The Production of Cognitive and Non-cognitive Human Capital in the Global Economy

Chong Xiang, Purdue University
Stephen Ross Yeaple, Penn State University and NBER

March 2017

Abstract

The quality of a country's educational infrastructure is a crucial determinant of economic well-being. Therefore, the comparisons of the relative strength and weakness of educational systems across countries are critical for both academic research and policy-making. A common approach measures the comparative quality of educational systems directly using international test scores. Aspects of educational quality that are ill-measured by exams, however, are neglected in such analyses. In this paper, we develop a general equilibrium framework that allow educational outcomes to vary in the extent to which they are readily quantified on exams. Our framework allows inference along multiple dimensions of educational quality and provides a method for aggregating over these dimensions to construct a single measure of institutional quality. Many countries that score well on international exams fair poorly according to our measure. Our comparative static results suggest important tradeoffs across eductional dimensions, and spell out the implications of educational-institution qualities for aggregate output.

1 Introduction

Human capital is central to both economics and other social sciences, and the key institution that produces human capital is the educational system. Therefore, the comparisons of the relative strength and weakness of educational systems across countries are critical for both academic research and policy-making. Unfortunately, while it is easy to measure the quantity of education (i.e. years of schooling or expenditures per pupil), it is difficult to determine the quality of education provided by a country's school system. Recently, greater attention has been paid to the performance of students on international assessment tests, like PISA, to assess the quality of education provided by country.

Judged by these test scores, the U.S. educational system does a poor job, generating low international scores despite having one of the highest levels of per-capita educational spending in the world. Not surprisingly, U.S. policy makers are alarmed. President Obama said that the nation that "out-educates us today will out-compete us tomorrow." The U.S. also implemented major policy changes, such as No Child Left Behind of 2001 and Race to the Top of 2009, that specifically rely on students' test scores for performance evaluation and rewards for their teachers and schools. Like the U.S., many other countries around the globe (e.g. U.K., Canada, Slovakia and Qatar) worry about low test scores. For example, in February 2014, Elizabeth Truss, the U.K. education minister, visited Shanghai, China, whose test score is much higher than the U.K.'s, to "learn a lesson a math".

Oddly, many countries whose students excel in international exams are worried that their students spend too much time studying for exams! For example, the Wall Street Journal reports that "A typical East Asian high school student often must follow a 5 a.m. to midnight compressed schedule, filled with class instruction followed by private institute courses, for up to six days a week, with little or no room for socializing" (February 29, 2012), and that "many students prepare for [the national college] entrance exams from an early age, often studying up to 16 hours a day for years to take these tests" (November 10, 2011). This concern has influenced policy: the Education Ministry in China declared a ban on homework assignments for young children in August 2013, and South Korea declared a 10 pm curfew on private tutoring.

The fear is that the educational systems emphasize testing to such a degree that students do not effectively develop other useful skills, such as leadership, co-operation, and communication. While the importance of these non-cognitive skills has been clearly established in academic research (e.g. Heckman and Rubinstein 2001), their quantification and measurement remain challenging, because many of them do not show up in test scores (e.g. Heckman and Kautz 2012). Hanushek and Woessmann (2011) recognize that "the systematic measurement of such skills has yet to be possible in international comparisons".

In this paper, we present a general equilibrium (GE) framework that we have developed to quantify the quality of educational systems along multiple dimensions. The starting point of our framework is the observation that peoples' occupational choices reveal information about their skills at different types of tasks, and part of these skills have been developed through their education. For example, a manager issues directions and guidance to subordinates, a secretary follows these orders, while an engineer uses the knowledge in math and science to solve problems. We follow previous research (e.g. Autor, Levy and Murname 2003) and classify occupations as non-cognitive and cognitive. Because the people in non-cognitive (cognitive) occupations are primarily drawing on their non-cognitive (cognitive) human capital, by observing people's occupational choices we can quantify the qualities of a country's educational system along these dimensions.

To be specific, we model the educational system as production functions of cognitive and non-cognitive human capital, and use the TFP's (Total Factor Productivity) of these production functions to measure the qualities of the educational system. We call them cognitive and non-cognitive productivities. Our inspiration is the strong and intuitive intellectual appeal of TFP and its ubiquitous uses to measure the qualities of production technologies for countries, industries and firms.

In addition, researchers have long recognized that incentives matter for educational outcome.[1] We accommodate incentives in our model by having heterogeneous workers make optimal occupational choices given their own comparative advantages in non-cognitive and cognitive skills, as in Willis and Rosen (1979). These comparative advantages, in turn, are determined by innate abilities at birth and human capital accumulated through education. Therefore, when workers make educational choices, they factor in the returns of human capital on the labor market, recognizing that non-cognitive and cognitive occupations require different types of human capital. This implies that in our model, educational outcome is affected by occupational choices, which, in turn, depend on the non-cognitive and cognitive productivities of the educational system.

We contribute to the literature that compares educational outcomes using international test scores. For cognitive productivities we use test scores as the starting point, leveraging on the widely available test-score data and building on the insight of the em-

[1] In empirical studies using micro data, researchers have long recognized that incentives, in the form of money or even candy, improve the scores of IQ tests (Heckman and Kautz 2012). In a recent large-scale field experiment in Mexico, Behrman, Parker, Todd, and Wolpin (2015) show that providing monetary incentives to students has substantial and immediate effects on their test scores.

pirical literature on test scores (e.g. Hanushek and Woessmann 2011). We then peel back the confounding factors of resources and incentives under the guidance of our GE model, to reveal the underlying proficiency of the educational systems in fostering cognitive human capital. We show that countries' cognitive-productivity rankings are substantially different than their PISA-score rankings. In particular, those with the highest test scores do not necessarily have the highest cognitive productivities (e.g. S. Korea, Hong Kong).

Our non-cognitive productivities are a novel dimension of the quality of the educational system that is invisible from test scores. They carry into our GE model the insight of the large empirical literature that examines non-cognitive skills using micro data (e.g. Kuhn and Weinberger 2005, Heckman and Kautz 2012). We show that countries' non-cognitive-productivity rankings have zero correlation with their PISA score rankings. Many countries with low test scores have high non-cognitive productivities (e.g. the U.S. and U.K.).

Our model then allows us to condense the multi-dimensional differences in cognitive and non-cognitive productivities into a single metric for the overall educational quality. This metric is the weighted power mean of cognitive and non-cognitive productivities, the weights being the employment shares of cognitive and non-cognitive occupations. The power coefficients of this metric depend on the following three parameters: the dispersion of workers' innate abilities, which govens the supply-side elasticity of the economy; the substitution-elasticity across different types of human capital in aggregate production, which governs the demand-side elasticity of the economy; and the output elasticity in the production of human capital. To identify these key parameters, we draw on the parsimonious relationships predicted by our model among publicly available data, such as test score, output per worker, and employment shares of non-cognitive and cognitive occupations. The simple and transparent ways we identify our parameters, and our unique focus on educational quality and the production of human capital, distinguish our work from the quantitative literature on worker heterogeneity and income dispersion (e.g. Ohsornge and Trefler 2007, Hsieh, Hurst, Jones and Klenow 2016, Burnstein, Morales and Vogel 2016).

More broadly, we speak to the production technologies of non-cognitive and cognitive human capital using macro data, complementing the studies that do so using micro data (e.g. Cunha, Heckman and Schennach 2010). Since the educational system has deep historic roots for many countries, it is an important part of these countries' institutions. We thus also contribute to the institutions literature (e.g. Hall and Jones 1999, Acemoglu, Johnson and Robinson 2001) by quantifying key characters of the educational institution and drawing out their implications for aggregate output.

Ever since the 1983 report by the National Commission on Excellence in Education, there have been heated debates in the U.S. about the pros and cons of focusing on test scores. We bring the rigor of economic modeling and quantitative analyses into these discussions. In our model, education policies that focus on test scores tend to increase cognitive human capital, but create dis-incentives against non-cognitive human capital and may decrease its quantity in the aggregate economy. Our model quantifies these pros and cons and calculates the net effect on aggregate output. e.g. we show that if the U.S. were to have Hong Kong's educational system, the U.S. test score would rise but U.S. aggregate output would fall. Such calculations also provide a benchmark for the cost effectiveness and payoffs of education policies, and clarify that aggregate output is a better goal for education policies than test score. In doing so, we contribute to a large empirical literature using micro data to evaluate the effects of education policies on individual outcome (e.g. Figlio and Loeb 2011). While GE models are widely used for policy analyses in public, macro and international economics, they have not yet been used for education policies. Therefore, our GE model provides a novel and useful tool.

The remainder of this paper is organized as follows. Section 2 discusses the key facts that motivate our theoretical framework. Section 3 sketches this theoretical framework in a closed economy setting. Section 4 outlines the identification of our structural parameters. Section 5 draws out the implications of our non-cognitive and cognitive productivities. Section 6 explores the quantitative implications of our model. Section 7 extends the model to an open economy setting. Section 8 concludes.

2 Test Scores and Educational Spending, and Non-cognitive and Cognitive Occupations

A simple way to assess the productivity of a country's educational system is to use internationally comparable PISA test scores with educational spending per student, as is shown in Figure 1. This figure shows that more input (spending) leads to more output (test score), with substantial deviations from the best linear predictor (crude measure of productivity).

Missing from this naïve assessment is that the non-cognitive skills that are important in a modern work place are not well assessed by examinations, and that the ability of educational systems to foster these skills will be hard to compare internationally. Moreover, to the extent that a school system emphasizes easily measured skills, the educational system will look productive along this dimension, in part because students will have emphasized this part of their education more at the expense of less quantifiable skills.

We now demonstrate that occupations differ in the extent to which performance on test scores matters for workplace productivity. We use leadership to measure non-cognitive occupations. If the O*NET characteristic "providing guidance and direction to subordinates ..." is important for an occupation, we classify it as non-cognitive, and we classify all the other occupations as cognitive. We focus on leadership because it gives us intuitive and plausible correlation patterns in the micro data used by previous studies and also in our own micro data. To be specific, Kuhn and Weinberger (2005) use U.S. data to show that those who have leadership experiences during high school have higher wages later in their lives. In addition, we show below, in Table 1, that the wages of leadership occupations are less correlated with test scores than those of the other occupations, using the framework of Neal and Johnson (1996).

The data used in Table 1 is the 1979 NLSY (National Longitudinal Survey of Youth). The dependent variable is the log of individuals' wages in 1991, and the main explanatory variable is their AFQT score (Armed Force Qualification Test) in 1980, before they enter the labor force. Column 1 shows that the coefficient estimate of AFQT score is positive and significant, and this result replicates Neal and Johnson (1996).[2] Columns 2 and 3 show that AFQT score has a smaller coefficient estimate for the subsample of non-cognitive occupations than for the subsample of cognitive occupations.[3] To show this pattern more rigorously, we pool the data in column 4 and introduce the interaction between AFQT score and the non-cognitive-occupation dummy. The coefficient estimate of this interaction term is negative and significant.[4] In column 5 we use the O*NET characteristic of enterprising skills as an alternative measure for leadership. The interaction

[2] We include both men and women in Table 1, while Neal and Johnson (1996) do the estimation separately for men and women. We have experimented with this and obtained very similar results. We also use the same sample cuts as Neal and Johnson (1996) (see the Appendix for the details).

[3] Note that the coefficient estimates for AFQT square are not significant.

[4] Note that (1) we include the non-cognitive dummy itself, plus the college dummy and its interaction with AFQT score; (2) the non-cognitive dummy itself has a positive and significant coefficient estimate, consistent with Kuhn and Weinberger (2005).

between enterprising skills and AFQT score is negative but not significant. [5]

Having classified occupations as non-cognitive and cognitive using the U.S. O*NET, we next bring in employment-by-occupation data from the International Labor Organization (ILO). We keep only the countries whose raw data are in ISCO-88 (International Standard Classification of Occupations), because O*NET occupations can be easily mapped into ISCO-88 occupations but the mappings among other occupation codes are very scarce (e.g. we cannot find the mapping between Canadian and U.S. occupation codes). This leaves us with a single cross-section of 37 countries, and most of them are in 2000. Examples of non-cognitive occupations include business professionals (ISCO-88 code 2410), managers of small enterprises (1310), building frame and related trades workers (7120), nursing and midwifery professionals (3230), etc. Examples for cognitive occupations include architects, engineers and related professionals (2140), finance and sales professionals (3410), secretaries (4110), motor vehicle drivers (8320), etc.

We then merge in mean PISA scores in reading, math and science from the official PISA website, and the ratios of private plus public expenditures on education to GDP in 2004 from the UNESCO Global Education Digest of 2007.[6][7] Finally, we add other variables, such as labor-force size and aggregate output, from standard sources, such as NIPA (National Income and Product Account) and PWT (Penn World Tables). Because we do not have physical capital in our model, as we show in section 3, we use labor

[5] We have also experimented with using the following O*NET characteristics to measure non-cognitive occupations: investigative skills, originality, social skills, and artistic talents. The results for originality, investigative skills and social skills are counter-intuitive, and artistic talents account for a very small fraction of the labor force. See the Appendix for more details.

[6] When PISA first started in 2000, only the reading test was administered, and only a small set of countries participated (e.g. the U.K. and Netherlands did not participate). In order to obtain PISA scores in all three subjects for every country in our sample, we calculate simple averages over time by country by subject, using all years of available data; e.g. Germany's PISA math score is the simple average of 03, 06, 09 and 2012, U.K.'s reading score the average of 06, 09 and 2012, etc. In the Appendix we show that PISA scores have small over-time variation but large cross-section variation.

[7] We use the scores of PISA, which tests high-school students, rather than the scores of adult tests, for two reasons. First, scores of high-school students are highly correlated with those of adults (e.g. Brown et al. 2007, Hanushek and Zhang 2009, Heckman and Kautz 2012). When we regress the 2012 PISA scores on the 2013 PIAAC (Program for the International Assessment of Adult Competencies) scores, we obtain coefficient estimates close to 1 and high R2 (Appendix Table A3). Second, adult tests cover fewer countries than PISA, and would cut our sample size by at least 25%. See the Appendix for more details.

income, or compensation of employees from NIPA, as our measure for aggregate output.[8] In addition, in developing countries (e.g. Egypt), large fractions of the labor force are engaged in subsistence farming, and it is unclear how to think about the contributions of human capital to subsistence farming. We thus drop all the developing countries, reducing our sample to 28 high-income countries. They account for 42.94% of world GDP in 2000. Table 2 provides summary statistics of our main variables of interest, and Table 3 lists the countries and years in our sample. Note that all our data come from public sources.

3 Educational Quality in the Closed Economy

In this section we develop our model for the closed-economy and illustrate the intuition of our key parameters. We also show how the model can make contact with observable country outcomes with an eye toward identification and quantification, in preparation for section 4.

3.1 Model Assumptions

There are K countries, indexed by k, each endowed with L^k units of heterogeneous labor. Workers are endowed with non-cognitive and cognitive attributes ε_n and ε_c, drawn from the following Frechet distribution

$$F(\varepsilon_n, \varepsilon_c) = \exp\left(-\left(T_c\varepsilon_c^{-\theta} + T_n\varepsilon_n^{-\theta}\right)^{1-\rho}\right), \ \theta \equiv \frac{\widetilde{\theta}}{1-\rho} \tag{1}$$

As we discussed in sub-section 2.1, we think about the attributes n and c as two distinct packages of skills, rather than two individual skills. In equation (1), the parameter ρ captures the degree to which non-cognitive and cognitive attributes are correlated. When $\rho = 0$, they are independent; when $\rho > 0$, they have positive correlation; and when $\rho \to 1$, they become perfectly collinear. The parameter θ captures the dispersion of attributes across workers. As θ rises, the distribution becomes more compressed, and so there is less worker heterogeneity. Note that for the distribution to have finite variance, we require $\theta > 1$. Finally, T_c and T_n, both positive, capture the locations of the

[8] We experimented with stripping capital from GDP by assuming a Cobb-Douglas production function and using the parameter values from the macro literature (e.g. Klenow and Rodriguez-Clare 1997). The aggregate output of this second approach has a correlation of 0.9994 with our main output variable.

attributes distribution; e.g. as T_c rises, the distribution of cognitive abilities shifts to the right, so that the average worker has better innate cognitive abilities. We assume that ρ, θ, T_c, and T_n do not vary across countries.[9]

To minimize the number of moving parts, we follow Hsieh et al. (2016) and model the education system as a human-capital production machine. Workers accumulate human capital of type i, $i = n$ (non-cognitive) or c (cognitive), according to the technology

$$h_i(e) = h_i^k e^{\eta},\ i = c, n. \tag{2}$$

In equation (2), e is an individual worker's spending on education, in units of the final good (we specify its production below). The parameter η captures decreasing returns in the production of human capital, and guarantees an interior solution for workers' optimal educational spending. We assume that η is common across countries. The parameters h_n^k and h_c^k are country k's productivities in non-cognitive and cognitive human capital, and they capture the strength of country k's educational system along these two dimensions, net of resources inputs.

We treat h_n^k and h_c^k as exogenous, because the educational institution has deep historic roots in many countries. For example, in the U.S., private universities and colleges are a main feature of the educational institution, and their legal rights and status were enshrined by the Supreme Court in 1819 in Dartmouth-College-vs-Woodward.[10] In S. Korea, and many other East Asian countries, the national exam has been a cornerstone of the educational institution for over 1,000 years.[11] We capture, and quantify, such

[9]The assumption over ρ and θ is standard in the literature using Roy model models. The assumption that the Ts are same is that there are no inherent genetic differences across countries.

[10]In 1816, New Hampshire enacted state law to convert Dartmouth College from a private institution to a state institution. The case went to the U.S. Supreme Court, the legal issue being whether Dartmouth's original charter with the King of England should be upheld after the American Revolution. In 1819, the Supreme Court sided with Dartmouth, and this decision also guaranteed the private status of other early colonial colleges, such as Harvard, William and Mary, Yale, and Princeton (e.g. Webb, Metha, and Jordan 2013).

