Education Choices in Competing Neighborhoods: Rational vs Coarse Expectations*

Simon GLEYZE† Philippe JEHIEL‡

Abstract

We develop a model in which students compete for seats at elite colleges and differ in ability and application cost. In disadvantaged communities, students estimate their admission chance based on the aggregate acceptance rate in the community. In advantaged communities, students have rational expectations. Competition between advantaged and disadvantaged (otherwise similar) communities results in more seats being allocated to the advantaged community. A different conclusion may arise when students of disadvantaged communities are less well informed about their characteristics but use rational expectations. The case of competition between disadvantaged communities is also discussed as well as the effect of quotas.

*We thank Roland Bénabou, Francis Bloch, Gabriel Carroll, Gabrielle Fack, Renato Gomes, Julien Grenet, Marc Gurgand, Yinghua He, Ronny Razin, Al Roth, Olivier Tercieux as well as the editor (Navin Kartik) and several reviewers for helpful comments. Jehiel thanks the European Research Council for funding (grant n25e6 742816).

†Uber Technologies, Inc. Email: gleyze.simon@gmail.com

‡Paris School of Economics and University College London. Email: jehiel@enpc.fr

1 Introduction

We are concerned in this paper with the competition of students for elite colleges. Students are assumed to differ in how much they would benefit from higher education (which is influenced by their ability type and possibly other factors) as well as how costly it is for them to apply (which may include the effort costs for admission tests in addition to the application fee). In particular, we consider the competition between students coming from different communities or neighborhoods, which may differ in their access to statistical knowledge about how the odds of admissions depend on objective characteristics of the applicant.

An empirical observation (see Hoxby and Turner (2015) and Goodman (2016)) is that high ability students from underrepresented communities seem to apply little to elite colleges, which seems paradoxical given that their admission odds are very high and such students would typically not have to incur large application costs (the financial part of these is heavily subsidized in many countries). Possible explanations that have been proposed in the literature include that such students may have less good outside options than their counterpart in more advantaged communities, thereby making them more conservative in their application strategy (Akbarpour et al. (2022) or Calsamiglia et al. (2021)) or that students from less advantaged communities may have less accurate view of their ability thereby leading them to be over-pessimistic about how much benefit they would get from higher education and/or about their odds of being accepted (see Hoxby and Turner (2015) and Ali and Shorrer (2025) for a theoretical exploration of this).

We propose a different explanation. In one of our main results, we consider two communities which are symmetric in all respects except that in one community (that should be thought as more advantaged), there is access to fine statistics that relate the observable characteristics of the student to the admission odds and in the other community (that should be thought of as less advantaged), only the coarse statistic describing the overall probability of admission among applicants in the community is available. When students in the disadvantaged community take the coarse statistic about the aggregate admission odds at face value as if it applied indiscriminately to all students in the community, we show that the disadvantaged community ends up having fewer seats at elite colleges than the advantaged community. In particular, this follows because in the disadvantaged community, some high ability students do not apply to elite college even though they would be admitted for sure.[1] This does not happen in the advantaged community.

It should be mentioned that in our model, students in the disadvantaged community do not have correct expectations about the admission odds because they do not correlate those with the ability type as they should. Several reasons may explain such erronenous expectations. First, disadvantaged students with high ability types (or high Grade Point Average GPA) may not be aware that admission odds are highly correlated with those (and believe instead that admission depends on other unrelated factors).[2] Second and relatedly, disadvantaged students may not have access to the precise data that would allow them to construct such ability-related statistics. In particular, they may sample from the past generation of students in their community those who applied and see among those who got admitted without observing precisely the ability characteristics of applicants. From this, they can construct the coarse statistic mentioned above but no more. To put it differently, the misconception we attribute to disadvantaged students takes the form of correlation neglect and while correlation neglect has been documented in a number of studies (see for example Enke and Zimmerman (2019)), we believe the form we consider here (that is consistent with key observations reported by Goodman in the context of education choices) has not been the subject of previous analysis, not to speak of how students with such misconceived expectations may interact with students holding rational expectations (that may result from access to better statistics in such more advantaged communities).[3]

[1] As a mirror image, we also have in the disadvanted community that some lower ability students apply even if their admission chances are null when their application costs are small enough.

[2] This has been documented by Goodman (2016), which has led her to famously advocate for mandatory admission college tests. (In this paper, we either assume that such tests are not mandatory, which is the case in many countries, or when referring to disadvanted students we focus on the residual students who may not take the result of the tests sufficiently into account.)

[3] Another approach concerned with education choices that considers correlation neglect

We also consider in our analysis an alternative scenario in which students from the disadvantaged community know less precisely how much they would benefit from higher education but have otherwise rational expectations, and we highlight a number of differences that this formulation would imply as compared to our leading (behavioral) approach. In particular, we do not necessarily have, in this alternative scenario, that the less informed community receives fewer seats, even assuming communities are otherwise symmetric, in contrast to the finding with the behavioral approach highlighted above. Of course, in reality, disadvantaged students may combine both inferior information and coarser expectations, but we believe our analysis helps identify a key role coarse expectations may have in explaining the education choices of disadvantaged students.

In the rest of the paper, we develop a stylized model of education choice in which (a continuum of) students differ in two characteristics, their benefit from attending higher education (which we decompose into an ability term and a match term that summarizes other aspects that could affect how much benefit the student could derive from being enrolled) and an application cost. In the main part, we assume students know these characteristics, but when considering the case of less informed students, we assume that they only observe their ability type and the application cost but not the match type.

We start by describing the analysis when there is only one community or neighborhood for which we characterize the equilibrium. When students base their application decision only on the aggregate admission rate, we refer to it as the sampling equilibrium for reasons related to the interpretation suggested above. Otherwise, when students have correct expectations, we refer to it as rational expectations equilibrium. We assess the obtained equilibria in light of the welfare they generate as well as the induced average (benefit-related) quality of admitted students. While the first-best would require that only students with maximal net benefit (i.e., benefit minus application cost) apply and none gets rejected, we note several forms of inefficiency in the

is Rees-Jones et al. (2024), even though in that paper the missperception bears about how rejection decisions at different schools are correlated (and correlation neglect leads to suboptimal hedging decisions in terms of portfolio of applications), which is very different from and complementary to our focus.

various equilibria. In the sampling equilibrium, high benefit students do not apply when the application cost is not low enough and some low benefit students apply and get rejected. In the Rational expectations equilibrium, when students know fully their benefit type, no applicant gets rejected in equilibrium but the outcome is inefficient because students do not sort according to their net benefit (but essentially according to their gross benefit). We approximate the welfare and the quality of admitted students so obtained as the mass of available seats gets small.

We next move to our main subject of interest, the study of competition for seats at elite colleges between several communities or neighborhoods. We start with the competition between asymmetrically knowledgeable neighborhoods. We first consider the case of two symmetric neighborhoods except that one relies on coarse expectations and one relies on rational expectations. We show that the neighborhood using coarse expectations receives fewer seats. We make the same analysis when neighborhoods use rational expectations but one neighborhood is less informed. We note that the less informed neighborhood need not end up with fewer seats in general, even though that is the case when the total mass of available seats is small enough. We also study the potential benefit of using quotas, which we identify here with assigning a mass of seats at elite colleges in proportion to the size of the community. When the total mass of seats is small, we show that quotas are beneficial in the case one community relies on coarse expectations and the other uses rational expectations. But it is neutral when all communities use rational expectations and they are asymmetrically informed.

