilibrium 5 (CSE5) Market clearing.

Competitive separating equilibrium 6 (CSE6) Non-negative expected profits for schools of either type.

From conditions CSE4 and CSE5 we obtain the following initial characterization of the equilibrium price function.

Lemma 2.1 *At $t = 1$ the prices charged do not depend on b.*

Proof. *Schools A must be indifferent among all students in order to meet condition CSE5. Therefore, since no school chooses $r > 0$ at $t = 1$, the price charged must be independent of b.* ■

From now on, we refer to the price charged at $t = 1$ as $p^1(r)$.

From condition CSE1 we obtain the following characterization of the signal function $r_\tau(b)$.

Lemma 2.2 *If a school of type τ that serves a student with ability b' chose r^*, then all type τ schools that serve students with $b > b'$ also choose r^*.*

Proof. *A school of type τ that serves a student with ability b chooses r^* if and only if*

$$p^0(b) - c_\tau(r, b) + \delta p^1(r^*) \geq p^0(b) + \delta p^1(0) \tag{2.11}$$

$$\Leftrightarrow \delta\left(p^1(r^*) - p^1(0)\right) \geq c_\tau(r, b). \tag{2.12}$$

In view of the fact that $\frac{\partial c_\tau(r,b)}{\partial b} < 0$, if condition (2.12) is satisfied for b', it must also be satisfied for all $b > b'$. ■

Lemma 2.3 *The price charged at $t = 0$ must be non-increasing in b.*

Proof. *Schools A must be indifferent among all students in order to meet condition CSE5. Therefore, as the cost of the signal is decreasing in b, $p^0(b)$ must also be decreasing in b for those students who attend schools that choose r^*.* ■

It can be shown that any equilibrium with some schools B choosing r^* do not satisfy the Cho-Kreps Intuitive criterion.[2] Therefore, in what follows we consider equilibria where all type B schools choose $r = 0$.

If type A schools that serve students with $b \geq b'$ choose r^*, and since no type B

[2] See Appendix A.

school chooses r^*, Bayesian updating implies that equilibrium beliefs are:

$$\mu(r) = \begin{cases} 1 & \text{if } r = r^*, \\ \frac{F_b(b')\lambda}{F_b(b')\lambda + (1-\lambda)} & \text{if } r = 0. \end{cases} \tag{2.13}$$

Hence, a school that chooses r^* cannot charge a higher price than $p_R(1)$, and a school that chooses $r = 0$ cannot charge a higher price than $p_R(\mu(0)) = p_R\left(\frac{F_b(b')\lambda}{F_b(b')\lambda + (1-\lambda)}\right)$ at $t = 1$. We assume that the prices charged at $t = 1$ are:

$$p^1(r) = \begin{cases} p_R(1) & \text{if } r = r^*, \\ p_R(\mu(0)) & \text{if } r = 0. \end{cases} \tag{2.14}$$

Therefore $p^1(0)$ is increasing in b'. As an example, if all type A schools choose r^* (that is, if $b' = \underline{b}$), then $p^1(0) = p_R(0) = 0$; on the other hand, if no school chooses r^* (that is, if $b' = \bar{b}$), then $p^1(0) = p_R(\lambda)$. To emphasize the dependence on b', we denote the price charged at $t = 1$ by schools that choose $r = 0$ as $p^1(0; b')$.

Let us define $r_B(b')$ as the value of r that solves:

$$\delta\left(p^1(r) - p^1\left(0; b'\right)\right) = c_B\left(r, \bar{b}\right) \tag{2.15}$$

for each b', as Figure 2.3 shows for $b' = \underline{b}$ and $b' = \bar{b}$.

Define $r_A(b')$ as the value of r that solves

$$\delta\left(p^1(r) - p^1\left(0; b'\right)\right) = c_A\left(r, b'\right) \tag{2.16}$$

for each b', as Figure 2.4 shows for $b' = \underline{b}$ and $b' = \bar{b}$.

Now we can prove the existence of a separating equilibrium.

Theorem 2.1 *There are ability levels* $b' \in \left[\underline{b}, \bar{b}\right]$ *that satisfy:*

$$r_A\left(b'\right) \geq r_B\left(b'\right), \tag{2.17}$$

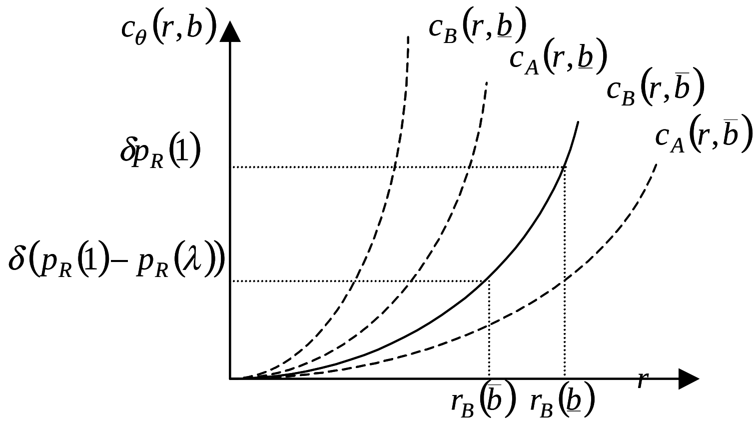

Figure 2.3: The level of $r_B(b')$

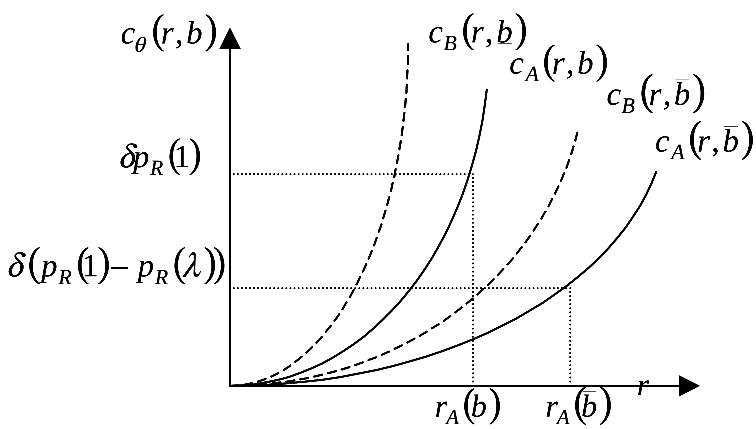

Figure 2.4: The level of $r_A(b')$

and can be part of a competitive (fully or partially) separating equilibrium where all type A schools that serve students with $b \geq b'$ at $t = 0$ choose $r^* = r_A(b')$, while the other schools choose $r = 0$.

Proof. *Since $r^* = r_A(b')$, we know that:*

$$\delta\left(p^1(r^*) - p^1\left(0; b'\right)\right) \geq c_A(r^*, b) \quad \text{for all } b \geq b'. \tag{2.18}$$

Therefore, all type A schools that serve students with $b \geq b'$ choose r^*. Since $r^* \geq r_B(b')$, we know that:

$$\delta\left(p^1(r^*) - p^1\left(0; b'\right)\right) \leq c_B\left(r^*, \overline{b}\right) \tag{2.19}$$

and therefore,

$$\delta\left(p^1(r^*) - p^1\left(0; b'\right)\right) < c_B(r^*, b) \quad \text{for all } b < \overline{b}. \tag{2.20}$$

Hence, all type B schools choose $r = 0$.

We know that $r_A\left(\overline{b}\right) > r_B\left(\overline{b}\right)$. Continuity of F_b implies that $p^1(0; b')$ is continuous, and therefore, that $r_A(b') - r_B(b')$ is continuous as well. Hence, we can find $b^* < \overline{b}$ such that $r_A(b') \geq r_B(b')$ for all $b' \in \left[b^*, \overline{b}\right]$.

Consider an ability level b' that satisfies condition 2.17, and $r^* = r_A(b')$. From condition CSE2, the equilibrium students' beliefs are $\mu(\phi) = \lambda$ and $\mu(r)$ as in (2.13). From conditions CSE3 and CSE4, the price charged at $t = 0$ must satisfy:

$$p^0(b) = \begin{cases} p^0(b') \leq p_R(\lambda) & \text{for all } b < b', \\ p^0(b') - c_A(r^*, b') + c_A(r^*, b) & \text{for all } b > b'. \end{cases} \tag{2.21}$$

The assignment of students is random, since all students are indifferent among schools and all type A schools are indifferent among students. Therefore, condition CSE5 is met.

Assume that $p^0(b') = p_R(\lambda)$. *Type A schools profits are therefore:*

$$p_R(\lambda) + \delta p^1\left(0; b'\right) = p_R(\lambda) - c_A\left(r^*, b'\right) + \delta p^1(r^*), \tag{2.22}$$

which are strictly positive.

In turn, type B schools' profits are:

$$p^0(b) + \delta p^1\left(0; b'\right) = \begin{cases} p_R(\lambda) + \delta p^1(0; b') & \text{for all } b < b', \\ p_R(\lambda) - c_A(r^*, b') + c_A(r^*, b) + \delta p^1(0; b') & \text{for all } b > b'. \end{cases} \tag{2.23}$$

Hence, the expected gain for type B schools depends on b'. *If* $b' = \bar{b}$, *the expected gain is* $(1+\delta)\, p_R(\lambda)$, *strictly positive. Continuity of* F_b *implies that* $p^1(0; b')$ *is continuous, and therefore, that the expected gain for schools B is continuous as well. Hence, we can find* $b^{**} < \bar{b}$ *such that the expected gain for schools B is positive for all* $b \in \left[b^{**}, \bar{b}\right]$. *Therefore, there are ability levels* $b' \in \left[\underline{b}, \bar{b}\right)$ *such that* $r^* = r_A(b')$, *and the assignment of students and the prices satisfy conditions CSE1 to CSE6.* ■

Under any equilibrium such as the equilibria described in Theorem (2.1), higher ability students receive tuition discounts, and they attend the schools with better results at $t = 0$, that is, the schools that choose $r^* > 0$. Since test results are the only observable school characteristic, and only high quality schools obtain high test results under those equilibria, this result may be interpreted as "stratification by ability". But as the assignment of students is random, this stratification has no consequence over the expected educational achievement of the students: the expected educational achievement differ among students with different ability, but only due to this difference in his/her own characteristics, and not due to differences among the schools they attend in equilibrium.

2.1.4 The consequence of restricting tuition discounts and selection of students

Now assume that schools cannot charge different prices to higher ability students, nor can they select them, either because schools do not observe students' ability or due to some regulation. If this is the case, the assignment of students at $t = 0$ must be random in equilibrium, and schools cannot charge a price higher that $p_R(\lambda)$ at $t = 0$. A random assignment of students at $t = 0$ implies that the equilibria described in Theorem (2.1) are still equilibria in the present scenario, except for the tuition discounts given to higher ability students. Theorem (2.2) characterizes the resulting separating equilibria:

Theorem 2.2 *If schools cannot charge different prices to higher ability students or select them, there are ability levels* $b' \in \left[\underline{b}, \overline{b}\right]$ *that satisfy:*

$$r_A\left(b'\right) \geq r_B\left(b'\right) \tag{2.24}$$

and can be part of a competitive (fully or partially) separating equilibrium where all type A schools that serve students with $b \geq b'$ *at* $t = 0$ *choose* $r^* = r_A(b')$*, while the other schools choose* $r = 0$*.*

Equilibrium students' beliefs are $\mu(\phi) = \lambda$ *and* $\mu(r)$ *is as in (2.13). The prices charged at* $t = 0$ *satisfy:*

$$p^0 \leq p_R(\lambda) \quad \text{for all } b \in \left[\underline{b}, \overline{b}\right]. \tag{2.25}$$

The assignment of students is random, since all students are indifferent among schools and all type A schools are indifferent among students.

Proof. *Assume that* $p^0(b') = p_R(\lambda)$. *Type A schools' profits are therefore:*

$$p_R(\lambda) + \delta p^1(0; b') \quad \text{for all } b < b', \text{ and} \tag{2.26}$$

$$p_R(\lambda) - c_A(r^*, b) + \delta p^1(r^*) \quad \text{for all } b > b'. \tag{2.27}$$

But the signal r^* *satisfies:*

$$\delta\left(p^1(r^*) - p^1(0; b')\right) = c_A(r^*, b'), \tag{2.28}$$

and therefore,

$$\begin{aligned} p_R(\lambda) - c_A(r^*, b) + \delta p^1(r^*) &> p_R(\lambda) - c_A(r^*, b') + \delta p^1(r^*) \\ &= p_R(\lambda) + \delta p^1(0; b') \quad \text{for all } b > b'. \end{aligned}$$

Hence, type A schools earn strictly positive profits. ■

Type B schools' profits are:

$$p_R(\lambda) + \delta p^1(0; b') \quad \text{for all } b \in [\underline{b}, \overline{b}], \tag{2.29}$$

also strictly positive.

As a result, if we compare two separating equilibria with the same $r^* = r_A(b')$, one with tuition discounts (as in Theorem (2.1)) and the other without tuition discounts (as in Theorem (2.2)), we find that the only difference between them is that both type A and B schools earn less in the former case. Hence, if schools cannot charge different prices to higher ability students or select them, the equilibrium assignment of students is not modified, and we still find that higher ability students attend schools with better results at $t = 0$, but as the competition among type A schools is restricted, students loose their tuition discounts (and are worse off).

Appendix A

Partially Separating Equilibrium

In this Appendix we prove that any equilibrium with some type B schools choosing r^* does not satisfy the Cho-Kreps Intuitive criteria.

Assume that type B schools that serve students with ability $b > b'$ choose r^* at $t = 0$. Then, schools A that serve students with ability $b > b'$ also choose r^*, since they face a lower cost and the same benefit as type B schools for choosing r^*.

Since type A schools are indifferent among all the students in equilibrium, the profits for schools A with the signal r^* are:

$$\pi_A^* = p^0(b) - c_A(b, r^*) + \delta p^1(r^*) \quad \text{for all } b \in \left[\underline{b}, \overline{b}\right]. \tag{A.1}$$

The profits for type B schools are:

$$\pi_B^*(b) = \begin{cases} p^0(b) - c_B(b, r^*) + \delta p^1(r^*) & \text{if } b > b', \\ p^0(b) + \delta p^1(0) & \text{if } b < b'. \end{cases} \tag{A.2}$$

Now, choose r^{**} such that $c_B(b'', r^{**}) - c_B(b'', r^*) = \delta\left(p^1(1) - p^1(r^*)\right)$ for some $b'' \geq b'$, and $c_B(b, r^{**}) - c_B(b, r^*) \geq \delta\left(p^1(1) - p^1(r^*)\right)$ for all $b \geq b'$.[1] That is,

$$p^0\left(b''\right) - c_B\left(b'', r^*\right) + \delta p^1(r^*) = p^0\left(b''\right) - c_B\left(b'', r^{**}\right) + \delta p^1(1), \text{ and} \tag{A.3}$$

$$p^0(b) - c_B(b, r^*) + \delta p^1(r^*) \geq p^0(b) - c_B(b, r^{**}) + \delta p^1(1) \quad \text{for all } b \geq b'. \tag{A.4}$$

[1] We can always find a level r^{**} that satisfies this condition, since $c_B(b, r^{**}) - c_B(b, r^*)$ is continuous, $c_B(b, r^{**}) - c_B(b, r^*) = 0$ for all b if $r^{**} = r^*$ and $\lim_{r^{**} \to \infty} \left(c_B(b, r^{**}) - c_B(b, r^*)\right) = \infty$ for all b.

For students with $b < b'$ we obtain:

$$p^0(b) + \delta p^1(0, b') = p^0(b) - c_B(b', r^*) + \delta p^1(r^*) \quad \text{(A.5)}$$

$$\geq p^0(b) - c_B(b', r^{**}) + \delta p^1(1) \quad \text{(A.6)}$$

$$> p^0(b) - c_B(b, r^{**}) + \delta p^1(1). \quad \text{(A.7)}$$

and therefore, no type B school would choose r^{**}, even if the students believe that a school that chooses r^{**} is of type A (i.e., even if they pay $p^1(1)$ at $t = 1$ to any school that chooses r^{**}).

But from Condition (2.9), we know that $c_B(b'', r^{**}) - c_B(b'', r^*) > c_A(b'', r^{**}) - c_A(b'', r^*)$. Therefore, this equilibrium does not satisfy the Cho-Kreps Intuitive Criterion, since a type A school that serves a student with ability b'' would choose r^{**} if the students believed that a school that chooses r^{**} is of type A.

Chapter 3

Competitive equilibrium in a reputation model with imperfect public monitoring

In this Chapter we present a reputation model with imperfect public monitoring in an otherwise perfectly competitive school market. The model considers a continuum of students (short run players) that differ in their income ($y \in Y$) and ability ($b \in B$), and a continuum of schools (long run players) characterized by their type (competent, C, or inept, I) and their reputation, μ. To isolate the analysis from possible peer effects, we assume that each school serves only one student each period.

We extend the model of reputation with imperfect public monitoring and exogenous replacements by Mailath and Samuelson (2001) introducing a continuum of producers. In their model, there is a unique firm and a continuum of consumers who repeatedly purchase an experience good from the firm. The utility level associated to the consumption of this good is a random variable: under a good outcome, the utility level is 1, and under a bad outcome, the utility level is 0. All consumers receive the same (public) realization of utility outcomes. There are two possible types of firms: competent and inept. A competent firm can choose high (costly) effort, which increases the probability of a good outcome, but inept firms always exert low effort. There is an exogenous probability λ that the firm exits the market, and an exogenous probability θ that a competent firm replaces the exiting firm. In the "high effort equilibrium", the competent firm always exerts high effort.

Hörner (2002) develops a model of reputation in a competitive setting. As Mailath and Samuelson (2001) do, he assumes that consumers experience a utility outcome stochastically related to the firm effort: the probability of a good outcome increases with effort. An important difference with our model, is that this utility outcome is not public, since all the clients of a given firm experience the same outcome, but this outcome is not observed by other consumers. Moreover, he focuses on nonrevealing high-effort equilibria, where all operating firms choose identical prices in each period on the equilibrium path. There are no capacity constraints, and consumers can switch from one firm to another without cost. As a result, if the customers of a given firm experience a bad outcome, they immediately switch to another firm, and the only surviving firms in a given period are those whose customers had never experienced a bad outcome. Therefore, the distribution of firms' reputations is degenerate: all operating firms in each period share the same reputation.

We assume that students value educational achievement, where $a = 0$ is referred to as low achievement, and $a = 1$ as high achievement. The schools provide educational achievement, but it is not observable in advance, nor contractible. Competent schools choose the educational achievement provided to their students by exerting effort. Neither effort nor educational achievement of former students are observable. But past results in a standardized test are observable, and the probability of obtaining a good result is higher when the student's educational achievement is high. The school's reputation is synthesized by the probability students assign to the schools being competent. We consider the infinitely repeated game, and examine a "high quality equilibrium", that is, an equilibrium where all competent schools provide a high educational achievement to their students. Within this equilibrium, the students' beliefs about the educational achievement provided by the school (or school type) are updated according to the school's past test scores.

The cost of providing high achievement is decreasing in b: the effort level necessary to reach $a = 1$ with a higher ability student is smaller. We denote the cost of high achievement for school C with a student with ability b as $c(b)$. Without effort (at zero cost), the educational achievement is 0 (independent of b). Competent schools also choose the students they serve. Inept schools cannot choose the educational achievement provided to their students: the educational achievement provided by inept schools is 0. If inept schools charge different prices or choose a different pool of students, they reveal their type. Hence, we will focus on pooling equilibria, where inept schools follow competent schools in the students chosen and prices charged.

We assume that there is a probability λ that the school leaves the market and is replaced by another school, which is of type C with probability θ; as these changes are unobservable to the students, the new school inherits the reputation of the school it is replacing. As a consequence, the school's reputation is continuously modified. Yet in the next Chapter we prove that there is an invariant, non-degenerate long run distribution of schools' reputations: each school's reputation changes every period even in the long run, but the population distribution of reputations remains constant. We consider the long run distribution of schools' reputation in the analysis of the present Chapter. An important characteristic of this long run distribution is that, when we consider the distribution of schools' reputation for competent and inept schools separately, the former first-order stochastically dominates the latter. That is, the probability that a school has a "good" reputation (i.e. a reputation better than any arbitrary level x) is lower for inept than for competent schools.

The main result of this Chapter is that in the high quality equilibrium, high ability students receive tuition discounts, and the assignment of students in private schools is increasing in income and ability: higher income/ability students attend schools with better reputation. This result is similar to the stratification by income and ability found by Epple and Romano (1998).

3.1 Assignment of students in the high quality equilibrium: a simple example

Assume that there are N students that attend N competent schools providing $a = 1$. We index students by i, and denote income and ability level of student i as y^i and b^i respectively. Students differ in their income and/or ability levels: the pair $\left(y^i, b^i\right)$ is different for all students, but for each level of income there are at least two students with different ability, and for each level of ability there are at least two students with different income (thus, if $N = 4$ there are two different levels of income and two different levels of ability).

For a concrete example, consider the following utility function:

$$u = z\left(a + 1\right), \tag{3.1}$$

where z is a consumption good and a is the educational achievement obtained in the school attended. In the high quality equilibrium, μ is the probability students assign to obtaining

$a = 1$ in a particular school. Therefore, the expected utility associated to attending a school with reputation μ that charges price p is:

$$E[u] = (y - p)(\mu + 1). \quad (3.2)$$

Schools differ in their reputations. Let j index schools, where a higher j index is assigned to a school with a higher reputation. All competent schools choose high effort. The one-period profit for a competent school j that serves a student with ability b^i and charges a price p_j^i to such student is therefore:

$$p_j^i - c\left(b^i\right). \quad (3.3)$$

The continuation payoff for the school is independent of the student chosen in the present period.

We say that an assignment of students is increasing in y and b if higher income students attend a school with higher reputation than lower income students (holding b constant), and if higher ability students attend a school with higher reputation than lower ability students (holding y constant).

Lemma 3.1 *The equilibrium assignment of students is increasing in income and ability.*

Proof. *i) Assume that student i attends school j and student i' attends school j' in equilibrium, with $y^{i'} > y^i$, $b^{i'} = b^i$ and $j' < j$. As the assignment must be optimal for* *students and schools, we obtain:*

$$\left(y^{i}-p_{j}^{i}\right)\left(\mu_{j}+1\right) \geq \left(y^{i}-p_{j'}^{i}\right)\left(\mu_{j'}+1\right), \tag{3.4}$$

$$\left(y^{i'}-p_{j'}^{i'}\right)\left(\mu_{j'}+1\right) \geq \left(y^{i'}-p_{j}^{i'}\right)\left(\mu_{j}+1\right), \tag{3.5}$$

$$p_{j}^{i} \geq p_{j}^{i'}, \text{ and} \tag{3.6}$$

$$p_{j'}^{i'} \geq p_{j'}^{i}. \tag{3.7}$$

But from (3.6) and (3.7) we get:

$$\left(y^{i'}-p_{j}^{i'}\right)-\left(y^{i}-p_{j}^{i}\right) \geq \left(y^{i'}-y^{i}\right) \geq \left(y^{i'}-p_{j'}^{i'}\right)-\left(y^{i}-p_{j'}^{i}\right)$$
$$\Rightarrow \left(y^{i'}-p_{j}^{i'}\right)\left(\mu_{j}+1\right)-\left(y^{i'}-p_{j'}^{i'}\right)\left(\mu_{j'}+1\right) > \left(y^{i}-p_{j}^{i}\right)\left(\mu_{j}+1\right)-\left(y^{i}-p_{j'}^{i}\right)\left(\mu_{j'}+1\right) \tag{3.8}$$

and from (3.4):

$$\left(y^{i'}-p_{j}^{i'}\right)\left(\mu_{j}+1\right)-\left(y^{i'}-p_{j'}^{i'}\right)\left(\mu_{j'}+1\right) > 0, \tag{3.9}$$

which contradicts (3.5). In other words, even if schools j and j' were indifferent between students i and i', and even if student i were indifferent between schools j and j', the higher income student (student i') would choose the higher reputation school (school j) in equilibrium.

ii) Assume that student i attends school j and student i' attends school j' in equilibrium, with $y^{i'} = y^{i}$, $b^{i'} > b^{i}$ and $j' < j$. As the assignment must be optimal for students *and schools, we obtain:*

$$\left(y^i - p_j^i\right)\left(\mu_j + 1\right) \geq \left(y^i - p_{j'}^i\right)\left(\mu_{j'} + 1\right), \tag{3.10}$$

$$\left(y^i - p_{j'}^{i'}\right)\left(\mu_{j'} + 1\right) \geq \left(y^i - p_j^{i'}\right)\left(\mu_j + 1\right), \tag{3.11}$$

$$p_j^i - c\left(b^i\right) \geq p_j^{i'} - c\left(b^{i'}\right), \text{ and} \tag{3.12}$$

$$p_{j'}^{i'} - c\left(b^{i'}\right) \geq p_{j'}^i - c\left(b^i\right). \tag{3.13}$$

But from (3.12) and (3.13) we get:

$$\begin{aligned} &\left(y^i - p_j^{i'}\right) - \left(y^i - p_j^i\right) \geq c\left(b^i\right) - c\left(b^{i'}\right) \geq \left(y^i - p_{j'}^{i'}\right) - \left(y^i - p_{j'}^i\right) \\ \Rightarrow &\left(y^i - p_j^{i'}\right)\left(\mu_j + 1\right) - \left(y^i - p_{j'}^{i'}\right)\left(\mu_{j'} + 1\right) > \left(y^i - p_j^i\right)\left(\mu_j + 1\right) - \left(y^i - p_{j'}^i\right)\left(\mu_{j'} + 1\right). \end{aligned} \tag{3.14}$$

And from (3.10):

$$\left(y^i - p_j^{i'}\right)\left(\mu_j + 1\right) - \left(y^i - p_{j'}^{i'}\right)\left(\mu_{j'} + 1\right) > 0, \tag{3.15}$$

which contradicts (3.11). ∎

Therefore, we obtain stratification by income and ability in equilibrium. This result follows from the form of the utility function considered. Nevertheless, we will show below that the only condition required to obtain this result in the general model is that perceived educational quality is a normal good.

