Ability, Length of Schooling, Growth and the Distribution of Human Capital

Christy Redmon & Robert Tamura

April 12, 2007

Abstract

A Mincer model of human capital with ability differences is presented. The optimal length of schooling by ability class is characterized, and the importance of school district composition for growth and distribution is presented.

INTRODUCTION

In this paper we present a model with heterogeneous ability among otherwise identical individuals. Adults care about their own consumption, their retirement leisure and the human capital of their children. We modify the model of Glomm and Ravikumer (1992) and Tamura (2001) in order to produce realistic schooling choice by different ability groups. The model is calibrated to fit the behavior of schooling choice by ability group observed in NLSY. The introduction of diminishing returns to schooling length allows for the model to produce both the observed heterogeneity in schooling length, but also the constancy of estimated rates of return per year of schooling. A standard Mincerian earnings regression would produce a rate of return estimate of 12.6 percent per additional year of schooling, despite heterogeneity in schooling length that ranges between 4.75 years and 20.7 years. These results are in complete agreement with Willis and Rosen (1979). Ability differences produces comparative advantage in schooling duration.

We then examine the question of school district composition. In a world where relative teacher quality is more important for human capital accumulation, Tamura (2001), there is a strong desire on the part of the highest ability individuals to segregate from the lower ability individuals. Thus we look at different possible levels of stratification of ability into different school districts. We examine economies with a single district, 2, 4, 5, 10, 20, 50 and 100 (full stratification) districts. Each configuration is symmetric in the population size of a district, and there is perfect sorting. For example in the 20 school district case, the lowest ability district is
composed of the first through fifth percentiles of the ability distribution and the top ability district is composed of the 96th to the 100th percentile. We show that in the long run, the economy grows fastest under full stratification. However during the short run many individuals strongly prefer more stratification. In the solutions we find that a world with 50 districts is tough to beat. As many as 80 percent of the population prefers this configuration to complete stratification. Our work on school district composition effects is quite in the spirit of Benabou (1996a,b) and Fernandez and Rogerson (1995, 1996, 1997).

The paper is organized as follows. The next section introduces the basic model. We show, through comparative statics that higher ability individuals choose longer schooling lengths. We document the fit of the model with the NLSY data. The choice of schooling choice is independent of the school district configuration, which makes the problem tractable. The following section presents the long run stationary distribution of human capital under the assumption of a single school district. We also show the essential constancy of the rate of return to schooling, but that Willis and Rosen (1979) comparative advantage holds. The next section introduces multiple school districts. We can show that for full stratification, the long run relative human capital distribution is more unequal than the single school district economy. We also show that higher balanced growth rates occur under full stratification. There exists an interesting assignment question early on, if there is not much inequality in human capital, then there exists excess demand for hiring the top human capital individuals as teachers. We use a rationing scheme and the fact that teacher quality and student quality are complements to assign the best teachers to the best students. Over time as inequality increases, it becomes possible to give all students access to the best teachers, albeit at different class sizes.

MODEL

In this section we detail the model that forms the heart of the paper. We introduce ability differences across individuals. We specify ability in a manner that is invariant to scale, and that influences the length of time an individual spends in school accumulating human capital. The model is consistent with a Mincerian earnings regression. The model is then solved using numerical methods and calibrated to fit contemporary US data. The importance of school districts are shown in the following section.

Assume that there are a continuum of individuals in the population with normalized population of 1. These individuals differ in their innate ability, $a$. An individual’s human capital depends on their innate ability, their schooling, their parent’s human capital and the quality of schooling they received. Consider the following
model of human capital accumulation for individual $i$:

$$h_{it} = A \exp(\exp(a_i^\theta)\beta S_{it}^\eta)h_{it-1} \left(\frac{1}{\text{class}_{t-1}}\right)^{\varepsilon \nu} \left(\frac{\bar{h}_t}{h_{it-1}}\right)^{(1-\varepsilon)\nu}$$  \hspace{1cm} (1)

