Mobility, Learning and School Choice

Claudia Herresthal

February 24, 2014

University of Oxford
Department of Economics
Manor Road Building
Manor Road
Oxford OX1 3UG
United Kingdom
claudia.herresthal@economics.ox.ac.uk

Abstract

This paper studies the interplay between match quality, learning and information quality in the context of school choice. Families value good schools, but also prefer closer schools. On the other hand, schools prefer higher-ability students. Schools are ranked based on the (exam) performance of their student intake. Performance depends both on the school’s quality and on the student intake’s ability. Families use the most recent ranking to determine which school to apply to. In this dynamic environment, feedbacks exist between families’ inferences about school quality and the matches between students and schools. Comparative statics analyses are performed on the steady state of this dynamic process. Lower transport costs improve the informativeness of rankings. However, allowing schools to select students based on ability has ambiguous effects on the informativeness of rankings. Higher informativeness makes families’ demand more sensitive to quality indicators but lowers welfare for uniformly distributed transport costs if school quality is fixed.
1 Introduction

Recent years have seen the introduction of “market-based” reforms to public school systems. Beginning with Friedman (1955, 1962), reformers have argued that market forces should induce competition between providers, forcing them to cater for the needs of their consumers more efficiently than any government regulation could achieve. Most existing reforms allow families to choose public schools outside their enrolment district or to receive vouchers or tax breaks such that fee-paying schools become an affordable alternative.¹ These reforms have proven quite popular: in the UK only half of students attend their nearest school and over 20% do not attend one of the nearest three schools (Burgess, 2006).² However, despite this (apparent) popularity, and the extensive empirical research done on the effects of school choice, there are few theoretical underpinnings.³

A primary motivation for expanding school choice for families is that it can lead to improvements in outcomes. However, this will only happen if families can distinguish between schools based on their educational quality. That information on school quality can have important effects on families’ choices is demonstrated in experiments by Hastings et al. (2007). In the UK, families can gain information about school quality from league tables, which rank secondary schools based on a summary statistic of outcomes in a national standardised exam.⁴ However, families face the problem that test outcomes are imperfect indicators of school quality as they are also affected by student ability. To address this problem, UK league tables include additional information about student characteristics.⁵ Nevertheless, while survey evidence shows that most parents consult league tables, it also shows that parents tend to base their choices on raw results only (Couldron et al., 2008).

This paper provides a framework for analysing how school choice affects outcomes when individuals are not perfectly informed about school quality. Further, the paper explores new avenues by which the informational value of league tables can be improved when information is imperfect, deriving from a dynamic link between rankings and student allocation that has not been previously studied. Families’ beliefs that performance-based rankings order schools according to their intrinsic quality depends on their beliefs of how students are allocated.

---

¹Competition between schools is also fostered by facilitating the entry of new types of schools, such as charter schools in the US or academies in the UK.
²In the UK half of students live within 2.4 km of three or more schools and for 3 in 5 students one of the three nearest schools is in the top third of the national league table (in 2001).
³For example, see Black (1999), Bradley et al. (2001), Belfield et al. (2002), Hoxby (2000, 2003), Dolton (2003), Bayer et al. (2010), Gibbons et al. (2008).
⁴Allen et al. (2010) estimate that choosing a school based on such statistics improves educational outcomes relative to choosing a school at random.
⁵see Wilson et al. (2008) for details
to schools. In turn, families’ inferences about school quality from rankings influence their application decisions and hence the future allocation of students to schools. Further, if some schools have to ration their places, the allocation of students is determined not only by families’ choice of school but also by how schools select among applicants. Analysing the process by which schools and families choose one another when information is imperfect can yield surprising insights about how policymakers can improve the likelihood that league tables rank schools according to their intrinsic quality. For example, while Allen et al. (2010) argue that regulating schools’ admission such that each school’s intake is balanced would simplify the signal extraction problem faced by parents, my analysis shows that such regulation of school admissions can inhibit the extent to which information about the quality ordering of schools is transmitted by the market. Further, enhancing the informational value of league tables increases expected educational outcomes if school quality and student ability are complements as the higher-ability students will be matched with better schools more frequently. However, enhancing informativeness leads to a decrease in total welfare for uniformly distributed transport costs because students more frequently attend schools which are further away from home and so the expenditure on students’ commute to school increases.

In the dynamic model of this paper, there are two schools, one of bad and one of good quality. Consecutive generations of families decide on applications to schools. Each generation consists of two families, living in different school districts. Each family can choose whether to apply to the school in its district (its “local” school) or the one in the other district (its “non-local” school). If a child is accepted at its non-local school, its family incurs a cost of transport, drawn at random each period from a stationary distribution. Before making their application decision, families observe their transport costs and a ranking of the two schools based on schools’ relative performance in educating children in the previous generation. Families compare schools based on their child’s expected educational outcome and any transport costs incurred, conditional on their child being accepted at the school. Each school has capacity for one student. Students attend their preferred school, if they are admitted to it and attend the other school, otherwise. They then receive a mark (educational outcome) that is influenced by both their ability and their school’s quality.

Families at any given time take the choices (the degree of sorting between schools and students) in the past as given. Having observed the rankings of schools from the previous period they then update their beliefs about which school is better and decide which to apply to. Their application decisions affect the amount of sorting in the current period. More information about schools’ relative quality is conveyed by their performance, the more likely

---

6Burgess et al. (2009) shows that parents care about distance to school
it is that, each time students are matched with schools, the best-performing school takes on
the most able students. To see why, imagine that students in the first generation are allocated
to schools at random, because they have no information about the schools' relative quality.
But in the second generation the stronger students attend the school which performed best in
the previous period. Families who observe the schools' relative performance with the second
generation’s students can infer more about relative quality than those who observed the
schools' relative performance with the first generation’s students. This is because students’
choices in the second generation are informative about the first period’s results. The good
school is more likely to have come top in the first period than the bad school. Therefore the
good school’s intrinsic quality advantage is reinforced by the sorting in the second period
and the unconditional probability of the good school performing best with students in the
second generation is higher.

Over time, the dynamic learning process will converge to a steady state, in which the
information extracted from a given ranking of schools is constant across generations, and
consistent with the fact that each generation chooses between schools optimally based on the
information they have. The steady state provides a tractable framework for analysing how
exogenously given factors impact on the degree of learning. I explore how policy measures
impact on the behavior of individuals in steady state, on how well informed they are and on
their welfare (from a utilitarian social planner’s point of view).7 In steady state, a downward
shift of the distribution of transport costs (in the sense of first order stochastic dominance)
yields higher equilibrium informativeness of school rankings. This is because with lower
transport costs families are more likely to find it worthwhile to apply to the best-performing
school, for any level of informativeness. As the best-performing school is more likely to
receive two applications, this increases chances that the high-ability student attends the
best-performing school which in turn, as argued above, increases informativeness.

Moreover, increasing the likelihood with which schools can select the high-ability student
also raises informativeness of rankings if student ability and school quality are complements.
When ability and quality are complements there are two effects which reinforce one another.
There is the direct effect that whenever both students apply to the best-performing school,
this school is more likely to take on the high-ability student. In addition, given the comple-
mentarity of ability and quality, students’ expected benefit from attending the better school

7Defining welfare as the joint surplus of both families stands in contrast to other welfare objectives
considered in the discussion of educational policies, such as providing students of different ability levels with
equal chances to access high-quality schools. My definition allows for the possibility that both families could
be made better off, even if educational gains are realised by only one of the students, as long as one student’s
gain outweighs the other student’s loss and compensatory monetary transfers could be implemented.
rises, inducing families to apply to the best-performing school more often. If ability and quality are substitutes, increasing the likelihood with which schools can select the high-ability student has ambiguous effects on the informativeness of rankings. The two effects analogous to those mentioned above counteract each other. While the direct effect is as above, students’ expected benefit from attending the better school decreases, inducing families to apply to the best-performing school less often.

