NBER WORKING PAPER SERIES

THE PRODUCTION OF COGNITIVE AND NON-COGNITIVE HUMAN CAPITAL IN THE GLOBAL ECONOMY

Chong Xiang
Stephen Yeaple

Working Paper 24524
http://www.nber.org/papers/w24524

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2018

Chong Xiang gratefully acknowledges financial support from the Purdue University Research Center of Economics, or PURCE (2015). Stephen Yeaple's contribution to this paper was funded in part by National Science Foundation Research Grant (SES-1360209, 2014-2017). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Production of Cognitive and Non-cognitive Human Capital in the Global Economy
Chong Xiang and Stephen Yeaple
NBER Working Paper No. 24524
April 2018
JEL No. F16,F63,F66,I21,I25,I26,O15,O43,O47

ABSTRACT

A country's welfare depends on its ability to accumulate cognitive and noncognitive human capital. However, we do not fully understand what makes some countries successful at producing human capital and even struggle with measurement. e.g. international test scores are informative about the cognitive dimension but neglect the non-cognitive dimension. In this paper, we develop a multi-country, open-economy general-equilibrium framework in which countries' ability to turn resources into human capital along the cognitive and non-cognitive dimensions is revealed by the endogenous educational and occupational choices of its citizens and their subsequent performance on international exams. Our model allows us to estimate countries' underlying productivities of cognitive and non-cognitive human capital. We find that high test scores do not necessarily imply high cognitive productivities (e.g. Switzerland, Hong Kong) and that many countries with low test scores have high non-cognitive productivities (e.g. the U.S. and U.K.).

We then aggregate over these two dimensions to construct a single educational quality index, and illustrate its intuition using an iso-education-quality curve. We use our model to decompose variation in output per capita across countries into a component involving the educational quality index and another involving output TFP. This exact decomposition shows that the differences in cognitive and noncognitive productivities across countries have large implications for differences in output per worker. These results help quantify the potential payoffs of education policies and clarify their objective; e.g. excessive attention to test scores may decrease aggregate output.

International trade plays an important role in our model because the gains from trade help to compensate a country for uneven productivity across human capital types. In counterfactual exercises, we show that if barriers to trade are completely eliminated, we would obtain a very different iso-education-quality curve. This implies large improvements of overall education quality, and large gains from trade, for the countries with strong comparative advantages in producing cognitive (e.g. S. Korea would gain 30.1% to 44.1% of its output) or non-cognitive human capital (e.g. the Netherlands would gain 18.8% to 55.6%).

Chong Xiang
Department of Economics
Purdue University
403 West State Street
West Lafayette, IN 47907-2506
cxiang@purdue.edu

Stephen Yeaple
Department of Economics
The Pennsylvania State University
520 Kern Building
University Park, PA 16802-3306
and NBER
sry3@psu.edu

1 Introduction

Human capital is central to both economics and other social sciences. Therefore, understanding how well countries produce their human capital is critical for both academic research and for policy. One way a country's performance is currently assessed is by looking at its students' scores on international assessment tests, like PISA. The U.S.' low PISA scores have alarmed policy makers[1] and motivated major policy changes (e.g. No Child Left Behind of 2001 and Race to the Top of 2009) in the U.S. Like the U.S., many other countries worry that their test scores are *too low* (e.g. U.K., Slovakia and Qatar).[2]

Oddly, many countries whose students excel in international exams worry that their test scores are *too high*! i.e. their educational systems overemphasize formal examination proficiency, and their students spend too much time studying for exams.[3] This concern has also influenced policy; e.g. 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 such soft skills as leadership, co-operation, and communication. While the importance of these soft, or non-cognitive, skills has been clearly established (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 develop a multi-country, open-economy general equilibrium (GE) framework to quantify how countries produce human capital along multiple dimensions. Our model is based on two intuitive premises. The first is that peoples' occupational

[1]e.g. President Obama said that the nation that "out-educates us today will out-compete us tomorrow."

[2]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".

[3]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).

choices reveal information about their skills at different types of tasks, and these skills reflect different types of human capital. For example, a manager issues directions and guidance to subordinates, while a secretary follows these orders and an engineer uses the knowledge in math and science to solve problems. Intuitively, this allows us to infer a country's comparative advantage in fostering non-cognitive relative to cognitive human capital from occupational choice data. The second is that countries' performances on international exams are informative about their accumulation of cognitive human capital. Our model then allows us to use this data to infer countries' absolute advantages in fostering cognitive human capital.

Specifically, we write down production functions of cognitive and non-cognitive human capital, and use the TFP's (Total Factor Productivity) of these production functions to quantify countries' productivities in accumulating cognitive and non-cognitive human capital. 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. Intuitively speaking, a country with a high cognitive (non-cognitive) productivity produces a large quantity of cognitive (non-cognitive) human capital, holding fixed resources inputs.

Researchers have long recognized that incentives matter for educational outcomes,[4] which is closely related to human capital production. We accommodate these incentives 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 individuals' comparative advantages, in turn, are determined by their innate abilities at birth and human capital accumulated. When workers decide how much human capital to invest in, 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, individual workers' human capital accumulation is affected by their occupational choices, which, in turn, depend on the non-cognitive and cognitive productivities of the economy.

[4] Heckman and Kautz (2012) survey earlier studies, which show that incentives, in the form of money or even candy, improve the scores of IQ tests. In more recent work, Behrman, Parker, Todd, and Wolpin (2015) show that providing monetary incentives to Mexican high-school students has substantial and immediate effects on their test scores. Researchers have also shown that instructor incentives matter for the scores of high-stake tests (e.g. Neal and Schanzenbach 2010). PISA, whose scores we use, is not a high-stake test.

Our model implies a set of empirical relationships that we connect with widely available data. 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 empirical literature on international test scores (e.g. Hanushek and Woessmann 2011). We then peel back the confounding factors of resources inputs and incentives under the guidance of our GE model, to reveal the countries' underlying productivities in fostering cognitive human capital. Importantly, this procedure remains the same whether our model setting is closed-economy, open-economy with free trade, or open-economy with positive trade costs. Our results 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).

For the values of non-cognitive productivities, we use occupation employment shares. Specifically, in our model, a country's comparative advantage for non-cognitive human capital, or the ratio of her non-cognitive productivity to cognitive productivity, drives workers' occupational choices. It then follows that this comparative advantage is revealed by the ratio of occupation employment shares. Intuitively speaking, the fact that many individuals in country k choose the non-cognitive occupation suggests that country k has a strong comparative advantage for non-cognitive human capital. If, in addition, country k has a high cognitive productivity, then this country must also have a high non-cognitive productivity.

Here, international trade plays an important theoretical role. In the closed-economy setting, the relative return of non-cognitive human capital depends on country k's comparative advantage for non-cognitive human capital, so that data on occupation employment shares are sufficient to back out this comparative advantage. With international trade, the relative return of non-cognitive human capital is determined globally. Intuitively, if country k is a large net importer of the service of non-cognitive human capital, it must have a strong comparative advantage in non-cognitive human capital, since the non-cognitive workers in k have chosen their occupation despite import competition. Our model delivers an analytical expression for the comparative advantage of human capital production, where the effects of trade are summarized by its factor content in terms of cognitive and non-cognitive human capital.

Looking at our data, we find small values of factor content trade. This result is consistent with previous studies,[5] and implies that the open- and closed-economy settings of our model deliver similar values of non-cognitive productivities. Using these values, we show that countries' non-cognitive-productivity rankings have zero correlation with their PISA score rankings, and many countries with low test scores have high non-cognitive productivities (e.g. the U.S. and U.K.). Therefore, non-cognitive productivities are a novel dimension of the quality of human-capital production that is not revealed by test scores.

Finally, our model allows us to condense the multi-dimensional differences in cognitive and non-cognitive productivities into a single metric, which we call educational quality index. This metric, which resembles an iso-quant, 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 only three key elasticities: the supply elasticity, which is governed by the dispersion of workers' innate abilities; the demand elasticity, which is the substitution-elasticity across different types of human capital in aggregate production; and the output elasticity in the production of human capital. To identify these elasticties, we draw on the parsimonious relationships, predicted by our model, among the variables of test score, output per worker, employment shares of non-cognitive and cognitive occupations, and factor content of trade. Our unique focus on cross-country differences in the production of human capital distinguishes 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, Lee 2017).

Graphing the iso-education-quality curve (the combinations of cognitive and non-cognitive productivities that produce the same educational quality index) allows us to visually assess the large differences in how countries produce their cognitive and non-cognitive human capital. It also shows that both cognitive and non-cognitive productivities are important for the educational quality index; e.g. the countries with imbalanced human capital productivities, such as Germany and Hong Kong, tend to have low overall educational quality. To draw out the economic significance of the educational quality index, we show that the ratio of output per worker between any pair of countries can be decomposed into a power function of this index, multiplied by a power function of the ratio of output TFP. Implementing this exact decomposition using raw data and our

[5] e.g. Trefler (1995), Davis and Weinstein (2001), and Costinot and Rodriguez-Clare (2014).

model parameters, we show that the differences in human-capital productivities across countries have large implications for output per worker. For example, Germany's output per worker is 62.96% of the U.S. level (data), of which 88.34% can be attributed to human-capital productivities (model parameters) and 71.26% to output TFP (model parameters).

The large differences in cognitive and non-cognitive productivities across countries are a source of comparative advantage, and we calculate how much countries' aggregate output would be if there were no frictions to trade. Under free trade, the iso-education-quality curve would have a very different shape and a very different slope, because imbalance in human capital productivities would help countries specialize and enjoy gains from trade. Deriving an analytical expression for output changes, we show large gains from trade liberalization, especially for the countries with strong comparative advantages in producing cognitive (e.g. S. Korea would gain 30.1% to 44.1% of its output) or non-cognitive human capital (e.g. the Netherlands would gain 18.8% to 55.6%). The magnitudes of these gains-from-trade calculations are comparable to the literature (e.g. Arkolakis, Costinot and Rodriguez-Clare 2012, or ACR 2012).

Our model quantifies the potential payoff of education policies. e.g. a 1% increase in U.S. cognitive productivity increases her aggregate output by 0.85%, if U.S. non-cognitive productivity remains unchanged. Our model also helps clarify the objective of education policies. e.g. a 2.5% increase in South Korea's non-cognitive productivity leads to a 0.33% *rise* in her aggregate output but a 0.07% *drop* in her test score. Test score and aggregate output may move in opposite directions in our model, because test score primarily reflects cognitive human capital, whereas both types of human capital are important for overall education quality and 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. Our model brings the rigor of economic modeling into these discussions by quantifying the pros and cons and calculating the net effect on aggregate output. e.g. the payoff of the 1% increase in U.S. cognitive productivity will be completely offset by a 2.95% reduction in U.S. non-cognitive productivity. Finally, the changes in cognitive and non-cognitive productivitites would have very different effects on aggregate output if there were no frictions to trade, as these changes may enhance or undermine countries' comparative advantages. e.g. a 1% increase in non-cognitive productivity for the Netherlands leads to a 0.35% increase in her aggregate output; in comparison, output would increase by

0.84% under free trade. These policy implications of our model and our results speak to a large empirical literature using micro data to evaluate the effects of education policies on individual outcome (e.g. Figlio and Loeb 2011).

Our paper also speaks to the literature that accounts for variation across countries in income per capita. This literature has focused on the appropriate way to aggregate human capital that varies in the number of years of education (e.g. Mankiw, Romer and Weil 1992, Klenow and Rodriguez-Clare 1997, Caselli 2005). Recent studies, such as Jones (2014) and Malmberg (2017), have improved on this literature by allowing different educational levels to be imperfect substitutes. A related body of work examines how international trade affects skill acquisition, focusing on, again, years of education (e.g. Findlay and Kierzkowski 1983, Atkin 2016, Li 2016, Blanchard and Olney 2017). These papers do not address the issues that years of education do not distinguish between cognitive versus non-cognitive skills, and we complement them by explicitly modeling how individuals optimally choose both the quantities and types of human capital to invest in.

On the other hand, an applied micro literature examines the formation of cognitive and non-cognitive skills using worker-level data (e.g. Kuhn and Weinberger 2005, Cunha, Heckman and Schennach 2010, Jackson, Johnson and Persico 2015). We take a macro perspective by quantifying the different ways in which different countries produce multiple types of human capital, and then clarifying the implications of such differences for aggregate output and for education policies in a general-equilibrium model.

More broadly, the ways countries produce their human capital are related to their educational systems, which often have deep historic roots and so are 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.

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. Section 4 shows how we obtain the values 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 collects robustness exercises, and section 8 concludes.

2 Non-cognitive and Cognitive Occupations, and Other Motivating Data Patterns

A simple way to assess a country's proficiency in human capital production 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 a country's ability to foster these skills will be hard to compare internationally. Moreover, to the extent that a country has a comparative advantage in producing easily measured skills, this country will look productive along this dimension, in part because workers will optimally choose to acquire these skills 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 use the framework of Neal and Johnson (1996) to show below that the wages of leadership occupations are less correlated with test scores than those of the other occupations.

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).[6] Columns

[6]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 Data Appendix).

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.[7] 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.[8] In column 5 we use the O*NET characteristic of enterprising skills as an alternative measure for leadership. The interaction between enterprising skills and AFQT score is negative but not significant. We will present additional robustness results involving other O*NET chracteristics in section 7 below.

Having classified occupations as non-cognitive and cognitive using the U.S. O*NET, we next bring in employment data by 3- or 4-digit occupations 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).[9] This leaves us with a single cross-section of 34 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.

One may wonder whether there are occupations for which human capital is not relevant. The last two columns of Table 1 show that the coefficient estimate of test score remains positive and significant after we control for the college dummy. Kuhn and Weinberger (2005) report that the marginal effect of leadership skills is as strong for low-education individuals as for high-education ones. When we tabulate the distribution of average schooling years across occupations using 2000 US Census, we find that

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

[8]Note that we have included the non-cognitive dummy itself, plus the college dummy and its interaction with AFQT score, as controls.

[9]The statistical agencies of Australia and New Zealand provided us with mappings from their (2-digit) national occupation codes to ISCO-88, and we include them in our sample as well. We obtain very similar results when we drop them.

this distribution is compressed in the left; e.g. the median is 12.8 years while the 5th percentile is 10.9 years. These results are intuitive, because the U.S. is a high-income country, where illiteracy, subsistence farming and the informal sector are not salient features of the economy. We therefore use the 28 high-income countries as our main sample, and examine the extended sample that also includes middle-income countries in section 7. These high-income countries account for 42.94% of world GDP in 2000. Table 2 provides summary statistics of the employment shares of cognitive and non-cognitive occupations, and Table 3 lists the countries and years in our sample. We discuss the other variables in Table 2 in section 4 below.

3 A Model of Human Capital Production with Heterogeneous Workers

In this section we develop our model for the production of human capital and illustrate the intuition of our key parameters. A key feature of our model is that heterogeneous workers optimally choose their investment in both the quantities and types of human capital. We also show how the model can make contact with observable country-level variables with an eye toward quantification, in preparation for section 4.

3.1 Model Structure

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

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

In the context of the correlation patterns that we discussed in section 2, we think about the attributes n and c as two distinct packages of skills, rather than two individual skills. These two packages may have common elements. In equation (1), the parameter ρ captures the degree to which non-cognitive and cognitive packages are correlated.[10] The parameter θ captures the dispersion of attributes across workers. As θ rises, the

[10] When $\rho = 0$, they are independent; when $\rho > 0$, they have positive correlation; and when $\rho \to 1$, they become perfectly collinear.

