



DEPARTMENT OF ECONOMICS
DISCUSSION PAPER SERIES

**INFERRING SCHOOL QUALITY FROM RANKINGS: THE
IMPACT OF SCHOOL CHOICE**

Claudia Herresthal

Number 747
June 2015

Manor Road Building, Oxford OX1 3UQ

Inferring School Quality from Rankings: The Impact of School Choice

Claudia Herresthal*

June 11, 2015

Abstract

School choice reforms allow families to apply to non-local schools and assign additional funding to schools based on families' demand. For these reforms to promote high-quality schools, families need to infer school quality from past performance, but past performance also depends on student ability. Because reforms alter the allocation of students to schools, it is unclear whether performance becomes more or less informative about quality. I model families as trading off estimated quality against proximity, and analyze a steady-state Bayesian-Nash equilibrium. I show that performance-based rankings become more informative about quality only if oversubscribed schools can choose whom to accept.

*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/clherresthal/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. I benefited from helpful comments and feedback from Pawel Dziewulski, Aytel Erdil, Peter Esö, Li Hao, Johannes Hörner, Emir Kamenica, Fuhito Kojima, Sujoy Mukerji, Collin Raymond, Tomasz Strzalecki, Alex Teytelboym, Adrien Vigier and the participants at BGSE Lunch Theory Workshop, Bonn; 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.

1 Introduction

School choice reforms allow families to apply to schools other than the ones to which they are zoned, and they also assign additional funding to schools based on families' demand for places.¹ In reality, families' demand for places exceeds capacity at many schools. Therefore schools also have a choice over applicants, which is often regulated by admission codes. The aim of these reforms is to assign additional funding to high-quality schools, i.e. schools which help children achieve, so that these schools can expand and cater to more students. To increase the success of these reforms, more families need to apply to high-quality schools.

Families do not observe school quality directly. To allow families to make more informed application choices, governments publish schools' performance on recent standardized tests.² However, past performance also depends on student ability, and hence a ranking of schools based on their past performance is not perfectly informative about their relative quality, i.e. a higher rank is not always assigned to a higher-quality school.³

Given families' informational constraints, school choice reforms could fail to promote high-quality schools if families find it harder to infer school quality from rankings. This may seem plausible considering that these reforms increase the variation in student intakes across schools, and therefore make it harder to attribute any differences in past performance to differences in school quality. Further, one may worry that rankings become less informative if schools are able to select between applicants based on their ability: a

¹In the US, 15% of students do not attend a district-assigned school [Greene et al., 2010], and open enrollment laws, as well as the introduction of magnet and charter schools, have increased the set of publicly-funded schools families can choose from. In England, over half of students do not attend their nearest state school [Burgess et al., 2006], and families can apply to any publicly-funded school, including privately-run free schools.

²In the US, school report cards were first required by the No Child Left Behind Act in 2002, and online searches for information about school quality increased with the introduction of school choice reforms [Lovenheim and Walsh, 2014]. In England, school rankings based on the most recent test outcomes have been published by the media annually since 1992, and are used by just below half of families [Coldron et al., 2008].

³Although statistical estimates of schools' value-added are often made available, these estimates are very imprecise [Leckie and Goldstein, 2009] and often not used by families [Coldron et al., 2008].

school which performed well once will then recruit a strong intake and thus increase its chances to perform well again, irrespective of its quality.

I develop a model to study how school choice reforms influence families' application choices when families are incompletely informed about school quality but observe a ranking of schools based on their latest performance. I find that school choice reforms cause more families to apply to high-quality schools, and make rankings become better indicators of schools' relative quality only if schools have some discretion over whom to accept. My results imply that families are less likely to identify high-quality schools if oversubscribed schools are restricted to allocate places to applicants by lottery.

This paper is the first to study how the informativeness of school rankings depends on the way in which families and schools choose one another.⁴ If we want to rely on market forces 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.

I introduce a model of two schools, one of good and one of bad quality, each of which is located in one of two districts. A priori, families in each generation are uninformed about both their child's ability and schools' qualities, but observe a ranking of schools based on their performance with students in the previous generation. Families apply to a school and incur a randomly drawn cost of transport if their child is accepted at the school in the other district (non-local school). Schools are capacity-constrained. If a school is oversubscribed, it selects based on applicants' ability. Students attend their preferred school if admitted and attend the other school, otherwise. Then student performance is realized; it is stochastically increasing in both school quality and student ability.

⁴Previous theoretical studies of families' application choices have not made explicit the dynamic inference problem associated with performance information (e.g. see De Fraja and Landeras [2006]), 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 (e.g. see Black [1999], Belfield and Levin [2002], Hoxby [2000], Bayer and McMillan [2005], Gibbons et al. [2006]).

In equilibrium, families' optimal application choice is described by a cut-off level for transport costs. In particular, families apply to their non-local school if and only if this school is ranked higher and the realized transport costs lie below this cut-off level. Families' optimal cut-off is based on the expected degree of sorting between students and schools in the previous generation. The more likely it is that strong students in the previous generation attended the good school, the more likely it is that the top-ranked school is of better quality. Hence, their expected gain from attending the top-ranked school is higher. In turn, a higher cut-off in the current generation will increase the expected degree of sorting and hence the informativeness of the observed rankings for future generations.

I study comparative statics on a Bayesian-Nash equilibrium in which families' optimal cut-off level is unchanged across generations. The introduction of school choice reforms, as captured by a downward shift of the distribution of transport costs (in the sense of first order stochastic dominance), makes rankings become more informative and increases families' optimal cut-off level for transport costs.⁵ This is because lower transport costs induce more families to apply to the school which they believe to be the good school, and this increases the likelihood that stronger students attend the better school. Hence, performance becomes a better indicator of quality and future generations are willing to incur higher costs to attend the better-performing school. This demonstrates that families' application choices and schools' admission choices impose an externality on the outcomes of future generations.

More generally, transport costs capture the degree of horizontal differentiation between schools as perceived by families. Schools differ in attributes unrelated to academic quality, and families' preferences over those attributes differ (e.g. some highly value sports grounds, others fine art supplies). All families trade off a schools' expected academic quality against these attributes. Thus, rankings become more informative about schools' relative quality if families perceive schools as less horizontally differentiated.

⁵With school choice reforms, attending a school outside their district does not require families to change their residence and therefore lowers the cost associated with attending a non-local school.

I also show that the informativeness of ranking increases with the extent to which oversubscribed schools can select between applicants, if student ability and school quality are complements. This is what policymakers can influence through the design of schools' admission codes.

The welfare implications of such policy interventions depend on whether student ability and school quality are complements or substitutes. Lower transport costs raise the degree of sorting between students and schools, and therefore they increase (decrease) expected performance in the case of complements (substitutes). Overall welfare effects are ambiguous. If a higher-ranked school accepts a strong non-local student, instead of a weak local student, then this causes the rejected local student to incur transport costs to attend a school of lower expected quality. 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.

While families' choice of schools is my leading example, the insights developed in this paper can also be applied to other context of matching with incomplete information; e.g. for patients to choose between hospitals, data on the outcomes of medical procedures are made available to inform patients about provider quality, and these outcomes depend on both the health of the patient and the quality of the provider.⁶

Formally, the results about learning in this paper are related to the results derived from the study of biased contests [Meyer, 1991]. The model also has some similarities to models in the social learning literature.⁷ My paper speaks to the literature on matching with incomplete information, e.g. Chade et al. [2014]. The assignment problem which arises when schools are

⁶E.g. see Gravelle and Sivey [2010], Gaynor et al. [2013, 2012], Cooper et al. [2011]

⁷E.g. see Bikhchandani et al. [1992]. In this model, agents observe performance indicators for each of their options, which are influenced by how their predecessors chose between options, instead of observing their predecessors' choices directly. My model shares with Lobel et al. [2007] 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.

oversubscribed is the subject of an influential literature on matching algorithms.⁸ 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 they submit to the algorithm.⁹

The paper is organized as follows. Section 2 introduces the model. Section 3 solves for equilibrium steady state, derives comparative statics and welfare results, and analyzes convergence to steady state. Section 4 discusses these findings. Section 5 concludes. All proofs and a table summarizing the notation used can be found in the Appendix.

2 The Model

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. Students differ in ability. In each generation, there is one student of the high-ability type (H) and one student of the low-ability type (L).¹⁰ There is one family with one child (student) 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

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

⁹More generally, this paper focuses how application choices interact across generations, rather than within a given generation, e.g. it abstracts away from peer effects as studied in [Epple and Romano, 1998].

¹⁰Comparative statics and welfare results continue to hold if we assume that each school's quality and that each student's ability are determined by an independent random draw.

school will select the high-ability student with probability $p \in (\frac{1}{2}, 1]$.¹¹

Families attach value $v \in \mathbb{R}^+$ 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).¹² These transport costs are drawn in each generation from a known distribution with continuous cumulative distribution function $F(c)$.

In period t , the expected utility of the family in district i if their child attends school in district j is given by:

$$E(U_t^{i,j}) = vP_t^{i,j}(h) - c_t^{i,j}$$

where $c_t^{i,j} = c_t$ if $i \neq j$ and $c_t^{i,j} = 0$ otherwise and where the probability of a high performance $P_t^{i,j}(h)$ depends on the quality of the school in district j and the ability of the student in district i .

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 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

¹¹Schools' payoffs are not explicitly modeled. 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 perfectly identifying students' relative ability or because schools are restricted in their selection of students by other priorities in their admission code, e.g. distance to school or siblings at the school. I assume $p > \frac{1}{2}$ so that schools select to some extent. If schools had to allocate all places by lottery or according to proximity this would correspond to $p = \frac{1}{2}$. See the Section Discussion for more detail.

¹²These costs are only incurred when a student gets accepted at his non-local school. Applications are costless.

know which school ranked higher in any period prior to $t - 1$.¹³

Further, families do not observe the ability of their own child, but they know that an oversubscribed school selects the high-ability student with probability $p \in (\frac{1}{2}, 1]$. 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.

Families and schools are matched. Then the educational performance of generation t is realized and a school ranking (W_t) is constructed.

Each generation of families corresponds to one period.

Production function - Educational performance can be either high or low.

Monotonicity - The production function, which maps a pair of school quality and student ability to the probability of a high performance, is assumed to be monotonic in the following sense:

- Any type of student is more likely to achieve a high performance at the good school than at the bad school.
- A high-ability student is more likely to achieve a high performance than a low-ability student, at any given type of school.

Denote the difference between the probability that a high-ability (low-ability) student obtains a high performance at the good school and the probability that he obtains a high performance at the bad school by Δ_H (Δ_L). By monotonicity, it is true that $\Delta_H, \Delta_L > 0$.

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

$$\Delta_H - \Delta_L \geq 0$$

and substitutes if

$$\Delta_H - \Delta_L \leq 0.$$

¹³It is time-consuming for families to look up rankings and process the information, and therefore it seems likely that families will not take all available past rankings into account. In particular, we assume families take only the most recent ranking into account because this is the one most easily obtained, e.g. in England this ranking is circulated in the media.