Welfare depends on how informative schools’ rankings are about their quality. A higher level of informativeness leads to a higher chance that the high-ability student attends the good school and that the low-ability student attends the bad school. If student ability and school quality are complements, then higher informativeness increases expected number of high educational outcomes. However, higher informativeness also leads to higher expected transport costs as students attend their non-local school more frequently. If transport costs are uniformly distributed then this effect on costs is so strong that total welfare decreases.\footnote{In light of these results, it seems that higher informativeness is not desirable. But this is a partial equilibrium result in the sense that school qualities are exogenously set and not allowed to adjust. As consumers become more informed about providers’ relative quality, providers’ incentives to invest in their quality in order to compete for consumers increase. Thus the overall impact of higher informativeness on welfare is more complicated. I treat school qualities as fixed despite these caveats because it is important first to understand the welfare cost associated with improving informativeness in the absence of quality improvements by providers.}

Formally, the results about learning in this paper are related to the results derived from the study of biased contests (Meyer, 1991).\footnote{In Meyer’s model, a firm wants to learn about the relative quality of two workers. It is optimal for the firm to assign a negative bias to the worker it believes to be of lower quality because by doing so it receives a stronger signal if this worker wins despite the disadvantage of the bias.} In this paper, differences in the abilities of admitted students effectively bias the ranking of schools’ performance. The model also has some similarities to models in the social learning literature in which agents extract information about the quality of different options from predecessors’ actions (e.g. Bikhchandani et al., 1992). My model shares with Lobel et al. (2007) the feature that agents have a limited window of observation but focuses on the steady state analysis rather than conditions for convergence.

This paper is organised as follows. Section 2 introduces the model. Section 3 defines the key objects of the analysis, namely mobility and informativeness, then it derives dynamic feedbacks using a steady state approach and, finally, it solves for the equilibrium steady state level of mobility and informativeness. In addition, comparative statics of the steady state equilibrium are analysed. Section 4 analyses the impact of exogenous changes on expected academic outcomes and on welfare. Section 5 concludes. All proofs and a table summarising the notation used can be found in the Appendix.
2 The Model

2.1 Players

There are two schools, which differ in their quality: one school is of good quality \((G)\), while the other is of bad quality \((B)\). Each generation consists of two families, whose children differ in their ability: one family’s child is of high ability \((H)\), while the other family’s child is of low ability \((L)\). The school that is local for one family is non-local for the other family. If a student attends his non-local school instead of his local school, he incurs a transport cost \(c\), which is drawn in each generation from a known distribution with continuous cumulative distribution function \(F(c)\).\(^{10}\) Families value the expected academic achievement a school can provide, but dislike the transport cost incurred. Schools value student ability.\(^{11}\) Each generation of families corresponds to one period. In each period, after the matching process between schools and students is completed, there will be exactly one student per school.

2.2 Schools’ production function

Educational outcomes can be either high \((h)\) or low \((l)\) and they depend on both student ability and school quality.

The probability a good school achieves a high result is given by \(P(h \mid \cdot^G)\) where \(\cdot\) is replaced either by \(H\) if the student is of high-ability, or by \(L\) if the student is of low ability :

<table>
<thead>
<tr>
<th>Student Ability</th>
<th>H</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P(h \mid \cdot^G))</td>
<td>1</td>
<td>(\alpha)</td>
</tr>
</tbody>
</table>

The probability a bad school achieves a low result is given by \(P(l \mid \cdot^B)\):

<table>
<thead>
<tr>
<th>Student Ability</th>
<th>H</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>(P(l \mid \cdot^B))</td>
<td>(\beta)</td>
<td>1</td>
</tr>
</tbody>
</table>

For a given student ability, the good school is at least as likely to achieve a high result as the bad school. For a given school quality, the high-ability student is at least as likely to achieve a high result as the low-ability student. The higher are \(\alpha\) and \(\beta\), the more important is

\(^{10}\)These costs are only incurred when a student gets accepted at his non-local school. Applications are costless.

\(^{11}\)Schools may value higher ability because they are rewarded based on academic achievement or based on the number of applications they receive. High-ability students may also be less costly to teach.
School quality relative to student ability in determining results. Furthermore, school quality and student ability are complements if $\beta \geq \alpha$, and substitutes if $\alpha \geq \beta$.

Schools’ results are ranked each period. The school with a higher result will come top in the ranking, and ties are broken at random. Denote the school that achieves the higher ranking by $W$. If the good school has the high-ability student while the bad school has the low-ability student, the good school will be the winner with certainty, i.e. $P(W^G|H^G) = 1$. By contrast, if the good school has the low-ability student while the bad school has the high-ability student, the good school will be the winner with probability $x$, i.e. $P(W^G|L^G) = \alpha \beta + \frac{1}{2} [\alpha (1-\beta) + (1-\alpha)\beta] \equiv x$.

2.3 Timing and Information

In each generation, student ability levels are randomly allocated among families such that the high-ability student is equally likely to have a local school of good or bad quality. Families have to decide whether to apply to their local or non-local school. Families observe their transport costs, but do not observe the ability of their own child. Further, families know that there is one good and one bad school and they have symmetric beliefs about which school is which. Each generation $t$ observes the ranking of schools based on the educational outcomes of generation $t-1$. The school that achieves the higher rank based on educational outcomes of generation $t-1$ is called the period $t-1$ winning school, denoted by $W_{t-1}$. Families do not observe the allocation of students to schools in period $t-1$ that generated this ranking. They are only aware that each generation in the past consisted of two types of students, that each type is equally likely to have had a local school of good quality and that transport costs were drawn independently from a stationary distribution with cdf $F(c)$. In addition, families know how many generations before them had access to league tables.

A family’s payoff is given by the expected probability of a high outcome ($h$), based on their posterior beliefs about school quality, less their expenditure on transport costs.$^{13}$ Once applications have been submitted, schools with more applications than places choose one student. The rejected student has to attend the other school. If a school chooses between

---

$^{12}$I could relax the assumption that $P(W|G, H) = 1$ and assume instead that $P(W|G, H) < 1$. As long as $x < P(W|G, H)$, the key insights would still hold.

$^{13}$Burgess et al. (2009) show that parents value educational attainment, travel costs and socioeconomic composition when choosing a school. In the context of Tiebout choice, Black (1999) estimates that parents are willing to pay a premium for living in the catchment area of a school with higher academic performance. However, Hastings et al. (2008) find preferences dominated by finding a school with student that share the same socio-economic background.
two applicants, it admits the high-ability student with probability $p \geq \frac{1}{2}$.\textsuperscript{14}

3 Equilibrium

Dynamic feedbacks exist between students’ willingness to travel to the school which they observe to have the best performance and the accuracy with which performance-based rankings signal which school is of better quality. The current generation’s willingness to travel to the school with better observed performance increases in their level of confidence that better performance signals higher quality. In addition, the current generation’s confidence that better performance signals higher quality increases in their conjecture of past generations’ willingness to travel to schools with better performance.

This section derives how much information rankings convey about relative school quality. The analysis uses two key objects, informativeness and mobility. Informativeness characterises families’ posterior beliefs, whereas mobility characterises families’ behaviour. The dynamic feedbacks described above are expressed formally in terms of these objects. To examine how the current generation’s posterior beliefs about school quality depend on the current generation’s conjecture about past generations’ mobility, we focus on a steady state of the dynamic process. In a steady state, the information conveyed by rankings about which school is best is constant across generations. In an equilibrium steady state, each generation best responds to the mobility of past generations. The equilibrium steady state provides a tractable framework for analysis, allowing for a straight-forward examination of how exogenously given factors impact on the degree of learning about school quality.

Period $t$ informativeness captures how much more confident generation $t$ is that the period $t-1$ winning school, denoted by $W_{t-1}$, is the good rather than the bad school. Further, mobility of generation $t$ captures the willingness of individuals in period $t$ to travel to the period $t-1$ winning school. Formally, mobility is the likelihood with which transport costs lie below some cut-off level that depends on families’ beliefs about which school is more likely to be good.

Definition:

Denote the school which wins with students in generation $t-1$ by $W_{t-1}$.

\textsuperscript{14}Schools prefer high-ability to low-ability applicants. $p < 1$ could arise because schools are not capable of identifying relative ability perfectly or because schools are restricted in their selection of students by other priorities in their admission code, e.g. distance to school, siblings at the school.
(1) Period $t$ informativeness is defined as

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

where generation $t$’s posterior beliefs are based on their symmetric prior and the ranking of schools derived from educational outcomes of generation $t - 1$.

(2) Mobility of generation $t$, denoted by $m_t$, is defined as the ex ante probability with which both families in generation $t$ apply to the school which performed best with students in generation $t - 1$ ($W_{t-1}$).

