School Rankings, Student Allocations and School Choice Reforms

Claudia Herresthal†

April 12, 2015

Working paper

Abstract

With school choice reforms, families choose where to apply and more public funding is allocated to highly-demanded schools. Given families’ informational constraints, it is unclear whether demand for places at high-quality schools increases. Families infer schools’ relative quality from performance-based rankings and trade off choosing a higher-ranked non-local school over their local school. I solve for a Bayesian-Nash equilibrium consistent with a steady-state level of informativeness of rankings. I find that rankings become more informative and more families apply to high-quality schools, if families can choose where to apply and schools can choose whom to accept. (JEL D82, D83, I21, I24, I28, H75)

*Additional Best Paper Prize awarded at the Warsaw International Economics Meeting. Earlier title: “Learning, Mobility and School Choice”.

†Department of Economics, University of Oxford, claudia.herresthal@economics.ox.ac.uk. The latest version of this paper is available at https://sites.google.com/site/dherresthal/research. This project was supported by funding from the Department of Economics, University of Oxford, and the Chellgren Scholarship, University College, University of Oxford. I would like to thank Meg Meyer for extraordinary advice and encouragement and Peter Eso, Sujay Mukerji, Pawel Dziewulski, Collin Raymond and Alex Teytelboym for guidance and feedback. I benefited from the helpful comments by Aytek Erdil, Li Hao, Johannes Horner, Emir Kamenica, Fuhito Kojima, Joel Sobel, Peter Sorensen, Tomasz Strzalecki, Adrien Vigier and the participants at INET Theory Lunch Workshop, Cambridge; CE2 workshop on Microeconomics, Warsaw; Warsaw International Economics Meeting; Public Economics UK conference, Bristol; York Symposium on Game Theory; CIREQ Montreal Matching conference; UECE Lisbon Meetings on Game Theory and Applications.
With school choice reforms, families choose where to apply instead of being assigned to schools based on their residence. These reforms are widespread: in the US, 15% of students attend parent-selected state schools rather than district-assigned schools (Greene et al., 2010); and in England, more than half of students attend a school that is not their nearest state school (Burgess et al., 2006). The aim of these “market-based” reforms is to identify high-quality schools based on families’ demand for places and then to assign more funding to these schools such that they can expand and cater to more students. In reality, families’ demand for places exceeds capacity at many schools. Therefore schools also have a choice over applicants, and this choice is often regulated by official admission codes.

Families are not perfectly informed about which schools add the most value to their child’s academic outcomes, i.e. they do not observe school quality. To allow families to make more informed application choices, governments have started to publish schools’ performance on standardized tests. In the US, school report cards were first required by the No Child Left Behind Act in 2002; and in England, school performance rankings were first published in 1992. While families have access to these performance indicators, they still do not observe school quality directly. This is because how well schools have performed on standardized tests depends not only on their own quality but also on the ability of their past student intake. What families infer about a school’s quality therefore also depends on their belief about how able the school’s intake was.

A recent debate has focused on which indicators would help families to identify good schools more easily. Statistical estimates of schools’ value-added are now available in many US school districts and are also published alongside raw exam scores in English league tables. One problem with these estimates is that they are very imprecise, making it hard to distinguish between schools (Leckie and Goldstein, 2009). Another problem is that families tend to base their choices on raw exam scores, rather than value-added measures. Therefore it is important to understand how rankings based on (raw) performance can be better indicators of schools’ relative quality; in particular, how the informativeness of these rankings is influenced by school choice reforms.

This paper studies the consequences of school choice reforms, when families are imperfectly informed about school quality but observe a ranking of schools based on past performance. My aim is to analyze whether school choice reforms achieve their aim of promoting high-quality schools. In particular, I analyze how school choice reforms affect the probability that families apply to high-quality schools. This is a complex question to study because school choice reforms influence

---

1In the US, the most common form of implementation are open enrollment laws, which allow families to choose publicly-funded schools outside their attendance area (inter-district) or outside their school district (inter-district). In addition, the introduction of magnet and charter schools increased the options available to families. In England, families can apply to any publicly-funded school outside their catchment area and the introduction of free schools has expanded the range of schools to choose from.

2In England, just below half of families use so called league tables published annually on the department for education website (Coldron et al., 2008), more may be aware of those printed by the media.

3e.g. see Deming [2014], Wilson and Piebalga [2008]

4Out of those that use league tables only 36% are interested in value-added scores [Coldron et al., 2008].
families’ application choices through both a direct and an indirect channel. They influence application choices directly by increasing the set of schools families can choose from. Further, they influence application choices indirectly by affecting the performance indicators from which families draw inferences about schools’ relative quality. This is because these reforms affect a school’s chances of attracting able students and hence its chances of performing well.

Given families’ informational constraints, for school choice reforms to succeed in promoting high-quality schools, rankings must become more informative. But it may seem plausible that school choice reforms make rankings less informative for the following reasons: If school districts have a similar composition of students and students are assigned to their local schools, then superior performance at some of the schools would suggest that these schools contributed more to students’ achievements, i.e. that they are of higher quality. By contrast, if students are able to choose between schools, some schools may receive more applications than there are places. As capacity will not expand sufficiently to cater for the additional demand, these schools will be able to select between applicants and recruit a more able intake than other schools. If families do not know which of the schools were oversubscribed in the past, then they cannot infer whether schools with better performance have a higher underlying quality or whether they happened to recruit better students. Consequently, it would be harder for families to identify high-quality schools. Further, one may worry that school choice reforms cause rankings to become self-fulfilling in the following sense: once a school performs well, families apply to this school, which allows the school to recruit able students and to perform well again, irrespective of its quality.

These arguments suggest that school choice reforms could fail to identify and promote high-quality schools, if oversubscribed schools are able to choose between students. If these arguments were correct, schools’ admissions should be regulated such that schools cannot select on ability, which is already the case for many schools in England and the US; e.g. some US schools allocate places by lottery. In England, banding systems have been used in the past to spread schools’ intake across the distribution of prior attainment.

However, this paper shows that these arguments are false. In fact, more families apply to high-quality schools, and rankings become better indicators of schools’ relative quality, if students choose where to apply and schools choose whom to accept. My results imply that school choice reforms are more likely to achieve their aim of promoting high-quality schools if we do not restrict oversubscribed schools to allocate places to applicants by lottery.

This paper is the first to study how the informativeness of school rankings depend on the way in which families and schools choose one another. Previous theoretical studies of families’ application choices have not made explicit the inference problem associated with rankings, and empirical studies have focused on how sensitive families’ demand for places is to schools’ performance, rather than to their quality, which is not directly observable. If we want to rely on market forces

---

5See De Fraja and Landeras [2006]
6For example, see Black [1999], Bradley et al. [2001], Belfield and Levin [2002], Hoxby [2000, 2003, ?], Bayer
to identify high-quality schools we need to gain a better understanding of how sensitive families’ demand for places is to underlying school quality. To develop this understanding, it is essential that we acknowledge families’ informational constraints and how their inference problem is influenced by the proliferation of performance data.

This paper models these complex relations among rankings, schools’ admissions and families’ choice of school. It delivers a tractable comparative statics analysis on the probability that families apply to high-quality schools. In particular, I perform comparative statics analysis on a Bayesian-Nash equilibrium consistent with a steady-state level of informativeness of performance-based rankings. First, I examine the impact of a reduction in families’ transport costs associated with attending a non-local school, which is effectively what school choice reforms achieve. Second, I study the effect of an increase in the probability that a school with a larger applicant pool can recruit a stronger intake, which is what policymakers can influence through the design of schools’ admission codes and student assignment mechanisms. For both forms of policy intervention, I analyze welfare implications. Further, I use these results to reveal unintended implications of ongoing reforms aimed at “leveling the playing field” for families from disadvantaged socioeconomic backgrounds, who are currently less likely to engage in school choice.

In the model, there are two schools, one of good and one of bad quality. Each is allocated in one of two school districts. Children differ in ability and the prior distribution of ability is the same in the two districts. A priori, families believe that each school is equally likely to be good. If a child is accepted at the school in the other district (non-local school), its family incurs a cost of transport, drawn at random from a stationary distribution. Before choosing where to apply, families observe their transport costs and a ranking of the two schools based on their students’ relative performance in the previous generation. Families compare schools based on their child’s expected educational outcome and any transport costs incurred, conditional on their child being accepted at the school. Schools are capacity constrained. Families can choose to apply to either their local or non-local school. If a school is oversubscribed, it selects (imperfectly) based on applicants’ ability. Students attend their preferred school if admitted and attend the other school, otherwise. Then student performance is realized which is increasing both in school quality and in student ability.

In equilibrium, families take the informativeness of rankings (signal quality) as given. Having observed the ranking of schools based on the previous generation’s performance (signal realization), they then update their beliefs about which school is better and decide to which one they apply. Their application decision is characterized by a cut-off level for transport costs. In particular, they apply to their non-local school if and only if this school has a higher rank than their local school and their transport costs realization lies below the cut-off level. The more informative the ranking is, the higher families’ optimal cut-off level and the more families apply to the higher-ranked school. The higher-ranked school can then select from a larger pool of applicants and will recruit a stronger student intake. In expectation, this raises the degree of sorting between students

and McMillan [2005], Gibbons et al. [2006]
and schools, i.e. how likely it is that the good school recruits stronger students. The higher the degree of sorting, the more likely it is that the good school will outperform the bad school because the good school’s underlying quality advantage is reinforced by a strong student intake. For these reasons, families’ application choices ultimately affect how informative future rankings are.

An equilibrium steady state is an equilibrium in which the level of informativeness of rankings (signal quality) is unchanged across generations. In particular, families make their optimal application choices taking as given how informative the observed ranking and then this causes the next ranking to be just as informative. I perform comparative statics on this steady-state equilibrium, first with respect to changes in the distribution of transport costs and second with respect to changes in schools’ capability to select among applicants.

In equilibrium steady state, a downward shift of the distribution of transport costs (in the sense of first order stochastic dominance) yields a higher equilibrium informativeness of school rankings and a higher optimal cut-off level for transport costs. Therefore school choice reforms, which effectively reduce transport costs, cause rankings to become more informative.