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 probability that the good school comes top of the ranking

- by α if a high-ability student attends the good school and a low-ability student attends the bad school,
- by β if a low-ability student attends the good school and a high-ability student attends the bad school.

The monotonicity assumption implies that $\alpha > \beta$ and $\alpha > 1 - \beta$. For details see the Appendix.

3 Results

3.1 Equilibrium Steady State

Families' optimal application strategies are described by a cut-off level for transport cost: families in generation t apply to their non-local school if and only if this school is the period- $t-1$ winning school and their realization for transport costs lies below their cut-off level.¹⁴ We consider a *steady state* in which this cut-off level is constant across generations. We then focus on those steady states which are consistent with equilibrium behavior. An *equilibrium steady state* is a steady state in which all generations of families behave optimally, taking as given the strategies played by other families.

We define two objects that are helpful for characterizing an equilibrium steady state: mobility and informativeness.

Definition: *Period t -informativeness*, $I_t \in [0, 1]$, is defined as

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

¹⁴In equilibrium, the winning school is more likely to be the good than the bad school. Therefore families would be willing to incur transport costs to attend their non-local school only if this school is the winning school, otherwise they would prefer to save transport costs and apply to their local school.

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.

Definition: 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 period- $t-1$ winning school.

An equilibrium steady state is characterized by a constant mobility level across generations, denoted by m^* .

Proposition 1 [Equilibrium]: An equilibrium steady-state level of mobility is characterized by

$$m^* = F(V \cdot I(m^*)) \quad (1)$$

and the corresponding equilibrium steady-state level of informativeness is given by

$$I(m^*) = \frac{\alpha + \beta - 1}{(1 - m^*(\alpha - \beta)(2p - 1))}. \quad (2)$$

Such an equilibrium level of mobility (and informativeness) always exists.

To solve for equilibrium steady state, consider a steady state characterized by mobility level $\hat{m}$ and then find the optimal cut-off of families in any given generation t . In equilibrium steady state, generation t 's optimal strategy is consistent with remaining in steady state: $m_t = \hat{m}$.

The optimal cut-off level of families in generation t is equal to their expected gain from attending the non-local winning school rather than the local losing school. It can be decomposed into i) the likelihood that the winning school is the good rather than the bad school (period- t informativeness) and ii) their expected gain from attending the winning school conditional on the winning school being the good school.

If families in generation t conjecture that the period- $t-1$ ranking was generated in a steady state characterized by mobility level $\hat{m}$ then their

period- t informativeness is given by

$$I_t(\hat{m}, p, \alpha, \beta) = \frac{\alpha + \beta - 1}{(1 - \hat{m}(\alpha - \beta)(2p - 1))}. \quad (3)$$

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

Families' gain from attending the good school rather than the bad school depends on their child's ability. They draw inferences about this ability conditional on their child having been selected out of two applicants, because if a family applies to the non-local winning school then this school will be oversubscribed in equilibrium. Hence, the expected gain from attending the non-local winning school conditional on this school being the good school, denoted by $V \in \mathbb{R}^+$, is given by

$$V(p, v, \Delta_H, \Delta_L) \equiv v((1 - p)(\Delta_L) + p(\Delta_H)) \quad (4)$$

where v is families' valuation of high performance, p is the probability that an oversubscribed school selects the high-ability student and Δ_H (Δ_L) is the increase in expected performance for the high-ability (low-ability) student defined by the production function.

Consequently, the family whose non-local school is the period- $t-1$ winning school is willing to incur their realized transport costs if and only if they lie below their optimal cut-off level, given by $V(p, v, \Delta_H, \Delta_L) \cdot I_t(\hat{m}, p, \alpha, \beta)$, henceforth $V \cdot I_t$. Hence, their mobility, m_t , is given by:

$$m_t = F(V \cdot I_t). \quad (5)$$

The mobility level that characterizes optimal application choices by generation t , m_t , can be expressed as a best response to the conjectured mobility of past generations, $\hat{m}$:

$$m_t(\hat{m}) = F(V \cdot I_t(\hat{m})) \quad (6)$$

In equilibrium steady state, generation t 's optimal strategy is consistent with remaining in steady state: $m_t = \hat{m}$. Such an equilibrium steady-state

level of mobility (and informativeness) always exists because mobility is a weakly increasing function of conjectured mobility and mobility is bounded above by 1. However, the equilibrium steady-state level is not necessarily unique.

3.2 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*}). The reason is the following: 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 equilibrium steady-state level of mobility (henceforth equilibrium level of mobility). Further detail will be given in the section on Convergence (see section 3.4).

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 of transport costs, $F(\cdot)$,*
- (b) *an increase in families' valuation of high performance, v ,*
- (c) *an increase in the influence of school quality on performance,*
- (d) *an increase in the probability, p , that an oversubscribed school selects a high-ability student if school quality and student ability are complements, i.e. if $\Delta_H - \Delta_L \geq 0$.*

If lower transport costs become more likely, in the sense that the distribution $F(\cdot)$ is subject to a negative shift in FOSD, then mobility increases in equilibrium steady state. Suppose we fix families' posterior beliefs about school quality at their initial level, I , then this change causes their 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 and

informativeness have to increase.

Intuitively, lower transport costs have both a direct and an indirect effect on mobility. The direct effect is an increase in the likelihood that transport costs fall below families' optimal cut-off level, holding their optimal cut-off level fixed. The indirect effect is that informativeness and hence the expected gain from attending the winning school increase, causing families' optimal cut-off level to increase. This indirect effect is easier to understand if we consider how lower transport costs in generation $t - 1$ affect period- t informativeness. As more families in generation $t - 1$ apply to the winning school, this 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 an equilibrium steady state, both effects get absorbed into one increase in mobility.

If we were to look at the sequence of mobility levels in the process of converging to their equilibrium steady-state level, it would be possible to highlight these different effects more clearly. In particular, even if a reduction in transport costs is experienced only by one generation then all subsequent generations will have a higher level of informativeness.

Such effects exist for other exogenous changes which cause more families to apply to the winning school. These changes include i) an increase in families' valuation of high performance and ii) a higher impact of school quality on performance, both of which increase families' expected benefit of attending the good school.

An increase in schools' capability to select also causes more families to apply to the winning school if student ability and school quality are complements. To illustrate this, consider families in generation t . In the case of complements, there are two effects which reinforce one another. There is the direct effect of higher sorting between student ability and school quality,

which arises because the winning school is more likely to recruit a stronger intake. In a steady state with a higher expected degree of sorting, generation t observes a more informative ranking. In addition, given the complementarity of ability and quality, students in generation t have a higher expected benefit from attending the good school conditional on being selected out of two applicants. If ability and quality are substitutes, increasing the likelihood with which schools can select the more able student has ambiguous effects on families' application choices. While generation t observes a more informative ranking, their expected benefit from attending the good school falls. Hence, the equilibrium steady-state level of informativeness may fall, causing families to apply to the winning school less often.

3.3 Performance and Welfare

This section studies the effects of exogenous changes on the expected educational performance and total welfare in equilibrium steady state. We take *expected educational performance* to be the ex-ante expected number of high educational performances in any given generation.

Proposition 3 [Performance]: *(1a) A negative shift in FOSD of the distribution of transport costs, $F(\cdot)$, and (1b) an increase in families' valuation of high performance, v , both raise (lower) ex-ante expected total performance if student ability and school quality are complements (substitutes), i.e. if $\Delta_H - \Delta_L \geq 0$ ($\Delta_H - \Delta_L \leq 0$).*

(2) An increase in the probability, p , that an oversubscribed school selects a high-ability student raises the ex-ante expected performance if student ability and school quality are complements, i.e. if $\Delta_H - \Delta_L \geq 0$.

Any exogenous change that increases the degree of sorting between student ability and school quality will increase (decrease) expected performance if school quality and student ability are complements (substitutes). This is because a higher degree of sorting leads to a performance gain for the high-ability student and a performance loss for the low-ability student. The gain

will exceed (fall short of) the loss if student ability and school quality are complements (substitutes).

If student ability and school quality are complements, then the degree of sorting increases with a negative shift in FOSD of the distribution of transport costs, $F(\cdot)$, with a rise in families' valuation of high performance, v , and with a rise in the probability, p , that an oversubscribed school selects a high-ability student. If school quality and student ability are substitutes, lower transport costs and a higher valuation of high performance induce families to apply to the winning school less frequently, causing the degree of sorting to decrease. A rise in schools' capability to select has ambiguous effects on how frequently families apply to the winning school.

An exogenous change that causes expected performance to increase does not necessarily increase welfare. This may arise in the case of complements if the higher inequality in students' performance has a sufficiently negative impact on welfare such that it outweighs the gains from higher total performance. However, welfare does not necessarily increase even if evaluated from a utilitarian social planner's point of view, i.e. if we take *total welfare* to be the sum of both families' ex-ante expected utility in any given generation.¹⁵

The reason is the following: a family will only be willing to incur their transport costs if their expected gain from attending the winning school in the other district exceeds these 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. Therefore the negative welfare impact of the increase in total expenditure by both families may outweigh the welfare gain derived from higher expected performance.¹⁶

For a similar reason, a change that lowers expected performance does not necessarily lower welfare. Suppose student ability and school quality are

¹⁵Detailed examples can be found in the Appendix.

¹⁶This would be 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.

substitutes and there is a negative shift in FOSD of the distribution of transport costs. Further, assume that families are already perfectly mobile (i.e. apply to the observed winner for all realizations of transport costs). In this situation, the degree of sorting between students and schools is unaffected by the reduction in transport costs, and hence expected performance is unaffected. However, the expected expenditure on transport costs decreases and so welfare increases.

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 modeled. 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, and close down those schools that are identified as bad. This would increase overall school quality. 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 more sensitive to their quality. Consequently, we may expect that supply-side responses cause additional benefits by improving overall school quality. With these benefits accounted for, welfare is more likely to increase with a policy intervention such as lower transport costs.

3.4 Convergence

This section analyzes how the informativeness of school rankings evolves starting from the first ranking, given each generation of families responds optimally to the mobility levels of all previous generations. The sequence of rankings corresponds to a sequence of levels of informativeness, which is increasing and converges over time. Its limit is the smallest level of informativeness consistent with equilibrium steady state.

Definition *The dynamic learning process is defined by the sequence $\{I_t\}_{t=0,1,\dots}$, where*

$$I_0 = 0 \tag{7}$$

and

$$I_t = \alpha + \beta - 1 + F(V \cdot I_{t-1})(\alpha - \beta)(2p - 1)I_{t-1} \quad (8)$$

for all $t \geq 1$.