[11]China used archery competitions to help make promotion decisions for certain bureaucrateic positions before 256 B.C.E. and established the imperial examination system as early as 605 A.D. and this remained in use for over 1,000 years. In this system, one's score in the national exam determines whether or not he is appointed to a government official, and if so, his rank. Through trade, migration, and cultural exchanges, China's imperial examination system spread to neighboring countries; e.g. Korea established a similar system in 958 A.D. (Seth, 2002).

cross-country differences in educational institutions as h_n^k and h_c^k, and so we place no restriction on their values.

Both non-cognitive and cognitive tasks are needed to produce the final good. When a worker chooses task i, or occupation i, her output is

$$h_i(e)\varepsilon_i, \;\; i = n, c \tag{3}$$

where $h_i(e)$ is the worker's human capital, accumulated according to the technology (2), and ε_i her attribute, drawn from the distribution in (1).[12] The representative firm hires workers in both cognitive and non-cognitive occupations to maximize output

$$y^k = \Theta^k \left(A_c \left(L_c^k \right)^{\frac{\alpha-1}{\alpha}} + A_n \left(L_n^k \right)^{\frac{\alpha-1}{\alpha}} \right)^{\frac{\alpha}{\alpha-1}} \tag{4}$$

In equation (4), Θ^k is country k's total-factor productivity (TFP), and A_c and A_n common technological parameters. The parameter $\alpha > 0$ is the substitution elasticity between non-cognitive and cognitive skills. L_n^k and L_c^k are the sums of individual workers' outputs of non-cognitive and cognitive tasks, which are specified in equation (3). L_n^k and L_c^k can also be interpreted as country k's aggregate supplies of non-cognitive and cognitive human capital. We use the final good as the numeraire, and assume, for now, that it cannot be traded. We also assume that workers are immobile across countries. These assumptions imply that no trade takes place. In section 7 we allow for free trade in the services of human capital.

3.2 Equilibrium

We begin this section by deriving labor supply by occupation. We analyze workers' occupational choices in two steps. We first solve for workers' optimal choices for education, given their occupational choices. This allows us to characterize the workers' highest income levels by occupation and so allows us to solve for workers' occupational choices. We then aggregate over the skills of these workers supplied to the labor market.

Let w_n^k and w_c^k denote, respectively, the earning of one unit of non-cognitive and cognitive human capital. By equations (2) and (3), a worker with attributes ε_i, $i =$

[12] Equation (3) assumes that occupation i uses skill i. We have experimented with having occupations use both skills, with occupation i being more intensive in skill i. This alternative specification adds little insight but much complexity so we have gone with the simpler set up.

n, c, receives income $w_i^k h_i^k e^\eta \varepsilon_i$. This worker, then, chooses the quantity of education to maximize $w_i^k h_i^k e^\eta \varepsilon_i - e$. With $0 < \eta < 1$, we are guaranteed the following interior solution, which is given by

$$e_i = (\eta w_i^k h_i^k \varepsilon_i)^{\frac{1}{1-\eta}} \tag{5}$$

Equation (5) says that a worker in country k accumulates more education if she is talented (high ε_i), if the skill she learns through education pays well in the job market (high w_i^k), or if country k's educational system provides high quality education (high h_i^k).

We now plug the worker' optimal educational choice in (5) into her maximization problem, and obtain the following expression for her highest net income in occupation i

$$I_i(\varepsilon_i) = (1-\eta)\eta^{\frac{\eta}{1-\eta}} (w_i^k h_i^k \varepsilon_i^k)^{\frac{1}{1-\eta}} \tag{6}$$

Equation (5) and (6) show that net income, $I_i(\varepsilon_i)$, is proportional to educational spending, $e_i(\varepsilon_i)$. This result will be handy for our analyses below. In addition, equation (6) implies that the worker chooses occupation n if and only if $w_c^k h_c^k \varepsilon_c^k \leq w_n^k h_n^k \varepsilon_n^k$. This is a classic discrete-choice problem (e.g. McFadden 1974). Using the Frechet distribution (1) we show, in the Appendix, that

Proposition 1 *The employment share of occupation i equals*

$$p_i^k = \frac{T_i (w_i^k h_i^k)^\theta}{T_c (w_c^k h_c^k)^\theta + T_n (w_n^k h_n^k)^\theta},\ i = c, n. \tag{7}$$

Equation (7) says that the non-cognitive employment share, p_n^k, is high, if workers have a strong comparative advantage in non-cognitive innate abilities (high T_n/T_c), non-cognitive skills have a high relative return in the labor market (high w_n^k/w_c^k), or country k's educational system has a strong comparative advantage in fostering non-cognitive human capital (high h_n^k/h_c^k). Note the role of θ in equation (7). As θ rises and workers become more homogeneous, smaller changes in wages or educational efficiencies lead to bigger shifts in the proportion of workers that opt to work in different occupations.

To solve the model, we start by calculating the average net income of non-cognitive and cognitive workers, which analytically involves taking the expected value of equation (6), with respect to ε_i, conditional on type i, $i = n, c$. We show, in the Appendix, that

Proposition 2 *The average net income is the same for non-cognitive and cognitive workers; i.e.*

$$I_n^k = I_c^k = \gamma(1-\eta)\eta^{\frac{\eta}{1-\eta}} \left(T_c(w_c^k h_c^k)^\theta + T_n(w_n^k h_n^k)^\theta\right)^{\frac{1}{\theta(1-\eta)}}, \gamma = \Gamma(1 - \frac{1}{\theta(1-\rho)(1-\eta)}) \quad (8)$$

Proposition 2 is a common feature of the solution to discrete choice problems where the underlying distribution is Frechet (e.g. Eaton and Kortum 2002). In equation (8), the term in the square brackets is the denominator of the employment-share expression, (7). $\Gamma(.)$ is the Gamma function and so γ is a constant. Proposition 2, together with equations (5) and (6), implies that

Corollary 1 *The average educational expenditure is the same for non-cognitive and cognitive workers and is equal to*

$$E_n^k = E_c^k = \gamma\eta^{\frac{1}{1-\eta}} \left(T_c(w_c^k h_c^k)^\theta + T_n(w_n^k h_n^k)^\theta\right)^{\frac{1}{\theta(1-\eta)}} \quad (9)$$

By the Corollary we now use E^k, without an occupation subscript, to denote the average educational spending in country k. Proposition 2 and its corollary will prove useful in pinning down the elasticity of human capital accumulation with respect to educational expenditure, as we show in section 4.

We next derive the aggregate supplies of non-cognitive and cognitive human capital in country k, $i = n, c$. It is the number of workers in occupation i, $L^k p_i^k$, times the average output within occupation i. By equations (2) and (3), the output of an individual occupation-i worker is $h_i^k e^\eta \varepsilon_i$, where e is given by equation (5). The average output within occupation i is then $E\left(h_i^k e^\eta \varepsilon_i | \text{Occupation } i\right)$, where the conditional expectation is with respect to ε_i. By equations (2), (3) and (5), $h_i^k e^\eta \varepsilon_i = (\eta w_i^k)^{\frac{\eta}{1-\eta}} \left(h_i^k \varepsilon_i\right)^{\frac{1}{1-\eta}}$. We show, in the Appendix, that

Proposition 3 *The aggregate supply of type-i human capital is*

$$L_i^k = L^k p_i^k E\left(h_i^k e^\eta \varepsilon_i | \mathit{Occupation}\ i\right) = \gamma L^k p_i^k \left[h_i^k (\eta w_i^k)^\eta \left(\frac{T_i}{p_i^k}\right)^{\frac{1}{\theta}}\right]^{\frac{1}{1-\eta}} \quad (10)$$

where the occupational employment shares, p_i^k, are given by (7).

To complete our characterization of labor supply, we use equations (7) and (10) to derive the relative supply of non-cognitive labor, which is given by

$$\frac{L_n^k}{L_c^k} = \frac{T_n}{T_c}\left(\frac{h_n^k}{h_c^k}\right)^{\theta}\left(\frac{w_n^k}{w_c^k}\right)^{\theta-1} \tag{11}$$

Intuitively, non-cognitive labor supply is increasing in the availability of raw talent in the country, the comparative advantage of that country in non-cognitive education, and the relative wage of the non-cognitive occupation. As foreshadowed by our discussion of Proposition 1, it is clear from equation (11) that θ is the supply elasticity: as workers' skills become more homogeneous their labor is more substitutable across occupations.

Having completed our analysis of the supply side for cognitive and non-cognitive skills, we turn our attention to the demand side. Cost minimization by final goods producers facing technology (4) determines the demand for cognitive and non-cognitive labor. The first order conditions imply that the relative demand for non-cognitive labor is given by

$$\frac{L_n^k}{L_c^k} = \left(\frac{A_c}{A_n}\frac{w_n^k}{w_c^k}\right)^{-\alpha} \tag{12}$$

Equation (12) is a standard labor demand equation where the key demand elasticity is given by α. When $\alpha > 1$ an increase in the relative wage of non-cognitive labor results in sufficiently large substitution that the cost share of non-cognitive labor in GDP falls, whereas when $\alpha < 1$ substitution is so low that the cost share rises. Of course, in the Cobb-Douglas case of $\alpha = 1$ the cost shares are fixed. The magnitude of α, which we will estimate using the model and data, plays an important qualitative role in the model.

Using labor supply, given by (11), and labor demand, given by (12), we can solve for equilibrium relative wages and relative labor quantities by country. We have

$$\frac{w_n^k}{w_c^k} = \left[\frac{T_c}{T_n}\left(\frac{h_c^k}{h_n^k}\right)^{\theta}\left(\frac{A_n}{A_c}\right)^{\alpha}\right]^{\frac{1}{\phi}} \tag{13}$$

where $\phi \equiv \theta + \alpha - 1 > 0$, and

$$\frac{L_n^k}{L_c^k} = \left(\frac{A_c}{A_n}\frac{p_n^k}{p_c^k}\right)^{\frac{\alpha}{\alpha-1}} = \left[\frac{T_n}{T_c}\left(\frac{h_n^k}{h_c^k}\right)^{\theta}\left(\frac{A_n}{A_c}\right)^{\theta-1}\right]^{\frac{\alpha}{\phi}} \tag{14}$$

From (13) and (14) it is clear that the countries with a comparative advantage in non-cognitive schooling will have lower relative non-cognitive occupation costs and a greater relative quantity of this type of labor.

Having solved for the equilibrium allocations of labor to different occupations as a function of countries' educational characteristics, we can now solve for output per worker across countries as a function of countries' TFP differences in final good production and differences in the quality of their educational infrastructure. To do so, we first define a base country 0 against which any particular country can be compared. As we show in the appendix, by substituting labor allocations (10) into (4) and by simplifying using (7) and (18), we can write GDP per capita in country k relative to the base country 0 as

$$\frac{y^k/L^k}{y^0/L^0} = \left[\frac{\Theta^k}{\Theta^0}\right]^{\frac{1}{1-\eta}} \times \left[\Omega^k\right]^{\frac{1}{1-\eta}} \tag{15}$$

where

$$\Omega^k \equiv \left(p_c^0 \left(\frac{h_c^k}{h_c^0}\right)^{\frac{\theta(\alpha-1)}{\phi}} + p_n^0 \left(\frac{h_n^k}{h_n^0}\right)^{\frac{\theta(\alpha-1)}{\phi}}\right)^{\frac{\phi}{\theta(\alpha-1)}} \tag{16}$$

The first term in equation (15), $\left[\Theta^k/\Theta^0\right]^{1/(1-\eta)}$, is the variation across countries in their GDP per capita that is due to Hick's neutral productivity differences. The sources of these differences could be associated with many factors, such as efficient court systems and business regulations. Note that because higher productivity increases the return to education, the effect of TFP is amplified by the power $1/(1-\eta)$.

The second term in parentheses $[\Omega^k]^{1/(1-\eta)}$ is the portion of per capita income differences across countries that is due to the quality of their educational system. The term Ω^k is a weighted power mean of the cognitive and non-cognitive productivities, with the weights being the occupational employment shares of the base country. This measure is akin to an index, summarizing the multi-dimensional differences in educational quality into a single numerical value, and it captures the contribution of the overall quality of the educational institution to output per capita.

Because the powers in the Ω^k are determined by the demand and supply elasticities, they play important roles in determining how the quality of a country's educational system, Ω^k, depends on the two types of educational TFP. As both θ, $\alpha \to \infty$, Ω^k goes to the maximum of the two educational qualities. This is intuitive as workers become equally capable at both perfectly substitutable tasks. In this case, having an education system that has highly uneven quality has few consequences for a country's well-being. As $\alpha \to -\infty$, however, so that production becomes Leontief, then Ω^k goes to the minimum of the two educational TFPs, and excelling along a single dimension does little good for national well-being. For the more empirically relevant case found in our data (see below) Ω^k is reasonably well approximated as a geometric mean. In this case, the relative importance of the two dimensions of educational TFP are determined by the occupational shares.[13]

3.3 Theory and Measurement

In this subsection, we show how the model can make contact with observable country outcomes with an eye toward identification and quantification in the next section. We begin by making an additional assumption that allows international test scores, when combined with parameter values, to reveal a country's absolute advantage in cognitive education. Specifically, we assume that the average test score of country k, S^k, is proportional to the average cognitive human capital in that is accumulated in a country; i.e.

$$S^k = b\frac{L_c^k}{L^k},\, b > 0. \tag{17}$$

We make this assumption because a large body of empirical work (see, e.g. Hanushek and Woessman 2011 for a survey) suggests that scores of international assessment tests, such as PISA, are good measures of cognitive skills.[14]

Proposition 3 and equation (17) imply that a country k's test score, relative to a reference country 0, equals

$$\frac{S^k}{S^0} = \left(\frac{E^k}{E^0}\right)^{\eta} \left(\frac{p_c^k}{p_c^0}\right)^{1-\frac{1}{\theta}} \left(\frac{h_c^k}{h_c^0}\right) \tag{18}$$

As expression (18) makes clear, a good showing on international tests can happen for multiple reasons. First, a high test score could be obtained by a high level of spending on education per capita, E^k. The effect of E^k on cognitive human capital, and so test score, is raised to the power of η, because the production technology of human capital, (2), is subject to diminishing returns.

The second term in (18) captures the effects of incentives and selection, and they arise in general equilibrium because heterogeneous individuals make optimal educational

[13]Note that the less substitutable are the two types of skills in the population (smaller α) the bigger the penalty toward poor performance in jourt one of the dimensions of educational investments.

[14]Using micro data, Cunha, Heckman, and Schennach (2010) show that indviduals' test scores are informative about their cognitive skills.

choices. To see these effects, suppose p_c^k is high in country k; i.e. a larger fraction of the labor force favors the cognitive occupation over the non-cognitive occupation. This could be because the relative return to cognitive skills in country k, w_c^k/w_0^k, is high, or country k's educational institution has a strong comparative advantage in fostering cognitive human capital (i.e. h_c^k/h_n^k is high). In either case, the cognitive occupation is an attractive career choice in country k, and so individuals have strong incentives to accumulate cognitive human capital. This incentive effect implies high average test score for country k, and its magnitude is raised to the power of 1.

On the other hand, workers are heterogeneous, and so a high p_c^k implies that many individuals with low innate cognitive abilities have self-selected into the cognitive occupation. Their presence tends to lower the average cognitive human capital, and so the test score. The magnitude of this selection effect is p_c^k raised to the power of $-1/\theta$. If θ is large, the distribution of innate abilities becomes more compressed. This means less individual heterogeneity and so the selection effect is weaker. Note that because $\theta > 1$ the incentive effect always dominates. We allow the data to steer us to the most appropriate value for θ, and it will turn out that the value does indeed exceed one.

Finally, cognitive productivity, h_c^k, soaks up all the other reasons why the test score is high for country k, net of the effects of resources, and incentives minus selection. In this sense, is the TFP of the educational institution along the cognitive dimension. An important implication of equation (18) is that a country's absolute advantage can be identified given a measure of the dispersion of skills in the population, the elasticity of human capital with respect to educational spending, and data on test scores, educational expenditures, and the share of the population that opts to work in cognitive occupations.

Now, exploiting equations (7) and (13) and measuring country k relative to base country 0, we can derive an expression that yields country k's relative advantage in cognitive education:

$$\frac{p_c^k/p_n^k}{p_c^0/p_n^0} = \left(\frac{h_c^k/h_n^k}{h_c^0/h_n^0}\right)^{\theta\frac{\alpha-1}{\phi}} \tag{19}$$

Expression (15) shows that given dispersion and substitution parameters and occupational choices by country, we can identify a country's comparative advantage. This result has the flavor of revealed comparative advantage. Given the endogenous choices of workers and the optimal hiring decisions of the final goods producers, the observed shares of workers in each occupation reveals the relative qualities of the educational system.

Note the importance of the magnitude of the parameter α. If $\alpha > 1$, the size of the change in relative demand to a shift in the relative cost of skills is small so that occupational choices dominate and a larger proportion of cognitive skills reflects a greater comparative advantage in cognitive education. When $\alpha < 1$, the skills are strong complements and high wages induced by poor educational efficiency reverses the relationship between comparative advantage and the proportion of workers in cognitive occupations. Finally, in the knife-edge case of $\alpha = 1$, the proportion of workers in each occupation is independent of comparative advantage, implying that non-cognitive employment shares are completely un-informative about relative quantities of non-cognitive human capital. We allow the data to steer us to the most appropriate value for α.

4 Identification of Structural Parameters

As we have shown in the previous section, given the elasticities, θ, η, and α, and data L^k, labor-force size, y_k, aggregate output, p_n^k, non-cognitive employment share, and p_c^k, cognitive employment share by country, we can identify country-level TFP, Θ^k, and educational TFPs, h_n^k and h_c^k. As these data are readily available, the challenge is to obtain parameter values for these elasticities. This section discusses our estimation of these elasticities.

