Entry in School Markets: Theory and Evidence from Brazilian Municipalities

Naercio Menezes-Filho* Rodrigo Moita† Eduardo Andrade‡

May 22, 2012

WORKING PAPER

Abstract

This paper develops a theoretical model of private school entry and estimates it using data from Brazilian municipalities. The school market is different from other markets because students are both consumers and inputs in the production fuction of education. There is a benefit to study among good peers. The theoretical model predicts a segregated equilibrium where the better students attend the private schools, rendering these with a better (endogenous) quality than the public ones. Hence, a private institution only needs to attract the best local students to be better than the existing public schools. The model's main prediction is that educational inequality induces entry. We use a panel data of private school entry in Brazilian municipalities between 1995 and 2000 to estimate an entry model. The econometric results confirm the main theoretical finding: education inequality has a positive effect on entry. A higher degree of inequality increases the private schools' ability to cream skim the best students. We also observe a decrease in the quality of the public schools, as measured by math and reading test scores, when a private school enters a town.

Key words: entry, school market, education, private education.
JEL Classification: I21

*Insper Institute and University of Sao Paulo. Email: naercioamf@insper.edu.br
†Insper Institute. Email: rodrigomsm@insper.edu.br
‡Insper Institute. Email: eduardo.andrade@insper.edu.br

1 Introduction

School choice has received a great deal of attention by economists recently. Many researchers have modeled the competition between public and private schools, the role of peer effects and the effect of school vouchers on academic results and stratification.[1] But, as far as we know, no study so far has modeled the decision of entry of a private school in a market. This is the aim of this paper.

We develop a theoretical model, where school entry decisions depend on market profitability. We analyze the case of N private schools entering a market where there exists one public school, which all students attend. Profits are a function of market fundamentals, such as population size, income, costs, income distribution and the distribution of abilities across the student population. Our model builds on the classical Hotelling's (1929) linear city model. In his model, firms' product differentiation depend on their locations on the line and consumers are uniformly distributed over it. We extend Hotelling's model by assuming that the quality of the firm depends on who consumes the good. In the case of schools, the mean quality of the students attending the school is what defines the quality of the school. Quality is endogenously determined in our model.

The equilibrium of the model shows that the private school cream skims the best students of the market, so that the private school has a higher quality than the public school. The cream skimming of good students makes it easier for entry to happen than it is in other markets. For example, an entrant on the bakery market needs to invest in differentiation if it wants to enter as a high quality bakery. A private school needs only to convince students that it will have a high quality, and then sellect the good students. If we use the definition of an entry threshold, as in Bresnahan and Reiss (1991), we can say that the entry threshold of a school is smaller than a firm with a similar cost structure operating in a industry in which quality is exogenous.[2]

In equilibrium, the private school sets a price that is higher than it would if the quality of the students were not important. The reason is that the higher its price, the more selective it becomes and consequently, the higher its quality. All these features of the equilibrium are a consequence of the fact that students are consumers as well as are inputs in the school's production function.

In terms of theory, the paper most closely related to ours is Epple and Romano (1998). Their model also has students with different abilities, and ability increases the student's academic achievement and that of the peers in the school attended, so that school quality is determined by the mean ability of its student population. Their aim, however, is to model the competition between several private schools and the public sector alternative to examine the effects of vouchers in the welfare equilibrium. In our case, we analyze how the cream skimming of good students affects the entry behavior and the pricing of

[1] See Epple and Romano (1998), Altonji, Huang and Taber (2010) and Mcwean, Urquiola and Vegas (2008).

[2] An entry threshold is a specific market size such that a firm enters a market when the market size is greater than, or crosses, the threshold.

the private school. We also look at how changes in the market fundamentals affect the profitbility of the private school, something useful in our empirical analysis.

Mcewan, Urquiola and Vegas (2008) investigate the effect of the introduction of vouchers in Chile on school quality and inequality. In order to achieve this aim, the authors compare municipalities with and without private schools, especially those where the size of the student population is just big enough to allow private schools to operate profitably. They do not, however, develop a formal theoretical model of entry nor estimate the determinants of entry using econometric techniques.

Altonji, Huang and Taber (2010) also aim at investigating the effects of the introduction of a voucher program, concentrating on the impact it may have on the quality of public schools. The authors model the cream skimming effect and show that it depends on the degree of heterogeneity within schools, the importance of the peer effect and the quality of former public school students. While we do not estimate the impact of cream skimming, we try to show in this paper that a model of school entry is fundamentally different from entry in other markets, precisely because of the cream skimming effect.

This paper is also related to the entry models of the empirical Industrial Organization literature . Since the seminal contribution of Bresnahan and Reiss (1991), several papers have examined entry in different sectors. Berry (1992), Seim (2004), among several others, have used data on firm's entry decisions to estimate profitability and the extent of competition in specific markets. However, to our knowledge, there is no use of these techniques to investigate the effect of cream skimming on entry decisions. Contrary to the case of market power, cream skimming induces facilitates entry, since it reduces investment and production costs.

In this paper we develop a theoretical entry model and estimate it using data from Brazilian municipalities. The econometric results confirm the main prediction of the theoretical model: the probability of entry of a private school is positively related to the education inequality of the municipality.

The paper is organized in the following way. Section 2 develops the theoretical model and presents the main predictions to be taken to the data. Section 3 describes the properties of our dataset and presents some preliminary evidence on the relationship between entry, population, income and inequality. Section 4 estimates the econometric models and section 5 concludes.

2 Theory: Market Structure with Social Interactions

2.1 General Description of the Model

We model the entry decision of private schools in a market as a two stage game. On the first stage the private school decides if it enters the market or stays out. On the second stage, conditional on entry on the first, it competes in the market and receive a payoff. A firm decides to enter if it expects a positive profit in equilibrium. The goal of this theoretical model is understand how the equilibrium in school markets affects market structure. This market is peculiar: the quality of a school depends on the quality of its students.

We model the problem of one free public school and N private schools competing in prices in a market. They compete for students with different abilities. We assume a continuum of students, each one indexed by its ability, or type. Students choose the school that gives them the highest utility among all schools available. Each school has a price and a quality, and the students' choice is based on these variables.

We make the assumption that high ability students benefit more from good schools. It incorporates two stylized facts from school markets. First, there is the traditional peer effect assumption that a student benefits from being among good students. If this is the case, we assume that better students gets more utility from being among good students. The empirical evidence regarding this issue is dubious. [3]

Epple and Romano (1998) are concerned with this assumption, and do not assume that higher ability students benefit more from a high average school. Instead, they assume that wealthier students are more willing to pay for school quality. They derive an optimal price schedule conditional on students' income and ability. A private school selects high income-low ability, low income-high ability and some students in between. Since tuition is conditional on income and ability, they find an equilibrium where the high income-low ability cross subsidize the low income-high ability. This is possible since high income students are more willing to pay for school quality. Also, the school acts as a first degree price discriminator, by extracting all the students´surpluses. In equilibrium, all students are indifferent among all schools. Segregation holds in Epple and Romano's model as well, but through cross-subsidization among students. This paper has a different purpose. Our goal is to derive an Bertrand-Nash oligopolistic equilibrium, where the schools charge an uniform price, quality depends on the consumers purchasing the good, and see how it relates to market structure.

Second, there can be asymmetric information in the post-school labor market, in the sense that ability is not perfectly observed - MacLeod and Urquiola (2009) use a similar assumption. In this case, attending a high average school works as a signal of the student's ability when she goes to the labor market. High ability students derive more utility from good schools because these schools require a higher effort to complete, and these students have a smaller loss of utility on exerting this effort. This is the same signaling argument used in Spence (1973). The multiplicative assumption in our utility function is a reduced form of the cost of exerting effort to complete the studies.

School quality is endogenous; it is the mean ability of the students attending it. For simplicity, we assume that this quality is the only characteristic that

[3]See Hanushek (1996) for a survey of the relationship between ability and education. Epple and Romano (1998, pg 36) presents the discussion that exists in the literature regarding peer effect and its relation to ability.

matters on the students' choice. Other things that affect school quality, such as facilities and quality of teachers, could be easily incorporated into the model.[4] In essence, a good school is one wich attracts high ability students.

The only stable equilibrium of this model is one in wich the students are segregated by ability across schools (see, for example, Benabou (1993)). The market converges to a situation in which schools are vertically differentiated, with an ordering in quality, and the public school having the lowest quality. Which school will be high or low quality is not decided within the model. In equilibrium, differences in school prices are a function of the difference in their qualities.

We derive some important conclusions from this equilibrium. First, the profitability of the private schools depends on the degree of heterogeneity in students' ability. More heterogeneous markets support a higher number of private schools in equilibrium. A market with homogeneous students will have no room for a private school, since students are willing to pay for the private school to benefit from the higher quality it offers.

Second, this heterogeneity is important since the private schools cream skims the best students when it enters a market. Therefore, the entry of a private school is detrimental to the public school's quality.

Third, the entry pattern of private schools in a market, or how entry relates to market size, is different from the one described by Bresnahan and Reiss (1991). These authors describe a homogeneous good market, where firms compete *a la* Cournot. The entry of the first firm is made easy by the cream skimming. Keeping the range of abilities constant, the market size needs to increase to accommodate further entrants into the market, as in Bresnahan and Reiss. They describe an equilibrium where the increase in market size necessary for the entry of one more firm in equilibrium converges to a certain level.[5] The difference is that in our case the market needs to increase at always increasing rates for the subsequent entry to happen. It reflects the loss of ability of further entrants to extract the gains of attracting good students to their schools.

Several papers analyze similar problems, which Glaeser and Scheinkman (2003) call 'non-market interactions'. In essence, it is a problem of positive externality in consumption.[6] They provide a general treatment of this class of problems. Benabou (1993) analyzes the problem of a city where citizens agglomerate in communities with different and endogenous levels of human capital.