In what follows we present the general reputation model and prove the main results. We assume that there is a continuum of students and schools. This assumption simplifies the analysis, since we know the proportion of schools that obtain high or low test results, and the fraction of schools that changes their type. Therefore, the evolution of the distribution of schools' reputation is known. Under this scenario, there is a "long run distribution of schools' reputations" in the high quality equilibrium, as we demonstrate in the following Chapter, and this is the distribution we consider in the analysis of the equilibrium in the present Chapter.

3.2 The Model

3.2.1 Schools

There is continuum with measure $(1-\kappa)$ of long lived, profit-maximizing private schools, that can serve at most one student. Schools are heterogeneous in their types $\tau \in \{C, I\}$ (competent –type C– or inept –type I–, private information) and their prior reputations ($\bar{\mu} = \Pr(C|\cdot)$, public information). The fraction of competent schools is $\theta \in (0,1)$, constant across time. G_t denotes the cdf of $\bar{\mu}$ at time t, with support U_t. G_t^C and G_t^I are the cdfs of the subpopulations of each private school type, so that $G_t = \theta G_t^C + (1-\theta) G_t^I$.

On the other hand, there is one inept[1] public school of unlimited capacity[2] that would accept any student and charge 0. This is a passive player. κ denotes the fraction of students that attend the public school, and $H_t(b, y)$ denotes their distribution.

We follow Mailath and Samuelson (2001) in introducing long run reputational effects in a model with imperfect public monitoring. Students are never completely sure about the school type: there is an exogenous probability $\lambda \in (0,1)$ that a school leaves the market each period and is replaced by another school. The probability of being replaced by a competent school is θ, and the new school inherits the reputation of the old school it

[1] We think that adding a small, limited mass of free, competent public schools would not alter substantially the results.

[2] One could alternatively think of a continuum of public schools of the same size as their private counterparts. What matters is that unlike the private sector, the public sector has no capacity constraints.

replaces. In other words, λ is the probability that an existing school changes its owner or management, and θ is the probability that the new manager is competent. Students do not observe the replacement, but they know it may occur (and they take this into account when updating beliefs). Under this scenario the uncertainty about types is continually replenished, since there is always a positive probability that a school changes its type. Without a mechanism such as this possibility of replacement, there is no room for permanent or long run reputations in a model with imperfect public monitoring, as Cripps, Mailath and Samuelson (2004) proved. This occurs because as the firm's reputation converges to one (i.e., consumers get convinced that the firm is competent), a bad signal has no effect on the posterior reputation, and hence the cost of choosing low effort converges to zero. In Hörner's (2002) model, the existence of competition among firms is a sufficient condition to assure that there is always a cost associated to low effort: as all firms charge the same price, if the customers of a given firm experience a bad outcome they switch to another firm. In our model, however, schools charge different prices according to their history of (public) test scores. Therefore, a school that obtained a low test result can still attract students in subsequent periods.

Schools observe students characteristics, and they can choose which student to serve. Therefore, private schools choose an admission policy (a set of "admitted" students, $P \subset B$). We restrict the sets P to belong to the family $\mathcal{B}(B)$ of compact, convex sets in B (i.e., intervals). Admission policies are observable, so they are potentially (very) informative to students. In most of the analysis that follows the equilibrium admission policies will be trivial (accept everyone) because the price function $p(\mu, b)$ will take care of the differences among students.

If a student is served by a competent school (which not only depends on P but also on student's equilibrium strategies), the school must choose whether to exert high or low effort. The effort level deterministically leads to an academic achievement level $a \in \{0,1\}$. For this reason we refer to low/high effort and low/high achievement interchangeably. $\alpha_t(\mu)$ denotes the fraction of type C schools with reputation μ that chooses $a = 1$ at time t, and it is referred to as a *production strategy* for it summarizes the choice of production of high and low quality education (effort). α_t is the average over the whole population of type C schools. The cost of providing high achievement is decreasing in b: the effort level required to reach $a = 1$ with a higher ability student is lower. $c(b)$ denotes the cost of high achievement for a student with ability b. Without effort (at zero cost), the educational achievement is 0 (independent of b). Inept schools always provide $a = 0$.[3]

After graduation, all students take a standardized test. The test results r are publicly observed. It so becomes commonly known which schools got high scores (H) and which ones got low scores (L), but nobody knows (up to equilibrium strategies) the characteristics of students that attended each school. The probability of obtaining a high score at t depends on the educational achievement provided by a school, in the following sense:

$$\Pr(r_t = H | a_t) = \begin{cases} \pi_1 \in (0,1) & \text{if } a_t = 1, \\ \pi_0 \in (0,1) & \text{if } a_t = 0, \end{cases} \tag{3.16}$$

where $\pi_1 > \pi_0$. Therefore, competent schools can increase the probability of obtaining a high score by exerting sufficient effort in order to provide $a = 1$. If the school accepts no

[3] They cannot choose higher effort, or the productivity of their effort is zero: they don't know how to improve student's achievement.

student, next period's reputation is $\mu_{t+1} = (1-\lambda)\,\mu_t + \lambda\theta$ –unless doing so would reveal its type. Otherwise, students beliefs are updated according to the test scores: the posterior belief after a high score is denoted μ_H, and after a low score μ_L.

Both decisions are made taking as given the personalized prices $p^*(\mu, b) \in \Phi$.

Remark 3.1 *Generally speaking, each school could make a decision contingent on the past history of observable play, including:*

- *Its own public record of scores* h_t
- *Its own reputation* μ_t
- *The history of equilibrium price functions* $\{p_k^*(\mu, b)\}_{k=0}^{t-1}$
- *The history of distributions of reputations* $\{G_k\}_{k=0}^{t}$
- *The history of equilibrium assignments* $\{\mu_k^*(b, y)\}_{k=0}^{t-1}$
- *The history of equilibrium production of quality education* $\{\alpha_k\}_{k=0}^{t-1}$

We focus, however, on strong Markov strategies where schools make decisions taking as the sole state variable their own reputation μ *and type* τ *(as well as the price function* p^**). This is justified by the perfect competition assumption, which rules out any sort of strategic behavior among schools, and the fact that students only care about the probability of receiving a high quality education (which is linked to the school type) and can only base their judgments about school types on their score records.*

We thus restrict attention to perfect Markov strategies σ of the form:

$$\begin{aligned} \sigma^C &: U \times \Phi \to \mathcal{B}(B) \times \{0,1\} \\ &: \sigma^C(\bar{\mu}, p) = (P, a) \end{aligned} \tag{3.17}$$

for competent schools, and:

$$\begin{aligned} \sigma^I &: U \times \Phi \to \mathcal{B}(B) \\ &: \sigma^I(\bar{\mu}, p) = P \end{aligned} \tag{3.18}$$

for inept schools since they can't produce high achievement.

Observe that the strategy has one observable component P, and for competent schools, another unobservable component a. The admission policy specifies which students *would* be accepted in case they applied (even if they don't). $P^{\tau}(\bar{\mu})$ is the set of admitted students by a school of type τ and prior reputation $\bar{\mu}$. The *interim* probability of being competent (denoted by μ_t) conditions on the observable strategies. In a separating equilibrium $\mu_t = 0$ if the observable behavior of a school within the class of schools with the same prior $\bar{\mu}_t$ differs from the equilibrium behavior for competent schools. In this case, since inept schools reveal their type, they charge at most 0 (because of the competition with the free public school).[4] We focus on a pooling equilibrium, where inept schools choose the same admission policy as competent schools, and $\mu_t = \bar{\mu}_t$. However, if they mimic competent schools, some inept schools inexorably will envy others –those who got the less able students, because they charge higher prices. Still, there is nothing they could do to attract those students without revealing their type.

[4]Observe, however, that if an inept school reveals its type at t, its $(t+1)$ prior would not be 0 but $\lambda\theta$: there exists a possibility of being replaced by a competent school.

3.2.2 Students

There is a sequence of (non-overlapping) generations of short-lived students, each of unit measure. Students $s \in S$ are heterogeneous in ability $(b \in B)$ and income $(y \in Y)$. The sets B and Y are closed intervals, and $S = B \times Y$. F denotes the joint cdf over (b, y), identical every period. Each student observes his own characteristics, but not the income or ability level of other students.

Students face prices $p^* (\mu, b)$ and admission policies $\{P^\tau (\bar{\mu})\}$ resulting in a set of schools that would admit them $P^{-1} (b) = \{\bar{\mu} \in U : \ b \in P^\tau (\bar{\mu}) \text{ for some } \tau\}$. Recall that the public school (with $\mu = 0$) would accept any applicant, so that regardless of the equilibrium under consideration these sets are nonempty: $0 \in P^{-1} (b)$ for all $b \in B$. Students choose a school from the set $P^{-1} (b)$, and spend whatever is left of their income on the consumption good: $z = y - p^* (\mu, b)$.

All students have the same bounded Bernoulli function $u (a, z)$, increasing in the unique consumption good, z, and the student's educational achievement, $a \in \{0, 1\}$.[5,6] Students do not observe the level of educational achievement provided by each school in advance: it is an "experience good". We will refer to he high educational achievement level as "high school quality".

Since students are short run players, they play static best replies, regardless of the history of the game. The only elements of the (public) history that may take into account

[5] Note that this implies that the utility function of the students is independent of b. That is, the valuation of educational achievement does not depend on b. With this assumption we isolate the analysis of the equilibrium assignment of students from possible differences in the valuation of educational achievement across students with different abilities.

[6] We also assume that the expected utility is zero if the student does not attend any (public or private) school.

are those that are judged to be useful in ascertaining school types. Students' beliefs are common, since all students observe the same history of test scores. Hence, a student of income y and ability b chooses μ (the only school attribute that matters to him) so as to:

$$\begin{aligned} \max_{\mu \in P^{-1}(b)} E\left[u\left(a,z\right)|\cdot\right] &= \Pr\left(a=1|\cdot\right)u\left(1,z\right)+\Pr\left(a=0|\cdot\right)u\left(0,z\right) \\ s/t \quad z &= y-p^{*}\left(\mu,b\right), \end{aligned} \tag{3.19}$$

from the set of schools that would accept him $P^{-1}(b)$. Replacing $\Pr\left(a=1|\cdot\right)=\Pr\left(a=1|\tau=C,\mu\right)$ $\Pr\left(\tau=C|\cdot\right)=\alpha\left(\mu\right)\mu$ leaves:

$$\max_{\mu \in P^{-1}(b)} \mu\alpha\left(\mu\right)u\left(1,y-p^{*}\left(\mu,b\right)\right)+\left(1-\mu\alpha\left(\mu\right)\right)u\left(0,y-p^{*}\left(\mu,b\right)\right). \tag{3.20}$$

We abuse terms and refer to a particular school with reputation μ (or group of them) as "the school μ", although the school's reputation is generally not constant across time, and despite the fact that schools are not countable (atomless economy).[7]

Summarizing, the timeline for the stage game is the following:

1. **Initial conditions.** Each school is endowed with a (public) history of test scores h_t and a (private) history of own types and ability of each former student, which determines $\bar{\mu}_t \in U_t \subset [0,1]$, the time-$t$ **prior probability of being competent** (simply called **reputation)**.

2. **Nature's moves**.

   (a) Nature announces privately its type to each school, drawing from the conditional

[7]Because there are uncountable many schools and students, we should qualify almost every phrase with an "a.e." provision, specifying that what we say is true only in positive-measure sets; this could become annoying, so we avoid every explicit comment of this sort, in the understanding that when we say "all" we mean "almost all".

distributions:

$$\begin{array}{l|cc} & \tau_t = C & \tau_t = I \\ \hline \Pr\left(\tau_{t+1} = C|\tau_t = \cdot\right) & \lambda\theta + (1-\lambda) & \lambda\theta \\ \Pr\left(\tau_{t+1} = I|\tau_t = \cdot\right) & \lambda(1-\theta) & 1-\lambda\theta \end{array} \quad (3.21)$$

The draws are independent within schools of the same last-period's type.

3. At the same time, Nature announces publicly to the (recently born) students their income and ability, $(y, b) \in Y \times B$.

Matching Stage. A Walrasian market for education opens.

1. Schools choose admission policies taking as given the price list $p^*(\mu, b)$. Admission policies are observable. The *interim* probability of being competent, μ_t, conditions on the observable behavior of the schools.

2. Students choose schools taking as given the price list $p^*(\mu, b)$ and the admission policy of each school.

3. Students and schools are correspondingly matched (at most one student per school); the equilibrium assignment is described by $\mu = \mu^*(y, b)$.

Education Production Stage. Each competent school chooses whether to spend $c(b)$ to offer a high quality education to its student or not, where $c'(b) < 0$. High (low) quality education translates itself into high (low) achievement level for the student. Note that quality is a dichotomous variable, so it wouldn't make sense to spend anything different from $c(b)$ or 0.

Information Stage. After graduation, all students take a standardized test. The test results r are publicly observed. The (common) beliefs about school types are the Bayesian upgrade of interim reputations. These posteriors are next period's priors.

These choices result collectively in:

- An **equilibrium price function** $p_t^*(\mu, b)$
- An **equilibrium assignment** or matching of students and schools $\mu_t^*(b, y)$
- An equilibrium production strategy $\alpha_t(\mu) \in [0, 1]$, the fraction of competent schools with reputation μ that chooses $a = 1$. This fraction determines the **equilibrium aggregate production of high quality education** or academic excellence at t. Because of the continuum assumption, exactly a fraction $\theta\alpha_t$ of students receives a high quality education (i.e., high achievement), and a fraction $\pi_1\theta\alpha_t + \pi_0((1 - \alpha_t)\theta + (1 - \theta))$ of students gets an H score at time t.

An assignment $\mu(b, y)$ is **feasible** if $\mu(b, y) \in P^{-1}(b)$ for all $b \in B$. In addition to the requirement that the assignment $\mu_t^*(b, y)$ be feasible, the corresponding masses of students and schools must "add up" for the assignment to be part of an equilibrium.

3.3 Equilibrium

The equilibrium on a single stage of the game cannot be analyzed in complete isolation because decisions at any given stage may have long run consequences for schools. For instance, the production decision of the school at date t will be $a = 1$ only if there is a future reward for this effort, because providing high achievement is unobservable, and hence there can be no present reward. On the other hand, the continuation value depends on the way students update their beliefs upon learning test scores, and on the sequence of price lists $\{p_t^*(\mu, b)\}$. Therefore, schools' production decision can only be understood within the repeated game. For this reason, we are forced to treat it as an exogenous variable when first analyzing the stage game, and endogenize it when looking at the repeated game.

There are two taxonomies for conceivable equilibria that are relevant to our discussion. The first is based on the industry-wide level of quality (or academic achievement) production, that is, in competent schools' *unobservable* behavior. Two polar cases can be distinguished:

1. **Low quality equilibrium** (LQE). An equilibrium where all competent schools choose null effort: all competent private schools provide $a = 0$ to their students, so that $\alpha_t = 0$. In this equilibrium, all students –by their knowledge of the equilibrium strategies– assign zero probability to the event $\{a = 1\}$. Therefore, schools' reputations are completely irrelevant, for students see no advantage in attending a competent school that offers the same level of educational achievement than inept schools.[8] Since the public school also provides $a = 0$, students are indifferent between the public school and any private school, as long as the latter are also free. Hence all schools μ must charge $p(\mu, b) = 0$ for every b, and consequently there is no incentive to provide high quality. This equilibrium would be the only possible outcome of the game if, for instance, the cost of $a = 1$, $c(b)$, were too high relative to the student's willingness to pay for it.

[8]Reputations are not constant over time because of the exogenous probability of type change. All reputations converge to θ in the long run, whatever their initial value.

2. **High quality equilibrium** (HQE). An equilibrium where the industry operates at full capacity, that is, all competent schools choose effort so as to provide $a = 1$ to their students: $\alpha_t = 1$. Observe that for this to be true, $a = 1$ must be optimal for competent schools regardless of their current reputation μ.

The second taxonomy distinguishes whether in the stage game different types of schools' *observable* behavior is different or not. In a (fully) **separating equilibrium** inept schools use different admission policies than competent schools:

$$P^C(\bar{\mu}) \neq P^I(\bar{\mu}) \quad \text{for all } \bar{\mu}. \tag{3.22}$$

Since admission policies are observable, school types become commonly known. This is to say, interim reputations are $\mu = 1$ for competent schools, and $\mu = 0$ for inept schools. As a consequence, in a separating equilibrium inept schools cannot charge a positive price because it is commonly known that they provide $a = 0$ just as the public school does, and the public school does it for free. Hence, the only way that an inept school could hope to make a profit is by imitating competent schools, i.e., exactly following them in whatever (observable) actions they choose. This occurs in a **completely pooling equilibrium:**

$$P^C(\bar{\mu}) = P^I(\bar{\mu}) \quad \text{for all } \bar{\mu}. \tag{3.23}$$

Of course, partial separation or pooling is conceivable too, in which inept schools of given reputations imitate alike competent schools and inept schools of other reputations behave differently from their competent counterparts.

Claim 3.1 *The repeated game doesn't have a (non-trivial, completely) separating equilibrium.*

If there was separation and competent schools charged a positive price, inept schools would rather mimic competent schools. If all competent schools always charged $p = 0$, no competent school would choose $a = 1$, hence $\alpha = 0$ and students would be indifferent between any private and the public school, so long as they all charged $p^* = 0$. At this price, competent and inept schools would be indifferent among students and hence among all possible admission policies. Competent and inept schools could choose different admission policies that separate them, but this would be a trivial separating equilibrium because they might just as well pool themselves under a unique admission policy (which could range from "accept everyone" to "accept none" because the public school by itself is capable of attending all students.).

Oddly enough, type observability combined with action unobservability is the worst case in this game from a social welfare viewpoint.

We will restrict attention to a completely pooling equilibrium, where competent and inept schools choose the same admission policy $P^*(\bar{\mu})$. As with the production decision, this choice's optimality can only be judged within the repeated game because it depends on the equilibrium sequence of price lists.

3.3.1 The pooling equilibrium

The pooling equilibrium is defined by the strategies the players choose along with a belief system that assigns to each information set a probability measure over the set of histories it contains. In particular, the belief system defines what is the interim reputation of a school that chooses any given admission policy $P'(\bar{\mu}) \neq P^*(\bar{\mu})$ –this is to say, the interpretation of out-of-equilibrium moves–.

We will consider a forward-induction kind of refinement for the belief system. The purpose of this refinement is to rule out equilibria where competent schools do not choose their preferred admission policy just because if they did, students would believe they are inept. To this end, we ask what admission policy would be preferred by competent and inept schools with prior reputation $\bar{\mu}$ if P did not affect the students beliefs about the school type, and we assume that the students beliefs are revised only if they observe an admission policy $P'(\bar{\mu}) \neq P^*(\bar{\mu})$ such that it would not be strictly preferred by competent schools, but it would be preferred by an inept one.

Let $\succsim_{\tau,\mu}$ denote the preference relation over $\Phi \times B$ for a school of type τ with **interim** reputation μ. Then, we will consider a belief system where the interim reputation of a school with **prior** reputation $\bar{\mu}$ is defined as follows:

$$\mu = \begin{cases} 0 \text{ if } P'(\bar{\mu}) \neq P^*(\bar{\mu}) \wedge \\ \qquad (p^*, P') \not\succ_{C,\bar{\mu}} (p^*, P^*) \wedge (p^*, P') \succ_{I,\bar{\mu}} (p^*, P^*), \\ \bar{\mu} \text{ otherwise.} \end{cases} \tag{3.24}$$

Observe that as all schools choose the same admission policy $P^*(\bar{\mu})$ then $\mu = \bar{\mu}$ for all $\bar{\mu} \in U$ in equilibrium.

3.3.2 Equilibrium in the stage game

Consider the stage game at date t, when competent and inept schools choose the same admission policy $P_t^*(\bar{\mu})$, the belief system is defined as in (3.24), and the production decision of competent schools is $\alpha_t(\mu)$. In the definition of the pooling equilibrium of the stage game we take the production decision as given, and we assume that the interim reputation equals the prior reputation.

Given the production strategy $\alpha_t(\mu)$, the equilibrium in the stage game is defined as follows.

Definition 3.1 *A* ***Walrasian*** *(completely)* ***pooling equilibrium*** *(WPE) of the stage t for the non-atomic economy $\langle F_t, G_t^C, G_t^I \rangle$ with the production strategy $\alpha_t(\mu)$ is a price function $p_t^*(\mu, b)$, a feasible assignment function*[9] *$\mu_t^*(b, y)$, and an admission policy function shared by competent and inept schools $P_t^*(\bar{\mu})$ such that:*

CE1 *$\mu_t^*(b, y)$ is optimal for the* ***students****, given the price function $p_t^*(\mu, b)$, the admission policy function $P_t^*(\bar{\mu})$ and the production strategy $\alpha_t(\mu)$. That is, no student with characteristics (b, y) wants to attend a school with a reputation different from $\mu_t^*(b, y)$. This entails that $b \in P_t^*(\mu_t^*(b, y))$.*

CE2 *$P_t^*(\bar{\mu})$ is optimal for* ***competent*** *schools with interim reputation $\mu = \bar{\mu}$, given the price function $p_t^*(\mu, b)$ and the production strategy $\alpha_t(\mu)$.*

CE3 *Market clearing. This is to say, given the assignment $\mu_t^*(b, y)$, the proportion of students who attend schools with a smaller reputation than μ among those who attend private schools equals $G_t(\mu)$ for all $\mu \in U_t$, and*

$$\kappa H_t(b, y) + (1 - \kappa) G_t(\mu^*(b, y)) = F(b, y) \qquad \forall b, y, \tag{3.25}$$

where κ is the fraction of students that attend the public school (i.e., the measure of the set $\{(b, y) \in B \times Y : \quad \mu_t^*(b, y) = 0\}$) and $H_t(b, y)$ is their corresponding cdf.

If $\langle p_t^*(\mu, b), \mu_t^*(b, y), P^*(\bar{\mu}) \rangle$ is a WPE for the production strategy $\alpha_t(\mu)$, then $p_t^*(\mu, b)$ is called a **supporting price function** for $\mu_t^*(b, y)$.

[9] Strictly speaking, $\mu^*(b, y)$ could be a correspondence, but in the cases we are interested in it is not, and it greatly simplifies the presentation to treat it like a function.

Let

$$B_t(\mu) \equiv \{b \in B : \mu = \mu_t^*(b, y)\} \tag{3.26}$$

be the set of abilities of those students that according to the equilibrium assignment attend schools with reputation μ. It must be the case that competent schools with reputation μ are indifferent among all students in $B_t(\mu)$, otherwise some of them would be rejected and consequently $\mu_t^*(b, y)$ would not be the equilibrium assignment.

Observe that in equilibrium there are price quotes for every pair (μ, b), yet the equilibrium assignment generally speaking defines a strict subset of $U_t \times B \times Y$. We denote by E_t such subset:

$$E_t \equiv \{(\mu, b, y) \in U_t \times B \times Y : \quad \mu = \mu_t^*(b, y)\}, \tag{3.27}$$

and by $[E_t]$ its projection over $U_t \times B$, i.e.,

$$[E_t] = \{(\mu, b) \in U_t \times B : \ b \in B_t(\mu)\}. \tag{3.28}$$

Then $p^*_{[E_t]}$ denotes the restriction of p^* to $[E_t]$.[10]

There is a great deal of redundancy outside $[E_t]$, in the sense that there are many price functions such that student choices $\mu_t^*(b, y)$ and school choices $(P_t^*(\bar{\mu}))$ are mutual best replies and yield the same outcome (no trade). To see this, consider the particular segment of students with a common ability level b (and whose incomes differ) and a school with (prior and interim) reputation μ. There are different situations that are coherent with the fact that we don't observe transactions between them, i.e., $(\mu, b) \notin [E_t]$:

[10] There could be infinitely many price functions $p(\mu, b)$ that coincide with $p^*_{[E]}(\mu, b)$ in $[E]$, and in that case any of them would be an equilibrium price function provided CE1 and CE2 are satisfied. In cases like this we will regard the pair $\langle p, \mu^* \rangle$ for any member of that class as the same equilibrium.

1. The price $p_t^*(\mu, b)$ is strictly unattractive to all students of ability b, regardless of their income. In this case, what strategies specify schools do is irrelevant; they could either reject or accept these students.

2. The price $p_t^*(\mu, b)$ is weakly attractive to some student of ability b and strictly unattractive to competent schools with reputation μ, so that their optimal policy is to reject his application: $b \notin P_t^*(\mu)$.

If, on the other hand, the price $p_t^*(\mu, b)$ were weakly attractive to some student of ability b, weakly attractive to school μ but strictly attractive to at least one of them, then they would trade so that $(\mu, b) \in [E_t]$. In the case of complete indifference on both sides, (μ, b) may or may not be part of $[E_t]$.