Proposition 4 [Convergence] *The dynamic learning process is characterized by the informativeness of the first generation with access to rankings (7) and by the recurrence equation linking informativeness of subsequent generations (8). Its limit as $t \rightarrow \infty$, denoted by I^{1*} , is the smallest equilibrium level of informativeness. As $\{I_t\}_{t=0,1,\dots}$ converges so does the sequence of mobility levels $\{m_t\}_{t=0,1,\dots}$. The sequence $\{m_t\}_{t=0,1,\dots}$ converges to the smallest equilibrium level of mobility, denoted by m^{1*} , where*

$$m^{1*} = F(V \cdot I^{1*}). \quad (9)$$

Starting from any level of informativeness, the sequence of levels of informativeness that arises from families' optimal application choices is increasing and will converge to an equilibrium level of informativeness. Starting from a symmetric prior across schools, i.e. a level of informativeness equal to zero, such a sequence will converge to the smallest equilibrium level of informativeness. As equilibrium levels of informativeness and mobility are jointly ordered, this also implies that the sequence of mobility levels will converge to the smallest equilibrium level of mobility.¹⁷

Studying how informativeness evolves over time gives further insights into why informativeness in steady state depends positively on conjectured mobility (see equation (3)). As an example, compare the inference made by the first two generations. Generation 1 is the first generation with access to rankings and conjectures correctly that in the previous generation each

¹⁷Note that even if families do not respond optimally to past generations' mobility, but simply apply to the winning school with a fixed probability, then the sequence of levels of informativeness converges. However, it may not converge to an equilibrium steady-state level.

student attended his local school. Then generation 1 believes that the period-0 winning school is the good school, denoted by $\overline{W}_0$, with probability,

$$P(\overline{W}_0) = \frac{1}{2}(\alpha + \beta) > \frac{1}{2},$$

and hence,

$$I_1 = \alpha + \beta - 1.$$

Note that α and β characterize the intrinsic quality advantage the good school has over the bad school. The higher are α or β , the more informative is the ranking that results from a random assignment of students to schools.

Generation 2 observes the period-1 winning school, but not the period-0 winning school. Given the conjecture 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 period-0 winning school. Conditional on the good (bad) school being the period-0 winning school, denoted by $\overline{W}_0$ ($\underline{W}_0$), generation 2 can then infer the probability of the event that the good (bad) school is the period-1 winning school, denoted by $\overline{W}_1$ ($\underline{W}_1$).¹⁸

The higher is $\hat{m}_1$, the more likely the period-0 winning school will also be the period-1 winning school. This is because the period-0 winning school 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 than the bad school of winning in period 1, conditional on having won in period 0:

$$P(\overline{W}_1|\overline{W}_0) > P(\underline{W}_1|\underline{W}_0).$$

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

¹⁸Schools' 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 would have to randomize between applications, i.e. $p = \frac{1}{2}$.

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, but the good school has a higher absolute probability of winning for any given increase in conjectured mobility.

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

$$\begin{aligned} P(\overline{W}_1) &= P(\overline{W}_1|\overline{W}_0) P(\overline{W}_0) + P(\overline{W}_1|W_0) P(W_0) \\ &= \frac{(\alpha + \beta)}{2} + \hat{m}_1 \frac{1}{2} (2p - 1) (\alpha - \beta) (\alpha + \beta - 1) \end{aligned}$$

and hence,

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

The good school is more likely to be the period-0 winner than the bad school, and an increase in conjectured mobility makes it more likely that period-0 winner is also the period-1 winner. Hence, a higher conjectured mobility of generation 1 increases the unconditional probability that the good school is the period-1 winner.

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 equilibrium steady-state level.

4 Discussion

This section discusses how policy changes can improve what families infer about schools' relative quality from the latest performance-based ranking. It is important that rankings identify schools which add value to children's education, so that families apply to these schools, and so that these schools are expanded. 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.

4.1 The distribution of transport costs and its implications for policy

We found that lowering transport costs makes it easier for families to identify good schools based on their recent performance if oversubscribed schools select applicants based on their ability. A reduction in transport costs causes families to attach more importance to schools' estimated quality relative to schools' location when deciding where to apply.

School choice - School choice reforms lower families' costs associated with sending their child to a non-local school. Families are no longer required to move to this school's attendance area (or school district); instead they can simply commute. As is the case with a reduction in transport costs, this induces families to attach relatively less importance to schools' location.

Common standards in schools - The cost realization, c , can more generally be interpreted as a cost of choosing one school over another if both were of the same quality, i.e. it can be a measure of the horizontal differentiation of schools as perceived by families. Such costs can arise if schools vary in attributes unrelated to academic quality, e.g. in terms of their sports facilities or their curricula. Families differ in their preferences over these attributes and for these reasons would perceive different schools as more desirable (or more costly). By enforcing common standards in dimensions unrelated to quality, policymakers can reduce the horizontal differentiation of schools as perceived by families and thereby cause families to place relatively more weight on estimated quality (vertical differentiation).

Easily accessible rankings - The cost realization, c , can also be thought of as the time and effort incurred by families to research how schools compare in the latest performance-based ranking. Families weigh up this cost against the benefit from observing which school is ranked higher, given the informativeness of the ranking. By making rankings more easily accessible, policymakers can induce more families to look up relative performance and cause more families to apply to higher-ranked schools. That access to performance information can increase applications to well-performing schools has been shown by Hastings and Weinstein [2007].

Student incentive-compatible assignment mechanisms - This paper has assumed a simple application procedure: an oversubscribed school selects between applicants and the unsuccessful applicant is assigned to the other school. With this application procedure, it is optimal for any student to apply to the school he most prefers. In reality, application procedures are more complicated, and families may find it optimal to express a preference for a school to which their child is likely to be admitted (safe school) rather than express a preference for the one with the highest perceived quality.¹⁹ The cost of applying to the school with the highest perceived quality (the “riskier” school) can be captured by the cost realization, c . By using an assignment mechanism that is student incentive-compatible, policymakers lower this cost and hence increase the number of applications received by high-quality schools.

4.2 Schools’ capability to select and its implications for policy

This paper yields insights on how regulating the selection by oversubscribed schools affects the informativeness of performance-based rankings.

No lottery - If students were admitted based on a lottery among applicants, i.e. $p = \frac{1}{2}$, then the informativeness of rankings is lower than if over-

¹⁹Calsamiglia 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 each schools. In particular, they found that many families switched their first choice school after neighborhood boundaries, and hence priorities, were reassigned.

subscribed schools were to some extent able to select among applicants based on their ability, i.e. if $p > \frac{1}{2}$. This refutes the conjecture that admission by lottery would make it easier to identify good schools from performance in a setting in which families have limited observations of past performance.²⁰

What is more, if an admission lottery is used, then the policies described in section 4.1 are ineffective.²¹ The reason is that these policies cause high-ranked schools to receive a larger number of applications, and for this to be an advantage for recruiting a strong intake, oversubscribed schools need to be able to select.²²

Fewer priorities in admission codes - By reducing regulation on schools' admission, i.e. raising p , policymakers can improve the informativeness of rankings, if school quality and student ability are complements. This demonstrates a disadvantage from assigning priorities to students based on criteria such as their residence. Debates surrounding the design of school admission codes has so far not paid sufficient attention to how its design influences the information families have about schools and their application choices.²³

4.3 Student allocation as a performance advantage and its implications for policy

The interventions discussed so far have all targeted the allocation of students to schools. In particular, we have shown that families are more likely to identify high-quality schools if schools with better performance take on more able students in the next generation. This identification relies on the feature

²⁰This setting seems realistic given that it is time-consuming for families to look up rankings and process the information.

²¹This implication arises because families with high- and low-ability children are equally likely to apply to the school with higher rank. If these policies affect families with high-ability children more than families with low-ability children, then these policies would remain effective even if an admissions lottery was used, but they would be less effective than if schools could select among applicants.

²²A similar argument is made by Gavazza and Lizzeri [2009], but their argument focuses on what this implies for schools' incentives to improve.

²³Hatfield et al. [2011] analyze which assignment mechanisms ensure that a school is assigned a set of students that is weakly better for that school whenever it is more preferred by students. Schools' preferences over students are based on admission priorities rather than on how much these students contribute to performance.

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.

Performance-based rewards - Policymakers can create other channels through which better-performing schools get an advantage for future performance; e.g. they could assign additional funding to schools with better performance. By contrast, assigning funding to those schools with low performance compromises the identification of good schools in the long run.

4.4 Application to health care

The framework developed in this paper also provides useful insights for health economics. In the context of health care, patients are matched with hospitals and the success of treatment depends on both the condition of the patient and the quality of the hospital. Patients can choose between hospitals and for many procedures patients can also access performance indicators, such as mortality rates.²⁴ 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.

There is evidence that sicker patients benefit more from seeking out a good hospital (substitutes between hospital quality and patient health) and sicker patients respond more strongly to indicators of hospital quality.²⁵ Evidence for selection by hospitals is limited, but Dranove et al. [2003] find that the introduction of cardiac surgery report cards led to elective surgery being carried out on less severely-ill patients.²⁶

If there is no selection by hospitals, this would suggest that sicker patients attend better hospitals, which would correspond to the case where $p < \frac{1}{2}$.

²⁴In the UK, a policy in 2006 allowed patients who needed specialized treatment to choose between hospitals. In 2007 this was complemented by a website which supplied information on hospital quality, including mortality rates.

²⁵e.g. see Cutler et al. [2004], Gaynor et al. [2012], Cooper et al. [2011]

²⁶Mortality rates are usually adjusted for observable risk factors, but providers may still have reasons to select patients based on their health status: they may have additional information about patients' health or may be risk-averse [Dranove et al., 2003].

Given substitutes between patient health and hospital quality, our model would predict that any reduction in transport costs (or increase in patient choice) would make mortality rates less informative about hospital quality. However, it would increase overall expected patient outcomes. Therefore, a trade-off exists between the short-term gains from better overall patient outcomes and the long-term gains from channeling funding towards good hospitals, so they can expand and treat more patients.

If hospitals select between patients based on their health status then we would expect sicker patients to be matched with worse hospitals, which would correspond to the case where $p > \frac{1}{2}$. Our model would predict that any reduction in transport costs (or increase in patient choice) would make mortality rates more informative about hospital quality, but decrease overall expected patient outcomes. This matches findings by Dranove et al. [2003]: the introduction of report cards lowered the cost of researching information on quality, which corresponds to a reduction in transport costs in the framework of this paper, and this led to small gains for healthy patients and large deterioration of outcomes for sick patients in the short term.

5 Conclusion

This paper studies the impact of school choice reforms on families' application choices when families infer school quality from performance-based rankings and trade off estimated school quality against proximity. I find that the introduction of school choice reforms causes rankings to become more informative only if oversubscribed schools can to some extent choose whom to accept. Despite each generation of families being equally uninformed ex ante, school choice reforms cause more families to apply to high-quality schools over time. My findings are important in light of the recent debates on both school choice reforms and the design of admission procedures.

The analysis of this dynamic learning process about school quality is made tractable by using the concept of equilibrium steady state, i.e. families' optimal strategy is unchanged across generations. Comparative statics show that an increase in 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.

A motivation for implementing school choice reforms is the idea that these reforms will induce schools to compete for applicants. My findings are an important first step in better understanding the quasi-market forces generated by the implementation of school choice reforms, because they explore the link between families' demand for places and school quality. I view the framework developed as a building block for future research about schools' investments. In particular, if schools are allowed to invest in their quality, we can explore how schools' incentives to improve their performance interact with families' beliefs and families' application strategies in equilibrium steady state. Such analysis could yield further insights into whether school choice reforms are a policy instrument that works as a "tide to lift all boats", i.e. whether they cause each type of school to invest in the quality of their provision. This links in with the industrial organization literature on how greater competition affects managers' incentives for effort (e.g. Scharfstein [1988], Hart [1983]).