We also explore competition between two less knowledgeable communities, this time allowing for asymmetric distributions of types across communities. In this part, we particularize the analysis to the case of a small mass of seats, and we study how seats are distributed as a function of the asymmetries across communities and whether quotas could be beneficial for total welfare and/or for improving the quality of admitted students. In the case neighborhoods rely on coarse statistics, we note a key role played by the mean ability in the neighborhood. More precisely, a neighborhood enjoying a higher mean ability ends up getting comparatively fewer seats. The mean ability plays no role in the case neighborhoods use rational expectations but have incomplete information. We also characterize when quotas can have a positive welfare effect and observe that the variance of the ability type has an effect in the case neighborhoods rely on the coarse statistics. By contrast, quotas are welfare neutral when neighborhoods have incomplete information on their benefit characteristic.

Related Literature At some broad level, our insights allow to shed light on empirical observations such as Hastings and Weinstein (2008) Hoxby and Turner (2015) or Kapor et al. (2020), which explicitly document various mistakes in expectation formation in the field of school choice. On the theory side, our paper relates to several approaches in behavioral game theory (see in particular Eyster and Rabin (2005), Jehiel (2005)) to the extent that sampling students can be interpreted as missing the correlation between acceptance rate and ability. The approach closest to ours in the context of education is Manski (1993) who postulates an additive log-income equation, and assumes that students infer the returns to schooling by taking the conditional expectation of log-income. When students omit to condition on ability—e.g., because they do not observe the ability of their peers—he shows that more low-ability and less high-ability students enroll in college. Our model is different in that in our environment there is a strategic interaction across students due to the scarcity of seats in elite colleges, and students' expectations concern their admission chances which are affected by the application strategy of other students, but our analysis of the sampling equilibrium in the one-neighborhood case shares some similarity with Manski (1993) to the extent some high ability students refrain from applying and some low ability students apply and get rejected. However, our main contribution lies in the multi-neighborhood version of our model and the interaction of neighborhoods either using rational expectations or the sampling heuristic, which has no counterpart in Manski (1993).[4]

[4]Another reference that studies biased belief formation in education choices is Streufert (2000). While the mechanism he studies and the induced bias are different in the two papers (in Streufert, it is related to the change of composition of the neighborhood after the education choice), another key difference is the equilibrium aspect of our approach, which makes our approach closer to Manski than Streufert.

2 SETUP

We introduce a stylized model of career choice with strategic students and rationing at elite colleges. In this Section, we consider a single neighborhood with a mass of students normalized to 1 and a mass $q < 1$ of seats at elite colleges. Students are parameterized by θ that represents how much the student would benefit (in expectation) from being enrolled at an elite college and c that represents the application cost. The benefit can be decomposed into two terms $\theta = \phi + \omega$ where ϕ is an ability score (say, GPA) of the student and ω is an extra match parameter (as reflected in teacher recommendations, college essays, extra curricular activities, etc.). In all scenarios analyzed below, we assume that the student knows his ability score ϕ and his application cost c. In a large part of the paper, we also assume that he observes his match parameter ω, but in some parts we also discuss the case in which ω is not observed by the student.

For expositional simplicity, we will assume that ϕ, ω and c are independently distributed with smooth densities f, h and g on $[\underline{\phi}, \overline{\phi}]$, $[\underline{\omega}, \overline{\omega}]$ and $[0, \overline{c}]$, respectively. We will denote by F, H and G the corresponding cumulative, and we will assume that $f(\overline{\phi})$, $h(\overline{\omega})$ and $g(0)$ are all strictly positive (this will be used in our approximations when we consider the case of a small mass q of seats). When the student fully observes θ (i.e. both ϕ and ω), we will refer to $k(\cdot)$ as the density of θ with support $[\underline{\theta}, \overline{\theta}] = [\underline{\phi} + \underline{\omega}, \overline{\phi} + \overline{\omega}]$ where obviously $k(\cdot)$ can be derived from $f(\cdot)$ and $h(\cdot)$.

Students choose among two occupations: going to a non-selective college (or directly on the labor market) L, or applying to a selective college C. That is, their action set is $A = \{L, C\}$. For a student with benefit type θ, the gross utility of attending an elite college is $U^C(\theta) = \theta$, whereas we assume that the utility of going to a non-selective college is $U^L(\theta) = 0$.[5] The net utility of a student with type (θ, c) is thus given by:

- 0 if the student goes to a non-elite college L.
- $\theta - c$ if the student applies to C and obtains a seat.

[5]One can interpret this as saying that θ is the extra net benefit of attending an elite college compared to a non-elite college and that if the student gets rejected from the elite colleges he can eventually enjoy the benefit corresponding to the non-elite colleges.

- $-c$ if the student applies to C but does not get a seat.

The description of the model is completed as follows. First, students decide whether or not to apply to C as a function of what they observe. Then elite colleges after observing the benefit types θ of the applicants fill in their q seats by accepting the applicants with highest θ.[6]

A crucial aspect is how students form their expectation about their admission chance. In one approach, we assume students are not able to fine tune their expectation to their type, leading them to reason as if the acceptance rate was the same irrespective of the type. In equilibrium, we require this expected acceptance rate to match the (equilibrium) aggregate acceptance rate among applicants. We refer to the resulting outcome as a local sampling equilibrium, as it can be thought of as a steady state in which new generations of students would be informed (through local sampling) about acceptance rate of past generations without further details. We provide several interpretations of the local sampling equilibrium later on after the formal definition has been introduced.

Our second main approach is the traditional rational expectations setup for which we will distinguish between the cases in which the student only observes ϕ or observes both ϕ and ω in addition to c.

2.1 Local sampling equilibrium

We present our description of the local sampling equilibrium assuming students observe fully (θ, c).

A strategy profile $\sigma : N \longrightarrow \Delta A$ is a (measurable) function mapping the set of types into application probabilities. We let $\sigma(\theta, c) \in [0, 1]$ denote the probability that student (θ, c) applies to C.

For any strategy profile, let $\theta(\sigma)$ denote the cutoff such that any student with ability $\theta \geq \theta(\sigma)$ who applies to C is admitted. It is defined as follows:

[6] Colleges are assumed to fuly observe θ (i.e. ϕ and ω). The extension of our analysis to the case in which colleges screen applicants based on imperfect signals about applicants' type θ is left for future research.

$\theta(\sigma) = \underline{\theta}$ when

$$\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{c}}^{\overline{c}} \sigma(\theta, c)k(\theta)g(c) \, \mathrm{d}c \, \mathrm{d}\theta < q$$

Otherwise, $\theta(\sigma)$ is uniquely defined as the largest θ^* such that

$$\int_{\theta^*}^{\overline{\theta}} \int_{\underline{c}}^{\overline{c}} \sigma(\theta, c)k(\theta)g(c) \, \mathrm{d}c \, \mathrm{d}\theta = q$$

A key object that drives the choice of student (θ, c) is the subjective probability this student assigns to obtaining a seat at an elite college conditional on applying to C. In the sampling equilibrium, this probability p is constant. Based on p, student (θ, c) applies to C whenever:[7]

$$p\theta - c \geq 0$$

Moreover, in a sampling equilibrium, this p should match the aggregate acceptance rate of applicants given the strategy σ. This leads to the following definitions.