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 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.[11]

To minimize the number of moving parts, we follow Hsieh et al. (2016) and specify the following human-capital production function. 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 human capital accumulation, 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 choice of e. We assume that η is common across countries. The parameters h_n^k and h_c^k are country k's TFP's in the production functions of non-cognitive and cognitive human capital, and they capture country k's human capital productivities along these two dimensions, net of resources inputs.

We treat h_n^k and h_c^k as exogenous, because the educational institution, an important contributor to human capital production, 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.[12] In S. Korea, and many other East Asian countries, the national exam has been a cornerstone of the educational institu-

[11] The assumption over ρ and θ is standard in the general-equilibrium literature using the Frechet distribution. The assumption that the Ts are same is that there are no inherent genetic differences across countries. It is reasonable to imagine that in countries in which severe malnutrition could change Ts but it is not clear how. As we will consider a sample of primarily middle and upper income countries, this issue is less of a concern.

[12] 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).

tion for over 1,000 years.[13] We capture, and quantify, such 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 quantity of the worker's human capital, accumulated according to the technology (2), and ε_i her attribute, drawn from the distribution in (1).[14] The educational and occupational choices made by workers lead to aggregate supplies of cognitive and non-cognitive humam capital in country k of L_c^{kS} and L_n^{kS} (hence the S superscript), respectively.

The representative firm hires workers in both cognitive and non-cognitive occupations to maximize output

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

In equation (4), Θ^k is country k's output 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^{kD} and L_c^{kD} are the aggregate levels of non-cognitive and cognitive human capital demanded (hence the D superscript) by final goods producers in country k.

To introduce trade into our model, we assume that the individuals can sell the services of their human capital as intermediate inputs around the globe. The services of cognitive and non-cognitive labor are embodied in traded intermediates. Countries can costlessly export intermediates to an international clearinghouse for factor content and

[13]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).

[14]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 produces similar expressions for the aggregate quantities of human capital in country k. Its empirical implementation, however, is difficult, because it is unclear how to identify the shares of cognitive and non-cognitive skills by occupation in the data.

import intermediates at iceberg trade cost τ^k from this clearinghouse. In order to relate the aggregate quantities of cognitive and non-cognitive human capital to trade and occupation employment shares, we define net exports as quantity ratios

$$x_i^k = \frac{L_i^{kS} - L_i^{kD}}{L_i^{kS}}, i = c, n. \tag{5}$$

For example, if $x_c^k = -0.5\%$, country k imports cognitive human capital, the quantity of which is 0.5% of its aggregate supply.On the other hand, we assume that the final good itself is non-tradeable, because it is used as inputs in the production of human capital.

The key prices in country k are the price of an effective unit of cognitive human capital w_c^k, the price of an effective unit of non-cognitive human capital w_n^k, and the price of the final output, P^k. Given cost minimization of the perfectly competitive final goods producers, the price of the final good (4) is given by

$$P^k = \frac{1}{\Theta^k} \left((A_c)^\alpha \left(w_c^k\right)^{1-\alpha} + (A_n)^\alpha \left(w_n^k\right)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}. \tag{6}$$

Equation (6) says that P^k varies across countries for two reasons. First, w_c^k and w_n^k may vary across countries, because of trade costs. In addition, even if w_c^k and w_n^k have no cross-country variation (e.g. under free trade), the final good remains non-tradeable and output TFP varies across countries.

All markets are perfectly competitive. The timing happens as follows. First, workers choose how much and what type (cognitive or non-cognitive) of human capital to obtain. Second, final goods producers choose how many workers of each type to employ and how much of each type of intermediate input to import or export. Finally, all markets clear.

3.2 Equilibrium Conditions

We first analyze individual workers' optimal choices for the quantity and type of human capital accumulation. We then aggregate across individuals to obtain the total quantities of cognitive and non-cognitive human capital in country k, which characterizes the supply side of the economy. We then bring in the demand side and characterize the equilibrium.

Recall that human capital investment is in terms of final output. This means that the proper maximization problem facing an individual 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 human capital investment, after substituting for the price index and accounting for the normalization, is then

$$e(\epsilon_i) = \left(\eta \frac{w_i^k}{P^k} h_i^k \epsilon_i\right)^{\frac{1}{1-\eta}}. \tag{7}$$

In equation (7), $e(\epsilon_i)$ is the quantity of human capital investment. Equation (7) says that, intuitively, gifted individuals, of both occupations, make large quantities of human capital investment. In addition, individuals in country k also have large quantities of human capital investment if real wages are high.

We now plug the worker' optimal choice in (7) into her maximization problem, and obtain the following expression for her optimal net income in occupation i,

$$I_i(\varepsilon_i) = (1-\eta)\eta^{\frac{\eta}{1-\eta}} \left(\frac{w_i^k}{P^k} h_i^k \varepsilon_i^k\right)^{\frac{1}{1-\eta}}. \tag{8}$$

Equation (7) and (8) show that net income, $I_i(\varepsilon_i)$, is proportional to human-capital spending, $e_i(\varepsilon_i)$. In addition, (7) and (8) show that the final-good price index, P^k, has the same effects on $e(\epsilon_i)$ and $I_i(\varepsilon_i)$ for both occupations. This means that P^k does not affect individuals' occupational choices, which are based on their comparative advantages. To be specific, equation (8) 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*

$$p_i^k = \frac{T_i(w_i^k h_i^k)^\theta}{T_c(w_c^k h_c^k)^\theta + T_n(w_n^k h_n^k)^\theta},\ i = c, n. \tag{9}$$

Equation (9) 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 has a strong comparative advantage in fostering non-cognitive human capital (high h_n^k/h_c^k). In (9), θ plays an important role. As θ rises and workers become more homogeneous, given changes in w_i^k or h_i^k lead to bigger shifts in the proportion of workers that opt to work in different occupations. Equation (9) characterizes individuals' optimal choices for the types of human capital accumulated, and plays a key role in our model.

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 (8), 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.*

$$I_n^k = I_c^k = \gamma(1-\eta)\eta^{\frac{\eta}{1-\eta}} \left[T_c \left(\frac{w_c^k}{P^k} h_c^k \right)^{\theta} + T_n \left(\frac{w_n^k}{P^k} h_n^k \right)^{\theta} \right]^{\frac{1}{\theta(1-\eta)}}, \quad (10)$$

$$\text{where } \gamma = \Gamma\left(1 - \frac{1}{\theta(1-\rho)(1-\eta)}\right).$$

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 (10), the term in the square brackets is proportional to the denominator of the employment-share expression, (9). $\Gamma(.)$ is the Gamma function and so γ is a constant.

Equations (9) and (10) imply that:

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

$$E_n^k = E_c^k = \gamma \left[\eta \left(T_c \left(\frac{w_c^k}{P^k} h_c^k \right)^{\theta} + T_n \left(\frac{w_n^k}{P^k} h_n^k \right)^{\theta} \right)^{\frac{1}{\theta}} \right]^{\frac{1}{1-\eta}}. \quad (11)$$

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 η, the elasticity of the output of human capital with respect to input, as we show in section 4.

We now solve for the aggregate supply of human capital of type i, L_i^{kS}, $i = n, c$. We show in the Appendix that

Proposition 3 *Given occupational and educational choices of workers, the aggregate supply of locally provided human capital of type i in country k is*

$$L_i^{kS} = L^k p_i^k E(h_i^k e^{\eta} | Occp.i) = \frac{L^k p_i^k}{w_i^k} \left(\eta^{\eta} (P^k)^{1-\eta} \left(T_c \left(\frac{w_c^k}{P^k} h_c^k \right)^{\theta} + T_n \left(\frac{w_n^k}{P^k} h_n^k \right)^{\theta} \right)^{\frac{1}{\theta}} \right)^{\frac{1}{1-\eta}} \gamma. \quad (12)$$

Equation (12), together with $w_c^k L_c^k + w_n^k L_n^k = P^k Y^k$, implies that the income shares of cognitive and non-cognitive workers are given by

$$\frac{w_i^k L_i^{kS}}{P^k Y^k} = p_i^k. \tag{13}$$

To complete our characterization of the supply side of the economy, we use equations (9) and (12) to derive the relative supply of non-cognitive human capital, which is given by

$$\frac{w_n^k}{w_c^k}\frac{L_n^{kS}}{L_c^{kS}} = \frac{p_n^k}{p_c^k} \text{ where } \frac{p_n^k}{p_c^k} = \frac{T_n (w_n^k h_n^k)^\theta}{T_c (w_c^k h_c^k)^\theta} \tag{14}$$

Equation (14) says that the relative supply of non-cognitive labor, L_n^{kS}/L_c^{kS}, is increasing in the availability of raw talent in the country, the comparative advantage of that country in non-cognitive human capital, h_n^k/h_c^k, and the relative return of non-cognitive human capital, w_n^k/w_c^k. As foreshadowed by our discussion of Proposition 1, it is clear from equation (14) that θ is the supply elasticity: as workers' skills become more homogeneous, a given change in h_n^k/h_c^k or w_n^k/w_c^k affects the occupational choices of more workers, and so solicits a larger response in L_n^{kS}/L_c^{kS}.

We now turn our attention to the demand side. Cost minimization by final goods producers facing technology (4) determines the demand for cognitive and non-cognitive human capital, implying that the cost share of input $i = c, n$ is given by

$$s_i^k = \frac{(A_i)^\alpha \left(w_i^k\right)^{1-\alpha}}{(A_c)^\alpha \left(w_c^k\right)^{1-\alpha} + (A_n)^\alpha \left(w_n^k\right)^{1-\alpha}}. \tag{15}$$

It follows immediately that the relative demand for non-cognitive human capital is given by

$$\frac{w_n^k}{w_c^k}\frac{L_n^{kD}}{L_c^{kD}} = \frac{s_n^k}{s_c^k} \text{ where } \frac{s_n^k}{s_c^k} = \frac{(A_n)^\alpha \left(w_n^k\right)^{1-\alpha}}{(A_c)^\alpha \left(w_c^k\right)^{1-\alpha}} \tag{16}$$

Equation (16) is a standard relative demand equation where the demand elasticity is given by α.

With relative supply, (14), and relative demand, (16), in hand, we can combine them with international trade, (5), to characterize factor market clearing in country k

$$\frac{L_n^{kD}}{L_c^{kD}} = \frac{L_n^{kS}}{L_c^{kS}}\frac{1 - x_n^k}{1 - x_c^k}. \tag{17}$$

Finally, international equilibrium requires that countries' exports of cognitive human capital must be equal to other countries' imports of cognitive human capital. Defining

M_c as the set of countries that import cognitive labor (i.e. $x_c^k < 0$) and X_c as the set of countries that export cognitive labor (i.e. $x_c^k > 0$). International factor market clearing requires that

$$\sum_{k \in X_c} x_c^k L_c^{kS} + \sum_{k \in M_c} x_c^k L_c^{kS} \tau^k = 0, \tag{18}$$

where L_c^{kS} must satisfy (12). Let w_c and w_n denote the prices of cognitive and non-cognitive human capital on the international factor market clearinghouse. Then factor prices in country k are given by

$$\begin{aligned} w_c^k &= w_c,\ w_n^k = w_n \tau^k \text{ if } k \in X_c, \\ w_c^k &= w_c \tau^k,\ w_n^k = w_n \text{ if } k \in X_n. \end{aligned} \tag{19}$$

We are now in a position to define the equilibrium of our model.

Definition 1 *An equilibrium to our model is a set of international factor prices w_c and w_n that imply local factor prices via (19) and that imply quantities of factors supplied locally, given by (14), and factors demanded, given by (16). These quantities clear domestic factor markets, given by (17), and the associated factor trades clear the international market for cognitive human capital, given by (18) in conjunction with (12).*

In the next sub-sections we show how to use the equilibrium conditions of our model to first obtain the values of cognitive and non-cognitive productivities, h_c^k and h_n^k, and then draw out their implications for output per worker.

3.3 Cognitive and Non-cognitive Productivities: Meaurement

In this sub-section, we lay down preparatory work for quantification in section 4 below, by showing how our model can make contact with observables in the data. We begin by backing out a country's comparative advantage in human capital production from data. Rearranging equation (9) we can solve for a country's comparative advantage in human capital production:

$$\frac{h_c^k}{h_n^k} = \left(\frac{p_c^k / T_c}{p_n^k / T_n} \right)^{\frac{1}{\theta}} \frac{w_n^k}{w_c^k}.$$

Intuitively, if a large fraction of the population is employed in cognitive occupations despite a low relative wage in that occupation, it must be that accumulating cognitive human capital is relatively easy. We can solve for the relative wage using (14), (16) and (17), plug the solution into the expression above, and then divide country k's expression relative to a base country 0. These steps yield

$$\frac{h_c^k/h_n^k}{h_c^0/h_n^0} = \left(\frac{p_c^k/p_n^k}{p_c^0/p_n^0}\right)^{\frac{1}{\theta}+\frac{1}{\alpha-1}} \left(\frac{1-x_c^k}{1-x_n^k}\right)^{\frac{1}{\alpha-1}}. \tag{20}$$

To see the intuition of equation (20), we start with autarky, when it simplifies to

$$\frac{p_c^k/p_n^k}{p_c^0/p_n^0} = \left(\frac{h_c^k/h_n^k}{h_c^0/h_n^0}\right)^{\frac{\theta(\alpha-1)}{\phi}}, \tag{21}$$

where $\phi \equiv \theta + \alpha - 1 > 0$. Equation (21) shows the importance of the key demand-side elasticity, α. Suppose h_c^k/h_n^k increases; i.e. country k has a stronger comparative advantage in producing cognitive human capital. By equation (9), this has a direct effect on the relative employment share of cognitive occupations, p_c^k/p_n^k, as well as an indirect effect on it, through the movement of the relative return to cognitive human capital, w_c^k/w_n^k. Equation (21) says that the net effect depends on α. If $\alpha > 1$, demand is elastic, and so the movement in the relative return is small, and the direct effect dominates. Therefore, a larger relative employment share of cognitive occupations reflects a greater comparative advantage for cognitive human capital. When $\alpha \leq 1$, however, this result does not hold. We show, in section 4 below, that data indicates $\alpha > 1$.

We now go back to equation (20). It has the flavor of revealed comparative advantage: we can back out a country's comparative advantage for cognitive human capital, the left-hand side of (20), using the data and parameter values on the right-hand side of (20). The first term there captures the effects of the endogenous choices of workers and the optimal hiring decisions of the final goods producers. If we observe, in the data, that many have chosen the cognitive occupation in country k, we can infer that country k has a strong comparative advantage for cognitive human capital. The second term on the right-hand side of (20) captures the effects of international trade. If we observe, in the data, that country k imports, in the net, the service of cognitive human capital (i.e. $x_c^k < 0$), we can infer that country k has a stronger comparative advantage for cognitive human capital than its employment shares suggest, because the cognitive workers in k have chosen their occupation despite import competition.