The logic behind this result is as follows: Consider a dynamic process in which different generations have different realizations of transport costs. In a generation, in which transport costs lie above families’ optimal cut-off levels, each student attends their local school. In this case, how informative future rankings will be depends only the distribution of student ability across school districts and is independent of how informative the observed ranking was. By contrast, in a generation, in which transport costs lie below families’ optimal cut-off levels, all families apply to the higher-ranked school. This school can then recruit better students. In this case, future rankings will be more informative than the observed ranking. This is because families’ application choice, together with schools’ admission choices, raise the likelihood that stronger students attend the school which performed better with students in the previous generation. This implies a higher expected degree of sorting between students and schools and hence causes future rankings to be more informative.

Using this reasoning, I identify an externality that families’ application choices and schools’ admission choices have on the outcomes of future generations. As more families apply to the better-performing school and this school selects able students, the rankings observed by future generations become more informative. Hence, the degree of sorting between students and schools also becomes stronger in future generations. Therefore even if only one generation were to benefit from a reduction in transport costs, future generations would observe more informative rankings and the degree of sorting between students and schools would be higher.

More generally, transport costs capture the degree of horizontal differentiation between schools as perceived by families. Schools differ in attributes which are unrelated to academic quality, and families’ preferences over those attributes differ. Another example, besides location, is schools’ facilities, e.g. some families highly value good sports grounds while others value fine art supplies. All families trade off a schools’ expected academic quality against these attributes. The more
general message of this paper is that if families perceive schools as less horizontally differentiated then in the equilibrium of the dynamic process, rankings become more informative about schools’ relative quality.

I also analyze the implications of increasing the likelihood with which schools can select the more able student in steady state. I show that, too, raises the informativeness of rankings if student ability and school quality are complements. When ability and quality are complements, there are two effects which reinforce one another. There is the direct effect that whenever both students apply to the the best-performing school, it is more likely to recruit a stronger student intake. In addition, given the complementarity of ability and quality, students’ expected benefit from attending the better school rises, inducing families to apply to the best-performing school more often. Both effects raise the expected degree of sorting between students and schools and hence the informativeness of rankings. If ability and quality are substitutes, increasing the likelihood with which schools can select the more able student has ambiguous effects on the informativeness of rankings. The degree of sorting between students and schools does not necessarily increase because students’ benefit from attending the better school falls, causing families apply to the best-performing school less often.

The welfare implications of such policy interventions depend on whether student ability and school quality are complements or substitutes. Lower transport costs raise informativeness and the degree of sorting between students and schools. In the case of substitutes, a higher degree of sorting lowers overall student performance. By contrast, in the case of complements, a higher degree of sorting raises overall student performance, but has ambiguous welfare effects. This is because a strong student who gets accepted at their non-local better-performing school imposes two kinds of costs on a weaker student who lives in the district with the better-performing school. This weaker students not only has a lower expected performance because he attends a school of lower expected quality but also has to incur the cost of commuting to this school. Complementarities ensure that the weaker student’s reduction in performance is outweighed by the gain in the stronger student’s performance. However, it is ambiguous whether the welfare gain from higher overall student performance outweighs the loss due to additional transport costs incurred.

Formally, the results about learning in this paper are related to the results derived from the study of biased contests [Meyer, 1991].\(^7\) In this paper, differences in the abilities of admitted students effectively bias the ranking of schools’ performance. The model also has some similarities to models in the social learning literature in which agents extract information about the quality of different options from predecessors’ actions (e.g. Bikhchandani et al. [1992]). However, in my model agents observe performance indicators for each of their options, which are influenced by how their predecessors chose between options, but they do not observe predecessor’s actions directly. My

\(^7\)In Meyer’s model, a firm wants to learn about the relative quality of two workers. It is optimal for the firm to assign a negative bias to the worker it believes to be of lower quality because by doing so it receives a stronger signal if this worker wins despite the disadvantage of the bias.
model shares with Lobel et al. [2007] the feature that agents have a limited window of observation but focuses on the steady state analysis rather than conditions for convergence. Callander and Hörner [2009] propose a steady state analysis but focus on agents inferring information from the relative frequency with which actions were taken by predecessors. The assignment problem which arises when schools are oversubscribed is the subject of an influential literature on matching algorithms. Although my paper does not form part of this literature, it adds the important insight that the allocation of students to schools influences what future generations believe about schools’ relative quality and hence which preferences over schools they submit to the algorithm.

The paper is organized as follows. Section 1 introduces the model and solves for equilibrium steady state. Section 2 derives comparative statics and analyzes welfare, then it discusses these findings. Section 3 gives detail on the convergence to steady state. Section 4 concludes. All proofs and a table summarizing the notation used can be found in the Appendix.

1 The Model

1.1 Players and Payoffs

Consecutive generations of families are matched with schools. In each generation, families choose where to apply and oversubscribed schools choose whom to accept. Payoffs depend on the match outcome.

Schools differ in quality. There is one good school \( G \) and one bad school \( B \), each associated with a school district. Families’ children differ in ability. In each generation, there is one child of the high-ability type \( H \) and one child of the low-ability type \( L \). There is one family per school district. Schools are randomly allocated to districts at the start. In each generation families are also randomly allocated to districts, independent of schools’ qualities.

Each family chooses which school to apply to. If a school receives applications from both families it can select which student to accept. The rejected student will attend the other school. I will assume that an oversubscribed school will select the high-ability student with probability \( p \).

Families attach value \( (v > 0) \) to a high educational performance \( (h) \). Educational performance depends on both their child’s ability and the school’s quality. In addition, families incur transport costs \( c \) if their child attends the school in the other district (non-local school). These transport costs are drawn in each generation from a known distribution with continuous cumulative distribution function \( F(c) \).

---

8 e.g. see Abdulkadiroglu and Sönmez [2003], Erdil and Ergin [2008].

9 Schools’ payoffs are not explicitly modelled. We can think of schools aiming to select the high-ability student because he helps the school achieve well or because he is less costly to teach. \( p < 1 \) could arise because schools are not capable of identifying relative ability perfectly or because schools are restricted in their selection of students by other priorities in their admission code, e.g. distance to school, siblings at the school.

10 These costs are only incurred when a student gets accepted at his non-local school. Applications are costless.
In period $t$, the expected utility of the family in district $i$, if their child attends school in district $j$, is given by:

$$E (U_{i,j}^t) = v \cdot P (h_t^j) - c_t^{i,j}$$

where $c_t^{i,j} = c_t$ if $i \neq j$ and $c_t^{i,j} = 0$ otherwise.

### 1.2 Timing and Information

It is common knowledge what types of students and schools exist, and that these are distributed independently and symmetrically across districts. In addition, the production function for educational performance and the distribution of transport costs, $F(c)$, are known.

Families do not observe the quality of schools. Each generation $t$ observes a school ranking (signal realization) based on the educational performance of generation $t - 1$. The school that achieves the higher rank based on educational outcomes of generation $t - 1$ is called the period-$t - 1$ winner, denoted by $W_{t-1}$.

Families do not observe the allocation of students to schools in the past, nor do they know which school ranked higher in any period prior to $t - 1$.

Families do not observe the ability of their own child, but they know that an oversubscribed school selects the high-ability student with probability $p \in \left(\frac{1}{2}, 1\right]$. Families form an expectation of their child's educational outcome, conditional on being accepted, based on the observed ranking ($W_{t-1}$). In addition, families observe the transport costs $c$ that they would incur if their child attended their non-local school.

Having observed $W_{t-1}$ and $c$, each family applies to one school. If oversubscribed, a school selects the high-ability student with probability $p$ and the rejected student attends the other school. Once the match between students and schools is determined, educational performance of generation $t$ is realized and a school ranking ($W_t$) is constructed.

Each generation of families corresponds to one period.

### 1.3 Production function

Educational performance can be either high ($h$) or low ($l$). Throughout the paper, let an overline (underline) denote association with the good (bad) school, e.g., let $\overline{h}$ ($\underline{h}$) denote the event that the student at the good (bad) school gets a high performance outcome. The probability of a high performance depends on both school quality and student ability.

The probability of high performance for a high-ability student

- at the good school is given by $P (\overline{h} | \overline{H}) = \alpha$
- at the bad school is given by $P (h | H) = 1 - \alpha$
The probability of high performance for the low-ability student

\[ \begin{align*}
\text{at the good school is given by } & P(h|L) = \bar{b} \\
\text{at the bad school is given by } & P(h|L) = 1 - \bar{b}
\end{align*} \]

where \( \bar{a}, \bar{a}, \bar{b}, \bar{b} \in [0, 1] \). Denote the vector characterizing the production process by \( A = (\bar{a}, \bar{a}, \bar{b}, \bar{b}) \).

**Monotonicity** - This production function is assumed to be monotonic in the following sense: Any type of student is more likely to achieve a high performance outcome at the good school than at the bad school:

\[ \Delta_{h} \equiv P(h|H) - P(h|L) = \bar{a} - (1 - \bar{a}) > 0 \]

and

\[ \Delta_{l} \equiv P(h|L) - P(h|L) = \bar{b} - (1 - \bar{b}) > 0 \]

In addition, the high-ability student is at least as likely to achieve a high outcome as the low-ability student, for a given school quality:

\[ \Delta_{g} \equiv P(h|H) - P(h|L) = \bar{a} - \bar{b} > 0 \]

and

\[ \Delta_{B} \equiv P(h|H) - P(h|L) = 1 - \bar{a} - (1 - \bar{b}) > 0 \]

**Complements and Substitutes** - Student ability and school quality are complements if

\[ \Delta_{H} - \Delta_{L} \equiv \Delta_{G} - \Delta_{B} \geq 0 \]

and substitutes if

\[ \Delta_{H} - \Delta_{L} \equiv \Delta_{G} - \Delta_{B} \leq 0 \]

### 1.4 Rankings

Schools are ranked each period based on their students’ performance. The school with a higher performance will come top in the ranking, and ties are broken at random. Denote the winning school by \( W \).