DEFINITION 1. *σ is optimal given p if*

$$\sigma(\theta, c) = \begin{cases} 1 & \textit{when} \quad c \leq p\theta \\ 0 & \textit{when} \quad c > p\theta \end{cases}$$

DEFINITION 2. *σ^S is a local sampling equilibrium if σ^S is optimal given p^S where*

$$p^S = q \Big/ \int_{\underline{\theta}}^{\overline{\theta}} \int_{0}^{\overline{c}} \sigma^S(\theta, c)k(\theta)g(c) \, \mathrm{d}c \, \mathrm{d}\theta.$$

If the acceptance rate were believed to be 1, every student (θ, c) such that $c < \theta$ would apply, which would define a sampling equilibrium whenever $\int_{\underline{\theta}}^{\overline{\theta}} k(\theta)G(\theta) \, \mathrm{d}\theta \geq q$. However, when $\int_{\underline{\theta}}^{\overline{\theta}} k(\theta)G(\theta) \, \mathrm{d}\theta < q$, some rationing must occur. A sampling equilibrium is shown to exist by using the intermediate

[7]For completeness, we assume that the student applies to C when indifferent, but how indifferences are resolved plays no role in the analysis.

value theorem applied to the the composition of the continuous functions that first maps the acceptance rate p to the corresponding optimal strategy and then computes the resulting aggregate acceptance rate from the formula $q \Big/ \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{\overline{c}} \sigma^S(\theta, c) k(\theta) g(c) \, \mathrm{d}c \, \mathrm{d}\theta$. To see this more concretely, note that if the acceptance rate were very small, only those with very small c would apply, they would all be accepted, thereby leading to an effective acceptance rate equal to 1, and as just seen when the acceptance rate is close to 1, there would be rationing resulting in an effective acceptance rate smaller than 1.

Observe that in a sampling equilibrium students observe their type (θ, c) but their expectation about the admission chance is taken to be the aggregate admission rate of applicants independently of the benefit type θ. This expectation is not correct as the true admission outcome depends on θ. It is also different from the rational expectations equilibrium that we consider later in a scenario in which students would be partially informed of θ. In such an alternative scenario, the application strategy could not depend finely on θ (as the benefit type would not be fully observed).

Several interpretations of the sampling equilibrium can be proposed. First, students while knowking their ability type (GPA) may wrongly believe that admissions are predominantly based on unrelated factors.[8] The resulting steady state would be our above sampling equilibrium viewed as a Berk-Nash equilibrium (Esponda-Pouzo, 2016) in such a setting with misspecified prior. Second, and maybe giving extra support to such a form of misspecification, the sampling equilibrium can be viewed as a steady state of a multi-generation education choice setup in which students from a given generation in addition to observing their type (θ, c)would also have access to the admission outcomes of applicants in the previous generation, but would not have access to the ability characteristics of those past applicants. Such a feedback would allow the students from the current generation to form an estimate of the aggregate acceptance rate, but not of how the acceptance rate depends on the benefit characteristic.[9] The sampling equilibrium then assumes that this

[8]This is in agreement with the narrative in Goodman (2016) according to which students with high GPA seem unaware they have high chances of being accepted to elite colleges (when they are not subject to a mandatory admission test).

[9]To simplify matters, we have in mind that the samples are representative so that there is no variation across students and no estimation error about the aggregate acceptance rate.

aggregate statistics is used as a proxy to estimate the admission chance irrespective of the benefit type and that a steady state has been reached. This is similar to the approach developed in Manski (1993) when modeling how students would form their estimate of the returns to schooling when the ability of others is not observed.

From a related sampling perspective, one can alternatively assume that students have access to the benefit characteristics of past applicants (at least partially). With this view, the sampling equilibrium can be interpreted as modeling a form of correlation neglect in which students would miss the correlation between admission chance and the benefit characteristic and treat the two variables as if they were independently distributed. While correlation neglect has been documented and studied in the literature (see for example Enke and Zimmermann (2019) in an abstract experimental setting and Rees-Jones and Shorrer (2023) for a study of correlation neglect in the context of multiple school choice applications), the kind of correlation neglect considered here has not been studied in the literature to the best of our knowledge.[10]

We believe the sampling equilibrium is particularly appealing in contexts in which students would have to rely on their own experience to form their expectations, since in this case it sounds plausible students would not be able to understand finely how the acceptance probability depends on the benefit type (and students could more naturally maintain their misconceived perception on how admissions at elite colleges relate to ability). As such, we believe the sampling equilibrium is likely to capture how students from disadvantaged neighborhoods make their education choices whereas students from more advantaged neighborhoods would likely have access to more precise statistics, which we will model by endowing the corresponding students

[10] In the working paper we discuss a variant with less extreme form of correlation neglect. More precisely, we assume that to form an estimate of the admission chance a student with benefit type θ considers only those applicants with an ability characteristic in a window around θ where the window is adjusted so that the mass of data considered by the student reaches a pre-defined threshold τ. The sampling equilibrium corresponds to a scenario with large τ (so that the entire neighborhood enters the sample) while the rational expectations equilibrium is approached as τ converges to 0. While appealing in a number of respects (such a modeling can be viewed as formalizing in a simple way the bias-variance trade-off), this more flexible formulation is less tractable than the one adopted here.

with rational expectations.

Getting to the economics of a local sampling equilibrium, we observe that two types of mistakes can arise: On the one hand, we see high benefit students $\theta > \theta(\sigma)$ not applying when $c > p^S\theta$ whereas they would have been admitted had they applied. On the other hand, we see low benefit students $\theta < \theta(\sigma)$ applying when $c < p^S\theta$ whereas their application is a pure waste given that they are rejected.

2.2 Rational expectations equilibrium

When students observe fully (θ, c), rational expectations require that a (θ, c)-student expects that if he applies he will be accepted with probability $p^R(\theta)$ where

$$p^R(\theta) = \begin{cases} 1 & \text{when} \quad \theta \geq \theta(\sigma) \\ 0 & \text{when} \quad \theta < \theta(\sigma) \end{cases}$$

and $\theta(\sigma)$ is as before the benefit threshold above which applicants get accepted when students follow strategy σ.

Thus, in a Rational expectations equilibrium, students with benefit $\theta \geq \theta^R$ apply whenever $\theta > c$ and those with $\theta < \theta^R$ do not apply where θ^R is charactcterized by $\theta^R = \underline{\theta}$ if

$$\int_{\underline{\theta}}^{\overline{\theta}} k(\theta)G(\theta)\,\mathrm{d}\theta < q.$$

And otherwise, θ^R is uniquely characterized by

$$\int_{\theta^R}^{\overline{\theta}} k(\theta)G(\theta)\,\mathrm{d}\theta = q.$$

In a rational expectations equilibrium, every applicant is accepted. Yet, there are inefficiencies because it would have been more efficient to have only those students with maximal $\theta - c$ applying whereas in a rational expectations equilibrium, a student with $\theta > \theta^R$ applies (and gets accepted) as long as $c < \theta^R$ (inducing, for example, a low net benefit when c is close to θ).

We also consider the scenario in which a student would be less informed of θ and only observe ϕ together with the application cost c. In this case, let $\theta^C \geq \underline{\theta}$ denote the benefit threshold such that those applicants with $\theta \geq \theta^C$ get accepted. When knowing ϕ, the student expects that by applying he will be accepted and get $\phi + \omega$ whenever $\omega \geq \theta^C - \phi$, which in turn leads him to apply whenever

$$c \leq \int_{\omega > \theta^C - \phi} (\phi + \omega) l(\omega) \, \mathrm{d}\omega.$$

Integrating over ϕ to equate demand and supply yields that $\theta^C = \underline{\theta}$ whenever

$$\int_{\phi} G \left(\int_{\omega > \theta^C - \phi} (\phi + \omega) l(\omega) \, \mathrm{d}\omega \right) f(\phi) \, \mathrm{d}\phi < q.$$

And otherwise $\theta^C > \underline{\theta}$ is uniquely determined by

$$\int_{\phi} G \left(\int_{\omega > \theta^C - \phi} (\phi + \omega) l(\omega) \, \mathrm{d}\omega \right) \left(1 - H(\theta^C - \phi)\right) f(\phi) \, \mathrm{d}\phi = q.$$

Remark. We could have considered a scenario in which students do not observe ω and as in the sampling equilibrium rely on the aggregate acceptance rate to make their education choices (rather than rational expectations). We suspect that similar insights as those shown for the sampling equilibrium with perfect observation of θ would arise there, hence our modeling choice to simplify notation.[11]

To quantify inefficiencies and measure the quality of admitted students in the various scenarios considered below, we introduce two aggregate measures for arbitrary application strategy $\sigma(\cdot)$.