We now turn to backing out a country's absolute advantage in producing human capital from data. To do this, we assume that country k's average score on international exams, such as PISA, is informative about its aggregate supply of cognitive human capital. Country k's cognitive productivity can then be obtained by appropriately adjusting this test score for educational expenditures and occupational choices. Specifically, for some positive constant b, we assume that

$$S^k = b\frac{L_c^{kS}}{L^k}, \tag{22}$$

where L_c^{kS} is given by (12). We then combine this assumption with equations (9), (11), (12) and (13), to obtain the following expression for country k's test score relative to a base country:

$$\frac{S^k}{S^0} = \left(\frac{E^k}{E^0}\right)^\eta \left(\frac{p_c^k}{p_c^0}\right)^{1-\frac{1}{\theta}} \left(\frac{h_c^k}{h_c^0}\right) \tag{23}$$

As expression (23) 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 (23) captures the effects of incentives and selection, and they arise in general equilibrium because heterogeneous individuals make optimal choices for human capital investment. To see these effects, suppose that many choose the cognitive occupation in country k; i.e. p_c^k is high. This means that the cognitive occupation is an attractive career choice, 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, h_c^k is country k's TFP in producing cognitive human capital. An important implication of equation (23) is that, in order to isolate h_c^k, one has to adjust test scores for the convoluting factors of educational expenditures and occupational employment shares.

3.4 Output per Worker, Educational Quality Index and Output TFP

In this sub-section we clarify the connections between h_c^k and h_n^k and income differences across countries. We derive an analytical expression that decomposes the differences in output per worker into a component that reflects differences in human-capital productivities, and another that reflects differences in output TFP. To do so, we first define a base country 0 against which any particular country can be compared. We show, in the Appendix, that output per worker in country k relative to the base country 0 is

$$\frac{Y^k/L^k}{Y^0/L^0} = \left[\frac{\Theta^k}{\Theta^0}\Omega^k\right]^{\frac{1}{1-\eta}}, \tag{24}$$

where

$$\Omega^k = \left(p_c^0\left(\left(\frac{p_c^0(1-x_c^0)}{p_c^k(1-x_c^k)}\right)^{\frac{1}{\alpha-1}}\frac{h_c^k}{h_c^0}\right)^{\theta} + p_n^0\left(\left(\frac{p_n^0(1-x_n^0)}{p_n^k(1-x_n^k)}\right)^{\frac{1}{\alpha-1}}\frac{h_n^k}{h_n^0}\right)^{\theta}\right)^{\frac{1}{\theta}}. \tag{25}$$

Equation (24) shows how income per capita across countries can be decomposed into a component that is due to an e*ducational quality index*, $\left(\Omega^k\right)^{\frac{1}{1-\eta}}$, and to output TFP, $\left[\Theta^k/\Theta^0\right]^{1/(1-\eta)}$. Equation (25) says that the educational quality index, Ω^k, is a weighted power mean of the ratios of cognitive and non-cognitive productivities, with the weights dependent on the occupational employment shares and the shares of the factor services that are traded. This index summarizes the multi-dimensional differences in cognitive and non-cognitive productivities into a single numerical value, and it captures the contribution of the overall quality of the educational institution to output per capita. The output TFP measure, akin to a residual, $\left[\Theta^k/\Theta^0\right]^{1/(1-\eta)}$, could vary across countries due to things like efficiency of court systems and business regulations. Both Ω^k and

Θ^k/Θ^0 are amplified by the power $1/(1-\eta)$, because higher output per worker lowers the relative price of the final good and makes it easier to produce both types of human capital.

To provide additional intuition for the educational quality index, Ω^k, we consider the special cases of autarky and free trade. In these special cases, equation (25) simplifies substantially, while the decomposition (24) remains unchanged.

Closed Economy In this case, $x_c^k = x_n^k = 0$, so that we can use equations (9) and (14)-(18) to solve for relative wages:

$$\frac{w_n^k}{w_c^k} = \left[\frac{T_c}{T_n} \left(\frac{h_c^k}{h_n^k} \right)^{\theta} \left(\frac{A_n}{A_c} \right)^{\alpha} \right]^{\frac{1}{\phi}}, \tag{26}$$

where $\phi \equiv \theta + \alpha - 1 > 0$. Given this information, the educational quality index simplifies (see the Appendix for the details) to

$$\Omega^k \equiv \left(p_c^0 \left(\frac{h_c^k}{h_c^0} \right)^{\frac{\theta(\alpha-1)}{\phi}} + p_n^0 \left(\frac{h_n^k}{h_n^0} \right)^{\frac{\theta(\alpha-1)}{\phi}} \right)^{\frac{\phi}{\theta(\alpha-1)}}. \tag{27}$$

Because the powers in Ω^k are determined by the demand and supply elasticities, θ and α, they play important roles in determining how overall education quality, Ω^k, relates to cognitive and non-cognitive productivities, h_c^k and h_n^k. As both θ, $\alpha \to \infty$, Ω^k goes to the maximum of h_c^k and h_n^k. This is intuitive, as workers become equally capable at both perfectly substitutable tasks. In this case, being strong in producing one type of human capital but weak in producing the other type has few consequences for a country's well-being. As $\alpha \to -\infty$, however, the aggregate production function becomes Leontief, Ω^k goes to the minimum of h_c^k and h_n^k, and excelling along a single dimension in human-capital production 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, where the relative importance of cognitive and non-cognitive productivities is determined by the occupational shares. In this case, both cognitive and non-cognitive productivities are important, and so a country with high productivity along one dimension but low productivity along the other tends to have low overall education quality.[15] We will further illustrate this intuition in section 5, by drawing a curve along which Ω^k stays constant. We call this the iso-education-quality curve.

Free Trade In the case of free trade, there is a single global price per unit of cognitive (w_c) and non-cognitive (w_n) human capital. Given effective factor price equalization, it immediately follows that a common numeraire for all countries can be defined. It is useful to define this numeraire as a bundle of inputs into final-good production, i.e. $\left((A_c)^{\alpha}(w_c)^{1-\alpha} + (A_n)^{\alpha}(w_n)^{1-\alpha}\right)^{\frac{1}{1-\alpha}} = 1$. Then the price of final output in country k is given by $P^k = \left(\Theta^k\right)^{-1}$.

As a result, the relative demand for non-cognitive labor is the same across countries and the relative supply depends only on h_n^k/h_c^k, by equations (14) and (16). This means that the educational quality index simplifies to

$$\Omega^k = \left(p_c^0\left(\frac{h_c^k}{h_c^0}\right)^{\theta} + p_n^0\left(\frac{h_n^k}{h_n^0}\right)^{\theta}\right)^{\frac{1}{\theta}}. \tag{28}$$

Comparing the educational quality index with that of the closed economy, given by equation (27), we see that the key difference is in the power coefficients in the construction of the power mean of cognitive and non-cognitive productivities. Critically, the power coefficients under free trade 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 relative to the closed economy case; i.e. as if $\alpha \to \infty$ in equation (27). As a result, being relatively inefficient at producing one type of human capital is less of a drag on output per worker. Intuitively, in a world of free trade, imbalance in human capital productivities helps countries specialize and is a source of welfare gains. We will further illustrate this intuition in section 6, by showing how the iso-education-quality curve changes its shape in the movement from closed economy to free trade.

[15] Note that as α becomes smaller (i.e. cognitive and non-cognitive human capital become less substitutable in aggregate production), equation (27) assigns a bigger penalty for given imbalances in cognitive and non-cognitive productivities.

3.5 The Gains from Trade

In the previous sub-section, we have used decomposition (24) for cross-country comparisons, for one given equilibrium. In this sub-section we show how to use it to do comparative statics. We focus on changes in trade costs, since the results we derive will be useful for our counterfactuals in section 6 below.

Specifically, we follow Deckle, Eaton, and Kortum (2008), or DEK 2008, and let $\widehat{z}$ denote the value of variable z at the subsequent equilibrium relative to its value at the initial equilibrium; i.e. $\widehat{z} = z'/z$. We then re-interpret the base country 0 in equation (24) to represent country k's initial equilibrium. Assuming that there is no change in L^k, Θ^k, h_c^k or h_n^k, we use equations (24) and (13) to obtain

$$\widehat{Y}^k = \left[\left(p_c^k \left(\frac{\widehat{w_c^k}}{P^k} \right)^{\theta} + p_n^k \left(\frac{\widehat{w_n^k}}{P^k} \right)^{\theta} \right)^{\frac{1}{\theta}} \right]^{\frac{1}{1-\eta}}, \tag{29}$$

where

$$\widehat{P}^k = \left(s_c^k \left(\widehat{w_c^k} \right)^{\theta} + s_n^k \left(\widehat{w_n^k} \right)^{\theta} \right)^{\frac{1}{1-\alpha}}.$$

Equation (29) says that the change in output depends on the weighted power mean of the changes in cognitive and non-cognitive workers' real earnings. This weighted mean is amplified by the power $1/(1-\eta)$ because changes in real earnings affect workers' human capital investment, by equation (7).

Equation (29) applies to any change in trade costs, and we now make it more specific, with an eye towards our counterfactuals in section 6. Suppose we know the values of the variables of the initial equilibrium, which can be data or a free-trade equilibrium we compute (in section 6). Suppose the subsequent equilibrium is autarky. We can then simplify equation (29) by repeatedly exploiting the internal factor market relationship given by equation (17), to obtain:

$$\widehat{Y}^k = \left(p_c^k (1 - x_c^k)^{\frac{\theta}{\theta+\alpha-1}} + p_n^k \left(1 - x_n^k\right)^{\frac{\theta}{\theta+\alpha-1}}\right)^{\frac{\alpha-1+\theta}{(\alpha-1)\theta} \frac{1}{1-\eta}}. \tag{30}$$

Equation (30) computes the output loss in the hypothetical movement from the initial equilibrium, with trade, back to autarky, and it uses the employment shares and trade values at the initial equilibrium, p_c^k, p_n^k, x_n^k and x_c^k, plus parameter values, all of which are observables. In this sense, it is the counterpart of the ACR-2012 formula for our model. Like ACR 2012, equation (30) says that gains from trade (which is $1/\widehat{Y}^k$) are large if there is a lot of trade. Relative to ACR 2012, equation (30) extends the results beyond one single input, and says that the cognitive and non-cognitive contents of trade should be weighed, respectively, by cognitive and non-cognitive employment shares. We show in the Appendix that a country must gain from trade as long as it is a net importer of at least one type of factor service (i.e. $\widehat{Y}^k < 1$ if $x_i^k < 1$ for at least one i).

4 Values of Structural Parameters

Table 2 reports the summary statistics of our data. We already discussed the occupation employment shares, p_c^k and p_n^k, in section 2. For test score, S^k, we obtain mean PISA scores in reading, math and science from the official PISA website.[16] [17] For aggregate output, Y^k, we use labor income, or compensation of employees from NIPA (National Income and Product Account), since we do not have physical capital in our model.[18] For the trade flows, x_c^k and x_n^k, we calculate the numbers of cognitive and non-cognitive workers embedded in the net export flows relative to the numbers of these workers in country k's labor force.[19] Table 2 shows that the absolute values of these trade flows are

[16] We use PISA scores because they are widely reported in the media, and have influenced education policies in many countries. In addition, PISA samples students in a nationally representative way, covers many countries, and controls qualities of the final data (e.g. the 2000 UK scores and 2006 US reading scores are dropped because of quality issues). Finally, while PISA scores are for high-school students, they are highly correlated with the scores of adult tests (e.g. Hanushek and Zhang 2009). Compared with PISA, adult tests cover substantially fewer countries (they would cut our sample size by at least 25%) and also have lower response rates (e.g. Brown et al. 2007). See also the Data Appendix.

[17] When PISA first started in 2000, only the reading test was administered, and only a small set of countries participated (e.g. the 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 Data Appendix we show that PISA scores have limited over-time variation.

[18] We experimented with stripping capital from GDP, obtained from PWT (Penn World Tables), 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.

[19] In computing the trade flows we follow similar steps as Costinot and Rodriguez-Clare (2014). See the Data Appendix for the details.

small, consistent with the findings of the trade literature. Finally, we obtain the ratios of private plus public expenditures on education to GDP in 2004 from the UNESCO Global Education Digest of 2007. Note that all our data come from public sources.

In this section, we use these data and our model to extract the values of the elasticities, θ, η, and α, and final-good TFP, Θ^k, and the TFPs of human capital production, h_n^k and h_c^k.

4.1 Elasticity of Human Capital Production, η

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

Proposition 4 *Country k spends fraction η of its aggregate output on education; i.e.*

$$E^k L^k = \eta Y^k. \tag{31}$$

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 (31), 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 2);[20] i.e. $\eta = 0.1255$.

4.2 Supply Elasticity, θ

We now turn to θ, which measures the dispersion of innate abilities across workers and also governs the elasticity of the aggregate supplies of human capital. Using the results of the previous proposition and equation (23), we obtain

$$\ln \left(\frac{S^k}{(Y^k/L^k)^\eta} \right) = D + \left(1 - \frac{1}{\theta} \right) \ln p_c^k + \ln h_c^k \tag{32}$$

where D is a constant. Equation (32) 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.[21]

[20] The standard deviation of this share is low, at 0.0194. In addition, the mean share remains almost unchanged when we expand our sample (see sub-section 7.3).

[21] Relative to (22), (32) 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.

Equation (32) is also a falsifiable prediction of our model that can be taken to the data. It instructs us to construct novel 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.[22]

Figure 2 clearly illustrates that, consistent with equation (32), the countries in which workers are clustered in cognitive occupations are the countries that score well on tests (normalized by resources inputs), 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 (32). 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.

Table 4 shows the results of fitting our data using (32), 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,[23] and in

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

[23] In the raw data of Australia and New Zealand, the occupation classification codes are their own national codes rather than ISCO 88. See also note 9.

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 (5) of Table 1.

Table 4 produces a range of values for θ, 2.0877 ~3.4965. In comparison, Hsieh et al (2016)'s model also features a Frechet distribution of innate abilities, 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 our range.[24] We use $\theta = 3.4965$ in the rest of the paper and show, in section 7, that we get very similar results if we use other values for θ instead.

Finally, we calculate the residuals and construct cognitive productivities, h_c^k, according to (32). These values are relative, and so we normalize the U.S. value to 1.

4.3 Demand Elasticity, α

For the value of α, the substitution elasticity on the demand side, we combine equations (4), (14), (16), (17) and (22) to obtain

$$\log\left(\frac{Y^k}{S^k L^k}\frac{1}{1-x_c^k}\right) = F + \frac{\alpha}{\alpha-1}\log\left(1+\frac{p_n^k\left(1-x_n^k\right)}{p_c^k\left(1-x_c^k\right)}\right) + \log\Theta^k \tag{33}$$

where the constant F has no cross-country variation.

Equation (33) 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 of country k's residents by equation (22). We adjust it by $\left(1-x_c^k\right)$ to take into account net export of the services of cognitive human capital. The second input is the relative quantity of non-cognitive human capital used in production, and it is a monotonic function of $\frac{p_n^k\left(1-x_n^k\right)}{p_c^k\left(1-x_c^k\right)}$. Therefore,

[24] Note that in equation (32), p_c^k depends on country k's comparative advantage, by equation (9), while h_c^k is country k's absolute advantage. Because it is unclear how absolute and comparative advantages might be correlated, it is also unclear how $\ln p_c^k$ and $\ln h_c^k$ are correlated.

equation (33) 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 (33), then, is similar to the estimation of the aggregate production function. The coefficient of $\log\left(1+\frac{p_n^k(1-x_n^k)}{p_c^k(1-x_c^k)}\right)$ gives us α, and the residuals give us Θ^k, the output TFP. For our model, (33) is another fasifiable prediction that can be taken to the data. It 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.