Generation $t$’s level of mobility ($m_t$) positively depends on period $t$ informativeness ($I_t$). The maximum level of transport costs that students in generation $t$ are willing to pay to attend the observed winner depends on the expected premium of being admitted at the good school rather than the bad school, denoted by $V$, weighted by their posterior beliefs that the period $t - 1$ winning school is the good rather than the bad school ($I_t$). The family in the winning school’s district incurs no transport costs to attend the winner. Hence they always apply. The family in the losing school’s district is willing to incur their realised transport costs if and only if they lie below their cut-off level $V \cdot I_t$. Hence the ex ante probability of both families applying to the winning school is $F(V \cdot I_t)$. Mobility of generation $t$ is given by

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

where $V \equiv v(\alpha (1 - p) + p\beta)$ and $v > 0$ denotes families’ valuation of high results.

Period $t$ informativeness depends on generation $t$’s conjecture about the mobility level of past generations. A steady state concept is used to analyse generation $t$’s inference. Assume generation $t$ conjectures that all past generations had the same level of mobility, denoted by $\hat{m}$. The process generating ranking outcomes becomes a time-homogenous Markov process. In steady state, the probability the good school wins at any given point in time is constant. Generation $t$ updates their beliefs that rankings reflect relative school quality based on this stationary distribution. Their posterior beliefs characterised by steady state informativeness $I(\hat{m})$ given by

$$I(\hat{m}) = \frac{x}{1 - \hat{m} (1 - x) (2p - 1)}$$

$I(\hat{m})$ is weakly increasing and convex in mobility $\hat{m}$. A higher conjectured mobility level
implies the conjecture of a stronger link between a school’s performance with any given generation and the ability of its student intake in the subsequent generation.\textsuperscript{15} Intuitively, a stronger link implies that students’ allocation is more informative about past performance and therefore more information is transmitted from one generation to the next. The good school has an intrinsic quality advantage over the bad school (if $x > 0$) which is enhanced by the sorting of students over time.\textsuperscript{16}

In steady state, mobility of the current generation can be expressed as a best response to the conjectured mobility of past generations:

$$m = F(V \cdot I(m))$$

An equilibrium steady state is a steady state such that each generation makes a best response to past generations’ mobility. Such an equilibrium state level of mobility (and informativeness) always exists. This is because mobility generated by optimal behaviour is a continuous function of conjectured mobility and mobility is bounded above by 1.

**Proposition 1 [Equilibrium]:**

*An equilibrium steady state level of mobility is characterised by

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

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

To illustrate this with an example, suppose $F$ is uniformly distributed on the support $[0, v\bar{c}]$. Then the number of equilibrium levels of mobility is either one or three. See Figure 1. Which case occurs depends both on the spread of the distribution for costs (i.e. the upper limit of the support $v\bar{c}$) and the convexity of $I(\hat{m})$. The mobility level which characterises families’ best response is proportional to informativeness given conjecture $\hat{m}$, subject to the constraint that mobility cannot exceed 1, and the constant of proportionality is inversely related to the size of the spread. If the spread is sufficiently large, only one fixed point exists:

\textsuperscript{15}This holds provided schools do not have to randomise when selecting between applicants, i.e. $p > \frac{1}{2}$.

\textsuperscript{16}Further insights on the relationship between $\hat{m}$ and $I(\hat{m})$ can be found in the section on convergence (see Section 5).
$m$ increases sufficiently slowly with $\hat{m}$ such that $m = \hat{m}$ at some $\hat{m} < 1$ and thereafter $\hat{m} > m$. Similarly, if the spread is sufficiently small, only one fixed point exists: $m$ increases sufficiently rapidly with $\hat{m}$ such that $m > \hat{m}$ at all $\hat{m} < 1$ and $m = \hat{m}$ at $\hat{m} = 1$. For intermediate levels of spread, three fixed points exists if and only if $I(\hat{m})$ is sufficiently convex. In this case, $m$ increases sufficiently slowly with $\hat{m}$ such that $m = \hat{m}$ at some $\hat{m} < 1$, then due to the convexity of $I(\hat{m})$, $m = \hat{m}$ again at some higher level $\hat{m} < 1$ and finally, $m = \hat{m}$ at the upper bound at $\hat{m} = 1$.

For general transport cost distributions $F$, comparative statics can be performed on any equilibrium steady state level of mobility. The following analysis will focus on the smallest such equilibrium steady state level of mobility ($m_{1*}$) and the corresponding equilibrium steady state level of informativeness ($I_{1*}$). The reason is that starting with the first generation who gains access to rankings, all subsequent generations’ optimal behaviour leads mobility to converge to the smallest steady state equilibrium level of mobility (henceforth equilibrium level of mobility). Further detail will be given in the section on Convergence (see section 5). Any exogenous change which increases the mobility level that characterises families’ best response to a given conjecture $\hat{m}$ will increase the equilibrium level of mobility, due to the positive relationship between mobility and conjectured mobility. Since such an exogenous change not only increases informativeness for any given level of mobility but also increases the equilibrium level of mobility, the equilibrium level of informativeness will also increase.

A downward shift in the distribution of transport costs, in the sense of FOSD, increases chances that in any generation $t$, families’ draw of transport costs lies below the maximum they are willing to pay to travel to the period $t - 1$ winning school, given their child is admitted. Hence families’ best response to a given conjecture $\hat{m}$ is characterised by a higher mobility level and equilibrium levels of mobility and informativeness increase. 2.

An increase in families’ valuation of high results increases the expected benefit for students in any generation $t$ of attending the period $t - 1$ winning school, conditional on being accepted. Hence the cut-off level for transport costs of generation $t$ increases and, prior to the realisation of transport costs, families have a higher likelihood of applying to the period $t - 1$ winning school (higher $m$) for any given level of conjectured mobility $\hat{m}$. Consequently, equilibrium levels of mobility and informativeness increase. See Figure 2: Note that for uniformly distributed transport cost, a negative FOSD shift in the distribution for transport costs affects the mapping between mobility and conjectured mobility in the same way as an increase in the valuation for high results. Increasing the extent to which educational outcomes are determined by school quality rather than student ability, i.e. higher $\alpha$ or $\beta$, has two effects on mobility, which reinforce
Figure 1: Equilibrium levels of mobility for $\alpha = 0.2$, $\beta = 0.4$, $p = 1$
Figure 2: $F^2$ first order stochastically dominates $F^1$
where $F^1 = U[0, \tilde{\tau}_1]$ s.t. $\frac{\alpha}{\tilde{\tau}_1} = 0.5$ and for $F^2 = U[0, \tilde{\tau}_2]$ s.t. $\frac{\alpha}{\tilde{\tau}_2} = 1$
and $\alpha = 0.2$, $\beta = 0.4$, $p = 1$

each other. Firstly, informativeness $I(\hat{m})$ increases at any level of conjectured mobility $\hat{m}$. Intuitively, for any given allocation of students to schools rankings are more likely to order schools according to their quality. Secondly, students’ expected benefit of attending the good school increases. Both effects increase the cut-off level for transport costs of families in any generation $t$, causing them to apply to the period $t - 1$ winning school more often. As above, equilibrium levels of mobility and informativeness increase. See Figure 3.

If ability and quality are complements, increasing schools’ capability to select students based on ability $p$ also has two effects on mobility, which reinforce one another. There is the direct effect that, whenever both students apply to the best-performing school, this school is more likely to take on the high-ability student. This causes informativeness to increase at a given level of conjectured mobility, for similar reasons for why an increase in conjectured mobility increases informativeness. In addition, given the complementarity of ability and quality, being accepted at the period $t - 1$ winning school is better news about the premium of attending this school for students in any generation $t$. Both effects raise families’ cut-off level for transport costs. As above, equilibrium levels of mobility and informativeness increase. If ability and quality are substitutes, increasing the likelihood with which schools select the high-ability student has ambiguous effects on the informativeness of rankings. The two effects are analogous to those mentioned above, but counteract each other. While the direct effect is as above, being accepted at the period $t - 1$ winning school is now worse news about the premium of attending this school for students in any generation $t$. So the overall effect on families’ cut-off level for transport costs is ambiguous and hence the effect on the equilibrium levels of mobility and informativeness is ambiguous. See Figure 4.
Figure 3: $\alpha^1 > \alpha^2$ where $\alpha^1 = 0.2$ and $\alpha^2 = 0.5$ and $\beta = 0.4$, $p = 1$ and $F = U \left[0, \bar{c}\right]$ s.t. $\frac{v}{\bar{c}} = 1$

Figure 4: $p^1 > p^2$ where $p^1 = 0.5$ and $p^2 = 0.7$
1) complements where $\alpha = 0.2$ and $\beta = 0.6$
2) substitutes where $\alpha^1 = 0.7$ and $\alpha^2 = 0.9$ and $\beta = 0.5$
and for 1) and 2) $F = U \left[0, \bar{c}\right]$ s.t. $\frac{v}{\bar{c}} = 1$
Proposition 2 [Comparative Statics]