[11] It should also be mentioned that in such an approach, students should also form expectations about the ω they will get if admitted. We could assume that they consider the mean value of ω for this purpose (even if that would require that students know this mean value) or else more in line with the spirit of our approach, we could require that students have access to the mean value of ω among admitted students in their neighborhood and that they take this empirical mean as a proxy for the average ω they will get if admitted (this would result in an extra selection bias effect of the sort studied in Jehiel (2018)). An exact analysis of this is left for future research.

First, we define the aggregate welfare as

$$W(\sigma) = \int_{\theta(\sigma)}^{\overline{\theta}} \int_0^{\overline{c}} \theta\sigma(\theta, c)k(\theta)g(c)\,\mathrm{d}c\,\mathrm{d}\theta - \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{c} c\sigma(\theta, c)k(\theta)g(c)\,\mathrm{d}c\,\mathrm{d}\theta.$$

The first term describes the aggregate benefit of admitted students. The second term reflects the aggregate application costs of applicants.

Second, we define the average quality of admitted students as

$$M = \frac{\int_{\theta(\sigma)}^{\overline{\theta}} \int_0^{\overline{c}} \theta\sigma(\theta, c)k(\theta)g(c)\,\mathrm{d}c\,\mathrm{d}\theta}{\int_{\theta(\sigma)}^{\overline{\theta}} \int_0^{\overline{c}} \sigma(\theta, c)k(\theta)g(c)\,\mathrm{d}c\,\mathrm{d}\theta}.$$

2.3 Small Number of Seats

In a number of our formal results, we consider the case in which the mass q of seats is small, and, to simplify notation, we normalize $\overline{\theta} = 1$. In this case, we are able to provide closed form approximations to the sampling equilibrium, the rational expectations equilibrium as well as the associated welfare and average quality of admitted students.

We start with the sampling equilibrium. Let $p(q)$ denote the aggregate acceptance rate when the mass of seats is q. Given that a (θ, c)-student applies if $c < p\theta$, this acceptance rate $p(q) = p$ is characterized by

$$p = q \Big/ \int_{\underline{\theta}}^{1} k(\theta)G\left(p\theta\right)\,\mathrm{d}\theta.$$

In the next Proposition, we establish that the sampling equilibrium is unique when q is small and we provide approximations to the admission benefit cutoff type $\theta^S(q)$ above which students when applying are accepted, the admission rate $p(q)$ as well as the welfare $W(q)$ and the average ability of admitted students $M(q)$ in terms of the mass q of seats for arbitrary densities f, h and g. All proofs appear in Appendix.

Proposition 1. *When q is small, there is a unique sampling equilibrium. More-*

over

$$
\begin{aligned}
p(q) &= \left(\frac{q}{g(0)E(\theta)}\right)^{1/2} + o(q^{1/2}) \\
\theta^S(q) &= 1 - \left(\frac{1}{f(\overline{\phi})h(\overline{\omega})}\right)^{1/2}\left(\frac{2E(\theta)q}{g(0)}\right)^{1/4} + o(q^{1/4}) \\
W(q) &= \left(1 - \frac{E(\theta^2)}{2E(\theta)}\right) q + o(q) \\
M(q) &= 1 - \frac{2}{3}\left(\frac{1}{f(\overline{\phi})h(\overline{\omega})}\right)^{1/2}\left(\frac{2E(\theta)q}{g(0)}\right)^{1/4} + o(q^{1/4})
\end{aligned}
$$

To give some roadmap to understand Proposition 1, note first, that $p(q)$ must be small as q gets small, as otherwise, a significant mass of students would apply (all those (θ, c) with $c < p\theta$) leading to a small acceptance rate when q is small. Second, it is readily verified that $\theta^S(q)$ must be close to 1 as otherwise $p = \int_{\theta^S}^1 h(\theta)G(p\theta)\, \mathrm{d}\theta / \int_{\underline{\theta}}^1 h(\theta)G(p\theta)\, \mathrm{d}\theta$ could not be close to 0. We next use this identity to approximate p in terms of $(1-\theta^S)$ (approximating $k(\theta)$ by $\sqrt{2}f(\overline{\phi})h(\overline{\omega})(1-\theta)$ when θ is close to 1). The rest of the derivations follow similar approximations. Important features of these approximations include the observations that $\theta^S(q)$ gets away from 1 according to a term of the order of $q^{1/4}$ and that there is a role for $E(\theta)$ and $E(\theta^2)$ in these approximations. That $E(\theta)$ plays a role can be seen on the technical side when we approximate $\int_{\underline{\theta}}^1 k(\theta)G(p\theta)\, \mathrm{d}\theta$ by $\int_{\underline{\theta}}^1 pk(\theta)\theta g(0)\, \mathrm{d}\theta = pg(0)E(\theta)$ when p is small. At a more economic level, in a sampling equilibrium, even small θ students apply when their c is small enough, and this is what makes the distribution of θ outside at the very top relevant for the equilibrium characterization.

Consider now the rational expectations equilibrium approach, and let $\theta^R(q)$ and $\theta^C(q)$ denote the corresponding admission thresholds when there are q seats and students observe θ or when they do not observe ω, respectively. We have:

PROPOSITION 2. *When q is small and students observe θ, we have in the unique rational expectations equilibrium that $\theta^R(q) = 1 - \left(\frac{2}{f(\overline{\phi})h(\overline{\omega})G(1)}q\right)^{1/2} + o(q^{1/2})$. When students do not observe ω, the rational expectations equilibrium is unique and* *satisfies:*

$$\begin{aligned}
\theta^C(q) &= 1 - \left(\frac{3}{f(\overline{\phi})h(\overline{\omega})^2 g(0)}q\right)^{1/3} + o(q^{1/3}) \\
W(q) &= \frac{q}{2} + o(q) \\
M(q) &= 1 - \frac{2}{3}\left(\frac{3}{f(\overline{\phi})h(\overline{\omega})^2 g(0)}q\right)^{1/3} + o(q^{1/3})
\end{aligned}$$

Note that in the REE when ω is not observed, no student with $\phi < \theta^C(q) - \overline{\omega}$ applies, since such students understand that they have no chance of being accepted when applying. This is in contrast with the sampling equilibrium in which a student with type (ϕ, ω, c) applies no matter how small ϕ is as long as $c < p(q)\theta$. With a symmetric argument, the same logic implies that a student with a sufficiently high benefit type applies for a much wider range of c in the REE than in the sampling equilibrium.[12] This in turn explains why for a given mass of seats q, the acceptance threshold is higher in the REE than in the sampling equilibrium since the mass of high benefit types is higher among applicants in the REE than in the sampling equilibrium.[13]

We note also that the welfare in the REE when students do not observe ω is (approximately) $\frac{q}{2}$ irrespective of the distribution of ϕ and ω in contrast with the finding in the sampling equilibrium (in which $E(\theta)$ and $E(\theta^2)$ were shown to play a role). The welfare approximation in the REE can be understood simply. For any $\phi > \theta^R(q) - \overline{\omega}$, the expected chance of being accepted is (approximately) $h(\overline{\omega})(\overline{\phi} - \phi)$ and thus a student with ϕ characteristic applies when $c < h(\overline{\omega})(\overline{\phi} - \phi)$. On average, this results in a net benefit (benefit diminished by the application cost) equal to $\frac{1}{2}h(\overline{\omega})(\overline{\phi} - \phi)$ (because c spans almost uniformly values from 0 to $h(\overline{\omega})(\overline{\phi} - \phi)$ when $g(0) > 0$). The approximation of $W(q)$ follows. Since $E(\theta^2) < E(\theta)$, we have that the welfare is bigger in the sampling equilibrium than in the REE when q is small. This will be used

[12] In the case in which the student would observe θ in the REE, any student with type $\theta > \theta^R$ applies as long as $c < \theta^R$ since such a student would be sure to be accepted when applying.