We begin with η, the elasticity of human capital attainment with respect to educational expenditure. Corollary 1 and Proposition 3 imply that

Corollary 2 *Country k spends fraction η of its aggregate output on education; i.e.*

$$E^k L^k = \eta y^k. \tag{20}$$

Proof. By equations (9) and (10), $w_i^k L_i^k = L^k p_i^k E^k / \eta$, and so $\eta \left(\sum_i w_i^k L_i^k\right) = L^k E^k \left(\sum_i p_i^k\right) = E^k L^k$. In our model aggregate output equals aggregate income, and so $\eta \left(\sum_i w_i^k L_i^k\right) = \eta y^k$. ■

By equation (20), η is the ratio of aggregate educational spending, $E^k L^k$, to aggregate output, y_k. Therefore, we set its value to match the mean share of public plus private educational expenditure in output, 0.1255 (see Table 3); i.e. $\eta = 0.1255$.

We now turn to our estimation of θ, which measures the dispersion of innate abilities across workers and also governs the elasticity of the aggregate supplies of human capital. Using equation (18), we obtain

$$\ln\left(\frac{S^k}{(y^k/l^k)^\eta}\right) = D + \left(1 - \frac{1}{\theta}\right) \ln p_c^k + \ln h_c^k \tag{21}$$

where D is a constant. Equation (21) decomposes the cross-country variation in the average test score, S^k, into resource inputs, $(y^k/L^k)^\eta$, incentives (minus selection), p_c^k, and cognitive productivity, h_c^k.[15]

Equation (21) also instructs us to construct variables and to look for novel correlation patterns that previous research has not examined. We follow these instructions in Figure 2. The vertical axis is log PISA math score, normalized by the logarithm of output per worker raised to the power of η. The horizontal axis is log cognitive employment share. We weigh the data in the scatterplot using aggregate output.[16]

Figure 2 clearly illustrates that, consistent with equation (21), the countries in which workers are clustered in cognitive occupations are the countries that score well on tests, which can measure primarily cognitive achievement. The best-fit line has $R^2 = 0.288$ and a slope coefficient of 0.717. This novel correlation pattern provides an important validation that incentives indeed matter for the accumulation of human capital, a key mechanism of our general-equilibrium model.

Figure 2 also allows us to interpret the correlation pattern as structural parameters of our model, because it follows the exact specification of equation (21). The slope coefficient of the best-fit line corresponds to the coefficient of log cognitive employment share, $\left(1 - \frac{1}{\theta}\right)$, implying that $\theta = 3.4965$. This estimate for θ provides yet another validation of our model, which, as we discussed in section 3, requires $\theta > 1$. The countries' deviations from the best-fit line then correspond to the log of their cognitive productivities, h_c^k.

Furthermore, Figure 2 illustrates the intuition for the identification of θ. As we discussed earlier, with individual heterogeneity, selection moderates the effect of incentives on average cognitive human capital. A small θ implies high heterogeneity and strong selection effect. This means we should observe limited variation in the normalized test scores despite substantial variation in cognitive employment shares; i.e. log cognitive employment share should have a small slope coefficient in Figure 2. Therefore, we identify θ through the strength of the selection effect, the magnitude of which is $-1/\theta$ according to our model.

One may wonder whether cognitive employment share is correlated with cognitive

[15] Relative to (18), (21) has output per worker rather than educational expenditure per worker, because we have more data points on output per worker than for average educational expenditure per capita.

[16] The countries in our sample vary a lot in their size (e.g. Switzerland, Germany, and the United States.)

productivity in (21), and whether this correlation is an issue for the way we calibrate our structural parameters. We use equations (7) and (13) to show that

$$p_c^k = \frac{1}{1 + \left(\frac{A_n}{A_c}\right)^{\frac{\alpha\theta}{\phi}} \left(\frac{T_n}{T_c}\left(\frac{h_n^k}{h_c^k}\right)^{\theta}\right)^{\frac{\alpha-1}{\phi}}}$$

This expression clarifies that cognitive employment share is determined by the ratio of cognitive productivity to non-cognitive productivity, h_c^k/h_n^k, or the comparative advantage of the educational institution. Therefore, cognitive employment share is uncorrelated with cognitive productivity if cognitive productivity, a measure of the absolute advantage of the educational institution, is uncorrelated with its comparative advantage. We also look at alternative specifications below, and compare our estimates with those from the literature.

Table 4 shows the results of fitting our data using (21), implemented as a regression with aggregate output as weight. Column (1) corresponds to the best-fit line in Figure 2. In column (2) we add Australia and New Zealand but dummy them out,[17] and in column (3) we use labor-force size as weight. The results are very similar to column (1). In column (4) we use PISA reading score. The coefficient becomes smaller, 0.521, and remains significant, implying that $\theta = 2.0877$. Column (5) has PISA science score and the results are similar to column (4). Column (6) uses the O*NET characteristic of enterprising skills as an alternative measure of leadership, and so non-cognitive occupations. The coefficient is positive but not significant, and this pattern echoes column (3) of Table 2.[18]

Table 4 produces a range of values for θ, 2.0877˜3.4965. We use $\theta = 3.4965$ in the rest of the paper and show, at the end of this section, that our estimates are very similar to the literature, and that we get very similar results if we use other values for θ (e.g. 2.0877) instead.

We then calculate the residuals and construct cognitive productivities, h_c^k, according to (21). Like the TFP estimates in the growth literature (e.g. Hall and Jones, 1999), our estimates for cognitive productivities are relative, and so we normalize the U.S. value to 1.

[17] As discussed in subsection 2.2, these countries have different occupation classification codes in their raw data.

[18] We present the results of alternative measures of non-cogntivie occupations in Appendix Table 4A.

We estimate α, the substitution elasticity on the demand side, using the aggregate production function (4). Specifically, we substitute out the quantities of human capital, L_c^k and L_n^k, using equations (10), (17) and (14). After some algebra we obtain

$$\ln\left(\frac{y^k}{L^k S^k}\right) = F + \frac{\alpha}{\alpha - 1}\ln\left(1 + \frac{p_n^k}{p_c^k}\right) + \ln\Theta^k \tag{22}$$

where the constant F has no cross-country variation.

Equation (22) is an input-output relationship. The output is y^k, and there are two inputs. The first is the quantity of cognitive human capital, represented by $L^k S^k$, since test score, S^k, represents average cognitive human capital by equation (17). The second input is the relative quantity of non-cognitive human capital, which can be represented by p_n^k/p_c^k by equation (14). Therefore, equation (22) shows how aggregate output, normalized by the quantity of cognitive human capital, varies with the relative quantity of non-cognitive human capital, and this variation identifies α.

The estimation of (22), then, is similar to the estimation of the aggregate production function.[19] The coefficient of $\ln\left(1 + \frac{p_n^k}{p_c^k}\right)$ gives us α, and the residuals give us Θ^k, the output TFP. In the estimation, our data disciplines our model for two reasons. First, (22) instructs us to use the average test score as one input and the ratio of employment shares as the relative quantity of another input. These are novel ways to measure the quantities of human capital that previous research has not considered. In addition, our model needs $\alpha > 1$, as we discussed for equation (13), and to make this inference the coefficient of $\ln\left(1 + \frac{p_n^k}{p_c^k}\right)$ must exceed 1.

Table 5 shows the results of fitting our data using (22), implemented as a regression with aggregate output as weight. The structure of Table 5 is similar to Table 4 and so are the flavors of the results. Columns (1), (4) and (5) use PISA math, reading and science scores, respectively. Column (2) drops Austalia and New Zealand, and column (3) uses labor-force size as weight. The coefficients are all significant, ranging from 2.923 to 3.125. Using 3.125 we infer that $\alpha = 1.4706$. Column (6) uses enterprising skills as the alternative measure for non-cognitive occupations, and the coefficient is positive but not significant, echoing Tables 2 and 4.

[19] As in the growth literature, we implicitly assume that output TFP is uncorrelated with relative quantity, which in our case is determined by the comparative advantage of a country's educational system. While progress has been made in the micro literature with respect to identification, it has been slower in the cross-country macro literature.

We then calculate the residuals and construct the output TFP, Θ^k, according to (22), normalizing the U.S. value to 1. We check the correlation coefficients between our output TFP estimates and those reported in the literature. They are all positive and significant, ranging from 0.4674 (Klenow and Rodriguez-Clare 1997) to 0.6377 (PWT 8.0), and provide an external validation for our approach.[20]

Now we go back to equation (19) and use our estimates for θ and α to obtain h_n^k/h_c^k, the comparative advantage of the educational institution. We implement equation (19) as a regression with only the constant and no explanatory variable. The constant soaks up the variables of base-country 0, and the residuals allow us to calculate h_n^k/h_c^k. Here, we use the United States as the base country so that the value for the U.S. is 1. The values of h_n^k/h_c^k then allow us to compute h_n^k, the non-cognitive productivities.

Table 6 summarizes our parameter values and how we identify them. In comparison, Hsieh et al (2016)'s model features the same Frechet distribution of innate abilities as ours, but for identification they use worker-level data and explore wage dispersion within occupations and labor-force participation; i.e. their data and identification strategy are completely different from ours. Despite such differences, Hsieh et al (2016)'s θ estimate ranges from 2.1 to 4, matching ours. On the other hand, Burnstein et al. (2016) features a CES aggregate production function, like us, but for identification they use cross-section and over-time variations in occupational wages and employment in micro data. Although Burnstein et al. (2016)'s data and identification strategy are completely different from ours, their substitution-elasticity estimate ranges from 1.78 to 2, similar to ours.[21]

We now perform sensitivity analyses. We first use $\theta = 2.0877$ and PISA reading scores to obtain cognitive and non-cognitive productivities and rank countries using these alternative estimates. We then calculate the correlation coefficients of these alternative values and rankings with our main specification, and report them in Table 7. These correlation coefficients range from 0.9583 to 1.0000. We next consider $\theta = 2.0877$ and $\alpha = 2$. As Table 7 shows, the values and rankings of alternative cognitive and non-cognitive productivities are again highly correlated with our main specification.

[20]See Appendix Table A4 for all the pairwise correlation coefficients.

[21]The substitution-elasticity parameter is not identified in Hsieh et al. (2016). Burnstein et al. (2016), on the other hand, do not model the production of human capital.

5 Cognitive and Non-cognitive Productivities

Having identified the values of h_c^k and h_n^k in section 4, we present them in this section and draw out their potential implications for education policies.

5.1 Cognitive Productivities

Figure 3 plots the countries' rankings in h_c^k against their rankings in PISA math score, and Table 3 lists these rankings by country. These two rankings are positively correlated (0.5101), since both test score and cognitive productivity measure the quality of the educational system along the cognitive dimension. However, Figure 3 shows that they are quite different for many countries. We highlight these differences using the 45 degree line.

These differences arise because test score is an outcome, and so a noisy measure for the underlying quality of cognitive education. Equation (21) highlights two sources of noisiness. The first is resources, $(y^k/L^k)^\eta$. Other things equal, an educational system with more resources is expected to produce better outcome. The second is incentives (minus selection), $\left(1-\frac{1}{\theta}\right)\ln p_c^k$. The country where individuals are strongly incentivized to learn cognitive skills will perform well in international tests. Equation (21) then allows us to use test score, S^k, as the starting point, and remove the effects of resources and incentives, to arrive at our cognitive productivity, h_c^k. Therefore, cognitive productivity is a cleaner measure for the underlying quality of cognitive education than test score.

Consider, first, Poland, Czech Republic, Hungary and Slovakia. They have decent PISA scores, ranked outside of top 10. However, our model says that this outcome should be viewed in the context of low output per worker in these countries, and so limited educational resources. Therefore, the qualities of their educational systems are better than their test scores suggest, and they all rank within top 10 based on cognitive productivities.

Now consider Hong Kong, South Korea and Switzerland. They are superstars in PISA scores, all ranked within top 5. However, our model says that this outcome should be viewed in the context of high cognitive employment shares and so strong incentives to accumulate cognitive human capital. Therefore, the qualities of their educational systems are not as good as their test scores suggest, and their rankings drop to 10, 12 and 14, respectively, by cognitive productivities.

Finally, we look at the U.S. First, the U.S. has very high output per worker. The abundance of resources makes the low U.S. PISA scores even harder to justify. Second, the employment share of cognitive occupations is relatively low in the U.S., implying weak incentives to accumulate cognitive human capital. The effects of resources and incentives thus offset each other, leaving the U.S. ranking in cognitive productivities very close to its ranking in PISA scores, near the bottom in our set of 28 countries. In our Introduction, we discussed the worries and concerns about the quality of the U.S. educational system. Figure 3 quantifies these concerns and shows that they are well justified, when we look at the cognitive dimension. We now move on to the non-cognitive dimension.

5.2 Non-cognitive Productivity

Figure 4 plots the countries' rankings in h_n^k against their rankings in PISA math score, and Table 3 lists the rankings by country. Figure 4 clearly shows that the PISA-math rankings are simply not informative about non-cognitive productivity rankings (correlation $= -0.0602$ with p-value $= 0.7609$). Thus non-cognitive productivities allow us to compare countries' educational systems in a novel dimension, hidden from PISA scores.

In our Introduction, we discussed the concerns in S. Korea and many East Asian countries that the educational systems emphasize exams so much that students are unable to develop non-cognitive skills. Our results in Figure 4 quantify this issue and suggest that these concerns are well grounded. S. Korea and Hong Kong, super starts in terms of PISA scores, round up the very bottom among our 28 countries. Their very low non-cognitive productivities are because of low relative employment shares of non-cognitive occupations, and good-but-not-stellar cognitive productivities.

Figure 4 also shows that PISA-math rankings substantially understate the proficiency of the U.S. and U.K. educational systems in fostering non-cognitive skills. The U.S. ranks in the middle of our 28 countries and the U.K. ranks No. 4. Many in the U.S. have long argued against focusing exclusively on test scores in education. For example, the National Education Association states that in response to NCLB and RTT, "We see schools across America dropping physical education ... dropping music ... dropping their arts programs ... all in pursuit of higher test scores. This is not good education." Figure 4 provides quantifications for this argument, showing that the U.S. educational system has a comparative advantage for non-cognitive skills. As for the U.K., it ranks ahead of Hong Kong in both non-cognitive (Figure 4) and cognitive productivities (Figure 3), and it seems reasonable to assume that Hong Kong and Shanghai, China, have similar educational systems. If Elizabeth Truss had known about these rankings in 2014, would she have traveled to Shanghai to "learn a lesson in math"?

In summary, our estimates for cognitive and non-cognitive productivities provide better numerical metrics than test scores for the qualities of education. As another example, Figures 3 and 4 suggest that the educational systems of Finland, Netherlands and Belgium are far more worthy of emulation than those of South Korea and Hong Kong. Below we condense the multi-dimensional differences in cognitive and non-cognitive productivities into a single index for the overall educational quality, and quantify its contribution to output per worker.

6 Closed Economy Comparative Statics

We start by illustrating the overall educational quality, Ω^k, of equation (16), using Figure 5. This figure is a scatter plot of the values of cognitive productivities, h_c^k, against the values of non-cognitive productivities, h_n^k, for the countries in our sample. Since h_c^{US} and h_n^{US} are normalized to 1, by equation (16), the overall education quality of the U.S. is also 1, given that we use the U.S. as the base country (i.e. $\Omega^{US} = 1$). Inspired by isoquants, we plot the iso-education-quality curve for the U.S. in Figure 5; i.e. the combinations of cognitive and non-cognitive productivities that produce the same overall education quality as the U.S. This curve illustrates the trade-off between cognitive and non-cognitive productivities in maintaining the same level of overall education quality. It also illustrates the countries whose overall education qualities are similar to the U.S. (e.g. Sweden and Denmark), those with higher overall education qualities than the U.S. (e.g. the U.K. and Finland), and those with lower overall education qualities (e.g. Italy and S. Korea).

We then compute the numerical values of the decomposition (15) and (16) and report them in Table 8. Column (1) shows the countries' output per worker relative to the U.S. (i.e.$\frac{y^k/L^k}{y^0/L^0}$). Columns (2) shows the contribution of output TFP to differences in output per worker (i.e. $\left[\frac{\Theta^k}{\Theta^0}\right]^{\frac{1}{1-\eta}}$). Column (3) shows the contribution of overall education quality to output per worker (i.e. $\left[\Omega^k\right]^{\frac{1}{1-\eta}}$). Columns (2) and (3) are an *exact* decomposition of column (1), even though we have calculated column (1) using our data and columns (2) and (3) using our parameter values. This is because we follow the exact specifications of our model in parameter identification, (19), (20), (21) and (22).[22]

Table 8 shows how our sample countries compare with the U.S., in terms of overall educational quality, output TFP, and output per worker. Consider Germany. First, the overall quality of Germany's educational institution is lower than the U.S., the effect of which puts Germany's output per worker at 88.34% of the U.S. level (column (3)). On top of this, Germany also has lower output TFP than the U.S., the effect of which places its output per worker at 71.26% of the U.S. level (column (2)). Aggregating these two effects, Germany's output per worker is 62.96% (= 88.34% x 71.26%) of the U.S. level (column (1)).[23]

Table 8 also quantifies the large differences in overall educational qualities across countries that Figure 5 has illustrated. For example, although S. Korea's educational system delivers high test scores, it puts S. Korea's output per worker at 71.42% of the U.S. level, other things equal. Finland, on the other hand, has the strongest educational institution in our sample, which puts Finland's output per worker at 154.58% of the U.S. level, ceteris paribus. These results suggest that educational policies and reforms have very large potential payoffs, as well as danger, in terms of aggregate output.

We now calculate how changes in the qualities of the educational institution affect test scores and aggregate output, and how such calculations help inform the discussions of education policies and reforms. These comparative statics are very easy to implement using our model. Equations (15) and (16) provide closed form solutions for output per worker, and map changes in educational TFPs into changes in output per worker relative to any arbitrary base country which includes the initial equilibrium. Equations (19), (20) and (21), together with the identity $p_c^k + p_n^k = 1$, imply that the percentage change in test score is a linear function of the percentage changes in educational TFPs (see the Appendix for the proof)

$$(1-\eta)d\ln S^k = (1+Bp_c^k)d\ln h_c^k - (Bp_n^k)d\ln h_n^k, B = \frac{(\theta-1)(\alpha-1)-\alpha\eta}{\theta+\alpha-1} > 0, \quad (23)$$

[22] The decomposition is not exact for Australia and New Zealand because we dummy them out in implementing (21) and (22).

