Skill Misallocation and Education Quality

Alexander Ugarov*

January 2, 2018

Abstract

The paper uses PISA and PIAAC datasets to show that academic skills have little connection with occupational choice in countries with low education quality. This observation holds true both for expected occupations of 15-year old students and for current occupations of adult workers. I offer a theoretical explanation for this finding based on a labor signaling model. The model assumes that in countries with lower correlation between skills and occupations firms base their hiring decisions on a noisier signal of human capital. Firms respond to noise by increasing the wage sensitivity to perfectly observable years of education and by reducing the sensitivity to a noisy human capital signal. Following the change in wage structure households apply less unobservable learning effort, resulting in a decrease in skills conditional on education level.

JEL Classification: E24, I26, J24
Keywords: Labor Markets, Education, Misallocation, Sorting

* Department of Economics, University of Arkansas (e-mail: augarov@uark.edu). I am grateful to the participants of 2017 Midwest Economics Association Meeting and 2017 Southern Economic Meeting for comments and suggestions, but accept the full responsibility for the paper's shortcomings. The paper is preliminary and incomplete.

1 Introduction

Workers in poor countries tend to have less human capital for the same level of education than workers in rich countries. This difference in education quality explains more than 20% of cross-country productivity variatio, which is very close to the explanatory power of the difference in education quantity (Erosa and Koreshkova, 2010; Manuelli and Seshadri, 2014). Still there is no accepted explanation for the existing gap in education quality.

Numerous studies show that additional year of education brings consistent increases of about 10% in future earnings (for example, Psacharopoulos and Patrinos, 2004). Students can both overestimate and underestimate these returns, but educational decisions are sensitive to their beliefs (Jensen, 2010; Hastings, Neilson and Zimmerman, 2017). But do students believe that the skills resulting from the education process matter for their future earnings? In this paper I measure the student's beliefs about the importance of academic skills for future labor outcomes. As long as students can vary the quality of their learning, beliefs about the potential rewards for skill should be an important determinant of education quality.

There are several reasons to believe that perceived returns to skill vary between countries. The existing studies done for developing countries find a strong variation in actual returns to skill related to labor market and educational institutions (Hanushek et al, 2013; Hanushek et al, 2017). In developing countries different types of skill misallocation can also affect the returns to skill and job assignments. First, many companies are privately-held and managed by family members of owners, which results in lower productivity (Caselli and Gennaioli, 2013; Bloom and Van Reenen, 2010). Beaman and Magruder (2012) in the field experiment in India show that job referrals can often be biased towards less skilled worker's friends and relatives. Many developing countries also suffer from ethnic, religious and caste discrimination, for which India remains the most famous example (Banerjee and Knight, 1988; Hnatkovska, Lahiri and Sourabh, 2012).

In the empirical part of the paper I construct two country-level measures of perceived impact of skills. These variables reflect the statistical dependency between students' skills and their expected occupations in PISA 2015 dataset. Because both measures use occupational choice as an outcome variable, they also measure the strength of occupational sorting on skills. In other words, under relatively standard assumptions[1] a lower perceived return to skill indicates a higher skill misallocation. I show that the perceived impact of skills for students has a strong positive correlation with the returns to skill for adult workers.

The main finding of this paper is that students demonstrate lower levels of academic knowledge in countries with lower perceived correlation between skill and future occupations. On average the increase in perceived returns from the lowest level of Costa Rica to the highest level observed in Czech Republic corresponds to more than a half of gap in average mathematics

[1]Skills affect occupational choice in the hierarchical job assignment models (Sattinger, 1979; Costrell and Loury, 2004; Teuling and Gautier, 2004) or in occupational choice models based on comparative advantage (Roy, 1951; Heckman and Sedlacek, 1985; Hsieh, Hurst, Jones and Klenow, 2013).

score between these countries. The positive relationship between the perceived returns to skills and academic skills holds after controlling for level of economic development, school resources and socio-economic status of parents.

I construct a signaling labor market model to explain the positive correlation between the perceived returns to skill and education quality. The model's setup is based on the models of statistical discrimination (Phelps, 1972; Lundberg and Startz, 1983). In this model firms base their estimate of worker's human capital on actual years of schooling and a noisy signal of human capital. The noise in a human capital signal includes both the quality of the honest screening process and the corruption of hiring managers. Firms assign workers to job and pay wages below or above the marginal product of labor whenever they overestimate or underestimate the worker's skills. By my definition, the countries with a higher noise of the human capital signal have a higher skill misallocation. I argue that in response to higher skill misallocation workers invest less in human capital per year of schooling.

The main empirical contribution of the paper is a measurement of perceived impact of skills for a large set of developed and developing countries. Until 2012 most studies of returns to skill relied on small samples from developed countries (Neal and Johnson, 1996) In last five years the studies based on new international datasets demonstrated a large variation in returns to skill in developing countries for adult respondents. This paper, to my knowledge, is the first to study the returns to skill as perceived by students making educational decisions, which can differ from the actual returns. My analysis also covers a much larger sample of developing countries.

The paper also contributes to the literature by showing an empirical and theoretical link between skill misallocation and education quality. The literature tends to explain the low education quality in developing countries either through the supply or through the demand effects. The supply-side arguments rely on several studies showing that particular institutional features of educational system tend to improve the education quality (for example, Woessman, 2003; Glewwe and Kremer, 2006). These studies explain most of the cross-country variation in educational achievement scores. The causal identification is still complicated because educational institutions do not vary much over time and international educational achievement tests have limited country coverage.

The alternative demand-side explanation suggests that households choose education quality in response to the perceived demand for skill. Oster and Millet (2010) find that opening of call centers in India resulted in the significant increase in enrollment in English language schools. In papers of Manuelli and Seshadri (2014) and Erosa et al (2010) households choose a higher quality of schooling in response to higher expected skill prices. This paper follows the same logic in endogenizing the choice of schooling quality in response to expected returns to skills.[2]

[2]Skill price here is a wage per unit of human capital, but in my paper the human capital is not publicly observable and the sensitivity of wages to human capital corresponds to the skill price in models with perfectly observable human capital.

My approach still differs from the papers (Manuelli and Seshadri, 2014; Erosa et al, 2010) by assuming that the lower returns to skills in poor countries result from skill misallocation, but not from the lower total factor productivity.

Several recent papers construct and estimate theoretical models of skill misallocation, but to my knowledge none of them studies the connection between the skill misallocation and the education quality. Mora and Watts (2015) combine the Becker-Tomes human capital accumulation with a signaling model to explain low intergenerational mobility in highly meritocratic countries. The paper of (Hsieh, Hurst, Jones and Klenow, 2013) studies the talent misallocation between occupations due to gender or race barriers. They find that elimination of occupational barriers increased the workers productivity in US by 15-20% between 1960 and 2008. Singha (2014) argues that the budget constraints in developing countries lead to occupational mismatch, because workers tend to keep the occupations of their parents to save on educational expenses. The result of Singha (2014) implies that budget constraints lead to a lower education quality, but only for workers educated by their parents.

In the next section I describe the measurement of perceived returns to skill. It starts with explaining the logic of misallocation measures in the subsection 3.1. The second subsection explains the procedures and the data used to construct the variables. In the next subsection 3.3 I demonstrate the correlation between the PISA educational scores and the misallocation measures. The concluding subsection 3.4 describes the validation of constructed variables by comparing them with similar measures for adult workers based on a different data set. Section 3 sets up the theoretical model and describes its predictions.

2 The Importance of Skills

2.1 Intuition

The usual approach to estimate the returns to skill implies regressing labor incomes on measures of skill or intelligence. The downside of this approach is that the data on skills and labor incomes is available for only the small set of countries. In this paper to cover a much larger sample of countries I use occupations as a key measure of labor market outcome instead of wages.