[13] It departs from 1 with a distance of the order of $q^{1/4}$ in the sampling equilibrium and of the order of $q^{1/2}$ or $q^{1/3}$ in the REE depending on whether the student observes ω.

below when analyzing the welfare implications of quotas.[14]

3 Competing Neighborhoods

Our main object of interest lies in the study of multiple neighborhoods competing for the same positions. In a given neighborhood, we assume that students share the same level of knowledge/understanding of the functioning of the education system. They either form their application strategy based on the aggregate acceptance rate observed locally in their neighborhood. This is similar to the approach developed above for the local sampling equilibrium where it is the acceptance rate in the neighborhood the student belongs to that is used to form the student's expectation.[15] Or alternatively, students have rational expectations. In the latter case, we consider both the scenario in which students know fully their benefit characteristic θ and the scenario in which only ϕ is osberved about θ.

When a neighborhood either relies on the sampling heuristic or ω is not observed by students in this neighborhood (but rational expectations are used instead), we refer to the neighborhood as a less knowledgeable one. We will study the competition between asymmetrically knowledgeable neighborhoods, illustrating the different implications the two forms of imperfect knowledgeability have on the analysis. We will next study the competition between two imperfectly knowledgable neighborhoods. In each case, we will first provide a postitive analysis in the absence of any intervention. We will next discuss the effect of having quotas taking the form of allocating seats to neighborhoods in proportion of their size. In some of our results, we consider the case of a small mass of seats making use of the approximations derived at the end of Section 2.

[14]We have not provided the welfare approximation in the REE case when students observe θ. This welfare would depend on the entire distribution of c. However, given that the welfare would be approximately $q/2$ whenever θ is not perfectly observed, this is the welfare approximation we consider below when discussing the competition between a neighborhood relying on the sampling heuristic and a neighborhood relying on rational expectations. This could be approached formally by decomposing $\omega = \omega_1 + \cdots + \omega_n$ and requiring that not all componnents of ω are known by the student.

[15]Following the above sampling interpretation, we are assuming here that the sampling is restricted to students from the same neighborhood (which can also be viewed as how we define neighborhood in our context).

To formalize the questions of interest, we consider a two-neighborhood setup. Neighborhood $i = 1, 2$ consists of a mass m_i of students with benefit type $\theta = \phi + \omega$ and application cost type c defined as above. We let $f_i(\phi)$, $h_i(\omega)$ and $g_i(c)$ denote the mass-normalized distribution of (ϕ, ω, c). That is, (ϕ, ω, c) is distributed according to density $m_i f_i(\phi) h_i(\omega) g_i(c)$ in neighborhood i.

In neighborhood i, we assume that either students observe θ and rely on the (local) aggregate admission rate among applicants in that neighborhood. We then refer to such a neighborhood as having type $\tau_i = S$. Or else, expectations are rational and students either observe fully θ in which case we refer to the type of such a neighborhood as $\tau_i = FR$ or else students only observe ϕ in which case we refer to the type of such a neighborhood as $\tau_i = CR$. When $\tau_i = FR$, we say that neighborhood i is fully knowledgeable while when $\tau_i = S$ or CR we say that neighborhood i is imperfectly knowledgeable.

When the number of seats in neighborhood i is q_i, we refer to $\sigma^{\tau_i}(\cdot \mid q_i)$ and $\theta(\sigma^{\tau_i}, q_i)$ as an τ_i-equilibrium strategy and the corresponding benefit admission threshold in this neighborhood, which is determined by the sampling equilibrium conditions if $\tau_i = S$ or the Rational expectations equilibrium conditions if $\tau_i = FR$ or CR as defined in Section 2.[16]

An equilibrium with competing neighborhoods is formally defined as follows.

DEFINITION 3. *An equilibrium with competing neighborhoods $i = 1, 2$ with type τ_i and mass q of seats is a strategy profile (σ_1, σ_2) such that there exist q_1 and q_2 satisfying*

1. $\sigma_i = \sigma^{\tau_i}(\cdot \mid q_i)$ *for some τ^i-equilibrium $\sigma^{\tau_i}(\cdot \mid q_i)$ in neighborhood i with a mass q_i of seats;*
2. $q_1 + q_2 = q$ *and,*
3. $\theta^{\tau_1}(\sigma^{\tau_1}, q_1) = \theta^{\tau_2}(\sigma^{\tau_2}, q_2)$.

[16]Given that the mass of students was assumed to be 1 in Section 2, we need to use q_i/m as the size-normalized mass of seats in the expressions shown in Section 2 to describe σ^{τ_i} and θ^{τ_i}.

In equilibrium, given the mass q_i of seats assigned to neighborhood i, a corresponding τ_i-equilibrium must be played in neighborhood i (condition 1), the total mass of seats should add up to the total mass of available seats (condition 2), and the admission threshold should be the same in the two neighborhoods (condition 3). It is this third condition that results from the selection process being merged between the two neighborhoods that creates a link between the two neighborhoods. The rest of this Section is devoted to analyzing the implication of this link as well as whether some public interventions in the form of quotas could be beneficial in particular for welfare purposes.

3.1 Asymmetrically knowledgeable neighborhoods

We start our investigation by considering one neighborhood with type $\tau = S$ that relies on the sampling heuristics for expectation purposes and one neighborhood with type $\tau = FR$ being fully knowledgeable. We allow the two neighborhoods to have different sizes (as size may be thought of as being an important determinant whether the neighborhood has biased or rational expectations),[17] but we freeze other sources of asymmetry. We have:

PROPOSITION 3. *Assume neighborhood 1 with size m_1 relies on the local sampling heuristic (i.e., $\tau_1 = S$) and neighborhood 2 with size m_2 is fully knowledgeable (i.e., $\tau_2 = FR$). Assume the distributions k_i and g_i of θ_i and c_i are the same across the two neighborhoods (these are denoted k and g, respectively with G being the cdf of g). In equilibrium, neighborhood 1 receives fewer seats per heads than neighborhood 2. That is, $\frac{q_1}{m_1} \leq \frac{q_2}{m_2}$ with a strict inequality as soon as there is some scarcity of seats which arises when $q < (m_1 + m_2)\int_{\underline{\theta}}^{\overline{\theta}} G(\theta)k(\theta)\,\mathrm{d}\theta$.*

Proof. Equilibrium conditions write (letting $\theta^* \geq \underline{\theta}$ be the common benefit threshold for admission and p_1 the aggregate acceptance rate in neighbor-

[17] Think of a neighborhood as being disadvantaged because it is under-reperesented.

hood 1):

$$\begin{aligned} m_1 \int_{\theta^*}^{1} G(p_1\theta)k(\theta)\,\mathrm{d}\theta &= q_1 \\ m_2 \int_{\theta^*}^{1} G(\theta)k(\theta)\,\mathrm{d}\theta &= q_2 \end{aligned}$$

where $q_1 + q_2 = q$ and

$$m_1 p_1 \int_{\underline{\theta}}^{1} G(p_S\theta)k(\theta)\,\mathrm{d}\theta = q_1.$$

For any θ^* and any $p_1 < 1$ (some applicants are rejected in S), the above conditions trivially imply that $q_2/m_2 > q_1/m_1$. The rest of the proof is easily completed. **Q. E. D.**

The intuition for the result is very simple. Only students with benefit type θ above the common acceptance threshold θ^* get accepted. In the fully knowledgeable neighborhood 2, this implies that any student with $\theta > \theta^*$ applies whenever $c < \theta$. By contrast (when $\theta^* > \underline{\theta}$), in neighborhood 1, some students with $\theta < \theta^*$ apply (whenever $c < p_1\theta$) and get rejected. The acceptance rate p_1 is thus strictly smaller than 1 (this requires that $q < (m_1 + m_2)\int_{\underline{\theta}}^{\overline{\theta}} G(\theta)k(\theta)\,\mathrm{d}\theta$). Moreover, in neighborhood 1, students with $\theta > \theta^*$ only apply when $c < p_1\theta$, which happens less frequently than in the fully knowledgeable neighborhood due to the p_1 discounting. The comparative assignment of seats follows.