The purpose of this paper is closer to the pioneer work of Becker (1992) on restaurant pricing. That paper characterizes the equilibrium of a market where the utility customers receive from eating in a restaurant depends on

[4] Other dimensions not directly related to school quality, such as religious affiliation, may be important on students' choice. However, this more refined differentiation requires a minimum number of students with these preferences, only satisfied on larger markets. Since we analyze data on small to medium towns, we believe there is no scale for such schools to exist on these markets. Therefore, we assume these dimensions are not important in our context.

[5] This level of market increase is the one necessary to support one more firm in a perfectly competitive market, when margins have vanished.

[6] Economic theory is well more developed towards negative externality in consumption, e.g. price increase due to a higher demand.

total number of consumers at the restaurant. There is market differentiation in equilibrium, with some restaurants charging a higher price and being always full, and others charging a lower price and having empty tables. A crucial difference in relation to this paper is that here what matters is not the number of customers, but the type of customer a firm attract. If the school attracts good students, it has a higher quality and is able to charge a higher price. This difference has an implication on the pricing strategies. On Becker's case the successful restaurant's strategy is to charge a low enough price to generate excess demand. Here, the good private schools charge a high enough price to repel low types away.

2.2 The Model

A private school may decide to enter or not the market. We model this entry decision as a two stage game. On the first stage the school decides between entering the market or not. And on the second stage it sets its price and compete with the public school. We normalize the private school's opportunity cost to zero, so that the decision rule of the firm on the first stage is:

$$\begin{cases} enter & if & \pi \geq 0 \\ don't & if & \pi < 0 \end{cases}$$

where π is its profits.

On the second stage, upon entering the market, the school decides its price policy, which will depend on its cost structure and demand. It chooses the optimum tuition (p^*) in order to maximize the following profit function:

$$\max_{p} \pi(p) = pq(p) - c[q(p)]^2 - F$$

where F is the fixed cost, c is part of the marginal cost and q is the number of students.

The public school's problem is not modeled formally. Implicitly it is assumed that there are resources to finance the public education of the students who decide not to enroll themselves in the public school.

Demand

There is a continuum of students with ability a distributed on the interval $[\underline{\gamma}, \overline{\gamma}]$, with density $f(a)$. The utility of a student with ability a is $U(y-p, \theta, a)$, where y is income, p is the tuition and θ is the average ability of the school he/she attends. So, the only attribute of the school that matters for a student is the average quality of the students attending it. It is a simplifying assumption.

We further assume a multiplicative interaction between θ and a on the utility function, so that $U(y-p, \theta a)$. This interaction is a key assumption and it means that high ability students benefit more from high θ schools than students with lower abilities. Subsection 1.1 discusses in depth the validity of these assumptions.

There are $N+1$ schools in the market, with $s = 0, ..., N$. One public school, that we denote school 0, and N private schools. A students attend the school that gives him/her the higher utility.

Given the preferences and the choices students have, the set of students attending school j is

$$A_j(p., \theta.) = \{a_i | U(y_i - p_j, \theta_j a_i) \geq U(y_i - p_k, \theta_k a_i), \forall k \neq j\} \tag{1}$$

The share of students attending school j is given by:

$$s_j = \int_{A_j} f(a) da \tag{2}$$

And the quality of school j, θ_j - the mean ability of the students attending it - is given by:

$$\theta_j = \int_{A_j} a f(a) da \tag{3}$$

Note the simultaneity of equations 1 and 3: while A_j is the set of students who prefer school j, it depends on θ_j, which is equal the integral over A_j.

Schools

We assume that the public school charges zero tuition, and that the private schools compete in price. The public school has a passive role in the model, in the sense that it does not react to the private schools actions. The profit of private school j is given by:

$$\max_{p_j} p_j s_j M - C(s_j M) - F_j \tag{4}$$

Where M is the market size $C(s_j M)$ is variable cost and F_j is sunk entry cost.

Equilibrium

We characterize and study the equilibrium of the market game in two steps. First, we adopt an agnostic view of the pricing game, and say that at the right prices the equilibrium holds and is as described. Second, we look at the properties of the Bertrand-Nash equilibrium. The effect of the social interaction has important implications on pricing and at the overall equilibrium.

The following proposition states that in any equilibrium the private schools charge a higher price and have a higher average θ than the public school.

Proposition 1: In any equilibrium, it must be that $\theta_{pr} > \theta_{pu}$ and $p > 0$.

Proof: For simplicity, suppose there is one private school and one public only, and that both schools have the same average quality of the student body, $\bar{\theta}$. In this case, the private school charges $p = 0$, and all students are indifferent between the private and the public institution. It is not an equilibrium, since the private school needs to charge a positive price to have positive profits.

***Result 1**:* Once a private school first enters a market with public schools only, it cream skims the best students.

The following proposition states that there is an equilibrium with N private schools where segregation prevails.

Proposition 2: There is an segregated equilibrium where $\theta_0 < \theta_1 < ... < \theta_N$, and $0 < p_1 < ... < p_N$.

Proof: Suppose not. All private schools have the same average quality and charge the same prices. This situation is indeed an equilibrium, but it is an unstable one. We come back to this point later. Let's focus on two schools, i and j, such that, $\theta_i^0 = \theta_j^0$ and $p_i^0 = p_j^0$. Suppose one of them, let's say j receives a positive shock on its quality, such that $\theta_j^1 > \theta_i^0$. Now if both schools charge the same price, all students strictly prefer j. In this situation it is optimal for j to raise its price: $p_j^1 > p_j^0 = p_i^0$. All the students in school j have an utility loss due to the higher price but, by the multiplicativity assumption, the students with higher ability benefit more from the higher average quality. Some of the students with lower ability find it optimal to transfer to the cheaper school i. When this process ceases, the new equilibrium will have $\theta_j > \theta_i$ and $p_j > p_j$. If we apply this argument to all pairs of private schools, we get the complete ordering for the N schools.

***Result* 2**: The model predicts a market with quality differentiation among the private schools, but similar qualities among the public schools.

The following corollary describes the indifferent (or 'boundary') students between neighbor schools.

Corollary: There are points $\underline{\gamma} < \gamma_{01} < \gamma_{12} < ... < \gamma_{N-1N} < \overline{\gamma}$, such that if student i attends school j then $a_i \in [\gamma_{j-1j}, \gamma_{jj+1}]$. Also, $U(y-p_{k-1}, \theta_{k-1}\gamma_{k-1k}) = U(y-p_k, \theta_k\gamma_{k-1k})$, for any γ_{k-1k} for $k = 1, ..., N$.

Proposition 2 and this last corollary implies that we can re-write equation 3 as

$$\theta_j = \int_{\gamma_{j-1j}}^{\gamma_{jj+1}} af(a)da \tag{5}$$

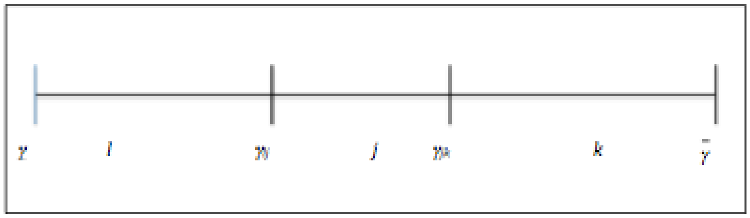
Now, we have a deeper look at the pricing game played by the schools. From equation 4, we derive the best reply function for firm j:

$$s_j(\theta, p)M + p_j \frac{ds_j}{dp_j}M - \frac{dC(s_j, M)}{dp_j} = 0$$

Using the fact that the derivative of the cost function is $C_{p_j} = \frac{dC}{dq}\frac{ds_j}{dp_j}M$, and rearranging terms we find the following expression:

$$s_j(\theta, p) + \left(p_j - \frac{dC}{dq_j}\right)\frac{ds_j}{dp_j} = 0 \tag{6}$$

Since in any equilibrium we have $p \geq C_q$, a necessary condition for an interior equilibrium is $\frac{ds_j}{dp_j} < 0$. A negatively sloped demand curve is a standard assumption in traditional demand analysis, but it is not a correct one in this context. In order to see this, lets assume there are three schools on the market, k, l and m, as shown in the next figure.



The total derivative of school j's market share with respect to its price is:

$$\frac{ds_j(\theta, p)}{dp_j} = \frac{\partial s_j}{\partial \theta_j}\frac{\partial \theta_j}{\partial p_j} + \frac{\partial s_j}{\partial \theta_l}\frac{\partial \theta_l}{\partial p_j} + \frac{\partial s_j}{\partial \theta_k}\frac{\partial \theta_k}{\partial p_j} + \frac{\partial s_j}{\partial p_j} \tag{7}$$

Since this is a comparative statics on the Bertrand-Nash equilibrium, we consider an unilateral deviation of firm j. The effect of a change in price goes four ways: the first term on the right hand side is the effect through the own quality θ_j, the second and third terms are through the quality of the neighbors, θ_l and θ_k, and the last term is the direct price effect. This last term is always negative, but the other ones may not be. We need to find the sign of the other partial effects in equation 7.

Define an intermediate school as one which quality is nor the lowest - the public is the lowest - neither the highest. In this case, school j is the intermediate. The following theorem under which the segregated equilibrium exists.

Theorem: There exists a segregated equilibrium if and only if $\frac{\partial s_j}{\partial \theta_j}\frac{\partial \theta_j}{\partial p_j} + \frac{\partial s_j}{\partial \theta_k}\frac{\partial \theta_k}{\partial p_j} < -\frac{\partial s_j}{\partial p_j} - \frac{\partial s_j}{\partial \theta_l}\frac{\partial \theta_l}{\partial p_j}$.