For each country I measure the dependency between academic skills and future occupations for the representative sample of high school students. I label the two resulting variables as perceived returns to skill, but it should be noted that they differ from the standard measures in two key aspects. First, instead of labor incomes these variables use occupations as labor market outcome variable. Second, these measures rely on expected self-reported outcomes instead of actual outcomes. The expected outcomes are arguably more relevant for student's decisions to accumulate skills in contrast to potentially different actual outcomes. Despite these differences from standard measures, I call these variables "returns to skill" for the rest of the paper to avoid verbose wording.

Both measures of returns to skill also describe the intensity of sorting across occupations based on skills. In many standard models of job assignments for heterogeneous workers a weaker sorting implies a higher misallocation or mismatch of workers across jobs by skill levels. For example, one can assume that the productivity of surgeon is more sensitive to his intellectual skills than a productivity of a night guard. If in some country A low-skilled individuals choose to become surgeons, while high-skilled individuals choose to become night guards, the output of country A reduces relative to the potential output because of this skill misallocation. My measures of returns to skills will be low in country A as skills there have only a small impact on occupations choice.

2.2 Data

My main data comes from the Program for International Student Assessment (PISA) 2015 micro dataset. The Program conducts the survey of skills, background and attitudes of 15-year old high school students. The 2015 dataset covers 72 countries, including at least 40 developing countries. On average each country's sample contains a nationally representative sample of on average 7500 students with a maximum number of 32330 students for Spain and 1398 for Puerto-Rico. The sample is stratified by school with average of 140 students coming from each school.

My measures of skill impact utilize the student's self-reported expected occupation and the data on cognitive and non-cognitive skills. The future occupation variable comes the responses to the PISA question "What kind of job do you expect to have when you are about 30 years old?". Almost 80% of students have indicated some future occupation with the remaining 20% either giving a vague description, stating no future employment (housewife, student, unemployed) or answering that they do not know the answer.

PISA also provides the measurement of abilities both through the PISA subject scores (mathematics, reading and science) and through the psychological self-assessment. For each subject score PISA reports 10 plausible values. Each plausible value constitute one random draw from the conditional distribution of score based on student's responses. I average plausible values to get a a measure of skill for each domain.

The dataset also contains three metrics constructed from different self-assessment questions, which I use to proxy for non-cognitive skills. "Collaboration and Teamwork disposition" metric shows the degree to which students enjoy cooperation. "Student Attidudes, Preferences and Self-related beliefs: Achieving motivation (WLE)" metrics describes the student's drive for achievement. Finally the third measure "Subjective well-being: Sense of Belonging to School (WLE)" can proxy both for interpersonal skills and for the school learning atmosphere.

2.3 Construction of Returns to Skill Measures

I construct two measures of dependency between skills and occupational choice:

1. **Single-dimensional perceived skill misallocation**: Spearman rank correlation between skill (first principal component of PISA reading and math scores) and occupational prestige score

2. **Multidimensional perceived skill misallocation**: chi-square (Cramer V) for the correlation between reading, mathematics scores, motivation and occupations:

The first approach relies on the assumption that both skill and occupational assignment can be described by single-dimensional indexes. The first principal component of student's reading and mathematics score describes the aggregate skill. The ISEI occupational prestige score proxies for hierarchy of occupations and for future rewards on the labor market. The single-dimensional perceived misallocation is equal to the Pearson correlation between the percentile of a student by skill in the national sample distribution and the percentile of student by the prestige of expected occupation in the national sample.

Most other studies of returns to skill also assume that both skills and labor outcome are single-dimensional. In these studies numeracy skills or aptitude tests often describe the skill (Neal and Johnson, 1996; Hanushek et al, 2013), while the wage rate is the outcome variable. The cross-country comparisons also require an assumption that countries have a similar ranking of occupations by sensitivity of productivity to skill (same occupations ladder).

If skill is actually multi-dimensional, then using the single-dimensional indexes might indicate a strong skill mismatch in cases when sorting is perfectly optimal (Lindenlaub, 2016). The second approach relaxes this assumption and uses several dimensions to describe skill and do not assume a particular ordering of occupations. It measures the dependency between a skill category of students and a predicted occupation. I use the vector of reading, mathematics scores to describe cognitive skills and motivation to describe non-cognitive skills. Then for each of the three skill measures I separate a national sample into 4 quintiles. The skill category of a student is a combination of her reading, mathematics and motivation quintiles as well as gender, giving in total 128 categories. The algorithm then separates all the reported expected occupations into 10 aggregate occupations corresponding to the first number in ISCO-08 classification. The value of the multidimensional index is equal to the χ^2 statistics of dependency between skill and occupation categories scaled to 0-1 range according to the sample size (Cramer V statistic):

$$V = \sqrt{\frac{\chi^2}{N \min(k-1, r-1)}}$$

In this equation N corresponds to the sample size, k is the number of rows (skill categories) in the correspondence table and r is the number of columns or occupations.

In contrast to the single-dimensional measure, the multidimensional index does not rely on the assumption of common occupations ladder between countries. If, for example, a job of computer programmer in Poland is more skill-intensive than the job of a doctor, the multidimensional measure will still be high as long as high-skilled students want to become programmers rather than doctors. On other hand, this measure hardly relates to actual returns to skill. Even if best workers sort into the least demading jobs, the multidimensional index can still be very high. Both measures vary from 0 to 1 with 1 indicating the perfect dependency between skills and occupational choice. For both variables a higher level of dependency indicates the lower level of skill misallocation.

Perceived skill misallocation strongly varies between countries in the PISA sample. Czech Republic has the highest values of both single-dimensional and multidimensional measures, indicating highest impact of skills on occupational choice or the lowest skill misallocation. The correlation between the rank of ability and occupational prestige rank is equal to 0.58, while the multi-dimensional index is equal to 0.24. Costa Rica lies on the other side of the spectrum with the single-dimensional measure equal to 0.05 and multi-dimensional measure equal to 0.096. Surprisingly United States lies in the middle of distribution for both a single-dimensional measure and for the multi-dimensional one. Two measures of talent misallocation are highly correlated with Pearson correlation equal to 0.87 (Table 1). It implies that the variation in the first single-dimensional measure of perceived skill misallocation does not result from the variation in prestige of particular occupations or in the role of non-cognitive skills, because the multi-dimensional measure does not make these assumptions.

2.4 Validity

Before proceeding to further analysis of effects and correlations between my measures of returns to skill and educational outcomes I need to make sure that my measures do reflect some inherent characteristics of students sorting to occupations. There are several concerns regarding these measures I need to address in this section:

- Can the noise in skills measurement explain the variation in cross-country variation of returns to skills measures?
- Do perceived returns to skill reflect the actual returns to skill?

Measurement noise. First, I address the question of noise in ability measurement. Some studies suggest that the part of cross-country variation in country average PISA scores results from lower student's effort on tests. If some students put less effort, their score do not reflect their ability.

I use the time data provided for 2015 test to measure the average time of cognitive test to completion, as the average time spent can proxy for effort. The analysis based on timing data for knowledge-testing questions does not reveal any systematic pattern in the number of skipped answers or in the number answers answered in less than 2 seconds. The average time to complete the cognitive part (educational testing) tends to be higher in more misallocated countries, which is the opposite of what one should expect if trying to explain higher misallocation measures through the lack of effort in answering cognitive questions.