The smallest equilibrium level of mobility, denoted by $m_1^*$, and the smallest equilibrium level of informativeness, denoted by $I_1^*$, both increase with

(a) a negative shift in FOSD of the distribution for transport costs $F(c)$,

(b) an increase in valuation for high results $v$,

(c) an increase in the influence of school quality on educational outcomes, i.e. an increase in $\alpha$ or $\beta$,

(d) an increase with the capability of schools to select students based on ability $p$ if school quality and student ability are complements ($\beta > \alpha$)

(Potentially separate section labelled Discussion)

This analysis suggests avenues by which policymakers can increase the likelihood that school league tables rank schools according to their relative quality. By improving transport links or subsidising the cost for commuting, a policymaker can reduce the costs that families incur if their child attends a school further away from home. In addition, a policymaker could raise families’ valuation of superior educational outcomes, e.g. by educating them about the skills premium achieved in the labour market. Both interventions will increase the number of families who find it worthwhile to apply to the school with the highest ranking. If schools have some ability to select between applicants, a larger pool of applicants for the school with the highest ranking will translate into a stronger student intake. Over time, the outcome will be an increase in sorting between students and schools and hence a higher likelihood that league tables reflect relative school qualities.

More generally, the analysis suggests that allowing school choice rather than assigning students a place at their local school implies a higher likelihood that league tables reflect relative school qualities. If families were automatically assigned a place at their local school, families can still choose where to live (Tiebout choice). As the demand for places at the better performing school increases, house prices in its enrolment district increase (see Black, 1999). Only families with a sufficiently high income can afford to move to the popular schools’ catchment area and those with insufficient income will have to move elsewhere. If there is a positive correlation between income and ability, better performing schools may attract an intake of higher ability than their competitors. However, if families faced a lower cost for taking up a place at a popular school, e.g. only costs of transport rather than a house price premium, the school would have a larger applicant pool to select from and its advantage in terms of student ability would be higher. Other schemes that reduce the cost of
applying to a non-default school, such as voucher schemes, can also increase informativeness by strengthening the link between a school’s performance and its advantage in student ability.

A further channel by which a policymaker could improve how well school league tables signal school quality concerns schools’ admission process. A policymaker could enhance schools’ capability to select students based on their ability, either by imposing fewer restrictions on admission regulations or by increasing financial means to develop entry exams which are informative about applicants’ ability. Such an intervention would again facilitate sorting between students and schools. It is clear from this analysis that schools’ admission process impacts on the information conveyed by league tables and this effect should be considered in the design of admission regulations. Allen et al. (2010) suggest that the composition of student intakes should be balanced across schools in order to learn more about school quality from educational outcomes. This analysis suggests that balancing student intakes is not the optimal way to achieve an accurate ordinal ranking of schools based on their quality.

By facilitating sorting between students and schools, even small changes in school quality (measured here by parameters $\alpha$ and $\beta$) can have larger than expected effects on informativeness over time. Not only do rankings become more informative at any given allocation of students to schools, but the increased sorting between students and schools will further reinforce how well rankings reflect school quality. Therefore changes in relative school quality cause a larger demand response than predicted if these dynamic aspects are ignored. School qualities have been treated as fixed in the preceeding analysis. However, this finding suggests that, if schools were to invest in their quality, their incentives to make up for small quality difference are large.

These findings apply to other public services for which the outcomes of the service provision depend on provider quality as well as the characteristics of its consumers. Relaxing constraints on consumer choice and reducing consumers’ costs associated with choosing non-default providers can improve how well performance-based rankings reflect relative provider quality. A necessary and sufficient condition is that providers have the ability to select, at least to some extent, those consumers that have a favourable impact on their future performance whenever they face excess demand. In the case of hospitals, this would mean that doctors need to be allowed to choose patients to some degree based on the state of their health.
4 Welfare

Policy measures employed to increase the informativeness of league tables in equilibrium steady state also affect educational outcomes and welfare. Consider the case of uniformly distributed transport costs, i.e. \( F = U [0, \bar{c}] \). The expected number of high results, denoted by \( R \), is given by

\[
R \equiv \alpha + 1 - \beta + P (H^G) (\beta - \alpha)
\]

where \( P (H^G) \) denotes the probability that the high-ability student attends the good school in steady state. If the high-ability student is at the good school, only one high result is realised; whereas if the high-ability student attends the bad school, the expected number of high results is \( R = \alpha + 1 - \beta \). If student ability and school quality are complements, i.e. \( \beta \geq \alpha \), then out of the two possible student allocations, the one in which the high-ability student attends the good school has a larger expected number of high results. If ability and quality are substitutes, i.e. \( \alpha \geq \beta \), the student allocation in which the high-ability student attends the bad school has a larger number of high results.

Lower transport costs, in the sense of a negative FOSD shift in the distribution for transport costs, or an increase in families’ valuations of high results, increase the equilibrium level of mobility and hence raise the probability that the high-ability student attends the good school. If student ability and school quality are complements (substitutes), these changes increase (decrease) the expected number of high results.

A higher capability of schools to select students based on ability (\( p \)) increases the likelihood that the high-ability student attends the good school in two ways, which reinforce each other, given student ability and quality are complements. Not only does an increase in \( p \) increase the equilibrium level of mobility, it also increases the likelihood that the high-ability student is at the good school at any given level of mobility. Given the assumption of complements, this causes the expected number of high results to increase.

A utilitarian social planner will weigh up the number of expected results, \( R \), against any expected expenditure on transport costs, denoted by \( TC \). Hence total welfare, denoted by \( W \), will be given by

\[
W \equiv vR - TC
\]

A negative FOSD shift in the distribution for transport costs will decrease welfare unless
mobility is perfect, i.e. unless \( m^{1\ast} = 1 \). This is because a negative FOSD shift in the distribution for transport costs leads to a higher equilibrium level of mobility, unless mobility has already reached its upper bound at 1. As the equilibrium level of mobility increases, expected expenditure on transport costs increases. Families pay for transport more frequently and conditional on paying for transport, they pay higher amounts. Further, this increase in expenditure is so large that it outweighs the increase in the number of expected results which arises if student ability and school quality are complements. If families are perfectly mobile, \( m^{1\ast} = 1 \), they are willing to pay any possible realisation of transport costs, conditional on their child being accepted at the winning school. In this case, a negative FOSD shift in transport costs does not change the equilibrium level of mobility, but it decreases expected expenditure on transport, causing welfare to increase.

**Proposition 3 [Welfare]:**

Let \( F = U [0, \bar{c}] \).

1. If student ability and school quality are complements, i.e. \( \beta \geq \alpha \), then the expected number of high results increases
   - (a) with a negative shift in FOSD of the distribution for transport costs \( F (\cdot) \).
   - (b) with an increase in the valuation of high results \( v \).
   - (c) with an increase in schools’ capability to select students based on ability, \( p \).

2. If equilibrium level of mobility is strictly less than 1, total welfare decreases with a negative shift in FOSD of the distribution for transport costs \( F (\cdot) \).

A policymaker could invest in better infrastructure or impose fewer constraints on families’ choice between schools in order to improve informativeness. If school quality and student ability are complements, the resulting increase in sorting between students and schools would increase total academic outcomes. However, even if the cost of the policy intervention is ignored, welfare would decrease because the additional expenditure on transport is too large in relation to the gain in academic achievement. This stark result is alleviated to some degree if the social planner is assumed to value educational achievement more than families, due to the positive externalities generated by education for society. In this case, the valuation of educational outcomes could be sufficiently large such that a FOSD shift in the distribution for transport costs increases welfare in the presence of complementarities between ability and quality, because the increase in educational outcomes is weighted more than the increase in expenditure on transport costs.

It is important to bear in mind that this analysis is only a partial equilibrium analysis in
the sense that it does not take into account schools’ investment in quality. In a market, where schools are rewarded for attracting students, improving informativeness will make families’ demand for schools more sensitive to schools’ quality and hence increase schools’ incentives to invest in their quality. Then welfare may increase with a policy intervention such as improved infrastructure, because the increase in educational outcomes is sufficiently large due to the increase in schools’ quality.

5 Convergence

The interaction of mobility and informativeness has the interesting property that, from the point in time at which rankings become available, informativeness increases and converges over time. The learning process describes the evolution of informativeness over time, if each generation acts optimally given the mobility of previous generations. The limit of the learning process is the steady state informativeness associated with the smallest equilibrium level of mobility.