References

- A. Abdulkadiroglu and T. Sönmez. School choice: A Mechanism Design Approach. *The American Economic Review*, 93(3):729–747, 2003.
- R. Allen, S. Burgess, and L. McKenna. School Performance and Parental Choice of School: Secondary Data Analysis. 2014.
- P. Bayer and R. McMillan. Choice and Competition in Local Education Markets. Working Paper 11802, National Bureau of Economic Research, November 2005.
- C. R. Belfield and H. M. Levin. The Effects of Competition between Schools

- on Educational Outcomes: A Review for the United States. *Review of Educational Research*, 72(2):279–341, 2002.
- S. Bikhchandani, D. Hirshleifer, and I. Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, pages 992–1026, 1992.
- S. E. Black. Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, 114(2):pp. 577–599, 1999.
- S. Burgess, A. Briggs, B. McConnell, and H. Slater. *School choice in England: Background facts*. Centre for Market and Public Organisation, University of Bristol Bristol, 2006.
- S. Burgess, E. Greaves, A. Vignoles, and D. Wilson. What Parents Want: School Preferences and School Choice. *The Economic Journal*, 2014.
- S. Callander and J. Hörner. The Wisdom of the Minority. *Journal of Economic Theory*, 144(4):1421–1439, 2009.
- C. Calsamiglia and M. Güell. The Illusion of School Choice: Empirical Evidence from Barcelona. 2014.
- H. Chade, G. Lewis, and L. Smith. Student Portfolios and the College Admissions Problem. *The Review of Economic Studies*, 81(3):971–1002, 2014.
- J. Coldron, E. Tanner, S. Finch, L. Shipton, C. Wolstenholme, B. Willis, S. Demack, and B. Stiell. Secondary School Admissions. *DCSF research report*, 2008.
- Z. Cooper, S. Gibbons, S. Jones, and A. McGuire. Does Hospital Competition Save Lives? Evidence From The English NHS Patient Choice Reforms. *The Economic Journal*, 121(554):F228–F260, 2011.

- D. M. Cutler, R. S. Ickman, and M. B. Landrum. The Role of Information in Medical Markets: An Analysis of Publicly Reported Outcomes in Cardiac Surgery. *American Economic Review*, 94(2):342–346, 2004.
- G. De Fraja and P. Landeras. Could do better: The Effectiveness of Incentives and Competition in Schools. *Journal of Public Economics*, 90(1):189–213, 2006.
- D. Dranove, D. Kessler, M. McClellan, and M. Satterthwaite. Is More Information Better? The Effects of "Report Cards" on Health Care Providers. *Journal of Political Economy*, 111(3):555–588, June 2003.
- D. Epple and R. E. Romano. Competition between Private and Public Schools, Vouchers, and Peer-Group Effects. *The American Economic Review*, 88(1):pp. 33–62, 1998.
- A. Erdil and H. Ergin. What's the matter with Tie-breaking? Improving Efficiency in School Choice. *The American Economic Review*, 98(3):669–689, 2008.
- A. Gavazza and A. Lizzeri. Transparency and Economic Policy. *Review of Economic Studies*, 76(3):1023–1048, 2009.
- M. Gaynor, C. Propper, and S. Seiler. Free to choose? Reform and Demand Response in the English National Health Service. Technical report, National Bureau of Economic Research, 2012.
- M. Gaynor, R. Moreno-Serra, and C. Propper. Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service. *American Economic Journal: Economic Policy*, 5(4):134–66, 2013.
- S. Gibbons, S. Machin, and O. Silva. Choice, Competition and Pupil Achievement. IZA Discussion Papers 2214, Institute for the Study of Labor (IZA), July 2006.
- H. Gravelle and P. Sivey. Imperfect Information in a Quality-competitive Hospital Market. *Journal of Health Economics*, 29(4):524–535, 2010.

- J. Greene, T. Loveless, W B. MacLeod, T. Nechyba, P. Peterson, M. Rosenthal, and G. Whitehurst. Expanding Choice in Elementary and Secondary Education. *A Report on Rethinking the Federal Role in Education, Brown Center on Education Policy at Brookings*, 2, 2010.
- O. D. Hart. The Market Mechanism as an Incentive Scheme. *The Bell Journal of Economics*, pages 366–382, 1983.
- J. S. Hastings and J. M. Weinstein. Information, School Choice, and Academic Achievement: Evidence from Two Experiments. NBER Working Papers 13623, National Bureau of Economic Research, Inc, November 2007.
- J.W. Hatfield, F. Kojima, and Y. Narita. Promoting School Competition Through School Choice: A Market Design Approach. Working Papers 2011-018, Human Capital and Economic Opportunity Working Group, September 2011.
- C. M. Hoxby. Does Competition among Public Schools Benefit Students and Taxpayers? *American Economic Review*, 90(5):1209–1238, 2000.
- D. P. Kessler and M. B. McClellan. Is Hospital Competition Socially Wasteful? *The Quarterly Journal of Economics*, 115(2):577–615, 2000.
- G. Leckie and H. Goldstein. The Limitations of using School League Tables to Inform School Choice. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(4):835–851, 2009.
- I. Lobel, D. Acemoglu, M. Dahleh, and A. Ozdaglar. Preliminary Results on Social Learning with Partial Observations. In *Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools*, 2007.
- M. F. Lovenheim and P. Walsh. Does Choice Increase Information? Evidence from Online School Search Behavior. 2014.

- W. B. MacLeod and M. Urquiola. Anti-Lemons: School Reputation and Educational Quality. Working Paper 15112, National Bureau of Economic Research, June 2009.
- M. A. Meyer. Learning from Coarse Information: Biased Contests and Career Profiles. *The Review of Economic Studies*, 58(1):15–41, 1991.
- P. Milgrom and J. Roberts. Comparing Equilibria. *The American Economic Review*, pages 441–459, 1994.
- D. Scharfstein. Product-market Competition and Managerial Slack. *The RAND Journal of Economics*, pages 147–155, 1988.

A Appendix

A.1 Notation

G, B	school quality (good, bad)
H, L	student ability (high, low)
h, l	performance outcome (high, low)
$v > 0$	families' valuation of high performance outcome
c_t ; cdf $F(c)$	transport costs realization in t : drawn from stationary cdf $F(c)$
$p \in (\frac{1}{2}, 1]$	probability that a school with two applicants selects high-ability student
$\bar{H} (\underline{H})$	high-ability student at good (bad) school
$W_t : \bar{W}_t, \underline{W}_t$	school which won with students in generation t (period- t winner): good, bad school
$\bar{a} = P(\bar{h} \bar{H})$	probability that high-ability student achieves h at good school
$1 - \underline{a} = P(h \underline{H})$	probability that high-ability student achieves h at bad school
$\bar{b} = P(\bar{h} \bar{L})$	probability that low-ability student achieves h at good school
$1 - \underline{b} = P(h \underline{L})$	probability that low-ability student achieves h at bad school
$\alpha = \frac{\bar{a} + \bar{b}}{2} = P(\bar{W} \bar{H})$	probability good school wins with high-ability student
$\beta = \frac{\underline{a} + \underline{b}}{2} = P(\bar{W} \underline{H})$	probability good school wins with low-ability student
$\Delta_H = \bar{a} - (1 - \underline{a})$	difference in probability of high performance outcome at good versus bad school for high-ability student
$\Delta_L = \bar{b} - (1 - \underline{b})$	difference in probability of high performance outcome at good versus bad school for low-ability student
$\Delta_H - \Delta_L \geq (\leq) 0$	student ability and school quality are complements (substitutes)
$I_t \in [0, 1]$	period- t informativeness
$I^{1*} \in [0, 1]$	smallest equilibrium steady-state level of informativeness
$m_t \in [0, 1]$	mobility of generation t
$m^{1*} \in [0, 1]$	smallest equilibrium steady-state level of mobility
$V(v, p, \Delta_H, \Delta_L) > 0$	expected gain from attending the non-local winning school conditional on the winning school being the good school
$H_t^{W_{t-1}}$	period- $t-1$ winner takes on high-ability student in generation t
$H_t^{(non)localW_{t-1}}$	period- $t-1$ winner is (non-)local school for high-ability student in generation t

A.2 Production Function

The production function maps a pair of student ability and school quality to a probability of a high performance. We introduce the following notation.

Let an overline (underline) denote association with the good (bad) school, e.g. let $\bar{h}$ ($\underline{h}$) denote the event that the student at the good (bad) school gets a high performance outcome. Then define the production function by the parameters $\bar{a}$, $\underline{a}$, $\bar{b}$ and $\underline{b}$ as follows:

- The probability of high performance for a *high-ability* student
 - at the *good* school is given by $P(\bar{h}|\bar{H}) = \bar{a}$
 - at the *bad* school is given by $P(\underline{h}|\underline{H}) = 1 - \underline{a}$
- The probability of high performance for the *low-ability* student
 - at the *good* school is given by $P(\bar{h}|\bar{L}) = \bar{b}$
 - at the *bad* school is given by $P(\underline{h}|\underline{L}) = 1 - \underline{b}$

where $\bar{a}, \underline{a}, \bar{b}, \underline{b} \in [0, 1]$.

By monotonicity,

$$\Delta_H \equiv P(\bar{h}|\bar{H}) - P(\underline{h}|\underline{H}) = \bar{a} - (1 - \underline{a}) > 0 \quad (10)$$

$$\Delta_L \equiv P(\bar{h}|\bar{L}) - P(\underline{h}|\underline{L}) = \bar{b} - (1 - \underline{b}) > 0 \quad (11)$$

$$P(\bar{h}|\bar{H}) - P(\bar{h}|\bar{L}) = \bar{a} - \bar{b} > 0 \quad (12)$$

$$P(\underline{h}|\underline{H}) - P(\underline{h}|\underline{L}) = 1 - \underline{a} - (1 - \underline{b}) > 0. \quad (13)$$

Further, the probability that the good school ranks top is given by the following expressions:

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

$$P(\bar{W}|\bar{H}, \underline{L}) \equiv P(\bar{W}|\bar{H}) = \bar{a}\underline{b} + \frac{1}{2}(\bar{a}(1 - \underline{b}) + (1 - \bar{a})\underline{b}) \equiv \alpha$$

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

$$P(\overline{W}|\underline{H}, \overline{L}) \equiv P(\overline{W}|\underline{H}) = \underline{a}\overline{b} + \frac{1}{2}((1 - \underline{a})\overline{b} + \underline{a}(1 - \overline{b})) \equiv \beta.$$

By (10) and (11), it follows that

$$\begin{aligned} \overline{a} - (1 - \underline{a}) &> 0 > 1 - \underline{b} - \overline{b} \Leftrightarrow \\ \frac{\overline{a} + \underline{b}}{2} &> \frac{2 - \underline{a} - \overline{b}}{2} \Leftrightarrow \\ \alpha &> 1 - \beta. \end{aligned}$$

By (12) and (13), it follows that

$$\begin{aligned} \overline{a} - \overline{b} &> 0 > \underline{a} - \underline{b} \\ \frac{\overline{a} + \underline{b}}{2} &> \frac{\underline{a} + \overline{b}}{2} \\ \alpha &> \beta. \end{aligned}$$

A.3 Proof Proposition 1 [Equilibrium]

Proposition 1 [Equilibrium]: *An equilibrium steady-state level of mobility is characterized by*

$$m^* = F(V \cdot I(m^*)) \quad (1)$$

and the corresponding equilibrium steady-state level of informativeness is given by

$$I(m^*) = \frac{\alpha + \beta - 1}{(1 - m^*(\alpha - \beta)(2p - 1))}. \quad (2)$$

Such an equilibrium level of mobility (and informativeness) always exists.