When the price $p_t^*(\mu, b)$ is strictly attractive to some student of ability b and strictly unattractive to competent schools with reputation μ there still could exist mutually beneficial trades at higher prices. A situation like this does not emerge as a competitive equilibrium in a standard model, because it would entail an excess demand. It is the fact that students apply only to schools that would accept them (i.e, they choose among $\mu \in P^{-1}(b)$) what prevents the existence of an excess demand in our model. But if students were allowed to make (attractive) counter offers to the schools that reject them, a situation like this would not emerge as an equilibrium outcome in our model either. For this reason, we will consider the following refinement: $\langle p_t^*(\mu, b), \mu_t^*(b, y), P^*(\mu) \rangle$ is a *refined* WPE if there is no alternative contract preferred by some student and some competent school to their equilibrium contract, with strict preference for one of them. If $\succsim_{b,y}$ denotes the preference relation for a student $\mathbf{s} \in S$ with ability b and income y, the latter condition can be written as follows: $\nexists\,(p, \mathbf{s}, \mu) \in \Phi \times S \times U_t$ such that:

- $(p, \mu) \succsim_{b,y} (p_t^* (\mu_t^* (b, y), b), \mu_t^* (b, y))$
- $(p, b) \succsim_{C,\mu} (p_t^* (\mu_t^* (b', y), b'), b')$ for all $b' \in B(\mu)$

with strict preference for either the student or the school.

Lemma 3.2 *If there is $(p, \mathbf{s}, \mu) \in \Phi \times S \times U_t$ such that $(p, \mu) \succsim_{b,y} (p_t^* (\mu_t^* (b, y), b), \mu_t^* (b, y))$ and $(p, b) \succsim_{C,\mu} (p_t^* (\mu_t^* (b', y), b'), b')$ for all $b' \in B(\mu)$ with strict preference for either the student or the school, then, in this alternative contract the student* $\mathbf{s}$ *attends a school $\mu \neq \mu_t^* (b, y)$.*

Proof. Suppose that $\mu = \mu_t^* (b, y)$, and hence $b \in B(\mu)$. Then, for $(p, \mu) \succ_{b,y} (p_t^* (\mu, b), \mu)$ it must be the case that $p < p_t^* (\mu, b)$; but this implies that $(p, b) \prec_{C,\mu} (p_t^* (\mu, b), b)$. Similarly, for $(p, b) \succ_{C,\mu} (p_t^* (\mu, b), b)$ it must be the case that $p > p_t^* (\mu, b)$, and therefore $(p, \mu) \prec_{b,y} (p_t^* (\mu, b), \mu)$. Hence, any alternative contract weakly preferred to the equilibrium contract by both sides and strictly preferred by one of them must satisfy $\mu \neq \mu_t^* (b, y)$. ■

Lemma 3.3 *$\langle p_t^* (\mu, b), \mu_t^* (b, y), P^* (\mu)\rangle$ is a* refined *WPE iff there is no extension of $p^*_{[E_t]}$ to $U_t \times B$, denoted by p', such that:*

i) the policy $P' (\mu) = B$ for all μ is optimal for competent schools given the price function p', and

ii) $\mu_t^* (b, y)$ is not optimal for the students given P' and p'.

Proof. Suppose that there is an extension of $p^*_{[E_t]}$ to $U_t \times B$, denoted by p', such that conditions i) and ii) are satisfied. Then, $(p', b) \sim_{C,\mu} (p_t^* (\mu_t^* (b', y), b'), b')$ for all

$b' \in B(\mu)$ and for all μ. As $(p_t^*(\mu_t^*(b,y),b), \mu_t^*(b,y))$ is still available for the students under p', condition ii) implies that the optimal assignment given P' and p' is strictly preferred to $\mu_t^*(b,y)$ by some student. Therefore, there is an alternative contract that is weakly preferred by some school and strictly preferred by this student. We conclude that if $\langle p_t^*(\mu,b), \mu_t^*(b,y), P^*(\mu)\rangle$ is a *refined* WPE, then there is no extension of $p_{[E_t]}^*$ to $U_t \times B$ such that conditions i) and ii) are satisfied.

Now suppose that there is no extension of $p_{[E_t]}^*$ to $U_t \times B$ such that condition i) and ii) are simultaneously satisfied. Consider an extension of $p_{[E_t]}^*$ to $U_t \times B$, such that condition i) is satisfied. Then, $(p',b) \sim_{C,\mu} (p_t^*(\mu_t^*(b',y),b'),b')$ for all $b' \in B(\mu)$ and for all μ. Besides, as condition ii) is not satisfied, we know that no student strictly prefer a contract (p',μ) to $(p_t^*(\mu_t^*(b,y),b), \mu_t^*(b,y))$. Therefore, any $(p,\mathbf{s},\mu) \in \Phi \times S \times U_t$ such that $(p,\mu) \succ_{b,y} (p_t^*(\mu_t^*(b,y),b), \mu_t^*(b,y))$ for student $\mathbf{s}$ must satisfy that the price charged by the school μ to the student $\mathbf{s}$ under p is lower than $p'(\mu,b)$, and hence $(p,b) \prec_{C,\mu} (p',b)$. But since $\succsim_{C,\mu}$ is transitive, we know that $(p,b) \prec_{C,\mu} (p_t^*(\mu_t^*(b',y),b'),b')$ for all $b' \in B(\mu)$. Similarly, if $(p,b) \succ_{C,\mu} (p_t^*(\mu_t^*(b',y),b'),b')$ for all $b' \in B(\mu)$, then $(p,\mu) \prec_{b,y} (p_t^*(\mu_t^*(b,y),b), \mu_t^*(b,y))$ and we conclude that if there is no extension of $p_{[E_t]}^*$ to $U_t \times B$ such that conditions i) and ii) are satisfied, then $\langle p_t^*(\mu,b), \mu_t^*(b,y), P^*(\mu)\rangle$ is a *refined* WPE. ■

Therefore, we only consider equilibria $\langle p_t^*(\mu,b), \mu_t^*(b,y), P_t^*(\mu)\rangle$ with the property that $p_t^*(\mu,b)$ is a supporting price function for $\mu_t^*(b,y)$ under the admission policy of accepting everyone: $P_t^*(\mu) = B$. With this requirement we rule out situations where the equilibrium assignment would not survive the students making attractive counter offers to the schools that formerly rejected them.

3.3.3 Belief Updating in Pooling Equilibria of the Repeated Game

Individual reputations

The score-generating process at any given stage of the game is as follows: Natures chooses a type τ for a given school (C or I); a type C school (competent) chooses an effort level and hence a level of academic achievement a (0 or 1) , while a type I school can only "choose" $a = 0$; schools that choose $a = 1$ have a probability of getting a high score (H) of π_1, while those that chose $a = 0$ have a probability of π_0, where $\pi_0 < \pi_1$.

Students form beliefs about a particular school's current type τ_t in a Bayesian fashion taking into account the information contained in its history h_t, the equilibrium production strategies $\{\alpha_t(\mu)\}$ and the school's date-0 reputation μ_0. For a fixed sequence of production strategies $\{\alpha_t(\mu_t)\}$, the school reputation follows a Markov process, where $\Pr\left(\tau_t = C|h_t, \{\alpha_k(\mu_k)\}_{k=0}^{t}, \mu_0\right)$ is defined recursively as follows:

$$\begin{aligned}
\mu_0 &= \Pr(\tau_0 = C) \quad \text{(given as an initial condition), and} \\
\mu_1 &= \Pr(\tau_1 = C|h_1, \alpha_0(\mu_0)) \\
&= \begin{cases} \lambda\theta + (1-\lambda)\frac{\mu_0(\pi_1\alpha_0(\mu_0)+\pi_0(1-\alpha_0(\mu_0)))}{\mu_0(\pi_1\alpha_0(\mu_0)+\pi_0(1-\alpha_0(\mu_0)))+\pi_0(1-\mu_0)} & \text{if } h_1 = H, \\ \lambda\theta + (1-\lambda)\frac{\mu_0((1-\pi_1)\alpha_0(\mu_0)+(1-\pi_0)(1-\alpha_0(\mu_0)))}{\mu_0((1-\pi_1)\alpha_0(\mu_0)+(1-\pi_0)(1-\alpha_0(\mu_0)))+(1-\pi_0)(1-\mu_0)} & \text{if } h_1 = L, \end{cases}
\end{aligned} \tag{3.29}$$

with:

$$\Pr(r_0 = H|\mu_0, \alpha_0(\mu_0)) = \pi_1\alpha_0(\mu_0)\mu_0 + \pi_0(1-\alpha_0(\mu_0)\mu_0), \tag{3.30}$$

and from then on, is governed by the stochastic difference equation

$$\begin{aligned} \mu_{t+1} &= \Pr\left(\tau_t = C | h_t, \{\alpha_k\}_{k=0}^t\right) \\ &= \begin{cases} \lambda\theta + (1-\lambda)\frac{\mu_t(\pi_1\alpha_t(\mu_t)+\pi_0(1-\alpha_t(\mu_t)))}{\mu_t(\pi_1\alpha_t(\mu_t)+\pi_0(1-\alpha_t(\mu_t)))+\pi_0(1-\mu_t)} & \text{if } h_{t+1} = h_t H, \\ \lambda\theta + (1-\lambda)\frac{\mu_t((1-\pi_1)\alpha_t(\mu_t)+(1-\pi_0)(1-\alpha_t(\mu_t)))}{\mu_t((1-\pi_1)\alpha_t(\mu_t)+(1-\pi_0)(1-\alpha_t(\mu_t)))+(1-\pi_0)(1-\mu_t)} & \text{if } h_{t+1} = h_t L, \end{cases} \end{aligned} \tag{3.31}$$

where

$$\Pr\left(H|\mu_t\right) = \pi_1\alpha_t\left(\mu_t\right)\mu_t + \pi_0\left(1 - \alpha_t\left(\mu_t\right)\mu_t\right). \tag{3.32}$$

The sequence $\{\alpha_t(\mu_t)\}$ is determined in equilibrium together with the price and assignment sequences. However, prices and assignments do not intervene directly in the evolution of reputations.

If at all times the **low quality equilibrium** obtains, whereby the industry never uses any of its capacity, $(\forall t, \quad \alpha_t = 0)$, the above equations become:

$$\mu_H = \lambda\theta + (1-\lambda)\mu, \text{ and} \tag{3.33a}$$

$$\mu_L = \lambda\theta + (1-\lambda)\mu. \tag{3.33b}$$

This is to say, reputations evolve deterministically according to:

$$\mu_{t+1} = \lambda\theta + (1-\lambda)\mu_t, \tag{3.34}$$

converging in the long run to the unique, globally stable steady state $\mu = \theta$. No matter what the initial reputations, since both types of school choose the same a, they are both equally likely to get good scores, so in their score history there is no information to tell one from the other. As time passes by, the chances that the initial school type has changed converges to 1 and the probability that any given school is of type C converges to the population average.

In the **high quality equilibrium** where the industry always works at full capacity ($\forall t, \quad \alpha_t = 1$), the above equations become:

$$\mu_H(\mu) = \lambda\theta + (1-\lambda)\frac{\pi_1\mu}{\pi_1\mu + \pi_0(1-\mu)}, \tag{3.35}$$

and

$$\mu_L(\mu) = \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\mu}{(1-\pi_1)\mu + (1-\pi_0)(1-\mu)} \tag{3.36}$$

respectively, where time subscripts have been dropped for the sake of readability. Hence each school's reputation evolves stochastically according to the binomial process:

$$\mu_{t+1} = \begin{cases} \lambda\theta + (1-\lambda)\frac{\pi_1\mu_t}{\pi_1\mu_t+\pi_0(1-\mu_t)} & \text{if } r_t = H, \\ \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\mu_t}{(1-\pi_1)\mu_t+(1-\pi_0)(1-\mu_t)} & \text{if } r_t = L, \end{cases} \tag{3.37}$$

with

$$\Pr(H|\mu_t) = \pi_1\mu_t + \pi_0(1-\mu_t). \tag{3.38}$$

This is illustrated in Figure 3.1. It is apparent in it that both functions $\mu_H(\mu)$ and $\mu_L(\mu)$ are continuous, one-to-one, and have continuous inverse functions.

Observe that the image of both functions is strictly smaller than the interval $[0,1]$ where original reputations are defined:

$$\begin{aligned} \mu_L([0,1]) &= \mu_H([0,1]) \\ &= [\lambda\theta, \lambda\theta + (1-\lambda)] \subset [0,1]. \end{aligned} \tag{3.39}$$

Moreover, the fact that μ_L lies above the identity for small values of μ means that for too low reputations even after the worst results next period's reputation would be better, because

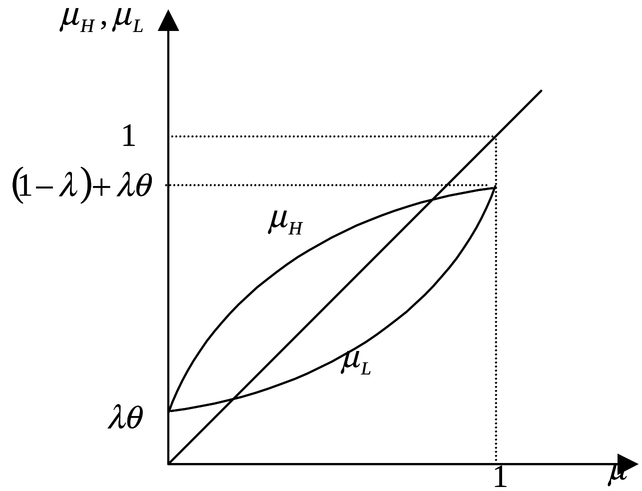

Figure 3.1: μ_H and μ_L as a function of μ.

of the probability of type change. A similar comment applies to very high reputations. Define $\mu_{\min}$ and $\mu_{\max}$ as the values that solve the following equations:

$$\mu_{\min} = \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\,\mu_{\min}}{(1-\pi_1)\,\mu_{\min} + (1-\pi_0)(1-\mu_{\min})}, \text{ and} \tag{3.40}$$

$$\mu_{\max} = \lambda\theta + (1-\lambda)\frac{\pi_1\mu_{\max}}{\pi_1\mu_{\max} + \pi_0(1-\mu_{\max})}. \tag{3.41}$$

Then, $\mu_{\min}$ and $\mu_{\max}$ are bounds for the support of G_t^τ, U_t^τ, if t is sufficiently large. Indeed, $\lim_{n\to\infty}\Pr\left(\mu_{t+n} < \mu_{\min}\right) = 0$ and $\lim_{n\to\infty}\Pr\left(\mu_{t+n} > \mu_{\max}\right) = 0$. Therefore, $\lim_{t\to\infty} U_t^\tau \subseteq [\mu_{\min}, \mu_{\max}] \subset [\lambda\theta, \lambda\theta + (1-\lambda)]$.

Observe that in this case (and unlike any other conceivable equilibrium) choosing $a = 1$ and being of type C are the same event.

Population reputations in a sequence of pooling HQE

Looking at the population as a whole, unlike individual reputations and because of the continuum assumption, the evolution of the cdf of reputations is deterministic. For competent schools, when $\alpha_t(\mu) = 1$ for all μ, t, we have:

$$\begin{aligned} G^C_{t+1}(x|\mu_t) &= \Pr\left(\mu_{t+1} < x|\mu_t\right) \\ &= (1-\lambda+\lambda\theta)\Pr\left(\mu_{t+1} < x|\mu_t, C\right) + \frac{\lambda\theta(1-\theta)}{\theta}\Pr\left(\mu_{t+1} < x|\mu_t, I\right) \\ &= (1-\lambda+\lambda\theta)\left[I_{\{\mu_H < x|\mu_t\}}\pi_1 + I_{\{\mu_L < x|\mu_t\}}(1-\pi_1)\right] \\ &+ \lambda(1-\theta)\left[I_{\{\mu_H < x|\mu_t\}}\pi_0 + I_{\{\mu_L < x|\mu_t\}}(1-\pi_0)\right], \end{aligned} \tag{3.42}$$

where $I_{\{\cdot\}}$ is an indicator function, taking the value 1 if event $\{\cdot\}$ occurs, and 0 otherwise.

Using the definitions of μ_H and μ_L in equations (3.35) and (3.36) respectively, we get:

$$\begin{aligned} \mu_H(\mu_t) < x \Leftrightarrow \mu_t &< \frac{(x-\lambda\theta)\pi_0}{((1-\lambda+\lambda\theta-x)\pi_1 + (x-\lambda\theta)\pi_0)} = \mu_H^{-1}(x)\text{, and} \\ \mu_L(\mu_t) < x \Leftrightarrow \mu_t &< \frac{(x-\lambda\theta)(1-\pi_0)}{((1-\lambda+\lambda\theta-x)(1-\pi_1) + (x-\lambda\theta)(1-\pi_0))} = \mu_L^{-1}(x). \end{aligned} \tag{3.43}$$

Therefore, we obtain:

$$\begin{aligned} G^C_{t+1}(x|\mu_t) &= (1-\lambda+\lambda\theta)\left[I_{\left\{\mu_t < \mu_H^{-1}(x)\right\}}\pi_1 + I_{\left\{\mu_t < \mu_L^{-1}(x)\right\}}(1-\pi_1)\right] \\ &+ \lambda(1-\theta)\left[I_{\left\{\mu_t < \mu_H^{-1}(x)\right\}}\pi_0 + I_{\left\{\mu_t < \mu_L^{-1}(x)\right\}}(1-\pi_0)\right]. \end{aligned} \tag{3.44}$$

By the law of iterated expectations, we obtain the (unconditional) distribution function of reputations for competent schools at $t+1$ as:

$$\begin{aligned}
G^C_{t+1}(x) &\equiv \Pr\left(\mu_{t+1} < x\right) \\
&= (1-\lambda+\lambda\theta)\left[\pi_1 \int_0^{\mu_H^{-1}(x)} \mathrm{d}G_t^C + (1-\pi_1)\int_0^{\mu_L^{-1}(x)} \mathrm{d}G_t^C\right] \\
&+ \lambda(1-\theta)\left[\pi_0 \int_0^{\mu_H^{-1}(x)} \mathrm{d}G_t^I + (1-\pi_0)\int_0^{\mu_L^{-1}(x)} \mathrm{d}G_t^I d\mu\right] \\
&= (1-\lambda+\lambda\theta)\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right) + (1-\pi_1) G_t^C\left(\mu_L^{-1}(x)\right)\right] \\
&+ \lambda(1-\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right) + (1-\pi_0) G_t^I\left(\mu_L^{-1}(x)\right)\right]. \qquad (3.45)
\end{aligned}$$

A similar procedure yields the cdf for inept schools:

$$\begin{aligned}
G^I_{t+1}(x) &\equiv \Pr\left(\mu_{t+1} < x\right) \\
&= (1-\lambda\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right) + (1-\pi_0) G_t^I\left(\mu_L^{-1}(x)\right)\right] \\
&+ \lambda\theta\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right) + (1-\pi_1) G_t^C\left(\mu_L^{-1}(x)\right)\right]. \qquad (3.46)
\end{aligned}$$

The overall cdf of μ at $t+1$ is then:

$$\begin{aligned}
G_{t+1}(x) &= \theta G^C_{t+1}(x) + (1-\theta) G^I_{t+1}(x) \\
&= \theta\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right) + (1-\pi_1) G_t^C\left(\mu_L^{-1}(x)\right)\right] \\
&+ (1-\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right) + (1-\pi_0) G_t^I\left(\mu_L^{-1}(x)\right)\right]. \qquad (3.47)
\end{aligned}$$

Summary 3.1 *In a sequence of high-quality equilibria for a given pair of distributions*

$\left(G_t^C, G_t^I\right)$, *next period's distributions* $\left(G_{t+1}^C, G_{t+1}^I\right)$ *are given by:*

$$\begin{aligned} T\left[\left(G_t^C(x), G_t^I(x)\right)\right] &= \Big((1-\lambda+\lambda\theta)\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right)+(1-\pi_1)\,G_t^C\left(\mu_L^{-1}(x)\right)\right] \\ &\quad +\lambda(1-\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right)+(1-\pi_0)\,G_t^I\left(\mu_L^{-1}(x)\right)\right], \\ &\quad (1-\lambda\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right)+(1-\pi_0)\,G_t^I\left(\mu_L^{-1}(x)\right)\right] \\ &\quad +\lambda\theta\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right)+(1-\pi_1)\,G_t^C\left(\mu_L^{-1}(x)\right)\right]\Big). \end{aligned} \tag{3.48}$$

The operator T defines a first-order difference equation on the product of two cdf spaces. Notice that T is constructed under the assumption that $\alpha_t(\mu_t)=1$ at all times, so it defines the dynamic of population reputations in an infinite sequence of HQE.

3.3.4 Equilibria in the Repeated Game

An equilibrium of the repeated game is formed by a sequence of *WPE* that satisfies two conditions. The first is that individual and population reputations evolve in accordance to Bayes' rule and the equilibrium strategies of each stage, starting from the initial distributions $\left(G_0^C, G_0^I\right)$. The second condition is that the choice variables that were taken as exogenous variables in the definition of the equilibrium of each stage game are in fact optimal for the schools, given the sequence of price lists and the belief system in the repeated game. That is, the admission policy $P_t^*(\bar{\mu})$ shared by competent and inept schools must be optimal for both types of school, despite the fact that inept schools would prefer another admission policy if this did not affect the students' beliefs. Likewise, the production decision $\alpha_t(\mu)$ must be optimal for competent schools, given the Bayesian updating of students' beliefs.

Definition 3.2 *A* ***Markov sequential pooling equilibrium*** *(MSPE) for the repeated* *game of the economy* $\left\langle F, G_0^C, G_0^I \right\rangle$ *with the belief system defined in (3.24) is a sequence of price lists* $\{p_t^*(\mu, b)\}$*, a sequence of assignment correspondences* $\{\mu_t^*(b, y)\}$*, a sequence of admission policies* $\{P_t^*(\bar{\mu})\}$*, a sequence of Markov production strategies* $\{\alpha_t^*(\mu)\}$*, and a sequence of reputation distribution pairs* $\left\{\left(G_t^C, G_t^I\right)\right\}$ *such that for all* t *and* μ *:*

1. *Each tuple* $\langle p_t^*, \mu_t^*, P_t^* \rangle$ *conforms a WPE of the stage game given* $\alpha_t^*(\mu)$*.*

2. $P_t^C(\bar{\mu}) = P_t^I(\bar{\mu}) = P_t^*(\bar{\mu})$ *is optimal for competent and inept schools with prior reputation* $\bar{\mu}$ *given the belief system.*

3. $\alpha_t^*(\mu)$ *is optimal for competent schools with interim reputation* μ*.*

4. *Individual reputations and the population cdfs of reputations* $\left(G_t^C, G_t^I\right)$ *evolve in accordance to Bayes' rule and equilibrium strategies.*

When in a MSPE the industry works at all times at full capacity (i.e., the equilibrium production strategy is $\alpha_t^*(\mu) = 1$ for all t) we call such equilibrium a **high-quality Markov sequential pooling equilibrium** (HQ-MSPE). From the discussion in the previous section, in a HQ-MSPE we have:

$$\left(G_{t+1}^C, G_{t+1}^I\right) = T\left[\left(G_t^C, G_t^I\right)\right]. \tag{3.49}$$

Observe that if there were a pair of distributions $\left(G^C, G^I\right)$ such that $\left(G^C, G^I\right) = T\left[\left(G^C, G^I\right)\right]$, and if there was a tuple $\langle p^*, \mu^*, P^* \rangle$ that constituted a WPE of the economy $\left\langle F, G^C, G^I \right\rangle$, then the constant sequence $\left\langle \{p^*\}, \{\mu^*\}, \{P^*\}, \{\alpha^*\}, \left\{G^C\right\}, \left\{G^I\right\} \right\rangle$ would be a MSPE of the repeated game for the economy with initial reputations $\left(G^C, G^I\right)$. In this case, we would say that $\langle p^*, \mu^*, P^*, \alpha^* \rangle$ is a **steady-state Walrasian pooling equilibrium** (SS-WPE) of $\left\langle F, G^C, G^I \right\rangle$.

In the next Chapter, we analyze the dynamics of μ_t and of the distribution of schools' reputation in a HQ-MSPE. Theorem (3.2) below guarantees the existence of such pair of distribution functions in the case of a HQ-MSPE –provided the existence of a supporting price function p^* is granted. That is, the distribution of schools' reputation converges to a "long run distribution", the fixed point of the dynamical system that describes the evolution of reputation distributions. We also find that the support of the long run distribution, U, is infinite, and that the distribution of competent schools' reputation first-order stochastically dominates that of inept schools.

Definition 3.3 *An operator T on a Banach space is said to be an $N-$contraction if there exists a finite $N \in \mathbb{N}$ such that its N-th iterate is a contraction, i.e., if $f : X \to X$ defined by:*

$$f(x) = T^N x$$

is a contraction in X.