where $a_i$ is the ability of individual $i$ as described by the percentile score, $\theta > 0$, $S_{it}$ is the choice of schooling length of the individual, $1 > \eta > 0$, $\text{class}_{t}$ is the average class size attended by the individual, and $\bar{h}_t$ is the average human capital of the teachers of individual $i$, and $\beta$ is the return to a skill adjusted year of schooling. A child perfectly inherits the ability level of his parent. There is no regression to the mean in raw ability. We choose this stark assumption to focus on the effects of stratification of school districts by ability. Observe that perpetual growth is possible as all the terms exclusive of parental human capital can produce growth as long as $A$ is sufficiently large. Ignoring the exponential terms, this specification contains a tradeoff between teacher quality and class size as in Tamura (2001). It is a technology that induces diminishing returns to educational expenditures and often induces convergence for $\varepsilon < \frac{1}{2}$. However because there exists ability differences, a single school district world will produce a stationary relative human capital distribution, but not full convergence in levels of human capital. When there are multiple school districts, then there will be no convergence.

A school district with tax base $\bar{h}_{t-1}$ hires $N_{t-1}$ teachers from the human capital distribution in society:

$$\tau_{t-1}\bar{h}_{t-1} = N_{t-1}\bar{h}_t$$  \hspace{1cm} (2)

Now class size is given by \(\frac{1}{N_{t-1}}\) or:

$$\frac{1}{\text{class}_{t-1}} = \frac{\tau_{t-1}\bar{h}_{t-1}}{\bar{h}_t}$$  \hspace{1cm} (3)

Under the assumption that teacher quality is a more important input to accumulation than class size, $\varepsilon < \frac{1}{2}$, the school district wishes to hire the highest human capital individuals to be teachers. This will have important consequences when there are multiple school districts, but for now we ignore this issue.

Assume that the typical individual chooses schooling length, $S_{it}$, retirement age, $R_{it}$, and the level of taxation for the public schooling of their child. Each adult has a single child, and that child has identical ability as the parent.\(^1\) Preferences are

\(^1\)In future editions we would like to examine a world in which there exists regression to the mean. This would clearly induce an ergodic distribution of relative human capital, but for now we examine the starker case of no regression to the mean in ability.
given by:
\[
\max_{\{S_{it}, R_{it}, \tau_t\}} \{ \ln c_{it} + \gamma \ln(T - R_{it}) + \delta \ln h_{it+1} \} \quad (4)
\]

For simplicity assume that \( c_{it} \) is the PDV of lifetime earnings:
\[
c_{it} = \int_{S_{it}}^{R_{it}} h_{it}(\sigma) e^{-r \sigma} d\sigma (1 - \tau_t) \quad (5)
\]

We abstract of life cycle considerations of human capital accumulation from on the job training or learning by doing, hence \( h_{it}(\sigma) = h_{it} \) for all \( \sigma > S_{it} \). Thus consumption can be written as:
\[
c_{it} = \frac{h_{it}}{r} \left[ e^{-r S_{it}} - e^{-r R_{it}} \right] (1 - \tau_t) \quad (6)
\]

Replacing for consumption, using (1) to solve for the value of \( h_{it} \) on the accumulation of \( h_{it+1} \) and collecting \( h_{it} \) terms, and ignoring terms that do not include choice variables produces an optimization problem facing the typical individual:
\[
\max_{\{S_{it}, R_{it}, \tau_t\}} \left\{ \left( 1 + \delta [1 - (1 - \varepsilon) \nu] \right) \ln h_{it} + \ln \left[ e^{-r S_{it}} - e^{-r R_{it}} \right] + \ln (1 - \tau_t) + \gamma \ln(T - R_{it}) + \delta \varepsilon \nu \ln \tau_t \right\} \quad (7)
\]

Finally replacing for \( h_{it} \) as a function of \( S_{it} \) yields:
\[
\max_{\{S_{it}, R_{it}, \tau_t\}} \left\{ \left( 1 + \delta [1 - (1 - \varepsilon) \nu] \right) \exp(a_t^0) \beta S_{it}^\eta + \ln \left[ e^{-r S_{it}} - e^{-r R_{it}} \right] + \ln (1 - \tau_t) + \gamma \ln(T - R_{it}) + \delta \varepsilon \nu \ln \tau_t \right\} \quad (8)
\]