Proof: See appendix.

Result 3: Otherwise, one of the two situations holds:

(i) Without switching costs, the students enter a game of 'following the smart'. The students are attending the public and one of the private schools. The other one is empty. It is worth for the above average students to transfer to the empty school, and so they do. On the next period, the remaining students, the below average group, find it worth to transfer to the same school the smart students went, and they do so. This process keep going *ad infinitum.*

(ii) With high enough switching costs, the schools converge to a pooling equilibrium, where the schools have the same quality. Despite being an unstable equilibrium, the switching costs bring stability to this situation.

Entry

Our goal is to see the effect of the social interactions on the market structure. Specifically, we want to show the role of the range of abilities on entry, and to see how the entry threshold ratios (see Bresnahan and Reiss, 1991) are changed with the existence of social interactions among the students. In order to be able to analyze some specific details of these problems we need to make some parametric assumptions.

First, the indirect utility function of a student is $U = y - p + \theta a$. So, for a student indifferent between school j and $j+1$ we have $y - p_j + \theta_j \gamma_{jj+1} = y - p_{j+1} + \theta_{j+1}\gamma_{jj+1}$. It implies the following relationship between prices and quality on both schools:

$$p_{j+1} - p_j = (\theta_{j+1} - \theta_j)\,\gamma_{jj+1} \tag{8}$$

Equation 8 says that the difference in prices between any two schools is proportional to the difference in the qualities they have. This result has an important implication: the profitability of private schools in a market depends on the extension of the range of abilities $\left[\underline{\gamma}, \overline{\gamma}\right]$. For instance, more homogeneous societies - where students have similar abilities - give little or no room for a private school to enter the market and attract the best students to positively differentiate itself from the public school. On the other hand, in heterogeneous societies - with longer range of abilities - private schools are able to attract the good student and have a substantial differentiation in relation to the public schools. The same reasoning is for among the private schools: a longer range of ability imply in a higher profitability for them.

***Result* 4**: The profitability of the private schools is positively related to the extension of the range ability. *Ceteris paribus*, a market with a wider range of abilities weakly have more firms in equilibrium.

In order to be able to analiticaly solve the model, we further assume that students' ability are uniformly distributed on the interval $[0, 1]$, that there is no variable costs, so that $C_j = F$, and M is the size of the market. The following proposition characterizes the equilibrium of this market.

Proposition 3: The price and boundary student of the n-th school in this market is $p_n = \frac{1}{2^{2(N-n+1)}}$ and $\gamma_{nn-1} = \frac{1}{2^{N-n+1}}$.

Proof: See appendix.

We use this result to compute table 1A.

The first column displays the entry order. We assume the that the entry order defines the quality: the first entrants becomes the best schools. For convenience, now we invert the notation and call the top school #1. From the second to the fifth column, it shows the lower boundary student, the price that school charges in equilibrium, the market share and the profit, respectively. For example, the boundary students for school #2 are $\gamma_{12} = 0.5$ and $\gamma_{23} = 0.25$, it charges a tuition of 0.63, has a market share equal to 0.25 and has profit equal to $0.0156M - F$. These columns are used to compute the entry threshold and the entry threshold ratio, on columns sixth and seventh. The entry threshold is the per firm market size M_n/n need for the last firm to break even. For example, for the second entrant to break even we need $M_2/2 = 32F$.

The last column displays the entry threshold ratios: $(M_{n+1}/n+1)/(M_n/n)$. We want to compare our results to the entry pattern described by Bresnahan and Reiss (1991). Their results are about an entry game in an homogeneous good's market where firms play Cournot on the competition stage. They find greater than one entry threshold ratios for the first firms to enter, and the thresholds decrease down to one as competition intensifies. The rationale behind this important result is that market size needs to more than double for the second entrant to enter, due to the loss in the ability that a duopoly has, as compared to a monopoly, to extract rent from consumers. As the firms enter the market, it approaches perfect competition, the margins vanished and the market increase necessary for one more firm to enter converges to a constant number. Therefore, entry threshold ratios converges to one.

We find a very different pattern for the entry threshold ratios. They start equal to 4, but increase after that. This diffence is due to the loss in ability to cream skim the best students. Those were taken by the first entrants. The last entrants have a significant lower quality than the first ones. Since the price is proportional to the difference in qualities, these last entrants need to charge a significantly lower tuition, and so they need a large number of students to break even. The key point is that we are holding constant the range of abilities. If we let the range of abilities grow as the market grows. we could observe a pattern similar to the one found by Bresnahan and Reiss.

This result has an important implication: the market never fully converges to monopolistic competition.[7] As the market grows, entry happens on the lower end of the quality distribution. At this lower end, the entry of an extra school make their qualities more similar, increasing price competition and decresing margins. It has little or no effect on the top schools. The fourth column of table 1A shows the market share of the firms. Note that firm 1 has half of the market. It still has half of the market irrespective of the number of later entrants. This total non-responsive of firm 1 to the entry of firms 2, 3 etc, is due to the uniform distribution assumption, but has little change when we simulate the market with a normal distribution, for example. This is accordance to the result of Shaked and Sutton (1983), that a market with vertical differentiation may never converges to a competitive environment.

Result 5: The entry of a private school may not increase the overall level of competition in a market. The entry of the *n-th* entrant has a larger impact on the competitive environment of the lower quality school, and a relatively lower impact on the schools located on the upper end of quality distribution.

This result resembles the one found in Becker (1983) where the equilibrium of the market is to have differentiation in equilibrium. Becker analyzes the case where firms are in the good or in the bad equilibrium. In the good equilibrium, firms charge a high price and have high demand. In the bad equilibrium, firms charge a low price and have excess capacity. Here, the equilibrium is a sequence of firms, ranging from high (good) to low (bad) quality.

Result 6: The school market is endogenously segmented by quality. At the high end, the high quality schools charge a high enough price to drive low quality students away, attract only the high ability students and have large profits. At the low end, low quality schools charge a lower price, attract the lower ability schools and have near zero profits.

[7] Since it is a differenciated goods market, it does not make sense to converge to perfect competition.

3 Data

The data we use come from two sources. The first source is the School Census, which is carried out annually by the Ministry of Education and has information on the number of schools and enrollments by municipality. We use the 2009 School Census for our main data set. The information on average income, income inequality and education inequality come from the Demographic Census, carried out by the Brazilian Census Bureau every ten years. Since the information from the 2010 Census is not available yet, we draw our municipality data from the 2000 census. Table 1 in the appendix presents the main variables and descriptive statistics.

Figure 1 plots the relationship between the percentiles of student population and the share of municipalities with private high schools, to start describing the basic features of the data.[8] From this preliminary inspection it seems that the size of the market is a factor that schools considerer when thinking about entry in a new market. Cities in the bottom third of the cumulative distribution of student population are unlikely to have a private school, with the share of cities with private schools increasing almost linearly after that.

Figure 2 describes the behavior of the share of municipalities with at least one private school as average income increases. We can see that, while there is a positive relationship between the two factors, as we would expect, the likelihood of a private school declines when income rises from percentile 33 to the median of the income distribution, which is counter-intuitive. It could be that municipalities in this range of the income distribution have some other characteristic that is correlated with income and with the likelihood of having a private school. The econometric exercises below will shed more light on this issue.

Figures 3 and 4 plot the relationship between the share of municipalities with private schools and income and education inequality, respectively. One can note that, while both indicators are positively related with private education, the relationship is much clearer in the case of the education inequalities. While the probability of having a private school rises slowly when inequality increases from very low values, the relationship is much stronger for the municipalities in the top 20% of the inequality distribution. This variable is a proxy for the range of abilities in the municipalities, which has a crucial role in the theoretical model developed above.

4 Results

4.1 Entry Probability

Table 2 presents the results of the entry model for secondary education. The table reports marginal effects from Probit estimation. Column 1 shows that school

[8]The results for elementary private schools are very similar to the one for high schools and are omitted for space considerations.

population, average income and education inequality (a proxy for the range of abilities in a city) are positively related with the probability of entry in the Brazilian municipalities and all attract statistically significant coefficients. The coefficient on income inequality is not significant at conventional levels though. Column (2) shows, however, that is the interaction between income and the size of the population that is driving the main results. After including the interaction both the individual effects become negative. Column (3) then includes the state dummies, to control for unobserved state effects and the estimation results are maintained. But now, the inequality coefficient is statistically significant at 10% and its interaction with income show that the effect of inequality decreases for richer municipalities, as theory predicted.

As the empirical model includes interactions between variables, it is easier to grasp their effects by means of graphs that simulate their impact on the average probability of entry, using the estimated marginal effects and keeping the other variables at their actual values. Figure 1 uses the estimated marginal effects to describe the behavior of entry probability as population increases, for different values of income. It is clear that entry rises significantly when the population passes the 900 students threshold. When income is low, the rate of increase is lower and keeps increasing until reaching the 70% probability when population reaches its 90th percentile (25,000). When income is high, the entry probability reaches its maximum value when the student population is about 7,000.

Figure 2 performs the same exercise for changes in income, for different population values. It shows firstly that the effect of income is smaller than that of population. For low values of population (around 900), income rises have a small effect on the entry probability. For bigger municipalities, however, income rises quickly make the probability reach its maximum value.

Figure 3 shows that income inequality has a very small effect on the entry probability in the secondary education, for different values of income. Figure 4, however, depicts the important effect that education inequality has on entry, especially at high levels of inequality. It seems, therefore, that for the secondary education, a wider range of abilities is more important for entry than a wider income distribution.