The probability that the good school wins with the high-ability student (against the bad school with the low-ability student) is given by:

\[ P(W|H, L) \equiv P(W|H) = \bar{a} \bar{b} + \frac{1}{2} (\bar{a} (1 - \bar{b}) + (1 - \bar{a}) \bar{b}) \equiv \alpha \]
The probability that the good school wins with the high-ability student (against the bad school with the low-ability student) is given by:

\[
P (W|H, L) \equiv P (W|H) = a\tilde{b} + \frac{1}{2} \left((1-a) \tilde{b} + a(1-\tilde{b})\right) \equiv \beta
\]

### 1.5 Equilibrium Steady State

This section derives how likely it is that families apply to the good school using the concept of equilibrium steady state. This will provide the basis for studying how exogenously determined factors influence the relationship between families’ application choices and school quality.

To analyze this dynamic process, we consider a steady state. A steady state is defined by a constant probability that the good school is ranked high. We then focus on those steady states which are consistent with equilibrium behavior. An equilibrium steady state is a steady state in which families behave optimally, taking as given other families’ behavior and the steady state probability that the good school is ranked high.

To start with, we define two objects that are helpful for characterizing an equilibrium steady state: mobility and informativeness. Period-\(t\) informativeness (\(I_t\)) captures generation \(t\)’s posterior beliefs about schools’ relative quality, which is based on both their symmetric prior beliefs and the observed period-\(t-1\) ranking. In addition, mobility in generation \(t\) (\(m_t\)) is defined by the likelihood with which both families in generation \(t\) apply to the period-\(t-1\) winning school, prior to the realization of their transport costs.

**Definition:** Denote the period-\(t-1\) winning school by \(W_{t-1}\).

(1) **Period \(-\)informativeness**, \(I_t \in [\frac{1}{2}, 1]\), is defined as

\[
I_t \equiv P (G|W_{t-1}) - P (B|W_{t-1}) \equiv 2P (G|W_{t-1}) - 1
\]

where \(P (G|W_{t-1}) \) \((P (B|W_{t-1}))\) denotes generation \(t\)’s posterior beliefs that the period-\(t-1\) winning school is the good (bad) school. These beliefs are based on their symmetric prior and the ranking of schools derived from performance with generation \(t-1\).

(2) **Mobility of generation \(t\)**, \(m_t \in [0, 1]\), is defined as the ex ante probability with which both families in generation \(t\) apply to the school which performed best with generation \(t-1\) \((W_{t-1})\).

For the process to be in a steady state, the level of mobility needs to be constant across generations. The reasoning is as follows: by definition, the likelihood that the good school is ranked high (wins) is constant in steady state. As rankings depend on the match between school
quality and student ability, the likelihood that the high-ability student is matched with the good school must also be constant in steady state. Families’ application choices are conditional on which school is the winning school. Their application choices are consistent with steady state, if and only if these application choices result in a constant probability that the high-ability student attends the winning school. This requires mobility to be constant across generations.

To solve for equilibrium steady state, consider a steady state characterized by mobility level $\hat{m}$ and then find the optimal choices of families in any given generation $t$. These families conjecture that the realization of period-$t-1$ rankings was generated as part of a steady state characterized by a constant mobility level $\hat{m}$ across all past generations. Their posterior beliefs about school quality, i.e. period-$t$ informativeness, is given by

$$I_t(\hat{m}; \alpha, \beta, p) = \frac{\alpha + \beta - 1}{(1 - \hat{m} (\alpha - \beta) (2p - 1))} (1)$$

Informativeness is weakly increasing in the conjectured level of mobility $\hat{m}$.\(^{11}\)

Families in generation $t$ decide where to apply, given their posterior beliefs. The family in the winning school’s district incurs no transport costs to attend the winner. Hence they always apply. The family in the losing school’s district will apply if and only if their transport costs lie below a cut-off level. Their optimal cut-off level for transport costs can be decomposed into i) their expected gain from attending the good rather than the bad school and ii) the likelihood that the winning school is the good rather than the bad school (period-$t$ informativeness). Families incur an expected gain of applying to the good school rather than the bad school, denoted by $V : \mathbb{R}^4 \rightarrow \mathbb{R}^+$, where

$$V(\Delta_H, \Delta_L, p, v) \equiv v ((1 - p) (\Delta_L) + p (\Delta_H))$$

This gain depends on their child’s ability, which is unknown. However, families draw inferences about their child’s ability conditional on its acceptance at the oversubscribed winning school. Consequently, the family in the losing school’s district is willing to incur their realized transport costs if and only if they lie below their cut-off level $V(\Delta_H, \Delta_L, p, v) \cdot I_t(\hat{m}; \alpha, \beta, p)$, henceforth $V \cdot I_t$. Hence the ex ante probability of both families applying to the winning school is given by

$$m_t = F (V \cdot I_t) (2)$$

where $F : V \cdot I_t \in \mathbb{R}^+ \rightarrow [0, 1]$.

The mobility level that characterizes optimal application choices by generation $t$ can be expressed as a best response to the conjectured mobility of past generations:\(^{11}\)

\(^{11}\)Further insights on the relationship between $\hat{m}$ and $I(\hat{m})$ can be found in the section on convergence (see Section 3).
\[ m_t(\hat{m}) = F(V \cdot I_t(\hat{m})) \]

An equilibrium steady state is a steady state such that generation \( t \)'s optimal choices are consistent with remaining in steady state: \( m_t = \hat{m} \). This implies that any generation’s optimal behavior is consistent with steady state. Such an equilibrium state state level of mobility (and informativeness) always exists. This is because mobility is a weakly increasing function of conjectured mobility and mobility is bounded above by 1.

**Proposition 1 [Equilibrium]:**

An equilibrium steady state level of mobility is characterized by

\[ m^* = F(V \cdot I(m^*)) \]  

(3)

and the corresponding equilibrium steady state level of informativeness is given by \( I(m^*) \). Such an equilibrium level of mobility (and informativeness) always exists.

How likely it is that both families apply to the good school in equilibrium steady state is given by \( \frac{m^*(I(m^*)+1)}{2} \).

2 Results

2.1 Comparative Statics

The following comparative statics analysis will focus on the smallest equilibrium steady state level of mobility \( (m^1) \) and the corresponding equilibrium steady state level of informativeness \( (I^1) \).

I will focus on the smallest equilibrium steady-state level of mobility for the following reasons. Consider a dynamic process which starts with the generation who first gains access to rankings. If each subsequent generation responds optimally to the mobility of all previous generations, then this dynamic process converges to the smallest steady state equilibrium level of mobility (henceforth equilibrium level of mobility). Further detail will be given in the section on Convergence (see section 3).

**Proposition 2 [Comparative Statics]**

The smallest equilibrium level of mobility, denoted by \( m^1 \), and the smallest equilibrium level of informativeness, denoted by \( I^1 \), both increase with

(a) a negative shift in FOSD of the distribution for transport costs \( F(c) \),
(b) an increase in valuation for high performance \( v \),
(c) an increase in the influence of school quality on educational performance, i.e. an increase in at least one of \( a, \bar{a}, b, \bar{b} \),
(d) an increase in the probability \( p \) that an oversubscribed school selects a high-ability student, if school quality and student ability are complements \( (\Delta_H - \Delta_L > 0) \).

If lower transport costs become more likely, in the sense that the distribution \( F(c) \) is subject to a negative shift in FOSD, then mobility increases in equilibrium steady state. Fix families’ posterior beliefs about school quality at their initial level, \( I \), then this change causes their unchanged best response (cut-off level \( V \cdot I \)) to be characterized by a higher level of mobility. To make their best response consistent with steady state, mobility has to increase.

Intuitively, lower transport costs have both a direct and an indirect effect on mobility. The direct effect is that more families apply to the winning school, holding fixed their expected gain from attending the winning school. The indirect effect is that informativeness and hence the expected gain from attending the winning school increase, also causing more families to apply to the winning school. The indirect effect is easier to understand if we consider how lower transport cost in generation \( t-1 \) affect period-\( t \) informativeness. As more families in generation \( t-1 \) apply to the winning school, the winning school is more likely to take on the high-ability student. Therefore the good school is more likely to be matched with the high-ability student. As the expected degree of sorting between students and schools increases, the good school’s quality advantage is more likely to be reinforced by the high-ability student. In brief, as more families in generation \( t-1 \) apply to the school they observe winning, families in generation \( t \) are more likely to observe the good school winning (higher informativeness). Because we consider a equilibrium steady state, both effects get absorbed into one increase in mobility.

An important general implication is that such effects exist for other exogenous changes which cause more families apply to the winning school. Such changes include i) an increase in families’ valuation for high performance or ii) a higher impact of school quality on performance or iii) an increase in school’s capability to select, if student ability and school quality are complements. All of the above lead to a higher expected benefit from attending the good school, conditional on their child being selected out of two applicants.

All of the comparative statics above also hold for the likelihood with which both families apply to the good school, which is given by \( \frac{m^*(I(m^*)+1)}{2} \).

2.2 Welfare

This section studies the welfare implications of changes to the distribution of transport costs, \( F(c) \), and to changes in schools’ capability to select, \( p \), in equilibrium steady state. We separately discuss effects on educational performance and total welfare.
We take *expected educational performance* to be the ex ante expected number of high educational performances in any given generation. Since we evaluate welfare from a utilitarian social planner’s point of view, we take *total welfare* to be the sum of both families’ ex ante expected utility in any given generation.

**Proposition 3 [Welfare]:**

(1) If and only if student ability and school quality are complements, i.e. $\Delta_G - \Delta_B \geq 0$, then the ex ante expected number of high performances increases with

(a) a negative shift in FOSD of the distribution for transport costs $F(\cdot)$.

(b) an increase in the valuation of high outcomes $v$.

(c) an increase in schools’ capability to select students based on ability, $p$.

(2) The utilitarian welfare function may decrease (increase) with (a), (b) and (c), even if student ability and school quality are complements (substitutes).