A.3.1 Deriving steady-state informativeness I in terms of mobility m

$$I_t(\widehat{m}, p, \alpha, \beta) = \frac{\alpha + \beta - 1}{(1 - \widehat{m}(\alpha - \beta)(2p - 1))} \quad (3)$$

Denote the event that the good school wins by $\overline{W}$ and the event that the bad school wins by $\underline{W}$. Define the state of the system realized in period t by $\overline{W}_t$ and $\underline{W}_t$ respectively.

We will first derive the probability that the high-ability student in generation t attends the period- $t-1$ winning school. The probability is a function of mobility level, m , and schools' capability to select, p . The period- $t-1$ winning school is equally likely to be the high-ability student's local or non-local school (events denoted by $H_t^{localW_{t-1}}$ and $H_t^{non-localW_{t-1}}$ respectively). With probability $1 - m$ each student applies to his local school and gets accepted. With probability m both students apply to the period- $t-1$ winning school 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

$$\begin{aligned} P\left(H_t^{W_{t-1}}\right) &= P\left(H_t^{W_{t-1}}|H_t^{localW_{t-1}}\right)P\left(H_t^{localW_{t-1}}\right) \\ &\quad + P\left(H_t^{W_{t-1}}|H_t^{non-localW_{t-1}}\right)P\left(H_t^{non-localW_{t-1}}\right) \\ &= \frac{1}{2}(1 - m + 2mp). \end{aligned}$$

Next, we will derive the probability that the good (bad) school is the period- t winning school, conditional on being the period- $t-1$ winning school. These transition probabilities between events $\overline{W}$ and $\underline{W}$ are functions of the mobility level, m , schools' capability to select, p , and the production function. In particular, the production function defines the probability that the good school wins with the high-ability student (against the bad school with the low-ability student), denoted by $P(\overline{W}_t|\overline{H}_t) = \alpha$, and the probability that the good school wins with the low-ability student (against the bad school with the high ability student), denoted by $P(\overline{W}_t|\underline{L}_t) = \beta$.

Let the transition probabilities from state j to state i be denoted by

$P(i|j)$:

$$\begin{aligned} P(\overline{W}_t|\overline{W}_{t-1}) &= P(\overline{W}_t|\overline{H}_t) P(H_t^{W_{t-1}}) + P(\overline{W}_t|\overline{L}_t) P(L_t^{W_{t-1}}) \\ &= \alpha \left(\frac{1}{2} (1 - m + 2mp) \right) + \beta \left(\frac{1}{2} (1 + m - 2mp) \right) \end{aligned}$$

and

$$\begin{aligned} P(W_t|W_{t-1}) &= (1 - P(\overline{W}_t|\overline{L}_t)) P(H_t^{W_{t-1}}) \\ &\quad + (1 - P(\overline{W}_t|\overline{H}_t)) P(L_t^{W_{t-1}}) \\ &= (1 - \beta) \left(\frac{1}{2} (1 - m + 2mp) \right) \\ &\quad + (1 - \alpha) \left(\frac{1}{2} (1 + m - 2mp) \right). \end{aligned}$$

The transition matrix, T , is given by:

$$T = \begin{pmatrix} P(\overline{W}_t|\overline{W}_{t-1}) & P(W_t|\overline{W}_{t-1}) \\ P(\overline{W}_t|W_{t-1}) & P(W_t|W_{t-1}) \end{pmatrix}.$$

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

$$\begin{pmatrix} P(\overline{W}) & P(W) \end{pmatrix} = \begin{pmatrix} P(\overline{W}) & P(W) \end{pmatrix} T$$

and

$$P(\overline{W}) + P(W) = 1.$$

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

$$\lim_{k \rightarrow \infty} T^k = \mathbf{1} \begin{pmatrix} P(\overline{W}) & P(W) \end{pmatrix}$$

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 characterized by

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

The Markov chain is not irreducible if and only if $\alpha = 1$ and $m = 1$ and $p = 1$.²⁷ In this case, the state $\overline{W}$ is an absorbing state and I assign $P(\overline{W}) = 1$. So

$$P(\overline{W}) = \frac{\alpha + \beta - (\alpha - \beta)m(2p - 1)}{2(1 - m(\alpha - \beta)(2p - 1))} \quad (14)$$

for all parameter configurations.

Families update their beliefs that the winning school is the good school, denoted by $P(G|W)$:

$$P(G|W) = \frac{P(\overline{W})P(G)}{P(\overline{W})P(G) + P(\underline{W})P(B)}$$

where $P(G)$ denotes the prior belief that a given school is of good quality.

Families' updated beliefs are based on their prior beliefs about school quality, which are symmetric with respect to school districts. Further, families' updated beliefs depend on the stationary probability that the good school is the winning school in steady state, denoted by $P(\overline{W})$ and given by (14).

By symmetry of prior beliefs, $P(G) = \frac{1}{2}$,

$$P(G|W) = \frac{P(\overline{W})}{P(\overline{W}) + P(\underline{W})} = P(\overline{W}).$$

By definition of informativeness,

$$I(m, p, \alpha, \beta) \equiv P(G|W) - P(B|W) = \frac{\alpha + \beta - 1}{(1 - m(\alpha - \beta)(2p - 1))}.$$

²⁷Order of limits matters, first let p tend to 1 then let m tend to 1.

Informativeness is weakly increasing in mobility, m :

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

as $\alpha > \beta$ and $\alpha > 1 - \beta$ and $\frac{1}{2} < p \leq 1$ and $0 \leq m \leq 1$.

A.3.2 Deriving mobility m_t in terms of informativeness I_t

$$m_t = F(V \cdot I_t) \tag{5}$$

Note that in equilibrium winning is good news about school quality. Therefore either each student applies to their local school or both apply to the winning school. If a student is accepted at their non-local winning school, they have to incur their cost of transport. Therefore a student will apply if and only if the expected benefit of attending their non-local winning school conditional on being accepted is higher than their realization of transport costs. Let this optimal cut-off level for generation t be denoted by $C_t \in \mathbb{R}$. It can be decomposed into the expected gain from attending the non-local winning school, conditional on this school being the good school, denoted by $V \in \mathbb{R}$, weighted by the level of informativeness $I_t \in \mathbb{R}$. Families' gain from attending the good school rather than the bad school depends on their child's ability. They draw inferences about this ability conditional on their child having been selected out of two applicants, because if a family applies to the non-local winning school then this school will be oversubscribed in equilibrium. Hence, the expected gain from attending the non-local winning school conditional on this school being the good school, denoted by $V \in \mathbb{R}^+$, is given by

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

Then C_t is given by

$$\begin{aligned} C_t(v, p, \Delta_L, \Delta_H, I_t) &= V(v, p, \Delta_L, \Delta_H) I_t \\ &= v((1 - p) \Delta_L + p \Delta_H) I_t. \end{aligned}$$

Both families apply to the winning school if realized transport costs lie below this cut-off C_t . Given transport costs are distributed according to $F(c)$, the ex-ante probability that both families apply to the winning school is given by

$$\begin{aligned} m_t(v, p, \Delta_L, \Delta_H, I_t, F(c)) &= F(V(v, p, \Delta_L, \Delta_H) I_t) \\ &= F(v((1-p)\Delta_L + p\Delta_H) I_t). \end{aligned}$$

A.3.3 Equilibrium

Denote the level of mobility in equilibrium steady state by m^* . Let families conjecture that rankings are generated as part of a steady state characterized by mobility level $\hat{m}$ and denote their optimal level of mobility by m . Then by definition of equilibrium steady state we look for a fixed point such that,

$$m^* \equiv m = \hat{m}.$$

By equation (3) and (5),

$$m(\hat{m}, p, \alpha, \beta, v, \Delta_L, \Delta_H, F(c)) = F(V(v, p, \Delta_L, \Delta_H) \cdot I(\hat{m}, p, \alpha, \beta))$$

where $\alpha, \beta, \Delta_L, \Delta_H$ are defined by the production function.

Simplifying notation yields

$$m(\hat{m}) = F(V \cdot I(\hat{m})).$$

Hence, the equilibrium level of mobility m^* is characterized by

$$m^* = F(V \cdot I(m^*)),$$

where m^* is a function of $p, \alpha, \beta, v, \Delta_L, \Delta_H, F(c)$, and $\alpha, \beta, \Delta_L, \Delta_H$ are defined by the production function. And the equilibrium level of informativeness is given by $I(m^*)$.

A.3.4 Existence of equilibrium

An equilibrium level of mobility satisfies condition (1). Let $G(\hat{m}) \equiv F(V \cdot I(\hat{m}))$. G is monotone increasing and $\hat{m} \in [0, 1]$. By Tarski's fixed point theorem, there exists an m^* such that $G(m^*) = m^*$.

A.4 Proof Proposition 2 [Comparative Statics]:

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 of transport costs, $F(c)$,*
- (b) *an increase in families' valuation of high performance, v ,*
- (c) *an increase in the influence of school quality on educational performance,*
- (d) *an increase in the probability, p , that an oversubscribed school selects a high-ability student if school quality and student ability are complements, i.e. if $\Delta_H - \Delta_L \geq 0$.*

Let informativeness, I , be our variable of interest. In equilibrium steady state, both (1) and (2) hold. Therefore, an equilibrium level of informativeness satisfies

$$I = \alpha + \beta - 1 + F(V(v, p, \Delta_L, \Delta_H) \cdot I) (\alpha - \beta) (2p - 1) I \quad (15)$$

where $\alpha, \beta, \Delta_L, \Delta_H$ are defined by the production function.

Denote by I^{1*} the smallest equilibrium level of informativeness satisfying condition (15).

Define a function $Z(p, v, F(c), \alpha, \beta, \Delta_L, \Delta_H, I) \rightarrow [0, 1]$ by

$$Z \equiv \alpha + \beta - 1 + F(V(v, p, \Delta_L, \Delta_H) \cdot I) (\alpha - \beta) (2p - 1) I.$$

In what follows, I will first fix the parameter subject to exogenous change. To explain the method, denote this parameter by $t \in T$, where T is a partially ordered set. To simplify notation, I will write $Z(t, I) : T \times [0, 1] \rightarrow [0, 1]$.