Table 5 shows the results of fitting our data using (33), 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 2.605 to 2.802. Using 2.802 we infer that $\alpha = 1.5549$. Column (6) uses enterprising skills as the alternative measure for non-cognitive occupations, and the coefficient is positive but not significant, echoing Tables 1 and 4. In column (7), we consider the case of autarky, setting $x_n^k = x_n^k = 0$ in (33). The results are very similar to column (1), implying that $\alpha = 1.4706$. In comparison, 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.[25] We explore alternative values of α in section 7 below.

We then calculate the residuals and construct the output TFP, Θ^k, according to (33), 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 0.4466 (Klenow and Rodriguez-Clare 1997) to 0.6687 (PWT 8.0), and provide an external validation for our approach.[26]

[25] 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.

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

5 Cognitive and Non-cognitive Productivities, and Overall Education Quality

The elasticities estimated in the previous section, combined with the data and equations (20) and (32), imply the full set of values of h_c^k and h_n^k. The h_n^k values are also relative, with the U.S. value set to 1. Table 6 summarizes our parameter values and how we obtain them. In this section we present the values of h_c^k and h_n^k, discuss their policy implications and draw out their economic significance.

5.1 Cognitive Productivity

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 human capital production 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-human-capital production. Equation (32) highlights two sources of noisiness. The first is resources, $(Y^k/L^k)^\eta$. Other things equal, a country with more resources inputs 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 (32) 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 resources for human capital production. Therefore, the qualities of their educational institutions 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 institutions 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 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 institution. 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 institutions 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. They have low non-cognitive productivities for two reasons. First, they have decent, but not stellar, cognitive productivities, as shown in Figure 3. In addition, in these countries, many choose the cognitive occupations, implying that these countries have weak comparative advantages for non-cognitive human capital, by equation (20).

Figure 4 also shows that PISA-math rankings substantially understate the proficiency of the U.S. and U.K. 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.[27] Figure 4 provides quantifications for this argument, showing that the U.S. indeed 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 the former U.K. education minister, 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 educational quality index, and quantify its contribution to output per worker.

5.3 Overall Educational Quality

We compute the educational quality index, Ω^k, using equation (27) and the parameter values obtained under the closed-economy setting, where we set factor content trade to 0. We do so because equation (27) relates Ω^k to exogenous parameters of our model, and closed-economy parameter values are very similar to the ones we have used so far (see sub-sections 4.3 and 7.2).

Figure 5 plots the value of h_c^k against the value of h_n^k, and provides a 2-dimensional illustration of the differences in h_c^k and h_n^k across our sample countries. It also serves as our canvas for the iso-education-quality curve, the combinations of h_c^k and h_n^k that yield a constant level of overall educational quality Ω^k. We draw the iso-education-quality curve through the United States, our benchmark country for which $h_c^k = h_n^k = 1$. This curve 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

[27] 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."

and S. Korea).

The curvature and shape of the iso-education-quality curve are determined by equation (27), and they tell us the trade-offs of increasing cognitive productivity for non-cognitive productivity. Intuitively speaking, given the values of α and θ, equation (27) says that both cognitive and non-cognitive productivities are important in overall education quality. We see, in Figure 5, that the countries with high productivity along one dimension but low productivity along the other (e.g. Germany and Hong Kong) lie below the iso-education-quality curve, meaning that they have lower overall education quality than the U.S. This is because the imbalance of their productivities holds down their overall educational quality. There is, however, a silver lining: this imbalance would be a very useful asset under free trade, as we show in section 6.

We now spell out the economic significance of the differences in h_c^k and h_n^k across countries and the connection between overall education quality and output per capita, by implementing the decomposition of equations (24) and (27) in Table 7. Column 1 shows output per capita by country, relative to the United States (i.e.$\frac{Y^k/L^k}{Y^0/L^0}$). Columns 2 and 3 show, respectively, the contribution of output TFP, $\left[\frac{\Theta^k}{\Theta^0}\right]^{\frac{1}{1-\eta}}$, and of overall educational quality, $\left[\Omega^k\right]^{\frac{1}{1-\eta}}$. To interpret these results, consider Germany. 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).[28] The decomposition in columns 1-3 is exact for every country, even though column 1 is obtained directly from data while columns 2-3 are calculated using our model parameters.

Column 3 shows the large differences in overall educational qualities across countries. 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

[28] In Table 7, $\theta = 3.4965$ and $\alpha = 1.4706$. Table 10 shows that we obtain very similar values and country rankings for overall education quality under alternative values of α and θ.

as danger, in terms of aggregate output.[29]

Finally, columns 4 and 5 report the decomposition based on equations (24) and (25), using the parameter values obtained under the trade-cost setting (e.g. the h_c^k and h_n^k values used in Figures 3 and 4), which accounts for the observed levels of factor content trade. The numbers in columns 2 and 3 are very close to columns 4 and 5, suggesting that the closed economy setting is a very good approximation for the data we observe. This is because the observed level of factor content trade is very small relative to the differences in occupation employment shares across countries.

6 Comparative Statics and Policy Implications

Having demonstrated the economic significance of the differences in human capital productivities, h_c^k and h_n^k, in section 5, we now draw out the policy implications of having multiple types of human capital in our GE model.

6.1 Closed Economy

We start from the closed-economy setting, and derive analytical expressions for how changes in the human capital productivities, h_c^k and h_n^k, affect test scores and aggregate output for any country k. These exercises are easy to implement using our model. First, the base "country 0" in equations (27) can be specified as the initial equilibrium of country k itself. Assuming that there is no change in the labor-force size, L^k, or output TFP, Θ^k, equation (27) simplifies to

$$\widehat{Y^k} = [p_c^k(\widehat{h_c^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}} + p_n^k(\widehat{h_n^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}}]^{\frac{\theta+\alpha-1}{\theta(\alpha-1)}\frac{1}{1-\eta}} \tag{34}$$

Second, equations (21), (33) and (32), together with the identity $p_c^k + p_n^k = 1$, imply that the change in test score is (see the Appendix for the derivation)

$$\widehat{S^k} = (\widehat{h_c^k})^{\frac{1}{1-\eta}} \frac{(\widehat{h_c^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}G}}{[p_c^k(\widehat{h_c^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}} + p_n^k(\widehat{h_n^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}}]^G}, G = \frac{1}{1-\eta}\left(1 - \frac{1}{\theta} - \frac{\eta\alpha}{\alpha-1}\right),$$

[29] The countries' rankings in overall education quality are very similar to their rankings in non-cognitive productivity (correlation is 0.9425). The correlation between overall-education-quality rankings and PISA-math rankings is 0.0909 (p value = 0.6456).

where the constant $G = 0.3676$ according to our parameter values. For small changes, this expression can be approximated as

$$(1-\eta)d\ln S^k = (1 + Bp_c^k)d\ln h_c^k - (Bp_n^k)d\ln h_n^k, B = \frac{(\theta-1)(\alpha-1)-\alpha\eta}{\theta+\alpha-1}, \quad (35)$$

where $B = 0.2496$ according to our parameter values.[30] Equation (35) 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 pushes most people away from choosing the non-cognitive occupation, by equation (9). 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.

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 (35) tells us that U.S. cognitive productivity rises by 2.12%, and equation (34) 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. Suppose S. Korea places less emphasis on test scores and more emphasis on non-cognitive skills in her education system. Assume that these policies increase S. Korea's non-cognitive productivity by 2.5%, and leave her cognitive productivity unchanged. Looking at test score, we might conclude that these policies are unsuccessful, because by equation (35), S. Korea's PISA math score would drop, slightly, by 0.07%. Looking at output, however, we would likely draw the opposite conclusion, because by equation (34), S. Korea's aggregate output would increase by 0.33%. This exercise shows that in our model, test score and output may move in the opposite direction, because there are multiple types of human capital and heterogeneous individuals respond to policies by changing the types

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

of human capital they invest in.

This exercise also illustrates that a reduction in test score could mask an increase in overall education quality. As a result, aggregate output is a better objective for education policies than test score. Indeed, many educational reforms that are promoted to raise test scores (e.g. No Child Left Behind of 2001 for the U.S.) 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. Using equation (34), we compute, for our sample countries, the marginal output changes following a 1% increase in h_c^k or h_n^k,[31] and report these results in columns (6) and (7) of Table 7. For the U.S., for example, the payoff of a 1% increase in h_c^{US} is 0.85% of aggregate output, holding h_n^{US} fixed. However, this payoff shrinks to 0 if h_n^{US} decreases by 2.95% (=1% $\times$ 0.85%/0.29%) instead.

Comparing columns (6) and (7), we see that a 1% increase in h_c^k is associated with larger output changes than a 1% increase in h_n^k. This is because in the data, the employment share of the cognitive occupation is higher than the non-cognitive occupation. This result does *not* imply that improving the quality of cognitive education carries a higher payoff than non-cognitive education, because it is unclear how much resources are needed for the 1% increase in h_c^k or h_n^k.

6.2 Closed Economy to Free Trade

As we know from equations (27) and (28), a country's educational quality index depends on the extent of openness to factor service trade. While the data suggest that the world is closer to autarky than to free trade, it is instructive to see what overall education quality and output per worker would be across countries were international trade frictionless.[32]

To compute the (counterfactual) free-trade equilibrium, we assume that the data

[31] We first set $\widehat{h_c^k} = 1.01$ and $\widehat{h_n^k} = 1$, and then $\widehat{h_c^k} = 1$ and $\widehat{h_n^k} = 1.01$ in (34).

[32] Our motivations for this world-is-flat counterfactual are as follows. First, free trade might be more useful for regional differences within countries, complementing the economic geography literature (e.g. Krugman 1991, Davis and Weinstein 2002, Redding and Sturm 2008, Allen and Arkolakis 2014). Second, internationally, while service trade has been growing faster than goods trade (e.g. wto.org), it has seen less liberalizations than goods trade, and so has more scope for further liberalization. Finally, new technology is rapidly decreasing the cost of service trade; e.g. in the U.S., the employment share of contract-firm workers reaches 14% in 2015 (the Wall Street Journal, A9, Sep. 15, 2017), surpassing the employment share of the manufacturing sector.

we observe (e.g. L^k, Y^k, p_c^k and p_n^k) are well approximated by the closed-economy equilibrium. We also use the parameter values obtained under the closed-economy setting (i.e. the ones we used in sub-sections 5.3 and 6.1). We show, in the Appendix, that under free trade, the international equilibrium condition (18) can be rewritten as

$$\sum_k \frac{H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta}{H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^k}{h_n^0}\right)^\theta} \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^k}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'}\left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(H\left(\frac{h_c^{k'}}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^{k'}}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}} = \frac{(\widetilde{\omega})^{1-\alpha}}{(\widetilde{\omega})^{1-\alpha}+1}, \tag{36}$$

where

$$H=\left(\frac{p_c^0}{p_n^0}\right)^{\frac{\theta+\alpha-1}{\alpha-1}}, \widetilde{\omega}=\left(\frac{A_c}{A_n}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{w_c}{w_n}\right).$$

In equation (36), the superscript "0" denotes the same base country as used in equation (21) (which is the U.S. in our computation). The only unknown variable in (36) is $\widetilde{\omega}$; all the other variables are known, either data or patameters. This means we can solve equation (36) for $\widetilde{\omega}$, recover relative demand from the expression (where the superscript "T" denotes the free-trade equilibrium)

$$s_c^T=\frac{(\widetilde{\omega})^{1-\alpha}}{(\widetilde{\omega})^{1-\alpha}+1}$$

and equation (16), and then recover relative supply using

$$p_c^{kT}=\frac{H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta}{H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^k}{h_n^0}\right)^\theta}$$

and equation (14). Finally, factor service trade can be recovered from (17).

Note that in our computation of the free-trade equilibrium, we are unable to apply the standard DEK 2008 techniques, because by assumption, our data and parameters are for the closed-economy equilibrium, where the trade flows are $x_c^k = x_n^k = 0$. However, this assumption allows us to subsume the unidentified parameters of A_c/A_n and T_c/T_n into $\widetilde{\omega}$, occupation employment shares, and cognitive and non-cognitive productivities (see the Appendix for the details). This means that our computation has the flavor of

DEK 2018, in that we do not need to make additional assumptions about unidentified parameters.

We can then draw on equation (30) to compute the change in output. We report, in column (4) of Table 8, the value $1/\widehat{Y^k}$ in (30), which is the gains in aggregate output in moving from autarky to free trade. To help visualize the patterns of these gains, we plot them against non-cognitive and cognitive productivities in Figure 6. In this 3D plot, the countries in the middle, who have balanced cognitive and non-cognitive productivities, have limited gains from trade. However, gains from trade are large for the countries on the edges of the figure, who have strong comparative advantages in either cognitive or non-cognitive human capital. For example, Hong Kong would see a 17.1% gain in output, S. Korea 44.1%, the Netherlands 18.8%, and Belgium 18.2%.

To explore the intuition for these countries' large gains from trade, we go back to the educational quality index. While imbalance in human capital productivities contributes to low overall education quality under closed economy, by equation (27), it helps countries specialize and so is a source of welfare gains under free trade, by (28). To illustrate the change in overall education quality, we plot, in Figure 7, the iso-educational-quality curve under free trade. Figure 7 has the same values of cognitive and non-cognitive productivities as Figure 5. In addition, the U.S. occupation employment shares under free trade are similar to their values under closed economy. However, the iso-educational-quality curve is generated by equation (28) in Figure 7, vs. (27) in Figure 5. As a result, this curve bends sharply towards the origin in Figure 7, in contrast to Figure 5. We also see, in Figure 7, that several countries that are exceptional along the cognitive dimension in human capital production (e.g. S. Korea and Hong Kong) lie above the iso-education-quality curve, meaning that they would have higher overall education quality than the U.S. under free trade. The results are highly intuitive: countries that have highly unbalanced educational productivities benefit dramatically from being able to specialize in the occupations in which they excel.

We now clarify the connection between changes in overall education quality and output gains from trade using the following decomposition:

$$\widehat{Y}^k = \widehat{Y}^0 (\widehat{\Omega}^k)^{\frac{1}{1-\eta}}. \tag{37}$$

Equation (37) says that country k's gains from trade (which is $1/\widehat{Y}^k$) is equal to the change in its educational quality index (i.e. $1/\widehat{\Omega}^k$) multiplied by the base country's (here the U.S.) gains from trade (i.e. $1/\widehat{Y}^0$). This is intuitive, since the U.S. is our benchmark country for the educational quality index.[33] Column (1) of Table 8 reports the education quality index in closed economy, and is the same as column (5) of Table 7. Column (2) shows the educational quality index under free trade, and column (3) shows the ratios of free-trade values to closed-economy values (i.e. $1/\widehat{\Omega}^k$ in equation (37)). We see that column (3) of Table 8 is very similar to column (4), because U.S. gains from trade are small; i.e. the large gains from trade we saw in Figure 6 are mostly driven by changes in overall education quality.

To be clear, U.S. gains from trade are small because the U.S. has the largest labor-force size and second highest output TFP in our sample, where several large countries, such as Japan and China, are missing. The inclusion of these countries may imply larger gains for the U.S. Since we lack occupation employment data for Japan, we assume that Japan has the same occupation shares and the same cognitive and non-cognitive productivities as S. Korea. We then use Japan's output-per-worker data and equation (27) to calculate Japan's output TFP. Likewise, we assume that China is the same as Hong Kong except for labor-force size and output TFP. We then re-compute the free-trade equilibrium and gains from trade using equations (36) and (30).