Table 3 presents the marginal effects of the probit regressions for high school education. The results are pretty similar to the ones for secondary education. The interaction between the size of the student population (between 15 and 25 years of age) and average income in column 3 is an important determinant of entry. Differently from the secondary education, however, the separate effects of these two variables are not statistically significant once the interaction is included. Income inequality seems much more important for entry in the case of high school education than for secondary education, a results that deserved further consideration. Moreover, the interaction between average income and income inequality is also more important here than in the case of secondary education. The results indicate that the impact of income inequality is stronger for poorer municipalities, which fits nicely with the prediction of the theoretical model.

Figures 9 to 12 simulate the impact of each of these variables on the probability of high school entry. The results are similar to the ones reported for the case of secondary education, although there are some relevant differences. The size of the student population has an important effect on the likelihood of entry, especially after the number of students is higher than 5,000. For poorer cities, the minimal size for profitable entry is 10,000. The effect of income depends heavily on the size of the population. When the number of high school students is in the range of 3,000, income has basically no effect on entry. As soon as the population reaches 10,000 students, incomes higher than R$86 rise the entry likelihood. In bigger and richer municipalities, there is 100% probability of existence of a high school.

Income inequality impacts entry in way that is heavily dependent on average income, as Figure 11 shows. When average income is low, inequality must be high so that the elite can sustain a private school. As soon as income rises, that are enough rich people in the community so as to make a private school profitable, so that inequality does not play any role, as predicted by the theoretical model. Education inequality in turn has an important impact on entry probability, especially after the ratio of the share of population with high school or more and the share with less then eight years is higher than 5%.

4.2 Long Differences

Table 4 presents the results of the long-differences regressions, relating entry of private schools in a city between 1995 and 2000 with the change in the explanatory variables between 1990 and 2000. These results are important since by estimating the model in differences we are controlling for city fixed effects that might be correlated with the explanatory variables of interest. Columns (1) and (2) report the results for the secondary education and show that both school population and education inequality are important determinants of entry, even after controlling for the municipalities' unobserved heterogeneity. Moreover, the results of column (2), which includes interactions between income and population and also between income and inequality, show that the first is also (marginally) statistically significant. Columns (3) and (4) report the results of the same specifications for high schools and show that the results are very similar to those of the secondary schools, with population and education inequality remaining statistically significant. The fact that income and income inequality loose significance in the long-differences specification may reflect the well-known problem of measurement errors in income, which are magnified in the differences specifications. Alternatively, the estimates of the levels' specifications (Tables 2 and 3) may suffer from omitted variable bias, due to the omission of the city fixed effects.

4.3 School Quality

Table 5 examines the effect of entry on the quality of the public schools in the municipality, as measured by its math's and readings test scores. The test scores are measured by the results of Prova Brasil, a national exam that is carried out amongst all the country´s public school students. In this exercise we are comparing the average score of public schools in the municipalities with a private school with that of the cities with public schools only.

Column (1) shows that math's test score in public schools are significantly lower in municipalities where there is a private school, even after controlling for other possible determinants of quality, such as income, size and inequality. Column (2) includes state dummies, to control for regional effects and show that the result remains qualitatively the same, even if the private school effect is halved. This result is consistent with private schools entering the market to capture the best students of the public system, although it is also possible that entry is facilitated by a poor performing school system. Columns (3) and (4) repeat the specifications for the reading test scores and show that the results remain the same. These results are consistent with our theoretical model, although they should be subjected to further robustness tests, especially with the use of panel data to confirm that they are not due to reverse causality.

We then examine the impact of private school entry on school quality using a different measure of quality, the results of the ENEM exams. The ENEM is a voluntary exam carried out at the end of high school, and it is the only examination that is also performed by private school students. Columns (1) and (2) of Table 6 use ENEM data to examine the effect of private school entry on the average quality of education of the public system. The results are similar to the ones presented in Table 5, even with the use of a completely different data set. It shows that municipalities with a private school have on average scores that are significantly lower than their public schools only counterparts, even after controlling for other possible determinants of test scores.

But does private school entry improve the average quality of education in the municipality or only increases inequality? Columns (3) and (4) examine this issue by estimating the impact of private school entry on the municipalities' average test scores, including those of the private schools themselves. The results show that there is no correlation between the overall scores of the municipalities and the presence of a private school, despite the fact that private schools are associated with lower grades in the public system. It seems therefore that private schools indeed attract the best students of the public system, but this mechanism does not improve the city overall standards. The gains brought about by the fact that the best students are studying together (peer effects), seems to be counterweighted by the loss imposed on the public school students, which no longer have the best students as references and are much bigger in number. Therefore, the results could be indicating that private school entry is generating inequality, without improving overall quality. But, we should be careful before jumping to important conclusion, since we must first submit these results to further robustness tests, especially with the use of panel data on test scores and private school entry.

5 Conclusions

While many papers have modeled the competition between private and public schools and the process of school choice by students and families, no study thus far has examined the determinants of entry of a private school in a city where all students are attending public schools. This paper develops a model of entry and tests its predictions using data from Brazilian municipalities. The model predicts that both the cream skimming of good students as well as income inequality facilitates entry.

The empirical part of the paper confirms most of the predictions of the theory. Entry is positively related to the interaction between the size of the market and income, to education and income inequalities and the effect of income inequality declines with average income. We find very similar results when we model secondary and high schools.

While promising, the results presented in this paper are still preliminary. We intend to collect data on school costs by municipality, which basically consist of teachers' wages, and on the quality and number of the public schools. Moreover, we intend to collect panel data on the number of schools by municipality, to examine whether these results are robust to the inclusion of municipality fixed effects. Finally, we intend to model the market share of the private firm once it decides to entry the market.

References

[1] ALTONJI, G. J., HUANG, C. and TABER, C. (2010) Estimating the Cream Skimming Effect of School Choice. NBER Working paper 16579.

[2] BRESNAHAN, T AND REISS, P. (1991) Entry and Competition in Concentrated Markets. Journal of Political Economy, vol. 99 (5), pp. 977-1009.

[3] EPPLE, D. and Romano, R. (1998) Competition between Private and Public Schools, Vouchers and Peer-Group Effects. American Economics Review. Vol. 88 (1), pp. 33-62.

[4] HOTELLING, H. (1929) Stability in Competition. Economic Journal. Vol. 39, pp. 41-57.

[5] MACEWAN, P., URQUIOLA, M. and VEGAS, EMILIANA (2008). School Choice, Stratification and Information on School Performance: lessons from Chile. Economia, Spring.

Appendices

Appendix 1 - Existence of Segregated Equilibrium

Lemma 1: Let $\frac{\partial \gamma_{lj}}{\partial p_j}$ and $\frac{\partial \gamma_{jk}}{\partial p_j}$ be the change on both bounds of school j due to a change in p_j. The following holds: $\left|\frac{\partial \gamma_{lj}}{\partial p_j}\right| > \left|\frac{\partial \gamma_{jk}}{\partial p_j}\right|$.

Proof: The multiplicative assumption makes utility equal to $U(y-p, \theta a)$, and the marginal rate of substitution between $(y-p)$ and school quality θ to be $1/a$. It implies that a student with lower ability a, attending school j, needs a higher increase in school quality θ to remain indifferent to an increase in tuition, for example. Therefore, an increase (decrease) in tuition leads to greater loss (gain) in utility to the lower ability students than to the more skilled students of school j. Since the change in quality θ_j is the same for all students, a price increase (decrease) leads more low ability students to switch to school l than students switching to k: $\left|\frac{\partial \gamma_{lj}}{\partial p_j}\right| > \left|\frac{\partial \gamma_{jk}}{\partial p_j}\right|$.

Lemma 2: $\frac{\partial \theta_j}{\partial p_j} > 0$ iff

$$-\frac{\partial \gamma_{lj}/\partial p_j}{\partial \gamma_{jk}/\partial p_j} < \frac{f(\gamma_{jk})(\gamma_{jk} - \theta_j)}{f(\gamma_{lj})(\theta_j - \gamma_{lj})} \tag{9}$$

Lemma 3: $\frac{\partial \theta_k}{\partial p_j} > 0$ iff

$$-\frac{\partial \gamma_{lj}/\partial p_j}{\partial \gamma_{jk}/\partial p_j} < \frac{f(\gamma_{jk})(\gamma_{jk} - \theta_j)}{f(\gamma_{lj})(\theta_j - \gamma_{lj})} \tag{10}$$

Lemma 4: $\frac{\partial \theta_l}{\partial p_j} > 0$ iff

$$-\frac{\partial \gamma_{lj}/\partial p_j}{\partial \gamma_{jk}/\partial p_j} < \frac{f(\gamma_{jk})(\gamma_{jk} - \theta_j)}{f(\gamma_{lj})(\theta_j - \gamma_{lj})} \tag{11}$$

Lemma 5: $\frac{ds_j}{dp_j} < 0$ iff $\frac{\partial s_j}{\partial \theta_j}\frac{\partial \theta_j}{\partial p_j} + \frac{\partial s_j}{\partial \theta_k}\frac{\partial \theta_k}{\partial p_j} < -\frac{\partial s_j}{\partial p_j} - \frac{\partial s_j}{\partial \theta_l}\frac{\partial \theta_l}{\partial p_j}$.

Appendix 1 - Equilibrium with uniform distribution

We start the case of one public and two private school, so $n = 0, 1, 2$.