For any $t \in T$, $Z(t, I)$ is continuous but for upward jumps in I . Then I will show that $Z(t, I)$ is monotone nondecreasing in t for all $I \in [0, 1]$. After Corollary 1, (p. 446), in Milgrom and Roberts [1994], (henceforth MR), this implies that the smallest fixed point of $Z(t, I)$, given by $I^{1*}(t)$, is monotone nondecreasing in t . Finally, I will show that $m^{1*}(t)$ is also monotone nondecreasing in t using condition (5).

A.4.1 (a) A negative shift in FOSD shift of the distribution of transport costs $F(\cdot)$

To simplify notation, write Z as a function of $F(\cdot)$ and I :

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

where V can be treated as a constant because it is neither a function of $F(\cdot)$ nor I .

For any I and any $F(\cdot)$ and $\bar{F}(\cdot)$, such that $F(\cdot)$ first-order stochastically dominates $\bar{F}(\cdot)$, i.e. for any c it holds that $F(c) \leq \bar{F}(c)$, we have

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

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

Using condition (5) with $I = I^{1*}(F(\cdot))$ and simplifying notation such that mobility is written as a function of $F(\cdot)$:

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

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

$$m^{1*}(F(\cdot)) = F(V \cdot I^{1*}(F(\cdot))) \leq \bar{F}(V \cdot I^{1*}(\bar{F}(\cdot))) = m^{1*}(\bar{F}(\cdot))$$

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

A.4.2 (b) An increase in families' valuation of high outcomes v

To simplify notation, write $V(v, p, \Delta_L, \Delta_H)$ as $V(v)$ and write Z as a function of v and I :

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

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

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

since $(\alpha - \beta) (2p - 1) I \geq 0$, $I \geq 0$, $V(v) \geq 0$ and $F(\cdot)$ is increasing. By MR, an increase in the valuation of high outcomes, v , increases the smallest fixed point, $I^{1*}(v)$.

Using condition (5) with $I = I^{1*}(v)$ and simplifying notation such that mobility is written as a function of v :

$$m^{1*}(v) = F(V(v) \cdot I^{1*}(v)).$$

An increase in v increases $m^{1*}(v)$:

$$m^{1*}(v') = F(V(v') \cdot I^{1*}(v')) \geq F(V(v) \cdot I^{1*}(v)) = m^{1*}(v)$$

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

A.4.3 (c) An increase in the impact of school quality on educational performance, i.e. an increase in either in any of $\bar{a}, \underline{a}, \bar{b}, \underline{b}$

i) Consider the probability that a high-ability student achieves a high outcome at the good school, denoted by $P(\bar{h}|\bar{H}) = \bar{a}$. The probability that the good school wins with the high-ability student, denoted by $\alpha(\bar{a}) = \frac{\bar{a} + \underline{b}}{2}$, is a function of $\bar{a}$. In addition, the difference between the probability that the high-ability student achieves a high outcome at the good school and

the probability he achieves a high outcomes at the bad school, denoted by $\Delta_H(\bar{a}) = \bar{a} - (1 - \underline{a})$, is a function of $\bar{a}$.

To simplify notation, write $V(v, p, \Delta_L, \Delta_H(\bar{a}))$ as $V(\bar{a})$ and write Z as a function of $\bar{a}$ and I :

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

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

$$\begin{aligned} Z(\bar{a}', I) - Z(\bar{a}, I) &= \left[\frac{\bar{a}' - \bar{a}}{2} \right] [1 + (2p - 1) F(V(\bar{a}) I) I] \\ &\quad + \left(\frac{\bar{a}' + \underline{a}}{2} - \beta \right) (2p - 1) \cdot \\ &\quad \left\{ F(V(\bar{a}') I) - F(V(\bar{a}) I) \right\} I \\ &\geq 0 \end{aligned}$$

since $V(\bar{a}') \geq V(\bar{a}) \geq 0$, $\alpha(\bar{a}') = \frac{\bar{a}' + \underline{a}}{2} > \beta$ by monotonicity of the production process, $\frac{1}{2} < p \leq 1$, $0 \leq I \leq 1$ and $F(c)$ is positive and increasing. By MR, an increase in the probability that a high-ability student achieves a high outcome at the good school, denoted by $P(\bar{h}|\bar{H}) = \bar{a}$, increases the smallest fixed point, $I^{1*}(\bar{a})$.

Using condition (5) with $I = I^{1*}(\bar{a})$ and simplifying notation such that mobility is written as a function of $\bar{a}$:

$$m^{1*}(\bar{a}) = F(V(\bar{a}) I^{1*}(\bar{a})).$$

An increase in $\bar{a}$ increases m_1^* :

$$m^{1*}(\bar{a}') = F(V(\bar{a}') I^{1*}(\bar{a}')) \geq F(V(\bar{a}) I^{1*}(\bar{a})) = m^{1*}(\bar{a})$$

since $I_1^*(\bar{a}') \geq I_1^*(\bar{a}) \geq 0$, $V(\bar{a}') \geq V(\bar{a}) \geq 0$ and $F(\cdot)$ is increasing.

ii) Consider the probability that a low-ability student achieves a high out-

come at the good school, denoted by $P(\bar{h}|\bar{L}) = \bar{b}$. The probability that the good school wins with the low-ability student, denoted by $\beta(\bar{b}) = \frac{\bar{b}+a}{2}$, is a function of $\bar{b}$. In addition, the difference between the probability that the low-ability student achieves a high outcome at the good school and the probability he achieves a high outcomes at the bad school, denoted by $\Delta_L(\bar{b}) = \bar{b} - (1 - \underline{b})$, is a function of $\bar{b}$.

To simplify notation, write $V(v, p, \Delta_L(\bar{b}), \Delta_H)$ as $V(\bar{b})$ and write Z as a function of $\bar{b}$ and I :

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

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

$$\begin{aligned} Z(\bar{b}', I) - Z(\bar{b}, I) &= \left[\frac{\bar{b}' - \bar{b}}{2} \right] [1 - (2p - 1) F(V(\bar{b}) I) I] \\ &\quad + \left(\alpha - \frac{\bar{b}' + a}{2} \right) (2p - 1) \cdot \\ &\quad \left\{ F(V(\bar{b}') I) - F(V(\bar{b}) I) \right\} I \\ &\geq 0 \end{aligned}$$

since $V(\bar{b}') \geq V(\bar{b}) \geq 0$, $\alpha > \beta(\bar{b}') = \frac{\bar{b}'+a}{2}$ by monotonicity of the production process, $\frac{1}{2} < p \leq 1$, $0 \leq I \leq 1$, and $F(c)$ is positive and increasing. By MR, an increase in the probability that a low-ability student achieves a high outcome at the good school, denoted by $P(\bar{h}|\bar{L}) = \bar{b}$, increases the smallest fixed point, $I^{1*}(\bar{b})$.

Using condition (5) with $I = I^{1*}(\bar{b})$ and simplifying notation such that mobility is written as a function of $\bar{b}$:

$$m^{1*}(\bar{b}) = F(V(\bar{b}) I^{1*}(\bar{b})).$$

An increase in $\bar{b}$ increases m_1^* :

$$m^{1*}(\bar{b}') = F(V(\bar{b}') I^{1*}(\bar{b}')) \geq F(V(\bar{b}) I^{1*}(\bar{b})) = m^{1*}(\bar{b})$$

since $I_1^*(\bar{b}') \geq I_1^*(\bar{b}) \geq 0$, $V(\bar{b}') \geq V(\bar{b}) \geq 0$ and $F(\cdot)$ is increasing.

iii) Consider the probability that a high-ability student achieves a high outcome at the bad school, denoted by $P(\underline{h}|\underline{H}) = 1 - \underline{a}$. The probability that the good school wins with the low-ability student, denoted by $\beta(\underline{a}) = \frac{\bar{b} + \underline{a}}{2}$, is a function of $\underline{a}$. In addition, the difference between the probability that the high-ability student achieves a high outcome at the good school and the probability he achieves a high outcomes at the bad school, denoted by $\Delta_H(\underline{a}) = \bar{a} - (1 - \underline{a})$, is a function of $\underline{a}$.

To simplify notation, write $V(v, p, \Delta_L, \Delta_H(\underline{a}))$ as $V(\underline{a})$ and write Z as a function of $\underline{a}$ and I :

$$Z(\underline{a}, I) \equiv \alpha + \beta(\underline{a}) - 1 + F(V(\underline{a}) \cdot I) (\alpha - \beta(\underline{a})) (2p - 1) I.$$

For any $\underline{a}, \underline{a}'$ such that $\underline{a}' > \underline{a}$ and any I ,

$$\begin{aligned} Z(\underline{a}', I) - Z(\underline{a}, I) &= \left[\frac{\underline{a}' - \underline{a}}{2} \right] [1 - (2p - 1) F(V(\underline{a}) I) I] \\ &\quad + \left(\alpha - \frac{\underline{a}' + \bar{b}}{2} \right) (2p - 1) \cdot \\ &\quad \left\{ F(V(\underline{a}') I) - F(V(\underline{a}) I) \right\} I \\ &\geq 0 \end{aligned}$$

since $V(\underline{a}') \geq V(\underline{a}) \geq 0$, $\alpha > \beta(\underline{a}') = \frac{\underline{a}' + \bar{b}}{2}$ by monotonicity of the production process, $\frac{1}{2} < p \leq 1$, $0 \leq I \leq 1$, and $F(\cdot)$ is positive and increasing. By MR, an decrease in probability that a high-ability student achieves a high outcome at the bad school, i.e. an increase in $\underline{a}$, increases the smallest fixed point, $I^{1*}(\underline{a})$.

Using condition (5) with $I = I^{1*}(\underline{a})$ and simplifying notation such that mobility is written as a function of $\underline{a}$:

$$m^{1*}(\underline{a}) = F(V(\underline{a}) I^{1*}(\underline{a})).$$

An increase in $\underline{a}$ increases m_1^* :

$$m^{1*}(\underline{a}') = F(V(\underline{a}') I^{1*}(\underline{a}')) \geq F(V(\underline{a}) I^{1*}(\underline{a})) = m^{1*}(\underline{a})$$

since $I_1^*(\underline{a}') \geq I_1^*(\underline{a}) \geq 0$, $V(\underline{a}') \geq V(\underline{a}) \geq 0$ and $F(c)$ is increasing.

iv) Consider the probability that a low-ability student achieves a high outcome at the bad school, denoted by $P(\underline{h}|\underline{L}) = 1 - \underline{b}$. The probability that the good school wins with the high-ability student, denoted by $\alpha(\underline{b}) = \frac{\bar{a} + \underline{b}}{2}$, is a function of $\underline{b}$. In addition, the difference between the probability that the low-ability student achieves a high outcome at the good school and the probability he achieves a high outcomes at the bad school, denoted by $\Delta_L(\underline{b}) = \bar{b} - (1 - \underline{b})$, is a function of $\underline{b}$.