We plot these gains from trade against non-cognitive and cognitive productivities in Figure 8. As compared with Figure 6, Figure 8 shows a similar overall pattern: the countries with strong comparative advantages see large gains, while those in the middle see small gains. Figure 8 is also different from Figure 6 in several aspects, due to the addition of Japan and China. First, the U.S. gains are larger, at 8.1%. Second, the countries with strong comparative advantages for cognitive human capital have smaller gains;[34] e.g. 6.6% for Hong Kong and 30.1% for S. Korea. Finally, the countries with strong comparative advantages for non-cognitive human capital have substantially larger gains; e.g. 40.5% for the U.K., 44.0% for Finland, and 55.6% for the Netherlands.

6.3 Free Trade

We now explore the policy implications under the free-trade equilibrium, by computing the marginal changes in aggregate output in response to a 1% increase in h_c^k or h_n^k. To do so, we extend the real output calculations in (29) to allow a country's cognitive and

[33] In our computation for gains from trade, the results obtained using (37) are identical to those obtained using (30).

[34] Japan's gains are the same as S. Korea's, and China's the same as Hong Kong's.

non-cognitive productivities to change. The resulting calculation is given by

$$\widehat{Y}^k = \left[\widehat{h}_c^k \left(\widehat{p}_c^k\right)^{-\frac{1}{\theta}} \left(\widehat{s}_c\right)^{-\frac{1}{\alpha-1}}\right]^{\frac{1}{1-\eta}}. \tag{38}$$

The intuition of equation (38) is as follows. The cost share of cognitive human capital, s_c, depends on the relative return, w_c/w_n, by equation (15). This means that the change in s_c in equation (38) represents the change in relative prices, or country k's terms of trade. On the other hand, the change in h_c^k represents the change in country k's absolute advantage in human capital production, while the change in p_c^k represents the change in k's comparative advantage. Thus these two terms in (38) represent the changes in country k's endowments and PPF (Production Possibility Frontier).

Using equations (36) and (38),[35] we calculate the output changes for our sample countries and report these changes in columns (6) and (7) of Table 8. Columns (8) and (9) show the output changes when we add China and Japan into our free-trade equilibrium. These output changes are quite different from the closed-economy results in Table 7, because comparative advantage plays a major role. For the countries with strong comparative advantages in non-cognitive human capital (e.g. Belgium, Netherlands, U.K. and U.S.), a 1% increase in h_n^k elicits much larger output responses under free trade than under closed economy, because this change reinforces these countries' comparative advantages. In comparison, the 1% increase in h_n^k leads to very limited output responses for the countries with strong comparative advantages in cognitive human capital (e.g. France, Hong Kong and S. Korea), because it erodes their comparative advantages.

7 Robustness Exercises and Discussions

7.1 Values of h_c^k and h_n^k

In section 4, we showed that our values of θ, α and Θ^k are similar to the literature. How about our values of h_c^k and h_n^k? Previous estimates of h_c^k and h_n^k do not exist, to the best of our knowledge, and so we explore whether our h_c^k and h_n^k values are sensible using industry level trade data. Our idea is that the h_c^k and h_n^k values reflect countries' relative abundance in non-cognitive human capital, a source of comparative advantage. It thus

[35] In practice, we obtain identical results whether we: (1) use (36) for the initial equilibrium and then apply (38) for the change; or (2) use (36) for both the initial and subsequent equilibria.

follows that non-cognitive abundant countries should be net exporters of the industries that use non-cognitive human capital intensively.

To take this prediction to the data, we follow the literature (e.g. Nunn 2007, Bombardini, Gallipoli and Pupato 2012) 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. To measure trade patterns, we calculate net export divided by the sum of import and export by industry by country. 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.[36] 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. These results suggest that our h_c^k and h_n^k values are useful for the variation of trade patterns by industry by country, even though we did not use such variation to obtain their values.

7.2 Alternative α and θ Values, and Alternative O*NET Characteristics

We now perform sensitivity analyses and report the results in Table 10. The 2nd and 3rd columns show the values of α and θ, and the rest of the table report the correlation coefficients of h_c^k, h_n^k, Θ^k and Ω^k with our benchmark values. The first row of Table 10 re-caps these benchmark values. The second row shows the values we obtain for the case of closed economy, where we set factor content of trade to zero. The third and fourth rows show how our parameter values change, relative to the benchmark, if we vary the value of θ to 2.0877 or vary both θ and α to 2.0877 and 2. We obtain similar results in

[36]See our Appendix for more details.

these cases.

We now experiment with using the first principal component of the following 8 O*NET characteristics to measure leadership. They include guiding and directing subordinates, which we have used so far, plus leadership in work style; coordinating the work and activities of others; developing and building teams; coaching and developing others; recruiting and promoting employees; monitoring and controlling resources and spending; and coordinate or lead others in work. We obtain similar results, as shown in the second-to-last row of Table 10.[37]

We have also experimented with using the following O*NET characteristics to measure non-cognitive occupations: investigative skills, originality, social skills, and artistic talents. As we show in the Data Appendix, when we use investigative skills, originality or social skills, we find that for the occupations important in these characteristics, wages have stronger correlations with test scores, as compared with the other occupations. These are opposite to our results in Table 1. When we use artistic talents, we find that the artists with high test scores earlier in life tend to have low wages later. However, artists account for less than 1% of the labor force.

7.3 Middle-Income Countries

In this sub-section we extend our sample to also include the middle-income countries that have 3- or 4-digit ISCO-88 occupation data. This increases the number of countries we have from 28 to 34. We focus on the results under the closed-economy setting, to keep our discussions succinct.

The mean value of education expenditure to output becomes 0.1228, very similar to the value of 0.1255 in Table 2. This implies that we obtain very similar value of η for the extended sample. Column (7) of Table 4 shows the implementation of equation (32) using the extended sample. The results are similar to column (2), implying that we obtain similar values for θ and h_c^k. Column (8) of Table 5 show the implementation of equation (33) using the extended sample. The results are similar to column (1), indicating that we obtain similar values for α and Θ^k, and so h_n^k. The last row of Table 10 reports our parameter values for the extended sample, and shows their high correlation coefficients with our benchmark values. To help visualize these similarities, we use these parameter

[37] Table A2 shows the results of the Neal-Johnson regressions for the leadership principal component. These results are similar to Table 1.

values of the extended sample to graph the iso-education-quality curve in Figure 9. We label the names of the middle-income countries in all capital letters. Figure 9 is similar to Figure 5. It also enriches Figure 5, showing that overall education quality varies substantially among the middle-income countries.

7.4 Discussions and Extensions

In this sub-section we discuss several elements that we have abstracted away from in our model. First, the different ways in which countries produce human capital may provide incentives for immigration; e.g. those educated in high relative-cognitive-productivity countries (e.g. S. Korea) may have incentives to migrate to where such relative productivities are low (e.g. the U.S.). This intuition for immigration is similar to trade, which we have in our model. In addition, in our model, the individuals who migrate at young ages are not distinguishable from native-borns, and it is unclear whether data for immigrants show how much human capital they accumulate in their birth countries. If such data is available for future research, it will be interesting to explore the implications of the cross-country differences in cognitive and non-cognitive productivities for the welfare gains of immigration.

In our model, international trade is driven solely by differences in relative productivities across countries, but we still obtain large gains from trade liberalization in our counterfactuals, because individuals switch occupations and also change the quantities and types of their human capital in response to trade shocks. Adding product differentiation and intra-industry trade will produce even larger gains from liberalization. We have abstracted away from these additional elements, because we want to obtain analytical solutions for our model, and also because the effects of these elements on welfare gains have been extensively studied in the literature (e.g. Costinot and Rodriguez-Clare 2014).

Finally, we have taken cognitive and non-cognitive productivities as exogenous parameters. Our motivation is to quantify these parameters and to draw out their implications for output per worker and gains from trade, given that previous estimates of their values do not exist. Could policies affect the values of cognitive and non-cognitive productivities? If so, what policies? How much resources do these policies require? Could there be optimal policies, and how might they vary in closed vs. open economies? We leave these questions for future research.

8 Conclusion

The measurement of human capital accumulation across countries is fraught with difficulties. Merely counting the number of students, years of education, or money spent does not correct for quality differences across countries. International test scores offer a degree of comparability of student outcomes, but only for easily codified, cognitive knowledge. Many are concerned that excessive attention paid to test scores not only results in resources that are wasted "teaching to the test" but that students enter the labor market with poorly developed non-cognitive human capital.

We adapt the Roy (1951) framework to measure cognitive and non-cognitive productivities of high- and middle-income countries and to quantify their contributions to real output per capita. This framework allows us to obtain the values of the full set of structural parameters using publically available data on international test scores, educational spending per capita, occupational choices, and international trade. Our stylized model is analytically tractable, and provides the following novel insights.

We show that hard-to-measure non-cognitive human capital is quantitatively important for overall educational quality. Many countries that perform well on international tests have below-the-average productivities on non-cognitive human capital, and this is large enough to drag down their overall educational quality. In addition, policy reforms that *increase* test score may *reduce* aggregate output, and vice versa. This points to the importance of spelling out the impacts of education policies on aggregate output when we formulate their objectives and conduct cost-benefit analyses. Here, our model provides a useful tool.

While we show that globalization and associated trade in factor services are critical in assessing the quality of a country's educational institutions, the data suggest that the world is closer to autarky than it is to free trade. For the moment at least, educational institutions that focus on one type of human capital to the great detriment of another are the source of substantially lower income per capita; i.e. imbalance is a source of weakness. However, the countries with imbalanced human-capital productivities would reap very large output gains if countries were to engage in free trade of the services of human capital; i.e. under free trade, imbalance would be a source of strength.

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] Allen, Treb, and Costas Arkolakis. "Trade and the Topography of the Spatial Economy." The Quarterly Journal of Economics 129.3 (2014): 1085-1140.

[3] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare. "New trade models, same old gains?" The American Economic Review 102.1 (2012): 94-130.

[4] Atkin, David, 2016. "Endogenous Skill Acquisition and Export Manufacturing in Mexico", American Economic Review 106(8), 2046-2085.

[5] 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).

[6] 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.

[7] Blanchard, Emily and William Olney, 2017. "Globalization and Human Capital Investment: Export Composition Drives Educational Attainment", Journal of International Economics, 106, 165-183.

[8] Bombardini, Matilde, Giovanni Gallipoli, and Germán Pupato. "Skill dispersion and trade flows." American Economic Review 102.5 (2012): 2327-48.

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

[10] Casselli, F. 2005. "Accounting for Cross-Country Income Differences." in Handbook of Economic Growth.

[11] Costinot, Arnaud and Andrés Rodríguez-Clare, 2014. "Trade Theory with Numbers: Quantifying the Consequences of Globalization", in Handbook of International Economics, edts. Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, vol. 4, chapter 4.

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

[13] Davis, Donald R., and David E. Weinstein. "An account of global factor trade." American Economic Review 91.5 (2001): 1423-1453.

[14] Davis, Donald and David Weinstein. "Bones, bombs, and break points: the geography of economic activity." The American Economic Review 92.5 (2002): 1269-1289.

[15] Deckle, Jonathan Eaton, and Samuel Kortum. 2008. "Global Rebalancing with Gravity." *International Monetary Fund Staff Papers* 55: 511-540

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

[17] 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.

[18] Findlay, Ronald and Henryk Kierzkowski, 1983. "International Trade and Human Capital: A Simple General Equilibrium Model", Journal of Political Economy, 91(6), 957-978.

[19] 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.

[20] Hanushek, E., "Education Production Functions", The New Palgrave Dictionary of Economics, 2008.

[21] Hanushek, Eric, and Ludger Woessman. 2011. "How Much do Educational Outcomes Matter in OECD Countries?" *Economic Policy* 26(67): 427-491.

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

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

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

[25] Jackson, Kirabo , Rucker C. Johnson, and Claudia Persico. 2016. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms," *The Quarterly Journal of Economics* 131(1): 157-218.

[26] Jones, Benjamin. 2014. "The Human Capital Stock: A Generalized Approach." *American Economic Review* 104(11): 3752-3777.

[27] Klenow, Peter, and Andres Rodriquez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" NBER Working paper.

[28] Krugman, Paul. "Increasing returns and economic geography." Journal of political economy 99.3 (1991): 483-499.

[29] Kuhn, Peter, and Catherine Weinberger. 2005. "Leadership Skills and Wages." *Journal of Labor Economics* 23(3): 395-436.

[30] Lee, Eunhee, 2017. "Trade, Inequality, and the Endogenous Sorting of Heterogeneous Workers", mimeo, University of Maryland.

[31] Li, Bingjing, 2016. "Export Expansion, Skill Acquisition and Industry Specialization: Evidence from China", mimeo, National University of Singapore.

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

[33] Malmberg, Hannes. 2017. "Human Capital and Development Accounting Revisited." mimeo Institute for International Economic Studies.

[34] 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.

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

[36] Ohnsorge, Franziska and Daniel Trefler, "Sorting it Out: International Trade and Protection with Heterogeneous Workers." Journal of Political Economy 115(5) (2007): 868-892.

[37] Redding, Stephen J., and Daniel M. Sturm. "The costs of remoteness: Evidence from German division and reunification." The American Economic Review 98.5 (2008): 1766-1797.

[38] Roy, Arthur. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3(2): 135-146.

[39] Trefler, Daniel. "The case of the missing trade and other mysteries." The American Economic Review (1995): 1029-1046.

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

9 Theory Appendix 1

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 (9).

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^*

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

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, (8), 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,

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

Therefore, the average net income of non-cognitive workers, I_n, equals $(1-\eta)\eta^{\frac{\eta}{1-\eta}} \gamma [T_c(\frac{w_c}{P} h_c)^\theta + T_n(\frac{w_n}{P} h_n)^\theta]^{\frac{1}{\theta(1-\eta)}}$.This is equation (10).

9.3 Proposition 3

We again drop the superscript k. The average real income of a worker in occupation i is I, so the total real income of workers in occupation i is $Lp_i I$. The real wage of a unit of effective labor of type i is w_i/P and the number of effective units is L_i but we must net out expenditure on education. Hence, we must have

$$\frac{w_i}{P} L_i (1-\eta) = L p_i I.$$

Substituting using (8), we obtain

$$\begin{aligned}
L_i &= \frac{Lp_i}{w_i} P \gamma \eta^{\frac{\eta}{1-\eta}} \left[T_c \left(\frac{w_c^k}{P^k} h_c^k \right)^\theta + T_n \left(\frac{w_n^k}{P^k} h_n^k \right)^\theta \right]^{\frac{1}{\theta(1-\eta)}} \\
&= \frac{Lp_i}{w_i} \gamma \eta^{\frac{\eta}{1-\eta}} \left[P^{1-\eta} \left(T_c \left(\frac{w_c^k}{P^k} h_c^k \right)^\theta + T_n \left(\frac{w_n^k}{P^k} h_n^k \right)^\theta \right)^{\frac{1}{\theta}} \right]^{\frac{1}{1-\eta}}.
\end{aligned}$$

9.4 Equation (23)

Using equations (22), (11) and (12), we can show that

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

We now use equation (9) to obtain that $T_c\left(w_c^k h_c^k\right)^\theta + T_n\left(w_n^k h_n^k\right) = \frac{T_i(w_i^k h_i^k)^\theta}{p_i}$. This expression allows us to substitute out the term $T_c\left(w_c^k h_c^k\right)^\theta + T_n\left(w_n^k h_n^k\right)$ in equation (12), giving us, together with equation (22), that

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

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

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

Taking the ratio of this expression with respect to country 0, we get equation (23).