Going beyond the symmetry assumptions made in Proposition 3 would make the comparison between $\frac{q_1}{m_1}$ and $\frac{q_2}{m_2}$ hard to derive in general. However, in the limit in which the total mass q of seats is small, no symmetry is needed to show that the fully knowledgeable neighborhood receives almost all seats. As it turns out, this conclusion holds true even if in neighborhood 2 no observation of ω is made by students, as long as students rely on rational expectations. Formally,

Proposition 4. *Assume that $\tau_1 = S$ and $\tau_2 = FR$ or CR. In the limit as q goes to 0, neighborhood 2 gets almost all seats as compared with neighborhood 1 in equilibrium. That is, $\lim_{q\to 0}\frac{q_1}{q_2} = 0$.*

This result derives from the approximations obtained at the end of Section 2 when the mass of seats is assumed to be small. In neighborhood 1 (with $\tau_1 = S$), the admission threshold deviates from 1 in a term proportional to $(q_1)^{1/4}$ where q_1 is the mass of seats obtained in this neighborhood. By contrast, in neighborhood 2 the admission threshold deviates from 1 in a term proportional to $(q_2)^{1/2}$ when $\tau_2 = FR$ and proportional to $(q_2)^{1/3}$ when $\tau_2 = CR$ where q_2 is the mass of seats obtained in neighborhood 2. In equilibrium, the admission threshold must be the same in both neighborhoods. Given that $x^{1/4}$ is much bigger than $x^{1/2}$ or $x^{1/3}$ when x is small, the only way to achieve this is to have that q_1 is much smaller than q_2, hence the claim in Proposition 4.

Consider next how seats are assigned between neighborhoods when they both rely on rational expectations but one of them has less information. That is, one neighborhood say neighborhood 1 has type $\tau_1 = CR$ and the other neighborhood 2 has type $\tau_2 = FR$. We note that in this case it is not clear that the less informed neighborhood necessarily ends up having fewer seats per capita than the fully informed neighborhood, even assuming that the two neighborhoods are otherwise symmetric as in Proposition 3.

To see this, let f, h and h denote the common distribution of ϕ, ω and c in the two neighborhoods. Let θ^* be the (common) admission threshold and q_1, q_2 the equilibrium mass of seats obtained by neighborhoods 1 and 2, respectively. We have:

$$\int_{\underline{\phi}}^{\overline{\phi}} \left(\int_{\theta^* - \phi}^{\overline{\omega}} G(\phi + \omega) h(\omega) \, \mathrm{d}\omega \right) f(\phi) \, \mathrm{d}\phi = q_2/m_2$$

since in neighborhood 2, student with ability type ϕ applies whenever $\omega > \theta^* - \phi$ and $c < \phi + \omega$ (hence the $G(\theta + \omega)$ term).

We also have

$$\int_{\phi} G \left(\int_{\omega > \theta^* - \phi} (\phi + \omega) h(\omega) \, \mathrm{d}\omega \right) (1 - H(\theta^* - \phi)) \, f(\phi) \, \mathrm{d}\phi = q_1/m_1$$

since in neighborhood 1 a student with ability type ϕ anticipates that by participating she will get an expected benefit $\int_{\omega>\theta^*-\phi}(\phi+\omega)h(\omega)\ \mathrm{d}\omega$ and thus participates only when $c < \int_{\omega>\theta^*-\phi}(\phi+\omega)h(\omega)\ \mathrm{d}\omega$ (hence the term $G\left(\int_{\omega>\theta^*-\phi}(\theta+\omega)l(\omega)\ \mathrm{d}\omega\right)$) and that she then gets a seat (she is selected) only when $\omega > \theta^* - \phi$ (hence the term $(1 - H(\theta^* - \phi))$).

In the special case, in which there would be no scarcity of seats (that is, $\theta^* = \underline{\theta}$ implying that $H(\theta^* - \phi) = 0$ for all ϕ), the second condition simplifies into:

$$\int_{\phi} G\left(\int_{\omega>\theta^*-\phi}(\phi+\omega)h(\omega)\ \mathrm{d}\omega\right) f(\phi)\ \mathrm{d}\phi = q_1/m_1.$$

And when G is concave, Jensen's inequality yields that $q_1/m_1 > q_2/m_2$, as promised.

So in the absence of scarcity of seats we would have that the less informed neighborhood has more seats per capita than the fully informed neighborhood when the cumulative distribution of the application cost is concave. Allowing for some scarcity of seats still leads to the same comparison between q_1/m_1 and q_2/m_2 (by a continuity argument) in some cases, thereby illustrating the sharp contrast with the insight derived in Proposition 3 when in neighborhood 1 students were assumed to observe their benefit type fully but were instead assumed to rely on the local aggregate admission rate to decide on their application strategy.

Moving to the case of a small number of seats, we can however establish:

PROPOSITION 5. *Assume that $\tau_1 = CR$ and $\tau_2 = FR$. In the limit as q goes to 0, neighborhood 2 gets almost all seats as compared with neighborhood 1 in equilibrium. That is, $\lim_{q\to 0}\frac{q_1}{q_2} = 0$.*

This result follows again from the approximations derived at the end of Section 2. Since for a given (small) mass of seats q_1 (resp.q_2) the admission threshold in neighborhood 1 (resp. 2) departs from 1 by a term that is proportional to $(q_1)^{1/3}$ (resp. $(q_2)^{1/2}$), the only way to equate the admission threshold in the two neighborhoods is to have that q_1 is much smaller than q_2, hence the claim in Proposition 5.

Thus, in the case of a small mass of seats, we obtain that a less knowledgeable neighborhood gets much fewer seats than a fully knowledgeable neighborhood whether less knowledgeable means less informed (Proposition 5) or relying on the sampling heuristic (Proposition 4). We have also established in this limit that a neighborhood using the sampling heuristic would get much fewer seats than a neighborhood having incomplete information but relying on rational expectations (Proposition 4).

To conclude this subsection, we now investigate, for the case of a small mass q of seats, the effects of quotas on welfare and on the quality of admitted students. By quotas, we mean here allocating a mass of seats to neighborhood i in proportion to its size, i.e., imposing that $\frac{q_1}{m_1} = \frac{q_2}{m_2}$. But any alternative definition of quotas that would favor the neighborhood receiving fewer seats per head in the laissez-faire case would yield the same conclusion. We have:

PROPOSITION 6. *When neighborhood 1 relies on the local sampling heuristic ($\tau_1 = S$) and neighborhood 2 relies on rational expectations (and is not perfectly informed of θ, say $\tau_2 = CR$), imposing quotas is good for total welfare and bad for the average quality of admitted students when the total mass of seats q is small.*

This results from the approximations derived at the end of Section 2. More precisely, for welfare, it is more efficient to allocate a bigger mass of seats to neighborhood 1 than neighborhood 2 because the marginal effect on welfare of increasing the mass of seats is bigger in neighborhood 1 than in neighborhood 2 (since $\frac{1}{2} < 1 - \frac{E(\theta_2^2)}{2E(\theta_2)}$). For the quality of admitted students, the comparison goes the other way because the admission threshold is further away from 1 in neighborhood 1 than in neighborhood 2 for comparable masses of seats in the two neighborhoods. Given that in laissez-faire, neighborhood 1 would get almost no seats (Proposition 4), the result of Proposition 6 follows. It should be mentioned that in Proposition 6 we have assumed that in neighborhood 2 students do not know their ability type fully. Since the result holds no matter how small the part they do not observe is, we feel fair to extend the insight of this Proposition to the case in which students would almost perfectly observe their type in neighborhood 2, thereby covering in the limit the case in which $\tau_2 = FR$.