Lower transport costs, in the sense of a negative FOSD shift in $F(\cdot)$, cause expected performance to increase, if and only if student ability and school quality are complements. This is because, in the case of complements, lower transport costs lead to higher sorting between student ability and school quality. The impact on sorting works via two different channels: i) families apply to the winning school more often and, hence, apply to the good school more often (direct effect) and ii) rankings become more informative and hence families apply to the good school even more often (indirect effect). A higher expected degree of sorting is associated with a gain for the high-ability student and a loss for the low-ability student. The gain for the high-ability student outweighs the loss for the low-ability student, if and only if student ability and school quality are complements.

A negative FOSD shift in the distribution for transport costs does not necessarily lead to higher total welfare, even if expected performance increases. A family applies to their non-local winning school if and only if their expected gain, conditional on being accepted, outweighs their costs. However, they do not take into account that, if their child is accepted at the winning school, they impose an externality on the family living in the winning school’s district. This other family will then have to incur a cost of transport to attend the losing school and their child’s expected performance will decrease. In the case of complements, overall expected performance increases but this may not outweigh the increase in expenditure on transport costs.

This stark negative implication of lower transport costs on welfare is alleviated to some degree if the social planner is assumed to value educational achievement more than families, due to the positive externalities generated by education for society. In this case, the increase in educational outcomes is weighted more than the increase in expenditure on transport costs.
An analogous reasoning holds for other changes which cause sorting between student ability and school quality to increase, such as an increases in the valuation for high performance or an increase in schools' capability to select.

It is important to bear in mind that this analysis only considers the short term, in which the benefit from identifying good schools is not explicitly modelled. However, in the long term we expect such benefits to exist, for several reasons. First, we can expand those schools which are identified as good, so that they can cater to more students. This would increase overall student performance. Second, schools may be able to invest in the quality of their provision in the long term. They have greater incentive to invest, if their demand for places is sensitive to their quality, also leading to higher overall standards of schooling. With these benefits accounted for, welfare is more likely to increase with a policy intervention such as lower transport costs.

If we take into account the gains derived from identifying a good school in the long run, it becomes clear that a social planner may face a trade-off between long-term and short-term gains. Imagine that student ability and school quality are substitutes. To maximize short-term gains a social planner would like to increase the chances that the low-ability student attends the higher-ranked school. However, this would imply that future rankings become less informative and therefore would imply that it is harder to identify the good school in the long-run.

2.3 Discussion

This section discusses how various policies can improve what we can learn about school quality from performance-based rankings. It is important that these indicators identify schools which add value to children’s education, so that families apply to these schools, and so that these schools are expanded. In addition, once we have identified good schools, we can analyze what they do differently than their competitors and their best practices can be adopted by other schools. We will draw on the analysis of previous sections and discuss more generally the roles played by both transport costs and schools' selection of students.

2.3.1 The role played by the distribution of transport costs $F(c)$ and is implications for policy

This paper has found that lowering transport costs makes it easier to identify good schools based on their performance, if oversubscribed schools select applicants based on their ability. We will discuss a range of policies which effectively act as a reduction in transport, because they induce more families to apply to high-ranked schools.

The logic for why these policies make performance-based rankings more informative is as follows: high-ranked schools have a larger applicant pool to choose from and this has the consequence that high-ability students attend better-performing schools more often. Therefore, sorting between student ability and school quality increases. The intrinsic performance advantage of high-quality
schools is reinforced by strong intakes, and, as a result, performance-based rankings become better signals of schools’ relative quality.

School choice - School choice reforms reduce families’ costs associated with sending their child to a non-local school. This is because families are no longer required to move to this school’s attendance area (or school district); instead they can simply commute. With school choice reforms, families therefore place less weight on school location and more weight on perceived quality when deciding where to apply.

Student incentive-compatible assignment mechanisms - The design of the student assignment process can also influence how much weight families attach to schools’ relative performance in their application choices. In some assignment mechanisms, families have to trade off expressing a preference for a school to which their child is likely to be admitted against expressing a preference for one which they believe to be of high quality.\textsuperscript{12} Changing the admission process such that families can report their true preferences over schools would cause more families to apply to a higher-ranked school over their “safe option”.

Common standards in schools - Families are going to place more importance on schools’ relative performance the more similar they perceive schools to be along dimensions that are unrelated to quality. Policymakers could induce more families to apply to the best-performing schools by enforcing common standards in dimensions unrelated to quality, e.g. by requiring all schools to adopt the same core curriculum or to offer the same facilities.

Easily accessible rankings - It will take families some time and effort to research how schools compare in the most recent performance-based ranking. The more easily families can obtain this information, the more families will consider this ranking in their choice of school. That access to rankings can influence application choices has been shown by Hastings and Weinstein [2007].

2.3.2 The role played by schools’ capability to select and its implications for policy

The paper yields several important insights on how schools’ capability to select influences how informative performance-based rankings are. First, if schools cannot select among applicants, the policies outlined above (section 2.3.2) may not increase the informativeness of rankings. Second, we learn more about school quality if we let schools choose whom to accept rather than assign students by lottery, and we may learn even more if we expand schools’ capability to select.

No lottery or balancing - The policies outlined above (section 2.3.2) increase the number of families who apply to high-ranked schools. If schools cannot select among applicants then it is not clear whether this implies that able students are matched with better-performing schools more often.\textsuperscript{13} In particular, it would need to be the case that families with higher-ability children

\textsuperscript{12}Calsamiglia and Güell [2014] find empirical evidence that, in a system with a Boston mechanism, families base their preferences over schools on how high a priority their child is assigned at the schools. In particular, they found that many families switched their first choice school after neighborhood boundaries, and hence priorities, were reassigned.

\textsuperscript{13}A similar argument is made by Gavazza and Lizzeri [2009], but their argument focuses on what this implies for
attach relatively more weight to schools’ rank than families with lower-ability children. There is empirical evidence that this is the case: families from stronger socio-economic backgrounds tend to value education more, tend to be more informed about quality indicators and tend to be less budget-constrained [Greene et al., 2010, Allen et al., 2014].

Further, the paper refutes the conjecture that policymakers could make performance a better indicator of school quality, if they were to balance intakes across schools or assign students to schools by lottery. In the model framework, assigning students to schools by lottery corresponds to a value of \( p = \frac{1}{2} \) and we show that informativeness is higher for \( p > \frac{1}{2} \).

Empirical research has used lottery assignment to identify school quality by comparing the outcomes of lottery winners to those of lottery losers.\(^{14}\) We can clearly learn about school quality by using these methods, but we can also achieve this by letting schools choose whom to accept and using performance-based indicators. This method has the advantage that i) the indicators that families use to choose between schools become more informative and ii) schools gain an advantage from performing well, thereby providing stronger incentives for performance.\(^{15}\)

**Fewer priorities in admission codes** - This paper shows that policymakers can improve the identification of good schools by giving schools more freedom to select, if school quality and student ability are complements. More selection by schools increases sorting given families’ application choices. In addition, in case of complements, more selection causes more families to apply to high-ranked schools, which further increases sorting.\(^{16}\) In particular, this demonstrates a possible disadvantage of restricting schools’ selection through assigning priorities to students, e.g. based on their residence, whether they have siblings at the school or whether their parents work at the school. Debates surrounding the design school admission codes has so far not paid sufficient attention to how this design influences the identification of high-quality schools.

### 2.3.3 Student allocation as a performance advantage and its implications for policy

What the policies listed above have in common is that they increase the chances that a better-performing school ends up with a stronger student intake. Essentially, a school with a better signal about quality is assigned an advantage. This advantage makes it more likely that this school performs better than its competitor in the future. Policymakers can also assign such an advantage directly, instead of manipulating the allocation of students to schools.

**Performance-based rewards** - The interventions discussed so far have all targeted the allocation

\(^{14}\)e.g. see Cullen et al. [2005]; Deming [2014] uses lottery assignment to test how well estimates of school quality predict student outcomes.

\(^{15}\)Hatfield et al. [2011] analyze which assignment mechanisms ensure that a school is assigned a the set of students that is weakly better for that school whenever that it is more preferred by students. Schools’ preferences over students are based on admission priorities rather than on how easy it is to teach these students.

\(^{16}\)If families knew their child’s ability, then their expected benefit from sending their child to the better school is independent of schools’ capability to select. Therefore higher selection would increase the degree of sorting even in the case of substitutes.
of students to schools. In particular, we have shown that it becomes easier to identify high-quality schools if schools with better performance take on better students in the next generation. This mechanism relies on the fact that a better-performing school gets some advantage for future performance. A stronger intake is one example of such an advantage, and it is assigned to better-performing schools if students can choose where to go and schools can choose whom to accept. However, policymakers can create other channels through which better-performing schools get an advantage for future performance; e.g., they could assign additional funding to those schools.

2.3.4 Welfare

It is controversial to employ school choice reforms, or any of the other policies mentioned above. If families choose where to apply and schools choose whom to accept then the sorting between school quality and student ability increases. A direct consequence is an increase in inequality. Further, the gain in performance experienced by stronger students may not outweigh the loss experienced by weaker students. Even if there are complementarities then there is a cost of pushing weaker students out of schools they would have otherwise applied for, e.g. weaker students may have to attend schools which are not only of worse quality but also further from home.

However, these arguments miss out on the dynamic effects triggered by this assignment of students to schools. Even if school choice reforms apply to only one generation, families in all subsequent generations are more likely to apply to high-quality schools. To see why, consider the model framework and imagine that only generation $t$ experiences a reduction in transport costs. Then in generation $t$, the high-ability student is more likely to attend the period-$t - 1$ winning school. Hence, generation $t + 1$ expects a higher degree of sorting between school quality and student ability in generation $t$, which increases their posterior belief that the period-$t$ winning school is the good school. Given these beliefs, families in generation $t + 1$ are more willing to apply to the period-$t$ winning school. This, in turn, affects generation $t + 2$’s posterior beliefs in such a way that they are more willing to apply to the period-$t + 1$ winning school.

The above reasoning shows that policymakers face a trade off in case school quality and student ability are substitutes. By implementing policies which increase the likelihood that more able students attend better-performing, policymakers improve their knowledge of which schools are good over time. However, in any given generation, such policies decrease the expected level of performance and welfare.