**Proposition 4 [Convergence]**

*Given the first ranking is based on educational outcomes of students in generation 0, the dynamic learning process, denoted by the sequence \( \{I_t : t = 0, 1, \ldots \} \) is characterised by informativeness of generation 0:*

\[ I_0 = 0 \]

*and by the recurrence equation linking informativeness of subsequent generations:*

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

*for all \( t \geq 1 \).

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

\[ m^{1*} = F (V \cdot I^{1*}) . \]
Studying how the learning process evolves over time gives further insights into the positive relationship between conjectured mobility and informativeness in steady state (see Proposition on Equilibrium). As an example, compare the inference made by the first two generations with access to rankings. Generation 1 is the first generation with access to rankings and observes schools’ performance with students of the previous generation. Further, generation 1 infers that in the previous generation each student attended his local school, as the previous generation’s posterior beliefs about school quality equal their symmetric prior beliefs. Generation 1 believes that the winner with students in the previous generation is the good school with probability:

\[ P(W_0^G) = \frac{1}{2} (1 + x) \geq \frac{1}{2} \]

Hence,

\[ I_1 = x \]

Note that \( x \) is the intrinsic quality advantage the good school has over the bad school. The higher is \( x \), the more informative is the ranking that results from a random assignment of students to schools. If educational outcomes are only determined by school quality, i.e. \( x = 1 \), then \( I_1 = 1 \).

Generation 2 observes schools’ performance with students in generation 1. Taking as given that generation 1 had a mobility level equal to \( \hat{m}_1 \), generation 2 can infer the probability that the high-ability student in generation 1 attended the school that won with students in generation 0, i.e. \( P(H_1^W|W_0) \). Based on \( P(H_1^W|W_0) \), generation 2 can infer the probability that the good (bad) school won with students in generation 1, conditional on having won with students in generation 0.

\[
P(W_1^G|W_0^G) = P(W_1^G|H_1^W, W_0^G) P(H_1^W|W_0^G) + P(W_1^G|L_1^W, W_0^G) P(L_1^W|W_0^G)
= \frac{(1 + x)}{2} + \frac{1}{2} \hat{m}_1 (2p - 1) (1 - x)
\]

Similarly,

\[
P(W_2^B|W_1^B) = \frac{(1 - x)}{2} + \frac{1}{2} \hat{m}_2 (2p - 1) (1 - x)
\]
Schools’ performance with students in generation 1 is independent of their performance with students in generation 0, if students in generation 1 apply to their local school independent of schools’ performance, i.e. if \( \hat{m}_1 = 0 \), or if schools have to randomise between applications, i.e. \( p = \frac{1}{2} \). Assuming \( p > \frac{1}{2} \), the higher is \( \hat{m}_1 \), the more likely the best-performing school with students in generation 0 will also be the best-performing school with students in generation 1. I will refer to this effect that higher conjectured mobility has on the conditional probability of winning as the reinforcement effect.

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

\[
P \left( W_1^G \right) = P \left( W_1^G | W_0^G \right) P \left( W_0^G \right) + P \left( W_1^G | W_0^B \right) P \left( W_0^B \right)
\]

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

Hence,

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

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

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

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

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

The convergence property of the learning process implies that there are no long term
Figure 5: Impulse Response Function of Informativeness;
\[ F^1(\alpha (1 - p) + \beta p) I_t = I_t \text{ for all } t \geq 0 \]
\[ F^2(\alpha (1 - p) + \beta p) I_t = 1.8I_t \text{ for } 0 \leq t \leq 2 \text{ and } F^2 = F^1 \text{ for } t \geq 3 \]
effects of a short-lived policy intervention, such as a temporary subsidy to transport costs or a limited time period in which parents can choose between schools more freely. Imagine a short-lived policy intervention in the form of a negative FOSD shift in the distribution for transport costs for one generation only. Figure 5 shows two possible trajectories of informativeness over time; one trajectory is based on the assumption of an unchanged distribution for transport costs while the other trajectory is based on a setting which deviates from the former in that the second generation faces a distribution which is first-order stochastic dominated by the original distribution.

However, a social planner who evaluates welfare as a discounted stream of each generation’s payoffs less the costs of the subsidy may still find such a short-lived policy intervention worthwhile, if student ability and school quality are complements. The intervention can only be costly if the introduction of the subsidy changes families’ application decisions such that in the generation eligible for the subsidy both families apply to the winning school instead of their local school and only if the non-local student gets accepted. If these condition hold, the likelihood that the high-ability in this generation attends the good school increases and leading to a higher sum of families’ expected payoffs. For the generation affected by the subsidy the benefit will not outweigh the costs. Families would only alter their application decisions, if the subsidy reduced the benefits. However, subsequent generations conjecture the mobility of the generation affected by the subsidy to be higher and hence mobility, informativeness and the degree of sorting will be higher than without a subsidy, potentially outweighing short-term costs.
6 Conclusion

This paper studies what determines the accuracy with which school league tables rank schools according to their quality, taking into account that schools’ performance depends on students’ ability. In a dynamic environment, the match between students and schools impacts on schools’ performance and schools’ performance in turn affects application decisions and hence the match between students and schools. Comparative static results about the informativeness of league tables are derived in an equilibrium steady state. In an equilibrium steady state, the information conveyed by league tables to families is constant over time and each generations best responds to this information. Lowering transport costs increases the equilibrium level of informativeness, as long as schools are to some extent able to select applicants based on their ability. Allowing schools greater freedom to select applicants increases the equilibrium level of informativeness, if student ability and school quality are complements. Higher informativeness increases educational outcomes, if ability and quality are complements. Higher informativeness also increases the expected expenditure on transport costs, causing overall welfare to decrease if transport costs are uniformly distributed.

The outcome of this dynamic learning process about school quality is analysed under the assumption that school quality is fixed. If schools were rewarded based on their popularity, higher informativeness implies that the demand for schools would be more sensitive to their quality and hence schools would face stronger incentives to improve their quality. An avenue for future research is to incorporate schools’ quality investment decisions into the model in order to gain further insights into the complex effects of market-based reform.

Findings in this paper apply to other sectors in which providers’ performance indicators are influenced by the characteristics of their consumers. How informative providers’ performance indicators are depends on the freedom with which consumers can choose between providers and also the freedom with which providers can choose between those consumers that demand their service. However, higher informativeness does not necessarily lead to welfare improvements with fixed provider quality.

7 Bibliography

References

Department of Economics, University of Bristol, UK, June 2010.


Appendix

8.1 Notation

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = h, l$</td>
<td>result</td>
</tr>
<tr>
<td>$q = G, B$</td>
<td>school quality</td>
</tr>
<tr>
<td>$a = H, L$</td>
<td>student ability</td>
</tr>
<tr>
<td>$m_t \in [0, 1]$</td>
<td>mobility of generation $t$</td>
</tr>
<tr>
<td>$p \in \left[ \frac{1}{2}, 1 \right]$</td>
<td>probability school with two applications selects $H$ student</td>
</tr>
<tr>
<td>$v \geq 0$</td>
<td>valuation for high result</td>
</tr>
<tr>
<td>$V \equiv v(\alpha + p(\beta - \alpha))$</td>
<td>expected payoff from attending good school</td>
</tr>
<tr>
<td>$W^q_t$</td>
<td>winner with students in generation $t$ is school of quality $q$</td>
</tr>
<tr>
<td>$a^q_t$</td>
<td>student of ability $a$ who attends school that won with students in generation $t-1$</td>
</tr>
<tr>
<td>$d^a_t(\text{nonlocal} W^q_{t-1})$</td>
<td>ability $a$-student's (non)local school won with students in generation $t-1$</td>
</tr>
<tr>
<td>$\alpha = P(h</td>
<td>L^G)$</td>
</tr>
<tr>
<td>$\beta = P(l</td>
<td>H^B)$</td>
</tr>
<tr>
<td>$x \equiv P(W^G</td>
<td>L^G) = \frac{\alpha + \beta}{2}$</td>
</tr>
<tr>
<td>$\alpha \geq \beta$</td>
<td>substitutes between student ability and school quality</td>
</tr>
<tr>
<td>$\beta \geq \alpha$</td>
<td>complements between student ability and school quality</td>
</tr>
</tbody>
</table>
8.2 Proof Proposition 1 [Equilibrium]

(1) Mobility in terms of informativeness

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

\[
C = v (\alpha (1 - p) + p\beta) I_t
\]

\( C \) is the maximum level of transport costs families are willing to pay to send their child to the winning school given it is their non-local school. Their realised transport costs lie below this with probability

\[
m_t = F (v (\alpha (1 - p) + p\beta) I_t).
\]

(2) Informativeness in terms of mobility

Denote the event that the good school wins by \( W^G \) and the event that the bad school wins by \( W^B \). Define the state of the system realised in period \( t \) by \( W^G_t \) and \( W^B_t \) respectively.