9.5 Equations (24) and (25)

We start by substituting equation (12) into equation (13), to obtain

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

and so relative to the base country we have

$$\frac{Y^k/L^k}{Y^0/L^0} = \left(\frac{P^0}{P^k}\right)^{\frac{\eta}{1-\eta}} \left[p_c^0 \left(\frac{w_c^k h_c^k}{w_c^0 h_c^0}\right)^{\theta} + p_n^0 \left(\frac{w_n^k h_n^k}{w_n^0 h_n^0}\right)^{\theta}\right]^{\frac{1}{\theta(1-\eta)}}$$

rearranging

$$\frac{Y^k/L^k}{Y^0/L^0} = \left(\frac{w_c^k/P^k}{w_c^0/P^0}\right)^{\frac{1}{1-\eta}} \left[p_c^0 \left(\frac{h_c^k}{h_c^0}\right)^{\theta} + p_n^0 \left(\frac{w_c^0}{w_c^k}\frac{w_n^k h_n^k}{w_n^0 h_n^0}\right)^{\theta}\right]^{\frac{1}{\theta(1-\eta)}} \tag{39}$$

Let $nx_i^k = p_i^k x_i^k$, $i = n, c$. We now use equations (14), (16) and (17) to show that

$$\frac{w_c^k}{w_n^k} = \left(\frac{A_c}{A_n}\right)^{\frac{\alpha}{\alpha-1}} \left(\frac{p_n^k - nx_n^k}{p_c^k - nx_c^k}\right)^{\frac{1}{\alpha-1}} \tag{40}$$

From the price index we have

$$\begin{aligned} P^k &= \frac{1}{\Theta^k}\left((A_c)^{\alpha}(w_c^k)^{1-\alpha} + (A_n)^{\alpha}(w_n^k)^{1-\alpha}\right)^{\frac{1}{1-\alpha}} \\ &= \frac{w_c^k}{\Theta^k}\left((A_c)^{\alpha} + (A_n)^{\alpha}\left(\frac{w_c^k}{w_n^k}\right)^{\alpha-1}\right)^{\frac{1}{1-\alpha}} \end{aligned} \tag{41}$$

Combine this expression with equation (40), we have

$$\begin{aligned} \frac{w_c^k}{P^k} &= \Theta^k \left((A_c)^{\alpha} + (A_n)^{\alpha}\left(\left(\frac{A_c}{A_n}\right)^{\frac{\alpha}{\alpha-1}}\left(\frac{p_n^k - nx_n^k}{p_c^k - nx_c^k}\right)^{\frac{1}{\alpha-1}}\right)^{\alpha-1}\right)^{\frac{1}{\alpha-1}} \\ &= \Theta^k \left((A_c)^{\alpha} + (A_c)^{\alpha}\frac{p_n^k - nx_n^k}{p_c^k - nx_c^k}\right)^{\frac{1}{\alpha-1}} \\ &= \Theta^k (A_c)^{\frac{\alpha}{\alpha-1}}\left(p_c^k - nx_c^k\right)^{-\frac{1}{\alpha-1}} \end{aligned}$$

where the last equality uses $p_n^k - nx_n^k + p_c^k - nx_c^k = 1$, which is implied by equation (9) and balance of trade. so relative real wage is

$$\frac{w_c^k/P^k}{w_c^0/P^0} = \frac{\Theta^k}{\Theta^0}\left(\frac{p_c^k - nx_c^k}{p_c^0 - nx_c^0}\right)^{-\frac{1}{\alpha-1}} \tag{42}$$

We substitute equations (40) and (42) into (39):

$$\begin{aligned}
\frac{Y^k/L^k}{Y^0/L^0} &= \left(\frac{\Theta^k}{\Theta^0}\left(\frac{p_c^k - nx_c^k}{p_c^0 - nx_c^0}\right)^{-\frac{1}{\alpha-1}}\right)^{\frac{1}{1-\eta}} \left[p_c^0\left(\frac{h_c^k}{h_c^0}\right)^\theta + p_n^0\left(\frac{\left(\frac{p_n^0-nx_n^0}{p_c^0-nx_c^0}\right)^{\frac{1}{\alpha-1}}}{\left(\frac{p_n^k-nx_n^k}{p_c^k-nx_c^k}\right)^{\frac{1}{\alpha-1}}}\frac{h_n^k}{h_n^0}\right)^\theta\right]^{\frac{1}{\theta(1-\eta)}} \\
&= \left[\frac{\Theta^k}{\Theta^0}\left(\frac{p_c^k - nx_c^k}{p_c^0 - nx_c^0}\right)^{-\frac{1}{\alpha-1}}\left(p_c^0\left(\frac{h_c^k}{h_c^0}\right)^\theta + p_n^0\left(\frac{\left(\frac{p_n^0-nx_n^0}{p_c^0-nx_c^0}\right)^{\frac{1}{\alpha-1}}}{\left(\frac{p_n^k-nx_n^k}{p_c^k-nx_c^k}\right)^{\frac{1}{\alpha-1}}}\frac{h_n^k}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}}\right]^{\frac{1}{1-\eta}} \\
&= \left[\frac{\Theta^k}{\Theta^0}\left(p_c^0\left(\left(\frac{p_c^k - nx_c^k}{p_c^0 - nx_c^0}\right)^{-\frac{1}{\alpha-1}}\frac{h_c^k}{h_c^0}\right)^\theta + p_n^0\left(\left(\frac{p_c^k - nx_c^k}{p_c^0 - nx_c^0}\right)^{-\frac{1}{\alpha-1}}\frac{\left(\frac{p_n^0-nx_n^0}{p_c^0-nx_c^0}\right)^{\frac{1}{\alpha-1}}}{\left(\frac{p_n^k-nx_n^k}{p_c^k-nx_c^k}\right)^{\frac{1}{\alpha-1}}}\frac{h_n^k}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}}\right]^{\frac{1}{1-\eta}} \\
&= \left[\frac{\Theta^k}{\Theta^0}\left(p_c^0\left(\left(\frac{p_c^0 - nx_c^0}{p_c^k - nx_c^k}\right)^{\frac{1}{\alpha-1}}\frac{h_c^k}{h_c^0}\right)^\theta + p_n^0\left(\left(\frac{p_n^0 - nx_n^0}{p_n^k - nx_n^k}\right)^{\frac{1}{\alpha-1}}\frac{h_n^k}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}}\right]^{\frac{1}{1-\eta}}
\end{aligned}$$

This is equations (24) and (25).

9.6 Equation (27)

We drop the superscripts "D" and "S", since in closed-economy, $L_i^{kD} = L_i^{kS}$, $i = n, c$. In addition, we normalize $P^k = 1$. By equation (4),

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

By equation (16),

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

Substituting this expression into the output equation yields

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

Substituting out p_c^k using equation (13), we obtain

$$w_c^k = \Theta^k \left(p_c^k\right)^{-\frac{1}{\alpha-1}} \left(A_c\right)^{\frac{\alpha}{\alpha-1}} \tag{43}$$

Rearranging equation (11),

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

Given that $E^k = \eta Y^k / L^k$, we can substitute w_c^k in equation (43) to obtain after rearranging

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

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

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

We then substitute equation (26) into this expression, to obtain an expression with no endogenous variables

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

Therefore,

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

Combining equations (9) and (26) 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 expressions for $\frac{Y^k/L^k}{Y^0/L^0}$, and simplifying, we arrive at our decomposition:

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

9.7 Equation (28)

Under free trade, $P^k = (\Theta^k)^{-1}$. This means that equation (12) implies

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

Combine this expression with equation (13), we can write real output per capita in country k relative to a base country 0 as

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

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 occupation shares from the base country, we obtain

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

9.8 Equations (29) and (30)

Re-interpret the country 0 in equation (39) as the initial equilibrium of country k, we have

$$\widehat{Y}^k = \left(\left[\frac{T_c\left(\frac{w_c^{k\prime}}{P^{k\prime}}h_c^k\right)^\theta + T_n\left(\frac{w_n^{k\prime}}{P^{k\prime}}h_n^k\right)^\theta}{T_c\left(\frac{w_c^k}{P^k}h_c^k\right)^\theta + T_n\left(\frac{w_n^k}{P^k}h_n^k\right)^\theta}\right]^{\frac{1}{\theta}}\right)^{\frac{1}{1-\eta}}$$

$$= \left[p_c^k\left(\frac{\widehat{w_c^k}}{\widehat{P^k}}\right)^\theta + p_n^k\left(\frac{\widehat{w_n^k}}{\widehat{P^k}}\right)^\theta\right]^{\frac{1}{\theta}\frac{1}{1-\eta}}$$

This is equation (29).

Now let the subsequent equilibrium be autarky, indicated by primes $'$. Re-write equation (29) as

$$\widehat{Y}^k = \left(\frac{\widehat{w_c^k}}{\widehat{P^k}}\right)^{\frac{1}{1-\eta}}\left[p_c^k + p_n^k\left(\frac{\widehat{w_n^k}}{\widehat{w_c^k}}\right)^\theta\right]^{\frac{1}{\theta}\frac{1}{1-\eta}}. \qquad (44)$$

By equation (41), we have

$$\frac{\widehat{w}_c^k}{\widehat{P}^k} = \frac{\left((A_c)^\alpha + (A_n)^\alpha\left(\frac{w_n^k}{w_c^k}\right)^{1-\alpha}\right)^{\frac{1}{1-\alpha}}}{\left((A_c)^\alpha + (A_n)^\alpha\left(\frac{w_n^{k\prime}}{w_c^{k\prime}}\right)^{1-\alpha}\right)^{\frac{1}{1-\alpha}}}$$

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

$$= \left(p_c^k(1-x_c^k) + p_n^k(1-x_n^k)\left(\frac{\widehat{w}_n^k}{\widehat{w}_c^k}\right)^{1-\alpha}\right)^{\frac{1}{\alpha-1}}$$

where the last equality uses the result $s_i^k = p_i^k(1-x_i^k), i = n, c$, which is implied by equations (14), (16), and (17). Substituting this expression into equation (44), we obtain

$$\widehat{Y}^k = \left(\left(p_c^k(1-x_c^k) + p_n^k(1-x_n^k)\left(\frac{\widehat{w}_n^k}{\widehat{w}_c^k}\right)^{1-\alpha}\right)^{\frac{1}{\alpha-1}}\left(p_c^k + p_n^k\left(\frac{\widehat{w_n^k}}{\widehat{w_c^k}}\right)^\theta\right)^{\frac{1}{\theta}}\right)^{\frac{1}{1-\eta}}. \qquad (45)$$

By equations (9) and (40), we have

$$\frac{w_n^k}{w_c^k} = \left(\frac{1 - x_c^k}{1 - x_n^k} \left(\frac{h_c^k}{h_n^k} \right)^\theta \left(\frac{A_n}{A_c} \right)^\alpha \right)^{\frac{1}{\theta+\alpha-1}}.$$

This means that

$$\frac{\widehat{w_n^k}}{\widehat{w_c^k}} = \left(\frac{1 - x_n^k}{1 - x_c^k} \right)^{\frac{1}{\theta+\alpha-1}},$$

where we have used $x_c^{k'} = x_n^{k'} = 0$ at the subsequent equilibrium of autarky. Substituting this back into (45), we obtain

$$\widehat{Y}^k = \left(\left(p_c^k(1 - x_c^k) + p_n^k(1 - x_n^k) \left(\frac{1 - x_n^k}{1 - x_c^k} \right)^{\frac{1-\alpha}{\theta+\alpha-1}} \right)^{\frac{1}{\alpha-1}} \left(p_c^k + p_n^k \left(\frac{1 - x_n^k}{1 - x_c^k} \right)^{\frac{\theta}{\theta+\alpha-1}} \right)^{\frac{1}{\theta}} \right)^{\frac{1}{1-\eta}}$$

To simplify the steps let $p_c^k(1 - x_c^k)^{\frac{\theta}{\theta+\alpha-1}} + p_n^k \left(1 - x_n^k\right)^{\frac{\theta}{\theta+\alpha-1}} = C$. Then we have

$$\begin{aligned}
(\widehat{Y}^k)^{1-\eta} = \quad & ((1 - x_c^k)^{\frac{\alpha-1}{\theta+\alpha-1}})^{\frac{1}{\alpha-1}} (p_c^k(1 - x_c^k)^{1+\frac{1-\alpha}{\theta+\alpha-1}} + p_n^k \left(1 - x_n^k\right)^{1+\frac{1-\alpha}{\theta+\alpha-1}})^{\frac{1}{\alpha-1}} \frac{C^{\frac{1}{\theta}}}{((1 - x_c^k)^{\frac{\theta}{\theta+\alpha-1}})^{\frac{1}{\theta}}} \\
= \quad & C^{\frac{1}{\alpha-1}+\frac{1}{\theta}} = (p_c^k(1 - x_c^k)^{\frac{\theta}{\theta+\alpha-1}} + p_n^k \left(1 - x_n^k\right)^{\frac{\theta}{\theta+\alpha-1}})^{\frac{1}{\alpha-1}+\frac{1}{\theta}}
\end{aligned}$$

This is equation (30).

9.9 Proof of the Gains from Trade, for (30)

To be specific, we prove that $\widehat{Y}^k < 1$ if $x_i^k < 1$ for at least one i in equation (30). This equation can be written as

$$(\widehat{Y}^k)^{(1-\eta)(\alpha-1)} = (p_c^k(1 - x_c^k)^{\frac{\theta}{\theta+\alpha-1}} + p_n^k \left(1 - x_n^k\right)^{\frac{\theta}{\theta+\alpha-1}})^{\frac{\theta+\alpha-1}{\theta}}$$

Substituting using equation (5), we have

$$(\widehat{Y}^k)^{(1-\eta)(\alpha-1)} = \left(p_c^k \left(\frac{L_c^{kD}}{L_c^{kS}} \right)^{\frac{\theta}{\theta+\alpha-1}} + p_n^k \left(\frac{L_n^{kD}}{L_n^{kS}} \right)^{\frac{\theta}{\theta+\alpha-1}} \right)^{\frac{\theta+\alpha-1}{\theta}}$$

From this expression it is clear that to gain from trade (i.e. $\widehat{Y}^k < 1$) it must be that $\frac{L_i^{kS}}{L_i^{kD}} < 1$ for at least one i. Trade balance requires that if $\frac{L_c^{kS}}{L_c^{kD}} > 1$, then $\frac{L_n^{kS}}{L_n^{kD}} < 1$, and vice versa.