The profit functions of schools 1 to 3 becomes:

$$\max_{p_1} p_1 (\gamma_{12} - \gamma_{01}) M - F$$

$$\max_{p_2} p_2 (\gamma_{23} - \gamma_{12}) M - F$$

$$\max_{p_3} p_3 (1 - \gamma_{23}) M - F$$

The equilibrium of the market just described becomes:$p_3 = \frac{1}{4}$, $\gamma_{23} = \frac{1}{2}$, $p_2 = \frac{1}{16}$, $\gamma_{12} = \frac{1}{4}$, $p_1 = \frac{1}{32}$, $\gamma_{01} = \frac{1}{8}$

Tables and Figures

Table 1 – Variables Definitions and Descriptive Statistics

Variable	Definition	Mean	Min	Max
Number of private Secondary Schools	Number of Private Schools in the Municipality	3.67	0	1,310
Number of private High School	Number of Private Schools in the Municipality	1.60	0	821
Income Inequality	Theil Index	0.521	0.185	1.271
Average Income	Average income per capita	170	28	954
Education Inequality	Ratio of adults with high school or more and adults with less than 8 years of education	0.046	0	1.03
Student Population – Secondary	Student Population – Younger than 15	9,127	218	2,592,829
Student Population – High School	Student Population – Between 15 and 25	30,833	795	10,400,000

Table 1A: Entry Thresholds under the Uniform Distribution

n	*γ*	*p*	*s*	*π*	*Mn/n* (*)	*entry threshold ratio* (**)
1	0.500	0.250	0.500	0.1250000 M-F	8F	
2	0.250	0.063	0.250	0.0156250 M-F	32F	4.0
3	0.125	0.016	0.125	0.0019531 M-F	171F	5.3
4	0.063	0.004	0.063	0.0002441 M-F	1024F	6.0
5	0.031	0.001	0.031	0.0000305 M-F	6554F	6.4
6	0.016	0.000	0.016	0.0000038 M-F	43691F	6.7
7	0.008	0.000	0.008	0.0000005 M-F	299593F	6.9
8	0.004	0.000	0.004	0.0000001 M-F	2097152F	7.0
9	0.002	0.000	0.002	0.0000000 M-F	14913081F	7.1
10	0.001	0.000	0.001	0.0000000 M-F	107374182F	7.2
15	0.000	0.000	0.000	0.0000000 M-F	2.34562E+11F	7.5
20	0.000	0.000	0.000	0.0000000 M-F	5.76461E+16F	7.6

(*) *Mn* is the market size that generates zero profits for the n firm. Mn/n is the per firm market size that generates zero profits.
Mn/n is the entry threshold of the n entrant.
(**) It is the ratio of the entry thresholds Mn/n.

TABLE 2 – ENTRY – FUNDAMENTAL EDUCATION

Dependent Variable: Private School Entry - 2009			
	(1)	**(2)**	**(3)**
Ln (school-age population)	0.337**	-0.510 **	-0.263 **
	0.007	*0.116*	*0.106*
Ln (average income)	0.056 **	-1.394 **	-0.511 **
	0.027	*0.197*	*0.186*
Education Inequality	1.863 **	2.308 **	1.401 **
	0.521	*0.324*	*0.334*
Income Inequality	0.020	-0.278	0.847
	0.065	*0.597*	*0.555*
Teachers Wages	-0.061**	-0.058**	0.016
	0.013	0.012	0.012
Ln(population) * Ln (income)	-	0.171 **	0.113 **
		0.023	*0.021*
Income Inequality* Ln (income)	-	0.059	-0.173 *
		0.118	*0.109*
State Dummies	*No*	*No*	*Yes*
Constant	-11.866	19.891	0.248
	0.797	*4.420*	*5.238*
Obs	3,309	3,309	3,309

Notes: Marginal Effects after Probit. Robust standard errors in italics.
Only Municipalities with less than 6 private schools
** and * Denote statistical significance at 5% and 10%.

TABLE 3 – ENTRY - HIGH SCHOOL

Dependent Variable: Private School Entry - 2009			
	(1)	**(2)**	**(3)**
Ln (school population)	0.267 **	-0.187	-0.041
	0.007	*0.122*	*0.129*
Ln (average income)	0.148 **	-0.472 **	-0.156
	0.032	*0.202*	*0.213*
Education Inequality	1.342 **	1.225 **	1.552 **
	0.469	*0.452*	*0.400*
Income Inequality	0.171 **	1.224 **	1.637 **
	0.051	*0.453*	*0.462*
Teachers Wages	-0.026**	-0.030**	-0.014
	0.011	0.011	0.011
Ln(population) * Ln (income)	-	0.089 **	0.060 **
		0.024	*0.025*
Income Inequality* Ln (income)	-	-0.205 **	-0.292 **
		0.088	*0.089*
State Dummies	*No*	*No*	*Yes*
Constant	-17.333	1.568	-8.589
	1.083	*6.144*	*6.957*
Obs	3,507	3,507	3,507

Notes: Marginal Effects after Probit. Robust standard errors in italics.
Only Municipalities with less than 6 private schools
** and * Denote statistical significance at 5% and 10%.

TABLE 4 – LONG DIFFERENCES:

Dependent Variable: Change in Number of Schools: 2000-1995				
	(1)	**(2)**	**(3)**	**(4)**
	Fundamental		**High School**	
D ln(school population)	0.566**	0.423 **	0.118 **	0.145**
(2000-1990)	*0.084*	*0.117*	*0.038*	*0.059*
D ln(average income)	-0.019	0.004	-0.048	-0.034
(2000-1990)	0.082	*0.083*	*0.042*	*0.046*
D Education Inequality	3.242**	3.197 **	1.851 **	1.852**
(2000-1990)	*0.314*	*0.316*	*0.161*	*0.161*
D Income Inequality	0.076	0.051	0.007	0.097
(2000-1990)	*0.142*	*0.207*	*0.074*	*0.108*
D(population) * D (income)	-	0.485 *	-	-0.076
(2000-1990)		*0.277*		*0.137*
D(Inequality)* D(income)	-	0.045	-	-0.274
(2000-1990)		*0.465*		*0.240*
Constant	0.139	0.140	-0.071	-0.074
	0.0317	*0.032*	*0.017*	*0.019*
Obs	4495	4495	4495	4495

Notes: Robust Standard Errors in Parenthesis
Only Municipalities with less than 6 private schools
** and * Denote statistical significance at 5% and 10%.

TABLE 5 – EFFECT OF ENTRY ON PUBLIC SCHOOL QUALITY

Dependent Variable: Average Test Sore in Prova Brasil – 2007 – 8th Grade				
	(1)	**(2)**	**(3)**	**(4)**
	Math	**Math**	**Portuguese**	**Portuguese**
Private School	-0.013**	-0.005 **	-0.013**	-0.005 **
	0.002	*0.002*	*0.002*	*0.002*
Ln (Average income)	0.088**	0.073**	0.078**	0.067**
	0.002	*0.003*	0.002	*0.003*
Education Inequality	-0.355**	-0.346 **	-0.198**	-0.176 **
	0.046	*0.050*	*0.044*	*0.048*
Income Inequality	0.041**	0.040**	0.030**	0.026**
	0.005	*0.005*	*0.005*	*0.005*
Ln (School Population)	-0.010**	-0.009**	-0.004**	-0.005**
	0.001	*0.001*	*0.001*	*0.001*
State Dumies	*No*	*Yes*	*No*	*Yes*
Constant	5.126	5.121	5.065	5.074
	0.015	*0.023*	*0.014*	*0.022*
Obs	4495	4495	4495	4495

Notes: Robust Standard Errors in Parenthesis
** and * Denote statistical significance at 5% and 10%.

TABLE 6 – EFFECT OF ENTRY ON SCHOOL QUALITY

Dependent Variable: Average ENEM Test Score – 2009 – 11th Grade				
	(1)	**(2)**	**(3)**	**(4)**
	Average Public School Score		Average School Score	
Private School	-0.007**	-0.006**	-0.0001	0.0001
	0.002	0.002	0.0026	0.0026
Ln (school population)	-0.002	0.002	-0.001	0.002*
	0.002	*0.002*	*0.001*	*0.001*
Ln (average income)	0.066 **	0.052 **	0.065 **	0.052**
	0.003	0.003	*0.002*	*0.003*
Education Inequality	0.038	0.001	0.107 **	0.078 **
	0.035	*0.036*	*0.036*	*0.060*
Income Inequality	-0.082 **	-0.044 **	-0.079**	-0.042 **
	0.008	*0.008*	*0.008*	*0.008*
State Dummies	*No*	*Yes*	*No*	*Yes*
Constant	5.956	5.931	5.954	5.924
	0.0189	*0.028*	*0.018*	*0.028*
Obs	4,618	4,618	4,618	4,618

Notes: Robust Standard Errors in Parenthesis

** and * Denote statistical significance at 5% and 10%.

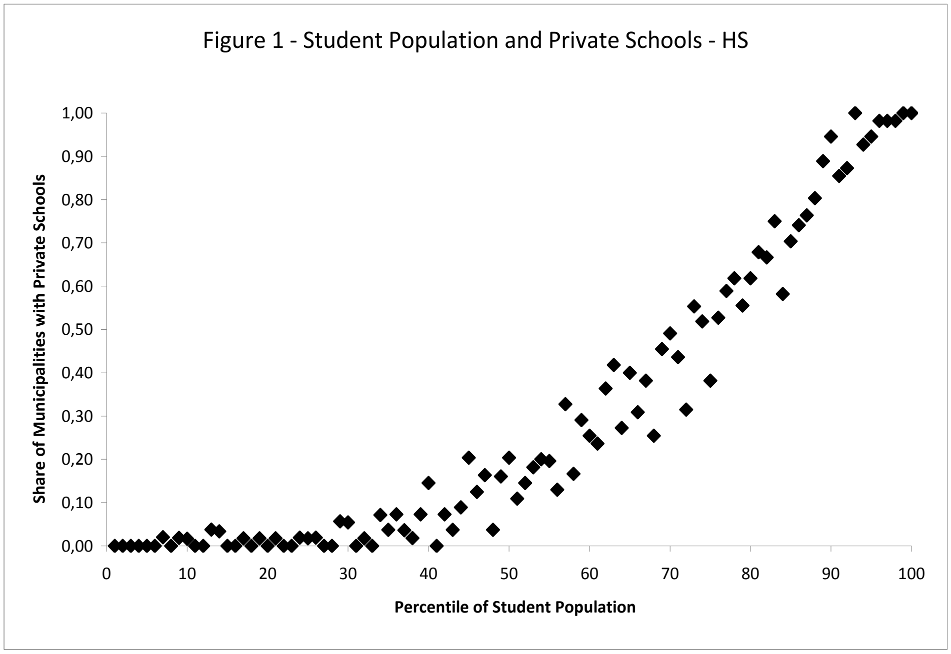


Figure 1 - Student Population and Private Schools - HS
Share of Municipalities with Private Schools
1,00
0,90
0,80
0,70
0,60
0,50
0,40
0,30
0,20
0,10
0,00
0
10
20
30
40
50
60
70
80
90
100
Percentile of Student Population