The following Theorem generalizes Banach's fixed point Theorem to $N-$contractions:

Theorem 3.1 *An $N-$contraction T on a Banach space has a unique fixed point.*

Proof. See Chapter 4. ■

Theorem 3.2 *T is an $N-$contraction in $\left(X^C \times X^I, \rho\right)$, the complete metric space of the set of all pairs of distribution functions $\left(G^C, G^I\right)$ with support contained in the interval*

$[\lambda\theta, \lambda\theta + (1-\lambda)]$ *endowed with the metric*[11]*:*

$$\rho\left(\left(G^C, G^I\right), \left(G^{C\prime}, G^{I\prime}\right)\right) = \max\left\{\sup_x \left|G^C(x) - G^{C\prime}(x)\right|, \sup_x \left|G^I(x) - G^{I\prime}(x)\right|\right\}. \quad (3.50)$$

Proof. The proof is the main content of Chapter 4. ■

The significance of this Theorem is that it ensures that if the economy is at all times in a HQE, the population-wide distributions of reputations for competent and inept schools converges –regardless of their date-0 value– to a unique steady state. We can offer the following characterization of the long run distribution of reputations G in a HQ-MSPE:

Lemma 3.4 *The support of the long run distribution of reputations is infinite. In particular, the long run distribution is non-degenerate.*

Proof. See Chapter 4. ■

Remark 3.2 *Recall that the support U of G is smaller than the set of possible reputations $[0,1]$. In fact, the following chain of inclusions verifies: $U \subseteq [\mu_{\min}, \mu_{\max}] \subset [\theta\lambda, 1-\lambda+\lambda\theta]$.*

Lemma 3.5 *G^C is stochastically larger than G^I.*

Proof. See Chapter 4. ■

Summary 3.2 *If the game as a whole has an equilibrium where at all stages there is full production of high quality, in the long run the distributions $\left(G_t^C, G_t^I\right)$ converge globally to a unique steady state pair $\left(G^C, G^I\right)$. Each G^τ is non-degenerate, has an infinite support*

[11] The metric under consideration declares the distance between any two pairs of cumulative distribution functions $\left(G^C, G^I\right)$ and $\left(G'^C, G'^I\right)$ to be the maximum between $\rho_\infty\left(G^C, G^{C\prime}\right)$ and $\rho_\infty\left(G^I, G^{I\prime}\right)$, where $\rho_\infty\left(G^\tau, G^{\tau\prime}\right)$ stands for the distance between the distribution functions G^τ and $G^{\tau\prime}$ under the sup metric.

U (strictly smaller than the unit interval), and G^C first-order stochastically dominates G^I, i.e., competent schools have "better" reputations than inept ones.

3.4 The Steady-State High-Quality Equilibrium

This Section analyzes a steady-state Walrasian pooling equilibrium (SS-WPE) of $\left\langle F, G^C, G^I \right\rangle$ in which the constant sequence $\left\langle \{p^*\}, \{\mu^*\}, \{P^*\}, \{\alpha^*\}, \{G^C\}, \{G^I\} \right\rangle$ is a Markov sequential pooling equilibrium (MSPE) such that $\alpha^* = 1$ for all $\mu \in U$. For short, we will refer to it as a **steady-state high quality equilibrium** (SS-HQE).

Imposing $a^* = 1$ for all μ yields the following Bellman equation for competent schools:

$$v_C(\mu) = \max_{b \in B} \{p(\mu, b) - c(b) + \delta(1-\lambda)(\pi_1 v_C(\mu_H(\mu)) + (1-\pi_1) v_C(\mu_L(\mu)))\}. \quad (3.51)$$

Observe that only the contemporaneous payoff in the right hand side of (3.51) depends on b. Hence, if $p(\mu, b) - c(b)$ were strictly increasing in b, P would contain just the maximum of B (which exists because B is compact); if it were strictly decreasing, it would contain the minimum of B. Otherwise, it would be:

$$P^*(\mu) = \{b \in B : p(\mu, b) - c(b) = k(\mu)\}. \quad (3.52)$$

All this assuming $v_C(\mu) \geq 0$, for otherwise the optimal policy would be $P^*(\mu) = \varnothing$.

The following is a necessary condition for $\mu^*(b, y)$ to be an equilibrium assignment:

$$B(\mu) \subset \arg\max_b v_C(\mu) = P^*(\mu). \quad (3.53)$$

Since $\alpha = 1$ in the HQE, the student's expected utility becomes:

$$E[u] = \mu u(1, z) + (1-\mu) u(0, z). \quad (3.54)$$

Totally differentiating (3.54) with respect to z and μ we get:

$$\mathrm{d}E\left[u\right]=\left(u\left(1,z\right)-u\left(0,z\right)\right)\mathrm{d}\mu+\left(\mu\frac{\partial u\left(1,z\right)}{\partial z}+\left(1-\mu\right)\frac{\partial u\left(0,z\right)}{\partial z}\right)\mathrm{d}z. \tag{3.55}$$

To shorten notation a bit, write $u_z\left(1,z\right)$ for $\frac{\partial u(1,z)}{\partial z}$ and $\Delta u\left(\cdot,z\right)=u\left(1,z\right)-u\left(0,z\right)$. Imposing $\mathrm{d}E\left[u\right]=0$, we obtain the marginal rate of substitution between perceived quality (i.e., reputation) and consumption:

$$MRS=\left|\frac{\mathrm{d}z}{\mathrm{d}\mu}\right|_{\mathrm{d}E[u]=0}=\frac{\Delta u\left(\cdot,z\right)}{E\left[u_z\left(\cdot,z\right)\right]}. \tag{3.56}$$

We assume perceived quality is a normal good, i.e., that the MRS is increasing in z. The assumption is therefore:

$$\frac{\partial MRS}{\partial z}=\frac{\Delta u_z\left(\cdot,z\right)}{E\left[u_z\left(\cdot,z\right)\right]}-\frac{\left(\Delta u\left(\cdot,z\right)\right)E\left[u_{zz}\left(\cdot,z\right)\right]}{\left(E\left[u_z\left(\cdot,z\right)\right]\right)^2}>0. \tag{3.57}$$

3.4.1 Pricing

Since $P^*\left(\mu\right)$ is optimal for competent schools with reputation μ, those schools must be indifferent among all students in $P^*\left(\mu\right)$. Therefore we obtain the following result:

Theorem 3.3 *Every SS-HQE supporting price function satisfies that for all $\mu\in U$ and $b\in B\left(\mu\right)$ (i.e., $\forall\left(\mu,b\right)\in\left[E\right]$):*

$$p^*\left(\mu,b\right)-c\left(b\right)=k\left(\mu\right). \tag{3.58a}$$

Proof. *In equilibrium $B\left(\mu\right)\subset P^*\left(\mu\right)$.* ■

Remark 3.3 *From Lemma (3.3) we know that this function $p^*\left(\mu,b\right)$ can be extended to all $U\times B$ if the admission policies take the simple form "admit everyone", i.e., $P^*\left(\mu\right)=B$ for* *all $\mu \in U$. This extension is the additively separable function $p^*(\mu, b) = k(\mu) + c(b)$. We will assume in the sequel that $p^*(\mu, b) = k(\mu) + c(b)$ is the actual equilibrium price function in $U \times B$, so that $P^*(\mu) = B$ and the optimal strategy for competent schools is:*

$$\sigma(\mu, C, p^*) = (B, 1). \tag{3.59}$$

Hence, the Bellman equation for competent schools is:

$$v_C(\mu) = k(\mu) + \delta(1-\lambda)(\pi_1 v_C(\mu_H(\mu)) + (1-\pi_1) v_C(\mu_L(\mu))). \tag{3.60}$$

Since inept schools choose the same admission policy as competent schools but they do not spend $c(b)$, their expected contemporaneous payoff is $E_{b \in B(\mu)}[p(\mu, b)] = k(\mu) + E_{b \in B(\mu)}[c(b)]$. Therefore, the Bellman equation for inept schools is:

$$v_I(\mu) = k(\mu) + E_{b \in B(\mu)}[c(b)] + \delta(1-\lambda)(\pi_0 v_I(\mu_H(\mu)) + (1-\pi_0) v_I(\mu_L(\mu))). \tag{3.61}$$

Lemma 3.6 *The value of reputation is nonnegative:*

$$v_C(\mu) \geq \max\{0, \delta(1-\lambda) v_C((1-\lambda)\mu + \lambda\theta)\}.$$

Proof. $v_C(\mu) \geq 0$ is a participation constraint. Accepting the applicant this period must be preferred to rejecting him and waiting till next period. This is so if and only if:

$$v_C(\mu) \geq \delta(1-\lambda) v_C((1-\lambda)\mu + \lambda\theta).$$

■

The next four lemmas follow from incentive compatibility constraints as well.

Lemma 3.7 *The cost of high effort is bounded above by:*

$$\delta(1-\lambda)(\pi_1 - \pi_0)(v_C(\mu_H) - v_C(\mu_L)).$$

Proof. For spending in high achievement to be preferred, it must be the case that for $b \in P^*(\mu)$:

$$p(\mu, b) - c(b) + \delta(1-\lambda)(\pi_1 v_C(\mu_H) + (1-\pi_1) v_C(\mu_L))$$
$$\geq p(\mu, b) + (1-\lambda)\delta(\pi_0 v_C(\mu_H) + (1-\pi_0) v_C(\mu_L)).$$

Solving for $c(b)$ yields the desired result. ■

Lemma 3.8 $v_C(\mu) \geq v_I(\mu) \geq 0$

Proof. $v_I(\mu)$ is not larger than the payoff that a competent school could get through the policy of never investing in $a = 1$. It could be smaller, because inept schools cannot choose $a = 1$, and they must imitate the admission policy while competent schools choose it. $v_I(\mu) \geq 0$ is a participation constraint. ■

Remark 3.4 *This is not to say that an inept school will not be willing to incur in a loss at a particular moment, but simply that it will be willing to do so only if the expected future profits makes it worthwhile.*

Lemma 3.9 $p^*(\mu, b)$ *is strictly increasing in* μ *for all* $\mu \in U$.

Proof. Since

$$v_C(\mu) = k(\mu) + \delta(1-\lambda)(\pi_1 v_C(\mu_H) + (1-\pi_1) v_C(\mu_L)),$$

we know:

$$\frac{\partial v_C(\mu)}{\partial \mu} = \frac{\partial k}{\partial \mu} + \delta(1-\lambda)\left(\pi_1 \frac{\partial v_C(\mu_H)}{\partial \mu_H}\frac{\partial \mu_H}{\partial \mu} + (1-\pi_1)\frac{\partial v_C(\mu_L)}{\partial \mu_L}\frac{\partial \mu_L}{\partial \mu}\right).$$

But since $\frac{\partial \mu_H}{\partial \mu}$ and $\frac{\partial \mu_L}{\partial \mu}$ are both positive, and $\frac{\partial k}{\partial \mu} = \frac{\partial p}{\partial \mu}$,

$$\text{sign}\left(\frac{\partial p}{\partial \mu}\right) = \text{sign}\left(\frac{\partial v_C}{\partial \mu}\right).$$

Finally, if $\frac{\partial v_C}{\partial \mu} < 0$, the expected value would be decreasing in the probability of a high score:

$$\left(\pi_1 v_C\left(\mu_H\right) + \left(1-\pi_1\right) v_C\left(\mu_L\right)\right) < \left(\pi_0 v_C\left(\mu_H\right) + \left(1-\pi_0\right) v_C\left(\mu_L\right)\right),$$

resulting in a negative gross benefit of investing in $a = 1$. ■

Lemma 3.10 $v_I\left(\mu\right) \geq \delta\left(1-\lambda\right) v_I\left(\lambda\theta\right).$

Proof. Choosing the same admission policy as competent schools must be preferred by inept schools to accepting a different pool of students and revealing their type. If a school chooses an admission policy $P'\left(\mu\right) \neq P^*\left(\mu\right)$, the interim reputation is $\mu = 0$, and the posterior reputation is $\lambda\theta$. Since inept schools cannot charge a positive price if they reveal their type, this condition is satisfied if and only if:

$$v_I\left(\mu\right) \geq \delta\left(1-\lambda\right) v_I\left(\lambda\theta\right).$$

■

Remark 3.5 *As $\lambda\theta < \mu$ for all $\mu \in U$ and the value function $v_C\left(\mu\right)$ is increasing, then $v_C\left(\mu\right) > \delta\left(1-\lambda\right) v_C\left(\lambda\theta\right)$. Therefore, all competent schools strictly prefer the admission policy $P^*\left(\mu\right)$ to any different admission policy.*

Summary 3.3 *In a SS-HQE and under the trivial admission policy $P^*\left(\mu\right) = B$ the pricing function $p^*\left(\mu, b\right)$ is additively separable, strictly increasing in μ and strictly decreasing in b,* *for all $\mu \in U$ and $b \in B$.*[12] *At the same time, the value functions for competent and inept schools $v_C(\mu)$ and $v_I(\mu)$ are non-negative and v_C is strictly increasing in $\mu \in U$.*

We now turn to the characterization of the equilibrium assignment.

3.4.2 Assignment

Private versus public schools

Definition 3.4 *We define student $(b, y)'s$* ***reservation price*** *p_R for a school μ as the price that makes him exactly indifferent between attending school μ and the public school:*

$$u(0, y) = \mu u(1, y - p_R) + (1 - \mu) u(0, y - p_R). \tag{3.62}$$

Lemma 3.11 *$p_R(\mu, b, y)$ is:*

1. *Increasing in μ and y, and*
2. *Independent of b.*

Proof. Totally differentiating (3.62) with respect to y, μ and p_R we get:

$$u_z(0, y)\,\mathrm{d}y = \Delta u(\cdot, y - p_R)\,\mathrm{d}\mu + E[u_z(\cdot, y - p_R)](\mathrm{d}y - \mathrm{d}p_R). \tag{3.63}$$

Hence, $p_R(\mu, b, y)$ must be increasing in μ for all (b, y) :

$$\frac{\partial p_R}{\partial \mu} = \frac{\Delta u(\cdot, y - p_R)}{E[u_z(\cdot, y - p_R)]} > 0. \tag{3.64}$$

On the other hand, it does not depend on b, and is increasing in y under the assumption that $\frac{\partial MRS}{\partial z} > 0$ because:

$$p_R = \int_0^{\mu} MRS|_{E[u]=u(0,y)}\,\mathrm{d}\mu, \tag{3.65}$$

[12] Other combinations of admission policies and price functions are also consistent with the same allocation: deny admission whenever the price is smaller than $p^*(\mu, b)$ above.

where $MRS|_{E[u]=u(0,y)}$ stands for the MRS evaluated at z', the level of z that solves:

$$\mu u\left(1, z'\right) + (1-\mu)\, u\left(0, z'\right) = u\,(0, y)\,. \tag{3.66}$$

Since z' is increasing in y, we obtain:

$$\frac{\partial p_R}{\partial y} = \int_0^{\mu} \frac{\partial\; MRS|_{E[u]=u(0,y)}}{\partial y} \mathrm{d}\mu \tag{3.67}$$

$$= \int_0^{\mu} \frac{\partial MRS}{\partial z}\frac{\partial z'}{\partial y} \mathrm{d}\mu > 0. \tag{3.68}$$

■

Theorem 3.4 *The public school receives students with lower income and/or ability than private schools, that is, the set of students $S = B \times Y$ can be divided in two disjoint sets S_1 and S_2 such that for all $\mathbf{s} \in S_1$ and for all $\mathbf{s}' \in S_2$, either $\mathbf{s}' \geq \mathbf{s}$ or they are not comparable under $\geq$ (where $\geq$ is the usual product order in $\mathbb{R}^2$). Moreover, $\bar{S}_1 = \operatorname{supp} G\left(\mu^*(b, y)\right)$ and $\bar{S}_2 = \operatorname{supp} H\,(b, y)\,.$*

Proof. Let $f\,(\mathbf{s}, \mu) = p_R\,(\mu, y) - p^*\,(\mu, b)\,.$ Student $\mathbf{s} = (b, y)$ prefers the public school if and only if $f\,(\mathbf{s}, \mu) \leq 0$ for all $\mu \in U$. If the set $\{\mathbf{s} \in S : f\,(\mathbf{s}, \mu) = 0 \text{ for all } \mu \in U\}$ is nonempty, it divides the set S in two: the lower set $S_1 \equiv \{\mathbf{s} \in S : f\,(\mathbf{s}, \mu) < 0 \text{ for all } \mu \in U\}$ is the set of students that strictly prefer the public school, and the upper set $S_2 \equiv \{\mathbf{s} \in S : f\,(\mathbf{s}, \mu) > 0 \text{ for some } \mu \in U\}$ is the set of students that strictly prefer to attend some private school (i.e., in the interior of the support of $G\left(\mu^*(b, y)\right)$). Hence, S_1 is the interior[13] of the support of $H\,(b, y)$, and S_2 the interior of the support of $G\left(\mu^*(b, y)\right)$.

On the other hand, since $p_R\,(\mu, y)$ is increasing in y, so is $p_R\,(\mu, y) - p^*\,(\mu, b)$; likewise, since $p^*\,(\mu, b)$ is decreasing in b, $p_R\,(\mu, y) - p^*\,(\mu, b)$ is increasing in b. Hence, $f\,(\mathbf{s}, \mu)$ is increasing in $\mathbf{s}$.

[13] Open relative to $B \times Y$.

If $\mathbf{s}' \in S_2$, then $f(\mathbf{s}', \mu) > 0$ for some $\mu \in U$; call it μ'. If $\mathbf{s} \in S_1$, $f(\mathbf{s}, \mu) < 0$ for all $\mu \in U$, and in particular for μ'. Hence, we have:

$$f(\mathbf{s}, \mu') < 0 < f(\mathbf{s}', \mu'),$$

where $f(\mathbf{s}, \mu)$ is increasing in $\mathbf{s}$; it follows that $\mathbf{s}' \not< \mathbf{s}$: either $\mathbf{s}' > \mathbf{s}$ or they are not comparable under $>$. ■

Assignment within private schools

Theorem 3.5 *In a SS-HQE $\mu^*(b, y)$ is strictly increasing in both arguments. This is to say, there is stratification by income and ability.*

Proof. Within the private sector, the students' choice is determined by:

$$\max_{\mu \in P^{-1}(b)} E[u] = (\mu u(1, y - p^*(\mu, b)) + (1 - \mu) u(0, y - p^*(\mu, b))), \tag{3.69}$$

where in the HQE under consideration, $P^{-1}(B) = U$ for all b.[14] Shortening notation,

$$\max_{\mu \in U} E[u] = u(0, z) + \mu \triangle u(\cdot, z), \tag{3.70}$$

where $\triangle u(\cdot, z) = u(1, z) - u(0, z)$ and $z = y - p^*(\mu, b)$. This yields a "demand curve" (in fact, an optimal application policy for students) $\mu^d = \mu^d(b, y|p^*)$ that coincides with the equilibrium assignment $\mu^*(b, y)$ because it is feasible under the equilibrium admission policies. By the maximum Theorem, $\mu^*(b, y)$ is a continuous function. An interior μ^* satisfies the FOC:

$$\frac{\partial E[u]}{\partial \mu} = E[u_z(\cdot, z)]\left(-\frac{\partial p^*}{\partial \mu}\right) + \triangle u(\cdot, z) = 0. \tag{3.71}$$

[14] In the analysis that follows we assume that U is a convex set; that is, $U = [\mu_{\min}, \mu_{\max}]$. A necessary condition for $U = [\mu_{\min}, \mu_{\max}]$ is that $\mu_H(\mu_{\min}) < \mu_L(\mu_{\max})$, which depends on the parameters λ, π_0, π_1 and θ.

Therefore, for an interior solution the price function must be increasing in μ:

$$\frac{\partial p^*}{\partial \mu} = \frac{\triangle u(\cdot, z)}{E[u_z(\cdot, z)]} > 0. \tag{3.72}$$

Observe that this is also a necessary condition for the HQE to exist (see Lemma (3.9)).

The SOC is:

$$\frac{\partial^2 E[u]}{\partial \mu^2} = -2\frac{\partial p(\mu, b)}{\partial \mu}\Delta u_z(\cdot, z) + E[u_{zz}(\cdot, z)]\left(\frac{\partial p(\mu, b)}{\partial \mu}\right)^2 - E[u_z(\cdot, z)]\frac{\partial^2 p(\mu, b)}{\partial \mu^2} < 0. \tag{3.73}$$

The stratification result is obtained from comparative statics.[15] Taking the derivative of the FOC with respect to y and b respectively, we obtain:

$$\frac{\partial \mu}{\partial y} = \frac{\left(\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\right)\frac{\partial p(\mu,b)}{\partial \mu}}{2\frac{\partial p(\mu,b)}{\partial \mu}\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 + E[u_z(\cdot, z)]\frac{\partial^2 p(\mu,b)}{\partial \mu^2}}, \text{ and}$$

$$\frac{\partial \mu}{\partial b} = \frac{\left(-\Delta u_z(\cdot, z) + E[u_{zz}(\cdot, z)]\frac{\partial p(\mu,b)}{\partial \mu}\right)\frac{\partial p(\mu,b)}{\partial b}}{2\frac{\partial p(\mu,b)}{\partial \mu}\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 + E[u_z(\cdot, z)]\frac{\partial^2 p(\mu,b)}{\partial \mu^2}}. \tag{3.74}$$

Replacing $\frac{\partial p(\mu,b)}{\partial \mu}$ with $\frac{\Delta u(\cdot,z)}{E[u_z(\cdot,z)]}$, and $\frac{\partial p(\mu,b)}{\partial b}$ with $\frac{\partial c(b)}{\partial b}$, we get:

$$\frac{\partial \mu}{\partial y} = \frac{\left(\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\right)\frac{\Delta u(\cdot,z)}{E[u_z(\cdot,z)]}}{2\frac{\partial p(\mu,b)}{\partial \mu}\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 + E[u_z(\cdot, z)]\frac{\partial^2 p(\mu,b)}{\partial \mu^2}}, \text{ and} \tag{3.75}$$

$$\frac{\partial \mu}{\partial b} = \frac{-\left(\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\frac{\Delta u(\cdot,z)}{E[u_z(\cdot,z)]}\right)\frac{\partial c(b)}{\partial b}}{2\frac{\partial p(\mu,b)}{\partial \mu}\Delta u_z(\cdot, z) - E[u_{zz}(\cdot, z)]\left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 + E[u_z(\cdot, z)]\frac{\partial^2 p(\mu,b)}{\partial \mu^2}}. \tag{3.76}$$

Under the assumption that perceived quality is a normal good, the numerators in (3.75) and (3.76) are positive. On the other hand, (3.75) and (3.76) have a common denominator, the negative of the expression in equation (3.73). Therefore, $\frac{\partial \mu}{\partial y}$ and $\frac{\partial \mu}{\partial b}$ are positive if and only if the SOC is met. ■

[15]See Appendix B

Remark 3.6 *The students' problem has an interior solution if and only if there is stratification by income and ability. Stratification by income and ability is the consequence of the combination of achievement being a "normal good", and prices being increasing in* μ *and decreasing in* b*, so it will be present in any equilibrium with those characteristics.*[16] *In particular, it has nothing to do with the fact that we are looking at a steady state.*

Solving for the equilibrium assignment of students

Given the joint distribution of income and ability, F, any assignment function $\mu(b, y)$ induces a distribution over μ, F_μ, where $F_\mu(x)$ tells the proportion of students who attend schools with reputation $\mu \leq x$. Therefore, if κ is the fraction of students attending the public school under an assignment function $\mu(b, y)$, then $F_\mu(0) = \kappa$.

Assume that $f_{b,y}$ is the probability density function (pdf) of F. Let μ^+ denote the positive values of μ, and f_{μ^+} the (truncated) induced pdf over μ^+ (i.e., for private schools). Then,

$$f_{\mu^+}(x) = \begin{cases} 0 & \text{if } \mu = 0, \\ \frac{f_\mu(x)}{1-\kappa} & \text{if } \mu > 0. \end{cases} \tag{3.77}$$

We will derive f_{μ^+} from the equilibrium assignment $\mu^*(y, b)$ and the joint pdf $f_{b,y}$.

Let $f^+_{b,y}$ denote the truncated joint pdf over income and ability with support $S^+ =$

[16] This result contrasts with that of Epple and Romano (1998). In the Epple and Romano model, stratification by ability is obtained if the lagrangian multipliers on the mean ability of the peer group are weakly ascending, which implies a weakly positive relationship between school quality and discounts to ability, in addition to the "weakly single crossing in ability" assumption about students preferences, which is similar to our assumption that the utility function of the students is independent on b. However, the authors claim that neither condition is necessary for stratification by ability, but they do not provide sufficient and necessary conditions.

$\{(b,y) \in B \times Y : \quad \mu_t^* (b,y) > 0\}$ that satisfies:

$$f_{b,y}^{+} (b,y) = \begin{cases} 0 & \text{if } \mu^* (b,y) = 0, \\ \frac{f_{b,y}(b,y)}{1-\kappa} & \text{if } \mu^* (b,y) > 0. \end{cases} \tag{3.78}$$

From condition CE3, the distribution F_{μ^+} induced by the equilibrium assignment function $\mu^* (b,y)$ must equal the distribution of reputations G_μ. To obtain F_{μ^+}, define the following bivariate transformation:

$$\mu = \mu^* (b,y)^+ \text{, and} \tag{3.79}$$

$$\beta = b. \tag{3.80}$$

where $\mu^* (b,y)^+$ is the restriction of $\mu^* (b,y)$ to S^+.

Since $\mu^* (b,y)^+$ is increasing in both income and ability, we can obtain y as a function of μ and b from equation (3.79), $y = h(\beta,\mu)$. Therefore, the inverse transformation:

$$y = h(\beta,\mu) \text{ and} \tag{3.81}$$

$$b = \beta \tag{3.82}$$

is one-to-one. Using this transformation, the joint distribution of μ^+ and β is:

$$f_{\beta,\mu^+} (\beta,\mu) = \begin{cases} \frac{1}{1-\kappa} f_{b,y} (\beta, h(\beta,\mu)) \frac{\partial h(\beta,\mu)}{\partial \mu} & \text{if } (\beta, h(\beta,\mu)) \in (B \times Y)^+, \\ 0 & \text{if not.} \end{cases} \tag{3.83}$$

We are interested in the marginal distribution over μ^+:

$$f_{\mu^+} (\mu) = \int_{-\infty}^{\infty} f_{\mu^+,\beta} (\beta,\mu) \, d\beta$$

$$= \frac{1}{1-\kappa} \int_{\beta_{\min}(\mu)}^{\beta_{\max}(\mu)} f_{b,y} (\beta, h(\beta,\mu)) \frac{\partial h(\beta,\mu)}{\partial \mu} d\beta, \tag{3.84}$$

where $\beta_{\min}(\mu)$ and $\beta_{\max}(\mu)$ are obtained from the condition $(h(\beta,\mu),\beta) \in (B \times Y)^{+}$.