[23] Columns (2) and (3) in Table 8 are based on $\theta = 3.4965$ and $\alpha = 1.4706$. Table 7 shows that we obtain very similar values and country rankings for overall education quality under alternative values of α and θ.

where $B = 0.2496$ according to our parameter values.[24] Equation (23) says that an increase in test score, S^k, can be achieved by either an increase in cognitive productivity, h_c^k, and/or a reduction in non-cognitive productivity, h_n^k. The latter works because an educational institution with a very low level of non-cognitive productivity simply denies most people the option of accumulating non-cognitive human capital through education. This creates very strong incentives to accumulate cognitive human capital, showing up as an increase in test score. As a result, a rise in test score may result from a better educational institution along the cognitive dimension, or a worse one along the non-cognitive dimension. While the former is a blessing, the latter is a curse in disguise, as we illustrate below.

Suppose the U.S. can implement some policy reform to boost its PISA score by 2.58%, in order to advance 5 places in PISA math rankings. This puts U.S. PISA math score at U.K's level. To illustrate the intended consequence of this policy, assume that U.S. non-cognitive productivity, h_n^{US}, remains unchanged. Equation (23) tells us that U.S. cognitive productivity rises by 2.12%, and equation (15) tells us that U.S. aggregate output rises by 1.81%. The increase in output provides an upper bound estimate for the amount of resources to be spent on the reform, or an estimate for the potential returns of the reform if we know the amount of resources spent. This exercise illustrates that our model is a useful tool for the cost-benefit analysis of education policies.

Our model is also useful for clarifying the objective of education policies. In the U.S., both No Child Left Behind of 2001 and Race To the Top of 2009 are motivated by the concern for low test scores, and both measure student performance using test scores. Our model shows that test score and output may move in the opposite direction, because there are multiple types of human capital and heterogeneous individuals respond to policy changes by changing their choices for education. Suppose that the U.S. implements less ambitious education reforms than in scenario 1 above, and succeeds in raising U.S. PISA score by 0.258%. To illustrate the unintended consequence of this policy, assume that U.S. cognitive productivity, h_c^{US}, remains unchanged. Then by equation (23), U.S. non-cognitive productivity, h_n^{US}, decreases by 4.08%, and by equation (15), U.S. aggregate output decreases by 1.02%. This exercise illustrates that an increase in test score could mask a reduction in the overall quality of the educational institution. As a result, aggregate output is a better objective for education policies than test score.

[24] This is based on $\theta = 3.4965$. If $\theta = 2.0877$, $B = 0.1279$.

Indeed, many educational reforms that are promoted to raise test scores have been criticised because of the fear that improvement along one dimension may come at the expense of decline along another. Our model quantifies the pros and cons of education policy reforms. For instance, many in the U.S. advocate emulating the heavily test-based educational systems and practices of east Asian countries, such as those in S. Korea and Hong Kong. While this may increase cognitive learning, it can also induce poor performance in non-cognitive human capital. Our calculations in section 4 show that Hong Kong's cognitive productivity is 1.13 times the U.S. level, but her non-cognitive productivity is 0.40 times the U.S. level. Equation (23) then says that should the U.S. get Hong Kong's educational system, test score would increase by 22.50%, putting the U.S. as the world champion in PISA scores. However, equations (15) and (16) tell us that despite this accomplishment in test scores, U.S. aggregate output would decrease by 11.13%!

7 Open Economy Extension

In the closed economy analysis, there were no interactions whatsover between countries. This assumption matters for assessing the welfare effects associated with uneven schooling quality, because comparative advantage allows countries to import those factor services in which they are not well endowed. In this section, we make the polar opposite assumption that intermediate inputs that are created by each of the two occupations are freely traded. This case is relatively easy to analyze because it allows us to impose effective factor price equalization while still allowing the real incomes of factors to vary across countries according to output TFP variation across countires as well as to differences in cognitive and non-cognitive productivities.

7.1 Model and Equilibrium

The production function continues to be given by equation (4), but an important distinction in the open economy is that the stocks of cognitive and non-cognitive labor used in final good production are no longer restricted to those supplied locally. This is because we now assume that L_c and L_n are freely traded on the world market (HOV-like assumption), but we continue to assume that the final good is non-traded. Otherwise, all of the assumptions made earlier continue to hold.

Because the services of human capital are freely traded, there must be a single global price per unit of cognitive (w_c) and non-cognitive (w_n) human capital. Given effective factor price equalization and common factor intensities for final good production, it immediately follows that labor demand continues to be given by (12) in all countries.

The fact that final goods are untraded and that the final good production technology varies across countries due to the Hick's neutral productivity shifters Θ^k means that the level of real wages does not equalize across countries and that the relative price of education varies across countries. The exact price index of final good production given technology (4) is given by

$$P^k = \frac{1}{\Theta^k}\left((A_c)^{\alpha}(w_c)^{1-\alpha} + (A_n)^{\alpha}(w_n)^{1-\alpha}\right)^{\frac{1}{1-\alpha}} \tag{24}$$

Defining the numeraire to be $\left((A_c)^{\alpha}(w_c)^{1-\alpha} + (A_n)^{\alpha}(w_n)^{1-\alpha}\right)^{\frac{1}{1-\alpha}} = 1$, the price of final output in country k will be given by $P^k = \left(\Theta^k\right)^{-1}$.

Now recall that the educational investment is in terms of final output so that the proper program facing a student that will choose occupation i is

$$\max_{e}\left\{w_i h_i^k e^{\eta}\epsilon_i - P^k e\right\}$$

and so the optimal choice of education, after substituting for the price index and accounting for the normalization, is then

$$e(\epsilon_i) = \left(\eta w_i \Theta^k h_i^k \epsilon_i\right)^{\frac{1}{1-\eta}}. \tag{25}$$

Here we see that education (and so average wages) obtained varies across countries for two reasons despite effective factor equalization. First, high TFP increases the real return of education relative to its cost. Second, higher educational TFP has an isomorphic effect. Turning to educational choices, the proportion of the population that chooses to become trained in cognitive skills continues to be given by (7) as comparative and not absolute advantage drives occupation choices.

To calculate aggregate educational expenditure, we sum $e(\epsilon_i)$ for all workers that select into each type.

$$E_c^k = \left(\eta\Theta^k\right)^{\frac{1}{1-\eta}}\left(T_c\left(w_c h_c^k\right)^{\theta} + T_n\left(w_n h_n^k\right)^{\theta}\right)^{\frac{1}{\theta(1-\eta)}}\gamma. \tag{26}$$

Comparing this expression with that of the closed economy, equation (9), we see that educational expenditure per capita varies across countries due to differences in productivity. Given the constant elasticity η, however, it continues to be the case that educational spending per capita is proportional to output per worker.

The variation in educational expenditure per capita has the effective of shifting absolute (but not relative) labor supplies across countries

$$L_c^k = L^k p_c^k E(h_c^k e^\eta | cognitive) = \frac{L^k p_c^k}{w_c} \left((\eta \Theta^k)^\eta \left(T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta \right)^{1/\theta} \right)^{1/(1-\eta)} \gamma. \tag{27}$$

Relative labor supply continues to be given by (11).

We show in the appendix that the open economy equivalent GDP per capita decomposition is given by

$$\frac{y^k/L^k}{y^0/L^0} = \left(\frac{\Theta^k}{\Theta^0} \left(p_c^0 \left(\frac{h_c^k}{h_c^0} \right)^\theta + p_n^0 \left(\frac{h_n^k}{h_n^0} \right)^\theta \right)^{\frac{1}{\theta}} \right)^{\frac{1}{1-\eta}}. \tag{28}$$

Comparing the open economy decomposition, with that of the closed economy given by equation (15), we see that the key difference is in the power coefficients in the construction of the power mean of educational obtainment. Critically, the power coefficients in the open economy do not include α as local labor market demand does not have to equal local labor market supply. This has the effect of increasing the size of these power coefficients. Under the condition that $\alpha > 1$, it is as if $\alpha \to \infty$ in the closed economy case and so being relatively inefficient at providing one type of education is less of a drag on the economy. Intuitively, in a world of free trade the ability to buy rather than make the comparative disadvantage good is the source of welfare gains.

7.2 Identification

In the open economy setting, the final good is nontraded and in terms of local prices. By equation (2), the production of human capital uses the final good, and so educational spending is also in local prices. The loca-price-index, P^k, affects output and educational spending in the same way, so that equation (20) continues to hold. This means that our identification of η remains unchanged. On the other hand, equations (26) and (27) say that, although P^k affects both educational spending, E^k, and the aggregate supply of cognitive human capital, L_c^k, it does not affect the ratio $L_c^k/(E^k)^\eta$. This implies, together with assumption (17), that equations (18) and (21) continue to hold. As a result, our identification of θ and h_c^k also remains unchanged. Summarizing these results we have

Proposition 4 *The identifications of η, θ and h_c^k, or equations (20) and (21), continue to hold in the open-economy setting.*
Proof. *See the Appendix.* ■

We now turn to the comparative advantage of the educatoinal institution, and show, in the Appendix, that the open-economy equivalent of equation (19) is

$$\frac{p_c^k/p_n^k}{p_c^0/p_n^0} = \left(\frac{h_c^k/h_n^k}{h_c^0/h_n^0}\right)^\theta \tag{29}$$

Intuitively, (29) differs from (19) for the same reason that (28) differs from (15) and (16). Relative to a closed economy, individuals sell the services of their human capital in the global labor market in an open economy, and so place more emphasis on the comparative advantage of the educational institution, h_c^k/h_n^k, in their occupational decisions. It then follows that, conversely, the data of employment-share ratio, p_c^k/p_n^k, implies smaller differences in h_c^k/h_n^k in the open-economy setting than in the closed-economy setting.

Proposition 4 and equation (29) allow us to identify non-cognitive productivity, h_n^k. We then use the decomposition (28) to calculate the overall educational quality, and to identify output TFP, Θ^k.

Columns (4) and (5) of Table 8 report the results of the decomposition (28). The variation of overall education quality across countries is reduced as compared with the closed-economy setting, consistent with the intuition of equations (28) and (29); i.e. the free global flows of ideas and talents in the open-economy setting allow countries to take advantage of the differences in the relative strength of their educational institutions. The third row of Table 7 reports the correlation coefficients between the open-economy values and country rankings with the closed-economy ones, in terms of non-cognitive productivity and overall education quality. They range from 0.7757 to 0.9160, suggesting that overall, our estimates for the productivities and qualities of educational institutions are broadly similar under closed- and open-economy settings. The last row of Table 7 shows that our closed- and open-economy estimates remain similar under alternative θ values.[25]

7.3 Patterns of Trade

Our open-economy model has the HO-like prediction that the countries with relative abundance in non-cognitive human capital are net exporters of its service. To take this prediction to the data, we follow the literature (e.g. Nunn 2007) and examine the correlation between the patterns of trade and the interactions between relative factor abundance and factor-use intensities. For each country in our sample, we collect aggregate import and export for the 31 NAICS manufacturing industries in the 2000 U.S. census, and the 9 1-digit service industries in the UN service-trade database. We measure trade patterns by revealed comparative advantage, or net export divided by the sum of import and export. For each country, we measure its relative abundance in non-cognitive human capital, physical capital and skilled labor as, respectively, the non-cognitive employment share, the ratio of physical capital stock to population, and the fraction of college-educated labor force. For each industry, we measure the intensities of non-cognitive human capital, physical capital and skilled labor using U.S. data. [26] Finally, we control for industry fixed effects and country fixed effects.

Table 9 reports the results. Column (1) includes only the interaction for non-cognitive human capital. We add the interaction for physical capital in column (2), and then the interaction for skilled labor in column (3). The interaction for non-cognitive human capital has positive and significant coefficient estimates in all specifications, consistent with the prediction of our open-economy model. These results provide another important validation of our model, because we did not use industry-level import and export data for parameter identification.

7.4 Comparative Statics

Comparative statics in the open economy case are somewhat more involved than in closed economy case because demand for cognitive and non-cognitive labor must be aggregated

[25] α does not affect open-economy parameter values, since the open-economy setting is essentially $\alpha \longrightarrow \infty$.

[26] See our Appendix for more details.

over countries. We begin this section by showing how the work of Deckle, Eaton, and Kortum (2008) can be used to solve for changes in global prices without information on technological and endowment parameters.

Defining changes to variable x as $\widehat{x} = x'/x$, we can use the procedures outlined in Deckle, Eaton, and Kortum (2008) to solve for changes in international relative prices due to a shock that enters anywhere in the world. We show in the appendix that the labor market clearing condition can be written:

$$\begin{aligned} & (\widehat{w}_c)^{\theta+\alpha-1} \sum_k \frac{p_c^k P_Y^k y^k}{\sum_j p_c^j P_Y^j y^j} \widehat{L}^k (\widehat{\Theta}^k)^{\frac{\eta}{1-\eta}} \left(\widehat{h}_c^k\right)^{\theta} \left(p_c^k \left(\widehat{w}_c \widehat{h}_c^k\right)^{\theta} + p_n^k \left(\widehat{w}_n \widehat{h}_n^k\right)^{\theta}\right)^{\frac{1}{\theta(1-\eta)}-1} \quad (30) \\ = & \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} \widehat{L}^k (\widehat{\Theta}^k)^{\frac{\eta}{1-\eta}} \left(p_c^k \left(\widehat{w}_c \widehat{h}_c^k\right)^{\theta} + p_n^k \left(\widehat{w}_n \widehat{h}_n^k\right)^{\theta}\right)^{\frac{1}{\theta(1-\eta)}} . \end{aligned}$$

This condition can then be combined with our normalization, written in changes as

$$(\widehat{w}_c)^{1-\alpha} \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} p_c^k + (\widehat{w}_n)^{1-\alpha} \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} p_n^k = 1$$

to pin down the price effects of any shock to deep model parameters. With these changes in hand, all of the reallocations can be solved as well as shifts in test scores, educational expenditures, and welfare per worker.

8 Conclusion

TBW

References

[1] Acemoglu, Daron, Simon Johnson, and James A. Robinson. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91.5 (2001): 1369-1401.

[2] Autor, David H., Frank Levy and Richard J. Murname, 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration”, *Quarterly Journal of Economics* 118(4).

[3] Behrman, J. R., Parker, S. W., Todd, P. E., & Wolpin, K. I. (2015). Aligning learning incentives of students and teachers: results from a social experiment in Mexican high schools. *Journal of Political Economy*, 123(2), 325-364.

[4] Burnstein, Ariel, Eduardo Morales and Jonathan Vogel, 2016. “Changes in Between-Group Inequality: Computers, Occupations and International Trade”, mimeo.

[5] Choi, Yochul, David Hummels, and Chong Xiang. 2009. “Explaining Import Quality: the Role of the Income Distribution”. J*ournal of International Economics*, 78, 293-303.

[6] Cunha, Flavio, James Heckman and Susanne Schennach. “Estimating the Technology of Cognitive and Non-cognitive Skill Formation”, *Econometrica* 78(3), May 2010, 883-931.

[7] Deckle, Jonathan Eaton, and Samuel Kortum.

[8] Eaton, Jonathan and Samuel Kortum, 2002. “Technology, Geography and Trade”, *Econometrica*, 70(5), 1741-1779.

[9] Figlio, David and Susanna Loeb, 2011. “School Accountability”, in *Handbook of the Economics of Education*, Volume 3, Edited by Eric Hanushek, Stephen Machin and Ludger Woessmann, Elsevier North-Holland: Amsterdam, 383-421.

[10] Hall, Robert E. and Charles I. Jones, "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, February 1999, Vol. 114, pp. 83-116.

[11] Heckman, James J., and Tim Kautz. "Hard evidence on soft skills." *Labour economics* 19.4 (2012): 451-464.

[12] Heckman, James J. and Yona Rubinstein, “The Importance of Noncognitive Skills: lessons from the GED Testing Program”, *American Economic Review* Papers & Proceedings 91(2), May 2001, pp. 145-149.

[13] Hsieh, Chang-Tai, Erik Hurst, Charles Jones and Peter Klenow. 2013. “The Allocation of Talent and U.S. Economic Growth”, NBER working paper 18693.

[14] Hummels, David, Rasmus Jorgensen, Jakob Munch, and Chong Xiang, 2011, "The wage and employment effects of outsourcing: evidence from Danish matched worker-firm data", NBER working paper 17496.

[15] Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang. "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data", *American Economic Review*, 104 (6), June 2014, 1597-1629.

[16] Liu, Runjuan and Daniel Trefler, 2011. A Sorted Tale of Globalization: White Collar Jobs and the Rise of Service Offshoring, NBER working paper 17559.

[17] Neal, Derek A., and William R. Johnson, "The Role of Premarket Factors in Black –White Wage Differences", *Journal of Political Economy* 104 (5), October 1966, 869-895.

[18] Nunn, Nathan. "Relationship-specificity, incomplete contracts, and the pattern of trade." *The Quarterly Journal of Economics* (2007): 569-600.

[19] Pierce, Justin and Peter Schott, 2009, "A Concordance Between Ten-Digit U.S. Harmonized System Codes and SIC/NAICS Product Classes and Industries". NBER working paper 15548.

[20] Willis, Robert and Sherwin Rosen. "Education and Self-Selection", J*ournal of Political Economy* 87(5), October 1979, S7-S36.