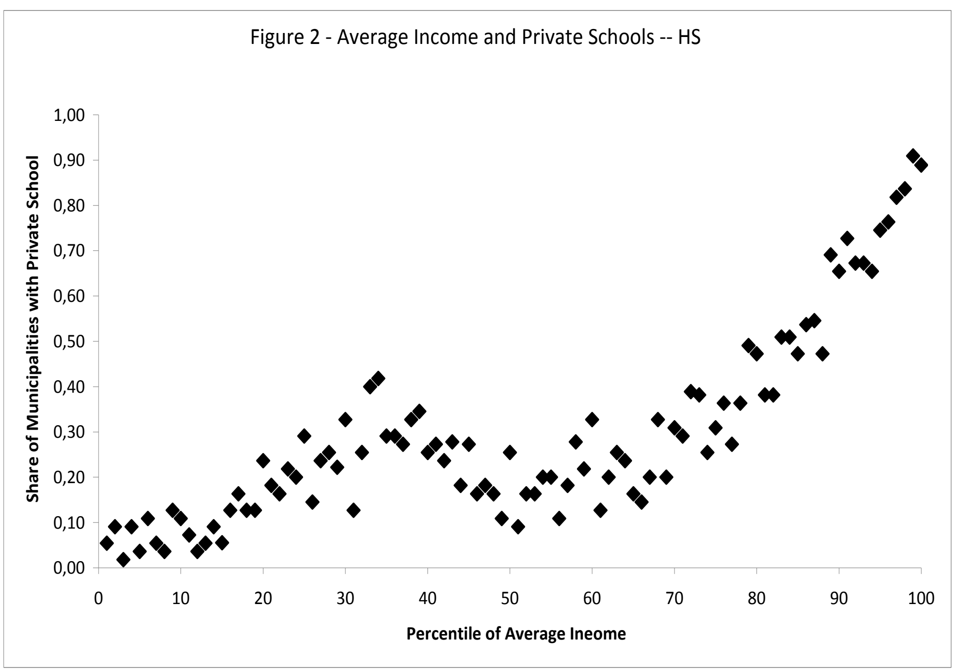


Figure 2 - Average Income and Private Schools -- HS
Share of Municipalities with Private School
1,00
0,90
0,80
0,70
0,60
0,50
0,40
0,30
0,20
0,10
0,00
0
10
20
30
40
50
60
70
80
90
100
Percentile of Average Ineome

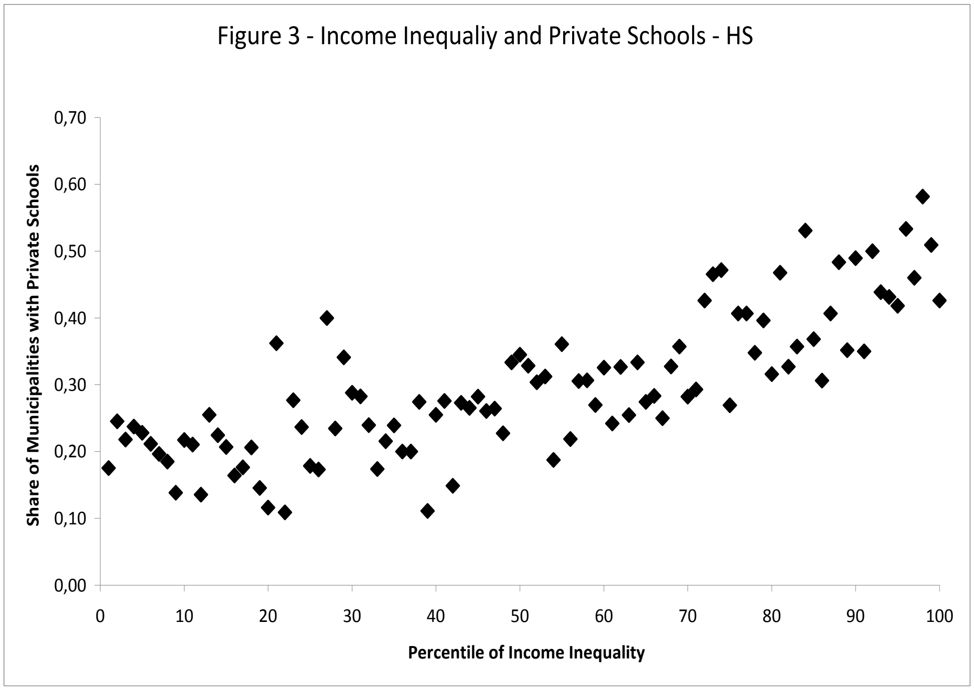


Figure 3 - Income Inequaliy and Private Schools - HS
Share of Municipalities with Private Schools
0,00
0,10
0,20
0,30
0,40
0,50
0,60
0,70
0
10
20
30
40
50
60
70
80
90
100
Percentile of Income Inequality

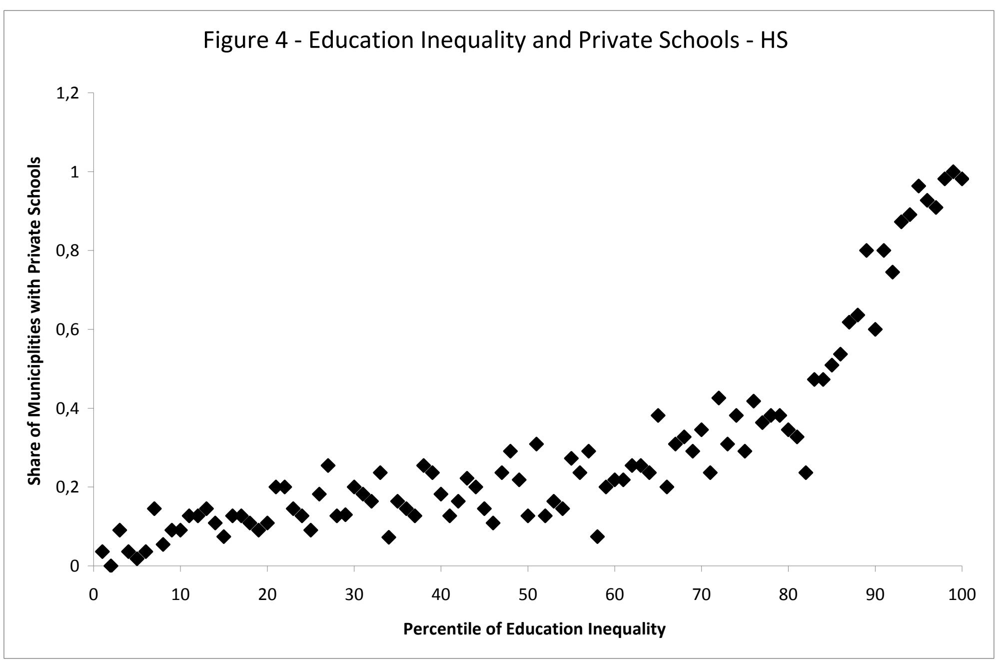


Figure 4 - Education Inequality and Private Schools - HS
Share of Municiplities with Private Schools
0
0,2
0,4
0,6
0,8
1
1,2
0
10
20
30
40
50
60
70
80
90
100
Percentile of Education Inequality

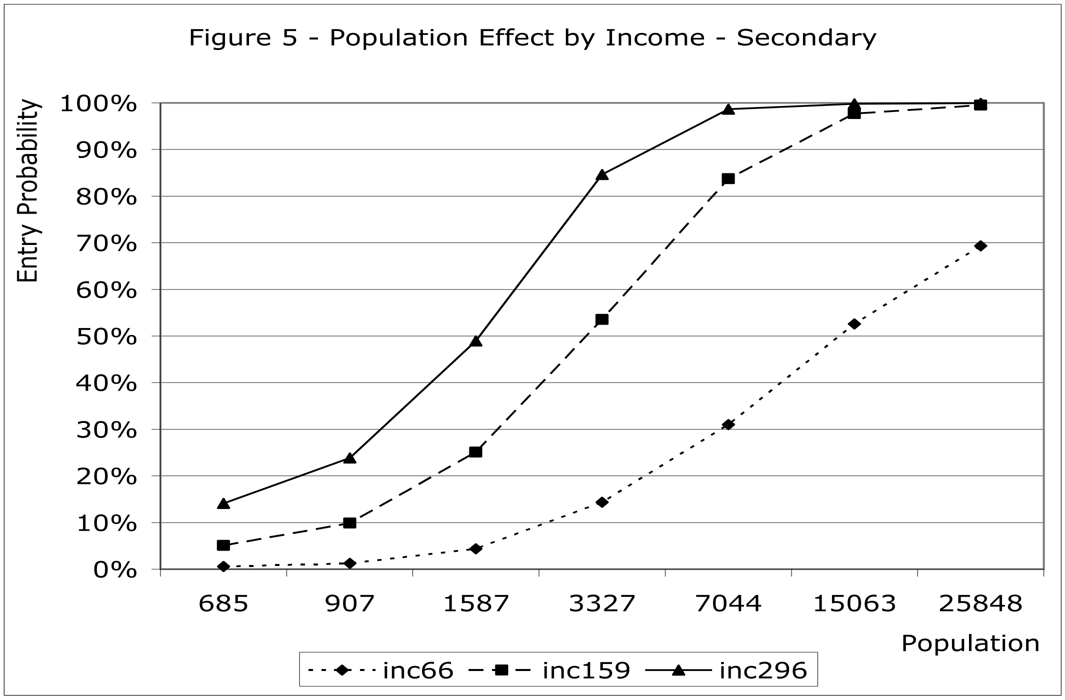


Figure 5 - Population Effect by Income - Secondary
Entry Probability
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
685
907
1587
3327
7044
15063
25848
Population
inc66
inc159
inc296