For a given level of effort the noise in skill measurement can still result from the fact that each student replies to only a small set of questions, which can not cover the potential knowledge expected from a high school student. To measure this noise I use the variation in plausible scores for each of the three tested academic subjects. I find a weak negative correlation between my measures of perceived returns to skill and the dispersion of plausible values for mathematics and a weak positive correlation for the reading plausible values dispersion. Overall, there is no evidence that the measurement of knowledge drives the cross-country variation in perceived returns to skill.

The observed variation in my measures also can result from the noise in reporting of future occupations. While there is no direct way to measure the deviations between the actually expected and reported occupations, the percentage of uncertain answers have a positive and statistically significant, but a weak correlation with measures of skill impact. The Pearson correlation is equal to 0.38 for the single-dimensional perceived return and 0.31 for the multi-dimensional perceived misallocation.

Occupational prestige as a measure of future rewards. Next, I address the concern that income levels by occupation might differ between countries, making an occupational prestige score a poor measure of future rewards. For example, doctors in Poland might have a much lower pay relative to national average incomes compared to the doctors in US, potentially leading to lower occupational prestige[3]. To address this concern, I estimate the average income level by occupation by averaging the income levels of parents in the same country for occupation code with highest occupational prestige score in the family. Family income levels in PISA 2015 are given in 6 country-specific intervals. Suppose, a student reports the highest income level (6) and the student's father is a doctor and the mother is a primary school teacher. In this case the income level of family is attributed to the occupation of doctor as this occupation has a highest occupational prestige score. This calculation does not account for the income generated by the second-highest occupational code, but the error should be small as long as there is either a strong marital sorting or low employment levels of mothers.

The income-based perceived skill misallocation measure is then equal to the correlation between the student's percentile by skill and the student's percentile of average income of expected occupation. The data allows to calculate the measure only for 15 countries. For this limited sample of countries the correlation between the old occupational prestige-based and the new income-based skill misallocation measures equals to 0.8. It implies that using

[3] For example, https://economix.blogs.nytimes.com/2009/07/15/how-much-do-doctors-in-other-countries-make/

the occupational prestige score as a uniform proxy for income in different countries does not strongly bias my results.

Perceived vs actual returns to skill. For a subset of mostly developed countries the Programme for the International Assessment of Adult Competencies (PIAAC) provides the data on occupations, earnings and cognitive skills of adult workers. This dataset allows me to construct the measure of actual returns to skill and contrast it with already calculated variables of perceived skill misallocation.

On the first calculation step, I recode the ISCO-8 occupation code into the occupational prestige index (ISEI) by using *ISCOISEI* routine for Stata[4]. Then I calculate the percentile of each worker in the country's distribution of occupational prestige to obtain a measure of job allocation. The conversion to percentiles pursues the same goal as the conversion done when calculating the misallocation based on PISA data: it allows to obtain a measure of job assignment which is free from cross-country differences in occupational distributions.

On the second step I construct the index of ability, which is equal to the first principal component of numeracy and literacy skills. The actual return to skill is equal to the Spearman rank correlation between ability and the occupational prestige score. I consider the resulting number as a single-dimensional measure of actual returns to skill and compare it with the perceived returns calculated from PISA dataset. Table 1 describes the pairwise correlations between the PISA-based misallocation measures and the PIAAC-based measure for adult workers.

There is a strong and positive connection between the perceived and actual returns to skill. For a limited sample of 22 countries for which the data is available in PIAAC and PISA, the Pearson correlation is equal to 0.53 and is significant at 5%. The correlation between the actual single-dimensional measure and perceived multi-dimensional is also positive, but is relatively weak and not statistically significant for this sample size. Overall, these calculations suggest that the perceived returns to skill actually measure some characteristics of actual labor market assignments, whether they result from employment or educational decisions.

Determinants of skill misallocation. In the first subsection I demonstrate that there is a wide variation in perceived returns to skill and in multidimensional skill misallocation across countries. What drives these differences? Next, I study the correlations between the two skill misallocation measures and different macroeconomic and institutional country characteristics (Table 2). The goal of this exercise is not to identify the causal link, but to limit the range of potential explanations of observed occupational sorting patterns. Pairwise correlations in these regard fulfill my goal and allow to avoid both multicollinearity and power issues.

First, I start with the characteristics of educational systems. The official graduation age describes the uncertainty in occupational reporting. As all students report their occupational choice at the age of 15, the difference in high school graduation age implies that some students are much closer to the moment of implementing their occupational decisions. It is then natural to assume that students which are closer to graduation, are going to report more deliberate

[4]Written by J. Hendrickx, https://ideas.repec.org/e/phe38.html

choices. The data shows an opposite pattern: countries with higher graduation age demonstrate a higher dependency between skills and occupational choice.

Next, I study macroeconomic characteristics. I use log GDP per capita, calculated from Penn World Tables 9.0 to proxy for economic development, economic growth in last 10 year to describe (following Hanushek et al, 2017), two measures of financial sector depth and the Gini coefficient. I also add the educational Gini coefficient, which I calculate based on the to describe the inequality in complementary factors.

Both misallocation measures have small correlation with the level of economic development as measured by GDP per capita. On average rich countries tend to have less talent misallocation, but due to the small coefficient and the small sample size the connection is not statistically significant even at 10%. Two measures of financial sector development also do not have any statistically significant correlation with misallocation.

Measures of perceived returns to education tend to be lower in countries experiencing fast economic growth in last 10 years. The correlation is marginally significant at 10% for the first measure and marginally insignificant for the second. The direction of correlation differs a strong positive correlation between economic growth and returns to skills for adult workers, observed in Hanushek et al (2017).

The income inequality as measured by the Gini coefficient has a very strong and negative correlation with both measures of skill allocation. The countries with higher inequality tend to have lower sorting on skills or higher perceived skill misallocation. The correlation coefficient is equal to -0.69 for the first measure and -0.73 for the second. In both cases it is significant at 1% level despite a small sample size of 43 countries for which the data is available. The observed positive correlation between inequality and skill misallocation suggests that the trade-off between inequality of opportunities and inequality of outcomes as suggested by Benabou (2000) is either weak or non-existent in my sample.

The connection between political institutions and skill misallocation is relatively weak. Table 2 reports pairwise correlation coefficients for 6 different measures of political institutions: polity as a difference between democratic and autocracy features in political system, democracy index, constraint on chief executive, executive recruitment index, political competition and control of corruption. The variables except for World Bank's Control of corruption, come from Polity IV dataset. All correlations have expected positive signs, but only democracy index is significant at 10%.

The development of contractual institutions also correlates with higher perceived returns to skill. Higher contract enforcement costs correspond to lower allocation measures with statistical significance at 5% for the first measure (rank correlation between skill and occupational prestige) and significance at 10% for the second multi-dimensional measure.

Table 2 demonstrates a strong correlation between trade openness and sorting based on ability. The share of foreign trade (sum of export and import) in GDP positively correlates with both measures, but is significant only at 10%. One of the reasons for low significance is a large variation in the share due to large variation in country sizes. The residual from the regression of trade share on log population is statistically significant at 5% for both measures. Both average trade costs per container and the applied weighted average tariff on all goods relate to lower perceived returns to skills and are highly statistically significant.

Intergenerational elasticity of incomes from Corak (2013) is also correlated with misallocation measures, but these correlations can follow from the known correlation between the intergenerational income elasticity and income inequality (Corak, 2006). The Inequality of Opportunities index (IoP), which is produced by Brunori (2016) for selected European countries, measures the variance in incomes explained by observable uncontrollable circumstances (such as parental education, parental occupations and gender). My calculations do not show any significant correlation between the IoP index and the misallocation measures.

Summing up, both measures of perceived returns to skill demonstrate strong and positive correlation with trade openness measures and strong and negative correlation with Gini coefficients. It implies that the theoretical explanation of skill misallocation patterns should also generate higher inequality in more misallocated countries. The strength of occupational sorting based on skills tends to be higher in countries with good political and business institutions.