Then, from CE3 the equilibrium assignment of students $\mu^{*}(b,y)$ satisfies:

$$F_{\mu}(x) = \frac{1}{1-\kappa}\int_{0}^{x}\int_{\beta_{\min}(\mu)}^{\beta_{\max}(\mu)} f_{b,y}(\beta, h(\beta,\mu))\frac{\partial h(\beta,\mu)}{\partial \mu} d\beta d\mu = G_{\mu}(x). \tag{3.85}$$

Example 3.1 *Assume that $u(z,a) = z(a+1)$. The consumer's problem is then:*

$$\max_{\{\mu\}} E[u] = z(\mu+1) = (y - p(\mu,b))(\mu+1). \tag{3.86}$$

The MRS is therefore:

$$MRS = \frac{z}{\mu+1}, \tag{3.87}$$

and solving $E[u] = u(0,y)$ we obtain:

$$z' = \frac{y}{\mu+1}, \tag{3.88}$$

and

$$MRS|_{E[u]=u(0,y)} = MRS|_{z=z'} = \frac{y}{(\mu+1)^{2}}. \tag{3.89}$$

The reservation price p_R is therefore:[17]

$$p_R = \int_{0}^{\mu}\frac{y}{(\mu+1)^{2}} = \frac{\mu y}{\mu+1}. \tag{3.90}$$

The FOC for the interior solution is:

$$-\frac{\partial k(\mu)}{\partial \mu}(\mu+1) + (y - p(\mu,b)) = 0 \tag{3.91}$$

$$\Rightarrow \frac{\partial k(\mu)}{\partial \mu} = \frac{y - k(\mu) - c(b)}{(\mu+1)}. \tag{3.92}$$

[17] Alternatively, p_R solves:

$$\begin{aligned} y &= (y - p_R)(\mu+1) \\ &\Rightarrow p_R = \frac{\mu y}{\mu+1} \end{aligned}$$

From the FOC (3.92), any combination of y and b such that $y - c(b)$ is constant results in the same choice of μ for the students who attend private schools. This means that the equilibrium assignment function takes the form $\mu^*(y,b) = \phi(y - c(b))$, where $\phi' > 0$. Therefore, the distribution of μ induced by the equilibrium assignment is:

$$f_{\mu^+}(\mu) = \frac{1}{1-\kappa} \int_{\beta_{\min}(\mu)}^{\beta_{\max}(\mu)} f_{b,y}\left(\beta, \phi^{-1}(\mu) + c(\beta)\right) \frac{\partial \phi^{-1}(\mu)}{\partial \mu} d\beta. \tag{3.93}$$

The function $\phi^{-1}(\mu)$ is obtained from the equilibrium condition $f_{\mu^+}(\mu) = g_\mu(\mu)$ as we describe below for a particular cost function and joint distribution of students (y,b) and reputations.

Assume that $c(b) = 1 - b$, the joint distribution of (b,y) is uniform in $[0,1] \times [5,10]$, and the distribution of μ is uniform in $\left(\frac{1}{4}, \frac{3}{4}\right)$. Then, $p_R\left(\mu = \frac{1}{4}, y = 5\right) = 1$. We assume that the student who attends the lower reputation school pays his reservation price. Therefore, in this case we can obtain:

$$k\left(\mu = \tfrac{1}{4}\right) = 0, \text{ and} \tag{3.94}$$

$$p\left(\mu = \tfrac{1}{4}, b = 0\right) = 1. \tag{3.95}$$

All students choose private schools, and $B \times Y = [0,1] \times [5,10]$. The one-period gain for the schools $\mu = \frac{1}{4}$ is $k\left(\mu = \frac{1}{4}\right) = 0$, and therefore, all schools obtain positive earnings in the high quality equilibrium.

We consider the following increasing transformation:

$$\mu = \phi(y - 1 + b) \text{ and} \tag{3.96}$$

$$b = \beta. \tag{3.97}$$

From (3.93) we obtain:

$$f_{\mu}(\mu) = \int_{\beta_{\min}(\mu)}^{\beta_{\max}(\mu)} f_{b,y}\left(\beta, \phi^{-1}(\mu) + 1 - \beta\right) \frac{\partial \phi^{-1}(\mu)}{\partial \mu} d\beta, \tag{3.98}$$

where $\beta_{\max}(\mu)$ *and* $\beta_{\min}(\mu)$ *are obtained from the conditions* $\left(\phi^{-1}(\mu) + 1 - \beta\right) \in [5, 10]$ *and* $\beta \in [0,1]$. *The condition* $\left(\phi^{-1}(\mu) + 1 - \beta\right) > 5$ *is relevant only if* $\phi^{-1}(\mu) - 4 < 1$, *that is, only if* $\mu < \phi(5)$. *The condition* $\left(\phi^{-1}(\mu) + 1 - \beta\right) < 10$ *is relevant only if* $\phi^{-1}(\mu) - 9 > 0$, *that is, if* $\mu > \phi(9)$. *Therefore, the marginal density of* μ *is:*

$$f_{\mu}(x) = \begin{cases} \frac{1}{5}\int_0^{\phi^{-1}(\mu)-4} \frac{\partial \phi^{-1}(\mu)}{\partial \mu} d\beta = \frac{1}{5}\left(\phi^{-1}(\mu) - 4\right) \frac{\partial \phi^{-1}(\mu)}{\partial \mu} & \text{if } \mu < \phi(5), \\ \frac{1}{5}\int_0^{1} \frac{\partial \phi^{-1}(\mu)}{\partial \mu} d\beta = 5 \frac{\partial \phi^{-1}(\mu)}{\partial \mu} & \text{if } \phi(5) < \mu < \phi(9), \\ \frac{1}{5}\int_{\phi^{-1}(\mu)-9}^{1} \frac{\partial \phi^{-1}(\mu)}{\partial \mu} d\beta = \frac{1}{5}\left(10 - \phi^{-1}(\mu^*)\right) \frac{\partial \phi^{-1}(\mu)}{\partial \mu} & \text{if } \mu > \phi(9). \end{cases} \tag{3.99}$$

Using the equilibrium condition $f_{\mu}(\mu) = g_{\mu}(\mu) = 2$, *we obtain three differential equations:*

$$\left(\phi^{-1}(\mu) - 4\right) \frac{\partial \phi^{-1}(\mu)}{\partial \mu} = 10 \quad \text{if } \mu < \phi(5), \tag{3.100}$$

$$\frac{\partial \phi^{-1}(\mu)}{\partial \mu} = 10 \quad \text{if } \phi(5) < \mu < \phi(9), \text{ and} \tag{3.101}$$

$$\left(10 - \phi^{-1}(\mu^*)\right) \frac{\partial \phi^{-1}(\mu)}{\partial \mu} = 10 \quad \text{if } \mu > \phi(9). \tag{3.102}$$

Solving (3.100), (3.101) and (3.102), we obtain the following equilibrium assignment of students, which is illustrated in Figure 3.2:

$$\mu^*(b, y) = \begin{cases} \frac{(y+b-5)^2+5}{20} & \text{if } \mu < 0.3 \\ \frac{y+b-3}{10} & \text{if } 0.3 < \mu < 0.7 \\ \frac{15-(y+b-11)^2}{20} & \text{if } \mu > 0.7 \end{cases} . \tag{3.103}$$

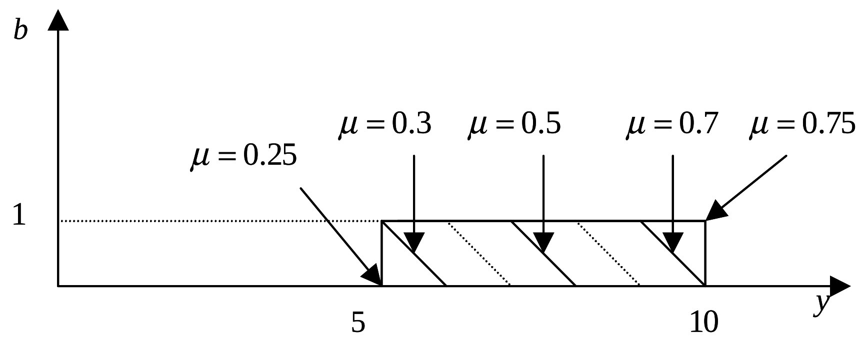

Figure 3.2: Equilibrium allocation of students in Example 3.1

Solving for the equilibrium price function

To obtain the equilibrium price function $p^*(\mu, b) = k(\mu) + c(b)$ we use the FOC of the students' maximization problem, evaluated at the equilibrium assignment.[18]

Example 3.2 *In the Cobb-Douglas example above, the price function $p(\mu, b) = k(\mu) + c(b)$ is obtained solving the differential equations obtained from the FOC:*

$$\frac{\partial k(\mu)}{\partial \mu} = \frac{\left(\phi^{-1}(\mu) - k(\mu)\right)}{(\mu+1)} = \begin{cases} \frac{\left(\sqrt{20\mu-5}+4-k(\mu)\right)}{(\mu+1)} & \text{if } \mu < 0.3, \\ \frac{(2+10\mu-k(\mu))}{(\mu+1)} & \text{if } 0.3 < \mu < 0.7, \\ \frac{\left(10-\sqrt{15-20\mu}-k(\mu)\right)}{(\mu+1)} & \text{if } \mu > 0.7. \end{cases} \tag{3.104}$$

Solving (3.104) with $k\left(\mu = \frac{1}{4}\right) = 0$, we obtain:

$$k(\mu) = \begin{cases} -\frac{1}{\mu+1}\left(1 - 4\mu - \frac{1}{6}\sqrt{5}(4\mu-1)^{\frac{3}{2}}\right) & \text{if } \mu < 0.3, \\ -\frac{1}{\mu+1}\left(\frac{2.45}{3} - 2\mu - 5\mu^2\right) & \text{if } 0.3 < \mu < 0.7, \\ -\frac{1}{\mu+1}\left(4 - 10\mu - \frac{1}{6}\sqrt{5}(3-4\mu)^{\frac{3}{2}}\right) & \text{if } \mu > 0.7. \end{cases} \tag{3.105}$$

[18]Note that when we consider an allocation of students increasing in b, the price function satisfies the SOC of the students' maximization problem.

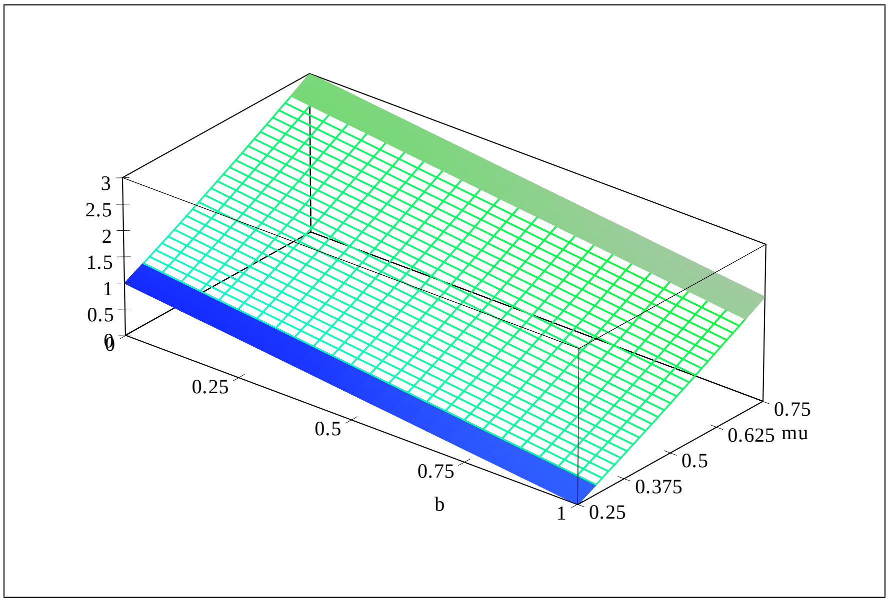

Figure 3.3: Equilibrium price function $p(\mu, b)$ in Example 3.2

Therefore, we obtain the following equilibrium price function, which is illustrated in Figure 3.3:

$$p(\mu, b) = \begin{cases} -\frac{1}{\mu+1}\left(1 - 4\mu - \frac{1}{6}\sqrt{5}\,(4\mu-1)^{\frac{3}{2}}\right) + 1 - b & \text{if } \mu < 0.3, \\ -\frac{1}{\mu+1}\left(\frac{2.45}{3} - 2\mu - 5\mu^2\right) + 1 - b & \text{if } 0.3 < \mu < 0.7, \\ -\frac{1}{\mu+1}\left(4 - 10\mu - \frac{1}{6}\sqrt{5}\,(3-4\mu)^{\frac{3}{2}}\right) + 1 - b & \text{if } \mu > 0.7. \end{cases} \quad (3.106)$$

3.4.3 Welfare

There is something peculiar about this economy: production costs depend on consumers characteristics: $c(b)$. This fact complicates a bit the partial equilibrium Pareto analysis. Without it, we would simply say that any *ex-post* efficient assignment of academic achievement would give the $\theta\%$ available of $a = 1$ to the ones that value it the most (the highest income students in our case, as academic achievement is assumed to be a normal good), and produce it at minimum cost (that is, with the $\theta\%$ highest ability students.) This is so because the game is of transferable utility, and consequently the Pareto optimal assignment is the one that maximizes total surplus. However, the previous assignment is not possible: there are low-ability students among those of higher willingness to pay. The maximization of total surplus then would consist of finding the subset of $B \times Y$ of measure θ that maximized the expected value of $p_R(y) - c(b)$. This is still a convex set in the upper part of $B \times Y$, whose lower frontier is given by the equation $p_R(y) - c(b) = \varpi$ for some constant $\varpi > 0$, i.e., $\{(b, y) : p_R(y) - c(b) \geq \varpi\}$.

It is easy to see that this is not the SS-HQE assignment, because it is based on reputation and not achievement. Since G^I has full support on U, there is always positive measure of the set $\{(b, y) : p_R(y) - c(b) \geq C\}$ that attends inept schools, thereby getting $a = 0$. However, if the social planner were as informationally constrained as students are, then it would choose the assignment $\mu^*(b, y)$. Hence, the SS-HQE assignment is constrained efficient in that sense.

3.4.4 Existence

The question of existence of a SS-HQE bears on the characteristics of the (decreasing, twice differentiable) cost function $c(b)$ of a particular economy. The game will not have a SS-HQE if $c(b)$ is too high, i.e., it does not respect the bound of Lemma (3.7):

$$c(b) \leq \delta(1 - \lambda)(\pi_1 - \pi_0)(v_C(\mu_H(\mu)) - v_C(\mu_L(\mu))) \tag{3.107}$$

for all $b \in B(\mu)$, and for all μ. Since $(\mu_H(\mu) - \mu_L(\mu)) > 0$ for all $\mu \in (0,1)$, and $\lim_{\mu\to 0}(\mu_H(\mu) - \mu_L(\mu)) = 0$ and $\lim_{\mu\to 1}(\mu_H(\mu) - \mu_L(\mu)) = 0$, the set of possible reputations U –which we already know to be bounded away from 0 and 1– must be *sufficiently* far away from those borders. Recall that Section 4 established that the long run reputation distribution G has a support $U \subseteq [\mu_{\min}, \mu_{\max}] \subset [\theta\lambda, 1-\lambda+\lambda\theta]$, where $\mu_{\min}$ and $\mu_{\max}$ were defined as:

$$\mu_{\min} = \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\mu_{\min}}{(1-\pi_1)\mu_{\min} + (1-\pi_0)(1-\mu_{\min})}, \text{ and} \tag{3.108}$$

$$\mu_{\max} = \lambda\theta + (1-\lambda)\frac{\pi_1\mu_{\max}}{\pi_1\mu_{\max} + \pi_0(1-\mu_{\max})}. \tag{3.109}$$

As $p(\mu,b)$ is increasing in μ, so is $v_C(\mu)$. Therefore, $(v_C(\mu_H(\mu)) - v_C(\mu_L(\mu))) > 0$ for all $\mu \in U$, and it follows that:

Theorem 3.6 *There exist differentiable, strictly decreasing cost functions $c(b) > 0$ such that for them SS-HQE exists.*

Summary 3.4 *The SS-HQE $\langle p^*(\mu,b), \mu^*(b,y), B, 1\rangle$ is defined by the following system of equations:*

$$\begin{aligned}
(P^*(\mu), a^*) &= (B,1) = \arg\max_{a,b}\{p^*(\mu,b) - c(b) + \delta(1-\lambda)(\pi_1 v_C(\mu_H) + (1-\pi_1)v_C(\mu_L))\}, \\
v_I(\mu) &= E_{b\in B(\mu)}[p^*(\mu,b)] + \delta(1-\lambda)(\pi_0 v_I(\mu_H) + (1-\pi_0)v_I(\mu_L)) \geq 0 \qquad \forall\mu\in U, \\
\mu_{t+1} &= \begin{cases} \mu_H \equiv \lambda\theta + (1-\lambda)\frac{\pi_1\mu_t}{\pi_1\mu_t + \pi_0(1-\mu_t)} & \text{if } r_t = H \\ \mu_L \equiv \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\mu_t}{(1-\pi_1)\mu_t + (1-\pi_0)(1-\mu_t)} & \text{if } r_t = L \end{cases}, \\
\mu^*(b,y) &= \arg\max_{\mu\in U} \mu u(1, y - p^*(\mu,b)) + (1-\mu)u(0, y-p^*(\mu,b)), \\
p^*(\mu,b) &= k(\mu) + c(b), \text{ and} \\
F(b,y) &= \kappa H(b,y) + (1-\kappa)G(\mu^*(b,y)) \qquad \forall b,y.
\end{aligned}$$

In the steady state of a HQE the price function $p^*(\mu, b)$ *is additively separable, increasing in* μ *and decreasing in* b*: reputations are market-priced, and so is ability. The assignment exhibits stratification by income and ability:* $\mu^*(b, y)$ *is monotonically increasing, so that the richer, more able students attend the better-reputed schools.*

3.5 Policy Effects in the Steady-State HQE

The prevailing view in the literature is that market-based educational systems exhibit stratification by income and ability because of the existence of *peer effects*, i.e., an externality among classmates according to which educational achievement is increasing in the average ability of classmates (e.g., Epple and Romano, 1998). Richer students would be willing to pay more to mix themselves with the more able students, leaving "behind" the poorer and less talented. In the extreme, private schools could be viewed as mere facilitators of this sorting.

Stratification is judged by many to be an undesirable outcome. There are many different reasons for this. Stratification at the school level could produce social disintegration. It might exacerbate or make endure income inequality if educational achievement is a major determinant of human capital accumulation. It might be judged to be unfair in itself, because economic opportunities would be unequally distributed from the beginning of humans' lives (and especially so if the availability of more resources at the early childhood makes the accumulation of educational ability more likely). As a consequence, different policy measures have been proposed to revert or preclude this outcome, ranging from designing differentiated demand-subsidies (vouchers) to statization of the whole educational system.

This Section analyzes some of those proposals. It is worth stressing that the consequences of such policies may be markedly different in a model like Epple and Romano's and ours. Within the limits of a model where schools produce nothing but sorting, each of these proposals will be judged according to its efficacy in attaining the desired social integration. However, in models like the one presented here, where schools can provide different levels of education quality at differential costs, efficiency becomes an important issue too.

Stratification by income and ability follows from the assumption that perceived quality is a normal good along with the shape of the price function faced by students. The first kind of stratification is the consequence of the price being increasing in μ, while the second is the consequence of the price being decreasing in b, as equations (3.75) and (3.76) show.

We will analyze two different proposals designated to avoid stratification in a voucher system. To avoid stratification by income, any additional co-payments from students to schools are forbidden in both cases: the voucher is the only revenue source for schools. The first proposal can be described as follows:

1. The State gives a uniform voucher p to every student,

2. Any additional co-payments from students to schools are forbidden.

3. Schools cannot select students, but they are matched to schools at random through a lottery system.

Lotteries then replace the market-based endogenous matching implemented by prices and admission policies.

It is easy to see what the consequence of such a policy is in our model: if schools cannot benefit from their reputation, either receiving better students or charging higher prices, they don't have an incentive to exert the costly effort required to provide a high educational achievement. The value function for competent schools would become:

$$\begin{aligned} v_C(\mu) &= \max_{a\in\{0,1\}} \{(1-a)\left[p+\delta(1-\lambda)\left[\pi_0 v_C(\mu_H)+(1-\pi_0)v_C(\mu_L)\right]\right] \\ &\quad +a\left[p-c(b)+\delta(1-\lambda)\left[\pi_1 v_C(\mu_H)+(1-\pi_1)v_C(\mu_L)\right]\right]\}. \end{aligned} \tag{3.110}$$

Since p no longer depends on μ, nor does the value function:

$$v_C = \max_{a\in\{0,1\}} \{(1-a)\left[p+\delta(1-\lambda)v_C\right]+a\left[p-c(b)+\delta(1-\lambda)v_C\right]\}. \tag{3.111}$$

Clearly $a=0$ is preferred to $a=1$:

$$p+\delta(1-\lambda)v_C > p-c(b)+\delta(1-\lambda)v_C \qquad \forall b\in B. \tag{3.112}$$

Hence, the economy would undoubtedly revert to the low-quality equilibrium, whereby $\alpha^*(\mu)=0$ for all μ, and private schools would participate if and only if the voucher is set high enough so that they can profit[19]:

$$\begin{aligned} v_C &= p+\delta(1-\lambda)v_C \\ &\Rightarrow \quad v_C = \frac{p}{\delta(1-\lambda)} > 0. \end{aligned} \tag{3.113}$$

[19]Recall that the cost of low quality was normalized to 0; p in this case in not the amount of the voucher itself, but the margin over (low quality) production costs. Hence, p may well be negative if the subsidy is too low. In that case, the private sector will not participate in education.

The second proposal we analyze is due to Epple and Romano (2002), who apply the model developed in Epple and Romano (1998) to study a voucher design that would reach the benefits of a voucher program without cream skimming, in order to provide high and equal-quality education to all students. They propose a type-dependent voucher with no extra charges allowed. That is, the voucher level is higher if the school serves low ability students, and the schools must accept the voucher as the only source of financing.

Assume that differences in the voucher level just reflects differences in production costs; that is, the voucher is $p(b) = k + c(b)$. Then, the value function for competent schools would become:

$$\begin{aligned} v_C(\mu) &= \max_{b \in B, a \in \{0,1\}} \{(1-a)[k + c(b) + \delta(1-\lambda)[\pi_0 v_C(\mu_H) + (1-\pi_0) v_C(\mu_L)]] \\ &\quad + a[k + \delta(1-\lambda)[\pi_1 v_C(\mu_H) + (1-\pi_1) v_C(\mu_L)]]\}. \end{aligned} \tag{3.114}$$

But again p no longer depends on μ, and therefore we obtain:

$$v_C = \max_{b \in B, a \in \{0,1\}} \{(1-a)[k + c(b) + \delta(1-\lambda) v_C] + a[k + \delta(1-\lambda) v_C]\}, \tag{3.115}$$

and $a = 0$ is preferred to $a = 1$:

$$k + c(b) + \delta(1-\lambda) v_C > k + \delta(1-\lambda) v_C \quad \forall b \in B. \tag{3.116}$$

As in the previous case, the economy would revert to the low-quality equilibrium, whereby $\alpha^*(\mu) = 0$ for all μ. Notice that if schools were to produce high quality, then schools would be indifferent among all students, but all students would strictly prefer the school with the highest reputation. However, as the economy reverts to the low quality equilibrium, students are indifferent among all schools, but all schools strictly prefer the less able students. This occurs because, as long as the cost of producing low quality does not depend on b, there is no reason to pay a higher voucher to the schools that serve low ability students in the low quality equilibrium.

It is clear then that both lotteries and type-dependent vouchers would increase equality, but most likely not in the way its supporters predict: if schools cannot charge fees in addition to the voucher, the economy would waste all of its capacity for producing academic excellence. This outcome is certainly inefficient, because the fact that the economy was in a HQE before means that the valuation of high achievement more than compensated its cost.

3.6 Concluding Remarks

The main result of this Chapter is that in the high quality equilibrium, the assignment of students in private schools is increasing in income and ability, a result similar to the stratification by income and ability found by Epple and Romano (1998). A well-known conclusion of Epple and Romano model is that the answer to the question “who gains, who loses, and how does it add up?” with a voucher program, may be that due to the “cream skimming” effect, students with relatively low income and ability who remain in the public sector lose, while richer and higher ability students that leave the public sector gain, relative to a market left alone without the voucher.

In this Chapter we presented a model of long run reputation with imperfect public monitoring, and we analyzed the “high quality equilibrium”, where all private competent schools make costly effort, and their students obtain high educational achievement. Under the scenario described in this model, the answer to the question “who gains, who loses, and how does it add up?" with a voucher program, may be very different to the conclusion arrived at by Epple and Romano. Assume that we start with an educational system with a large public sector, and the voucher program promotes the entry of private schools (both type C and type I schools) without additional public funds. Then, the high quality equilibrium described in this Chapter may be attained if the schools are allowed to charge fees in addition to the voucher. In this case, all students benefit from the possibility of attending high quality schools and nobody looses with the voucher program, since there are no peer effects.