9 Theory Appendix

9.1 Proposition 1

To simplify notation, we drop the superscript k. In addition, let $\omega_c = w_c h_c$, $\omega_n = w_n h_n$, $F_c = \frac{\partial F(.)}{\partial \varepsilon_c}$, and $F_{nc} = \frac{\partial^2 F(.)}{\partial \varepsilon_n \partial \varepsilon_c}$. Using the definition of p_n, we have

$$\begin{aligned} p_n &= \Pr(\omega_n \varepsilon_n \geq \omega_c \varepsilon_c) = \int_0^\infty \int_{\frac{\omega_c}{\omega_n}\varepsilon_c}^\infty F_{nc} d\varepsilon_n d\varepsilon_c \\ &= \int_0^\infty [F_c(\varepsilon_c, \varepsilon_n \to \infty) - F_c(\varepsilon_c, \varepsilon_n = \frac{\omega_c}{\omega_n}\varepsilon_c)] d\varepsilon_c \\ &= \int_0^\infty F_c(\varepsilon_c, \varepsilon_n \to \infty) d\varepsilon_c - \int_0^\infty F_c(\varepsilon_c, \varepsilon_n = \frac{\omega_c}{\omega_n}\varepsilon_c) d\varepsilon_c \end{aligned}$$

Using the Frechet distribution (1), we have

$$F_c(\varepsilon_c, \varepsilon_n) = AFT_c\varepsilon_c^{-\theta-1}, A = (1-\rho)\theta(T_n\varepsilon_n^{-\theta} + T_c\varepsilon_c^{-\theta})^{-\rho}$$

(1) When $\varepsilon_n \to \infty$, $A = (1-\rho)\theta(T_c\varepsilon_c^{-\theta})^{-\rho}$ and $F = \exp[-(T_c\varepsilon_c^{-\theta})^{1-\rho}]$. Therefore,

$$\begin{aligned} F_c(\varepsilon_c, \varepsilon_n &\to \infty) = (1-\rho)\theta(T_c\varepsilon_c^{-\theta})^{-\rho}\exp[-(T_c\varepsilon_c^{-\theta})^{1-\rho}][T_c\varepsilon_c^{-\theta-1}] \\ &= \theta(1-\rho)(T_c)^{1-\rho}\varepsilon_c^{-\theta(1-\rho)-1}\exp[-(T_c)^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}] \end{aligned}$$

and

$$\begin{aligned} \int_0^\infty F_c(\varepsilon_c, \varepsilon_n &\to \infty)d\varepsilon_c = \int_0^\infty \theta(1-\rho)(T_c)^{1-\rho}\varepsilon_c^{-\theta(1-\rho)-1}\exp[-(T_c)^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}]d\varepsilon_c \\ &= \int_0^\infty \frac{d\exp[-(T_c)^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}]}{d\varepsilon_c} = \left(\exp[-(T_c)^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}]\right)\Big|_0^\infty = 1 \end{aligned}$$

(2) When $\varepsilon_n = \frac{\omega_c}{\omega_n}\varepsilon_c$,

$$A = (1-\rho)\theta[T_n\varepsilon_c^{-\theta}(\frac{\omega_c}{\omega_n})^{-\theta} + T_c\varepsilon_c^{-\theta})^{-\rho} = (1-\rho)\theta(\varepsilon_c^{-\theta})^{-\rho}B^{-\rho}, B = T_n(\frac{\omega_c}{\omega_n})^{-\theta} + T_c$$

and,

$$F(\varepsilon_c, \varepsilon_n = \frac{\omega_c}{\omega_n}\varepsilon_c) = \exp\{-[T_n\varepsilon_c^{-\theta}(\frac{\omega_c}{\omega_n})^{-\theta} + T_c\varepsilon_c^{-\theta}]^{1-\rho}\} = \exp[-B^{1-\rho}(\varepsilon_c^{-\theta})^{1-\rho}]$$

Therefore,

$$\begin{aligned} F_c(\varepsilon_c, \varepsilon_n &= \frac{\omega_c}{\omega_n}\varepsilon_c) = (1-\rho)\theta(\varepsilon_c^{-\theta})^{-\rho}B^{-\rho}\exp[-B^{1-\rho}(\varepsilon_c^{-\theta})^{1-\rho}][T_c\varepsilon_c^{-\theta-1}] \\ &= (1-\rho)\theta T_c\varepsilon_c^{-\theta(1-\rho)-1}B^{-\rho}\exp[-B^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}] \end{aligned}$$

and

$$\begin{aligned} \int_0^\infty F_c(\varepsilon_c, \varepsilon_n &= \frac{\omega_c}{\omega_n}\varepsilon_c)d\varepsilon_c = \int_0^\infty (1-\rho)\theta T_c\varepsilon_c^{-\theta(1-\rho)-1}B^{-\rho}\exp[-B^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}]d\varepsilon_c \\ &= T_cB^{-1}\int_0^\infty \frac{d\exp[-B^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}]}{d\varepsilon_c} \\ &= \left(T_cB^{-1}\exp[-B^{1-\rho}\varepsilon_c^{-\theta(1-\rho)}]\right)\Big|_0^\infty = T_cB^{-1} \end{aligned}$$

(3) Using (1) and (2) above we have

$$p_n = 1 - T_c B^{-1} = \frac{T_n(\omega_c)^{-\theta}(\omega_n)^\theta}{T_c + T_n(\omega_c)^{-\theta}(\omega_n)^\theta} = \frac{T_n(\omega_n)^\theta}{T_c(\omega_c)^\theta + T_n(\omega_n)^\theta}$$

This is equation (7).

9.2 Proposition 2

To simplify notation, we drop the superscript k. We note that the Frechet distribution is max stable; i.e. the max of Frechet variables is still Frechet. To be specific, consider the random variable $\varepsilon^* = \max\{w_c h_c \varepsilon_c, w_n h_n \varepsilon_n\}$. By our discussions in section 3, $\varepsilon^* = w_n h_n \varepsilon_n$ if and only if the individual chooses occupation n.

We now obtain the cdf of the distribution of ε^*

$$\begin{aligned}
\Pr(\varepsilon^* &\leq y) = \Pr(w_c h_c \varepsilon_c \leq y \text{ and } w_n h_n \varepsilon_n \leq y) \\
&= F(\frac{y}{w_c h_c}, \frac{y}{w_n h_n}) \\
&= \exp[-B_1 y^{-\theta(1-\rho)}], B_1 = [T_c(w_c h_c)^\theta + T_n(w_n h_n)^\theta]^{1-\rho}
\end{aligned}$$

where we have used the Frechet distribution (1) in the second equality.

Consider the mean of non-cognitive workers' net income, I_n, conditional on choosing the non-cognitive occupation, n. By the expression of I_n, (6), we know that it is proportional to the mean of $(w_n h_n \varepsilon_n)^{\frac{1}{1-\eta}}$, conditional on choosing occupation n. This conditional mean is, by Bayesian rule, the mean of $(w_n h_n \varepsilon_n)^{\frac{1}{1-\eta}}$ for those choosing occupation n, divided by the employment share p_n. The mean of $(w_n h_n \varepsilon_n)^{\frac{1}{1-\eta}}$ for those choosing occupation n, in turn, is the mean of $(\varepsilon^*)^{\frac{1}{1-\eta}}$ for all workers times the employment share p_n. As a result, the conditional mean of I_n is proportional to the mean of $(\varepsilon^*)^{\frac{1}{1-\eta}}$, which equals

$$\int_0^\infty y^{\frac{1}{1-\eta}} \frac{d\exp[-B_1 y^{-\theta(1-\rho)}]}{dy} = \int_0^\infty y^{\frac{1}{1-\eta}} \exp[-B_1 y^{-\theta(1-\rho)}] B_1 \theta(1-\rho) y^{-\theta(1-\rho)-1} dy$$

We then use change-of-variables to calculate the value of this expression, because the Gamma function is defined as

$$\Gamma(a+1) = \int_0^\infty t^a e^{-t} dt,$$

where a is a constant. Let $x = B_1 y^{-\theta(1-\rho)}$. Then $y = (\frac{x}{B_1})^{-\frac{1}{\theta(1-\rho)}}$, and $dy = -\frac{1}{\theta(1-\rho)} B_1^{\frac{1}{\theta(1-\rho)}} x^{-\frac{1}{\theta(1-\rho)}-1} dx$. In addition, as $y \to 0$, $x \to \infty$; as $y \to \infty$, $x \to 0$. Therefore,

$$\begin{aligned}
& \int_0^\infty y^{\frac{1}{1-\eta}} \frac{d \exp[-B_1 y^{-\theta(1-\rho)}]}{dy} \\
= & \int_0^\infty y^{\frac{1}{1-\eta}} \exp[-B_1 y^{-\theta(1-\rho)}] B_1 \theta(1-\rho) y^{-\theta(1-\rho)-1} dy \\
= & \int_\infty^0 (\frac{x}{B_1})^{-\frac{1}{\theta(1-\rho)(1-\eta)}} e^{-x} B_1 \theta(1-\rho) (\frac{x}{B_1})^{\frac{1+\theta(1-\rho)}{\theta(1-\rho)}} [-\frac{1}{\theta(1-\rho)}] B_1^{\frac{1}{\theta(1-\rho)}} x^{-\frac{1}{\theta(1-\rho)}-1} dx \\
= & \int_0^\infty (\frac{x}{B_1})^{-\frac{1}{\theta(1-\rho)(1-\eta)} + \frac{1}{\theta(1-\rho)} + 1 - \frac{1}{\theta(1-\rho)} - 1} e^{-x} dx \\
= & B_1^{\frac{1}{\theta(1-\rho))(1-\eta)}} \int_0^\infty x^{-\frac{1}{\theta(1-\rho)(1-\eta)}} e^{-x} dx = B_1^{\frac{1}{\theta(1-\rho))(1-\eta)}} \Gamma(1 - \frac{1}{\theta(1-\rho)(1-\eta)}) \\
= & \gamma [T_c (w_c h_c)^\theta + T_n (w_n h_n)^\theta]^{\frac{1}{\theta(1-\eta)}}, \gamma = \Gamma(1 - \frac{1}{\theta(1-\rho)(1-\eta)})
\end{aligned}$$

Therefore, the average net income of non-cognitive workers, I_n, equals $(1-\eta)\eta^{\frac{\eta}{1-\eta}} \gamma [T_c (w_c h_c)^\theta + T_n (w_n h_n)^\theta]^{\frac{1}{\theta(1-\eta)}}$.This is equation (8).

9.3 Proposition 3

We again drop the superscript k. We start with the expression $h_i e^\eta \varepsilon_i = (\eta w_i)^{\frac{\eta}{1-\eta}} (h_i \varepsilon_i)^{\frac{1}{1-\eta}}$, which we show in the text, right above Proposition 3. Using this expression and the equation of net income, (6), we get

$$h_i e^\eta \varepsilon_i = (\eta w_i)^{\frac{\eta}{1-\eta}} \frac{I_i}{(1-\eta)\eta^{\frac{\eta}{1-\eta}} (w_i)^{\frac{1}{1-\eta}}} = \frac{I_i}{(1-\eta) w_i}$$

This means that

$$\begin{aligned}
L_i &= Lp_i E\left(h_i e^\eta \varepsilon_i | \text{Occupation } i\right) = Lp_i \frac{1}{(1-\eta) w_i} E(I_i | \text{Occupation } i) \qquad (31) \\
&= \frac{Lp_i}{(1-\eta) w_i} (1-\eta) \eta^{\frac{\eta}{1-\eta}} \gamma [T_c (w_c h_c)^\theta + T_n (w_n h_n)^\theta]^{\frac{1}{\theta(1-\eta)}}
\end{aligned}$$

where the last equality is by Proposition 2. To simplify this expression, we use Proposition 1 to get

$$T_c (w_c h_c)^\theta + T_n (w_n h_n)^\theta = \frac{T_i (w_i h_i)^\theta}{p_i}$$

This allows us to obtain

$$\begin{aligned}
L_i &= \frac{Lp_i}{(1-\eta)w_i}(1-\eta)\eta^{\frac{\eta}{1-\eta}}\gamma[T_c(w_ch_c)^\theta + T_n(w_nh_n)^\theta]^{\frac{1}{\theta(1-\eta)}} \\
&= \frac{Lp_i}{w_i}\eta^{\frac{\eta}{1-\eta}}\gamma[\frac{T_i(w_ih_i)^\theta}{p_i}]^{\frac{1}{\theta(1-\eta)}} = \eta^{\frac{\eta}{1-\eta}}\gamma Lp_i(h_i)^{\frac{1}{1-\eta}}(w_i)^{\frac{\eta}{1-\eta}}(\frac{T_i}{p_i})^{\frac{1}{\theta(1-\eta)}} \\
&= \gamma Lp_i\left[h_i(\eta w_i)^\eta\left(\frac{T_i}{p_i}\right)^{\frac{1}{\theta}}\right]^{\frac{1}{1-\eta}}
\end{aligned}$$

9.4 Closed Economy Productivity Decomposition,(15) and (16)

Starting with the final good production function, we have

$$\begin{aligned}
y^k &= \Theta^k\left(A_c\left(L_c^k\right)^{\frac{\alpha-1}{\alpha}} + A_n\left(L_n^k\right)^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}} \\
&= \Theta^k L_c^k\left(A_c + A_n\left(\frac{L_n^k}{L_c^k}\right)^{\frac{\alpha-1}{\alpha}}\right)^{\frac{\alpha}{\alpha-1}}
\end{aligned}$$

The first-order condition for optimal input choice requires

$$\frac{L_c^k}{L_n^k} = \left(\frac{p_c^k}{p_n^k}\frac{A_n}{A_c}\right)^{\frac{\alpha}{\alpha-1}}.$$

Substituting this expression into the output equation yields

$$y^k = \Theta^k L_c^k\left(\frac{A_c}{p_c^k}\right)^{\frac{\alpha}{\alpha-1}}$$

Educational and occupational choice requires that $w_c^kL_c^k = p_c^ky^k$. Substituting this expression into the output equation, we obtain

$$w_c^k = \Theta^k\left(p_c^k\right)^{-\frac{1}{\alpha-1}}(A_c)^{\frac{\alpha}{\alpha-1}} \tag{32}$$

Rearranging the educational expenditure equation,

$$E_c^k = \left(\eta w_c^kh_c^k\right)^{\frac{1}{1-\eta}}\left(\frac{T_c}{p_c^k}\right)^{\frac{1}{\theta(1-\eta)}}\gamma,$$

and substituting $E^k = \eta y^k/L^k$, we can substitute w_c^k in equation (32) to obtain after rearranging

$$\frac{y^k}{L^k} = \left(\Theta^kh_c^k\left(p_c^k\right)^{-\frac{\phi}{\theta(\alpha-1)}}(A_c)^{\frac{\alpha}{\alpha-1}}(T_c)^{\frac{1}{\theta}}\eta\right)^{\frac{1}{1-\eta}}\frac{\gamma}{\eta},$$

where we have defined $\phi \equiv \alpha + \theta - 1$. Substituting out p_c^k using its definition, we obtain

$$\frac{y^k}{L^k} = \left(\Theta^k h_c^k \left(1 + \frac{T_n \left(h_n^k\right)^\theta}{T_c \left(h_c^k\right)^\theta} \left(\frac{w_n^k}{w_c^k}\right)^\theta \right)^{\frac{\phi}{\theta(\alpha-1)}} (A_c)^{\frac{\alpha}{\alpha-1}} (T_c)^{\frac{1}{\theta}} \eta \right)^{\frac{1}{1-\eta}} \frac{\gamma}{\eta}.$$

Finally, factor market clearing implies

$$\frac{w_c^k}{w_n^k} = \left[\frac{T_n}{T_c} \left(\frac{A_c}{A_n}\right)^\alpha \left(\frac{h_n^k}{h_c^k}\right)^\theta \right]^{\frac{1}{\phi}}.$$

Substituting this expression into the GDP per capita equation and simplifying, we obtain an expression with no endogenous variables

$$\frac{y^k}{L^k} = \left(\Theta^k h_c^k \left(1 + \left(\frac{T_n \left(h_n^k\right)^\theta}{T_c \left(h_c^k\right)^\theta}\right)^{\frac{\alpha-1}{\phi}} \left(\frac{A_n}{A_c}\right)^{\alpha \frac{\theta}{\phi}} \right)^{\frac{\phi}{\theta(\alpha-1)}} (A_c)^{\frac{\alpha}{\alpha-1}} (T_c)^{\frac{1}{\theta}} \eta \right)^{\frac{1}{1-\eta}} \frac{\gamma}{\eta}.$$

Comparing GDP per capita in country k to a base country (or to the initial values for that country in a comparative static, we have

$$\frac{y^k/L^k}{y^0/L^0} = \left(\frac{\Theta^k}{\Theta^0} \frac{h_c^k \left(1 + \left(\frac{T_n \left(h_n^k\right)^\theta}{T_c \left(h_c^k\right)^\theta}\right)^{\frac{\alpha-1}{\phi}} \left(\frac{A_n}{A_c}\right)^{\alpha \frac{\theta}{\phi}} \right)^{\frac{\phi}{\theta(\alpha-1)}}}{h_c^0 \left(1 + \left(\frac{T_n \left(h_n^0\right)^\theta}{T_c \left(h_c^0\right)^\theta}\right)^{\frac{\alpha-1}{\phi}} \left(\frac{A_n}{A_c}\right)^{\alpha \frac{\theta}{\phi}} \right)^{\frac{\phi}{\theta(\alpha-1)}}} \right)^{\frac{1}{1-\eta}}.$$

Combining the occupational share equations and labor market clearing conditions for the base country, we have

$$\left(\frac{A_n}{A_c}\right)^{\frac{\theta\alpha}{\phi}} \left(\frac{T_n}{T_c}\right)^{\frac{\alpha-1}{\phi}} = \left(\frac{\left(h_c^0\right)^\theta}{\left(h_n^0\right)^\theta}\right)^{\frac{\alpha-1}{\phi}} \frac{p_n^0}{p_c^0}.$$

Substituting this expression into the relative GDP per capita expressions and simplifying, we arrive at our decomposition:

$$\frac{y^k/L^k}{y^0/L^0} = \left(\frac{\Theta^k}{\Theta^0} \left(p_c^0 \left(\frac{h_c^k}{h_c^0}\right)^{\frac{\theta(\alpha-1)}{\phi}} + p_n^0 \left(\frac{h_n^k}{h_n^0}\right)^{\frac{\theta(\alpha-1)}{\phi}} \right)^{\frac{\phi}{\theta(\alpha-1)}} \right)^{\frac{1}{1-\eta}}.$$

9.5 Derivations of Various Equations

9.5.1 Equation (18)

Using equations (17) and (10), we can show that

$$S^k = b\frac{L_c^k}{L^k} = b(p_c^k)^{1-\frac{1}{\theta(1-\eta)}}\gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(w_c^k)^{\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}} \tag{33}$$

Using equations (17), (9) and (31), we can show that

$$\frac{S^k}{E^k} = \frac{bp_c^k}{\eta w_c^k} \Leftrightarrow w_c^k = \frac{bp_c^k E^k}{\eta S^k}$$

Plugging this expression into equation (33), we have

$$\begin{aligned}
S^k &= b(p_c^k)^{1-\frac{1}{\theta(1-\eta)}}\gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(\frac{bp_c^k E^k}{\eta S^k})^{\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}} \\
&= \gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(p_c^k)^{1-\frac{1}{\theta(1-\eta)}+\frac{\eta}{1-\eta}}(\frac{E^k}{\eta})^{\frac{\eta}{1-\eta}}(S^k)^{-\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}} \\
\Leftrightarrow &\ (S^k)^{\frac{1}{1-\eta}} = \gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(p_c^k)^{\frac{1}{1-\eta}-\frac{1}{\theta(1-\eta)}}(\frac{E^k}{\eta})^{\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}} \\
\Leftrightarrow &\ S^k = \gamma^{1-\eta}\eta^{\eta}(T_c)^{\frac{1}{\theta}}(p_c^k)^{1-\frac{1}{\theta}}(\frac{E^k}{\eta})^{\eta}h_c^k
\end{aligned}$$

This expression then implies equation (18).