[INCOMPLETE]

2.5 Education Quality and Misallocation

I find that both talent misallocation measures of this paper have a strong correlation with average PISA scores on country level. Figure 1 in Appendix shows the scatter diagrams of single-dimensional misallocation measure (rank correlation) plotted versus average PISA scores in reading, math and science between countries. The last diagram on the bottom right plots the misallocation measure against the returns to domestic schooling for immigrants in US (calculated in Schoellman, 2012) as the alternative measure of education quality. All the graphs show that countries with higher skill misallocation tend to have lower educational achievement scores, but the predictive power of misallocation is limited as evidenced by high variability of outcomes. The increase in perceived returns to skill from the lowest level (Costa Rica) to the highest level (Czech Republic) corresponds to the increase in mathematics score by about 100 points. This difference corresponds to more than half of the gap in educational achievement scores between the two countries.

Obviously while univariate regressions can make nice pictures, they suffer from the omitted variables bias. I start addressing this issue by controlling for log GDP per capita, which is a strong predictor of educaton quality (Hendricks, 2002; Schoellman, 2012). I also control for corruption by using the World Bank's Control of Corruption measure as corruption correlates both with my measures of skill misallocation and potentially with the efficiency of spending school resources.

Table 32 reports the results of OLS estimation of PISA mean scores on the first perceived misallocation measure (skill-occupation prestige rank correlation) and both macroeconomic variables. Controlling for log GDP per capita and corruption reduces the coefficient estimate for perceived skill misallocation by about 30% depending on the academic subject. The coefficient remains strongly statistically significant.

Next, I explore whether the correlation between the average PISA test scores and misallocation measures can be explained through the characteristics of educational systems. Table 4 in Appendix reports the coefficient estimates after controlling for school characteristics. On top of previously added macroeconomic variables the list of controls includes the student-teacher ratio, class size, school type dummies, average socio-economic status of parents, average parental education and PISA-produced measures of school and curriculum autonomy. These variables often emerge in the literature as important predictors of education quality (for example, Hanushek, 2012). I estimate this regression on the school level so that school characteristics vary within country while the perceived skill misallocation and the macroeconomic variables remain constant.

After controlling for the characteristics of educational systems the coefficient in front of perceived skill misallocation slightly decreases in magnitude, but remains positive and highly significant. The increase in perceived skill misallocation from the lowest to the highest level corresponds to approximately 50 points gain in mathematics score and slightly less in other subjects. Most likely, this value underestimates the actual connection between the perceived skill misallocation and knowledge as any decrease in demand should also affect the educational institutions. Overall, these result suggest that my measures of perceived returns to skill proxy for some important latent factors which vary independently from the educational resources and most important institutional characteristics.

3 Model

My empirical analysis indicates that some other important factors beside skills affect the occupational choice. A higher influence of these factors correlates with lower average academic skills of students. This correlation imposes certain limitations on what role these factors can play. For example, the uncertainty of future labor outcomes by itself does not automatically result in lower educational investment (Levhari and Weiss, 1974). However the uncertainty can decrease the investment in education if individuals are risk-averse and the uncertainty increases with the level of education.

In this section I propose one of the possible theoretical frameworks consistent with the empirical observations. The model's setup is similar to the setup of several models of statistical discrimination (Phelps, 1972; Aigner and Cain, 1977) and particularly to (Lundberg and Startz, 1983). The skill misallocation in my model corresponds to larger noise in statistical discrimination models. The main differences from the model of Lundberg and Startz (1983) consist in adding years of formal schooling into the human capital production function and abandoning the assumption of linearity of worker's utility function. These modifications allow to analyze the changes in returns to schooling resulting from skill misallocation in a standard Mincer regression setup and to make several novel predictions.

The economy consists of a continuum of workers (indexed by i) and a continuum of firms. Denote each worker's productivity by its human capital H_i. The human capital of the worker i depends on the years of formal schooling s_i and the learning effort e_i. The effort and the human capital are private information of a worker. In the model workers accumulate human capital through schooling combined with learning effort. After education period, workers apply for jobs by sending the noisy signal of human capital and the perfect signal of formal schooling. The noisy signal of human capital can be distorted by the corrupt hiring manager. Firms pay wages equal to an estimate of human capital conditional on signals obtained. Only one period is considered in the model.

3.1 Workers

Each worker chooses the years of schooling s_i and the learning effort e_i. Years of schooling in the model can correspond to easily verifiable information on diplomas, certificates and actual years of education. In the paper I use schooling and years of schooling interchangeably. The learning effort e_i represents the unobservable investment in human capital in form of self-education or developing non-cognitive skills.

Worker accrues all the lifetime benefits of her own education, denoted as the market wage w_i. I assume that the utility function has a logarithmic form $u() \equiv log(\cdot)$, which simplifies the derivations. The parameter γ_i represents ability. The objective function of a worker is to maximize the utility from the market wage w_i minus the learning effort costs $e_i s_i \exp(-\gamma_i)$:

$$\max_{e_i}[u[w_i] - e_i s_i \exp(-\gamma_i)] \tag{1}$$

Ability in general is correlated with schooling and distributed normally conditional on schooling:

$$\gamma_i = E(\gamma|s) + \tilde{\gamma}, Var(\gamma|s) = Var(\tilde{\gamma}) = \sigma^2_{\gamma|s}$$

The actual human capital H, which is also equal to the worker's productivity, is given by:

$$h_i \equiv \log(H_i) = s_i \log(e_i)$$

This functional form implies multiplicative complementarities between unobserved learning effort e and years of schooling s. The increase in effort by 1% leads to the increase in human capital by s%. This functional form is a special case of the human capital production function used in Bils and Klenow (2000).

This model does not describe how do workers choose the years of schooling s to keep the model tractable. This assumption is justified if credit constraints determine the choice of schooling. The model though allows for arbitrary distribition of schooling levels, and I model some equilibrium effects of change in distribution of schooling later.

The model also assumes that workers have a full control over the unobserved effort e_i, which roughly corresponds to the education quality. The assumption of private control over the education quality is rather strong, as many schools can impose the minimum requirements for effort through the grading policies. This minimum requirements can affect the signaling value of formal schooling. At the same time, many aspects of learning are not usually reflected in schooling credentials, including social skills and the residual knowledge. The problem of poor signaling value is stronger for middle and high school graduates (Arcidiacono, Bayer and Hizmo, 2008), where curriculum and grading practices are often determined by the government.

3.2 Firms

The output is produced by a continuum of firms, each of them is employing only one worker. The firms can not observe the actual human capital of a worker, but can observe years of schooling and a human capital signal $\hat{h}_i$ instead. The human capital signal $\hat{h}_i$ equals to the sum of true log human capital h_i and error $s_i f_i$:

$$\hat{h}_i \equiv h_i + s_i f_i = s_i \log(e_i) + s_i f_i \tag{2}$$

The variable f_i is a misallocation error as it incorporates all the factors distorting the job assignment in the model. The misallocation error f_i is indepently and normally distrbuted with $E(f|s) = 0, Var(f|s) = \sigma_f^2$. The term $s_i f_i$ is the error, which firms obtain when they measure the worker's human capital. The standard deviation of this error $s_i \sigma_f$ increases proportionally with schooling to accomodate for the fact that it is harder to measure human capital of more educated workers. There are at least three interpretations of the variable f consistent with a model, and two of them involve corruption:

- The firm hires an agent (hiring manager) to conduct the interview. After the interview the hiring manager provides his own estimate of the applicant's i human capital $\hat{h}_i$ to the principal. This estimate has to be unbiased on average as any biased estimate is going to be corrected for bias by the principal. Each applicant can bribe a hiring manager to improve the signal. Alternatively the hiring manager can make more beneficial reports for candidates among friends and relatives (Beaman and Magruder, 2012).