Recall that the individual cares about the class size of their children’s school since it affects the human capital accumulation of their children. The first order condition for the preferred tax rate for individual becomes:
\[
\frac{1}{1 - \tau_t} = \frac{\delta \varepsilon \nu}{\tau_t} \quad (9)
\]

Notice that the preferred tax rate is independent of the ability of the individual, and his desired schooling, or retirement age. Thus there is unanimity amongst individuals on the desired school income tax rate:
\[
\tau_t = \frac{\delta \varepsilon \nu}{1 + \delta \varepsilon \nu} \quad (10)
\]
The optimal retirement age and the duration of schooling are solutions to:

\[
\frac{r e^{-rR_{it}}}{e^{-rS_{it}} - e^{-rR_{it}}} = \frac{\gamma}{T - R_{it}} \tag{11}
\]

\[
\frac{r e^{-rS_{it}}}{e^{-rS_{it}} - e^{-rR_{it}}} = (1 + \delta [1 - (1 - \varepsilon)\nu])\eta \exp(a_i^\theta)\beta S_{it}^{\eta-1} \tag{12}
\]

These two Euler equations can be simplified in order to solve numerically for the optimal \((R_{it}, S_{it})\) pair. Dividing numerator and denominator of the left hand side of the first equation by \(e^{-rR_{it}}\) and the numerator and denominator of the left hand side of the second equation by \(e^{-rS_{it}}\) produces:

\[
\frac{r}{e^{r(R_{it} - S_{it})} - 1} = \frac{\gamma}{T - R_{it}} \tag{13}
\]

\[
\frac{r}{1 - e^{r(S_{it} - R_{it})}} = (1 + \delta [1 - (1 - \varepsilon)\nu])\eta \exp(a_i^\theta)\beta S_{it}^{\eta-1} \tag{14}
\]

We can rewrite the second equation and then substitute from the first equation to produce:

\[
\frac{r e^{r(R_{it} - S_{it})}}{e^{r(R_{it} - S_{it})} - 1} = \frac{\gamma}{T - R_{it}} + r = (1 + \delta [1 - (1 - \varepsilon)\nu])\eta \exp(a_i^\theta)\beta S_{it}^{\eta-1}
\]

\[
\frac{\gamma}{T - R_{it}} = -r + (1 + \delta [1 - (1 - \varepsilon)\nu])\eta \exp(a_i^\theta)\beta S_{it}^{\eta-1} \tag{15}
\]

For a given value of \(S_{it}\), a unique value of \(R_{it}\) is determined. Thus using this value for \(R(S_{it})\) we can solve for the optimal \(S_{it}\). Let \(Q = \exp(a_i^\theta)\), then comparative statics reveals:

\[
\frac{\partial S}{\partial Q} > 0 \tag{16}
\]

\[
\frac{\partial R}{\partial Q} > 0 \tag{17}
\]

Thus greater ability individuals spend more time in school, and retire later in life. Furthermore notice that the length of schooling and retirement age are time independent as long as the life expectation is constant. Numerically solving the model we produce the following figure relating ability to schooling choice and retirement age. The table below presents our parameter valuations in order to solve the model.

**Table 1: Parameter values for numerical solution**

\[^2\text{As long as the right hand side never falls below } \gamma.\]
The values are chosen in order to roughly fit the data on years of schooling by percentile. \( T = 80 \) is the life span of the typical worker. The interest rate at which the individual discounts future earnings is 10 percent. The triple \((\delta, \varepsilon, \nu)\) is chosen in order that the model expenditure share on education, .0566, is close to the national average in the US, of .068. The value of \( \varepsilon = .4 \) is close the value estimated in Tamura (2001) of between .37 and .41. The values of \((A, \theta, \beta, \eta)\) are chosen in order to roughly fit long run growth behavior. Long run growth in the economy under a single school district is .62 percent per year (40 year generation), and 1.25 percent per year (20 year generation). Under full stratification, or 100 school districts, one for each ability percentile, long run growth is 1.5 percent per year (40 year generation), and 3.0 percent per year (20 year generation).