Education is essentially a social activity and consequently it is beyond any doubts that it is plagued by all sorts of externalities or peer effects (although their particular form might diverge substantially from what has been considered so far in the economics literature.) Nevertheless, we just don't know how important they are. Moreover, this Chapter shows that those patterns in the data that the predominant view in the literature interprets as evidence of peer effects are also consistent with the information-based explanation developed here. Unfortunately the outcome and welfare effects of popular State interventions in the educational system differ substantially under both theories, so it appears to be very valuable to attempt to disentangle reputational from externality effects.

Appendix A

Comparative Statics: Students

Taking the derivative of the FOC of the student's maximization problem (Equation (3.69))with respect to y we obtain:

$$\begin{aligned}0 = u_z(z,1) - u_z(z,0) - (\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)) \frac{\partial p(\mu,b)}{\partial \mu} \\ - 2\frac{\partial p(\mu,b)}{\partial \mu}(u_z(z,1) - u_z(z,0)) \frac{\partial \mu}{\partial y} \\ + (\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)) \left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 \frac{\partial \mu}{\partial y} \\ - (\mu u_{zz}(z,1) + (1-\mu) u_z(z,0)) \frac{\partial^2 p(\mu,b)}{\partial \mu^2} \frac{\partial \mu}{\partial y}. \quad \text{(A.1)}\end{aligned}$$

Define Θ as:

$$\begin{aligned}\Theta \equiv -2\frac{\partial p(\mu,b)}{\partial \mu}(u_z(z,1) - u_z(z,0)) \\ + (\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)) \left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 \\ - (\mu u_z(z,1) + (1-\mu) u_z(z,0)) \frac{\partial^2 p(\mu,b)}{\partial \mu^2}. \quad \text{(A.2)}\end{aligned}$$

Therefore, $\frac{\partial \mu}{\partial y}$ is:

$$\begin{aligned}\frac{\partial \mu}{\partial y} &= \frac{u_z(z,1) - u_z(z,0) - (\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)) \frac{\partial p(\mu,b)}{\partial \mu}}{-\Theta} \\ &= \frac{u_z(z,1) - u_z(z,0) - (\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)) \frac{u(z,1)-u(z,0)}{\mu u_z(z,1)+(1-\mu)u_z(z,0)}}{-\Theta} \quad \text{(A.3)}\end{aligned}$$

If the SOC is satisfied and under the assumption that a is a normal good, we obtain that $\frac{\partial \mu}{\partial y} > 0$ (see conditions (3.73) and (3.57)). Therefore, the assignment $\mu(y, b)$ is increasing in y (richer students choose schools with better reputation).

Taking the derivative of the FOC with respect to b we obtain:

$$\begin{aligned} 0 = &- \left(u_z(z,1) - u_z(z,0)\right) \frac{\partial p(\mu,b)}{\partial b} \\ &- 2\left(u_z(z,1) - u_z(z,0)\right) \frac{\partial p(\mu,b)}{\partial \mu} \frac{\partial \mu}{\partial b} \\ &+ \left(\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)\right) \frac{\partial p(\mu,b)}{\partial b} \frac{\partial p(\mu,b)}{\partial \mu} \\ &+ \left(\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)\right) \left(\frac{\partial p(\mu,b)}{\partial \mu}\right)^2 \frac{\partial \mu}{\partial b} \\ &- \left(\mu u_z(z,1) + (1-\mu) u_z(z,0)\right) \frac{\partial^2 p(\mu,b)}{\partial \mu \partial b} \\ &- \left(\mu u_z(z,1) + (1-\mu) u_z(z,0)\right) \frac{\partial^2 p(\mu,b)}{\partial \mu^2} \frac{\partial \mu}{\partial b}. \end{aligned} \tag{A.4}$$

And therefore $\frac{\partial \mu}{\partial b}$ is:

$$\begin{aligned} \frac{\partial \mu}{\partial b} = &\frac{\left(-\left(u_z(z,1) - u_z(z,0)\right) + \left(\mu u_{zz}(z,1) + (1-\mu) u_{zz}(z,0)\right) \frac{\partial p(\mu,b)}{\partial \mu}\right) \frac{\partial p(\mu,b)}{\partial b}}{-\Theta} \\ &+ \frac{\left(\mu u_z(z,1) + (1-\mu) u_z(z,0)\right) \frac{\partial^2 p(\mu,b)}{\partial \mu \partial b}}{\Theta}. \end{aligned} \tag{A.5}$$

Since $\frac{\partial p(\mu,b)}{\partial b} < 0$, conditions (3.73) and (3.57) imply that the first addend in (A.5) is positive. The sign of the second addend depends on $\frac{\partial^2 p(\mu,b)}{\partial \mu \partial b}$. If $\frac{\partial^2 p(\mu,b)}{\partial \mu \partial b} \leq 0$, then $\frac{\partial \mu}{\partial b} > 0$ since $\frac{\partial u(z,a)}{\partial z} > 0$.

Chapter 4

The evolution of schools' reputations in a competitive school market with imperfect public monitoring

In this Chapter we analyze the way schools create and maintain their reputation in a competitive school market with imperfect public monitoring. We examine the "high quality equilibrium", that is, an equilibrium where all competent schools provide a high educational achievement to their students. We describe the dynamics of each school's reputation, and of the distribution of schools' reputation as well. The main Theorem of this Chapter establishes that the distribution of schools' reputation converges to a "long run distribution", the fixed point of the dynamical system that describes the evolution of distributions.

4.1 Belief updating or the building of schools' reputation

The model considers a continuum of students (short run players) and schools (long run players). Schools are characterized by their type (competent, C, or inept I). The fraction of competent schools is $\theta \in (0,1)$. Competent schools choose the educational achievement provided to their students, $a \in \{0,1\}$. Inept schools, however, cannot choose the educational achievement: they always provide $a = 0$. In the high quality equilibrium, all competent schools choose $a = 1$.

Students do not observe schools' effort nor educational achievement of previous generations of students, but they observe the history of results in a standardized test. Those test scores act as an imperfect signal of the educational achievement provided by schools. The probability of obtaining a high test result is higher when the school provides $a = 1$. Indeed, the probability of obtaining a high test score at t is given by:

$$\Pr\left(H|a_t\right) = \begin{cases} \pi_1 \in (0,1) & \text{if } a_t = 1, \\ \pi_0 \in (0,1) & \text{if } a_t = 0, \end{cases} \tag{4.1}$$

where $\pi_1 > \pi_0$.

Students' beliefs, μ, refer to the probability they assign to each school being of type C, or the "school's reputation". Therefore, in the high quality equilibrium students' beliefs also refer to the probability they assign to obtaining a high educational achievement in a particular school. Since test scores are the only school characteristic that the students observe, μ can only depend on past test scores. All students (perfectly) observe the same test scores, and their prior beliefs are common, so their posterior beliefs are also common. But as schools can differ in their history (past results obtained), they differ in their reputations almost surely. We abuse terms and refer to a particular school with reputation μ (or group of them) as "the school μ", although reputation is almost surely not constant across time.

Students are never completely sure about the school type: there is an exogenous probability $\lambda \in (0,1)$ that a school leaves the market each period and is replaced by another school, which can be of either type. The probability of being replaced by a competent school is θ, also exogenous. Therefore, we assume that the fraction of competent schools is constant over time. The new school inherits the reputation of the old school it replaces. In other words, λ is the probability that a given school changes its owner or management, and θ is the probability that the new owner or management is competent. Under this scenario the uncertainty about types is continually replenished, since there is always a positive probability that a school changes its type. Without a mechanism such as this possibility of replacement, there is no room for permanent or long run reputations in a model with imperfect monitoring, as Cripps, Mailath and Samuelson (2004) establish. Students do not observe the replacement, but they know it can occur (and they take this into account when updating beliefs).

If μ is the prior belief about school type at any given time, μ_H denotes the posterior belief after a high test score, and μ_L the posterior belief after a low test score. Bayesian updating of beliefs implies:

$$\mu_H(\mu) = \lambda\theta + (1-\lambda)\frac{\pi_1\mu}{\pi_1\mu + \pi_0(1-\mu)}, \tag{4.2}$$

and

$$\mu_L(\mu) = \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\mu}{(1-\pi_1)\mu + (1-\pi_0)(1-\mu)}. \tag{4.3}$$

Both $\mu_H(\mu)$ and $\mu_L(\mu)$ are increasing in μ, but $\mu_H(\mu)$ is a concave function of μ, whereas $\mu_L(\mu)$ is convex, as Figure 4.1 shows. Furthermore, $\mu_H(0) = \mu_L(0) = \lambda\theta$, $\mu_H(1) = \mu_L(1) = \lambda\theta + (1-\lambda)$, and $\mu_H(\mu) > \mu_L(\mu)$ for all $\mu \in (0,1)$. Therefore, even if μ is 0 or 1 in the first period, the posterior belief in all subsequent periods is higher than $\lambda\theta$ and lower than $\lambda\theta + (1-\lambda)$. In other words, the possibility of replacements bounds the posterior beliefs away from 0 and 1, thus students are never completely sure about the school type.

If we analyze Figure 4.1 as a dynamical system, the prior belief is the reputation

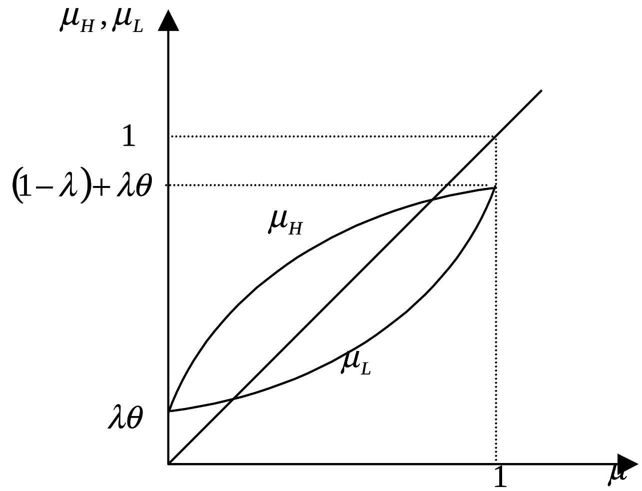

Figure 4.1: μ_H and μ_L as a function of μ.

at period t, and the posterior is the reputation at $t+1$. The reputation at $t+1$ is a random variable: it takes the value $\mu_{t+1} = \mu_H(\mu_t)$ in the event $\{r_t = H\}$, and $\mu_{t+1} = \mu_L(\mu_t)$ in $\{r_t = L\}$. As an example, if $\mu_t = x$ as in Figure 4.2, after a high test score the posterior belief will be $\mu_{t+1} = \mu_H(x)$; after two high test scores, $\mu_{t+2} = \mu_H(\mu_H(x))$, and so on.

Define $\mu_{\min}$ and $\mu_{\max}$ as the values that solve the following equations:

$$\mu_{\min} = \lambda\theta + (1-\lambda)\frac{(1-\pi_1)\mu_{\min}}{(1-\pi_1)\mu_{\min} + (1-\pi_0)(1-\mu_{\min})}, \text{ and} \tag{4.4}$$

$$\mu_{\max} = \lambda\theta + (1-\lambda)\frac{\pi_1\mu_{\max}}{\pi_1\mu_{\max} + \pi_0(1-\mu_{\max})}. \tag{4.5}$$

Then, $\lambda\theta < \mu_{\min} < \mu_{\max} < \lambda\theta + (1-\lambda)$. The posterior belief is $\mu_{t+1} > \mu_t$ for all $\mu_t < \mu_{\min}$, and it is $\mu_{t+1} < \mu_t$ for all $\mu_t > \mu_{\max}$, as can be seen in Figure 4.2. Therefore, if U_t is the support of the school reputations at t, then $\inf U_{t+1} > \inf U_t$ as long as $\inf U_t < \mu_{\min}$, and $\sup U_{t+1} < \sup U_t$ as long as $\sup U_t > \mu_{\max}$.

If $\mu_t \in (\mu_{\min}, \mu_{\max})$, then $\mu_{\min} < \mu_L(\mu_t) < \mu_t < \mu_H(\mu_t) < \mu_{\max}$. Therefore, the

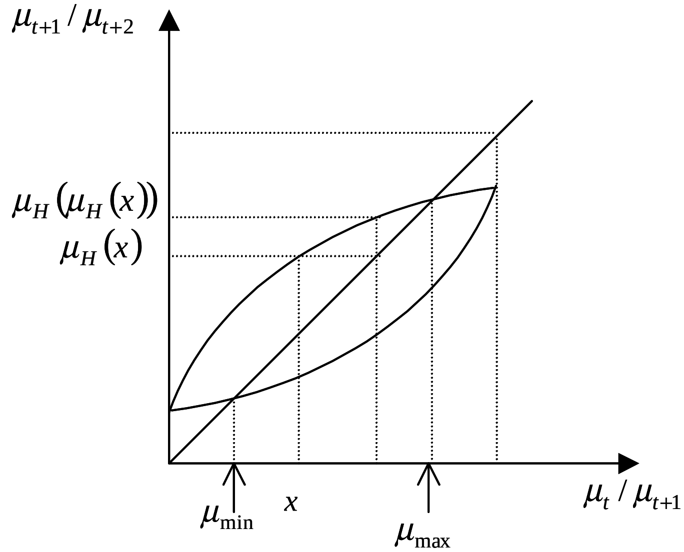

Figure 4.2: Dynamics of μ_t

school reputation will be improved after a high test score and will be deteriorated after a low test score.

4.2 Evolution of the population distribution of schools' reputation

We denote by $G_t^C(\mu)$ the distribution of reputations in period t for competent schools, and by $G_t^I(\mu)$ for inept schools. That is, $G_t^\tau(x)$ says what is the proportion of schools of type τ with reputation $\mu < x$ in period t. The overall distribution function of schools' reputations at t is then:

$$G_t(x) = \theta G_t^C(x) + (1-\theta) G_t^I(x) . \tag{4.6}$$

From t to $t+1$ beliefs are updated according to the test scores obtained by each school, and a fraction λ of existing schools leave the market and with probability θ are replaced by competent schools. Therefore, at $t+1$ a fraction $(1-\lambda)$ of competent schools remains, a fraction $\lambda\theta$ of competent schools leave the market but are replaced by new competent schools, and a fraction $\lambda\theta$ of inept schools leave the market and are replaced by new competent schools. Hence the pool of competent schools at $t+1$ includes previously existing competent schools, and new competent schools that replaced either old competent or old inept schools. The probability that the posterior belief is less than x for a competent school at $t+1$ with prior reputation μ_t is then:

$$\begin{aligned} G^C_{t+1}(x|\mu_t) &= \Pr\left(\mu_{t+1} < x|\mu_t\right) \\ &= (1-\lambda+\lambda\theta)\Pr\left(\mu_{t+1} < x|\mu_t, C\right) + \frac{\lambda\theta(1-\theta)}{\theta}\Pr\left(\mu_{t+1} < x|\mu_t, I\right) \\ &= (1-\lambda+\lambda\theta)\left[I_{\{\mu_H < x|\mu_t\}}\pi_1 + I_{\{\mu_L < x|\mu_t\}}(1-\pi_1)\right] \\ &\quad + \lambda(1-\theta)\left[I_{\{\mu_H < x|\mu_t\}}\pi_0 + I_{\{\mu_L < x|\mu_t\}}(1-\pi_0)\right], \end{aligned} \tag{4.7}$$

where $I_{\{E\}}$ is an indicator function, taking the value 1 if event E occurs, and 0 if not. Using the definitions of μ_H and μ_L in equations (4.2) and (4.3) respectively, we get:

$$\mu_H(\mu_t) < x \Leftrightarrow \mu_t < \frac{(x-\lambda\theta)\pi_0}{((1-\lambda+\lambda\theta-x)\pi_1 + (x-\lambda\theta)\pi_0)} = \mu_H^{-1}(x)\text{, and} \tag{4.8}$$

$$\mu_L(\mu_t) < x \Leftrightarrow \mu_t < \frac{(x-\lambda\theta)(1-\pi_0)}{((1-\lambda+\lambda\theta-x)(1-\pi_1) + (x-\lambda\theta)(1-\pi_0))} = \mu_L^{-1}(x)\,. \tag{4.9}$$

Therefore, we obtain:

$$\begin{aligned} G^C_{t+1}(x|\mu_t) &= (1-\lambda+\lambda\theta)\left[I_{\left\{\mu_t < \mu_H^{-1}(x)\right\}}\pi_1 + I_{\left\{\mu_t < \mu_L^{-1}(x)\right\}}(1-\pi_1)\right] \\ &\quad + \lambda(1-\theta)\left[I_{\left\{\mu_t < \mu_H^{-1}(x)\right\}}\pi_0 + I_{\left\{\mu_t < \mu_L^{-1}(x)\right\}}(1-\pi_0)\right]. \end{aligned} \tag{4.10}$$

By the law of iterated expectations, we obtain the (unconditional) cdf of reputations for the competent schools at $t+1$ as:

$$
\begin{aligned}
G_{t+1}^{C}(x) &\equiv \Pr\left(\mu_{t+1} < x\right) \\
&= (1-\lambda+\lambda\theta)\left[\pi_1 \int_0^{\mu_H^{-1}(x)} \mathrm{d}G_t^C + (1-\pi_1)\int_0^{\mu_L^{-1}(x)} \mathrm{d}G_t^C\right] \\
&\quad + \lambda(1-\theta)\left[\pi_0 \int_0^{\mu_H^{-1}(x)} \mathrm{d}G_t^I + (1-\pi_0)\int_0^{\mu_L^{-1}(x)} \mathrm{d}G_t^I\right] \\
&= (1-\lambda+\lambda\theta)\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right) + (1-\pi_1)\, G_t^C\left(\mu_L^{-1}(x)\right)\right] \\
&\quad + \lambda(1-\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right) + (1-\pi_0)\, G_t^I\left(\mu_L^{-1}(x)\right)\right]. \qquad (4.11)
\end{aligned}
$$

Similarly, for inept schools we obtain:

$$
\begin{aligned}
G_{t+1}^{I}(x|\mu_t) &= \Pr\left(\mu_{t+1} < x|\mu_t\right) \\
&= (1-\lambda\theta)\Pr\left(\mu_{t+1} < x|\mu_t, I\right) + \frac{\lambda(1-\theta)\,\theta}{(1-\theta)}\Pr\left(\mu_{t+1} < x|\mu_t, C\right) \\
&= (1-\lambda\theta)\left[I_{\left\{\mu_t<\mu_H^{-1}(x)\right\}}\pi_0 + I_{\left\{\mu_t<\mu_L^{-1}(x)\right\}}(1-\pi_0)\right] \\
&\quad + \lambda\theta\left[I_{\left\{\mu_t<\mu_H^{-1}(x)\right\}}\pi_1 + I_{\left\{\mu_t<\mu_L^{-1}(x)\right\}}(1-\pi_1)\right]. \qquad (4.12)
\end{aligned}
$$

Using the law of iterated expectations, we obtain:

$$
\begin{aligned}
G_{t+1}^{I}(x) &\equiv \Pr\left(\mu_{t+1} < x\right) \\
&= (1-\lambda\theta)\left[\pi_0 \int_0^{\mu_H^{-1}(x)} \mathrm{d}G_t^I + (1-\pi_0)\int_0^{\mu_L^{-1}(x)} \mathrm{d}G_t^I\right] \\
&\quad + \lambda\theta\left[\pi_1 \int_0^{\mu_H^{-1}(x)} \mathrm{d}G_t^C + (1-\pi_1)\int_0^{\mu_L^{-1}(x)} \mathrm{d}G_t^C\right] \\
&= (1-\lambda\theta)\left[\pi_0 G_t^I\left(\mu_H^{-1}(x)\right) + (1-\pi_0)\, G_t^I\left(\mu_L^{-1}(x)\right)\right] \\
&\quad + \lambda\theta\left[\pi_1 G_t^C\left(\mu_H^{-1}(x)\right) + (1-\pi_1)\, G_t^C\left(\mu_L^{-1}(x)\right)\right]. \qquad (4.13)
\end{aligned}
$$

The overall cumulative distribution function of μ at $t+1$ is then:

$$G_{t+1}(x) = \theta G^{C}_{t+1}(x) + (1-\theta) G^{I}_{t+1}(x)$$
$$= \theta \left[\pi_1 G^{C}_{t}\left(\mu_H^{-1}(x)\right) + (1-\pi_1) G^{C}_{t}\left(\mu_L^{-1}(x)\right)\right]$$
$$+ (1-\theta)\left[\pi_0 G^{I}_{t}\left(\mu_H^{-1}(x)\right) + (1-\pi_0) G^{I}_{t}\left(\mu_L^{-1}(x)\right)\right], \quad (4.14)$$

where θ is the fraction of competent schools.

4.3 Long run distribution of schools' reputation

Now we can prove the main Theorem of this Chapter. Let X^{τ} be the set of distribution functions G^{τ} with support contained in the interval $[\lambda\theta, \lambda\theta + (1-\lambda)]$, where $\tau \in \{C, I\}$ denotes the school type. We define the distance between two distribution functions G^{τ} and H^{τ} with the sup metric, that is:

$$\rho_{\infty}(G^{\tau}, H^{\tau}) = \sup_{x\in[\lambda\theta, \lambda\theta+(1-\lambda)]} |G^{\tau}(x) - H^{\tau}(x)|. \quad (4.15)$$

$X^C \times X^I$ is the set of all pairs of distribution functions $\left(G^C, G^I\right)$ with support contained in the interval $[\lambda\theta, \lambda\theta + (1-\lambda)]$. We define the distance between two elements of $X^C \times X^I$ as follows:

$$\rho\left(\left(G^C, G^I\right), \left(H^C, H^I\right)\right) = \max\left\{\rho_{\infty}\left(G^C, H^C\right), \rho_{\infty}\left(G^I, H^I\right)\right\}. \quad (4.16)$$

4.3.1 The existence of a "long run" distribution of schools' reputation

Let $T : X^C \times X^I \to X^C \times X^I$ be the operator defined as:

$$T\left(G^C_t, G^I_t\right) = \left(G^C_{t+1}, G^I_{t+1}\right), \quad (4.17)$$

with G^C_{t+1} and G^I_{t+1} defined as in equations (4.11) and (4.13) respectively.

Definition 4.1 *An operator T on a Banach space is said to be an $N-$contraction if there exists a finite $N \in \mathbb{N}$ such that its N-th iterate is a contraction, i.e., if $f : X \rightarrow X$ defined by:*

$$fx = T^N x$$

is a contraction in X. That is, if $\rho(fx, fy) \leq \beta\rho(x, y)$ for all $x, y \in X$, with $\beta \in (0, 1)$.

First, we enunciate Banach's fixed point Theorem, and we prove a generalization of this Theorem: if an operator in a complete metric space X is an $N-$contraction, then it has a unique fixed point in X. Then, we prove that T is an $N-$contraction in $X^C \times X^I$.

Banach's Fixed Point Theorem *Let (X, ρ) be a complete metric space and $f : X \rightarrow X$. If f is a contraction, then the sequence $\{x_n\}_{n=0}^{\infty}$ defined according to $x_n = f^n x$ is Cauchy, and f has a unique fixed point in X, where f^n is the nth iterate of the operator. Furthermore, if x_0 is any point in X, then the sequence $\{x_n\}_{n=0}^{\infty}$ converges to the fixed point.*[1]

The following Theorem generalizes Banach's fixed point Theorem to $N-$contractions:

Proposition 4.1 (generalization of Banach's Fixed Point Theorem) *Let (X, ρ) be a complete metric space and $T : X \rightarrow X$ an $N-$contraction. Then the sequence $\{x_n\}_{n=0}^{\infty}$ defined according to $x_n = T^n x$ is Cauchy, and T has a unique fixed point in X. Furthermore, if x_0 is any point in X, then the sequence $\{x_n\}_{n=0}^{\infty}$ converges to the fixed point.*

Proof. See Appendix C. ■

Now we prove that the operator T defined in equation (4.17) is an $N-$contraction.

[1] See for example Burkill and Burkill (1970, pp. 52).

Lemma 4.1 *The operator T defined in equation (4.17) satisfies:*

$$\rho\left(T\left(G_t^C, G_t^I\right), T\left(H_t^C, H_t^I\right)\right) \leq \rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right)$$

for any $\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right) \in X^C \times X^I$, with the metric defined as in equation (4.16).