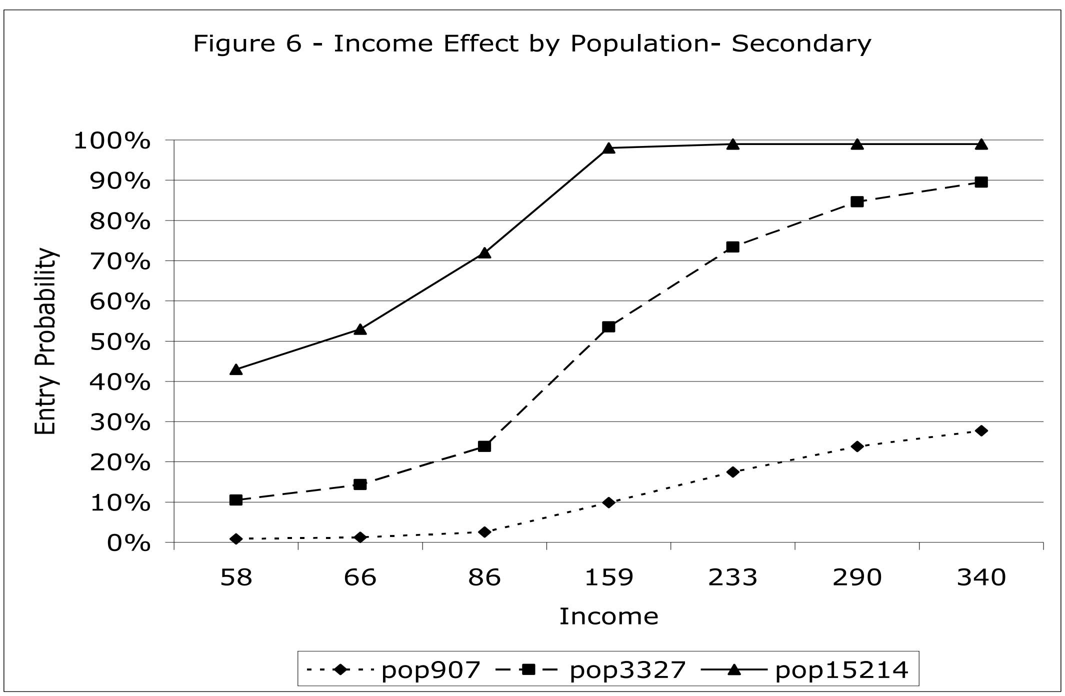


Figure 6 - Income Effect by Population- Secondary
Entry Probability
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
58
66
86
159
233
290
340
Income
pop907
pop3327
pop15214

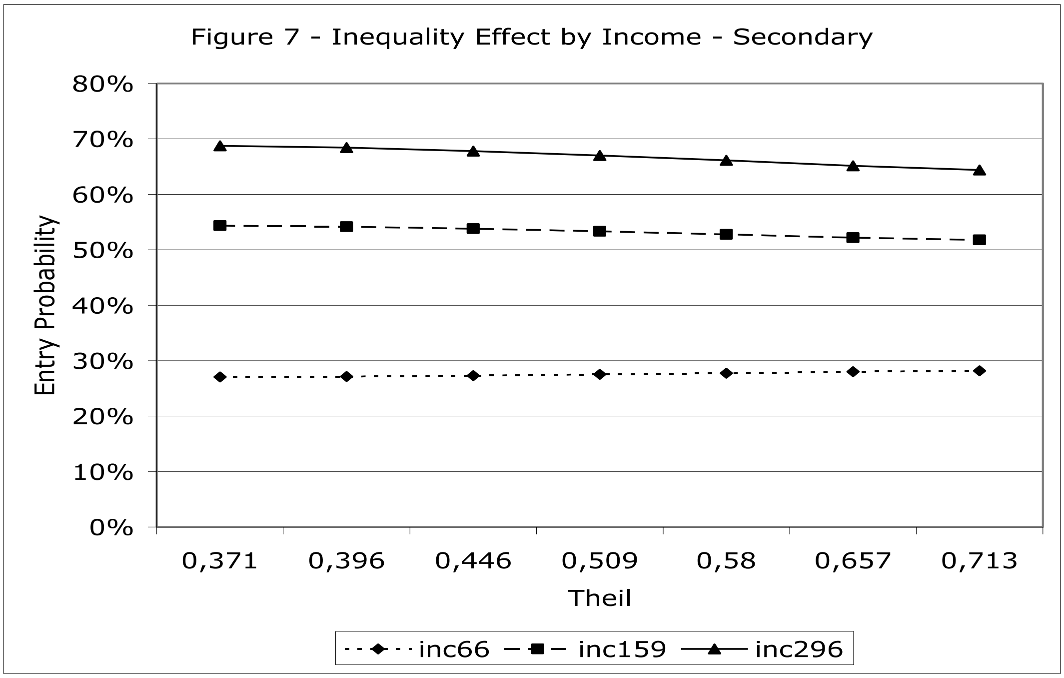


Figure 7 - Inequality Effect by Income - Secondary
80%
70%
60%
50%
40%
30%
20%
10%
0%
Entry Probability
0,371
0,396
0,446
0,509
0,58
0,657
0,713
Theil
inc66
inc159
inc296

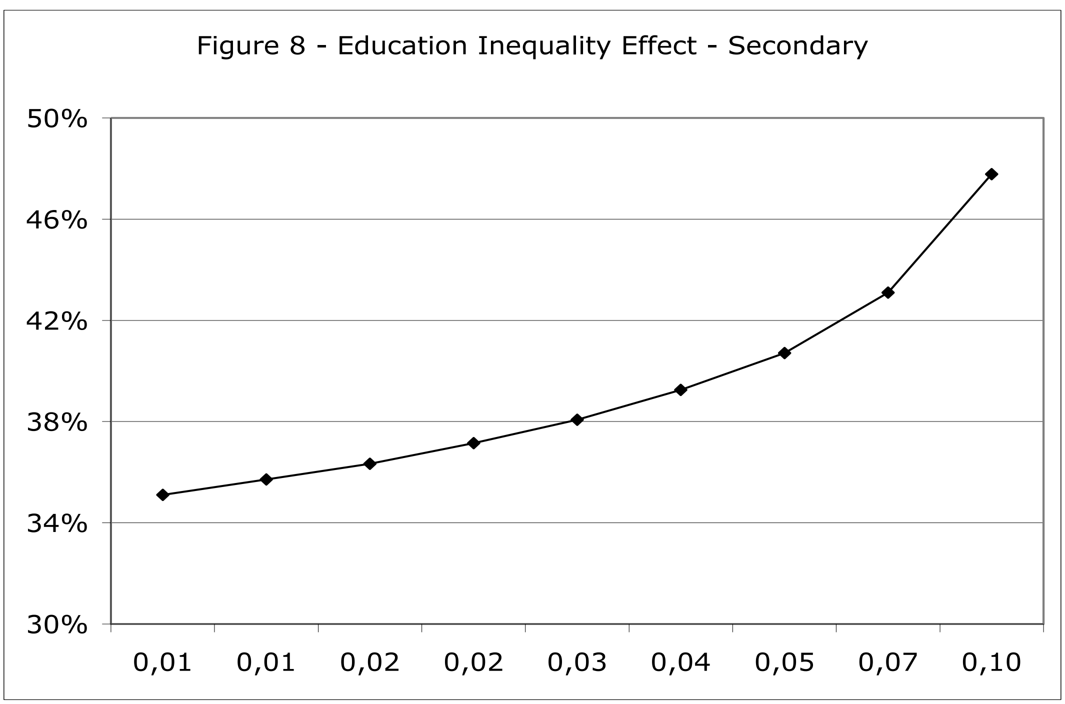


Figure 8 - Education Inequality Effect - Secondary
50%
46%
42%
38%
34%
30%
0,01
0,01
0,02
0,02
0,03
0,04
0,05
0,07
0,10