The first figure presents the optimal schooling length as a function of ability. It also includes the average year of schooling by ability percentile, measured from the AFQT, of whites 25 and older from the NLSY. It is clear from the picture that the model produces too much dispersion in years of schooling compared to the data. Part of the problem lies with mandatory schooling. Most states have a minimum age of dropping out of 15 or more. Thus an individual cannot drop out legally until at least 10 years of schooling. That is a binding constraint on the bottom 13 percent of the ability distribution in the model. Requiring one to stay in school for 11 years is binding on bottom 17 percent of the ability distribution. The table below presents the regression of schooling in the NLSY 25 and older white data on the model predictions. We present it for the entire ability distribution, for the ability distribution above the 13 percentile, and above the 17 percentile.\(^3\)

<table>
<thead>
<tr>
<th></th>
<th>full sample</th>
<th>above the 13th percentile</th>
<th>above the 17th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>model predicted schooling</td>
<td>.4106</td>
<td>.4199</td>
<td>.4405</td>
</tr>
<tr>
<td>(</td>
<td>(.0131)</td>
<td>(.0193)</td>
<td>(.0200)</td>
</tr>
<tr>
<td>constant</td>
<td>7.802</td>
<td>7.645</td>
<td>7.281</td>
</tr>
<tr>
<td>(</td>
<td>(.205)</td>
<td>(.317)</td>
<td>(.332)</td>
</tr>
<tr>
<td>N</td>
<td>97</td>
<td>84</td>
<td>81</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>.9105</td>
<td>.8506</td>
<td>.8584</td>
</tr>
</tbody>
</table>

The next graph presents the relationship between retirement age and ability.

\(^3\)There are only 97 ability percentiles in the data. Both the 60th and 83rd percentile cells are empty. The 100th percentile cell is also empty.
Fig. 1.
Fig. 2. Retirement age and ability
SINGLE SCHOOL DISTRICT

Consider now the problem in which all individuals are in a single school district. Furthermore we assume that individuals are treated identically, in other words we exclude the possibility of differential class size and teacher quality within the school district. Hence there is no tracking of individuals by ability. Thus a national school system would be an example of this equilibrium. In this case notice that there will not be human capital convergence, however there will exist a stationary distribution in relative human capital.

To see this, notice that under a single school district, all individuals are schooled in the same classes by the same average teacher quality. From the school district budget constraint, notice that the school district will choose to hire all individuals at the very top. Once the very top class is hired, the district then hires as many of the next highest human capital individuals as possible. This continues until the entire school budget is spent. Notice that all teachers are paid their opportunity wage, i.e. each individual is paid $h_{it}$. Now consider the relative human capital of any individual with any other individual. Without loss of generality, assume we normalize by the highest ability, $\bar{a}$, highest human capital individual, $\bar{h}_t$. Thus relative human capital is given by:

$$
\frac{h_{it+1}}{\bar{h}_{t+1}} = \frac{A \exp(\exp(a_i^\theta)\beta S_{it+1}) h_{it} \left( \frac{1}{\text{class}} \right)^\nu \left( \frac{\bar{a}}{h_{it}} \right)^{(1-\varepsilon)^\nu}}{A \exp(\exp(\bar{a}^\theta)\beta S_{t+1}) \bar{h}_t \left( \frac{1}{\text{class}} \right)^\nu \left( \frac{\bar{h}_t}{\bar{h}_{it}} \right)^{(1-\varepsilon)^\nu}}
$$

$$
\frac{h_{ut+1}}{\bar{h}_{t+1}} = \left( \frac{\exp(\exp(a_i^\theta)\beta S_{it+1})}{\exp(\exp(\bar{a}^\theta)\beta S_{t+1})} \right) \left( \frac{h_{ut}}{\bar{h}_t} \right)^{1-(1-\varepsilon)^\nu}
$$

Under our assumptions of identical treatment in school, and the perfect correlation of parental ability and child ability, the first parenthetical expression above is independent of time. The power on the second parenthetical expression is less than 1, thus there exists a stationary relative human capital distribution given by:

$$
\frac{h_i}{\bar{h}} = \left( \frac{\exp(\exp(a_i^\theta)\beta S_{i}^\theta)}{\exp(\exp(\bar{a}^\theta)\beta S_{t}^\theta)} \right)^{\frac{1}{1-(1-\varepsilon)^\nu}}
$$