Proof. From the definitions of G_{t+1}^C and G_{t+1}^I in equations (4.11) and (4.13) we get:

$$\begin{aligned}\sup_x \left|G_{t+1}^C(x) - H_{t+1}^C(x)\right| = \sup_x \Big|&(1-\lambda+\lambda\theta)\left[\pi_1\left(G_t^C\left(\mu_H^{-1}(x)\right) - H_t^C\left(\mu_H^{-1}(x)\right)\right)\right.\\ &\left.+(1-\pi_1)\left(G_t^C\left(\mu_L^{-1}(x)\right) - H_t^C\left(\mu_L^{-1}(x)\right)\right)\right]\\ &+\lambda(1-\theta)\left[\pi_0\left(G_t^I\left(\mu_H^{-1}(x)\right) - H_t^I\left(\mu_H^{-1}(x)\right)\right)\right.\\ &\left.+(1-\pi_0)\left(G_t^I\left(\mu_L^{-1}(x)\right) - H_t^I\left(\mu_L^{-1}(x)\right)\right)\right]\Big|. \qquad (4.18)\end{aligned}$$

and

$$\begin{aligned}\sup_x \left|G_{t+1}^I(x) - H_{t+1}^I(x)\right| = \sup_x \Big|&(1-\lambda\theta)\left[\pi_0\left(G_t^I\left(\mu_H^{-1}(x)\right) - H_t^I\left(\mu_H^{-1}(x)\right)\right)\right.\\ &\left.+(1-\pi_0)\left(G_t^I\left(\mu_L^{-1}(x)\right) - H_t^I\left(\mu_L^{-1}(x)\right)\right)\right]\\ &+\lambda\theta\left[\pi_1\left(G_t^C\left(\mu_H^{-1}(x)\right) - H_t^C\left(\mu_H^{-1}(x)\right)\right)\right.\\ &\left.+(1-\pi_1)\left(G_t^C\left(\mu_L^{-1}(x)\right) - H_t^C\left(\mu_L^{-1}(x)\right)\right)\right]\Big|. \qquad (4.19)\end{aligned}$$

Noticing that the supremum of the sum is weakly smaller than the sum of the supremum of each addend, and using it with the definition of ρ_∞, we obtain:

$$\begin{aligned}\sup_x \left|G_{t+1}^C(x) - H_{t+1}^C(x)\right| \leq &(1-\lambda+\lambda\theta)\left[\pi_1\rho_\infty\left(G_t^C, H_t^C\right) + (1-\pi_1)\rho_\infty\left(G_t^C, H_t^C\right)\right]\\ &+\lambda(1-\theta)\left[\pi_0\rho_\infty\left(G_t^I, H_t^I\right) + (1-\pi_0)\rho_\infty\left(G_t^I, H_t^I\right)\right], \qquad (4.20)\end{aligned}$$

and

$$\sup_x \left|G_{t+1}^I(x) - H_{t+1}^I(x)\right| \leq (1-\lambda\theta)\left[\pi_0\rho_\infty\left(G_t^I, H_t^I\right) + (1-\pi_0)\rho_\infty\left(G_t^I, H_t^I\right)\right]$$
$$+ \lambda\theta\left[\pi_1\rho_\infty\left(G_t^C, H_t^C\right) + (1-\pi_1)\rho_\infty\left(G_t^C, H_t^C\right)\right]. \quad (4.21)$$

Hence we obtain the desired result:

$$\rho\left(\left(G_{t+1}^C, G_{t+1}^I\right), \left(H_{t+1}^C, H_{t+1}^I\right)\right) \leq \max\left\{(1-\lambda+\lambda\theta)\rho_\infty\left(G_t^C, H_t^C\right) + \lambda(1-\theta)\rho_\infty\left(G_t^I, H_t^I\right)\right.$$
$$\left., (1-\lambda\theta)\rho_\infty\left(G_t^I, H_t^I\right) + \lambda\theta\rho_\infty\left(G_t^C, H_t^C\right)\right\}$$
$$\leq \max\left\{\rho_\infty\left(G_t^C, H_t^C\right), \rho_\infty\left(G_t^I, H_t^I\right)\right\}$$
$$= \rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right). \quad (4.22)$$

■

If the inequality in (4.22) is not strict for some distribution functions $\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right) \in X^C \times X^I$, then T is not a contraction. Two necessary conditions to obtain (4.22) with equality are:

Equality 1 (E1)

$$\arg\max_{\{x\}}\left|G_t^C\left(\mu_H^{-1}(x)\right) - H_t^C\left(\mu_H^{-1}(x)\right)\right| \cap \arg\max_{\{x\}}\left|G_t^C\left(\mu_L^{-1}(x)\right) - H_t^C\left(\mu_L^{-1}(x)\right)\right|$$
$$\cap \arg\max_{\{x\}}\left|G_t^I\left(\mu_H^{-1}(x)\right) - H_t^I\left(\mu_H^{-1}(x)\right)\right|$$
$$\cap \arg\max_{\{x\}}\left|G_t^I\left(\mu_L^{-1}(x)\right) - H_t^I\left(\mu_L^{-1}(x)\right)\right| \neq \phi. \quad (4.23)$$

If this condition is not met, it follows that the supremum of the sum is strictly smaller than the sum of the supremum of each addend in (4.18) and (4.19), because each addend is maximized in a different level of x.

Equality 2 (E2)

$$\rho_\infty\left(G_t^C, H_t^C\right) = \rho_\infty\left(G_t^I, H_t^I\right). \tag{4.24}$$

If this condition is not met, it follows that the average of $\rho_\infty\left(G_t^C, H_t^C\right)$ and $\rho_\infty\left(G_t^I, H_t^I\right)$ is strictly lower than $\max\left\{\rho_\infty\left(G_t^C, H_t^C\right), \rho_\infty\left(G_t^I, H_t^I\right)\right\}$ in (4.22).

Since there are no schools with $\mu < \lambda\theta$, we know that $G_t^\tau\left(\mu_H^{-1}(x)\right) = H_t^\tau\left(\mu_H^{-1}(x)\right) = 0$ for $x < \mu_H(\lambda\theta)$.[2] Similarly, for $x > \mu_L(\lambda\theta + (1-\lambda))$ we know that $G_t^\tau\left(\mu_L^{-1}(x)\right) = H_t^\tau\left(\mu_L^{-1}(x)\right) = 1$. Hence,

$$\begin{aligned} G_t^\tau\left(\mu_H^{-1}(x)\right) - H_t^\tau\left(\mu_H^{-1}(x)\right) &= 0 \text{ if } x < \mu_H(\lambda\theta)\text{, and} \quad (4.25) \\ G_t^\tau\left(\mu_L^{-1}(x)\right) - H_t^\tau\left(\mu_L^{-1}(x)\right) &= 0 \text{ if } x > \mu_L(\lambda\theta + (1-\lambda)) \quad (4.26) \end{aligned}$$

for $\tau \in \{C, I\}$.

As a consequence,

$$\arg\max_{\{x\}}\left|G_t^\tau\left(\mu_H^{-1}(x)\right) - H_t^\tau\left(\mu_H^{-1}(x)\right)\right| \subseteq \left[\mu_H(\lambda\theta), \lambda\theta + (1-\lambda)\right], \tag{4.27}$$

and

$$\arg\max_{\{x\}}\left|G_t^\tau\left(\mu_L^{-1}(x)\right) - H_t^\tau\left(\mu_L^{-1}(x)\right)\right| \subseteq \left[\lambda\theta, \mu_L(\lambda\theta + (1-\lambda))\right] \tag{4.28}$$

for $\tau \in \{C, I\}$. It follows that Condition E1 cannot be met if $\mu_L(\lambda\theta + (1-\lambda)) < \mu_H(\lambda\theta)$. In this case, we can say that T is a contraction mapping.

A case where $\mu_L(\lambda\theta + (1-\lambda)) > \mu_H(\lambda\theta)$ is described in Figure 4.3. In this case, Condition E1 is possible: for example, if $G_t^C = G_t^I$ are degenerate in $\lambda\theta$ and $H_t^C = H_t^I$ are degenerate in $\lambda\theta + (1-\lambda)$, then $\rho\left(\left(G_{t+1}^C, G_{t+1}^I\right), \left(H_{t+1}^C, H_{t+1}^I\right)\right) = \rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right)$, as Figure 4.4 shows. Therefore, if $\mu_L(\lambda\theta + (1-\lambda)) > \mu_H(\lambda\theta)$, T is not a contraction.

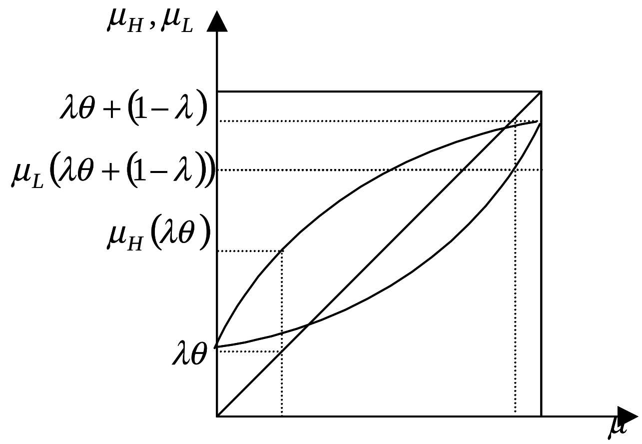

Figure 4.3: The case where $\mu_L(\lambda\theta + (1-\lambda)) > \mu_H(\lambda\theta)$.

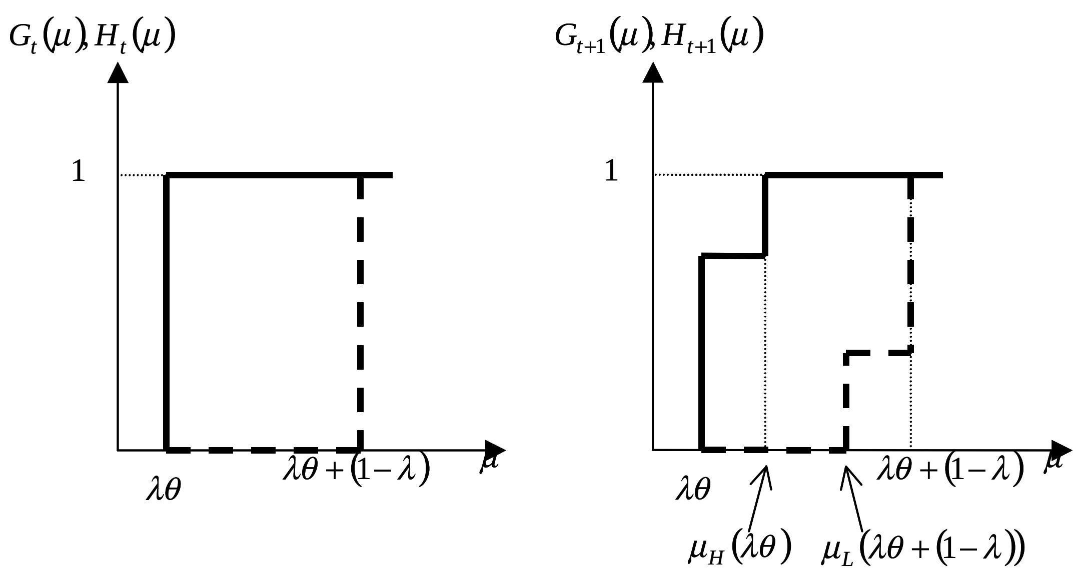

Figure 4.4: Example where $\rho\left(\left(G_{t+1}^C, G_{t+1}^I\right), T\left(H_{t+1}^C, H_{t+1}^I\right)\right) = \rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right)$.

Iterating the operator T and using lemma (4.1), we obtain:

$$\rho\left(T^{n}\left(G_{t}^{C},G_{t}^{I}\right),T^{n}\left(H_{t}^{C},H_{t}^{I}\right)\right)\leq\rho\left(\left(G_{t}^{C},G_{t}^{I}\right),\left(H_{t}^{C},H_{t}^{I}\right)\right). \tag{4.29}$$

We can establish two necessary conditions to obtain (4.29) with equality, as in E1 and E2. The condition analogous to E1 is now:

Equality 1' (E1')

$$\arg\max_{\{x\}}\left|G_{t}^{C}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)-H_{t}^{C}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)\right|\cap$$

$$...\cap\arg\max_{\{x\}}\left|G_{t}^{I}\left(\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)\right)-H_{t}^{I}\left(\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)\right)\right|\neq\phi. \tag{4.30}$$

We denote by $\mu_{H}^{n}\left(\mu_{\min}\right)$ the posterior belief after n periods with high test results starting from $\mu_{\min}$, and by $\mu_{L}^{n}\left(\mu_{\max}\right)$ the posterior belief after n low test scores starting from $\mu_{\max}$. Since there are no schools with $\mu<\lambda\theta$, we know that $G_{t}^{\tau}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)=H_{t}^{\tau}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)=0$ for $x<\mu_{H}^{n}\left(\lambda\theta\right)$ and $\tau\in\{C,I\}$. Similarly, for $x>\mu_{L}^{n}\left(\lambda\theta+\left(1-\lambda\right)\right)$ we know that $G_{t}^{\tau}\left(\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)\right)=H_{t}^{\tau}\left(\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)\right)=1$.[3] Hence,

$$G_{t}^{C}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)-H_{t}^{C}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)=0\text{ if }x<\mu_{H}^{n}\left(\lambda\theta\right)\text{ and} \tag{4.31}$$

$$G_{t}^{I}\left(\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)\right)-H_{t}^{I}\left(\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)\right)=0\text{ if }x>\mu_{L}^{n}\left(\lambda\theta+\left(1-\lambda\right)\right). \tag{4.32}$$

As a consequence,

$$\arg\max_{\{x\}}\left|G_{t}^{C}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)-H_{t}^{C}\left(\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)\right)\right|\subseteq\left[\mu_{H}^{n}\left(\lambda\theta\right),\lambda\theta+\left(1-\lambda\right)\right], \tag{4.33}$$

[2] Recall that G_{t}^{τ} and H_{t}^{τ} belong to X^{τ}, the set of all distribution functions with support contained in the interval $[\lambda\theta,\lambda\theta+(1-\lambda)]$.

[3] Notice that if $x<\mu_{H}^{n}\left(\mu_{\min}\right)$, we know that $\mu_{H}^{-1}\left(\mu_{H}^{-1}...\left(\mu_{H}^{-1}\left(x\right)\right)\right)<\mu_{\min}$. Similarly, if $x>\mu_{L}^{n}\left(\mu_{\max}\right)$, we know that $\mu_{L}^{-1}\left(\mu_{L}^{-1}...\left(\mu_{L}^{-1}\left(x\right)\right)\right)>\mu_{\max}$.

and

$$\arg\max_{\{x\}} \left| G_t^I \left(\mu_L^{-1} \left(\mu_L^{-1} ... \left(\mu_L^{-1} (x)\right)\right)\right) - H_t^I \left(\mu_L^{-1} \left(\mu_L^{-1} ... \left(\mu_L^{-1} (x)\right)\right)\right) \right| \subseteq \left[\lambda\theta, \mu_L^n \left(\lambda\theta + (1-\lambda)\right)\right]. \tag{4.34}$$

Therefore, Condition E1' cannot be met if $\mu_L^n (\lambda\theta + (1-\lambda)) < \mu_H^n (\lambda\theta)$.

In the following Lemma we prove that, even if $\mu_L (\lambda\theta + (1-\lambda)) > \mu_H (\lambda\theta)$, there is a finite number of periods N such that $\mu_L^N (\lambda\theta + (1-\lambda)) < \mu_H^N (\lambda\theta)$.

Lemma 4.2 *There is $N < \infty$ such that $\mu_L^N (\lambda\theta + (1-\lambda)) < \mu_H^N (\lambda\theta)$*

Proof. Since $\mu_H^n (x)$ converges to $\mu_{\max}$ and $\mu_L^n (x)$ converges to $\mu_{\min} < \mu_{\max}$, we know that there is $N < \infty$ such that $\mu_L^N (\lambda\theta + (1-\lambda)) < \mu_H^N (\lambda\theta)$. ■

Now we can prove that the operator T is an $N-$contraction.

Lemma 4.3 *The operator T defined in equation (4.17) satisfies that there is $N < \infty$ and such that*

$$\rho\left(T^N \left(G_t^C, G_t^I\right), T^N \left(H_t^C, H_t^I\right)\right) \leq \beta\rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right),$$

for any pair $\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right) \in X^C \times X^I$, with $\beta \in (0,1)$ and the metric defined as in equation (4.16). That is, T is an $N-$contraction.

Proof. From lemma (4.2) we know that there is $N < \infty$ such that $\mu_L^N (\lambda\theta + (1-\lambda)) < \mu_H^N (\lambda\theta)$. Since Condition E1' cannot be met if $\mu_L^N (\lambda\theta + (1-\lambda)) < \mu_H^N (\lambda\theta)$, we know that:

$$\rho\left(T^N \left(G_t^C, G_t^I\right), T^N \left(H_t^C, H_t^I\right)\right) < \rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right).$$

Furthermore, there is $\beta \in (0,1)$ such that:

$$\rho\left(T^N \left(G_t^C, G_t^I\right), T^N \left(H_t^C, H_t^I\right)\right) \leq \beta\rho\left(\left(G_t^C, G_t^I\right), \left(H_t^C, H_t^I\right)\right).$$

Therefore, the operator T is an $N-$contraction. ■

Example 4.1 *Assume that $N=1$, that is, $\mu_L(\lambda\theta+(1-\lambda))<\mu_H(\lambda\theta)$. Then we obtain:*

$$\rho\left(T\left(G_t^C,G_t^I\right),T\left(H_t^C,H_t^I\right)\right)=\max\left\{\sup_{x<\mu_H(\lambda\theta)}\left|G_{t+1}^C(x)-H_{t+1}^C(x)\right|\right.$$

$$,\sup_{x>\mu_L(\lambda\theta+(1-\lambda))}\left|G_{t+1}^C(x)-H_{t+1}^C(x)\right|$$

$$,\sup_{x<\mu_H(\lambda\theta)}\left|G_{t+1}^I(x)-H_{t+1}^I(x)\right|$$

$$\left.,\sup_{x>\mu_L(\lambda\theta+(1-\lambda))}\left|G_{t+1}^I(x)-H_{t+1}^I(x)\right|\right\}$$

$$\leq\max\left\{(1-\lambda+\lambda\theta)(1-\pi_1)\rho_\infty\left(G_t^C,H_t^C\right)+\lambda(1-\theta)(1-\pi_0)\rho_\infty\left(G_t^I,H_t^I\right)\right.$$

$$,(1-\lambda+\lambda\theta)\pi_1\rho_\infty\left(G_t^C,H_t^C\right)+\lambda(1-\theta)\pi_0\rho_\infty\left(G_t^I,H_t^I\right)$$

$$,(1-\lambda\theta)(1-\pi_0)\rho_\infty\left(G_t^I,H_t^I\right)+\lambda\theta(1-\pi_1)\rho_\infty\left(G_t^C,H_t^C\right)$$

$$\left.,(1-\lambda\theta)\pi_0\rho_\infty\left(G_t^I,H_t^I\right)+\lambda\theta\pi_1\rho_\infty\left(G_t^C,H_t^C\right)\right\}$$

$$\leq\rho\left(\left(G_t^C,G_t^I\right),\left(H_t^C,H_t^I\right)\right)$$

$$\max\left\{(1-\lambda+\lambda\theta)(1-\pi_1)+\lambda(1-\theta)(1-\pi_0)\right.$$

$$,(1-\lambda+\lambda\theta)\pi_1+\lambda(1-\theta)\pi_0$$

$$\left.,(1-\lambda\theta)(1-\pi_0)+\lambda\theta(1-\pi_1),(1-\lambda\theta)\pi_0+\lambda\theta\pi_1\right\}.$$

The following proposition states that $X^C\times X^I$ is a complete metric space.

Proposition 4.2 *The set of distribution functions G with support contained in the interval $[\lambda\theta,\lambda\theta+(1-\lambda)]$ with the sup metric is a complete metric space. Therefore, $\left(X^C\times X^I,\rho\right)$ is complete.*

Proof. See Appendix C. ■

As $\left(X^C \times X^I, \rho\right)$ is a complete metric space, we can apply the result established in Proposition 1, and we obtain the following Theorem.

Theorem 4.1 *The operator T has a unique fixed point in $X^C \times X^I$, that can be reached beginning from any pair of distribution functions $\left(G^C, G^I\right)$ in $X^C \times X^I$.*

The significance of this result is twofold. On the one hand, it implies that μ has a unique and globally stable ergodic distribution function G, which we refer to as the "long run distribution of schools' reputation" or the "long run distribution of beliefs".[4]

On the other hand, the equilibrium assignment of students and price function in the high quality equilibrium depend on the distribution of beliefs (and the joint distribution of income and abilities, that we assume constant). Therefore, if we consider the long run distribution of beliefs, there is a unique price function p^* faced by students and schools in equilibrium in all periods.

4.3.2 Characterization of the long run distribution of beliefs

Since $\left(G^C, G^I\right)$ is a fixed point of the operator T, we know that if the distribution of reputations at t is the long run distribution, G, then $G_{t+1} = G_t = G$. We use this fact to prove that G has the following property:

Lemma 4.4 *The support of the long run distribution of reputations, U, is contained in $[\mu_{\min}, \mu_{\max}]$. Furthermore, $\mu_{\min}$ and $\mu_{\max}$ are limit points of U.*

[4]Similarly, we will refer to the distribution of reputations for competent schools as to the "long run distribution of reputations for competent schools", and likewise for inept schools.

Proof. Consider a distribution G_t with support U_t, such that $U_t \nsubseteq [\mu_{\min}, \mu_{\max}]$. Then, either $\inf U_t < \mu_{\min}$ or $\sup U_t > \mu_{\max}$ (or both). But $\mu_{t+1} > \mu_t$ for all $\mu_t < \mu_{\min}$, and $\mu_{t+1} < \mu_t$ for all $\mu_t > \mu_{\max}$. Therefore, either $\inf U_{t+1} > \inf U_t$ or $\sup U_{t+1} < \sup U_t$ (or both), hence $U_{t+1} \neq U_t$, and G_t is not the long run distribution of reputations G. We conclude that $U \subseteq [\mu_{\min}, \mu_{\max}]$.

Now consider a distribution G_t with support U_t, such that $\mu_{\min}$ is not a limit point of U. Let define $U_t' = U_t \backslash \{\mu_{\min}\}$. As $\mu_{\min} < \mu_L(\mu_t) < \mu_t$ for all $\mu_t > \mu_{\min}$, then $\inf U_{t+1}' < \inf U_t'$, and G_t is not the long run distribution of reputations G. We conclude that $\mu_{\min}$ is a limit point of U. Similarly we can conclude that $\mu_{\max}$ is a limit point of U. ■

Lemma 4.5 *The support of the long run distribution of reputations is infinite. In particular, U is not degenerate.*

Proof. G cannot be degenerate because there is no μ such that $\mu_L(\mu) = \mu_H(\mu)$. To see that it must have an infinite support, consider a non-degenerate distribution G_t with finite support U_t and check that it cannot be the long-run distribution G as follows. Define m as:

$$m = \min \{\mu \in U_t : \mu > \mu_{\min}\}. \tag{4.35}$$

We know m exists because G_t is non degenerate and U_t is finite. Since $m > \mu_{\min}$, we know that $\mu_{\min} < \mu_L(m) < m$, and therefore G_{t+1} differs from G_t because $\mu_L(m) \in U_{t+1}$ and $\mu_L(m) \notin U_t$. Hence, G_t is not the long run distribution of reputations G. We conclude that U must be infinite. ■

As we pointed out before, schools' types are not even eventually (asymptotically) revealed. However, in the long run a good type has a better reputation than a bad type, in the sense that G^C first-order stochastically dominates G^I. We prove this result by using the fact that a pair $\left(G^C, G^I\right)$ of long run distribution functions for schools C and I is the limit point of the sequence defined from the iteration of the operator T starting from any pair of initial distribution functions in $X^C \times X^I$.

Lemma 4.6 *G^C first-order stochastically dominates G^I.*

Proof. From equations (4.11) and (4.13) we obtain:

$$\begin{aligned} G_{t+1}^I(x) - G_{t+1}^C(x) = (1-\lambda)\Big[&\pi_0 G_t^I\left(\mu_H^{-1}(x)\right) + (1-\pi_0)\, G_t^I\left(\mu_L^{-1}(x)\right) \\ &- \left(\pi_1 G_t^C\left(\mu_H^{-1}(x)\right) + (1-\pi_1)\, G_t^C\left(\mu_L^{-1}(x)\right)\right)\Big]. \end{aligned} \quad (4.36)$$

But $\mu_H^{-1}(x) < \mu_L^{-1}(x)$ for all $x \in [\lambda\theta, \lambda\theta + (1-\lambda)]$, so $G_t^\tau\left(\mu_H^{-1}(x)\right) \le G_t^\tau\left(\mu_L^{-1}(x)\right)$ for $\tau \in \{C, I\}$. Hence, if $G_t^I(x) \ge G_t^C(x)$ for all $x \in [\lambda\theta, \lambda\theta + (1-\lambda)]$, we obtain that $G_{t+1}^I(x) \ge G_{t+1}^C(x)$ for all $x \in [\lambda\theta, \lambda\theta + (1-\lambda)]$, since $\pi_1 > \pi_0$. Therefore, starting with the same distribution functions for competent and inept firms, $G_t^C = G_t^I$, we obtain that $G_{t+n}^I(x) \ge G_{t+n}^C(x)$ for all $x \in [\lambda\theta, \lambda\theta + (1-\lambda)]$ and for all $n > 0$. Taking the limit as n goes to infinity, we get:

$$G^I(x) \ge G^C(x) \text{ for all } x \in U.$$

■

Example 4.2 *Consider the following parameters: $\lambda = 0.1$; $\theta = 0.5$; $\pi_1 = 0.6$; $\pi_0 = 0.4$. Starting from a degenerate distribution of prior beliefs at θ, and with N=10000 schools, the*

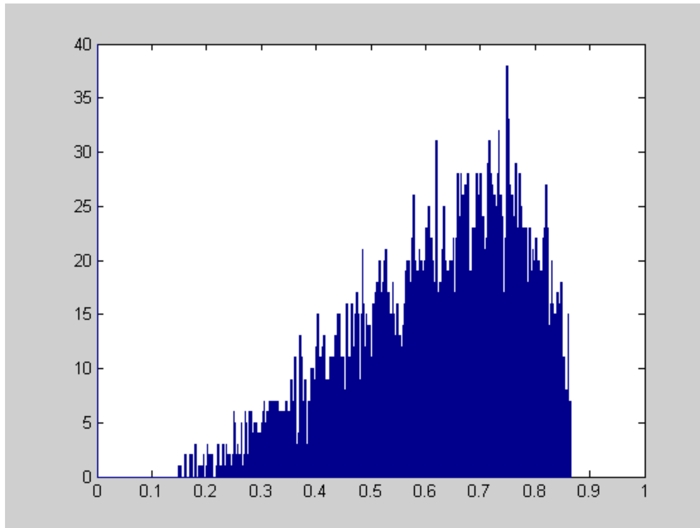

Figure 4.5: Distribution of schools' reputation. Competent schools.

histograms for μ after 1000 periods for schools C and I are illustrated in Figures 4.5 and 4.6 respectively. The Matlab program is provided in Appendix D.