We now prove by contradiction. Suppose that $\widehat{Y}^k > 1$. Then, we must have

$$p_c^k \left(\frac{L_c^{kD}}{L_c^{kS}} \right)^{\frac{\theta}{\theta+\alpha-1}} + p_n^k \left(\frac{L_n^{kD}}{L_n^{kS}} \right)^{\frac{\theta}{\theta+\alpha-1}} > 1 \tag{46}$$

Without loss of generality, let $X \equiv \frac{L_c^{kD}}{L_c^{kS}} > 1$ and $Y \equiv \frac{L_n^{kD}}{L_n^{kS}} < 1$. Then trade balance requires

$$p_c^k X + p_n^k Y = 1.$$

We can rewrite condition (46) as

$$\begin{aligned} p_c^k f(X) + p_n^k f(Y) &> 1 \\ p_c^k f(X) + \left(1 - p_c^k\right) f(Y) &> 1 \\ p_c^k \left(f(X) - f(Y)\right) &> 1 - f(Y) \end{aligned}$$

where f is continuous, increasing, and concave, i.e. $\alpha > 1$ and

$$f(z) = z^{\frac{\theta}{\theta+\alpha-1}}.$$

The trade balance condition can be written

$$p_c^k (X - Y) = 1 - Y$$

Now dividing the rearranged condition by the trade balance, we have

$$\frac{f(X) - f(Y)}{X - Y} > \frac{1 - f(Y)}{1 - Y}$$

But $X > 1 > Y > 0$ and $f(.)$ is increasing and concave, hence

$$\frac{f(X) - f(Y)}{X - Y} < \frac{1 - f(Y)}{1 - Y}.$$

This contradicts the assertion, so it must be the case that $\widehat{Y}^k < 1$.

9.10 Equation (33)

By equations (13) and (22)

$$\begin{aligned}
w_c^k S^k L^k &= b p_c^k P^k Y^k \\
\frac{Y^k}{S^k L^k} &= \frac{1}{b p_c^k} \frac{w_c^k}{P^k} \\
\frac{Y^k}{S^k L^k (1 - x_c^k)} &= \frac{1}{b p_c^k (1 - x_c^k)} \frac{w_c^k}{P^k} \\
&= \frac{\Theta^k}{b p_c^k (1 - x_c^k)} \frac{w_c^k}{\left((A_c)^\alpha (w_c^k)^{1-\alpha} + (A_n)^\alpha (w_n^k)^{1-\alpha} \right)^{\frac{1}{1-\alpha}}} \\
&= \frac{\Theta^k}{(A_c)^{\frac{\alpha}{1-\alpha}} b p_c^k (1 - x_c^k)} \left(\frac{(A_c)^\alpha \left(w_c^k\right)^{1-\alpha}}{(A_c)^\alpha (w_c^k)^{1-\alpha} + (A_n)^\alpha (w_n^k)^{1-\alpha}} \right)^{\frac{1}{1-\alpha}} \\
&= \frac{\Theta^k}{(A_c)^{\frac{\alpha}{1-\alpha}} b p_c^k (1 - x_c^k)} \left(s_c^k \right)^{\frac{1}{1-\alpha}}
\end{aligned}$$

Since

$$s_c^k = p_c^k (1 - x_c^k)$$

we have

$$\frac{Y^k}{S^k L^k (1 - x_c^k)} = \frac{\Theta^k}{(A_c)^{\frac{\alpha}{1-\alpha}} b} \left(p_c^k (1 - x_c^k) \right)^{\frac{\alpha}{1-\alpha}}$$

and finally

$$\begin{aligned}
1 + \frac{p_n^k (1 - x_n^k)}{p_c^k (1 - x_c^k)} &= \frac{p_c^k (1 - x_c^k) + p_n^k (1 - x_n^k)}{p_c^k (1 - x_c^k)} \qquad (47) \\
&= \frac{1 - p_c^k x_c^k - p_n^k x_n^k}{p_c^k (1 - x_c^k)} \\
&= \frac{1 - p_c^k x_c^k - p_n^k x_n^k}{p_c^k (1 - x_c^k)} \\
&= \frac{1}{p_c^k (1 - x_c^k)} = \frac{1}{s_c^k}
\end{aligned}$$

where the last expression follows from trade balance. hence

$$\begin{aligned}
\frac{Y^k}{S^k L^k (1 - x_c^k)} &= \frac{\Theta^k}{(A_c)^{\frac{\alpha}{1-\alpha}} b} \left(p_c^k (1 - x_c^k) \right)^{\frac{\alpha}{1-\alpha}} \\
&= \frac{\Theta^k}{(A_c)^{\frac{\alpha}{1-\alpha}} b} \left(1 + \frac{p_n^k (1 - x_n^k)}{p_c^k (1 - x_c^k)} \right)^{\frac{\alpha}{\alpha-1}}
\end{aligned}$$

This implies equation (33).

9.11 Equations (34) and (35)

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. Equations (21), (32), (33) and the identity $p_n^k + p_c^k = 1$ imply that

$$\widehat{S^k} = \widehat{h_c^k}(\widehat{Y^k})^{\eta}(\widehat{p_c^k})^{1-\frac{1}{\theta}}$$

$$\widehat{Y^k} = \widehat{S^k}(\widehat{p_c^k})^{-\frac{\alpha}{\alpha-1}}$$

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

$$\widehat{p_c^k} = \frac{1}{p_c^k} - \widehat{p_n^k}\frac{p_n^k}{p_c^k}$$

Solving these equations, we obtain

$$\widehat{S^k} = (\widehat{h_c^k})^{\frac{1}{1-\eta}}\frac{(\widehat{h_c^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}G}}{[p_c^k(\widehat{h_c^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}} + p_n^k(\widehat{h_n^k})^{\frac{\theta(\alpha-1)}{\theta+\alpha-1}}]^G}, G = \frac{1}{1-\eta}(1 - \frac{1}{\theta} - \frac{\eta\alpha}{\alpha-1}),$$

where our closed-economy parameter values imply that the constant $G = 0.3676$.

For small changes, 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 (21), (32) and (33) imply that

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

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

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

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 (35).

9.12 Derivation of (36)

With free trade, the international equilibrium condition (18) simplifies to

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

Using equation (13), the first-order condition for cost minimization, and factor price equalization, the international equilibrium condition becomes

$$\sum_k p_c^k \frac{P^k Y^k}{\sum_{k'} P^{k'} Y^{k'}} = s_c$$

where a country's output weight can be written

$$\frac{P^k Y^k}{\sum_{k'} P^{k'} Y^{k'}} = \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(T_c \left(\frac{w_c}{w_n} h_c^k\right)^\theta + T_n \left(h_n^k\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'} \left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(T_c \left(\frac{w_c}{w_n} h_c^{k'}\right)^\theta + T_n \left(h_n^{k'}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}.$$

Substituting for factor supplies, factor demands, and for income weights, the international equilibrium condition becomes

$$\sum_k \frac{\frac{T_c}{T_n}(h_c^k \frac{w_c}{w_n})^\theta}{\frac{T_c}{T_n}(h_c^k \frac{w_c}{w_n})^\theta + (h_n^k)^\theta} \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(\frac{T_c}{T_n} \left(\frac{w_c}{w_n} h_c^k\right)^\theta + \left(h_n^k\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'} \left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(\frac{T_c}{T_n} \left(\frac{w_c}{w_n} h_c^{k'}\right)^\theta + \left(h_n^{k'}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}$$

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

Defining $\widetilde{\omega} = \left(\frac{A_c}{A_n}\right)^{\frac{\alpha}{1-\alpha}} \left(\frac{w_c}{w_n}\right)$, we can substitute

$$\frac{w_c}{w_n} = \widetilde{\omega} \left(\frac{A_n}{A_c}\right)^{\frac{\alpha}{1-\alpha}}$$

into the international equilibrium condition to obtain

$$\sum_k \frac{\frac{T_c}{T_n}\left(\left(\frac{A_n}{A_c}\right)^{\frac{\alpha}{1-\alpha}} h_c^k \widetilde{\omega}\right)^\theta}{\frac{T_c}{T_n}\left(\left(\frac{A_n}{A_c}\right)^{\frac{\alpha}{1-\alpha}} h_c^k \widetilde{\omega}\right)^\theta + (h_n^k)^\theta} \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(\frac{T_c}{T_n}\left(\widetilde{\omega}\left(\frac{A_n}{A_c}\right)^{\frac{\alpha}{1-\alpha}} h_c^k\right)^\theta + \left(h_n^k\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'} \left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(\frac{T_c}{T_n}\left(\left(\frac{A_n}{A_c}\right)^{\frac{\alpha}{1-\alpha}} \widetilde{\omega} h_c^{k'}\right)^\theta + \left(h_n^{k'}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}$$

$$= \quad \frac{(\widetilde{\omega})^{1-\alpha}}{(\widetilde{\omega})^{1-\alpha}+1}$$

Defining

$$\Psi \equiv \frac{T_c}{T_n}\left(\frac{A_n}{A_c}\right)^{\frac{\theta\alpha}{1-\alpha}}$$

the international equilibrium condition becomes

$$\sum_k \frac{\Psi\left(h_c^k\widetilde{\omega}\right)^\theta}{\Psi\left(h_c^k\widetilde{\omega}\right)^\theta + (h_n^k)^\theta} \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(\Psi\left(h_c^k\widetilde{\omega}\right)^\theta + \left(h_n^k\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'} \left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(\Psi\left(h_c^{k'}\widetilde{\omega}\right)^\theta + \left(h_n^{k'}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}} = \frac{(\widetilde{\omega})^{1-\alpha}}{(\widetilde{\omega})^{1-\alpha}+1}$$

Rearranging this expression so that all human capital productivities appear as ratios, we obtain

$$\frac{(\widetilde{\omega})^{1-\alpha}}{(\widetilde{\omega})^{1-\alpha}+1} = \sum_k \frac{\Psi\left(\frac{h_c^0}{h_n^0}\right)^\theta\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta}{\Psi\left(\frac{h_c^0}{h_n^0}\right)^\theta\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta + \left(\frac{h_n^k}{h_n^0}\right)^\theta} \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(\Psi\left(\frac{h_c^0}{h_n^0}\right)^\theta\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta + \left(\frac{h_n^k}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'} \left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(\Psi\left(\frac{h_c^0}{h_n^0}\right)^\theta\left(\frac{h_c^{k'}}{h_c^0}\widetilde{\omega}\right)^\theta + \left(\frac{h_n^{k'}}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}$$

Now, let us assume that the data we observe can be well approximated by the closed economy equilibrium. We have, by equation (9),

$$\frac{p_c^0}{p_c^0} = \frac{T_c}{T_n}\left(\frac{h_c^0}{h_n^0}\frac{w_c^0}{w_n^0}\right)^\theta.$$

Substituting for the autarky equilibrium wages in the base country using equation (26), we obtain

$$\Psi\left(\frac{h_c^0}{h_n^0}\right)^\theta = \left(\frac{p_c^0}{p_n^0}\right)^{\frac{\theta+\alpha-1}{\alpha-1}}.$$

Substituting this expression back into the labor market clearing condition, we obtain equation (36):

$$\sum_k \frac{H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta}{H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^k}{h_n^0}\right)^\theta} \frac{\left(\Theta^k\right)^{\frac{\eta}{1-\eta}} L^k \left(H\left(\frac{h_c^k}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^k}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}}{\sum_{k'}\left(\Theta^{k'}\right)^{\frac{\eta}{1-\eta}} L^{k'} \left(H\left(\frac{h_c^{k'}}{h_c^0}\widetilde{\omega}\right)^\theta+\left(\frac{h_n^{k'}}{h_n^0}\right)^\theta\right)^{\frac{1}{\theta}\frac{1}{1-\eta}}} = \frac{(\widetilde{\omega})^{1-\alpha}}{(\widetilde{\omega})^{1-\alpha}+1}, H=\left(\frac{p_c^0}{p_n^0}\right)^{\frac{\theta+\alpha-1}{\alpha-1}}$$

9.13 Equation (38)

By equation (33),

$$Y^k = \Theta^k (A_c)^{\frac{\alpha}{\alpha-1}} L^k S^k b^{-1}\left(1-x_c^k\right)\left(1+\frac{1-x_n^k}{1-x_c^k}\frac{p_n^k}{p_c^k}\right)^{\frac{\alpha}{\alpha-1}}$$

We can now substitute out S^k using equation (23), to obtain

$$(\frac{Y^k}{L^k})^{1-\eta} = \Theta^k (A_c)^{\frac{\alpha}{\alpha-1}} \gamma^{1-\eta}\eta^\eta (T_c)^{\frac{1}{1-\eta}} h_c^k (p_c^k)^{1-\frac{1}{\theta}}\left(1-x_c^k\right)\left(1+\frac{1-x_n^k}{1-x_c^k}\frac{p_n^k}{p_c^k}\right)^{\frac{\alpha}{\alpha-1}}$$

Now substituting using equation (47), and comparing the expression above between the initial (denoted by the superscript "1") and subsequent (denoted by "2 ") equilibria, we have

$$\begin{aligned}
(\widehat{Y^k})^{1-\eta} &= \widehat{h_c^k}(\frac{p_c^{k2}}{p_c^{k1}})^{1-\frac{1}{\theta}}\frac{1-x_c^{k2}}{1-x_c^{k1}}\left(\widehat{s_c^k}\right)^{-\frac{\alpha}{\alpha-1}} \\
&= \widehat{h_c^k}(\widehat{p_c^k})^{-\frac{1}{\theta}}\frac{s_c^{k2}}{s_c^{k1}}\left(\widehat{s_c^k}\right)^{-\frac{\alpha}{\alpha-1}} \\
&= \widehat{h_c^k}(\widehat{p_c^k})^{-\frac{1}{\theta}}\left(\widehat{s_c^k}\right)^{-\frac{1}{\alpha-1}}
\end{aligned}$$

This is equation (38).

10 Data Appendix

10.1 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.

10.2 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". Leadership in work style = 1.C.2.b; coordinating the work and activities of others = 4.A.4.b.1; developing and building teams = 4.A.4.b.2; coaching and developing others = 4.A.4.b.5; recruiting and promoting employees = 4.A.4.c.2; monitoring and controlling resources and spending = 4.A.4.c.3; and coordinate or lead others in work = 4.C.1.b.1.g.

In addition, 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. Social skills involve "working with, communicating with, and teaching people". 1.B.1.d. 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. 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.

In raw O*NET data, there are two metrics, level and importance, for the characteristics we use. These two metrics are highly correlated. We use the raw importance metric, because its interpretation is the same across characteristics; e.g. a value of 3 or higher indicates that this characteristic is important for an occupation.

As reported 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.

Table A2 reports our results for the Neal-Johnson regressions when we use the first principal component of leadership. Columns (1) and (2) show that the AFQT score has a smaller coefficient estimate for the non-cognitive subsample than for the cognitive sub-sample. Column (3) shows that this difference is significant for the pooled sample when we include the interaction between AFQT score and the non-cognitive-occupation dummy. These results echo Table 1.

10.3 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 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.

10.4 Test Score Data

We have tabulated over-time changes of PISA scores within countries and found limited variation. For example, for the U.S. reading score the mean is 499.26 and the standard deviation is 3.93. The summary statistics by country is available upon request.

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. These results are reported in Table A3.

10.5 Industry Level Trade Data

We obtain the 6-digit HS (Harmonized System) import and export data for merchandise trade from COMTRADE, and convert the HS6 codes to 1997 NAICS codes using the mapping of Pierce and Schott (2009). We obtain the data for service trade from the United Nations Service Trade database. To convert the service-industry codes of NAICS 1997 into the 1-digit service-trade codes, we start from the mapping of Liu and Trefler (2011) and augment it with our own mapping.