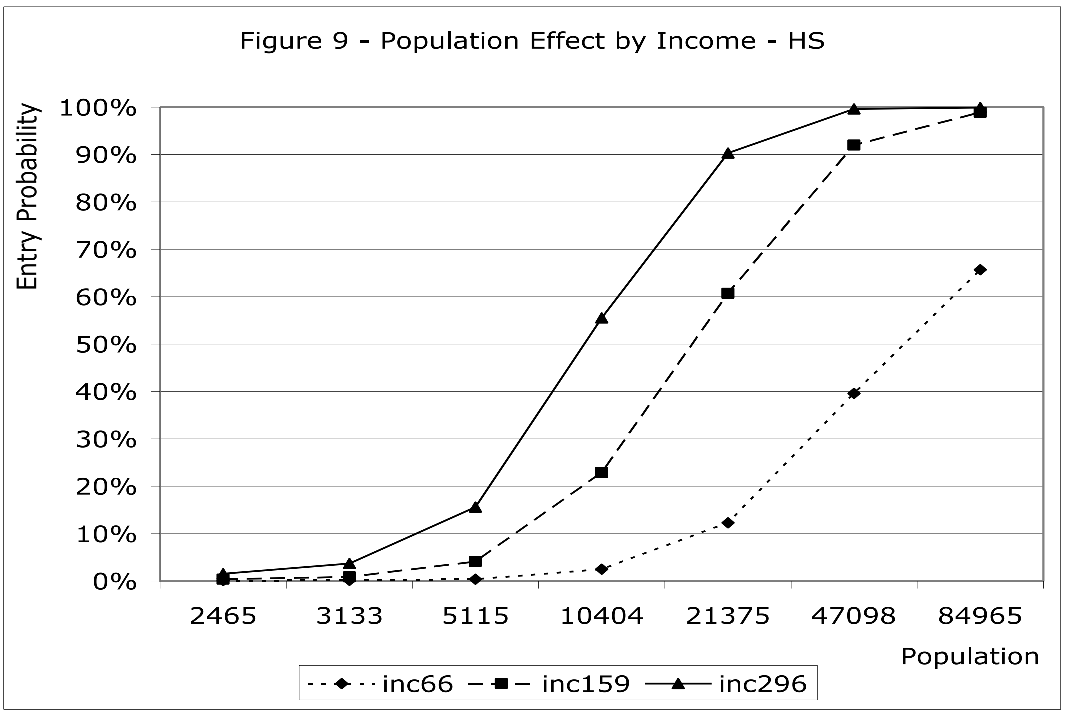


Figure 9 - Population Effect by Income - HS
Entry Probability
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
2465
3133
5115
10404
21375
47098
84965
Population
inc66
inc159
inc296

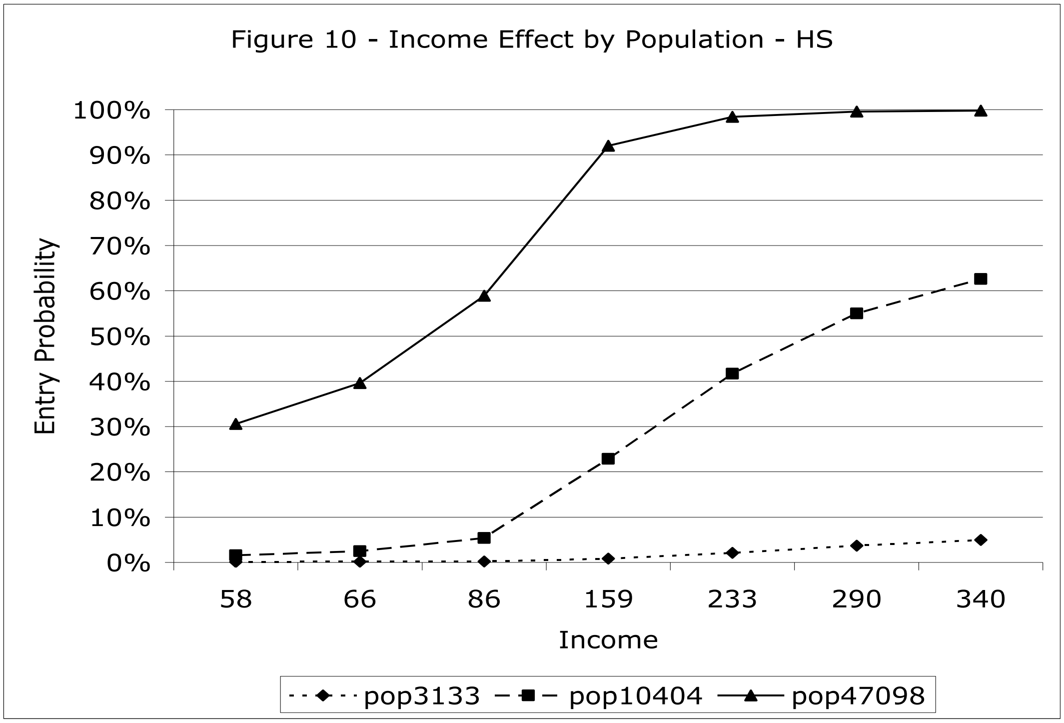


Figure 10 - Income Effect by Population - HS
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Entry Probability
58
66
86
159
233
290
340
Income
pop3133
pop10404
pop47098

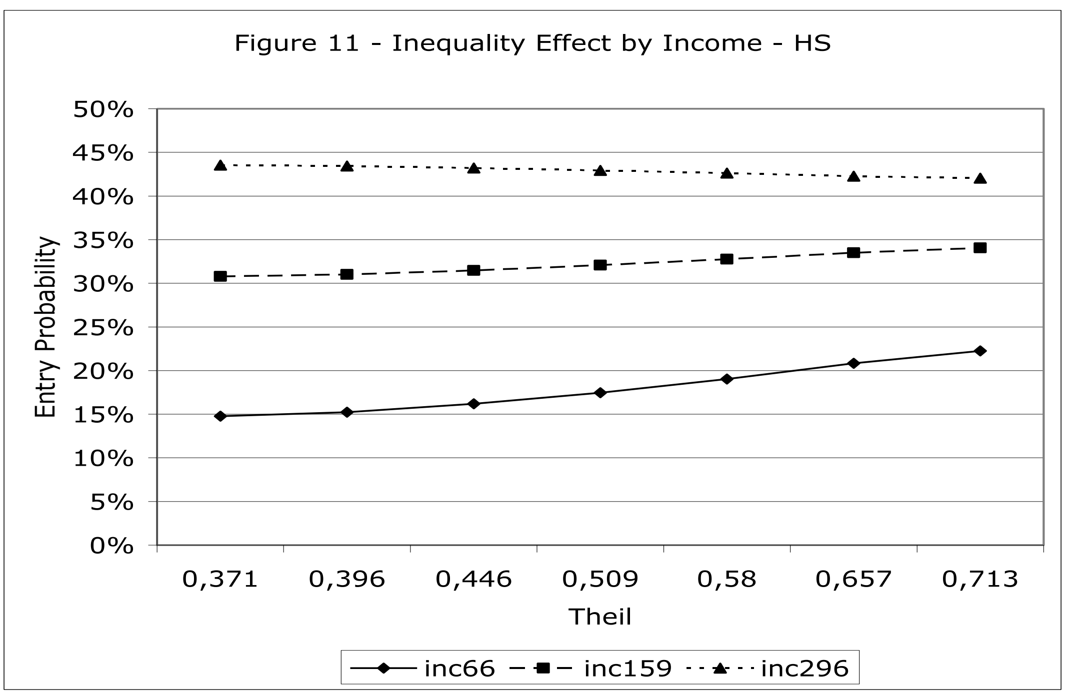


Figure 11 - Inequality Effect by Income - HS
50%
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
Entry Probability
0,371
0,396
0,446
0,509
0,58
0,657
0,713
Theil
inc66
inc159
inc296

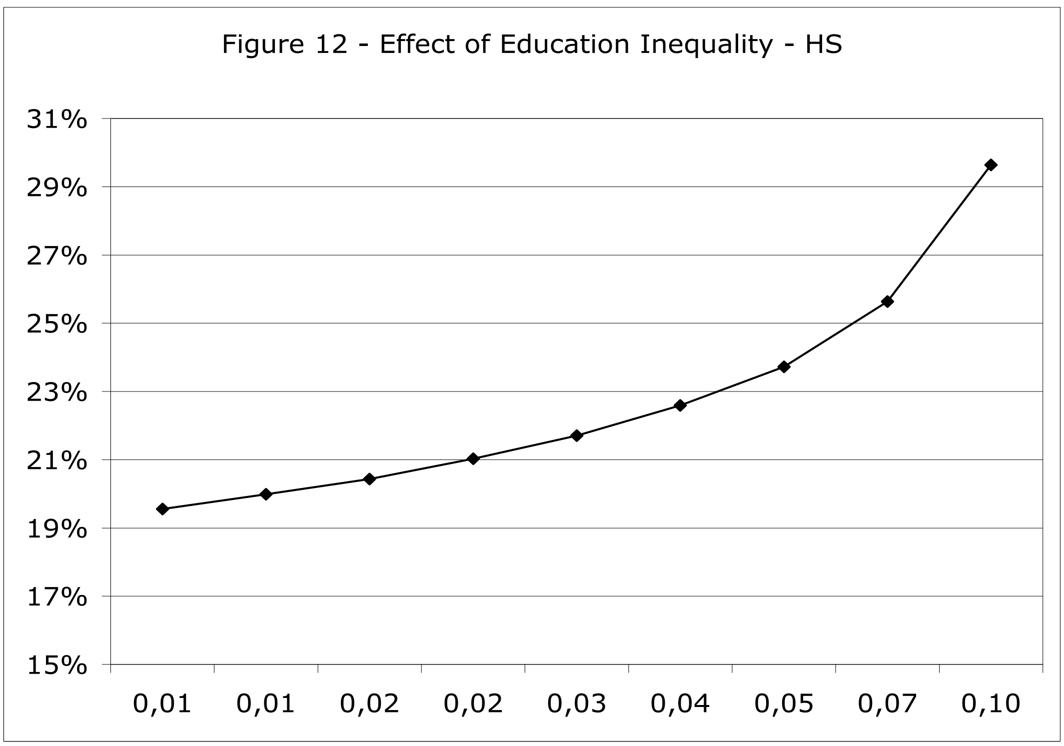


Figure 12 - Effect of Education Inequality - HS
31%
29%
27%
25%
23%
21%
19%
17%
15%
0,01
0,01
0,02
0,02
0,03
0,04
0,05
0,07
0,10