To simplify notation, write $V(v, p, \Delta_L(\underline{b}), \Delta_H)$ as $V(\underline{b})$ and write Z as a function of $\underline{b}$ and I :

$$Z(\underline{b}, I) \equiv \alpha(\underline{b}) + \beta - 1 + F(V(\underline{b}) \cdot I) (\alpha(\underline{b}) - \beta) (2p - 1) I$$

For any $\underline{b}, \underline{b}'$ such that $\underline{b}' > \underline{b}$ and any I ,

$$\begin{aligned} Z(\underline{b}', I) - Z(\underline{b}, I) &= \left[\frac{\underline{b}' - \underline{b}}{2} \right] [1 + (2p - 1) F(V(\underline{b}) I) I] \\ &\quad + \left(\frac{\underline{b}' + \bar{a}}{2} - \beta \right) (2p - 1) \cdot \\ &\quad \left\{ F(V(\underline{b}') I) - F(V(\underline{b}) I) \right\} I \\ &\geq 0 \end{aligned}$$

since $V(\underline{b}') \geq V(\underline{b}) \geq 0$, $\alpha(\underline{b}') = \frac{b' + \bar{a}}{2} > \beta$ by monotonicity of the production process, $\frac{1}{2} < p \leq 1$, $0 \leq I \leq 1$, and $F(\cdot)$ is positive and increasing. By MR, a decrease in the probability that a low-ability student achieves a high outcome at the bad school, i.e. an increase in $\underline{b}$, increases the smallest fixed point, $I^{1*}(\underline{b})$.

Using condition (5) with $I = I^{1*}(\underline{b})$ and simplifying notation such that mobility is written as a function of $\underline{b}$:

$$m^{1*}(\underline{b}) = F(V(\underline{b}) I^{1*}(\underline{b}))$$

An increase in $\underline{b}$ increases m_1^* :

$$m^{1*}(\underline{b}') = F(V(\underline{b}') I^{1*}(\underline{b}')) \geq F(V(\underline{b}) I^{1*}(\underline{b})) = m^{1*}(\underline{b})$$

since $I_1^*(\underline{b}') \geq I_1^*(\underline{b}) \geq 0$, $V(\underline{b}') \geq V(\underline{b}) \geq 0$ and $F(\cdot)$ is increasing.

A.4.4 (d) An increase in the probability, p , that an oversubscribed school selects a high-ability student if school quality and student ability are complements, i.e. if $\Delta_H - \Delta_L \geq 0$.

To simplify notation, write $V(v, p, \Delta_L, \Delta_H)$ as $V(p)$ and write Z as a function of p and I :

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

Assume complements, i.e. $\Delta_H \geq \Delta_L$, then for any p, p' such that $p' > p$:

$$V(p') = v\left(\left(1 - p'\right)(\Delta_L) + p'(\Delta_H)\right) \geq v\left((1 - p)(\Delta_L) + p(\Delta_H)\right) = V(p)$$

since $v > 0$.

Then for any I and any p, p' such that $p' > p$,

$$\begin{aligned} Z(p', I) - Z(p, I) &= I(\alpha - \beta) \left[(2p - 1) \left[F(V(p') I) - F(V(p) I) \right] \right. \\ &\quad \left. + 2(p' - p) F(V(p') I) \right] \\ &\geq 0 \end{aligned}$$

since $V(p') \geq V(p)$, $0 \leq I \leq 1$, $\alpha > \beta$, $\frac{1}{2} < p \leq 1$, and $F(c)$ is positive and increasing. By MR, if complements, i.e. $\Delta_H \geq \Delta_L$, an increase in schools' capability to select, denoted by p , increases the smallest fixed point, $I^{1*}(p)$.

Using condition (5) with $I = I^{1*}(p)$ and simplifying notation such that mobility is written as a function of p :

$$m^{1*}(p) = F(V(p) I^{1*}(p)).$$

If complements, i.e. $\Delta_H \geq \Delta_L$, then an increase in p increases mobility m^{1*} :

$$m^{1*}(p') = F(V(p') I^{1*}(p')) \geq F(V(p) I^{1*}(p)) = m^{1*}(p)$$

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

A.5 Proof Proposition 3 [Performance]

Proposition 3 [Performance]: (1a) A negative shift in FOSD of the distribution of transport costs, $F(\cdot)$, and (1b) an increase in families' valuation of high performance, v , both raise (lower) ex-ante expected total performance if student ability and school quality are complements (substitutes), i.e. if $\Delta_H - \Delta_L \geq 0$ ($\Delta_H - \Delta_L \leq 0$).

(2) An increase in the probability, p , that an oversubscribed school selects a high-ability student raises the ex-ante expected performance, if student ability and school quality are complements, i.e. if $\Delta_H - \Delta_L \geq 0$.

We will derive an expression for total expected performance in terms of the steady-state level of mobility. As a first step, we will derive the steady-state probability that the high-ability student attends the good school.

A.5.1 Expected degree of sorting between students and schools in steady state

$$P(\overline{H}) = \frac{1 + m(2\beta - 1)(2p - 1)}{2(1 - m(\alpha - \beta)(2p - 1))} \quad (16)$$

We will derive steady-state probability that the high-ability student attends the good school, denoted by $P(\overline{H})$, in the equilibrium steady state characterized by mobility level $m^{1*} \in \mathbb{R}$. Denote the event that the high-ability student is at the good school by $\overline{H}$ and the event that the high-ability student is at the bad school by $\underline{H}$. Further, denote the state realized in period t by $\overline{H}_t$ and $\underline{H}_t$ respectively.

The probability that the high-ability student in period t attends the period- $t-1$ winning school, denoted by $P(H_t^{W_{t-1}})$, is given by

$$P(H_t^{W_{t-1}}) = \frac{1}{2}(1 - m^{1*} + 2pm^{1*}).$$

It is a function of the level of mobility in equilibrium steady state, m^{1*} , and schools' capability to select, p . For derivation see Appendix Proof Proposition on Equilibrium.

Next we will derive the probability that the good (bad) school takes on the high-ability student in period t , conditional on having taken on the high-ability student in period $t - 1$. These transition probabilities between states $\overline{H}_t$ and $\underline{H}_t$ are functions of mobility level m^{1*} , schools' capability to select, p , and the production function. In particular, the production process defines the probability that the good school wins with the high-ability student (against the bad school with the low-ability student), denoted by $P(\overline{W}_t|\overline{H}_t) = \alpha$, and the probability that the good school wins with the low-ability student (against the bad school with the high ability student), denoted by $P(\overline{W}_t|\underline{H}_t) = \beta$. (See production process in section (2)).

Based on this we construct transition probabilities between state j and

state i denoted by $P(i|j)$:

$$\begin{aligned} P(\overline{H}_t|\overline{H}_{t-1}) &= P(H^{W_{t-1}}) P(\overline{W}_{t-1}|\overline{H}_{t-1}) \\ &\quad + (1 - P(H^{W_{t-1}})) (1 - P(\overline{W}_{t-1}|\overline{H}_{t-1})) \\ &= \frac{1}{2}\alpha (1 - m^{1*} + 2m^{1*}p) + \frac{1}{2}(1 - \alpha) (1 + m^{1*} - 2m^{1*}p) \end{aligned}$$

and

$$\begin{aligned} P(\underline{H}_t|\underline{H}_{t-1}) &= (1 - P(H^{W_{t-1}})) P(\overline{W}_{t-1}|\underline{H}_{t-1}) \\ &\quad + P(H^{W_{t-1}}) (1 - P(\overline{W}_{t-1}|\underline{H}_{t-1})) \\ &= \frac{1}{2}\beta (1 + m^{1*} - 2m^{1*}p) + \frac{1}{2}(1 - m^{1*} + 2m^{1*}p) (1 - \beta). \end{aligned}$$

The transition matrix, T , is given by:

$$T = \begin{pmatrix} P(\overline{H}_t|\overline{H}_{t-1}) & P(\underline{H}_t|\overline{H}_{t-1}) \\ P(\overline{H}_t|\underline{H}_{t-1}) & P(\underline{H}_t|\underline{H}_{t-1}) \end{pmatrix}.$$

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

$$\begin{pmatrix} P(\overline{H}) & P(\underline{H}) \end{pmatrix} = \begin{pmatrix} P(\overline{H}) & P(\underline{H}) \end{pmatrix} T$$

and

$$P(\overline{H}) + P(\underline{H}) = 1.$$

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

$$\lim_{k \rightarrow \infty} T^k = \mathbf{1} \begin{pmatrix} P(\overline{H}) & P(\underline{H}) \end{pmatrix}$$

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 characterized by

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

The Markov chain is not irreducible if and only if $\alpha = 1$ and $m^{1*} = 1$ and $p = 1$.²⁸ In this case, the state $\overline{H}$ is an absorbing state and I assign $P(\overline{H}) = 1$. So

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

for all parameter configurations.

A.5.2 Total expected performance

Total expected performance in equilibrium steady state, denoted by R , is a function of the production process and the probability that the high-ability student attends the good school in equilibrium steady state, denoted by $P(\overline{H})$.

If we treat $P(\overline{H})$ as given then total expected performance can be written as

$$\begin{aligned} R &= 1 - \underline{a} + \bar{b} + P(\overline{H}) (\Delta_H - \Delta_L) \\ &= 1 - \underline{a} + \bar{b} + P(\overline{H}) (\bar{a} + \underline{a} - \bar{b} - \underline{b}). \end{aligned} \tag{17}$$

The probability that the high-ability student attends the good school in equilibrium steady state, $P(\overline{H})$, is itself a function of the equilibrium level of mobility, m^{1*} , families' valuation of high outcomes, v , schools' capability to select, p , the distribution of transport costs, $F(\cdot)$, and the production process.

Substitute for (16) in (17),

²⁸Order of limits matters, first let p tend to 1 then let m^{1*} tend to 1.

$$R = 1 - \underline{a} + \bar{b} + \frac{1 + m^{1*} (2\beta - 1) (2p - 1)}{2 (1 - m^{1*} (\alpha - \beta) (2p - 1))} (\bar{a} + \underline{a} - \bar{b} - \underline{b}) \quad (18)$$

where the equilibrium level of mobility is a function of $p, \alpha, \beta, v, \triangle_L, \triangle_H, F(c)$.

Treat m^{1*} as an exogenously determined parameter. Using (18),

$$\frac{\partial R(m^{1*})}{\partial m_1^*} = (\bar{a} + \underline{a} - \bar{b} - \underline{b}) \frac{\partial}{\partial m_1^*} \frac{1 + m^{1*} (2\beta - 1) (2p - 1)}{2 (1 - m^{1*} (\alpha - \beta) (2p - 1))} \quad (19)$$

where

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

for all parameter configurations, since $\alpha > 1 - \beta$ and $\frac{1}{2} < p \leq 1$.