- The education system produces a noisy signal of human capital. More corrupt educational systems provide a signal of human capital of worse quality as both grades and diplomas are less reliable (Heyneman et al, 2008; Transparency International, 2013 for numerous examples). Moreover the quality of teachers can affect the quality of grading and make the educational signal more or less informative.

- Misallocation error emerges because of the inherent errors of the screening process.[5]

The first and second interpretations imply corruption either in employment or in education sector. My interpretation of firms extends to all the organizations which employ labor for productive purposes, including the public sector. Because the government sector represents the largest employer of skilled labor in many developing countries, corruption in public sector is likely to affect the labor market efficiency. The increase in rent seeking, of which corruption is a primary example, can result in even higher reallocation of skill towards the government jobs (Baumol, 1990; Murphy, Schleifer and Vishny, 1991)

All firms in the economy receive identical signals of worker's productivity. I also abstract from discussing the timing when the misallocation error f becomes known to the worker. Given the functional form for the lifetime utility, the knowledge of misallocation error does not affect the choice of the learning effort e.

It is important to differentiate between the human capital signal $\hat{h}_i$ (observed by firms) and the actual human capital h_i (unobserved):

$$\hat{h}_i = h_i + s_i f_i = s_i \log(e_i) + s_i f_i$$
$$h_i = s_i \log(e_i)$$

Firms compete for labor force in a Bertrand style competition and obtain zero expected profit. It implies that in equilibrium they pay the competitive wage, which is equal to the conditional expectation of marginal product y_i for each individual i, based on years of schooling s_i and the human capital signal $\hat{h}_i$.

$$w_i = w(s_i, \hat{h}_i) = E[y_i | s_i, \hat{h}_i]$$

The function $w(s, \hat{h})$ will be called a wage function. It maps two signals (years of schooling s_i and human capital signal $\hat{h}_i$), observed by firms into the wage offered to a worker.

Following Mora and Watts (2015), I include the misallocation penalty into expected marginal product.[6]. Misallocation penalty τ measures the losses in productivity resulting from the mismatch between the expected and actual productivity of the worker. The expected marginal product is given by the following expression:[7]

$$w(s_i, \hat{h}_i) = E[y_i | s_i, \hat{h}_i] = AE[\exp(h_i - \frac{\tau}{2}(h_i - E(h_i | s_i, \hat{h}_i))^2) | s_i, \hat{h}_i] \tag{3}$$

[5] I am not aware about any studies showing that screening process efficiency differs between countries, but it is reasonable to expect better screening in countries with long-term employment. On other hand, countries with short-term employment can reveal the human capital better through multiple spells of employment.

[6] My main result do not rely on misallocation penalties, but without it average wages in the economy will increase with misallocation

[7] This functional form for the expected product follows from the production function $y = \exp[h - \frac{\tau}{2}(h - a - Var[h|s, \hat{h}])^2)]$

Here the constant A is the economy-level productivity parameter or TFP. The parameter $\tau > 0$ measures the magnitude of the misallocation penalty. If $\tau = 0$ the output is proportional to the actual human capital of a worker ($y = AH_i$).

I consider the linear rational expectations equilibrium in this model. The wage function in equilibrium is going to depend on beliefs of firms about human capital of workers conditional on observed signal. The firms' belief is a cumulative distribution function $\mu(s, \hat{h}, x) = Prob(h < x|s, \hat{h})$, assigning the probability to the outcome that the actual log human capital of a worker with schooling s and human capital signal $\hat{h}$ is less than a number x. The formal definition of the equilibrium is given below.

3.3 Equilibrium

The equilibrium in this model is a combination of manager's belief function $\mu(s, \hat{h})$, workers decision rules $e(\gamma_i, p_i)$ and $s(\gamma_i, p_i)$, wage function $w(s, \hat{h})$, such that:

- Belief function $\mu(\cdot, \cdot, \cdot)$ is consistent with the distribution of workers by human capital, conditional on schooling and human capital signal.
- Wage is equal to the conditional expectation of human capital $w(s, \hat{h}) = E_\mu[AH|s, \hat{h}]$.
- Workers decision rule for human capital accumulation $e(\gamma, s)$ maximize the utility for a given wage function

I prove that there exists such a wage function, that is linear in human capital signal $\hat{h}$ and satisfies all the equilibrium conditions. I start with showing in Proposition 1 that if such equilibrium wage function exists, the optimal learning effort e^* is log-linear in ability and hence lognormal.

Proposition 1. *If equilibrium wage function is log-linear in the human capital signal* $\log(w(s, \hat{h})) = \theta_s s + \theta_h \hat{h} + G(s)$*, then the optimal level of learning effort is given by* $e_i = \theta_h \exp(\gamma_i)$.

Proof. After substituting the wage function and taking the expectation the utility is:

$$\max_{e_i}[\theta_s s + \theta_h(s_i(\log e_i + f)) - e_i s_i \exp(-\gamma_i)]$$

The first-order condition is:

$$\frac{\theta_h s_i}{e_i} - s_i \exp(-\gamma) = 0$$

Resulting in:

$$e_i^* = \theta_h \exp(\gamma_i)$$

First order condition describes the maximum solution as the objective function is strictly jointly concave. $\square$

From the Proposition 1 it follows that the log-human capital depends on the years of schooling and ability in the following way:

$$h_i = s_i(\log(\theta_h) + \gamma))$$

The equilibrium parameter θ_h measures the weight the firms place on the human capital to determine the wage. A higher θ_h corresponds to wage being more responsive to unobservable learning effort e as compared to years of schooling. The equation above implies that an increase in the weight θ_h increases the human capital conditional on schooling.

The Proposition 2 describes the equilibrium solution. In this solution the log wage is a linear function of schooling s and observed human capital signal $\hat{h}$. It then follows that the assumption of the Proposition 1 is satisfied in this solution.

Proposition 2. *The model has a unique linear rational expectations equilibrium in which the log wage is linear in the signal of human capital. In this equilibrium log wage is given by the following equation:*

$$\log(w(s,\hat{h})) = (1-\theta_h)s(\log(\theta_h) + E(\gamma|s)) + \theta_h\hat{h} + \frac{V(s)}{2(1+\tau V(s))} - 1/2\log(1+\tau V(s)) \quad (4)$$

Where θ_h is:

$$\theta_h \equiv \frac{\sigma^2_{\gamma|s}}{\sigma^2_{\gamma|s} + \sigma^2_f} \quad (5)$$

And:

$$V(s) = s^2\sigma^2_{\gamma|s}(1-\theta_h)$$

Proof. See Appendix. □

The variable $V(s)$ in this equation (4) defines the dispersion of firms' forecast of log human capital based on information available. It increases with years of schooling s because the dispersion of measurement error sf increases with schooling. The equilibrium value for θ_h in (5) is equal to the ratio of dispersion of ability $\sigma_{\gamma|s}$ conditional on schooling to the sum of dispersions of ability and misallocation error σ^2_f.

The following Proposition 3 studies the comparative statics of the equilibrium wage function coefficient. It states that the increase in misallocation results in lower human capital accumulation and lower variance of (log) wages.