While ability differences imply different schooling duration and different levels of human capital, there is no perpetual relative income divergence. Income inequality as expressed as human capital inequality is greater than the wealth inequality. Since higher ability individuals stay in school longer, their present discounted value of
earnings are not as large as the less able individuals. Though they work to an older age, this does not offset the lost early years of work. Recall that consumption is equal to the PDV of earnings, so relative wealth is the same as relative consumption:

\[
\frac{c_t}{\bar{c}_t} = \frac{h_{it}\left(e^{-rS_{it}} - e^{-rR_{it}}\right)}{\bar{h}_t\left(e^{-r\bar{S}} - e^{-r\bar{R}}\right)} \tag{20}
\]

While we do not prove that the lower ability worker has a greater relative wealth than given by earnings, we illustrate it in the numerical solutions below.

One final observation that we make in the numerical solutions. Suppose an econometrician observes earnings, \(h_{it}\), schooling, \(S_{it}\), parental earnings, \(h_{it-1}\), and school quality measured by class size and relative teacher quality, then a Mincerian earnings regression would look like:

\[
\ln h_{it} = \text{constant} + \hat{\beta}S_{it} + \varphi_1 \ln h_{it-1} - \varepsilon \nu \ln \text{class}_{it-1} + (1 - \varepsilon) \nu \ln \left(\frac{\bar{h}_t}{h_{it-1}}\right) \tag{21}
\]

This in fact is the type of regression run in Tamura (2001), except for the term \(\hat{\beta}S_{it}\). Instead of \(S_{it}\), lagged enrollment rates and lagged school year length was used. However notice that the problem with this regression at the state level, or at the individual level is that \(S_{it}\) is a function of ability. Furthermore the endogeneity of schooling cannot be corrected by simply adding an ability measure as an additional regressor. This is because the estimated return to schooling, \(\hat{\beta}\), is in fact equal to \(\exp(a_{\theta}^\beta)\beta S_{it}^{\eta-1}\). Examine the figure below, however for the actual value of \(\exp(a_{\theta}^\beta)\beta S_{it}^{\eta-1}\) for various values of ability. Notice the near constancy of the actual value! Thus while it appears that lower ability workers are mistakenly choosing not to continue their education, they are in fact choosing optimally. There are rising marginal returns to schooling by ability, but they are very small. The bottom 1 percent of the ability distribution receives a marginal return of 12.6 percent, whereas the 100 percentile ability individual receives a marginal return of 12.7 percent. To see the importance of ability and diminishing returns to schooling, we graph the marginal returns for all ability classes if all students were high school graduates and college graduates, respectively. The graph clearly shows the optimality of schooling choice made by each ability class. Forcing all students to become high school graduates produces below reservation returns to those in the bottom quartile of the ability distribution. It is clear that all individual with ability above the quartile receive greater than reservation returns for schooling beyond high school. Forcing all students to become college graduates produces below reservation returns to those at or below the median of the ability distribution. All individuals above the median receive greater reservations returns for schooling be-
returns to marginal year of schooling

Fig. 3. Marginal returns to schooling by ability
Fig. 4. Marginal returns to schooling
yond college. This model presents another demonstration of Willis and Rosen (1979) comparative advantage model of human capital accumulation.

**MULTIPLE SCHOOL DISTRICTS**

Now consider the polar opposite case. Here every ability group is its own school district. In other words there is perfect segregation by ability, human capital, earnings or wealth. The consequences for the lowest ability workers are nonlinear. During the early growth period in societies with low levels of human capital can exhibit falling levels of human capital amongst the lowest ability individuals. In contrast the highest ability workers accumulate human capital at a more rapid rate than if all workers were in a single school district. Eventually the lowest ability workers can hire sufficiently high quality teachers that their human capital accumulates. Furthermore more rapid human capital accumulation of the top workers eventually causes the human capital in the lowest ability groups to exceed that of the identical ability class under a single school district regime. However there appears to be perpetual divergence. Finally since all individuals wish to hire the highest human capital individuals as teachers, there is excess demand for the top ability groups to be teachers. We conjecture that this produces education rents. That is to say the top human capital workers are paid more than their opportunity wage in order to ration them in the economy. We do not prove that the proposed solution contained in this section is in fact an equilibrium. There is evidently work to be done!