This paper has focused on what policymakers can do to identify high-quality schools, but not explicitly modelled the advantages derived from knowing which schools are better than others. We may expect that good schools are expanded, while bad schools are closed down. In addition, as families’ demand for places becomes more sensitive to school quality, schools face stronger incentives to improve their quality. Consequently, we may expect that, in the long run, supply-side responses cause additional benefits through an overall improvement in school quality.
2.3.5 Application to health care

While this paper has focused on the context of education, some interesting parallels exist in the context of health care. Here too, we might expect that the outcome of medical treatment depends both on how sick the patient is and on how good the hospital or surgery is at treating the patient.

A 2006 UK policy allowed patients to choose between hospitals for specialized treatment. In 2007 this was complemented by a website which supplied information on provider quality, including mortality rates. If we apply the results of this paper to this context, then outcomes become a better indicator of hospital quality the more likely it is that better-performing hospitals take on healthier patients. This leads to a trade-off. To identify and expand high-quality hospitals we should give oversubscribed hospitals the option of selecting healthier patients. However, those patients benefit least from having been treated by high-quality hospitals.

In fact, in studies found that sicker patients responded more strongly to indicators of hospital quality and there was no evidence for selection by hospitals [Gaynor et al., 2012, Cooper et al., 2011]. In contrast to the education context, this self-selection by patients has adverse effects on the performance indicators of hospitals. This would suggest that the informativeness of performance indicators is reduced over time.

However, performance information was expressed by risk-adjusted mortality rates. Similar to value-added measures in the education context, these indicators depend not only on hospital quality but also on unobserved characteristics of patients’ health prior to treatment. Although there is no evidence that hospitals choose based on observable health characteristics, they may select healthier patients conditional on observable characteristics. Whether high-performing hospitals end up with relatively healthier patients (conditional on observable characteristics) depends on the relative strengths of the effect of patients’ self-selection rather than hospitals’ patient selection. An important difference between this context and education is that hospitals’ compensation varies with the difficulty of the patients’ treatment, while schools receive the same level of funding for each student. This may reduce incentives for hospitals to select patients based on their health status.\(^\text{17}\)

3 Convergence

The sequence of school rankings resulting from families’ optimal application choices has the interesting property that the informativeness of these rankings increases and converges over time. The limit of this sequence is smallest level of informativeness consistent with equilibrium steady state.

\textbf{Proposition 4 [Convergence]}

\(^\text{17}\)In the UK, schools receive additional funding when accepting students from poor backgrounds (pupil premium). It unclear as to whether this has led to lower covert selection by schools, but it has allowed schools to offer additional support to disadvantaged students.
Given the first ranking is based on educational outcomes of students in generation 0, the dynamic learning process, denoted by the sequence \( \{I_t : t = 0, 1, \ldots\} \) is characterized by informativeness of generation 0:

\[ I_0 = 0 \]

and by the recurrence equation linking informativeness of subsequent generations:

\[ I_t = \alpha + \beta - 1 + F(V \cdot I_{t-1}) (\alpha - \beta) (2p - 1) I_{t-1} \]

for all \( t \geq 1 \).

The limit of the dynamic learning process as \( t \to \infty \), denoted by \( I^{1*} \), is the smallest equilibrium level of informativeness. As \( I_t \) converges so does the sequence of mobility levels \( \{m_t : t = 0, 1, \ldots\} \). The sequence \( \{m_t : t = 0, 1, \ldots\} \) converges to the smallest equilibrium level of mobility, denoted by \( m^{1*} \):

\[ m^{1*} = F(V \cdot I^{1*}) . \]

Studying how informativeness evolves over time gives further insights into why informativeness in steady state depends positively on conjectured mobility (see 1). As an example, compare the inference made by the first two generations with access to rankings. Generation 1 is the first generation with access to rankings and observes schools’ performance with students of the previous generation. Further, generation 1 infers that in the previous generation each student attended his local school, as the previous generation’s posterior beliefs about school quality equal their symmetric prior beliefs. Generation 1 believes that the winner with students in the previous generation is the good school with probability:

\[ P(G|W_0) = P(W_0) = \frac{1}{2} (\alpha + \beta) > \frac{1}{2} \]

Hence,

\[ I_1 = \alpha + \beta - 1 \]

Note that \( \alpha \) and \( \beta \) characterize the intrinsic quality advantage the good school has over the bad school. The higher are \( \alpha \) or \( \beta \), the more informative is the ranking that results from a random assignment of students to schools. If educational performance outcomes are only determined by school quality, i.e. \( \alpha = \beta = 1 \), then \( I_1 = 1 \).
Generation 2 observes schools’ performance with students in generation 1. Taking as given that generation 1 had a mobility level equal to \( \hat{m}_1 \), generation 2 can infer the probability that the high-ability student in generation 1 attended the school that won with students in generation 0, i.e. \( P (H_1^W | W_0) \). Based on \( P (H_1^W | W_0) \), generation 2 can infer the probability that the good (bad) school won with students in generation 1, conditional on having won with students in generation 0.

\[
P (W_1 | W_0) = P (W_1 | H_1^W, W_0) P (H_1^W | W_0) + P (W_1 | L_1^W, W_0) P (L_1^W | W_0)
\]

\[
= \frac{(\alpha + \beta)}{2} + \frac{1}{2} \hat{m}_1 (2p - 1) (\alpha - \beta)
\]

Similarly,

\[
P (W_1 | W_0) = \frac{(2 - \alpha - \beta)}{2} + \frac{1}{2} \hat{m}_1 (2p - 1) (\alpha - \beta)
\]

Schools’ performance with students in generation 1 is independent of their performance with students in generation 0, if students in generation 1 apply to their local school independent of schools’ performance, i.e. if \( \hat{m}_1 = 0 \), or if schools have to randomize between applications, i.e. \( p = \frac{1}{2} \). Assuming \( p > \frac{1}{2} \), the higher is \( \hat{m}_1 \), the more likely the best-performing school with students in generation 0 will also be the best-performing school with students in generation 1. This is because the best-performing school with students in generation 0 is more likely to take on the high-ability student in generation 1 and any school is more likely to win with the high-ability student than the low-ability student, i.e. \( \alpha > \beta \). Importantly, for a given level of \( \hat{m}_1 \), the good school has a higher probability of being the best-performing school with students in generation 1 conditional on being the best-performing school with students in generation 0:

\[
P (W_1 | W_0) > P (W_1 | W_0)
\]

This is because the chances that the good school with a high-ability student wins against the bad school with a low-ability student are higher than the chances that the bad school with a high-ability student wins against the good school with a low-ability student, i.e. \( \alpha > 1 - \beta \). Consequently, a higher level of conjectured mobility raises the conditional probability of winning for both types of school by the same magnitude, but the good school has a higher absolute probability of winning for any given increase in conjectured mobility. I will refer to this effect that higher conjectured mobility has on the conditional probability of winning as the reinforcement effect.

In addition, generation 2 can infer the probability that the good (bad) school won with students in generation 0 (just like generation 1 did). Hence, generation 2 infers that the unconditional probability that the school winning with students in generation 1 is indeed the good school is
given by:

\[ P(\bar{W}_1) = P(\bar{W}_1|\bar{W}_0) P(\bar{W}_0) + P(\bar{W}_1|\bar{W}_0) P(\bar{W}_0) \]

\[ = \frac{(\alpha + \beta)}{2} + \hat{m}_1 \left(2p - 1\right) (\alpha - \beta) (\alpha + \beta - 1) \]

Hence,

\[ I_2 = (\alpha + \beta - 1) \left[1 + 2\hat{m}_1 (2p - 1) (\alpha - \beta)\right] \]

Due to the good schools’ intrinsic quality advantage, it is more likely to win with students in generation 0 than the bad school. An increase in conjectured mobility makes it more likely that the school that won with students in generation 0 is also the one that won with students in generation 1. In addition, conditional on having won with students in generation 0 the probability of winning with students in generation 1 is larger for the good school (reinforcement effect). Hence, a higher conjectured mobility of generation 1 means that the good school’s intrinsic quality advantage is reinforced and so the unconditional probability that the good school wins with students in generation 1 is higher.

In equilibrium, students in generation 1 apply to schools optimally given their level of informativeness and generation 2’s conjecture of mobility for generation 1 is correct, resulting in

\[ \hat{m}_1 = m_1 = F(V \cdot I_1) \]

So generation 1 is more mobile than generation 0 and hence generation 2 is better informed than generation 1. Mobility and informativeness increase over generations and eventually converge to the steady-state level.

4 Conclusion

This paper studies the impact of school choice reforms, when families infer school quality from performance-based rankings. I find that performance-based rankings become more informative, if families can choose where to apply and oversubscribed schools can choose whom to accept. Despite each generation of families being equally uninformed ex ante, school choice reforms cause more families apply to high-quality schools.

The analysis uses the concept of equilibrium steady state, i.e. it characterizes equilibrium strategies that are consistent with a constant informational value of rankings across generations. Comparative statics show that any change which increases the likelihood that a better-performing
school is matched with a strong intake raises the degree of sorting between student ability and school quality in equilibrium steady state. Hence, the performance advantage enjoyed by higher-quality schools is reinforced by strong intakes and therefore rankings are more likely to order schools according to their relative quality.

I discuss the impact of various policies on how informative performance-based rankings are, focusing on how they affect the degree of sorting between students and schools. Our findings are important in light of the recent debates on both school choice reforms and the design of admission procedures. In particular, I find that selection by schools is crucial, if we aim to identify high-quality schools based on families’ demand for places.

A motive for implementing school choice reforms is that these reforms will induce schools to compete for applicants. It is thought that in order to compete for applicants, schools will invest in the quality of their provision. My findings are an important first step in better understanding these aspects of school choice: I explore how families’ demand for places depends on school quality. However, I have not explicitly modelled schools’ payoffs from attracting more applications, and I have treated schools’ qualities as fixed.

I view the framework developed as a building block for future research about schools’ investments, when families infer school quality from performance-based rankings. I can modify the framework to explore how schools’ incentives to improve their performance interact with families’ beliefs and families’ application strategies in equilibrium steady state.