Hence, if complements, i.e. if $\bar{a} + \underline{a} - \bar{b} - \underline{b} \geq 0$, then

$$\frac{\partial R(m^{1*})}{\partial m_1^*} \geq 0. \quad (20)$$

Hence, if substitutes, i.e. if $\bar{a} + \underline{a} - \bar{b} - \underline{b} \leq 0$, then

$$\frac{\partial R(m^{1*})}{\partial m_1^*} \leq 0. \quad (21)$$

In addition,

$$\frac{\partial R(p)}{\partial p} = (\bar{a} + \underline{a} - \bar{b} - \underline{b}) \frac{\partial}{\partial p} \frac{1 + m^{1*} (2\beta - 1) (2p - 1)}{2 (1 - m^{1*} (\alpha - \beta) (2p - 1))} \geq 0 \quad (22)$$

since

$$\frac{\partial}{\partial p} \frac{1 + m^{1*} (2\beta - 1) (2p - 1)}{2 (1 - m^{1*} (\alpha - \beta) (2p - 1))} = \frac{(\alpha + \beta - 1) m^{1*}}{(1 - m_1^* (\alpha - \beta) (2p - 1))^2} \geq 0$$

as $\alpha > 1 - \beta$ and $0 \leq m^{1*} \leq 1$.

A.5.3 Proposition part (1a)

The distribution of transport costs, $F(\cdot)$, affects total expected performance, R , only through the equilibrium level of mobility, m_1^* . For simplification, we write total expected performance as $R(m^{1*}(F(\cdot)))$.

As the proposition on comparative statics shows, a negative FOSD shift in $F(\cdot)$ increases $m_1^*(F(\cdot))$. By the proposition on comparative statics and by (20), $R(m^{1*}(F(\cdot)))$ increases with a negative FOSD in $F(\cdot)$ if student ability and school quality are complements. By the proposition on comparative statics and by (21), $R(m^{1*}(F(\cdot)))$ decreases with a negative FOSD in $F(\cdot)$ if student ability and school quality are substitutes.

A.5.4 Proposition part (1b)

Families' valuation of high performance, v , affects total expected performance, R , only through the equilibrium level of mobility, m_1^* . For simplification, we write total expected performance as $R(m^{1*}(v))$.

As the proposition on comparative statics shows, a rise in v increases $m_1^*(v)$. By the proposition on comparative statics and by (20), $R(m^{1*}(v))$ increases with a rise in v if student ability and school quality are complements. By the proposition on comparative statics and by (21), $R(m^{1*}(v))$ decreases with a rise in v if student ability and school quality are substitutes.

A.5.5 Proposition Part (2)

Total expected performance, R , depends on the probability, p , that an oversubscribed school selects a high-ability student both directly as well as through the equilibrium level of mobility, m^{1*} . For simplification, we write total expected performance as $R(m^{1*}(p), p)$.

Take p' and p'' such that $p'' > p'$:

$$\begin{aligned} R\left(m^{1*}(p''), p''\right) - R\left(m^{1*}(p'), p'\right) &= R\left(m^{1*}(p''), p''\right) \\ &\quad - R\left(m^{1*}(p''), p'\right) \\ &\quad + R\left(m^{1*}(p''), p''\right) \\ &\quad - R\left(m^{1*}(p'), p''\right). \end{aligned}$$

Assume school quality and student ability are complements, i.e. $\bar{a} + \underline{a} - \bar{b} - \underline{b} \geq 0$. By the proposition on comparative statics, $m^{1*}(p'') \geq m^{1*}(p')$. Then using (20), we can show that the indirect effect is positive:

$$R\left(m^{1*}(p''), p''\right) - R\left(m^{1*}(p'), p''\right) \geq 0.$$

Further, using (22), we can show that the direct effect is positive:

$$R\left(m^{1*}(p''), p''\right) - R\left(m^{1*}(p''), p'\right) \geq 0.$$

Consequently, the total effect of an increase in p is positive if there are complements, i.e. $\bar{a} + \underline{a} - \bar{b} - \underline{b} \geq 0$.

A.6 Examples for welfare results

A.6.1 Welfare

A utilitarian social planner will weigh up total expected performance, R , valued at v , against any expected expenditure on transport costs, denoted by TC . Hence total welfare, denoted by TW , will be given by

$$TW \equiv v \cdot R - TC.$$

Expected expenditure on transport costs, TC , can be expressed as the expected level of costs conditional on costs being sufficiently low such that both families apply to the winning school, weighted by the probability that both

families apply to the winning school.

$$\begin{aligned}
TC &= m^{1*} E(c|c < V \cdot I^{1*}) \\
&= m^{1*} E\left(c|c < V \cdot \frac{\alpha + \beta - 1}{(1 - m^{1*}(\alpha - \beta)(2p - 1))}\right) \\
&= m^{1*} \int_0^{V \cdot \frac{\alpha + \beta - 1}{(1 - m^{1*}(\alpha - \beta)(2p - 1))}} cF(c) dc
\end{aligned}$$

where the equilibrium level of mobility is a function of $p, \alpha, \beta, v, \Delta_L, \Delta_H, F(c)$ and V is a function of v, p, Δ_L, Δ_H .

A.6.2 Example 1

Show that welfare may decrease with a negative shift in FOSD of the distribution of transport costs even if student ability and school quality are complements.

Assume $F(c)$ is a Uniform distribution on $[0, \bar{c}]$ and $\bar{c} = 1$ initially. Further assume $\bar{a} = \underline{a} = 1$, $\bar{b} = \underline{b} = \frac{3}{4}$, $\bar{v} = \underline{v} = \frac{1}{2}$, $p = 1$ and $v = 1$, hence $\Delta_H = 1$, $\Delta_L = \frac{1}{4}$, $V = 1$, $\alpha = \frac{7}{8}$, $\beta = \frac{3}{4}$, $m^{1*} = 4 - \sqrt{11}$, $I^{1*} = \frac{5}{4 + \sqrt{11}}$. A negative shift in FOSD in $F(c)$ corresponds to a decrease in $\bar{c}$. Consider a fall in $\bar{c}$ to $\bar{c}'$ such that

$$m^{1*}(\bar{c}') < 1.$$

Then the change in $\bar{c}$ only affects welfare through the equilibrium levels of informativeness and mobility.

Total expected expenditure as a function of m^{1*} :

$$TC(m^{1*}) = m^{1*} \frac{V}{2} \cdot \frac{\alpha + \beta - 1}{(1 - m^{1*}(\alpha - \beta)(2p - 1))}$$

Since a negative shift in FOSD of $F(c)$ affects $TC(m^{1*})$ only through m^{1*} ,

we can focus on how total expenditure varies with m^{1*} :

$$\begin{aligned} \frac{\partial TC(m^{1*})}{\partial m^{1*}} &= \frac{V}{2} \cdot \frac{\alpha + \beta - 1}{(1 - m^{1*}(\alpha - \beta)(2p - 1))} \\ &\quad + m^{1*} \frac{V(\alpha - \beta)(2p - 1)(\alpha + \beta - 1)}{2(1 - m^{1*}(\alpha - \beta)(2p - 1))^2}. \end{aligned}$$

By (19), the derivate of total expected performance, R , with respect to m^{1*} weighted by families' valuation of high performance, v , is given by:

$$v \frac{\partial R(m^{1*})}{\partial m^{1*}} = v \frac{(\alpha + \beta - 1)(2p - 1)}{2[(1 - m^{1*}(\alpha - \beta)(2p - 1))]^2} (\Delta_H - \Delta_L).$$

For an increase in mobility, m^{1*} , to lower welfare, it needs to raise total expenditure, TC , more than total expected performance, R , weighted by the valuation of high performance, v :

$$\begin{aligned} v \frac{\partial R(m^{1*})}{\partial m^{1*}} &< \frac{\partial TC(m^{1*})}{\partial m^{1*}} \\ v(\alpha + \beta - 1)(2p - 1)(\Delta_H - \Delta_L) &< V(\alpha + \beta - 1). \end{aligned}$$

This is satisfied for the given parameter values and $\bar{c} = 1$ with a strict inequality:

$$\frac{3}{4} < 1.$$

A.6.3 Example 2

Show that welfare may increase with a negative shift in FOSD of the distribution of transport costs even if student ability and school quality are substitutes.

$F(c)$ is a Uniform distribution on $[0, \bar{c}]$ and $\bar{c} = \frac{1}{8}$ initially. Further assume $\bar{a} = 1$, $\underline{a} = \frac{1}{4}$, $\underline{b} = \bar{b} = 1$, $p = 1$ and $v = 1$, hence $\Delta_H = \frac{1}{4}$, $\Delta_L = 1$, $V = \frac{1}{4}$, $\alpha = 1$, $\beta = \frac{5}{8}$, $m^{1*} = 1$ and $I^{1*} = 1$. Then total expected expenditure can be written as:

$$TC = \frac{\bar{c}}{2} = \frac{1}{16}.$$

A negative shift in FOSD in $F(c)$ corresponds to a decrease in $\bar{c}$. For a lower level of $\bar{c}$, it is still true that $m^{1*} = 1$. Then a negative shift in FOSD of $F(c)$ affects TC only through $\bar{c}$ and hence total expenditure decreases. Expected total performance, R , remains unchanged. Therefore the utilitarian welfare function increases even if school quality and student ability are substitutes.

A.7 Proof Proposition 4 [Convergence]

The dynamic learning process is characterized by the informativeness of the first generation with access to rankings (7) and by the recurrence equation linking informativeness of subsequent generations (8). Its limit as $t \rightarrow \infty$, denoted by I^{1} , is the smallest equilibrium level of informativeness. As $\{I_t\}_{t=0,1,\dots}$ converges so does the sequence of mobility levels $\{m_t\}_{t=0,1,\dots}$. The sequence $\{m_t\}_{t=0,1,\dots}$ converges to the smallest equilibrium level of mobility, denoted by m^{1*} , where*

$$m^{1*} = F(V \cdot I^{1*}). \quad (9)$$

A.7.1 Dynamic Learning Process

Generation 0 has no access to rankings and therefore holds symmetric posterior beliefs about 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 and evaluating them at m_{t-1} yields

$$\begin{aligned} P(\bar{W}_{t-1}) &= P(\bar{W}_{t-1} | \bar{W}_{t-2}) P(\bar{W}_{t-2}) \\ &\quad + P(\bar{W}_{t-1} | \underline{W}_{t-2}) (1 - P(\bar{W}_{t-2})) \\ &= \frac{1}{2} (\alpha + \beta) \\ &\quad + \frac{1}{2} m_{t-1} (\alpha - \beta) (2p - 1) (2P(\bar{W}_{t-2}) - 1). \end{aligned} \quad (23)$$

Expressing (23) in terms of informativeness, where $I_t \equiv 2P(\overline{W}_{t-1}) - 1$, gives:

$$I_t = \alpha + \beta - 1 + m_{t-1}(\alpha - \beta)(2p - 1)I_{t-1} \quad (24)$$

Using (5), then (24) is equivalent to:

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

A.7.2 Convergence to smallest equilibrium level of informativeness I^{1*}

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 $I'_t > I_t$

$$\begin{aligned} Z(I'_t) - Z(I_t) &= (\alpha - \beta)(2p - 1) \\ &\quad \left\{ \left[F(V \cdot I'_t) - F(V \cdot I_t) \right] I_t + F(V \cdot I'_t) [I'_t - I_t] \right\} \\ &\geq 0 \end{aligned}$$

since F is positive and increasing and $V \geq 0$ and $I_t \geq 0$.

Using $I_0 = 0$, equation 23, 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 (5),

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

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 informativenss I^{1*} corresponds to the smallest equilibrium level of mobility m^{1*} .