By contrast, we have that when both neighborhoods rely on rational expectations, total welfare is unaffected by how seats are assigned between the two neighborhoods (since the marginal effect of increasing the mass of seats is the same in both neighborhoods). In other terms, quotas seem overall beneficial for welfare purposes when the less knowledgeable neighborhood is modeled as relying on the sampling heuristic, but not when it is modeled as having inferior information.

3.2 Symmetrically knowledgeable neighborhoods

In this part, we consider the case of competition between two less knowledgeable neighborhoods and our interest lies in understanding how asymmetries in the distribution of characteristics affect how seats are distributed among the two neighborhoods. We will also analyze the effect of quotas in such environments. Throughout this part (and for tractability reasons), we restrict ourselves to the case of a small mass q of seats.

We start with the case of two neighborhoods relying on the local sampling heuristic. In the laissez-faire case, we have:

PROPOSITION 7. *As q gets small and $\tau_1 = \tau_2 = S$, the equilibrium with competing neighborhoods is unique and characterized by*

$$q_i = \frac{m_i \left(f_i(\overline{\phi})h_i(\overline{\omega})\right)^2 g_i(0)/E(\theta_i)}{m_1 \left(f_1(\overline{\phi})h_1(\overline{\omega})\right)^2 g_1(0)/E(\theta_1) + m_2 \left(f_2(\overline{\phi})h_2(\overline{\omega})\right)^2 g_2(0)/E(\theta_2)} q + o(q).$$

Moreoever, quotas in the form of allocating seats in proportion to the size of the neighborhood is welfare-improving whenever $\arg\min_i \left(f_i(\overline{\phi})h_i(\overline{\omega})\right)^2 g_i(0)/E(\theta_i) = \arg\min_i \frac{E(\theta_i^2)}{E(\theta_i)}$.

The first part of Proposition 7 describing seat allocations is derived straightforwardly using the approximations derived at the end of Section 2. While it is expected that the densities of f_i, h_i, and g_i around $\overline{\phi}$, $\overline{\omega}$ and 0 respectively would play a role (since q is assumed small), a distinctive feature of the competitive equilibrium when neighborhoods rely on the local sampling heuristics is the role played by the mean benefit $E(\theta_i)$ in the two neighborhoods. Freezing the differences in terms of $f_i(\overline{\phi})$, $h_i(\overline{\omega})$ and $g_i(0)$ across the two neighborhoods, Proposition 7 implies that the neighborhood with the larger mean benefit, i.e. $\arg\max_i E(\theta_i)$ obtains fewer seats per capita than the neighborhood with lower mean benefit. This is so because there are more applicants in the larger mean benefit neighborhood, which in turn results in a lower aggregate acceptance rate in this neighborhood. As a result, high benefit students apply at a lower rate in the larger mean benefit neighborhood, which overall leads this neighborhood to end up with a smaller mass of seats per capita. Regarding quotas, we can infer from the approximations in Section 2 that allocating more seats to the neighborhood i with lower value of $\frac{E(\theta_i^2)}{E(\theta_i)}$ is welfare-enhancing. The welfare effect of quotas follows straightforwardly from how seats are allocated in laissez-faire.[18]

In our next result, we consider the case of neighborhoods relying on rational expectations but in which students are uninformed of ω.

PROPOSITION 8. *As q gets small and $\tau_1 = \tau_2 = CR$, the equilibrium with competing neighborhoods is unique and characterized by*

$$q_i = \frac{m_i f_i(\overline{\phi}) h_i(\overline{\omega})^2 g_i(0)}{m_1 f_1(\overline{\phi}) h_1(\overline{\omega})^2 g_1(0) + m_2 f_2(\overline{\phi}) h_2(\overline{\omega})^2 g_2(0)} q + o(q).$$

Moroever, quotas have no effect on welfare.

Again, Proposition 8 follows from the approximations derived at the end of Section 2. We note that in this scenario, it is only the distributions of ϕ and ω around $\overline{\phi}$ and $\overline{\omega}$ respectively that matter. This is different from the finding in Proposition 7 in which the mean benefit in the various neighborhoods was shown to affect how seats are assigned between neighborhoods. That quotas have no effect on welfare is due to the observation made in the approximation results that no matter how ϕ, ω and c are distributed the marginal effect of increasing the mass of seats in a neighborhood with type $\tau = CR$ is invariably 1/2. In this case, assigning seats in proportion to the size of the neighborhood can be assessed for its redistributive merits and it is readily verified that quotas would also have a negative effect on the average quality of admitted students.

[18] In particular, we note that when $f_i(\overline{\phi})$, $h_i(\overline{\omega})$, $g_i(0)$ are the same across neighborhoods, quotas can be welfare-enhancing if the lower mean benefit neighborhood has a variance of benefit sufficiently higher than that of the larger mean benefit neighborhood.

4 Conclusion

Our study has revealed a novel channel through which disadvantaged communities (viewed as having coarse expectations about the admission odds because of lack of access to sufficiently detailed statistics) may end up being hurt relative to advantaged communities in terms of effective access to elite colleges. While standard policy instruments such as quotas may partially compensate for this and have a beneficial overall effect on total welfare, an alternative policy may consist in trying to de-bias the expectations used in such communities through appropriate interventions. Making admission tests mandatory as suggested by Goodman (2016) is such an intervention, and it has been shown to have a significant effect on education choices. Our analysis suggests that focusing on and improving such interventions is of vital importance to improve the self-selection of talented students into elite colleges.

PROOFS

Proof of Proposition 1

Step 1. ($p \to 0$ as $q \to 0$): By contradiction, suppose that there exists b with $p > b > 0$ for all q, then all (θ, c) such that $c < b\theta$ apply to C. The mass of applicants $m(q)$ is no smaller than $m^* = \Pr(c < b\theta)$ (and $m^* > 0$ where use is made of the full support assumption). But then $p(q) = \frac{q}{m(q)} \to 0$ and we get a contradiction.

Step 2. (Approximation of p in terms of θ^S): We have

$$\begin{aligned} p &= \frac{\int_{\theta^S}^1 k(\theta)G(p\theta)d\theta}{\int_0^1 k(\theta)G(p\theta)d\theta} \\ &\approx \frac{\int_{\theta^S}^1 k(\theta)p\theta g(0)d\theta}{\int_0^1 k(\theta)p\theta g(0)d\theta} \\ &= \frac{\int_{\theta^S}^1 k(\theta)\theta d\theta}{\int_0^1 k(\theta)\theta d\theta} \\ &= \frac{\int_{\theta^S}^1 k(\theta)\theta d\theta}{E(\theta)} \end{aligned}$$

using a 1st order Taylor approximation of G around 0. This approximation also implies that $\theta^S \to 1$ as $q \to 0$ to ensure that $p \to 0$ (from Step 1). This in turn implies that (using $k(\theta) \approx \sqrt{2} f(\overline{\phi})h(\overline{\omega})(1-\theta)$ when θ is in $(\theta^S, 1)$):

$$p \approx \frac{|k'(1)|}{E(\theta)} \frac{(1-\theta^S)^2}{2}$$

where $|k'(1)| = \sqrt{2} f(\overline{\phi})h(\overline{\omega})$.

Step 3. (Approximation of θ^* in terms of q):

$$\begin{aligned} q &= \int_{\theta^S}^1 k(\theta)G\left(p\theta\right)d\theta \\ &\approx \int_{\theta^S}^1 k(\theta)pg(0)\theta d\theta \\ &\approx pg(0)\left|k'(1)\right| \frac{(1-\theta^S)^2}{2} \end{aligned}$$

using the approximations $G(p\theta) \approx p\theta g(0)$, and $k(\theta) \approx |k'(1)|(1-\theta)$ for $\theta \in (\theta^S, 1)$.