We can imagine the following modification to the set-up: schools are able to improve their performance with the current intake by exerting effort, e.g. teachers are working longer hours to support students. School types differ in how costly it is for them to exert effort, and the effect of their effort depends on the ability of their student intake. Schools aim to maximize their long-term payoff, taking into account how their choices today affect their position in the future, e.g. how their effort level today affects who will apply to them in the future. In equilibrium, schools condition their effort choice on their type and their current rank. In equilibrium steady state, schools’ effort strategies form a best response to one another given the (constant) application strategies of families. Families are not informed about schools’ types. In equilibrium steady state, families’ application strategies are optimal given schools’ effort strategies and the current ranking of schools.

Such analysis could yield further insights into whether school choice reforms are a policy instrument that “lifts the tide for all boats”, i.e. whether it causes each type of school to exert more effort. This links in with the industrial organization literature on how greater competition affects managers’ incentives for effort (e.g. Scharfstein [1988], Hart [1983]).
5 Bibliography

References


Simon Burgess, Deborah Wilson, and Jack Worth. A natural experiment in school accountability: the impact of school performance information on pupil progress and sorting. The Centre for Market and Public Organisation 10/246, Department of Economics, University of Bristol, UK, October 2010.


6 Appendix

6.1 Notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G, B$</td>
<td>School quality (good, bad)</td>
</tr>
<tr>
<td>$H, L$</td>
<td>Student ability (high, low)</td>
</tr>
<tr>
<td>$h, l$</td>
<td>Performance outcome (high, low)</td>
</tr>
<tr>
<td>$v \geq 0$</td>
<td>Families' valuation for high performance outcome</td>
</tr>
<tr>
<td>$c_t: c \sim F(c)$</td>
<td>Transport costs realization in $t$: drawn from stationary cdf $F(c)$</td>
</tr>
<tr>
<td>$p \in (\frac{1}{2}, 1]$</td>
<td>Probability school with two applications selects high-ability student</td>
</tr>
<tr>
<td>$\bar{H}(H)$</td>
<td>High-ability student at good (bad) school</td>
</tr>
<tr>
<td>$W_t: W_t, W_{t-1}$</td>
<td>Winner with students in generation $t$ (period-$t$ winner): good, bad school</td>
</tr>
<tr>
<td>$\alpha = P(W</td>
<td>\bar{H})$</td>
</tr>
<tr>
<td>$\beta = P(W</td>
<td>H)$</td>
</tr>
<tr>
<td>$I_t \in [0,1]$</td>
<td>Period $t$-informativeness</td>
</tr>
<tr>
<td>$m_t \in [0,1]$</td>
<td>Mobility of generation $t$</td>
</tr>
<tr>
<td>$V \geq 0$</td>
<td>Families' expected premium from attending good rather than bad school</td>
</tr>
<tr>
<td>$\Delta_H - \Delta_L \geq (\leq) 0$</td>
<td>Complements (substitutes) between student ability and school quality</td>
</tr>
<tr>
<td>$H_{t-1}^{W_t}$</td>
<td>High-ability student in generation $t$ attends period-$t-1$ winner</td>
</tr>
<tr>
<td>$x_t^{(\text{non})\text{local}W_{t-1}}$</td>
<td>Period-$t-1$ winner is (non)local school for student of ability $x$</td>
</tr>
<tr>
<td>$x_t^{W_{t-1}}$</td>
<td>Student of ability $x$ attends period-$t-1$ winner</td>
</tr>
</tbody>
</table>

6.2 Proof Proposition 1 [Equilibrium]

1. Mobility in terms of informativeness

Families incur transport costs only if their child is accepted at their non-local school. Winning is good news about school quality so either each student applies at their local school or both apply to the winning school. A student’s expected gain ($C$) of attending the winning school, conditional on having been selected among two applicants, can be decomposed into a student’s expected gain of attending the good school, conditional on having been selected among two applicants, weighted by informativeness.

\[
C = v \left( (1-p) \left( \tilde{b} - (1-\tilde{b}) \right) + p (\tilde{a} - (1-a)) \right) I_t \\
= v \left( (1-p) \left( \tilde{b} + \tilde{b} \right) + p (\tilde{a} + a) - 1 \right) I_t
\]

$C$ is the maximum level of transport costs families are willing to pay to send their child to the winning school given it is their non-local school. Their realized transport costs lie below this with
\[ m_t = F \left( \left( 1 - p \right) \left( b + \bar{b} \right) + p \left( \bar{a} + a \right) - 1 \right) I_t \].

(2) **Informativeness in terms of mobility**

Denote the event that the good school wins by \( \overline{W} \) and the event that the bad school wins by \( W \).

Define the state of the system realised in period \( t \) by \( \overline{W}_t \) and \( W_t \) respectively.

(1) The winner with students in period \( t - 1 \) is equally likely to be the high-ability student’s local or non-local school (events denoted by \( H_{t\text{local}}^{W_{t-1}} \) and \( H_{t\text{non-local}}^{W_{t-1}} \) respectively). With probability \( 1 - m \) each student applies to his local school and gets accepted. With probability \( m \) both students apply to the with winner with students in period \( t - 1 \) and the high-ability student is accepted with probability \( p \). Denote the event that the high-ability student attends the winner in period \( t - 1 \) by \( H_{t}^{W_{t-1}} \). Then

\[
P \left( H_{t}^{W_{t-1}} \right) = P \left( H_{t}^{W_{t-1}} | H_{t\text{local}}^{W_{t-1}} \right) P \left( H_{t\text{local}}^{W_{t-1}} \right) \\
+ P \left( H_{t}^{W_{t-1}} | H_{t\text{non-local}}^{W_{t-1}} \right) P \left( H_{t\text{non-local}}^{W_{t-1}} \right) \\
= \frac{1}{2} \left( 1 - m + 2mp \right)
\]

(2) The probability that the good school wins with the high-ability student (against the bad school with the low-ability student) is \( P \left( \overline{W}_t | H_t \right) = \alpha \) and the probability that the good school wins with the low-ability student (against the bad school with the high ability student) is \( P \left( \overline{W}_t | L_t \right) = \beta \).

(See production process in section (1.3)).

(1) and (2) are combined to construct transition probabilities from state \( j \) to state \( i \), denoted by \( P \left( i | j \right) \):

\[
P \left( \overline{W}_t | \overline{W}_{t-1} \right) = P \left( \overline{W}_t | \overline{H}_t \right) P \left( H_{t}^{W_{t-1}} \right) + P \left( \overline{W}_t | L_t \right) P \left( L_{t}^{W_{t-1}} \right) \\
= \alpha \left( \frac{1}{2} \left( 1 - m + 2mp \right) \right) + \beta \left( \frac{1}{2} \left( 1 + m - 2mp \right) \right)
\]

and

\[
P \left( W_t | \overline{W}_{t-1} \right) = \left( 1 - P \left( \overline{W}_t | L_t \right) \right) P \left( H_{t}^{W_{t-1}} \right) + \left( 1 - P \left( \overline{W}_t | \overline{H}_t \right) \right) P \left( L_{t}^{W_{t-1}} \right) \\
= \left( \frac{1}{2} \left( 1 - m + 2mp \right) \right) + \left( \frac{1}{2} \left( 1 + m - 2mp \right) \right)
\]

The transition matrix \( T \) is given by:
\[ T = \begin{pmatrix} P(W_t | W_{t-1}) & P(W_t | W_{t-1}) \\ P(W_t | W_{t-1}) & P(W_t | W_{t-1}) \end{pmatrix} \]

If all transition probabilities are strictly positive the Markov chain is both irreducible and aperiodic and therefore has a unique stationary distribution which is characterised by the row vector \( \begin{pmatrix} P(W) & P(W) \end{pmatrix} \) that satisfies both

\[
\begin{pmatrix} P(W) & P(W) \end{pmatrix} = \begin{pmatrix} P(W) & P(W) \end{pmatrix} \begin{pmatrix} T \\ T \end{pmatrix}
\]

and

\[ P(W) + P(W) = 1 \]

Further, these conditions are sufficient for the transition probabilities to converge to the stationary distribution:

\[
\lim_{k \to \infty} T^k = \begin{pmatrix} P(W) & P(W) \end{pmatrix}
\]

where \( \mathbf{1} \) is the column vector with all entries equal to 1.

Therefore if all transition probabilities are positive then the stationary distribution is characterised by

\[ P(W) = \frac{\alpha + \beta - (\alpha - \beta) m (2p-1)}{2 (1 - m (\alpha - \beta) (2p-1))}. \]

The Markov chain is not irreducible if and only if \( \alpha = 1 \) and i) \( \beta = 1 \) or ii) \( m = 1 \) and \( p = 1 \) and \( \beta < 1 \). In both cases the state \( W \) is an absorbing state and I assign \( P(W) = 1 \).

So

\[ P(W) = \frac{\alpha + \beta - (\alpha - \beta) m (2p-1)}{2 (1 - m (\alpha - \beta) (2p-1))}. \]

for all parameter configurations.\(^{38}\)

Families’ updated beliefs are based on their prior beliefs about school quality as well as on their conjectured level of mobility \( \hat{m} \) and their observation of which school won last period.

\[ P(G|W) = \frac{P(W) P(G)}{P(W) P(G) + P(W) P(B)} \]

By symmetry of prior beliefs,

\(^{38}\)Take the limit as \( x \) tends to 0, then take the limit as \( m \) tends to 1.
\[ P(G|W) = \frac{P(W)}{P(W) + P(W|B)} = P(W) \]

By definition of informativeness,

\[ I(\hat{m}) \equiv P(G|W) - P(B|W) = \frac{\alpha + \beta - 1}{(1 - m(\alpha - \beta)(2p - 1))} \]

Further,

\[ \frac{\partial I(\hat{m})}{\partial \hat{m}} = \frac{(\alpha - \beta)(\alpha + \beta - 1)(2p - 1)}{2(1 - m(\alpha - \beta)(2p - 1))^2} \geq 0 \]

and

\[ \frac{\partial^2 I(\hat{m})}{\partial^2 \hat{m}} = \frac{(\alpha - \beta)^2(\alpha + \beta - 1)(2p - 1)^2}{2(1 - m(\alpha - \beta)(2p - 1))^3} \geq 0 \]

as \( \alpha \geq 1 - \beta \) and \( \frac{1}{2} < p \leq 1 \) and \( 0 \leq \hat{m} \leq 1 \).