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

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

(2) The probability that the good school wins with the high-ability student (against the bad school with the low-ability student) is \( P \left( W^G_t \mid H^G_t \right) = 1 \) and the probability that the
good school wins with the low-ability student (against the bad school with the high ability student) is $P \left( W_t^G | L_t^G \right) = x$. (See production process in section (2.2)).

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

$$P \left( W_t^G | W_{t-1}^G \right) = P \left( W_t^G | H_t^G \right) P \left( H_t^{W_{t-1}} \right) + P \left( W_t^G | L_t^G \right) P \left( L_t^{W_{t-1}} \right)$$

$$= \left( \frac{1}{2} (1 - m + mp) \right) + x \left( \frac{1}{2} (1 + m - mp) \right)$$

and

$$P \left( W_t^B | W_{t-1}^B \right) = (1 - P \left( W_t^G | L_t^G \right)) P \left( H_t^{W_{t-1}} \right) + (1 - P \left( W_t^G | H_t^G \right)) P \left( L_t^{W_{t-1}} \right)$$

$$= (1 - x) \left( \frac{1}{2} (1 - m + mp) \right) + 0 \left( \frac{1}{2} (1 + m - mp) \right)$$

The transition matrix $T$ is given by:

$$T = \begin{pmatrix}
P \left( W_t^G | W_{t-1}^G \right) & P \left( W_t^B | W_{t-1}^G \right) \\
P \left( W_t^G | W_{t-1}^B \right) & P \left( W_t^B | W_{t-1}^B \right)
\end{pmatrix}$$

The process is a time-homogeneous Markov chain and its unique stationary distribution is characterised by the row vector $\left( P \left( W^G \right) P \left( W^B \right) \right)$ that satisfies both

$$\left( P \left( W^G \right) P \left( W^B \right) \right) = \left( P \left( W^G \right) P \left( W^B \right) \right) T$$

and

$$P \left( W^G \right) + P \left( W^B \right) = 1$$

if and only if $x < 1$ and either $p < 1$ or $m < 1$ or both. Further, these conditions are sufficient for the transition probabilities to converge to the stationary distribution:

$$\lim_{k \to \infty} T^k = 1 \left( P \left( W^G \right) P \left( W^B \right) \right)$$

where 1 is the column vector with all entries equal to 1.
The stationary distribution is characterised by

\[ P(W_G) = \frac{1}{2} + \frac{x}{2(1 - m (1 - x) (2p - 1))}. \]

If \( x = 1 \) or if \( p = 1 \), \( m = 1 \) but \( x > 0 \), the state \( W_G \) is an absorbing state and I assign \( P(W_G) = 1 \). If \( p = 1 \), \( m = 1 \) and \( x = 0 \) then both \( W_G \) and \( W_B \) are absorbing states and I assign the distribution over states when students are first randomly allocated to schools: \( P(W_G) = \frac{1}{2} \).

So

\[ P(W_G) = \frac{1}{2} + \frac{x}{2(1 - m (1 - x) (2p - 1))} \]

for all parameter configurations.\(^{17}\)

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

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

By symmetry of prior beliefs,

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

By definition of informativeness,

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

Further,

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

and

\[ \frac{\partial^2 I(\hat{m})}{\partial^2 \hat{m}} = \frac{2 (1 - \hat{m} (1 - x) (2p - 1))(1 - x)^2 (1 - 2p)^2}{(1 + \hat{m} (1 - x) (1 - 2p))^4} \geq 0 \]

as \( 0 \leq x \leq 1 \) and \( \frac{1}{2} \leq p \leq 1 \) and \( 0 \leq \hat{m} \leq 1 \).

\(^{17}\)Take the limit as \( x \) tends to 0, then take the limit as \( m \) tends to 1.
(3) Equilibrium

By definition of equilibrium,

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

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

\[ m = F(VI(\hat{m})) \]

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

\[ m^* = F(VI(m^*)) \]

(4) Existence of equilibrium

An equilibrium level of mobility satisfies condition (3) in part (3). Let \( G(\hat{m}) \equiv F(VI(\hat{m})) \). \( G \) is a continuous function and \( G : [0,1] \rightarrow [0,1] \). By Brouwer’s fixed point theorem there exist an \( m^* \) such that \( G(m^*) = m^* \).

8.3 Proof Example: Equilibrium Steady States with Uniform distribution

By part 3) of the proposition on equilibrium, an equilibrium level of mobility satisfies condition (3). Given \( F \) is Uniform on \([0,v_c]\) and \( I(\hat{m}) \) as in equation (2), then \( F(VI(\hat{m})) \) is given by:

\[
F(VI(\hat{m})) = \begin{cases} 
\frac{1}{\hat{m}} \left( \frac{(\alpha(1-p)+p\beta)x}{1-\hat{m}(1-x)(2p-1)} \right) & \text{if } \frac{1}{\hat{m}} \left( \frac{\alpha(1-p)+p\beta)x}{1-\hat{m}(1-x)(2p-1)} \right) \leq 1 \\
1 & \text{if } \frac{1}{\hat{m}} \left( \frac{\alpha(1-p)+p\beta)x}{1-\hat{m}(1-x)(2p-1)} \right) > 1
\end{cases}
\]

\( m^* \) is the solution to a quadratic equation if this solution lies in the interval \([0,1]\):

\[-((m^*)^2 (1 - x) (2p - 1) + m^*) \bar{c} - (\alpha (1 - p) + p\beta) x = 0\]

The candidate solutions to this quadratic equation are:

\[
m^*_i = \frac{1 - \sqrt{1 - 4x \left( \frac{p\beta+(1-p)\alpha}{\bar{c}} \right) (1 - x) (2p - 1)}}{2 (1 - x) (2p - 1)}
\]
and

\[ m_{II}^* = \frac{1 + \sqrt{1 - 4x \left( \frac{p\beta + (1-p)\alpha}{\tau} \right) (1-x) (2p-1)}}{2 (1-x) (2p-1)}. \]

Note that as \( p \to \frac{1}{2} \) both \( m_{I}^* \) and \( m_{II}^* \) tend to \( \frac{x^2}{\tau} \) and as \( x \to 1 \) both \( m_{I}^* \) and \( m_{II}^* \) tend to \( \frac{1}{\tau} \).

\( m_{I}^* \) and \( m_{II}^* \) are real if and only if

\[ 1 - 4x \left( \frac{p\beta + (1-p)\alpha}{\tau} \right) (1-x) (2p-1) \geq 0 \]

\[ \iff 4x (1-x) (2p-1) (p\beta + (1-p)\alpha) \leq \tau \] (4)

If \( m_{I}^* \) is real and finite (condition 4 satisfied), \( m_{I}^* \geq 0 \) if and only if

\[ 1 - \sqrt{1 - 4x \left( \frac{p\beta + (1-p)\alpha}{\tau} \right) (1-x) (2p-1)} \geq 0 \]

\[ \iff 4x \left( \frac{p\beta + (1-p)\alpha}{\tau} \right) (1-x) (2p-1) \geq 0 \]

which holds for all parameter configurations.

If \( m_{I}^* \) is real (condition 4 satisfied), \( m_{I}^* \leq 1 \) if and only if

\[ \frac{1 - \sqrt{1 - 4x \left( \frac{p\beta + (1-p)\alpha}{\tau} \right) (1-x) (2p-1)}}{2 (1-x) (2p-1)} \leq 1 \]

\[ \iff \sqrt{1 - 4x \left( \frac{p\beta + (1-p)\alpha}{\tau} \right) (1-x) (2p-1)} \geq 1 - 2 (1-x) (2p-1) \] (5)

(5) holds if \( (1-x)(2p-1) \geq \frac{1}{2} \). This is because the expression on the LHS of (5) is positive given (4) is satisfied and the expression on the RHS of (5) is negative if \( (1-x)(2p-1) \geq \frac{1}{2} \). If \( (1-x)(2p-1) < \frac{1}{2} \), then (5) if and only if
\[ 1 - 4x \left( \frac{p\beta + (1 - p)\alpha}{c} \right) (1 - x) (2p - 1) \geq 1 - 4(1 - x)(2p - 1) + 4(1 - x)^2(2p - 1)^2 \]
\[ \Leftrightarrow \frac{x(p\beta + (1 - p)\alpha)}{1 - (1 - x)(2p - 1)} \leq \bar{c} \]

If \((1 - x)(2p - 1) < \frac{1}{2}\), then condition (5) implies condition (4):

\[ 4x(1 - x)(2p - 1)(p\beta + (1 - p)\alpha) \leq \frac{x(p\beta + (1 - p)\alpha)}{1 - (1 - x)(2p - 1)} \]
\[ \Leftrightarrow 4(1 - x)(2p - 1)(1 - (1 - x)(2p - 1)) \leq 1 \]

The LHS of (6) increases in \((1 - x)(2p - 1)\) if and only if \((1 - x)(2p - 1) \leq 1\). Therefore, if the LHS of (6) holds at \((1 - x)(2p - 1) = \frac{1}{2}\) then it holds for all \((1 - x)(2p - 1) < \frac{1}{2}\). In fact, (6) is satisfied at \((1 - x)(2p - 1) = \frac{1}{2}\).