Proposition 3. (Comparative Statics) *The linear rational expectations equilibrium has the following properties:*

1. *an increase in misallocation error σ^2_f leads to lower average human capital: $\frac{dEh}{d\sigma^2_f} < 0$*
2. *an increase in misallocation error σ^2_f leads to lower conditional variance of log wages $\frac{dVar(\log(w)|s)}{d\sigma^2_f} < 0$*

3. *a decrease in variance of ability conditional on schooling* $\sigma^2_{\gamma|s}$ *leads to lower returns to imperfectly observable human capital* $\frac{d\theta_h}{d\sigma^2_{\gamma|s}} < 0$

4. *a descrease in variance of ability conditional on schooling* $\sigma^2_{\gamma|s}$ *leads to lower conditional variance of log wages* $\frac{dVar(\log(w)|s)}{d\sigma^2_{\gamma|s}} < 0$

Proof. Obviously $\theta_h = \frac{\sigma^2_{\gamma|s}}{\sigma^2_{\gamma|s}+\sigma^2_f}$ is decreasing in σ^2_f. The (log)human capital is given as $s_i(log(\theta_h)+\gamma))$ and hence the decrease in θ_h decreases the human capital for each γ and each positive schooling level. For the part 2 of Proposition 3 I substitute the expression for the signal of human capital from (2) into the wage equation (4):

$$\log(w(s,\hat{h})) = \theta_h s_i(\log\theta_h + \gamma_i + f_i) + G(s)$$

where $G(s) = \log(A) + (\log(\theta_h) + E(\gamma_|s))(1-\theta_h)s + \frac{V(s)}{2(1+\tau V(s))} - 1/2\log(1+\tau V(s))$. The conditional variance is:

$$Var(\log(w(s,\log(\hat{h})))|s) = \theta_h^2 s^2(\sigma^2_{\gamma|s} + \sigma^2_f) = \frac{\sigma^4_{\gamma|s}}{(\sigma^2_{\gamma|s}+\sigma^2_f)^2}s^2(\sigma^2_{\gamma|s}+\sigma^2_f) = \frac{\sigma^4_{\gamma|s}s_i^2}{\sigma^2_{\gamma|s}+\sigma^2_f} \quad (6)$$

The conditional variance is a decreasing function of the misallocation error dispersion σ^2_f. The parts 3 is proved by direct differentiation of the equilibrium weight expression and the part 4 is proved by differentiation of the equation (6). □

Proposition 3 is the most important prediction of the model. Part 1 implies that a country with a higher misallocation of labor as measured by the variance of misallocation σ^2_f should have a lower unobservable investment in learning effort e and hence lower human capital for a given level of schooling. When the human capital signal $\hat{h}$ becomes less informative, firms rely less on the signal and more on years of schooling s. Workers respond by reducing the learning effort. It results in the decrease in the education quality and in lower human capital.[8] Part 2 of Proposition 3 provides another testable implication of the model. It predicts that the countries with a lower wage inequality conditional on schooling should have a lower education quality.

Part 3 states that a decrease in conditional variance have a negative effect on incentives for human capital accumulation by lowering the weight of human capital signal in the wage function. It implies that better sorting of students by ability can have a detrimental effect on human capital accumulation. Better sorting means that the students with higher ability obtain more schooling. The sign and the magnitude of the effect of better sorting depends on how the sorting improvement affects the conditional expectation of ability $E(\gamma|s)$. Assume that the schooling has a marginal distribution $\phi(\cdot)$ and the parameter ρ describes the sorting in such a way, that all the equilibrium model parameters are differentiable in ρ:

[8]This result mirrors the theoretical result in Lundberg and Startz (1983) who predict that groups with noisier human capital signals get lower returns to human capital and hence will optimally acquire less of it.

$$\frac{dEh}{d\rho} = \int s \left(\frac{d\theta_h}{d\sigma^2_{\gamma|s}} \frac{d\sigma^2_{\gamma|s}}{\rho} + \frac{dE(\gamma|s)}{d\rho} \right) \phi(s) ds$$

As long as the effect $\frac{dE(\gamma|s)}{d\rho}$ is small enough on average, the human capital stock goes down with better sorting. For example, this outcome is possible when the better sorting correspond to just lowering the conditional variance of ability. Then $\frac{dE(\gamma|s)}{d\rho} = 0$ and $\frac{dEh}{d\rho} < 0$. Another case is when ability and schooling have joint normal distribution:

$$(\gamma_i, s_i) \sim N\left(\begin{bmatrix} \gamma_0 \\ s_0 \end{bmatrix}, \begin{bmatrix} \sigma^2_\gamma & \rho\sigma_\gamma\sigma_s \\ \rho\sigma_\gamma\sigma_s & \sigma^2_s \end{bmatrix} \right)$$

In this example the integration gives the following total effect of increase in correlation ρ on human capital:

$$\frac{dEh}{d\rho} = \sigma_\gamma\sigma_s + (\sigma_\gamma/\sigma_s)s_0^2 - \frac{2\rho\sigma_f^2}{(1-\rho)^2\sigma^2_\gamma + \sigma^2_f} s_0$$

The effect of better sorting $\frac{dEh}{d\rho}$ may be negative for average human capital of workers if misallocation is high or the current level of sorting ρ is high or the ability variance is low. The effect also increases with schooling. The increase in correlation has a higher effect on average ability when schooling is high, because average ability in the example of joint normal distribution increases with schooling ($E(\gamma|s) = \gamma_0 + \frac{\rho\sigma_\gamma}{\sigma_s}(s_i - s_0)$). In contrast, the effect on (log) learning effort does not depend on current schooling or ability levels. Hence for higher educated workers the effect of better sorting is more likely to increase the human capital.

The prediction that the better sorting can decrease the human capital contradicts the existing wisdom on the effects of better sorting. Most studies suggest that eliminating the credit constraints to improve sorting in education should benefit human capital accumulation (De Gregorio, 1996; Christou, 2001; Cordoba, Ripoll, 2013). The difference in predictions arises because previous studies do not account for private learning effort, which decreases when the signaling value of formal education goes up.

Part 4 of Proposition 3 discusses another implication of better sorting. It implies that the decrease in conditional variance of ability (resulting, for example, from better sorting in education) reduces the wage inequality for any given level of schooling. This effect potentially complicates the empirical testing of the model, because sorting and misallocation are likely to move in opposite directions. Note, that Part 2 of Proposition 3 states that the increase in misallocation results in lower variance of wages. Suppose that countries with better ability sorting have less misallocation on labor market. Then their conditional variance of log wages might be lower or higher than the conditional variance of other countries depending on magnitudes of the differences in sorting and in skill misallocation.

4 Conclusion

Several previous studies demonstrate the strong effect of beliefs about the returns to education on educational decisions. In this paper I ask the question of whether the beliefs about the returns to quality of education have an impact on education quality. I construct a measure of perceived returns to skill for 52 different countries by using the data on occupational choice and abilities of high school students. My proxies for return to skill measure the occupational sorting of students based on their cognitive and non-cognitive skills.

I find a large cross-country variation in the role of cognitive and non-cognitive abilities in occupational choice. The measures of perceived skill misallocation for students highly correlate with similar measures of returns to skill constructed for adult workers in the same countries. Most importantly, the countries with a lower role of skills in occupational choice (higher skill misallocation) demonstrate significantly lower education quality as measured both by PISA scores and by the returns to domestic schooling for US immigrants.

I explain my empirical findings in a framework of labor signaling model. In the model firms receive distorted signals of worker's human capital as well as perfect information on years of formal education. An increase in signal noise results in the suboptimal job assignment (skill misallocation), but also reduces the sensitivity of equilibrium wages to human capital. The change in sensitivity reduces the incentives of workers to apply the unobservable learning effort and negatively affects the equilibrium stock of human capital for a given level of formal education. This prediction can explain lower returns to domestic education for migrants from most developing countries if these developing countries have higher skill misallocation.