(3) Equilibrium

By definition of equilibrium,

\[ m^* \equiv m = \hat{m} \]

By part (1) and (2) of this proposition,

\[ m = F(V \cdot I(\hat{m})) \]

Hence, the equilibrium level of mobility \( m^* \) is characterised by

\[ m^* = F(V \cdot I(m^*)) \]

(4) Existence of equilibrium

An equilibrium level of mobility satisfies condition (3) in part (3). Let \( G(\hat{m}) \equiv F(V \cdot I(\hat{m})) \). \( G \) is monotone increasing. By Tarski’s fixed point theorem there exist an \( m^* \) such that \( G(m^*) = m^* \).

6.3 Proof Proposition 2 [Comparative Statics]:

Using conditions (2) and (1), an equilibrium level of informativeness satisfies
\[ I = \alpha + \beta - 1 + F (V \cdot I) (\alpha - \beta) (2p - 1) I \tag{4} \]

Let

\[ Z (I) \equiv \alpha + \beta - 1 + F (V \cdot I) (\alpha - \beta) (2p - 1) I \]

Denote by \( I^{1*} \) the smallest equilibrium level of informativeness satisfying condition 4. Note that \( Z (I, t) : [0, 1] \times T \to [0, 1] \), where \( T \) is a partially ordered set. For all \( t \in T \), \( Z \) is continuous but for upward jumps. In addition \( Z \) is monotone nondecreasing in \( t \) for all \( I \in [0, 1] \). After Corollary 1, (p. 446), in Milgrom and Roberts (1994a), (henceforth MR), the function \( I^{1*} (t) \) is monotone nondecreasing in \( t \).

(a) FOSD shift in distribution for transport costs \( F \)

Consider

\[ Z (F, I) \equiv \alpha + \beta - 1 + F (V \cdot I) (\alpha - \beta) (2p - 1) I \]

For any \( I \) and any \( F, \overline{F} \) such that \( F \) first-order stochastically dominates \( \overline{F} \), i.e. for any \( c \)
\[ F (c) \leq \overline{F} (c) : \]

\[ Z (\overline{F}, I) - Z (F, I) = [\overline{F} (V \cdot I) - F (V \cdot I)] (\alpha - \beta) (2p - 1) I \geq 0. \]

since \( (\alpha - \beta) (2p - 1) I \geq 0 \). By MR, a negative shift in FOSD of \( F \) increases \( I^{1*} (F) \).

By condition 2,

\[ m^{1*} (F) = F (VI^{1*}) \]

A negative shift in FOSD of \( F \) increases \( m^{1*} \):

\[ m^{1*} (F) = F (VI^{1*} (F)) \leq F (VI^{1*} (\overline{F})) = m^{1*} (\overline{F}) \]

since \( I^{1*} (\overline{F}) \geq I^{1*} (F) \), \( F \) and \( \overline{F} \) are increasing and \( F \) first-order stochastically dominates \( \overline{F} \).

(b) Increase in valuation for high outcomes \( v \)

Consider

\[ Z (v, I) \equiv \alpha + \beta - 1 + F (V \cdot I) (\alpha - \beta) (2p - 1) I \]

32
For any $v, v'$ such that $v' > v$, then $V(v') \geq V(v)$. Hence for any $I$ and any $v, v'$ such that $v' > v$:

$$Z(v') - Z(v) = (\alpha - \beta)(2p - 1)I \left[ F\left(V\left(v'\right)I\right) - F\left(V\left(v\right)I\right) \right] \geq 0$$

since $(\alpha - \beta)(2p - 1)I \geq 0$, $I \geq 0$, $V(v) \geq 0$ and $F$ is increasing. By MR, $I^1(v)$ increases in $v$.

By condition 2,

$$m^1(v) = F\left(V(v)I^1\right)$$

An increase in $v$ increases $m^1$:

$$m^1\left(v'\right) = F\left(V\left(v'\right)I^1\left(v'\right)\right) \geq F\left(V\left(v\right)I^1\left(v\right)\right) = m^1(v)$$

since $I^1(v') \geq I^1(v)$, $V(v') \geq V(v)$ and $F$ is increasing.

(c) Increase in the impact of school quality on outcomes

Consider

$$Z(\bar{\alpha}, I) \equiv \alpha(\bar{\alpha}) + \beta - 1 + F(\bar{\alpha}) \cdot I \left( \alpha(\bar{\alpha}) - \beta \right)(2p - 1)I$$

where $\alpha(\bar{\alpha}) = \frac{\bar{\alpha} + b}{2}$.

For any $\bar{\alpha}, \bar{\alpha}'$ such that $\bar{\alpha}' > \bar{\alpha}$ and any $I$,

$$Z\left(\bar{\alpha}'\right) - Z\left(\bar{\alpha}\right) = \left[\frac{\bar{\alpha}' - \bar{\alpha}}{2}\right] \left[1 + (2p - 1)F\left(\bar{\alpha}\right)I\right]$$

$$+ \left(\frac{\bar{\alpha}' + b}{2} - \beta\right)(2p - 1) \left\{ F\left(V\left(\bar{\alpha}'\right)I\right) - F\left(V\left(\bar{\alpha}\right)I\right) \right\} I \geq 0$$

since $V(\bar{\alpha}') \geq V(\bar{\alpha}) \geq 0$, $\frac{1}{2} < p \leq 1$, $0 \leq I \leq 1$, $0 \leq F \leq 1$ and $F$ is increasing. By MR, $I^1(\bar{\alpha})$ increases in $\bar{\alpha}$.

By condition 2,

$$m^1(\bar{\alpha}) = F\left(V\left(\bar{\alpha}\right)I^1(\bar{\alpha})\right)$$

An increase in $\bar{\alpha}$ increases $m^1$:

$$m^1\left(\bar{\alpha}'\right) = F\left(V\left(\bar{\alpha}'\right)I^1\left(\bar{\alpha}'\right)\right) \geq F\left(V\left(\bar{\alpha}\right)I^1\left(\bar{\alpha}\right)\right) = m^1(\bar{\alpha})$$
since \( I^*_1(\bar{\sigma}) \geq I^*_1(\bar{\pi}) \geq 0, \quad V(\bar{\sigma}) \geq V(\bar{\pi}) \geq 0 \) and \( F \) is increasing.

A similar argument shows that an increase in \( \underline{b} \) increases \( I^{1*} \) and \( m^*_1 \).

Consider

\[
Z(a, I) \equiv \alpha + \beta(a) - 1 + F(V(a) \cdot I) (\alpha - \beta(a)) (2p - 1) I
\]

For any \( a, \ a' \) such that \( a' > a \) and any \( I \),

\[
Z\left(\frac{a' - a}{2}\right) \left[ 1 - (2p - 1) F(V(a) I) \right] + \left( \alpha - \frac{a' + b}{2} \right) (2p - 1) \left\{ F\left(V\left(\frac{a'}{2}\right) I\right) - F(V(a) I) \right\} I \geq 0
\]

since \( V(\bar{a}') \geq V(\bar{a}) \geq 0, \quad \frac{1}{2} < p \leq 1, \quad 0 \leq I \leq 1, \quad 0 \leq F \leq 1 \) and \( F \) is increasing. By MR, \( I^{1*}(a) \) increases in \( a \).

By condition 2,

\[
m^{1*}(a) = F\left(V(a) I^{1*}(a)\right)
\]

An increase in \( a \) increases \( m^*_1 \):

\[
m^{1*}\left(a'\right) = F\left(V\left(a'\right) I^{1*}\left(a'\right)\right) \geq F\left(V(a) I^{1*}(a)\right) = m^{1*}(a)
\]

since \( I^*_1(a') \geq I^*_1(a) \geq 0, \quad V(a') \geq V(a) \geq 0 \) and \( F \) is increasing.

A similar argument shows that an increase in \( \underline{b} \) increases \( I^{1*} \) and \( m^*_1 \).

(d) Increase in schools’ capability to select based on ability, \( p \)

Consider

\[
Z(p, I) \equiv \alpha + \beta - 1 + F(V \cdot I) (\alpha - \beta) (2p - 1) I
\]

For any \( I \) and any \( p, p' \) such that \( p' > p \)

\[
Z\left(p'\right) - Z(p) = I (\alpha - \beta) \left[ (2p - 1) \left\{ F\left(V\left(p'\right) I\right) - F(V(p) I) \right\} + 2 \left(p' - p\right) F\left(V\left(p'\right) I\right) \right]
\]

Assume complements, i.e. \( \bar{a} - \bar{b} \geq \underline{a} - \underline{b} \), then \( Z(p') - Z(p) \geq 0 \), since \( V(p') \geq V(p), \quad 0 \leq I \leq 1, \quad 0 \leq x \leq 1, \quad \frac{1}{2} \leq p \leq 1, \quad 0 \leq F \leq 1 \) and \( F \) is increasing. By MR, if \( \bar{a} - \bar{b} \geq \underline{a} - \underline{b} \), then \( I^{1*}(p) \) is increasing in \( p \).

By condition 2,
\[ m^{1*}(p) = F(V(p)I^{1*}(p)) \]

If \( \beta \geq \alpha \), then mobility increases with \( p \):
\[ m^{1*}(p') = F(V(p')I^{1*}(p')) \geq F(V(p)I^{1*}(p)) = m^{1*}(p) \]

since \( I^{1*}(p') \geq I^{1*}(p) \), \( V(p') \geq V(p) \) and \( F \) is increasing.

6.4 Proof Proposition 3 [Welfare]

Total expected performance in equilibrium steady state is given by:
\[ 1 - a + b + P(H)(\Delta_H - \Delta_L) \]

where \( P(H) \) denotes the probability that the high-ability student attends the good school in equilibrium steady state, which is endogenously determined.