In support of our thesis, we observe that individuals desire to maximize the relative teacher quality of the school district. Replacing for class size as above, produces:

\[
\frac{1}{\text{class}_{t-1}} = \frac{\tau_{t-1} \bar{h}_{t-1}}{\bar{h}'_{t-1}}
\]

(22)

Therefore the public contribution to individual \(i\)'s human capital accumulation is given by:

\[
\left(\frac{\tau_{t-1} \bar{h}_{t-1}}{\bar{h}'_{t-1}}\right)^{\epsilon \nu} \left(\frac{\bar{h}'_{t-1}}{h_{it-1}}\right)^{(1-\epsilon)\nu}
\]

\[
= \left(\frac{\tau_{t-1} \bar{h}_{t-1}}{h_{it-1}^{1-\epsilon}\nu} \bar{h}'_{t-1} \right)^{(1-2\epsilon)\nu}
\]

(23)

Under the maintained assumption that \(\epsilon < \frac{1}{2}\), every individual in a school district wishes to maximize the average human capital of teachers. We conjecture that this implies that if the highest human capital individual is in excess demand, then the rich-
est school district bids the most to keep them for their teachers. Furthermore the top human capital individuals can always withhold \( \tau \) proportion of their time/population to teach their own children. In the numerical solutions below we assign the top teachers to the top human capital districts sequentially. Only when the budget constraint of the top school district is exhausted do we allocate any remaining top teachers to the next richest school district. We continue in this manner until all top human capital teachers are assigned, then we move to the next highest quality teacher. Thus the top school district only hires from amongst themselves. The second richest school district hires only top teachers, and generally this continues until eventually all of the top teachers are assigned.

Everything remains the same from the previous section, except for the assumption of a single school district. In particular, nothing about the optimal duration of schooling and retirement age depends on whether there is a single school district or multiple school districts. Hence the optimal \((S_t, R_t)\) pair remains. What does differ is that there are now different school district budget constraints depending on the ability class, \( i \).

Given the above algorithm of teacher assignment, we present the numerical solutions for three cases, single school district, two school districts and 100 school districts, the latter corresponding with each percentile of the ability distribution. It turns out in the numerical solutions that there appears to be stationary relative human capital distributions in all three cases. This occurs, because the highest human capital individual becomes sufficiently rich relative to the entire GDP that they teach all children. Under the assumption that this holds, reconsider the relative human capital of a randomly chosen individual \( i \) with the top human capital individual in the case of perfect segregation by ability:

\[
\frac{h_{it+1}}{\bar{h}_{it+1}} = \frac{A \exp(\exp(a^\theta_i) \beta S^n_{it+1}) h_{it} \left( \frac{1}{\text{class}} \right)^{\epsilon \nu} \left( \frac{\bar{h}}{h_{it}} \right)^{(1-\epsilon) \nu}}{A \exp(\exp(\bar{\alpha}'i) \beta S^n_{it+1}) \bar{h} \left( \frac{1}{\text{class}} \right)^{\epsilon \nu} \left( \frac{\bar{h}}{h_{it}} \right)^{(1-\epsilon) \nu}}
\]

\[
\frac{h_{it+1}}{\bar{h}_{it+1}} = \frac{\left( \frac{\exp(\exp(a^\theta_i) \beta S^n_{it})}{\exp(\exp(\bar{\alpha}'i) \beta S^n)} \right) \left( \frac{h_{it}}{\bar{h}} \right) \left( \frac{\tau h_{it}}{\tau h_{t}} \right)^{\epsilon \nu} \left( \frac{h_{it}}{\bar{h}_{it}} \right)^{-(1-\epsilon) \nu}}{\left( 1 + 2 \epsilon \nu - \nu \right)}
\]

\[
\frac{h_i}{\bar{h}} = Z_i = \left( \frac{\exp(\exp(a^\theta_i) \beta S^n_i)}{\exp(\exp(\bar{\alpha}'i) \beta S^n)} \right)^{\frac{1}{\nu(1-2\nu)}}
\]

(24)