The model also predicts that improving the allocation of schooling towards the most able workers can have an adverse effect on education quality in countries with high skill misallocation. Better allocation of schooling improves the informativeness of formal credentials and decreases the return on the unobservable learning effort. If human capital signaling is already imprecise, the effect of lower incentives for learning can dominate the effect of better sorting by ability. The negative effect is stronger for workers with lower schooling.

References

[1] Altonji, J. G. and C. R. Pierret (2001). "Employer learning and statistical discrimination," *The Quarterly Journal of Economics*, 116(1), 313-350.

[2] Arcidiacono, P., Bayer, P, and A. Hizmo (2010) "Beyond Signaling and Human Capital: Education and the Revelation of Ability," *American Economic Journal: Applied Economics*, 2(4), 76-104.

[3] Banerjee B and J.B.Knight. (1985) "Caste Discrimination in the Indian Urban Labour Market," *Journal of Development Economics*, 17(3), 277-307.

[4] Beaman L., and J. Magruder (2012) "Who Gets the Job Referral? Evidence from a Social Networks Experiment," *American Economic Review*, 102(7), 3754-3593.

[5] Benabou R. (2000) "Meritocracy, Redistribution and the Size of the Pie," in *Meritocracy and Economic Inequality* ed. by K. Arrow, S. Bowles, S. Durlauf, Princeton University Press.

[6] Baumol, W. J. (1990) "Entrepreneurship: Productive, Unproductive, and Destructive," *The Journal of Political Economy*, 98(5 Part 1), 893-921.

[7] Bils, M. and P. J. Klenow (2000) "Does Schooling Cause Growth?," *American Economic Review*, 77, 1421-1449.

[8] Bloom, N., and J. Van Reenen (2010) "Why do management practices differ across firms and countries?" *The Journal of Economic Perspectives*, 24(1), 203-224.

[9] Brunori P. (2016) The Perception of Inequality of Opportunity in Europe, *The Review of Income and Wealth*, 63(3), 464-491.

[10] Caselli, F. and N. Gennaioli (2013), "Dynastic Management," *Economic Inquiry*, 51(1), 971-996.

[11] Christou, C. (2001) "Differential Borrowing Constraints and Investment in Human Capital," *Journal of Macroeconomics*, 23(2), 277-295.

[12] Corak M. (2006) "Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility," *Research on Economic Inequality*, 13 (1), 143-188.

[13] Corak M. (2013) "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 27(3), 79-102.

[14] Cordoba, J. C. and M. Ripoll (2013) "What Explains Schooling Differences across Countries?" *Journal of Monetary Economics*, 60(2), 184-202.

[15] Costrell R. and Loury G. (2004) "Distribution of Ability and Earnings in a Hierarchical Job Assignment Model," *Journal of Political Economy,* 112(6), 1322-1363.

[16] Erosa A., T. Koreshkova and D. Restuccia (2010) "How Important Is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *Review of Economic Studies*, 90(5), 1160-1183.

[17] Glewwe P. and M. Kremer (2006) Schools, Teachers, and Education Outcomes in Developing Countries. In *Handbook of the Economics of Education*, vol. 2, edited by E.Hanushek and F.Welch, 945-1017.

[18] Hendricks L. (2002) "How Important Is Human Capital for Development? Evidence from Immigrant Earnings,"*American Economic Review*, 92(1), 198-219.

[19] Hanushek E. and L. Woessmann (2012) "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation," *Journal of Economic Growth*, 17, 267-321.

[20] Hanushek E., Guido Schwerdt, S. Wiederhold, L. Woessmann (2013) "Returns to skills around the world: Evidence from PIAAC," *European Economic Review*, 73(C), 103-130.

[21] Hanushek E., Guido Schwerdt, S. Wiederhold, L. Woessmann (2017) "Coping with change: International differences in the returns to skills," *Economics Letters*, 153, 15-19.

[22] Hastings J., C. Neilson, and S. Zimmerman (2017), "The Effects of Earnings Disclosure on College Enrollment Decisions,", NBER working paper no. 21300.

[23] Heckman, J. and G. Sedlacek (1985), "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-selection in the Labor Market," *Journal of Political Economy*, 93(6), 1077-1125.

[24] Hnatkovska V., A. Lahiri and S. Paul (2012) "Castes and Labor Mobility," *American Economic Journal: Applied Economics*, 4(2), 274-307.

[25] Hsieh C., E. Hurst, C.I.Jones and P. Klenow (2013) "The Allocation of Talent and U.S. Economic Growth," NBER Working Paper No. 18693.

[26] Jensen R. (2010). "The (Perceived) Returns to Education and the Demand for Schooling," *Quarterly Journal of Economics*, 25(2), 515-548.

[27] Levhari D. and Y. Weiss (1974), "The Effect of Risk on the Investment in Human Capital," *The American Economic Review*, 64(6), pp. 960-963.

[28] Lundberg, S. J. and Startz, R. (1983). "Private discrimination and social intervention in competitive labor market," The American Economic Review, 73(3), 340-347.

[29] Manuelli R. E. and A. Seshadri (2014) "Human Capital and the Wealth of Nations," *American Economic Review*, 104(9), 2736-2762.

[30] Mayer, T. and Zignago, S. (2011) "Notes on CEPIIs distances measures: the GeoDist Database',' CEPII Working Paper 2011-25.

[31] Mora J.R. and M. Watts (2015) "The Rise of Meritocracy and the Inheritance of Advantage," working paper.

[32] Murphy, K. M., Shleifer, A. and Vishny, R. W. (1991). "The Allocation of Talent: Implications for Growth," *The Quarterly Journal of Economics*, 106(2), 503-530.

[33] Neal A. and W. Johnson (1996) "The Role of Premarket Factors in Black-White Wage Differences," *Journal of Political Economy*, 104(5), 869-895.

[34] Oster E. and B. Millett (2013) "Do Call Centers Promote School Enrollment? Evidence from Indiah," *Journal of Development Economics*, 104, 123-135.

[35] Phelps, E. S. (1972). "The Statistical Theory of Racism and Sexism," *The American Economic Review*, 62(4), 659-661.

[36] Psacharopoulos G. and H. Patrinos (2004) "Returns to investment in education: a further update," *Education Economics*, 12(2), 111-134.

[37] Roy, A. (1951). "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), new series, 135-146.

[38] Sattinger M. (1979) "Differential Rents and the Distribution of Earnings," *Oxford Economic Papers*, 31(1), 60-7.

[39] Schoellman T. (2012) "Education Quality and Development Accounting,",*Review of Economic Studies*, 79(1), 1-30.

[40] Singha R. (2014) "Intergenerational Occupational Mobility and Labor Productivity," Job Market Paper, Arizona State University.

[41] Sweetman, A. (2004) "Immigrant source country educational quality and Canadian labour market outcomes," Statistics Canada research paper.

[42] Teulings C. and P. Gautier (2004), "The Right Man for the Job," Review of Economic Studies, 71(2), 553-580.

[43] Woesmann (2003) "Schooling Resources, Educational Institutions and Student Performance: the International Evidence," *Oxford Bulletin of Economics and Statistics*, 65(2), 117-170.