If student ability and school quality are complements, i.e. if \( \Delta_H - \Delta_L \geq 0 \), then total expected performance increases if and only if \( P(H) \). I will first derive \( P(H) \) and then perform comparative statics with respect to exogenously determined (a) \( F(c) \), (b) \( v \) and (c) \( p \).

Derivation: probability that high-ability student is at the good school in equilibrium steady state
\[ P(H) = \frac{1 - (1 - \alpha - \beta) m^{1*}(2p - 1)}{2 - (2 - \alpha - \beta) m^{1*}(2p - 1)} \]

Denote the event that the high-ability student is at the good school by \( H \) and the event that the high-ability student is at the bad school by \( H_t \). Further, denote the state realised in period \( t \) by \( \overline{H}_t \) and \( H_t \) respectively.

1) The probability that the good school wins with the high-ability student (against the bad school with the low-ability student), is \( P(\overline{W}_t|\overline{H}_t) = \alpha \) and the probability that the good school wins with the low-ability student (against the bad school with the high ability student) is \( P(\overline{W}_t|H_t) = \beta \). (See production process in section 1.3).

2) The probability that in period \( t \) the high-ability student attends the school that won with students in period \( t - 1 \), denoted by \( P(H^{W_{t-1}}_t) \), is given by
\[ P(H^{W_{t-1}}_t) = \frac{1}{2} (1 - m + 2pm) \]
(see Appendix Proof Proposition on Equilibrium)

(1) and (2) are combined to construct transition probabilities between state $j$ and state $i$ denoted by $P(i|j)$:

$$
P(H_t|H_{t-1}) = P(H^{W_{t-1}}) P(W_{t-1}|H_{t-1}) + (1 - P(H^{W_{t-1}})) (1 - P(W_{t-1}|H_{t-1}))
= \frac{1}{2} \alpha (1 - m + 2mp) + \frac{1}{2} (1 - \alpha) (1 + m - 2mp)
$$

and

$$
P(H_t|H_{t-1}) = (1 - P(H^{W_{t-1}})) P(W_{t-1}|H_{t-1}) + P(H^{W_{t-1}}) (1 - P(W_{t-1}|H_{t-1}))
= \frac{1}{2} \beta (1 + m - 2mp) + \frac{1}{2} (1 - m + 2mp) (1 - \beta)
$$

The transition matrix $T$ is given by:

$$
T(m, p, x) = \begin{pmatrix}
P(H_t|H_{t-1}) & P(H_t|H_{t-1}) \\
(1 - P(H^{W_{t-1}})) P(W_{t-1}|H_{t-1}) & P(W_{t-1}|H_{t-1})
\end{pmatrix}
$$

The chain is irreducible and aperiodic unless $\alpha = 1$ and $m = 1$ and $p = 1$. Therefore this Markov chain has a unique stationary distribution characterised by

$$
P(H) = \frac{1 + m (2\beta - 1) (2p - 1)}{2 (1 - m (\alpha - \beta) (2p - 1))}
$$

If $p = 1$, $m = 1$ and $\alpha = 1$ then $H$ is an absorbing state and I assign $P(H) = 1$.

**Proposition Part 1 (a)**

$R$ depends on $F$ only through the fixed point $m^*_1$. As the proposition on comparative statics shows, a negative FOSD shift in $F$ increases $m^*_1$. Hence $R$ increases with a negative FOSD in $F$ if and only if $R$ increases with an increase in $m^{1*}$.

---

19 Order of limits matters, first let $p$ tend to 1 then let $m$ tend to 1.
\[
\frac{\partial R}{\partial m_1^*} = \frac{\partial P(H)}{\partial m_1^*} (\bar{\alpha} + \alpha - \bar{\beta} - \beta)
\]

(5)

where

\[
\frac{\partial P(H)}{\partial m_1^*} = \frac{(\alpha + \beta - 1)(2p - 1)}{2((1-m_1^*(\alpha - \beta)(2p - 1))^2 \geq 0}
\]

for all parameter configurations. \((P(H)\) is given in remark on welfare). Hence, \(R\) increases with a negative FOSD in \(F\) if there are complements, i.e. \(\bar{\alpha} + \alpha - \bar{\beta} - \beta \geq 0\) and decreases if there are substitutes, i.e. \(\bar{\alpha} + \alpha - \bar{\beta} - \beta \leq 0\).

**Proposition Part 1 (b)**

\(R\) depends on \(v\) only through the fixed point \(m_1^*\). As the proposition on comparative statics shows, an increase in \(v\) increases \(m_1^*\). By a similar argument as above, \(R\) increases with an increase in \(v\) if there are complements, i.e. \(\bar{\alpha} + \alpha - \bar{\beta} - \beta \geq 0\) and decreases if there are substitutes, i.e. \(\bar{\alpha} + \alpha - \bar{\beta} - \beta \leq 0\).

**Proposition Part 1 (c)**

\(R\) depends on \(p\) directly as well as through the fixed point \(m_{1^*}\). Take \(p'\) and \(p''\) such that \(p'' > p'\).

\[
R(m_{1^*}(p''),p'') - R(m_{1^*}(p'),p') = \\
\underbrace{R(m_{1^*}(p''),p'') - R(m_{1^*}(p'),p')}_{\text{direct effect}} \\
+ \underbrace{R(m_{1^*}(p''),p'') - R(m_{1^*}(p'),p'')}_{\text{indirect effect}}
\]

By the proposition on comparative statics and part 1(a), the indirect effect is positive if there are complements, i.e. \(\bar{\alpha} + \alpha - \bar{\beta} - \beta \geq 0 \geq 0\). Further, the indirect effect is 0 whenever \(p'\) is such that \(m_{1^*} = 1\). The direct effect can be expressed as

\[
\frac{\partial R}{\partial p} = (\bar{\alpha} + \alpha - \bar{\beta} - \beta) \frac{\partial P(H)}{\partial p}
\]

where
\[
\frac{\partial P (\mathcal{H})}{\partial p} = \frac{m_{1\ast} (\alpha + \beta - 1)}{(1 - m_{1\ast} (\alpha - \beta) (2p - 1))^2} \geq 0
\]

for all parameter configurations.

Consequently, the total effect of \( p \) is positive if there are complements, i.e. \( \bar{\pi} + a - \bar{b} - b \geq 0. \)

**Proposition Part 2**

To be completed.

### 6.5 Proof Proposition 4 [Convergence]

**Learning process**

Generation 0 has no access to rankings and therefore holds symmetric posterior beliefs about relative schools quality, hence \( I_0 = 0. \) Informativeness of generation \( t \) depends on how likely it is that the good school wins with students in generation \( t - 1. \) Using transition probabilities from the proof of the proposition on equilibrium, part 2), and evaluate them at \( m_{t-1} \) yields

\[
P (W_{t-1}) = P (W_{t-1}|W_{t-2}) P (W_{t-2}) + P (W_{t-1}|W_{t-2}) (1 - P (W_{t-2}))
= \frac{1}{2} (\alpha + \beta) + \frac{1}{2} m_{t-1} (\alpha - \beta) (2p - 1) (2P (\mathcal{H}_{t-2}) - 1)
\]  

(6)

Expressing (6) in terms of informativeness, where \( I_t \equiv 2P (\mathcal{H}_{t-1}) - 1, \) gives:

\[
I_t = \alpha + \beta - 1 + \frac{1}{2} m_{t-1} (\alpha - \beta) (2p - 1) I_{t-1}
\]  

(7)

Using (2), then (7) is equivalent to:

\[
I_t = \alpha + \beta - 1 + F (V \cdot I_{t-1}) (\alpha - \beta) (2p - 1) I_{t-1} \equiv Z (I_{t-1})
\]

**Learning process converges to smallest equilibrium level of informativeness** \( I_{1\ast} \) The sequence \( I_t \) converges to the smallest equilibrium level of informativeness. An increasing sequence converges to its least upper bound. Show 1) that the sequence is increasing and then 2) that the smallest fixed point is its least upper bound.

1) The sequence \( I_t \) is increasing because \( Z (I_t) \) increases in \( I_t \):

For any \( I_t, I_t' \) such that \( I_t' > I_t \)

\[
Z (I_t') - Z (I_t) = (\alpha - \beta) (2p - 1) \left\{ \left[ F (V \cdot I_t') - F (V \cdot I_t) \right] I_t + F (V \cdot I_t') \left[ I_t' - I_t \right] \right\} \geq 0
\]
since \( F \) is positive and increasing and \( V \geq 0 \) and \( I_t \geq 0 \).

Using \( I_0 = 0 \), equation 6, and \( I \in [0, 1] \), then

\[
I_1 = Z(I_0) \geq I_0 = 0
\]

and due to \( Z \) being increasing:

\[
Z(Z(I_{t-1})) = Z(I_t) \geq I_t = Z(I_{t-1})
\]

for all \( I_t \).

2) The smallest upper bound of the sequence \( I_t \) is given by the smallest fixed point \( \bar{I} \), where

\[
\bar{I} \equiv \inf \{ I : Z(I) \leq I \}
\]

If \( \bar{I} \) was not an upper bound then for some \( I' \leq \bar{I} \) it would be true that \( Z(I') > Z(\bar{I}) \). But \( Z \) is increasing and \( Z(\bar{I}) = \bar{I} \), so this is a contradiction. Further, if \( \bar{I} \) was not the least upper bound then for some \( I^{**} \), \( I^{**} < \bar{I} \). An increasing sequence converges to its least upper bound. In addition, the limit of the sequence is a fixed point such that: \( Z(I^{**}) = I^{**} \). But we assumed that \( \bar{I} \) was the smallest fixed point of \( Z \).

By definition 2,

\[
m_t = F(V \cdot I_t)
\]

and \( V \geq 0 \) and \( F \) is increasing. \( m_t \) increases monotonically with \( I_t \). Hence, as \( I_t \) converges so does \( m_t \). Further, the smallest equilibrium level of informativeness \( I^{1*} \) corresponds to the smallest equilibrium level of mobility \( m^{1*} \).