9.5.2 Equation (19)

By Proposition 1 we have

$$\frac{p_c^k}{p_n^k} = \frac{T_c(w_c^k h_c^k)^\theta}{T_n(w_n^k h_n^k)^\theta}$$

Substitute out the ratio w_c^k/w_n^k using equation (13), and we get equation (19).

9.5.3 Equation (22)

Substitute out the term L_c^k in the aggregate production function (4) using equation (17), and substitute out L_n^k in (4) using equations (14) and (17), we have

$$
\begin{aligned}
y^k &= \Theta^k\{A_c(bL^kS^k)^{\frac{\alpha-1}{\alpha}} + A_n[bL^kS^k(\frac{p_n^k}{p_c^k}\frac{A_c}{A_n})^{\frac{\alpha}{\alpha-1}}]^{\frac{\alpha-1}{\alpha}}\}^{\frac{\alpha}{\alpha-1}} \\
&= \Theta^k bL^kS^k[A_c + A_n\frac{p_n^k}{p_c^k}\frac{A_c}{A_n}]^{\frac{\alpha}{\alpha-1}} \\
&= \Theta^k bL^kS^k(A_c)^{\frac{\alpha}{\alpha-1}}(1+\frac{p_n^k}{p_c^k})^{\frac{\alpha}{\alpha-1}}
\end{aligned}
$$

The log of this expression is equation (22).

9.5.4 Equation (23)

The comparative static exercise involves changing h_c^k and h_n^k, holding the other parameters fixed, and tracing out the responses of the endogenous variables. First, the identity $p_n^k + p_c^k = 1$ implies that

$$d\ln p_n^k = -(d\ln p_c^k)\frac{p_c^k}{p_n^k}$$

Next, equations (19), (21) and (22) imply, respectively, that

$$(d\ln p_c^k) - d\ln p_n^k = \frac{\theta(\alpha-1)}{\theta+\alpha-1}(d\ln h_c^k - d\ln h_n^k)$$

$$d\ln S^k - \eta d\ln y^k = (1-\frac{1}{\theta})d\ln p_c^k + d\ln h_c^k$$

and

$$d\ln y^k - d\ln S^k = -\frac{\alpha}{\alpha-1}d\ln p_c^k$$

These four equations are all log linear, and we can solve for $d\ln y^k$, $d\ln S^k$, $d\ln p_c^k$, and $d\ln p_n^k$ in terms of $d\ln h_c^k$ and $d\ln h_n^k$. The solution for $d\ln S^k$ is equation (23).

9.6 Open Economy Productivity Decomposition

Suppose that factor prices are equalized (in nominal terms) so that we can talk about a single w_c and w_n that prevails everywhere. Then the value of output must be equal to the value of income and so

$$P_y^k y^k = w_c L_c^k + w_n L_n^k,$$

where P_Y^k is the price level of consumption in country k.

The supply of type i labor in country k is given by

$$L_i^k = \frac{L^k p_i^k}{w_i} \left((\eta \Theta^k)^\eta \left(T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta \right)^{1/\theta} \right)^{1/(1-\eta)} \gamma$$

So we can write real GDP per capita in country k relative to a base country 0 as

$$\frac{y^k/L^k}{y^0/L^0} = \frac{P_y^0}{P_y^k} \left(\left(\frac{\Theta^k}{\Theta^0} \right)^\eta \left(\frac{T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta}{T_c \left(w_c h_c^0 \right)^\theta + T_n \left(w_n h_n^0 \right)^\theta} \right)^{\frac{1}{\theta}} \right)^{\frac{1}{1-\eta}}$$

Normalizing $\left((A_c)^\alpha (w_c)^{1-\alpha} + (A_n)^\alpha (w_n)^{1-\alpha} \right)^{\frac{1}{1-\alpha}} = 1$, we have $P_y^k = \left(\Theta^k \right)^{-1}$ and so

$$\frac{y^k/L^k}{y^0/L^0} = \left(\frac{\Theta^k}{\Theta^0} \left(\frac{T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta}{T_c \left(w_c h_c^0 \right)^\theta + T_n \left(w_n h_n^0 \right)^\theta} \right)^{\frac{1}{\theta}} \right)^{\frac{1}{1-\eta}}$$

Rearranging, we obtain

$$\frac{y^k/L^k}{y^0/L^0} = \left(\frac{\Theta^k}{\Theta^0} \left(\frac{T_c \left(w_c h_c^0 \right)^\theta}{T_c \left(w_c h_c^0 \right)^\theta + T_n \left(w_n h_n^0 \right)^\theta} \left(\frac{T_c \left(w_c h_c^k \right)^\theta}{T_c \left(w_c h_c^0 \right)^\theta} \right) + \frac{T_n \left(w_n h_n^0 \right)^\theta}{T_c \left(w_c h_c^0 \right)^\theta + T_n \left(w_n h_n^0 \right)^\theta} \left(\frac{T_n \left(w_n h_n^k \right)^\theta}{T_n \left(w_n h_n^0 \right)^\theta} \right) \right)^{\frac{1}{\theta}} \right)$$

now replacing the expressions with occupations from the base country, we obtain

$$\frac{y^k/L^k}{y^0/L^0} = \left(\frac{\Theta^k}{\Theta^0} \left(p_c^0 \left(\frac{h_c^k}{h_c^0} \right)^\theta + p_n^0 \left(\frac{h_n^k}{h_n^0} \right)^\theta \right)^{\frac{1}{\theta}} \right)^{\frac{1}{1-\eta}}$$

So, the difference with the closed economy is in the exponents. Note, however, that since this is holding fixed relative prices it cannot be thought used to talk about comparative statics as was the case in the closed economy.

9.7 Proposition 4 and Equation (29)

9.7.1 Proposition 4, Part 1

Using equations (27) and (7), we can show that

$$\begin{aligned}
\eta w_c L_c^k &= \eta L^k p_c^k \left((\eta \Theta^k)^\eta \left(T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta \right)^{1/\theta} \right)^{1/(1-\eta)} \gamma \\
&= \eta L^k \frac{T_c \left(w_c h_c^k \right)^\theta}{T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta} \left((\eta \Theta^k)^\eta \left(T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta \right)^{1/\theta} \right)^{1/(1-\eta)} \gamma \\
&= \gamma L^k \eta^{\frac{1}{1-\eta}} (\Theta^k)^{\frac{\eta}{1-\eta}} T_c \left(w_c h_c^k \right)^\theta [T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta]^{\frac{1}{\theta}\frac{1}{1-\eta}-1}
\end{aligned}$$

By analogy we have

$$\eta w_n L_n^k = \gamma L^k \eta^{\frac{1}{1-\eta}} (\Theta^k)^{\frac{\eta}{1-\eta}} T_n \left(w_n h_n^k \right)^\theta [T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta]^{\frac{1}{\theta}\frac{1}{1-\eta}-1}$$

Adding up these equations we get

$$\begin{aligned}
\eta w_c L_c^k + \eta w_n L_n^k &= \gamma L^k \eta^{\frac{1}{1-\eta}} (\Theta^k)^{\frac{\eta}{1-\eta}} [T_n \left(w_n h_n^k \right)^\theta + T_c \left(w_c h_c^k \right)^\theta][T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta]^{\frac{1}{\theta}\frac{1}{1-\eta}-1} \\
&= \gamma L^k \eta^{\frac{1}{1-\eta}} (\Theta^k)^{\frac{\eta}{1-\eta}} [T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta]^{\frac{1}{\theta}\frac{1}{1-\eta}}
\end{aligned}$$

Using the output identity $P_y^k y^k = w_c L_c^k + w_n L_n^k$, where $P_y^k = \left(\Theta^k \right)^{-1}$, we have

$$\begin{aligned}
\eta y^k &= \eta \frac{w_c L_c^k + w_n L_n^k}{P_y^k} \\
&= \frac{1}{\left(\Theta^k \right)^{-1}} \gamma L^k \eta^{\frac{1}{1-\eta}} (\Theta^k)^{\frac{\eta}{1-\eta}} [T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta]^{\frac{1}{\theta}\frac{1}{1-\eta}} \\
&= \gamma L^k \eta^{\frac{1}{1-\eta}} (\Theta^k)^{\frac{1}{1-\eta}} [T_c \left(w_c h_c^k \right)^\theta + T_n \left(w_n h_n^k \right)^\theta]^{\frac{1}{\theta}\frac{1}{1-\eta}}
\end{aligned}$$

This expression and equation (26) imply that $E^k L^k = \eta y^k$; i.e. equation (20) still holds under open economy.

9.7.2 Proposition 4, Part 2

Using equations (17), (26) and (27), we can show that

$$\begin{aligned}
S^k &= b\frac{L_c^k}{L^k} \\
&= b\frac{p_c^k}{w_c}\left((\eta\Theta^k)^\eta\left(T_c\left(w_c h_c^k\right)^\theta + T_n\left(w_n h_n^k\right)^\theta\right)^{1/\theta}\right)^{1/(1-\eta)}\gamma \\
&= b\frac{p_c^k}{w_c}\frac{E^k}{\eta\Theta^k} \\
\iff & \quad w_c = b\frac{p_c^k}{S^k}\frac{E^k}{\eta\Theta^k} = b\frac{p_c^k}{S^k}\frac{E^k}{\eta}\frac{1}{\Theta^k}
\end{aligned}$$

We now use equation (7) to obtain that $T_c(w_c h_c)^\theta + T_n(w_n h_n)^\theta = \frac{T_i(w_i h_i)^\theta}{p_i}$. This expression allows us to substitute out the term $T_c(w_c h_c)^\theta + T_n(w_n h_n)^\theta$ in equation (27), giving us, together with equation (17), that

$$\begin{aligned}
S^k &= b\frac{L_c^k}{L^k} \\
&= bp_c^k\left(h_c^k(\eta\Theta^k w_c)^\eta\left(\frac{T_c}{p_c^k}\right)^{1/\theta}\right)^{1/(1-\eta)}\gamma \\
&= b(p_c^k)^{1-\frac{1}{\theta(1-\eta)}}\gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(\Theta^k w_c)^{\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}}
\end{aligned}$$

We then substitute out w_c using $b\frac{p_c^k}{S^k}\frac{E^k}{\eta}\frac{1}{\Theta^k}$ to obtain

$$\begin{aligned}
S^k &= b(p_c^k)^{1-\frac{1}{\theta(1-\eta)}}\gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(\Theta^k b\frac{p_c^k}{S^k}\frac{E^k}{\eta}\frac{1}{\Theta^k})^{\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}} \\
&= (\frac{1}{S^k})^{\frac{\eta}{1-\eta}}b^{\frac{1}{1-\eta}}\gamma\eta^{\frac{\eta}{1-\eta}}(T_c)^{\frac{1}{\theta(1-\eta)}}(p_c^k)^{1-\frac{1}{\theta(1-\eta)}+\frac{\eta}{1-\eta}}(\frac{E^k}{\eta})^{\frac{\eta}{1-\eta}}(h_c^k)^{\frac{1}{1-\eta}} \\
\Leftrightarrow & \quad S^k = b\gamma^{1-\eta}\eta^\eta(T_c)^{\frac{1}{(1-\eta)}}(p_c^k)^{1-\frac{1}{\theta}}(\frac{y^k}{L^k})^\eta h_c^k
\end{aligned}$$

where we have used the relationship $E^k L^k = \eta y^k$. The log of this expression is equation (21).

9.7.3 Equation (29)

Equation (7) implies that under open economy

$$\frac{p_c^k}{p_n^k} = \frac{T_c(w_c h_c^k)^\theta}{T_n(w_n h_n^k)^\theta}$$

This expression implies equation (29).

9.8 Open Economy Market Clearing Conditions

In this appendix, we derive the comparative static equilibrium conditions. Given free trade and Walras' law, we need total supply of cognitive labor to be equal to global demand of cognitive labor:

$$\sum_k L_c^{kS} = \sum_k L_c^{kD}$$

where L_c^{kS} and L_c^{kD} are supply and demand for cognitive labor in country k. Now consider an alternate equilibrium with variables denoted by primes $'$. In the comparative static, we have

$$\frac{\sum_k L_c^{\prime kS}}{\sum_k L_c^{kS}} = \frac{\sum_k L_c^{\prime kD}}{\sum_k L_c^{kD}} \Leftrightarrow \sum_k \frac{L_c^{kS}}{\sum_j L_c^{jS}} \left(\frac{L_c^{\prime kS}}{L_c^{kS}}\right) = \sum_k \frac{L_c^{kD}}{\sum_j L_c^{jD}} \frac{L_c^{\prime kD}}{L_c^{kD}}$$

For the weights, we know that in each country cognitive labor income is the product of its share in employment and GDP, so

$$L_c^{kS} = p_c^k \frac{P_Y^k y^k}{w_c}.$$

Furthermore, because all countries have access to the same technology (up to a Hick's neutral shifter) and face the same factor prices that the cost share of cognitive labor is the same in all countries:

$$L_c^{kD} = \frac{A_c^{\alpha} w_c^{1-\alpha}}{A_c^{\alpha} w_c^{1-\alpha} + A_n^{\alpha} w_n^{1-\alpha}} \frac{P_Y^k y^k}{w_c}.$$

Defining changes to variable x as $\widehat{x} = x'/x$, we can use these two expressions to write the equilibrium condition as

$$\sum_k \frac{p_c^k P_Y^k y^k}{\sum_j p_c^j P_Y^j y^j} \widehat{L}_c^{kS} = \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} \widehat{L}_c^{kD}$$

Using the expressions for labor supply in each countries and doing the hat algebra, we have after a series of straightforward simplifications

$$\widehat{L}_c^{kS} = \widehat{L}^k (\widehat{\Theta}^k)^{\frac{\eta}{1-\eta}} (\widehat{w}_c)^{\theta-1} \left(\widehat{h}_c^k\right)^{\theta} \left(p_c^k \left(\widehat{w}_c \widehat{h}_c^k\right)^{\theta} + p_n^k \left(\widehat{w}_n \widehat{h}_n^k\right)^{\theta}\right)^{\frac{1}{\theta(1-\eta)}-1},$$

and

$$\widehat{L}_c^{kD} = \widehat{L}^k(\widehat{\Theta}^k)^{\frac{\eta}{1-\eta}} \left(\widehat{w}_c\right)^{-\alpha} \left(p_c^k \left(\widehat{w}_c\widehat{h}_c^k\right)^\theta + p_n^k \left(\widehat{w}_n\widehat{h}_n^k\right)^\theta\right)^{\frac{1}{\theta(1-\eta)}},$$

where we have used the normalization $A_c^\alpha w_c^{1-\alpha} + A_n^\alpha w_n^{1-\alpha} = 1$ to arrive at the change in cognitive labor demand. Substituting these two expressions into the equilibrium condition, we have cognitive labor market clearing condition:

$$\left(\widehat{w}_c\right)^{\theta+\alpha-1} \sum_k \frac{p_c^k P_Y^k y^k}{\sum_j p_c^j P_Y^j y^j} \widehat{L}^k(\widehat{\Theta}^k)^{\frac{\eta}{1-\eta}} \left(\widehat{h}_c^k\right)^\theta \left(p_c^k \left(\widehat{w}_c\widehat{h}_c^k\right)^\theta + p_n^k \left(\widehat{w}_n\widehat{h}_n^k\right)^\theta\right)^{\frac{1}{\theta(1-\eta)}-1}$$

$$= \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} \widehat{L}^k(\widehat{\Theta}^k)^{\frac{\eta}{1-\eta}} \left(p_c^k \left(\widehat{w}_c\widehat{h}_c^k\right)^\theta + p_n^k \left(\widehat{w}_n\widehat{h}_n^k\right)^\theta\right)^{\frac{1}{\theta(1-\eta)}}.$$

This equation is (30) in the text. From the normalization, we must have

$$\Theta_c \left(\widehat{w}_c\right)^{1-\alpha} + \left(1-\Theta_c\right)\left(\widehat{w}_n\right)^{1-\alpha} = 1,$$

where

$$\Theta_c \equiv \frac{\left(A_c\right)^\alpha \left(w_c\right)^{1-\alpha}}{\left(A_c\right)^\alpha \left(w_c\right)^{1-\alpha} + \left(A_n\right)^\alpha \left(w_n\right)^{1-\alpha}}$$

is the cost share of cognitive labor used in global production. By definition, this is

$$\Theta_c = \frac{w_c \sum_k L_c^k}{\sum_k P_Y^k y^k}.$$

Because $w_c L_c^k = p_c^k P_Y^k y^k$, this can be rewritten

$$\left(\widehat{w}_c\right)^{1-\alpha} \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} p_c^k + \left(\widehat{w}_n\right)^{1-\alpha} \sum_k \frac{P_Y^k y^k}{\sum_j P_Y^j y^j} p_n^k = 1. \tag{34}$$

Equations (30) and (34) solve for global changes in cognitive and non-cognitive labor. The only data that is required to make these calculations are countries initial real GDPs ($P_Y^k y^k$) and the share of workers employed in cognitive occupations (p_c^k).