Combining with the approximation of p, we get

$$q \approx \frac{|k'(1)|^2 g(0)}{E(\theta)} \frac{(1-\theta^*)^4}{4}$$

Writing θ^* in terms of q, we get

$$\theta^* = 1 - \frac{1}{|k'(1)|^{1/2}} \left(\frac{E(\theta)}{g(0)} \right)^{1/4} (4q)^{1/4} + o(q^{1/4}).$$

Step 4. (Approximation of p in q): From steps 3 and 4 we get

$$p = \left(\frac{q}{g(0)E(\theta)} \right)^{1/2} + o(q^{1/2}).$$

Step 5. (Approximation of $W(q)$): We have

$$\begin{aligned}
W(q) &= \int_{\theta^S}^{1} \theta k(\theta) G(p\theta)\, d\theta - \int_0^1 k(\theta) \left[\int_0^{p\theta} c g(c) dc \right] d\theta \\
&\approx g(0) p \,|k'(1)|\, \frac{(1-\theta^S)^2}{2} - \int_0^1 g(0) \frac{p^2\theta^2}{2} k(\theta) d\theta \\
&\approx g(0) p \,|k'(1)|\, \frac{(1-\theta^*)^2}{2} - \frac{g(0)E(\theta^2)}{2} p^2 \\
&\approx q - \frac{E(\theta^2)}{2E(\theta)} q \\
&= \left(1 - \frac{E(\theta^2)}{2E(\theta)} \right) q + o(q)
\end{aligned}$$

using the above approximations.

Step 6. (Approximation of $M(q)$): We have

$$\begin{aligned} M(q) &= \frac{\int_{\theta^S}^1 \theta k(\theta) G(p\theta) d\theta}{\int_{\theta^S}^1 k(\theta) G(p\theta) d\theta} \\ &\approx 1 - \frac{2}{3}(1-\theta^S) \\ &= 1 - \frac{2}{3}\left(\frac{1}{f(\overline{\phi})h(\overline{\omega})}\right)^{1/2}\left(\frac{2E(\theta)q}{g(0)}\right)^{1/4} + o(q^{1/4}) \end{aligned}$$

using the above approximations. □

Proof of Proposition 2

When students observe fully their benefit, a (θ, c) student applies whenever $\theta > \theta^R$ and $c < \theta$. Equating supply and demand implies

$$q \approx \frac{(1-\theta^R)^2}{2} f(\overline{\phi}) h(\overline{\omega}) G(1)$$

from which we get

$$\theta^R(q) = 1 - \left(\frac{2}{f(\overline{\phi})h(\overline{\omega})G(1)} q\right)^{1/2} + o(q^{1/2}).$$

When students do not observe ω, let θ^C be the admission threshold and let $\phi^C = \theta^C - \overline{\omega}$. A (ϕ, c) student if applying expects to be admitted whenever $\omega > \theta^C - \phi$ and so applies (noting that when q is small θ^C must be close to 1 and thus ϕ and ω must be close to $\overline{\phi}$ and $\overline{\omega}$ to be admitted). whenever $\phi > \theta^C$ and $c < (\overline{\omega} - \theta^C + \phi)$. Equating demand and supply yields:

$$\int_{\phi^C}^{\overline{\phi}} f(\overline{\phi})(\overline{\omega} - \theta^C + \phi)^2 \left(h\left(\overline{\omega}\right)\right)^2 g(0) d\theta \approx q$$

which after integration yields

$$\frac{1}{3}(1-\theta^C)^3 f(\overline{\phi}) \left(h\left(\overline{\omega}\right)\right)^2 g(0) \approx q$$

and thus

$$\theta^C(q) = 1 - \left(\frac{3}{f(\overline{\phi})h(\overline{\omega})^2 g(0)} q\right)^{1/3} + o(q^{1/3}).$$

Welfare and the average benefit of the admitted student can then easily be obtained as:

$$\begin{aligned} W(q) &= \frac{q}{2} + o(q) \\ M(q) &= 1 - \frac{2}{3}\left(\frac{3}{f(\overline{\phi})h(\overline{\omega})^2 g(0)} q\right)^{1/3} + o(q^{1/3}) \end{aligned}$$

□

Proof of Proposition 4.

From Proposition 1, we have that $\theta^S(q_1) \approx 1 - \left(\frac{1}{f(\overline{\phi})h(\overline{\omega})}\right)^{1/2} \left(\frac{2E(\theta)q_1}{m_1 g(0)}\right)^{1/4}$ where q_1 is the mass of seats obtained in neighborhood 1.

When $\tau_2 = FR$, we have that $\theta^R(q_2) \approx 1 - \left(\frac{2}{f(\overline{\phi})h(\overline{\omega})G(1)m_2} q_2\right)^{1/2}$ where q_2 is the mass of seats obtained in neighborhood 2.

Equating $\theta^S(q_1) = \theta^R(q_2)$ and imposing that $q_1 = q_2$ (as required in equilibrium) implies that $\lim_{q \to 0} \frac{q_1}{q_2} = 0$, as required.

When $\tau_2 = CR$, we have that $\theta^C(q_2) \approx 1 - \left(\frac{3}{f(\overline{\phi})h(\overline{\omega})^2 g(0) m_2} q_2\right)^{1/3}$ and we obtain similarly that $\lim_{q \to 0} \frac{q_1}{q_2} = 0$ in this case. □

Proof of Proposition 7 and 8. The proofs of these results follow straightforwardly from the expressions found in Propositions 1 and 2 and in the laissez faire case the requirement that the admission threshold is the same in both neighborhoods. □

References

Akbarpour, M., Kapor, A., Neilson, C., Van Dijk, W., and Zimmerman, S. (2022). Centralized school choice with unequal outside options. *Journal of Public Economics*, 210.

Ali, S. N. and Shorrer, R. I. (2025). Hedging when applying: Simultaneous search with correlation. *American Economic Review*.

Calsamiglia, C., Martínez-Mora, F., and Miralles, A. (2021). School choice design, risk aversion and cardinal segregation. *The Economic Journal*.

Enke, B. and Zimmermann, F. (2019). Correlation neglect in belief formation. *The Review of Economic Studies*, 86(1):313–332.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, 73(5):1623–1672.

Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics*.

Hastings, J. S. and Weinstein, J. M. (2008). Information, School Choice, and Academic Achievement: Evidence from two Experiments. *The Quarterly journal of economics*, 123(4):1373–1414.

Hoxby, C. M. and Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review*, 105(5):514–517.

Jehiel, P. (2005). Analogy-Based Expectation Equilibrium. *Journal of Economic theory*, 123(2):81–104.

Jehiel, P. (2018). Investment Strategy and Selection Bias: An Equilibrium Perspective on Overoptimism. *American Economic Review*, 108(6):1582–97.

Kapor, A. J., Neilson, C. A., and Zimmerman, S. D. (2020). Heterogeneous Beliefs and School Choice Mechanisms. *American Economic Review*, 110(5):1274–1315.

Manski, C. F. (1993). Adolescent Econometricians: How do Youth Infer the Returns to Schooling? In *Studies of Supply and Demand in Higher Education*, pages 43–60. University of Chicago Press.

Rees-Jones, A. and Shorrer, R. (2023). Behavioral economics in education market design: A forward-looking review.

Rees-Jones, A., Shorrer, R., and Tergiman, C. J. (2024). Correlation neglect in student-to-school matching. *American Economic Journal: Microeconomics*.

Streufert, P. (2000). The effect of underclass social isolation on schooling choice. *Journal of Public Economic Theory*.