Hence \(m_*^I \in [0, 1]\) if and only if either

\[ (1 - x)(2p - 1) \geq \frac{1}{2} \text{ and } 4x(1 - x)(2p - 1)(p\beta + (1 - p)\alpha) \leq \bar{c} \] (7)

or

\[ (1 - x)(2p - 1) < \frac{1}{2} \text{ and } \frac{x(p\beta + (1 - p)\alpha)}{1 - (1 - x)(2p - 1)} \leq \bar{c}. \] (8)

If \(m_*^{II}\) is real (condition 4 satisfied), then \(m_*^{II} \geq m_*^I\). Hence \(m_*^{II} \geq 0\).

If \(m_*^{II}\) is real (condition 4 satisfied), then \(m_*^{II} \leq 1\) if and only if

\[ \sqrt{1 - 4x \left( \frac{p\beta + (1 - p)\alpha}{c} \right) (1 - x) (2p - 1)} \leq 1 \]
\[ \Leftrightarrow \sqrt{1 - 4x \left( \frac{p\beta + (1 - p)\alpha}{c} \right) (1 - x) (2p - 1)} \leq 2(1 - x)(2p - 1) - 1 \] (9)

If \((1 - x)(2p - 1) \leq \frac{1}{2}\) then (9) does not hold (by similar reasoning as above). If \((1 - x)(2p - 1) > \frac{1}{2}\) then (9) holds if and only if
1 - 4x \left( \frac{p\beta + (1 - p)\alpha}{c} \right) (1 - x) (2p - 1) \leq 1 - 4(1 - x)(2p - 1) + 4(1 - x)^2 (2p - 1)^2

\Leftrightarrow \frac{x(p\beta + (1 - p)\alpha)}{1 - (1 - x)(2p - 1)} \geq \bar{c}

Hence \( m^*_{II} \in [0, 1] \) if and only if

\((1 - x)(2p - 1) > \frac{1}{2} \) and \( 4x(1 - x)(2p - 1)(\beta(2p - 1) + 2x(1 - p)) \leq \bar{c} \leq \frac{x(p\beta + (1 - p)\alpha)}{1 - (1 - x)(2p - 1)} \) \hspace{1cm} (10)

Another candidate solution is \( m^* = 1 \) as \( F(VI(\hat{m})) \) is bounded above by 1.

\( m^* = 1 \) if and only if

\[ F(VI(1)) \geq 1 \]

\Leftrightarrow \frac{(\alpha(1 - p) + p\beta)x}{1 - (1 - x)(2p - 1)} \geq \bar{c} \hspace{1cm} (11)

By conditions (7), (8), (10) and (11) it follows that:

a) if \( m^*_{II} \in [0, 1] \) then \( m^*_I \in [0, 1] \).

b) if \( m^*_I \in [0, 1] \) then \( m^* = 1 \).

So necessary and sufficient for \( m^*_I \in [0, 1], m^*_{II} \in [0, 1] \) and \( m^* = 1 \) is \( m^*_{II} \in [0, 1] \). If \( m^*_I \notin [0, 1] \) then either \( m^*_I \in [0, 1] \) or \( m^* = 1 \) but not both.

### 8.4 Proof Proposition 2 [Comparative Statics]:

Using conditions (2) and (1), an equilibrium level of informativeness satisfies

\[ I = x + F(VI)(1 - x)(2p - 1)I \] \hspace{1cm} (12)

Let

\[ Z(I) \equiv x + F(VI)(1 - x)(2p - 1)I \]

Denote by \( I^{1*} \) the smallest equilibrium level of informativeness satisfying condition 12.
Note that $Z(I,t) : [0,1] \times T \rightarrow [0,1]$, where $T$ is a partially ordered set. For all $t \in T$, $Z$ is continuous but for upward jumps. In addition $Z$ is monotone nondecreasing in $t$ for all $I \in [0,1]$. After Corollary 1, (p.??), in Milgrom and Roberts (1994), (henceforth MR), the function $I^{1*}(t)$ is monotone nondecreasing in $t$.

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

Consider

$$Z(F,I) \equiv x + (1-x)(2p-1)F(VI)I$$

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

$$Z(\bar{F},I) - Z(F,I) = \left[\bar{F}(VI) - F(VI)\right] (1-x)(2p-1)I \geq 0.$$ 

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

By condition 1,

$$m^{1*}(F) = F(VI^{1*})$$

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

$$m^{1*}(F) = F(VI^{1*}(F)) \leq \bar{F}(VI^{1*}(\bar{F})) = m^{1*}(\bar{F})$$

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

(b) Increase in valuation for high results $v$

Consider

$$Z(v,I) \equiv x + (1-x)(2p-1)F(V(v)I)I$$

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

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

34
since \((1 - x)(2p - 1)I \geq 0, I \geq 0, V(v) \geq 0\) and \(F\) is increasing. By MR, \(I^{1*}(v)\) increases in \(v\).

By condition 1,

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

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

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

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

(c) Increase in the impact of school quality on results \(\alpha\) or \(\beta\)

Consider

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

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

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

since \(V(\alpha') \geq V(\alpha) \geq 0, \frac{1}{2} \leq p \leq 1, 0 \leq I \leq 1, 0 \leq F \leq 1\) and \(F\) is increasing. By MR, \(I^{1*}(\alpha)\) increases in \(\alpha\) and a similar argument holds for \(\beta\).

By condition 1,

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

An increase in \(\alpha\) increases \(m^{1*}\):

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

since \(I^{1*}(\alpha') \geq I^{1*}(\alpha) \geq 0, V(\alpha') \geq V(\alpha) \geq 0\) and \(F\) is increasing.

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

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

\[
Z\left(p', I\right) - Z\left(p\right) = I \left(1 - x\right) \left[\left(2p - 1\right) F\left(V\left(p'\right), I\right) - F\left(V\left(p\right), I\right)\right] + 2\left(p' - p\right) F\left(V\left(p'\right), I\right)
\]

Assume complements, i.e. \( \beta \geq \alpha \), then \( Z\left(p'\right) - Z\left(p\right) = 0 \), since \( V\left(p'\right) \geq V\left(p\right) \), \( 0 \leq I \leq 1 \), \( 0 \leq x \leq 1 \), \( \frac{1}{2} \leq p \leq 1 \), \( 0 \leq F \leq 1 \) and \( F \) is increasing. By MR, if \( \beta \geq \alpha \), then \( I^{1*}\left(p\right) \) is increasing in \( p \).

By condition 1,

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

If \( \beta \geq \alpha \), then mobility increases with \( p \):

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

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

### 8.5 Proof Proposition 3 [Welfare]

**Derivation A: Probability high-ability student at good school in steady state**

\( P(\text{H}^G) \)

\[
P\left(\text{H}^G\right) = \frac{1 - (1 - \alpha - \beta) m^{1*}\left(2p - 1\right)}{2 - (2 - \alpha - \beta) m^{1*}\left(2p - 1\right)}
\]

Denote the event that the high-ability student is at the good school by \( H^G \) and the event that the high-ability student is at the bad school by \( H^B \). Further, denote the state realised in period \( t \) by \( H^G_t \) and \( H^B_t \) respectively.

(1) The probability that the good school wins with the high-ability student (against the bad school with the low-ability student), is \( P\left(W^G_t|H^G_t\right) = 1 \) and the probability that the good school wins with the low-ability student (against the bad school with the high ability student) is \( P\left(W^G_t|H^B_t\right) = x \equiv \frac{\alpha + \beta}{2} \). (See production process in section (2.2) ).