10 Data Appendix

1. Sample Cuts for NLSY-79 Data

Following Neal and Johnson (1996) we: (1) use the 1989 version of AFQT and drop the observations with missing AFQT scores; (2) drop those whose wage exceeds

\$75 or below \$1 in 1991; and (3) drop those who are older than 17 when they take the AFQT.

2. O*NET Data

The following is the list of O*NET task ID's of the measures we discuss in the text. Leadership is 4.A.4.b.4, and enterprising 1.B.1.e. Enterprising skills involve "starting up and carrying out projects" and "leading people and making many decisions".

In addition, we have experimented with the following candidate measures. (1) Originality is about coming up with "unusual or clever ideas about a given topic or situation", or developing "creative ways to solve a problem". 1.A.1.b.2. (2) Social skills involve "working with, communicating with, and teaching people". 1.B.1.d. (3) Artistic talents show up when "working with forms, designs and patterns", where "the work can be done without following a clear set of rules". 1.B.1.c 2. (4) Investigative skills involve "working with ideas" and "searching for facts and figuring out problems mentally", and require "an extensive amount of thinking"; 1.B.1.b. The results are in Table A1.

When we use originality, social skills or investigative skills to measure non-cognitive skills, the AFQT coefficient of the non-cognitive sub-sample is larger than the cognitive sub-sample. This is counter-intuitive. On the other hand, for the artistic-talent sub-sample, the AFQT coefficient is negative, meaning that the artists with higher test scores have lower wages. However, out of the NLSY-79 sample of over 3000, there are only 30 artists, less than 1% of the sample size.

3. ILO Employment-by-Occupation Data

We map the O*NET occupation codes into the ISCO-88 codes using the crosswalk at the National Crosswalk center ftp://ftp.xwalkcenter.org/DOWNLOAD/xwalks/. We drop the following observations from the ILO raw data because of data quality issues. 1. All data from Cyprus, because the data source is official estimate (source code "E"). 2. Year 2000 for Switzerland, because over 1 million individuals, a large fraction of the Switzerland labor force, are "not classified". 3. Uganda, Gabon, Egypt, Mongolia, Thailand, Poland in 1994 and Romania in 1992, because the aggregate employment of the sub-occupation categories does not equal the number under "Total". 4. Estonia in 1998, S. Korea in 1995, and Romania in 2000, because the data is in 1-digit or 2-digit occupation codes.

Most countries have a single year of data around 2000. In Figure A1 we plot the non-cognitive employment share for all the countries that have multiple years of data. Within countries the non-cognitive employment share shows very limited variation over time. As a result, for this set of countries we keep the single year of data closest to 2000; e.g. 1990 for Switzerland, 2000 for U.S. and Australia, etc. By construction, the non-cognitive and cognitive employment shares sum to 1 by country.

4. Test Score Data

We have tabulated over-time changes of PISA scores within countries and found very little variation. For example, for the U.S. reading score the mean is 499.26 and the standard deviation is 3.93. We list these summary statistics by country by subject in Table A2.

There have been several international tests on adults: IALS (International Adult Literacy Survey), administered in 1994-1998, ALLS (Adult Literacy and Life Skills Survey), conducted in 2002-2006, and PIAAC (Program for the International Assessment of Adult Competencies), conducted in 2013. The response rate of IALS, 63%, is substantially lower than the initial wave of PISA in 2000, 89% (Brown et al. 2007). ALLS was designed as a follow-up to IALS, but only 5 countries participated. Of the 28 countries in our sample, only 18 participated in IALS, and only 21 in PIAAC. This would represent a 36% and 25% reduction in the number of observations, respectively.

We regress the 2012 PISA scores on 2013 PIAAC scores, for reading and math, for all the countries that participated in both tests, including those that are not in our sample. We obtain, respectively, the coefficient estimate of 0.938 and 1.067, and R-square of 0.508 and 0.527.

5. Correlation Coefficients of Output TFP Estimates

In Table A4 we report the full correlation table among our output TFP estimates, Θ^k, and those reported in the literature. Ours = our estimates for Θ^k; HJ98 = Hall and Jones (1998) TFP (A); KRC97 = Klenow and Rodriguez-Clare (1997); EK96 = Eaton and Kortum (1996); HR97 = Harrigan (1997); PWT_90 = Penn World Tables 8.0, current PPP, year 1990; PWT_00 = PWT 8.0, current PPP, 2000; EK 02 = Eaton and Kortum (2002). The correlation coefficients between our Θ^k and the literature's estimates, reported in the first column of Table A4 and in boldface, are comparable to those among the literature's estimates, reported in the rest of Table A4.

Figure 1 Test Score and Educational Spending Per Capita

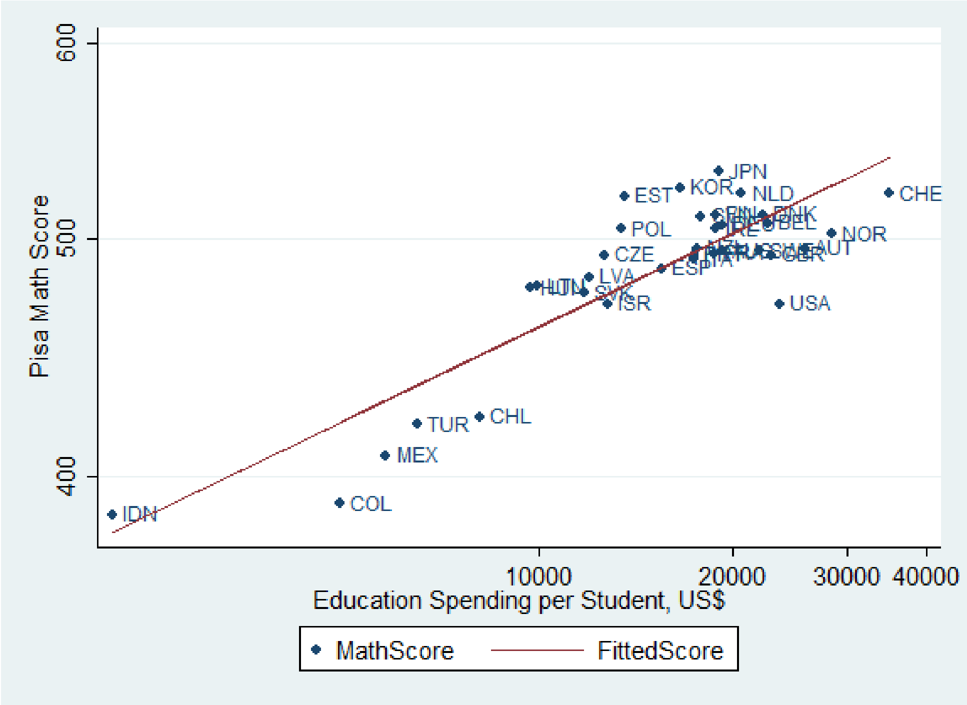

Figure 2 Figure 2 Normalized Test Scores and Cognitive Employment Shares

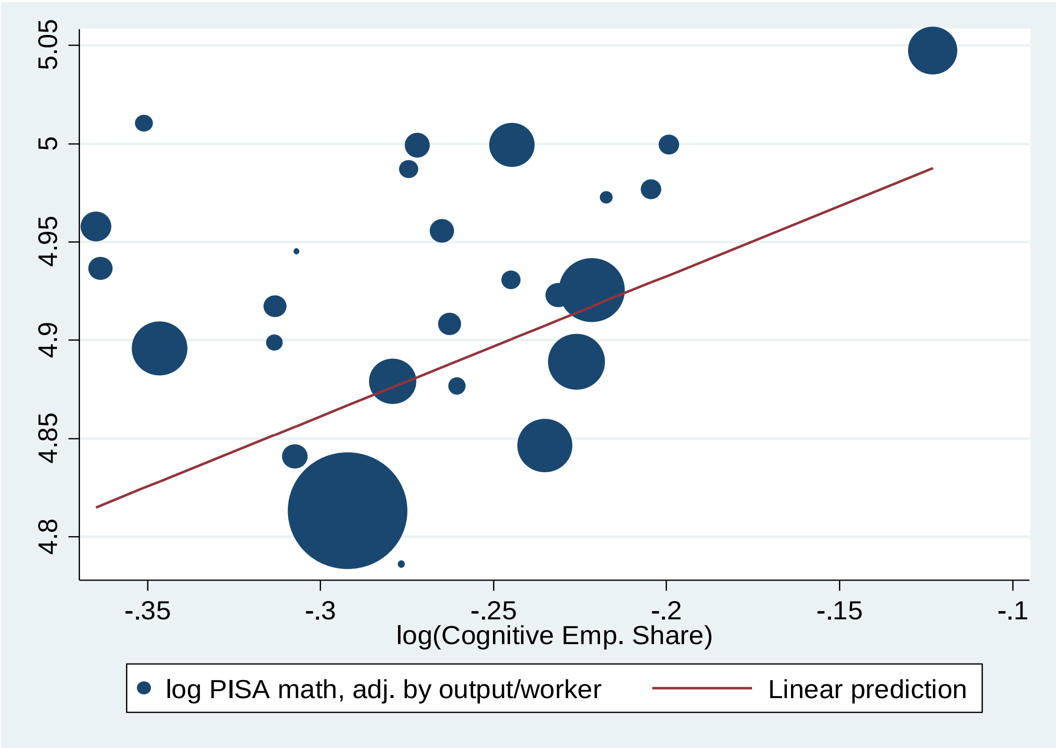

Figure 3 Cognitive-Productivity Ranking vs. PISA-Math Ranking

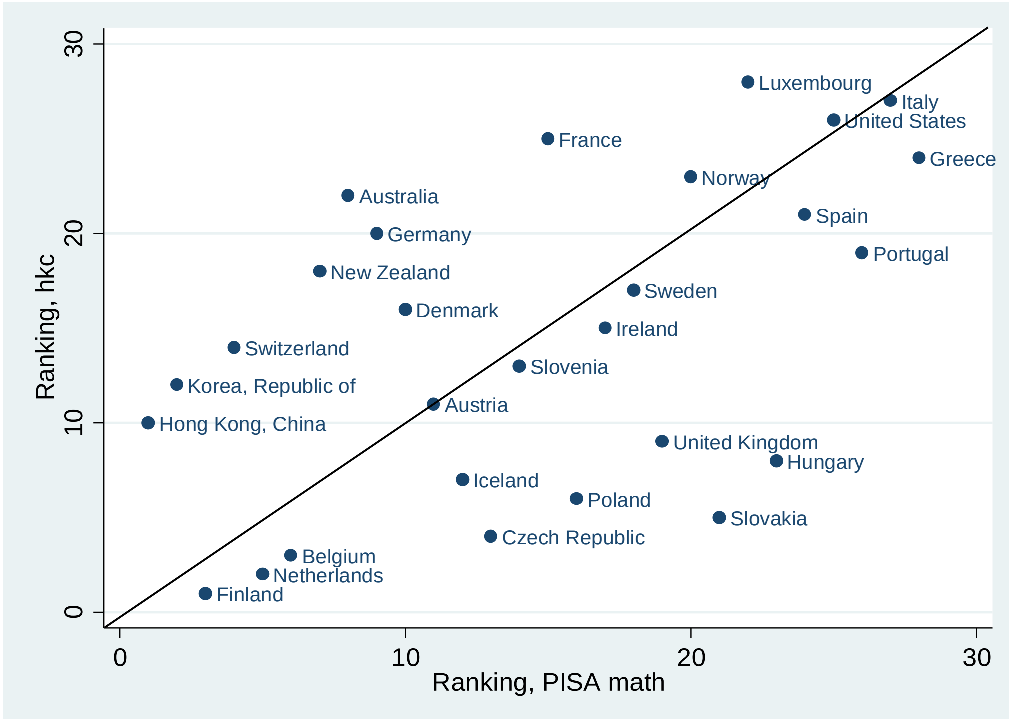

Figure 4 Non-Cognitive-Productivity Ranking vs. PISA-Math Ranking

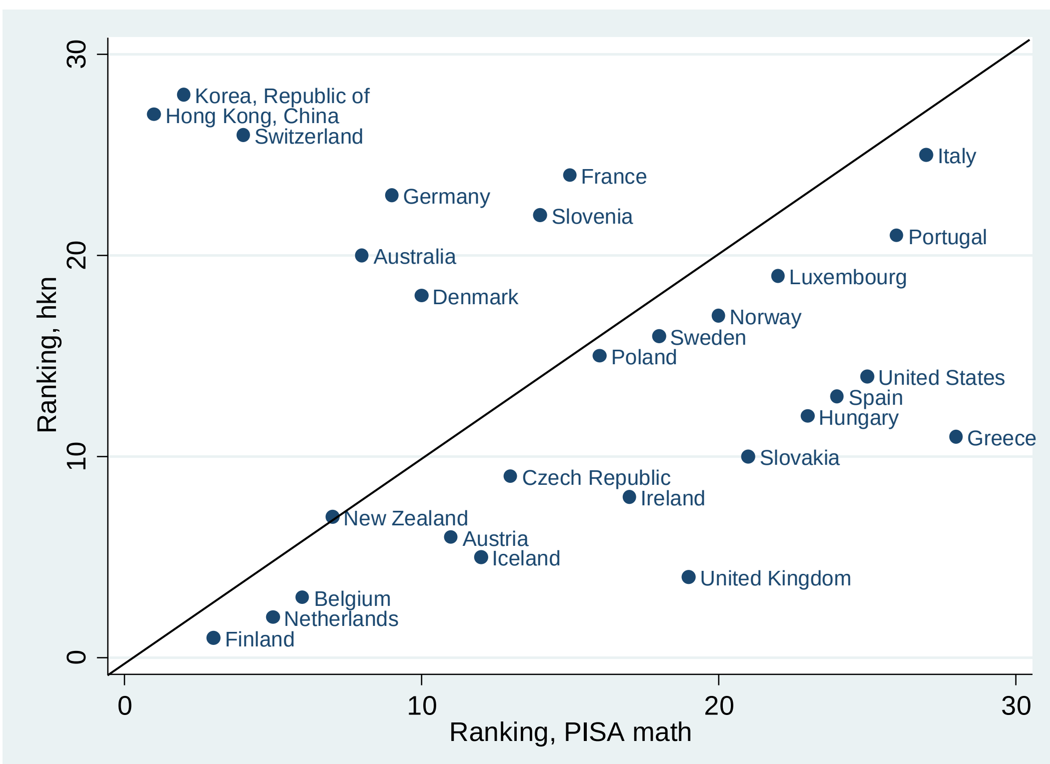

Figure 5 Overall Education Quality

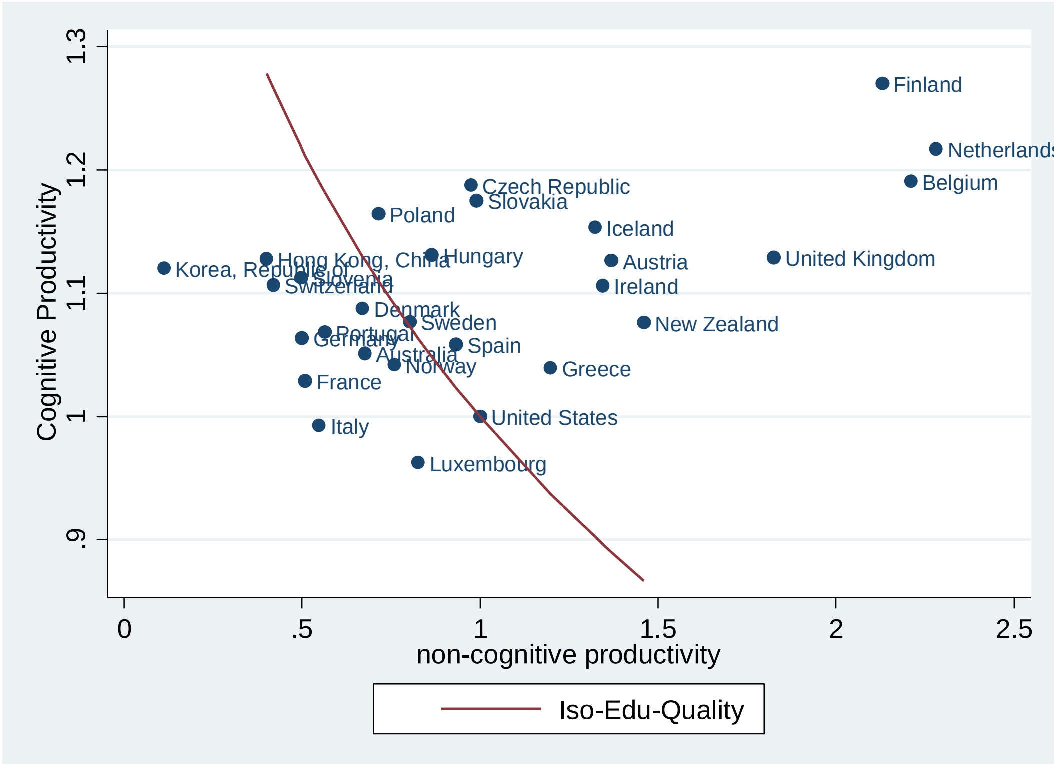

Table 1 Test Score and Wages of Non-cognitive and Cognitive Occupations

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Replicate	Non-Cog. SubSample	Cog. SubSample	Interaction	Alt. Leadership
Black	-0.0537***	-0.0937**	-0.0381*	-0.0661***	-0.0641***
	(0.0196)	(0.0365)	(0.0228)	(0.0191)	(0.0192)
Hispanics	0.0425**	0.0164	0.0482*	0.0413**	0.0414**
	(0.0211)	(0.0378)	(0.0251)	(0.0206)	(0.0206)
Age	0.0349***	0.0483***	0.0285***	0.0323***	0.0316***
	(0.00708)	(0.0129)	(0.00833)	(0.00689)	(0.00690)
Non-cog. Occp.				0.121***	0.127***
				(0.0163)	(0.0186)
College				0.187***	0.195***
				(0.0264)	(0.0263)
AFQT	0.183***	0.157***	0.183***	0.137***	0.125***
	(0.00964)	(0.0182)	(0.0113)	(0.0115)	(0.0113)
AFQT2	-0.0130	-0.0199	-0.00717	-0.0369***	-0.0358***
	(0.00802)	(0.0143)	(0.00961)	(0.00950)	(0.00956)
AFQT x Non-Cog.				-0.0345**	-0.00749
				(0.0159)	(0.0182)
AFQT x College				0.0525**	0.0495**
				(0.0245)	(0.0244)
Constant	6.233***	6.148***	6.281***	6.218***	6.232***
	(0.112)	(0.205)	(0.132)	(0.109)	(0.109)
Obs. No.	3,210	951	2,259	3,210	3,210
R2	0.168	0.151	0.163	0.214	0.211