A Proof of Proposition 2

Take a linear solution to the model as given for now. Then the optimal investments into unobservable skill e are given by the Proposition 1 and are linear in inverse learning costs γ_i. It implies that learning e_i is distributed lognormally.

$$\log(e_i) \sim N(\theta_h + \gamma_0, \sigma_\gamma^2)$$

Calculate the expected human capital conditional on schooling and use the expressions for optimal learning effort and schooling:

$$E(h_i|s_i) = E(s_i(\log(\theta_h) + \gamma_i)|s_i) = s_i(\log(\theta_h) + E(\gamma_i|s_i))$$

Then the joint distribution of $(h_i, \hat{h}_i)$ conditional on schooling s is given by:

$$\begin{bmatrix} h_i \\ \hat{h}_i \end{bmatrix} \sim N\left(\begin{bmatrix} s_i(\log(\theta_h) + E(\gamma_i|s_i)) \\ s_i(\log(\theta_h) + E(\gamma_i|s_i)) \end{bmatrix}, \begin{bmatrix} s_i^2\sigma_{\gamma|s}^2 & s_i^2\sigma_{\gamma|s}^2 \\ s_i^2\sigma_{\gamma|s}^2 & s_i^2(\sigma_{\gamma|s}^2 + \sigma_f^2) \end{bmatrix} \right)$$

Then the conditional expectation of (log) human capital, based on the observed schooling and the signal of skill is:[9]

$$E[h_i|s_i, \hat{h}_i] = s_i(\log(\theta_h) + E(\gamma|s)) + \frac{\sigma_{\gamma|s}^2}{\sigma_{\gamma|s}^2 + \sigma_f^2}(\hat{h}_i - s_i(\log(\theta_h) + E(\gamma|s))) =$$

$$= \big(1 - \frac{\sigma_{\gamma|p}^2}{\sigma_{\gamma|s}^2 + \sigma_f^2}\big)s_i(\log(\theta_h) + E(\gamma|s)) + \frac{\sigma_{\gamma|s}^2}{\sigma_{\gamma|s}^2 + \sigma_f^2}\hat{h}_i =$$

$$= (1 - \theta_h)s_i(\log(\theta_h) + E(\gamma|s)) + \theta_h\hat{h}_i$$

The conditional variance of (log) human capital is:

$$V \equiv Var[\log(h_i)|s_i, \hat{h}_i] = s_i^2\sigma_{\gamma|s}^2 - s_i^2\frac{\sigma_{\gamma|s}^2}{\sigma_{\gamma|s}^2 + \sigma_f^2}\sigma_{\gamma|s}^2 = s_i^2\sigma_{\gamma|s}^2(1 - \theta_h)$$

Based on the derivation from Mora and Watts (2015) the equation (3) for expected marginal product can be written as:

$$E[y_i|s_i, \hat{h}_i] = A\frac{1}{\sqrt{1 + \tau V}}\exp\left(\frac{V}{2(1 + \tau V)}\right)\exp(E[h_i|s_i, \hat{h}_i])$$

Then the equilibrium log-wage is:

$$\log(w(s_i, \hat{h}_i)) = \log(E[y_i|s_i, \hat{h}_i]) = \log(A) + E[h_i|s_i, \hat{h}_i] + \frac{V}{2(1 + \tau V)} - \frac{1}{2}\log(1 + \tau V)$$

Substitute the expression for the conditional expectation of human capital and the conditional variance

[9] Based on simple result in (DeGroot, 1970, p.62)

V to obtain:

$$\log(w(s_i, \hat{h}_i)) = \log(A) + (1-\theta_h)s_i(\log(\theta_h) + E(\gamma|s)) + \theta_h \hat{h}_i + \frac{s_i^2 \sigma_{\gamma|s}^2 (1-\theta_h)}{2(1+\tau s_i^2 \sigma_{\gamma|s}^2 (1-\theta_h))} -$$

$$-1/2 \log(1+\tau s_i^2 \sigma_{\gamma|s}^2 (1-\theta_h))$$

Therefore the wage function is linear in the signal of human capital and the condition for the Proposition 1 indeed holds.

B Tables

Table 1: Correlation matrix of misallocation measures

	Spearman rho (PIAAC)	Cramer V(PISA)	Spearman rho (PISA)
Spearman rho (PIAAC)	1		
Cramer V(PISA)	0.337	1	
Spearman rho (PISA)	0.533**	0.872***	1

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Probable Correlates of Perceived Misallocation Measures

	Rank correlation		Cramer V		Source
	Correlation	N obs	Correlation	N obs	
Educational System Characteristics					
Average high school graduation age	.442**	51	.489***	51	UNESCO
Gov. spending per tert. student	-.0771	43	-.0012	43	UNESCO
Macroeconomic Variables					
Log GDP per capita	.18	51	.14	51	PWT9.0
Econ. growth (2005-2014)	-.317*	51	-.266	51	PWT9.0
Stock market capitalization (% of GDP)	-.153	35	-.252	35	World Bank
Domestic credit to private sector	.0766	49	-.125	49	World Bank
Gini coefficient	-.687***	43	-.729***	43	World Bank
Education Gini coefficient	-.542***	50	-.546***	50	Calculation
Political Institutions					
Polity	.326	24	.162	24	Polity IV
Democracy	.428*	24	.285	24	Polity IV
Constraint on Chief Executive	.389	24	.252	24	Polity IV
Executive Recruitment	.191	24	.0274	24	Polity IV
Political Competition	.319	24	.144	24	Polity IV
Control of Corruption	-.234	51	-.217	51	Polity IV
Contract and Business Institutions					
Contract enforcement cost	-.379**	50	-.307*	50	World Bank
Conflict of interest regulation	-.0931	50	-.198	50	World Bank
Shareholder governance	.312*	50	.207	50	World Bank
International Trade Openness					
Trade(% of GDP)	.303*	51	.275	51	World Bank
Trade costs (USD per container)	-.627***	50	-.644***	50	World Bank
Applied weighted average tariff	-.428**	50	-.375**	50	World Bank
Economic Mobility					
Intergen. income elasticity	-.306	20	-.466*	20	Corak(2013)
Inequality of Opportunity	.355	18	.145	18	Brunori(2016)

Table 3: Perceived misallocation and test results

	(1) Reading	(2) Reading	(3) Math	(4) Math	(5) Science	(6) Science
Skill-occupational prestige correlation	191***	132***	36.9***	24.9***	253***	185***
	(4.1)	(3.2)	(3.7)	(2.8)	(5.1)	(4.3)
Log GDP per capita		11.6		8.26***		20.4
		(1.0)		(3.3)		(1.7)
Corruption (World Bank)		-16.9**		.358		-14.8**
		(-2.5)		(0.2)		(-2.1)
Constant	401***	289**	74.5***	-5.82	372***	174
	(21.6)	(2.5)	(19.1)	(-0.2)	(18.8)	(1.5)
Observations	48	48	48	48	48	48
Adjusted R^2	0.25	0.48	0.22	0.44	0.34	0.57

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Perceived skill misallocation and educational achievement scores: school-level

	Reading		Math		Science	
Skill-occupational prestige correlation	**77****	**(2.6)**	**111*****	**(3.4)**	**98.3*****	**(3.5)**
Log GDP per capita	18.7	(1.0)	36.1*	(1.8)	22.4	(1.4)
Corruption (World Bank)	-12.3	(-1.6)	-8.6	(-1.0)	-13.4*	(-1.9)
Student-teacher ratio	-.176	(-1.0)	-.496**	(-2.6)	-.301*	(-1.7)
Public school	27.7***	(3.2)	32.5***	(3.3)	26.7***	(3.0)
Private (government-dependent)	31.4***	(3.0)	35.3***	(2.8)	26.5**	(2.6)
Curriculum autonomy	4.52**	(2.3)	5.83**	(2.4)	5.36**	(2.4)
Mean parents' socio-economic status	1.99	(0.3)	-5.43	(-0.7)	.382	(0.1)
Observations	10139		10139		10139	
Adjusted R^2	0.59		0.66		0.66	

Other controls: class size, school autonomy, computer-student ratio, mean education of parents

t statistics in parentheses (* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$)

C Figures

Figure 1: Misallocation vs PISA scores