(2) The probability that in period \( t \) the high-ability student attends the school that won with students in period \( t - 1 \), denoted by \( P\left(H^{W_{t-1}}_t\right) \), is given by

\[
P\left(H^{W_{t-1}}_t\right) = \frac{1}{2} \left(1 - m + 2pm\right).
\]

36
(see Appendix Proof Proposition on Equilibrium, part 2)

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

$$
P (H_i^G | H_{t-1}^G) = (1 - P (H_{t-1}^W)) P (W_{t-1}^G | H_{t-1}^G) + (1 - P (H_{t-1}^W)) (1 - P (W_{t-1}^G | H_{t-1}^G))
$$

$$
= \frac{1}{2} (1 - m + 2mp)
$$

and

$$
P (H_i^B | H_{t-1}^B) = (1 - P (H_{t-1}^W)) P (W_{t-1}^G | H_{t-1}^B) + P (H_{t-1}^W) (1 - P (W_{t-1}^G | H_{t-1}^B))
$$

$$
= \frac{1}{2} (1 + m - 2mp) x + \frac{1}{2} (1 - m + 2mp) (1 - x)
$$

The transition matrix $T$ is given by:

$$
T (m, p, x) = \begin{pmatrix}
P (H_i^G | H_{t-1}^G) & P (H_i^B | H_{t-1}^G) \\
P (H_i^G | H_{t-1}^B) & P (H_i^B | H_{t-1}^B)
\end{pmatrix}
$$

For similar reasons as in proof of the proposition on equilibrium part 2), this Markov chain has a unique stationary distribution characterised by

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

if either $p < 1$ or $m < 1$ or both. If $p = 1$, $m = 1$ but $x > 0$, the state $H^G$ is an absorbing state and I assign $P (H^G) = 1$. If $p = 1$, $m = 1$ and $x = 0$ then both $H^G$ and $H^B$ are absorbing states and I assign $P (H^G) = \frac{1}{2}$ as students are initially allocated to schools at random.\(^{18}\) Hence,

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

for all parameter combinations.

\(^{18}\)Order of limits matters, first let $p$ tend to 1 then let $m$ tend to 1.
Derivation B: Total expenditure on welfare, \( TC \), for \( F = U \left[ 0, \bar{c} \right] \)

Conditional on both families applying to the same school, this school selects the non-local student with probability \( \frac{1}{2} \). In this case, both students will attend their non-local schools and the expected expenditure on transport costs for each families is given by:

\[
E(c | c < C^*) = \begin{cases} 
\frac{V m^1}{4} & \text{if } m^1 \leq 1 \\
\frac{\bar{c}}{2} & \text{if } m^1 > 1 
\end{cases}
\]

Both families apply to the same school with probability \( m^1 \). Further, \( I^1 \) can be expressed in terms of \( m^1 \) using condition (2). Hence the sum of unconditional expected expenditure for both families is given by:

\[
TC = \begin{cases} 
\frac{1}{2} \frac{V m^1}{1 - m^1 (1 - x)} & \text{if } m^1 \leq 1 \\
\frac{\bar{c}}{2} & \text{if } m^1 = 1 
\end{cases}
\]

Proposition Part 1 (a)

\( R \) depends on \( F \) only through the fixed point \( m_1^* \). As the proposition on comparative statics shows, a negative FOSD shift in \( F \) increases \( m_1^* \). Hence \( R \) increases with a negative FOSD in \( F \) if and only if \( R \) increases with an increase in \( m^1 \).

\[
\frac{\partial R}{\partial m^*_1} = \frac{\partial P \left( H^G \right)}{\partial m^*_1} (\beta - \alpha) \tag{13}
\]

where

\[
\frac{\partial P \left( H^G \right)}{\partial m^*_1} = \frac{2 (2p - 1) x}{[2 (1 - m^*_1 (1 - x) (2p - 1))]^2} \geq 0
\]

for all parameter configurations. \( P \left( H^G \right) \) is given in remark on welfare). Hence, \( R \) increases with a negative FOSD in \( F \) if there are complements, i.e. \( \beta - \alpha \geq 0 \) and decreases if there are substitutes, i.e. \( \alpha - \beta \geq 0 \).
Proposition Part 1 (b)

$R$ depends on $v$ only through the fixed point $m^*_1$. As the proposition on comparative statics shows, an increase in $v$ increases $m^*_1$. By a similar argument as above, $R$ increases with an increase in $v$ if there are complements, i.e. $\beta - \alpha \geq 0$ and decreases if there are substitutes, i.e. $\alpha - \beta \geq 0$.

Proposition Part 1 (c)

$R$ depends on $p$ directly as well as through the fixed point $m^{1*}$. Take $p'$ and $p''$ such that $p'' > p'$.

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

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

direct effect

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

indirect effect

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

$$\frac{\partial R}{\partial p} = (\beta - \alpha) \frac{\partial P (H^G)}{\partial p}$$

where

$$\frac{\partial P (H^G)}{\partial p} = \frac{m^{1*} x}{(1 - m^{1*} (1 - x)(2p - 1))^2} \geq 0$$

for all parameter configurations.

Consequently, the total effect of $p$ is positive if there are complements, i.e. $\beta - \alpha \geq 0$. 

39
Proposition Part 2

If \( \bar{c} \) is sufficiently large such that \( m^{1*} < 1 \) then total welfare \( W \) depends on \( F \) only through the fixed point \( m^{1*}_1 \). As the proposition on comparative statics shows, a negative FOSD shift in \( F \) increases \( m^{1*} \). Hence \( W \) increases with a negative FOSD in \( F \) for \( m^{1*} < 1 \) if and only if \( W \) increases with an increase in \( m^{1*} \).

\[
\frac{\partial W}{\partial m^{1*}} = v (\beta - \alpha) \frac{2x (2p - 1)}{[2 (1 - m^{1*} (1 - x) (2p - 1))]^2} - \frac{Vx}{2} \frac{1}{(1 - m^{1*} (1 - x) (2p - 1))^2}
\]

\[
= \frac{2x (v (\beta - \alpha) (2p - 1) - V)}{[2 (1 - m^{1*} (1 - x) (2p - 1))]^2}
\]

This is clearly negative for substitutes. Consider complements, i.e. \( \beta - \alpha \geq 0 \),

\[
\frac{\partial W}{\partial m^{1*}} \leq 0
\]

\[
\iff \beta - \frac{\beta}{p} \leq \alpha
\]

(14)

which holds for all parameter configurations. Hence total welfare decreases if \( \bar{c} \) is sufficiently large such that \( m^{1*} < 1 \).

If \( \bar{c} \) is sufficiently small such that \( m^{1*} = 1 \) then total welfare \( W \) depends on \( F \) directly and through the fixed point \( m^{1*}_1 \). This is because the boundary of the support of \( F \), denoted by \( \bar{c} \), enters total costs directly. A negative shift in FOSD of \( F \), i.e. a decrease in \( \bar{c} \), leaves \( m^{1*} = 1 \) unchanged. Therefore the valuation of expected results \( vR \) remains unchanged. However, a negative shift in FOSD of \( F \), i.e. a decrease in \( \bar{c} \), lowers total costs \( TC \) and hence total welfare increases.

8.6 Proof Proposition 4 [Convergence]

Learning process

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

Expressing (15) in terms of informativeness, where \( I_t \equiv 2P(W^G_{t-1}) - 1 \), gives:

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

Using (1), then (16) is equivalent to:

\[ I_t = x + F(VI_{t-1}) (1 - x) (2p - 1) I_{t-1} \equiv Z(I_{t-1}) \]

Learning process converges 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 \( I'_t > I_t \)

\[ Z(I'_t) - Z(I_t) = (1 - x) (2p - 1) \left\{ \left[ F(VI'_t) - F(VI_t) \right] I_t + F(VI'_t) \left[ I'_t - I_t \right] \right\} \geq 0 \]

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

Using \( I_0 = 0 \), equation 15, 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 \( \overline{I} \), where

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

If \( \overline{I} \) was not an upper bound then for some \( I' \leq \overline{I} \) it would be true that \( Z(I') > Z(\overline{I}) \).
But $Z$ is increasing and $Z(\overline{I}) = \overline{I}$, so this is a contradiction. Further, if $\overline{I}$ was not the least upper bound then for some $I^{**}$, $I^{**} < \overline{I}$. 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 $\overline{I}$ was the smallest fixed point of $Z$.

By definition 1,

$$m_t = F(VI_t)$$

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

42