4.4 Further remarks

The preceding analysis can be applied in different models of reputation with imperfect monitoring, provided that there is a continuum of firms.

Consider for example the model of reputation with imperfect public monitoring and exogenous replacements by Mailath and Samuelson (2001). In this model, there is a unique firm and a continuum of consumers who repeatedly purchase an experience good from the firm. Now assume that there is a continuum of firms in this market, instead of only one seller, and that the realization of utility outcomes for each firm is still public. Hence, consumers' beliefs are common, in the sense that they all assign the same probability to a

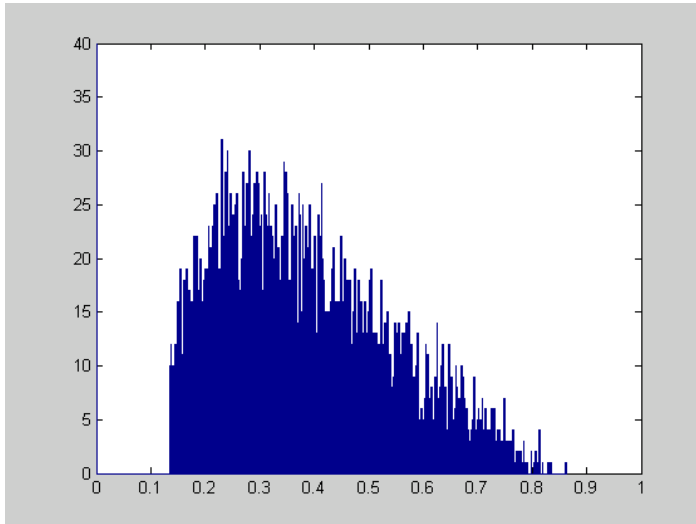

Figure 4.6: Distribution of schools' reputation. Inept schools.

particular firm being competent. However, as each firm has a different history of utility outcomes, it also has a different reputation under the high effort equilibrium. Hence, there is a non-degenerate distribution of firms' reputations. As in our model, it can be shown that the distribution of firms' reputation converges to a "long run distribution" (the fixed point of the dynamical system that describes the evolution of distributions). The relevance of this result is that the market equilibrium depends on the availability of firms with a given reputation (the "reputations supply"), which is stable across time under the long run distribution of firms' reputation.

Notice that this case differs from the model of reputation with imperfect private monitoring by Mailath and Samuelson (1998). In Mailath and Samuelson (1998) there is a (non-degenerate) distribution of consumers' posteriors that the firm is competent, since each consumer receives a different realization of the utility outcome, and they update beliefs accordingly. But this is a distribution of consumers (each facing a different history of utility outcomes from the same firm), not of firms (each with the same history of utility outcomes) as in our case.

Appendix A

Generalization of Banach's Fixed Point Theorem

A.1 Proof of Proposition 4.1

Take x_n and x_m such that $n > m \geq N$, and define:

$$a = \max\left\{a' \in \mathbb{N} : \left(a' \leq m\right) \wedge \left(a' = k * N \text{ for some } k \in \mathbb{N}\right)\right\}. \tag{A.1}$$

Then, $0 \leq m - a < N$ and $a \geq N$. We can find bounds on the distance between x_n and x_m as follows:

$$\begin{aligned}
\rho\left(x_n, x_m\right) &= \rho\left(T^a x_{n-a}, T^a x_{m-a}\right) \\
&\leq \beta\rho\left(T^{a-N} x_{n-a}, T^{a-N} x_{m-a}\right) \\
&\leq \beta^2\rho\left(T^{a-2N} x_{n-a}, T^{a-2N} x_{m-a}\right) \\
&\dots \\
&\leq \beta^{a/N}\rho\left(x_{n-a}, x_{m-a}\right).
\end{aligned} \tag{A.2}$$

Now, let:

$$d = \max\left\{\rho\left(x_i, x_j\right) : i, j \in [0, N]\right\}. \tag{A.3}$$

If $n - a > N$, define:

$$b = \max\left\{b' \in \mathbb{N} : \left(b' \leq n - a\right) \wedge \left(b' = k * N \text{ for some } k \in \mathbb{N}\right)\right\}, \tag{A.4}$$

then $0 \leq n - a - b < N$.

Using the triangular inequality, we obtain:

$$\rho\left(x_{n-a}, x_{m-a}\right) \leq \rho\left(x_{n-a}, x_b\right) + \rho\left(x_b, x_{b-N}\right) + ... + \rho\left(x_N, x_{m-a}\right). \tag{A.5}$$

Applying (A.2) to each addend bu the last one, we obtain:

$$\begin{aligned}
\rho\left(x_{n-a}, x_{m-a}\right) &\leq \beta^{b/N}\rho\left(x_{n-a-b}, x_0\right) + \beta^{(b-N)/N}\rho\left(x_N, x_0\right) + ... + \rho\left(x_N, x_{m-a}\right) \\
&\leq \left(\beta^{b/N} + \beta^{(b-N)/N} + \beta^{(b-2N)/N} + ... + 1\right) d \\
&= \sum_{i=0}^{b/N} \left(\beta^i\right) d \\
&\leq \sum_{i=0}^{\infty} \left(\beta^i\right) d \\
&= \frac{d}{(1-\beta)}.
\end{aligned} \tag{A.6}$$

If $n - a \leq N$,

$$\rho\left(x_{n-a}, x_{m-a}\right) \leq d < \frac{d}{(1-\beta)}. \tag{A.7}$$

Therefore, from (A.2), (A.6) and (A.7) we obtain:

$$\rho\left(x_n, x_m\right) \leq \beta^{a/N}\frac{d}{(1-\beta)}. \tag{A.8}$$

Since $\beta < 1$ we can take m sufficiently large in order to make $\beta^{a/N}$ arbitrarily small. Therefore, $\rho\left(x_n, x_m\right)$ can be made arbitrarily small taking m sufficiently large. That is, $(\forall\varepsilon > 0)\left(\exists N\left(\varepsilon\right)\right) : \rho\left(x_n, x_m\right) < \varepsilon$ if $n, m \geq N\left(\varepsilon\right)$ and so the sequence is Cauchy. As X is complete, the sequence $\{x_n\}$ converges to some element $x^*$of X.

As $\rho\left(Tx, Ty\right) \leq \rho\left(x, y\right)$, we know that T is continuous: taking $\delta = \varepsilon$, we obtain that $(\forall\varepsilon > 0)(\exists\delta > 0) : \rho\left(x, y\right) < \delta \Rightarrow \rho\left(Tx, Ty\right) \leq \rho\left(x, y\right) < \varepsilon$. Continuity of T implies that:

$$\lim_{n\to\infty} T\left(x_n\right) = T\left(\lim_{n\to\infty} x_n\right) = T\left(x^*\right), \tag{A.9}$$

where $x^* \equiv \lim_{n\to\infty} x_n$. But $\lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} x_{n+1} = x^*$. Therefore, x^* is a fixed point of T in X. The fixed point of T is unique, since if x and x' were two fixed points of T, we would know that $x = Tx = T^N x$ and $x' = Tx' = T^N x'$. Therefore,

$$\begin{aligned} \rho\left(x, x'\right) &= \rho\left(T^N x, T^N x'\right) \leq \beta\rho\left(x, x'\right) \qquad \text{with } \beta < 1 \\ &\Leftrightarrow \rho\left(x, x'\right) = 0, \end{aligned} \tag{A.10}$$

so that $x = x'$.

A.2 Proof of Proposition 4.2

Remember that we defined X^τ as the set of distribution functions G^τ with support contained in the interval $[\lambda\theta, \lambda\theta + (1-\lambda)]$, where $\tau \in \{C, I\}$ denotes the school type, and with the distance between two distribution functions G^τ and H^τ as:

$$\rho_\infty\left(G^\tau, H^\tau\right) = \sup_{x\in[\lambda\theta,\lambda\theta+(1-\lambda)]} \left|G^\tau(x) - H^\tau(x)\right|. \tag{A.11}$$

As distribution functions are bounded, if $\{G_n\}_{n=0}^{\infty}$ is a Cauchy sequence in X^τ with $\tau \in \{C, I\}$, it is also a Cauchy sequence in $\mathcal{B}(\mathbb{R}, [0,1])$, the space of bounded functions from $\mathbb{R}$ into $[0,1]$ equipped with the same metric ρ_∞ as X^τ. As $\mathcal{B}(\mathbb{R}, [0,1])$ is a complete metric space, then there is a bounded function $G \in \mathcal{B}(\mathbb{R}, [0,1])$ such that $G_n \to G$. We must show that $G \in X^\tau$, that is,

Condition 1 G is non-decreasing, i.e., $G(x_1) \geq G(x_0)$ whenever $x_1 > x_0$

Condition 2 G is right continuous, i.e., $\lim_{x\downarrow a} G(x) = G(a)$

Condition 3 G is normalized, i.e., $G(\lambda\theta) = 0$ and $G(\lambda\theta + (1-\lambda)) = 1$

Assume that G does not satisfy Condition 1. Then, there are two points $x_0, x_1 \in [\lambda\theta, \lambda\theta + (1-\lambda)]$ such that $x_1 > x_0$ and $G(x_1) < G(x_0)$. Define d as follows:

$$d \equiv G(x_0) - G(x_1) > 0, \tag{A.12}$$

and take $\varepsilon < \frac{d}{2}$. Then, if we consider any $G_n \in \{G_k\}_{k=0}^{\infty}$ such that $|G_n(x_0) - G(x_0)| < \varepsilon$, we know that

$$-\frac{d}{2} < G_n(x_0) - G(x_0) < \frac{d}{2}. \tag{A.13}$$

Since $G_n \in X^{\tau}$, G_n is non-decreasing, and therefore $G_n(x_1) \geq G_n(x_0)$. Then,

$$\begin{aligned} G_n(x_1) - G(x_1) &= G_n(x_1) - G(x_0) + d && \text{(A.14)} \\ &\geq G_n(x_0) - G(x_0) + d && \text{(A.15)} \\ &> \frac{d}{2} && \text{(A.16)} \\ &> \varepsilon. && \text{(A.17)} \end{aligned}$$

Therefore $|G_n(x_1) - G(x_1)| > \varepsilon$, and we conclude that $\rho_{\infty}(G_n, G) > \varepsilon$ for all $G_n \in \{G_n\}_{n=0}^{\infty}$. Hence, $G_n \nrightarrow G$, contradicting our hypothesis. We conclude that G is non-decreasing.

Now let x_0 be an arbitrary point of $[\lambda\theta, \lambda\theta + (1-\lambda)]$ and take $\varepsilon > 0$. Since $G_n \to G$, there is an n such that:

$$\rho_{\infty}(G_n, G) = \sup_{x \in [\lambda\theta, \lambda\theta + (1-\lambda)]} |G_n(x) - G(x)| < \varepsilon, \tag{A.18}$$

and since $G_n \in X$, G_n is right continuous and there is a $\delta > 0$ such that if $0 < x - x_0 < \delta$, then:

$$|G_n(x_0) - G_n(x)| < \varepsilon. \tag{A.19}$$

Using the triangular inequality, we obtain that if $0 < x - x_0 < \delta$, then:

$$\begin{aligned} |G(x_0) - G(x)| &\leq |G(x_0) - G_n(x_0)| + |G_n(x_0) - G_n(x)| + |G_n(x) - G(x)| && \text{(A.20)} \\ &< 3\varepsilon. && \text{(A.21)} \end{aligned}$$

Therefore, G is right continuous at x_0 and, since x_0 was chosen as an arbitrary point of $[\lambda\theta, \lambda\theta + (1-\lambda)]$, we conclude that G is right continuous.

Finally, assume that $G(\lambda\theta) \neq 0$ or $G(\lambda\theta + (1-\lambda)) \neq 1$. Let's define d as follows:

$$d \equiv \max \{|G(\lambda\theta)|, |1 - G(\lambda\theta + (1-\lambda))|\}, \tag{A.22}$$

and take $\varepsilon < d$. As $G_n \in X$, G_n is normalized, we have:

$$\rho_\infty(G_n, G) = \sup_{x \in [\lambda\theta, \lambda\theta + (1-\lambda)]} |G_n(x) - G(x)| \geq d > \varepsilon \tag{A.23}$$

for all $G_n \in \{G_n\}_{n=0}^{\infty}$. Hence, $G_n \nrightarrow G$. We conclude that G is normalized.

Summarizing, G is non-decreasing, right continuous and normalized, therefore $G \in X^\tau$, and (X^τ, ρ_∞) is a complete metric space.

Now we must show that $\left(X^C \times X^I, \rho\right)$ is also a complete metric space. Since the distance between two elements of $X^C \times X^I$ is defined as

$$\rho\left(\left(G^C, G^I\right), \left(H^C, H^I\right)\right) = \max\left\{\rho_\infty\left(G^C, H^C\right), \rho_\infty\left(G^I, H^I\right)\right\}, \tag{A.24}$$

then a sequence in $X^C \times X^I$ is Cauchy if and only if the sequences defined for each component (G_n^C and G_n^I) are Cauchy. Therefore, for any Cauchy sequence in $X^C \times X^I$ we know that the sequences defined for each component (G_n^C and G_n^I) converge in X^C and X^I respectively, and this in turn implies that the Cauchy sequence of pairs $\left(G_n^C, G_n^I\right)$ converges in $X^C \times X^I$. Therefore, $X^C \times X^I$ is a complete metric space.

Appendix B

Matlab program

```
N=10000;
syms a b
lamda=0.1;
theta=0.5;
pia=0.6;
pib=0.4;
aa=solve(a-((1-lamda)*((pia*a)/(pia*a+pib*(1-a)))+lamda*theta),'a>0','a<1',a);
bb=solve(b-((1-lamda)*(((1-pia)*b)/((1-pia)*b+(1-pib)*(1-b)))+lamda*theta),'b>0','b<1',b);
aa=double(aa)
bb=double(bb)

%Mu, Types and resuls in t=0
Mu0=0.5*ones(N,1);
T0=(rand(N,1));
i=1;
while i<N+1
    if T0(i,1)<0.5
        T0(i,1)=0;
```

```
        else
                T0(i,1)=1;
        end
        i=i+1;
end
%results (1: high)
R0=(rand(N,1));
i=1;
while i<N+1
        if T0(i,1)==0
                if R0(i,1)<(1-pib)
                        R0(i,1)=0;
                else
                        R0(i,1)=1;
                end
        else
                if R0(i,1)<(1-pia)
                        R0(i,1)=0;
                else
                        R0(i,1)=1;
                end
        end
```

```
        i=i+1;
    end

    %Mu, Types and results in t
    t=1;
    Mut=Mu0;
    Tt=T0;
    Rt=R0;
    while t<1000
        i=1;
        while i<N+1
            if Rt(i,1)==1
                Mut(i,1)=((1-lamda)*pia*Mut(i,1))/((pia*Mut(i,1))+(pib*(1-Mut(i,1))))+lamda*t
            else
                Mut(i,1)=((1-lamda)*(1-pia)*Mut(i,1))/(((1-pia)*Mut(i,1))+(1-pib)*(1-
Mut(i,1)))+lamda*theta;
            end
            i=i+1;
        end
        ddf
        Ttr=(rand(N,1));
        i=1;
        while i<N+1
```

```
    if Ttr(i,1)<(1-lamda)
        Tt(i,1)=Tt(i,1);
    else
        if Tt(i,1)<(1-(lamda*theta))
            Tt(i,1)=0;
        else
            Tt(i,1)=1;
        end
    end
    i=i+1;
end
Rtr=(rand(N,1));
i=1;
while i<N+1
    if Tt(i,1)==0
        if Rtr(i,1)<(1-pib)
            Rt(i,1)=0;
        else
            Rt(i,1)=1;
        end
    else
        if Rtr(i,1)<(1-pia)
```

```
                Rt(i,1)=0;
            else
                Rt(i,1)=1;
            end
        end
        i=i+1;
    end

    t=t+1;
end

hist(Mut,N/20)
title('Distribution of schools reputation')
xlim([bb aa])

MutA=Mut.*(Tt);
figure
hist(MutA,N/20)
title('Distribution of schools reputation. Competent schools')
xlim([bb aa])
MutB=Mut.*(ones(N,1)-Tt);
figure
hist(MutB,N/20)
title('Distribution of schools reputation. Inept schools')
```

xlim([bb aa])

Chapter 5

Conclusion

There are many ways to introduce incomplete information when modeling the school market. This dissertation examines two of them. In the signaling model developed in Chapter 2, school quality is a fixed characteristic of the school, and hence the only interesting questions are whether high quality schools can separate themselves from low quality schools using test scores as a signal, and how the competition among schools for high ability students affects the equilibrium price policy. In contrast with other models that emphasize peer effects, the possibility to differentiate prices according to students' ability do not benefit schools but only high ability students. Therefore, the consequence of a regulation that prevents tuition discounts or selection of students worsens high ability students and benefit schools, and it does not preclude the possibility of a separating equilibrium. The intuition of this result is that a regulation such as that limits the degree of competition among schools: a school that serves a high ability student faces lower costs, but the other schools cannot attract this student offering lower prices.

In the reputation model the school type refers to the possibility of providing high educational achievement to the students: only competent schools can provide high educational achievement. But as high educational achievement is costly and effort is unobservable, high quality is not assured. We analyze the "high quality equilibrium", an equilibrium where all competent schools make costly effort, and their students obtain high educational achievement. We focus in the long run equilibrium, in the sense that we consider the long run distribution of schools' reputations. But even in the long run schools' types are not revealed, and this allows for the possibility that there is always an incentive for competent schools to provide high quality.

When we analyze the equilibrium assignment of students in the high quality equilibrium, we find stratification by income and ability, as other theoretical models had found. In this model, this stratification pattern means that students with higher ability and/or income attend better reputation schools. But as the distribution of competent schools' reputation first-order stochastically dominates that of the inept schools, the probability that a school has a "good" reputation (i.e. a reputation better than a given level x) is lower for inept than for competent schools. Hence, an interesting implication of our model is that those schools that receive higher ability students are ex-ante different: loosely speaking, they are "more competent schools". This contrasts with others models such as Epple and Romano's (1998), where all schools are ex-ante identical but ex-post there is a strict hierarchy of schools' results.

Another important distinction between our model and those models that emphasize peer effects and do not take the moral hazard problem into consideration, is the role of allowing high quality schools (or high reputation schools) to charge higher fees. As an example, Epple and Romano (2002) propose a type-dependent voucher with no extra charges allowed to reach the benefits of a voucher program without "cream skimming". But to obtain a high quality equilibrium in our model, a necessary condition is that better reputation schools are allowed to charge higher fees: the cost of effort must be compensated with a benefit associated to obtaining a good test result. If the schools must accept the voucher as the only source of financing, competent schools would have no reward for their effort, since schools would not receive extra funds if they perform better. In this scenario, the goal to provide equal-quality education to all students would be reached, but with all students attending low quality schools.

Bibliography

[1] Abreu, D. and D. Pearce and E. Stacchetti, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring". Econometrica, Vol. 58,. No. 5, 1041-1063.

[2] Angrist, Joshua D. and Lang, Kevin, 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program" The American Economic Review, Vol. 95, No. 5, (Dec.), 1613-1634.

[3] Bar-Isaac, H., 2003. "Reputation and Survival: Learning in a Dynamic Signalling Model". Review of Economic Studies 70 (2), 231–251.

[4] Black, Sandra E., 1999. "Do better schools matter? Parental valuation of elementary education". Journal of Public Economics, 72, 2, May., pp. 577-599.

[5] Burkill, J.C., and H. Burkill, 1970. *A Second Course in Mathematical Analysis.* Cambridge University Press.

[6] Cripps, M., Mailath, G. and L. Samuelson, 2004. "Imperfect Monitoring and Impermanent Reputations". Econometrica, Vol. 72, No. 2 (March), 407–432.

[7] Cripps, M., Mailath, G. and L. Samuelson, 2004. "Disappearing Private Reputations in Long-Run Relationships, Second Version". PIER Working Paper No. 04-031.

[8] De Fraja, G. and P. Landeras, 2006. "Could do better: The effectiveness of incentives and competition in schools". Journal of Public Economics 90 189–213.

[9] Diamond, D. W., 1989. "Reputation Acquisition in Debt Markets". Journal of Political Economy, Vol. 97, No. 4, 828-862.

[10] Diamond, D. W., 1991. "Monitoring and Reputation: The Choice between Bank Loans and Directly Placed Debt". The Journal of Political Economy, Vol. 99, No. 4., 689-721.

[11] Epple, D. and R. Romano, 1998. "Competition between Private and Public schools, Vouchers and Peer Group Effects", The American Economic Review, Vol. 88, No.1 (Mar.), 32-62.

[12] Epple, D. and R. Romano, 2002. "Educational Vouchers and Cream Skimming", NBER Working Paper No.9354.

[13] Epple, D., D. Figlio and R. Romano, 2004. "Competition between private and public schools: testing stratification and pricing predictions". Journal of Public Economics 88, 1215-1245.

[14] Epple, D., R. Romano and H. Sieg, 2006. "Admission, Tuition and Financial Aid Policies in the Market for Higher Education". Econometrica, Vol. 74, No. 4 (July), 885–928.

[15] Fernández, R., 2001. "Sorting, Education and Inequality". NBER Working Paper 8101.

[16] Fudenberg, D. and J. Tirole, 1991. *Game Theory.* MIT Press.

[17] Fudenberg, D., D. Kreps, and E. Maskin, 1990. "Repeated Games with Long-run and Short-run Players". Review of Economic Studies, 57, 555-573.

[18] Fudenberg, D. and D. K. Levine, 1994. "Efficiency and Observability in Games with Long-Run and Short-Run Players". Journal of Economic Theory, 62 , 103-135.

[19] Fudenberg, D. and D. K. Levine, 1992. "Maintaining a Reputation when Strategies are Imperfectly Observed". Review of Economic Studies, 59: 561-579.

[20] Fudenberg, D. and D. K. Levine, 1989. "Reputation and Equilibrium Selection in Games with a Single Long-Run Player". Econometrica, Vol. 57, 759-778.

[21] Fudenberg, D., E. Maskin, and D.K. Levine, 1994. "The Folk Theorem in Repeated Games with Imperfect Public Information". Econometrica, Vol. 62, 997-1039.

[22] Hanushek, Eric A., 1989. "The impact of differential expenditures on school performance", Educational Researcher 18, No. 4, 41-51.

[23] Hanushek, E. A., J. F. Kain, J. M. Markman, and S. G. Rivkin, 2003. "Does peer ability affect student achievement?", Journal of Applied Econometrics, Vol. 18, No. 5, pp. 527-544.

[24] Holmström, B., 1999. "Managerial Incentive Problems: A Dynamic Perspective". Review of Economic Studies Vol. 66, No. 1, 169-182.

[25] Hörner, J., 2002. "Reputation and Competition", The American Economic Review, Vol. 92, No. 3., pp. 644-663.

[26] Hoxby, C., 1996. "Are Efficiency and Equity in School Finance Substitutes or Complements?". The Journal of Economic Perspectives, Vol. 10, No. 4, 51-72.

[27] Hoxby, C., 1999. “The productivity of schools and other local public goods producers”, Journal of Public Economics 74, 1-30.

[28] Kreps, D. and R. Wilson, 1982. “Reputation and Imperfect Information”. Journal of Economic Theory, 27:253-279.

[29] Mailath, G. and L. Samuelson, 1998. “Your Reputation Is Who You’re Not, Not Who You’d like To Be”, CARESS Working Paper 98-11.

[30] Mailath, G. and L. Samuelson, 2001. “Who Wants a Good Reputation?”, Review of Economic Studies 68, 415-441.

[31] Mailath, G. and L. Samuelson, 2006. *Repeated Games and Reputation. Long-Run Relationships.* Oxford University Press.

[32] Milgrom, P, and J. Roberts, 1986. “Price and Advertising Signals of Product Quality”. The Journal of Political Economy, Vol. 94, No. 4, 796-821.

[33] Nechyba, Thomas J., 2000. “Mobility, Targeting, And Private-School Vouchers”. The American Economic Review, Vol. 90, No.1, pp. 130-146.

[34] Osborne, Martin J. and Ariel Rubinstein, 1994. *A Course in Game Theory.* MIT Press.

[35] Rivkin, S., E. Hanushek and J. Kain, 2005. “Teachers, Schools, and Academic Achievement”, Econometrica, Vol. 73, No. 2, 417-458.

[36] Rockoff, J., 2004. “The impact of individual teachers on student achievement: evidence from panel data”, American Economic Review, Vol. 94, No. 2, 247-252.

[37] Rogerson, W.P., 1983. "Reputation and product quality", The Bell Journal of Economics, Vol. 14, No. 2, 508-516.

[38] Rothschild, M. and L. J. White, 1995. "The Analytics of the Pricing of Higher Education and Other Services in Which the Customers Are Inputs". Journal of Political Economy, vol. 103, issue 3, 573-86.

[39] Tadelis, S., 1999. "What's in a Name? Reputation as a Tradeable Asset", The American Economic Review, Vol. 89, No. 3, 548-563.

[40] Todd, Petra E. and Kenneth I. Wolpin, 2003. "On the Specification and Estimation of the Production Function for Cognitive Achievement". The Economic Journal 113 (485), F3–F33.