It is evident from above that there is more inequality in the fully stratified world
compared with the fully integrated world. Because teacher quality is the dominant input to human capital accumulation, long run growth for any ability class is greater in the fully stratified economy compared to the other economies. To see this notice that the budget constraint for the random school district becomes:

\[
\tau h_{it} = N \bar{h}_t
\]

\[
\tau \frac{h_{it}}{\bar{h}_t} = \tau Z_i = N
\]  

(25)

Therefore class sizes become constant, although they differ across school districts. The human capital growth rate of the random school district becomes:

\[
\frac{h_{it+1}}{h_{it}} = A \exp(\exp(a_i^\theta) \beta S_i^\eta) \tau^{\nu} \bar{Z}_i^{-(1-\nu)}
\]

\[
= A \exp(\exp(\bar{a}_i^\theta) \beta \bar{S}_i^\eta) \tau^{\nu}
\]  

(26)

Now compare this result with the long run growth rate in a fully integrated economy. This becomes:

\[
\frac{h_{it+1}}{h_{it}} = A \exp(\exp(a_i^\theta) \beta S_i^\eta) \left( \frac{\tau h_i}{\bar{h}_t} \right)^{\nu} \bar{Z}_i^{-(1-\nu)}
\]

\[
= A \exp(\exp(\bar{a}_i^\theta) \beta \bar{S}_i^\eta) \tau^{\nu} \bar{Z}^{\nu}
\]  

(27)

where \( \frac{\bar{h}_i}{h_i} = \bar{Z} < 1 \), the average relative human capital compared to teachers. Thus stratification has higher long term growth than integration. We show below in the numerical solutions, that this holds relative to symmetric school districts (in population) of 2, 4, 5, 10, 20, 25 and 50 districts compared to the 100 in full stratification.

The next figure contains the average human capital by number of districts. We plot the first 11 generations for each configuration.
It is clear from the figure that after the difference between full stratification, 100 districts, and 50 and 25 districts is difficult to see graphically until the 10th and 11th generations. Larger degrees of integration produce much quicker separation. The next figure presents the coefficient of variation by district type.
The next graph that we present asks how many groups prefer their existing school district to the full stratified world. In order to get this answer we compute the human capital for each percentile of the ability distribution under each of the district configurations, 1, 2, 4, 5, 10, 20, 25, 50 and 100. Then we compare the human capital for the x percentile in generation t with the human capital of the same x percentile in generation t in the fully stratified economy. This is not quite the same as asking, given the distribution of human capital, how many would prefer to switch to a different district configuration. However it does reveal how many generations before 50 percentiles would have higher human capital under stratification starting from the same initial human capital distribution.
What is evident is that the economy with 50 districts is strongly preferred to the full stratification world. Over three quarters of the ability distribution prefers this world to one with complete stratification on ability initially. The next figure is in fact a comparison of how many percentiles would prefer to move to fully stratified school districts over time. Thus we compute for every district configuration the time series of human capital for every percentile. In each generation we ask, given the distribution of human capital, how many percentiles would prefer to stay under the current structure compared with switching to full stratification.

Despite the fact that there is rising inequality in all of the economies, approaching their stationary relative income distributions, it is never the case that a majority of percentile groups prefer to switch to complete stratification. There are two configurations in which exactly 50 percentile groups would like to switch, the 10 and 50 district configurations. Now it may be possible to compensate enough of the losers from a switch to full stratification since human capital grows faster under complete segregation. We do not examine that issue here.

Finally we present in the next tables the relative income of quintiles in the stationary distribution for each school district configuration, and the income shares of the quintiles. We normalize by the highest income of the economy. Thus for each configuration we normalize by the income of the 100 percentile individual. Thus we
are not comparing incomes across district configurations.