Notes: The dependent variable is log wage, and the sample is NLSY 79. Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Table 2 Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Labor Force Size	28	12541.24	23132.62	156.43	120464.70
Non-cog. Emp. Share	28	0.2425	0.0514	0.1157	0.3775
Cognitive Emp. Share	28	0.7575	0.0514	0.6225	0.8843
Agg. Output ($000)	28	4.59E+08	1.18E+09	4130208	6.25E+09
Edu. Exp./Output	20	0.1255	0.0194	0.0985	0.1695
PISA Reading Score	28	498.96	18.30	468.93	539.34
PISA Math Score	28	503.73	22.17	455.80	553.40
PISA Science Score	28	506.81	19.70	470.07	554.28

Table 3 Sample Countries, Years and Rankings

Country	Year	Cog-Prod Ranking	PISA Math Ranking	Non-Cog Prod Ranking
Australia	2000	22	8	19
Austria	2000	11	11	6
Belgium	2000	3	6	2
Czech Republic	2000	4	13	12
Denmark	2000	16	10	20
Finland	2000	1	3	3
France	2000	25	15	23
Germany	2000	20	9	24
Greece	2000	24	28	9
Hong Kong, China	2001	10	1	27
Hungary	2000	8	23	14
Iceland	2000	7	12	8
Ireland	2000	15	17	7
Italy	2000	27	27	22
Korea, Republic of	2000	12	2	28
Luxembourg	2000	28	22	15
Netherlands	2000	2	5	1
New Zealand	1996	18	7	5
Norway	2000	23	20	17
Poland	2000	6	16	18
Portugal	2000	19	26	21
Slovakia	2000	5	21	11
Slovenia	2000	13	14	25
Spain	2000	21	24	13
Sweden	2000	17	18	16
Switzerland	1990	14	4	26
United Kingdom	2000	9	19	4
United States	2000	26	25	10

Table 4 Value of θ

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
lnp_c^k	0.717***	0.714***	0.696***	0.521***	0.512**	0.677*
	(0.230)	(0.224)	(0.223)	(0.165)	(0.201)	(0.357)
ASNZ		0.213**	0.208***	0.189***	0.189**	0.175**
		(0.0773)	(0.0574)	(0.0570)	(0.0695)	(0.0842)
Constant	5.076***	5.075***	5.072***	5.032***	5.040***	5.032***
	(0.0624)	(0.0607)	(0.0608)	(0.0448)	(0.0546)	(0.0784)
Observations	26	28	28	28	28	28
R-square	0.288	0.347	0.393	0.384	0.292	0.196

Notes: ASNZ is the dummy for Australia and New Zealand, whose raw occupation-employment data are in different classification codes as compared with the other countries in our sample.

Table 5 Value of α

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
$\ln(1 + p_n^k / p_c^k)$	3.125**	3.112**	3.046**	2.932**	2.923**	3.562*
	(1.224)	(1.259)	(1.205)	(1.170)	(1.210)	(1.846)
ASNZ	-1.094**		-1.093***	-1.070**	-1.070**	-0.971**
	(0.423)		(0.310)	(0.404)	(0.418)	(0.435)
Constant	3.465***	3.469***	3.486***	3.509***	3.501***	3.526***
	(0.332)	(0.342)	(0.329)	(0.318)	(0.328)	(0.405)
Observations	28	26	28	28	28	28
R-squared	0.282	0.203	0.354	0.283	0.269	0.212

Notes: ASNZ is the dummy for Australia and New Zealand, whose raw occupation-employment data are in different classification codes as compared with the other countries in our sample.

Table 6 Summary of Parameter Values and Identification

Parameters	Intuition	Values	Identification
η	Elasticity in Human Cap Prod	0.1255	Edu. spending as share of output, (20)
θ	Dispersion of Innate Ability	2.0877~3.4965	Strength of selection effect, (21)
α	Sub Elasticity in Agg Production	1.4706~1.5200	Agg. production function, (22)
Θ^k	Output TFP	Table 8	Output per worker, test score and relative emp. share, given α, (22)
h_c^k	TFP of Cognitive Education	Table 3	Normalized test score and cog. emp. share, given θ and η, (21)
h_n^k	TFP of Non-cognitive Education	Table 3	Revealed comp advantage by relative emp. share, given α and θ, (19)

Table 7 Alternative Parameter Values and Alternative Setting

	Cog Productivity		Non-cog Productivity		Overall Edu Quality	
	Value	Ranking	Value	Ranking	Value	Ranking
closed-economy, θ = 2.0887,						
	0.9844	0.9583	1.0000	0.9998	0.9888	0.9788
closed-economy, θ = 2.0887, α = 2						
	0.9844	0.9583	0.9906	0.9972	0.9798	0.9740
open-economy, θ = 3.4965						
	Identical	Identical	0.9160	0.8577	0.8365	0.7757
open-economy, θ = 2.0877						
	Identical	Identical	0.9562	0.9316	0.8824	0.8058

Notes: This table reports the correlation coefficients between the values and country rankings of cognitive productivity, non-cognitive productivity, and overall educational quality under our main specification and under alternative parameter values and settings.

Table 8 Contributions of Overall Education Quality to Output per Worker

		Closed-Econ Setting		Open-Econ Setting	
Countries	(1) Output Per Worker	(2) Contribution of Output TFP	(3) Contribution of Overall Edu Quality	(4) Contribution of Output TFP	(5) Contribution of Overall Edu Quality
Austria	0.6434	0.5297	1.2147	0.5576	1.1539
Belgium	0.6892	0.4636	1.4867	0.5516	1.2496
Czech Republic	0.3293	0.2860	1.1513	0.2724	1.2090
Denmark	0.5979	0.6187	0.9664	0.5517	1.0837
Finland	0.5037	0.3259	1.5458	0.3760	1.3398
France	0.7329	0.8517	0.8606	0.7250	1.0109
Germany	0.6296	0.7126	0.8834	0.6003	1.0487
Greece	0.5190	0.4761	1.0901	0.4942	1.0501
Hong Kong	0.6864	0.7724	0.8887	0.6166	1.1132
Hungary	0.3517	0.3292	1.0684	0.3082	1.1411
Iceland	0.5110	0.4168	1.2261	0.4320	1.1827
Ireland	0.6642	0.5583	1.1896	0.5879	1.1298
Italy	0.6761	0.7977	0.8476	0.6947	0.9733
S. Korea	0.4304	0.6027	0.7142	0.3996	1.0772
Luxembourg	1.4376	1.5674	0.9172	1.5092	0.9526
Netherlands	0.6712	0.4387	1.5300	0.5235	1.2822
Norway	0.7289	0.7589	0.9605	0.7028	1.0371
Poland	0.3045	0.2917	1.0438	0.2599	1.1715
Portugal	0.3845	0.4216	0.9121	0.3637	1.0573
Slovakia	0.2979	0.2600	1.1459	0.2491	1.1957
Slovenia	0.3929	0.4275	0.9191	0.3566	1.1020
Spain	0.6087	0.5913	1.0293	0.5730	1.0623
Sweden	0.5937	0.5917	1.0034	0.5508	1.0779
Switzerland	0.5855	0.6641	0.8816	0.5367	1.0908
United Kingdom	0.6349	0.4758	1.3345	0.5430	1.1694
United States	1.0000	1.0000	1.0000	1.0000	1.0000

Notes: Columns (2) and (3) are obtained using equations (15) and (16), and columns (4) and (5) obtained using equation (28).

Table 9 Patterns of Trade

	Dep. Var. = Revealed Comp Advantage		
	(1)	(2)	(3)
Non-cog abundance x non-cog intensity	15.989	15.979	10.615
	(2.92)	(2.92)	(2.02)
Cap abundance x cap intensity		0.000	0.000
		(0.10)	(0.22)
Skill abundance x skill intensity			9.173
			(4.71)
constant	-1.108	-1.113	1.976
	(-3.30)	(-3.28)	(2.77)
industry FE	yes	yes	yes
country FE	yes	yes	yes
R^2	0.369	0.369	0.401
# obs.	1103	1103	1103

Notes: the dependent variable is net export normalized by the sum of import and export values.

Figure A1 Non-Cognitive Employment Share Over Time for Select Countries

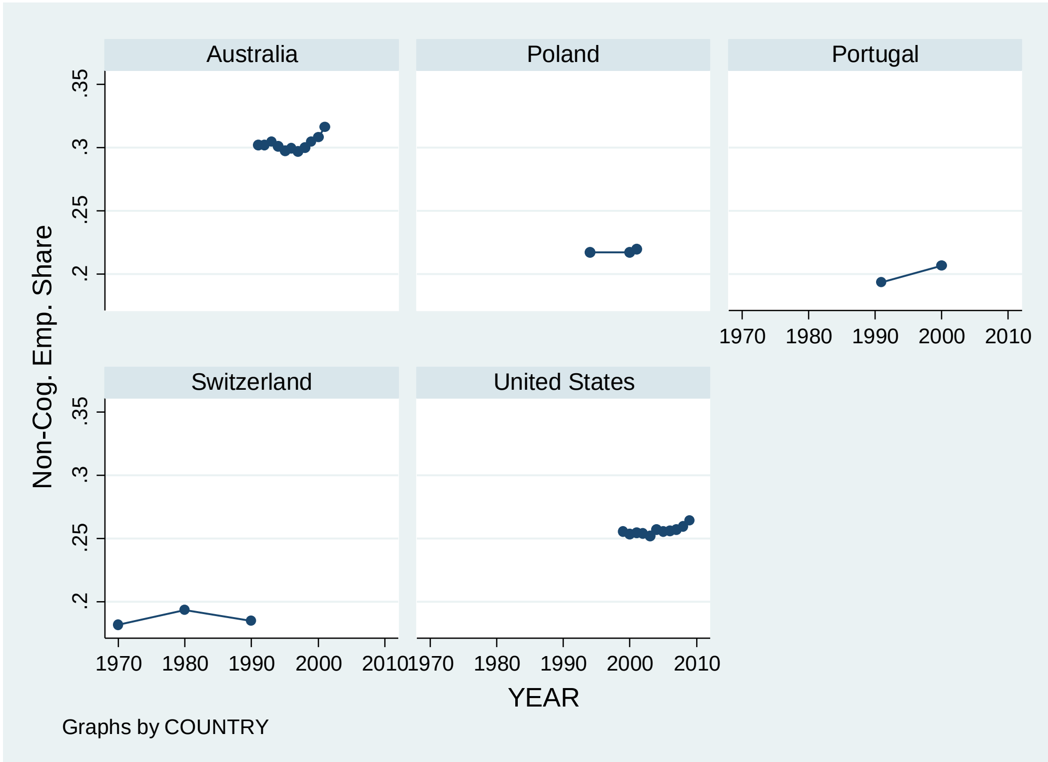

Table A1 Neal-Johnson Regressions for Alternative Measures of Non-Cognitive Skills

VARIABLES	Originality	Not Originality	Social-skill	Not Social-skill	Artistic	Not Artistic	Investigative	Not Investigative
Black	-0.0735*	-0.0463**	0.0238	-0.0515**	-1.490*	-0.0533***	0.010	-0.060***
	(0.0395)	(0.0216)	(0.0683)	(0.0202)	(0.799)	(0.0195)	(0.091)	(0.02)
Hispanics	0.0380	0.0398*	0.119	0.0364*	-0.586*	0.0422**	0.036	0.039*
	(0.0402)	(0.0240)	(0.0788)	(0.0215)	(0.331)	(0.0212)	(0.092)	(0.022)
Age	0.0569***	0.0220***	0.0557**	0.0325***	0.0752	0.0345***	0.027	0.036***
	(0.0136)	(0.00798)	(0.0254)	(0.00722)	(0.0844)	(0.00710)	(0.030)	(0.007)
AFQT	0.182***	0.154***	0.204***	0.185***	-0.713**	0.184***	0.188***	0.171***
	(0.0210)	(0.0109)	(0.0370)	(0.00979)	(0.333)	(0.00965)	(0.060)	(0.010)
$AFQT^2$	0.00428	-0.0382***	-0.00483	-0.0172**	0.299*	-0.0120	-0.043	-0.019**
	(0.0149)	(0.00996)	(0.0341)	(0.00807)	(0.150)	(0.00809)	(0.032)	(0.008)
Constant	5.942***	6.414***	5.732***	6.292***	6.061***	6.239***	6.642***	6.212***
	(0.216)	(0.126)	(0.403)	(0.114)	(1.357)	(0.112)	(0.481)	(0.114)
Obs. No.	1,096	2,114	382	2,828	30	3,180	158	3052
R^2	0.164	0.126	0.127	0.181	0.188	0.170	0.106	0.148

Table A2 Within-Country, Over-Time Variations of PISA Scores

	Mean			Std. Dev.		
Country	Math	Reading	Science	Math	Reading	Science
Australia	515.67	518.66	525.21	8.69	7.63	3.23
Austria	505.54	490.64	508.31	0.07	1.05	3.58
Belgium	519.86	505.92	507.27	6.80	2.96	2.81
Canada	526.10	527.32	529.54	5.96	4.37	4.57
Czech Republic	504.52	486.78	507.22	10.63	6.20	6.25
Denmark	507.66	494.94	497.90	7.08	1.75	1.79
Estonia	515.74	506.00	533.54	4.34	8.91	7.04
Finland	537.98	539.34	554.28	13.21	9.63	8.94
France	499.53	497.95	497.47	7.55	7.35	1.99
Germany	508.27	495.06	520.06	5.66	8.67	4.25
Greece	455.80	473.15	470.07	9.02	8.53	3.33
Hungary	487.04	485.37	500.29	6.68	5.86	5.23
Iceland	505.03	493.19	488.18	9.20	10.38	9.01
Ireland	498.24	515.65	512.77	7.43	12.05	8.00
Italy	473.90	481.49	485.92	11.95	9.04	9.41
Japan	530.66	515.24	539.19	5.92	17.19	7.68
S. Korea	547.42	537.98	532.64	4.78	11.42	9.09
Luxembourg	490.53	479.69	487.16	1.83	6.39	3.72
Netherlands	529.32	509.88	523.05	6.49	2.85	1.57
New Zealand	516.13	520.89	526.01	11.06	5.89	9.02
Norway	493.09	499.30	493.64	4.18	8.63	6.72
Poland	499.49	500.41	510.57	12.23	14.46	14.17
Portugal	476.53	479.42	485.51	12.06	8.77	9.87
Slovak Republic	492.15	468.93	483.30	7.47	6.25	10.52
Slovenia	502.35	486.27	514.91	1.83	7.10	3.59
Spain	483.22	480.58	491.04	2.27	12.12	4.69
Sweden	495.98	503.74	494.41	13.27	13.58	9.28
Switzerland	530.28	500.46	514.46	3.07	5.34	2.63
United Kingdom	493.93	496.19	514.20	1.52	2.74	0.53
United States	481.50	499.26	496.11	5.41	3.93	6.64

Table A3 Correlation between 2012 PISA and 2013 PIAAC scores

	PISA Reading	PISA Math
PIAAC Literacy	0.938	
	(5.18)	
PIAAC Numeracy		1.067
		(5.38)
Constant	249.047	215.948
	(5.13)	(4.13)
Obs. No.	28	28
R^2	0.508	0.527

Table A4 Correlation Coefficients for Output TFP Estimates

	Ours	HJ98	KRC97	EK 96	HR97	PWT_90	PWT_00
Ours	1						
HJ98	**0.5600**	1					
	0.0029						
KRC97	**0.4674**	0.8412	1				
	0.0327	0					
EK 96	**0.5669**	0.5348	0.7109	1			
	0.0220	0.0328	0.002				
HR97	**0.6054**	0.5841	0.5394	0.068	1		
	0.1117	0.1284	0.1677	0.8729			
PWT_90	**0.6214**	0.8792	0.7401	0.6976	0.6126	1	
	0.0004	0	0.0001	0.0027	0.1064		
PWT_00	**0.6377**	0.6878	0.2565	0.3856	-0.5382	0.7089	1
	0.0003	0.0001	0.2617	0.1402	0.1688	0	
EK 02	**0.5964**	0.4159	0.4828	0.7655	0.3538	0.6114	0.4646
	0.0115	0.0968	0.0496	0.0009	0.3899	0.0091	0.0602

Notes: Ours = our estimates for Θ^k; HJ98 = Hall and Jones (1998) TFP (A); KRC97 = Klenow and Rodriguez-Clare (1997); EK96 = Eaton and Kortum (1996); HR97 = Harrigan (1997); PWT_90 = Penn World Tables 8.0, current PPP, year 1990; PWT_00 = PWT 8.0, current PPP, 2000; EK 02 = Eaton and Kortum (2002).