<table>
<thead>
<tr>
<th>number of school districts</th>
<th>bottom quintile</th>
<th>second quintile</th>
<th>middle quintile</th>
<th>fourth quintile</th>
<th>top quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.0752</td>
<td>.1866</td>
<td>.3404</td>
<td>.5524</td>
<td>.8379</td>
</tr>
<tr>
<td>2</td>
<td>.0049</td>
<td>.0122</td>
<td>.2024</td>
<td>.5524</td>
<td>.8379</td>
</tr>
<tr>
<td>4</td>
<td>.0009</td>
<td>.0138</td>
<td>.0776</td>
<td>.2926</td>
<td>.8379</td>
</tr>
<tr>
<td>5</td>
<td>.0006</td>
<td>.0091</td>
<td>.0552</td>
<td>.2393</td>
<td>.8379</td>
</tr>
<tr>
<td>10</td>
<td>.0006</td>
<td>.0082</td>
<td>.0485</td>
<td>.2066</td>
<td>.7153</td>
</tr>
<tr>
<td>20</td>
<td>.0006</td>
<td>.0078</td>
<td>.0453</td>
<td>.1900</td>
<td>.6563</td>
</tr>
<tr>
<td>25</td>
<td>.0006</td>
<td>.0077</td>
<td>.0445</td>
<td>.1867</td>
<td>.6448</td>
</tr>
<tr>
<td>50</td>
<td>.0006</td>
<td>.0074</td>
<td>.0430</td>
<td>.1802</td>
<td>.6220</td>
</tr>
<tr>
<td>100</td>
<td>.0006</td>
<td>.0073</td>
<td>.0422</td>
<td>.1769</td>
<td>.6107</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>number of school districts</th>
<th>bottom quintile</th>
<th>second quintile</th>
<th>middle quintile</th>
<th>fourth quintile</th>
<th>top quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.0377</td>
<td>.0936</td>
<td>.1709</td>
<td>.2772</td>
<td>.4205</td>
</tr>
<tr>
<td>2</td>
<td>.0031</td>
<td>.0076</td>
<td>.1257</td>
<td>.3431</td>
<td>.5205</td>
</tr>
<tr>
<td>4</td>
<td>.0007</td>
<td>.0113</td>
<td>.0635</td>
<td>.2393</td>
<td>.6852</td>
</tr>
<tr>
<td>5</td>
<td>.0005</td>
<td>.0080</td>
<td>.0483</td>
<td>.2096</td>
<td>.7337</td>
</tr>
<tr>
<td>10</td>
<td>.0007</td>
<td>.0084</td>
<td>.0495</td>
<td>.2110</td>
<td>.7304</td>
</tr>
<tr>
<td>20</td>
<td>.0007</td>
<td>.0086</td>
<td>.0503</td>
<td>.2111</td>
<td>.7292</td>
</tr>
<tr>
<td>25</td>
<td>.0007</td>
<td>.0087</td>
<td>.0504</td>
<td>.2112</td>
<td>.7291</td>
</tr>
<tr>
<td>50</td>
<td>.0007</td>
<td>.0087</td>
<td>.0504</td>
<td>.2112</td>
<td>.7290</td>
</tr>
<tr>
<td>100</td>
<td>.0007</td>
<td>.0087</td>
<td>.0504</td>
<td>.2112</td>
<td>.7290</td>
</tr>
</tbody>
</table>

It is clear that the model produces more inequality than actually observed in the data except for one with a single school district. In future work we will again work to improve the fit of the model with actual income distribution statistics. It is our hope that we would be able to estimate the degree of stratification across countries and states by virtue at looking at the observed income distribution data.

**CONCLUSION**

In this paper we present a model with ability differences across individuals. Under constant ability heterogeneity, i.e. no regression to the mean in ability, we show that there is a unique choice of schooling length for each ability type. We then examine the resulting heterogeneous human capital distribution. We consider 9 different configurations of school districts, complete integration (1) district, 2, 4, 5, 10, 20, 25 50 and complete stratification (100) districts. We derive analytically for
the two limiting cases, one and 100, the relative human capital distribution. The other cases fall in between these two limiting cases. We show that while growth is fastest for the complete stratification economy, it is not likely to be chosen as the configuration for schooling. Only in two cases do 50 percent of the population ever prefers a switch to this configuration relative to their current organizational form.. In future versions we will attempt to find better parameterizations of the model in order to improve on the fit between schooling and ability. We also plan on estimating returns to schooling as corrected by the model’s prediction using non linear regressions. We will be presenting better justification for the assignment of teachers to school districts. The results of the paper are somewhat counterfactual with the evidence of
REFERENCES