10.6 Factor Content of Trade

Our computation of factor content of trade follows similar steps as Costinot and Rodriguez-Clare (2014). We first use US 2000 Census to get data for wage bill by industry for cognitive and non-cognitive type workers, where our industries are the same as in the previous sub-section. We then use the NBER Productivity Database to get data for output for manufacturing industries, and the United Nations UNIDO Database to get those for service industries. For each industry, we compute the value of cognitive (non-cognitive) type service embodied in trade as net export multiplied by the ratio of cognitive (non-cognitive) wage bill to output. We then sum across industries and divide the total by country k's aggregate output. These numbers do not correspond to the variable x_i^k in our model; rather, they correspond to $nx_i^k = w_i^k \left(L_i^{kS} - L_i^{kD}\right)/(P^k y^k)$, which is the value of net exports of type i human capital normalized by output. It is easy to show that $x_i^k = nx_i^k/p_i^k$, and this expression allows us to compute x_i^k using nx_i^k. In our computation, we have implicitly assumed that cognitive and non-cognitive types have the same cost shares across countries, because we only have cost-share data for the U.S. This assumption is also used in Costinot and Rodriguez-Clare (2014).

An alternative approach to calculate factor content of trade is to use industry employment as raw data, rather than wage bill (e.g. Davis and Weinstein 2001). We have expreimented with this approach, too. To be specific, we multiply the net export value of each industry by the ratio of cognitive (non-cognitive) employment to output, sum across industries, and then divide the total by country k's aggregate cognitive (non-cognitive) employment. These numbers correspond to the variables x_i^k. Under this alternative approach, results are very similar (i.e. θ and cognitive productivities are identical, $\alpha = 1.5054$, and the correlation coefficients with benchmark values are 0.9853 for non-cognitive productivity, 0.8979 for overall education quality, and 0.9621 for output TFP), except that Luxembourg has to be dropped from the analysis since her values of x_i^k exceed 1.

These results are based on direct factor requirements, and we now discuss the results based on total factor requirements. We start from the detailed industry-by-industry total requirement matrix, O, for the U.S. in 2007, with 389 IO industries, downloaded from the US BEA website. O_{11}, the element of the first row and first column, shows the value of industry 1's gross output needed to deliver \$1 of industry 1's net output. O_{21} shows the value of industry 2's gross output needed to deliver \$1 of industry 1's net output, and so on. On the other hand, O_{12} shows the value of industry 1's gross output needed to deliver \$1 of industry 2's net output, and so on. We next aggregate the IO industries into 31 NAICS-97 manufacturing industries and 9 1-digit service industries using the mapping from US BEA, plus our own mapping from the previous sub-section. In our aggregation we sum across the rows of O and compute means across its columns, obtaining the 40-by-40 total requirement matrix AO. We then multiply the transpose of AO by the matrix of direct requirements for cognitive and non-cognitive human capital, and obtain the data for their total requirements.

We obtain similar results; i.e. θ and cognitive productivities are identical, $\alpha = 1.7026$, and the correlation coefficients with benchmark values are 0.9922 for non-cognitive productivity, 0.9896 for overall education quality, and 0.9683 for output TFP.

10.7 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.

References

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

[2] 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.

Figure 1 Test Score and Educational Spending Per Capita

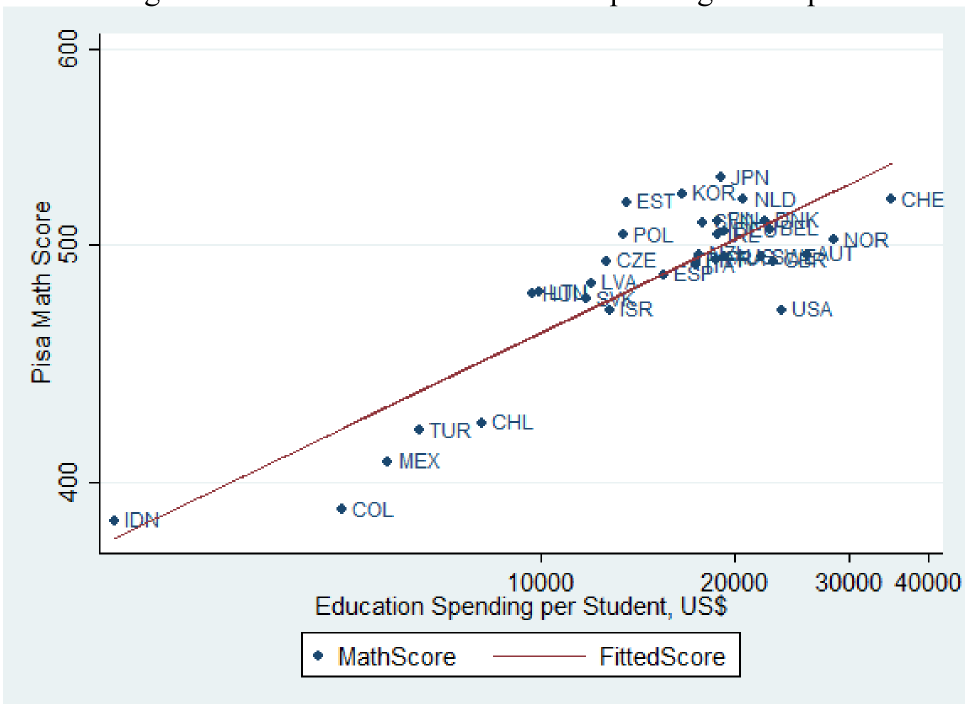

Figure 2 Normalized Test Scores and Cognitive Employment Shares

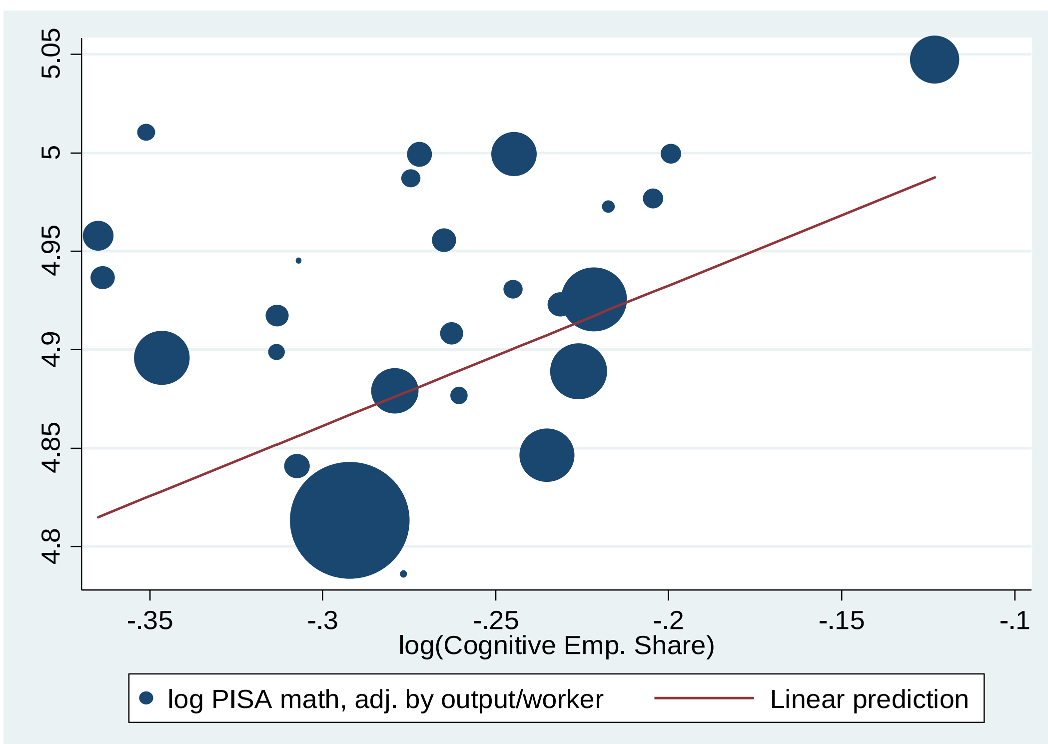

Figure 3 Cognitive-Productivity Ranking vs. PISA-Math Ranking

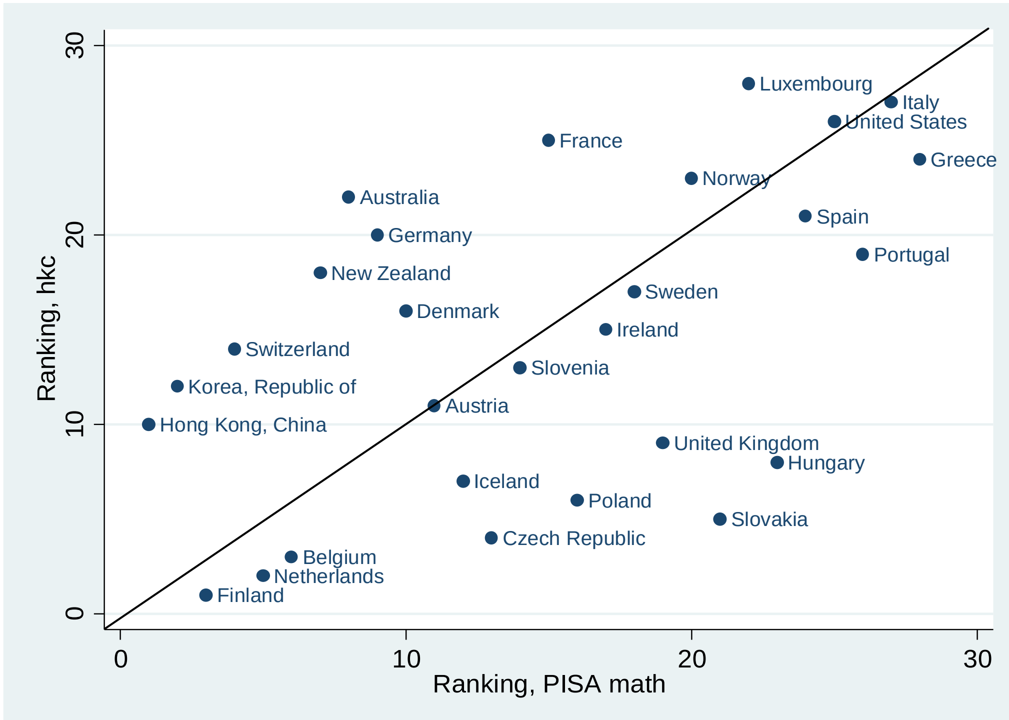

Figure 4 Non-Cognitive-Productivity Ranking vs. PISA-Math Ranking

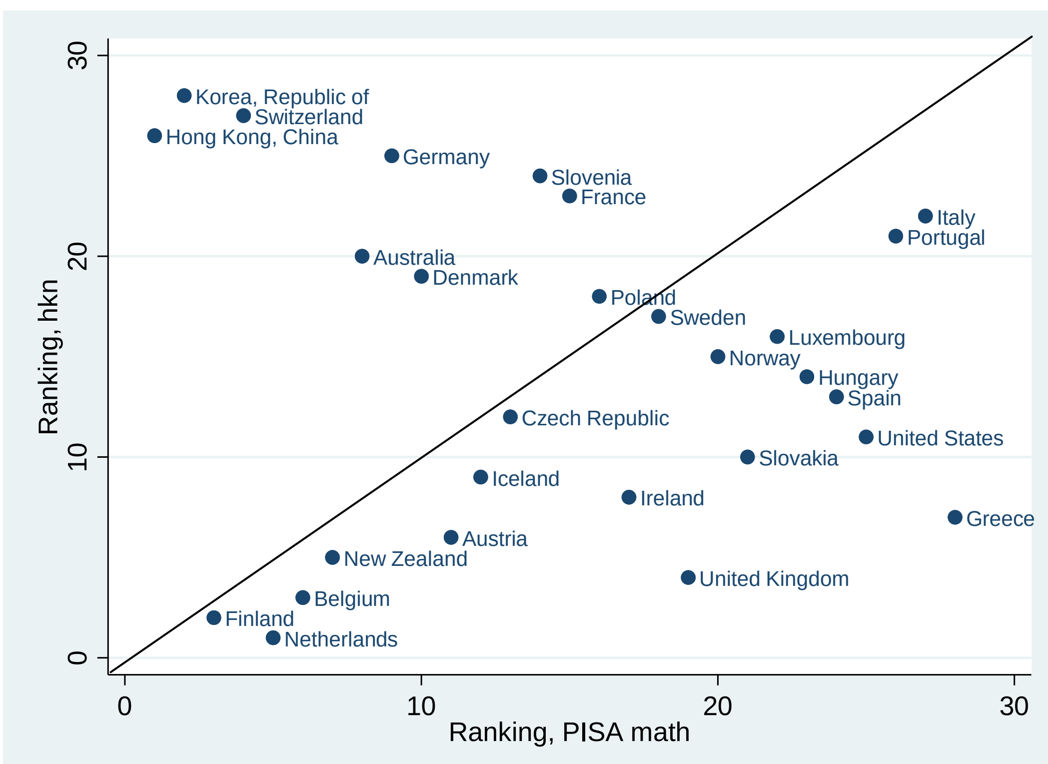

Figure 5 Overall Education Quality

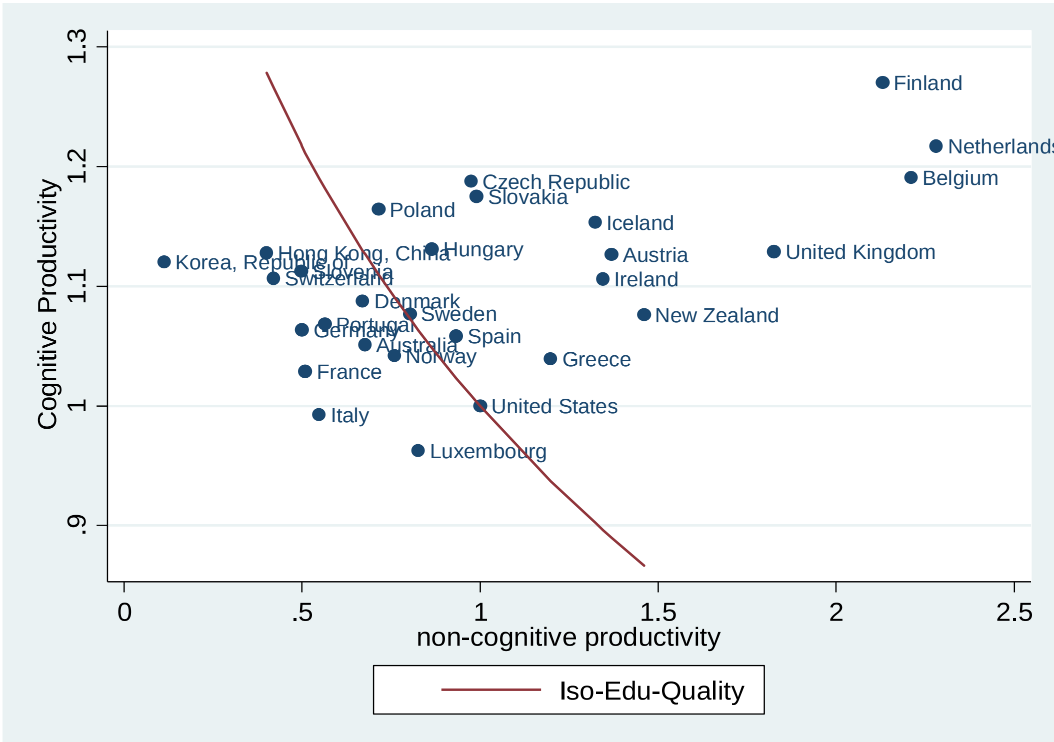

Figure 6 Output Gains, Autarky to Free Trade, Sample Countries

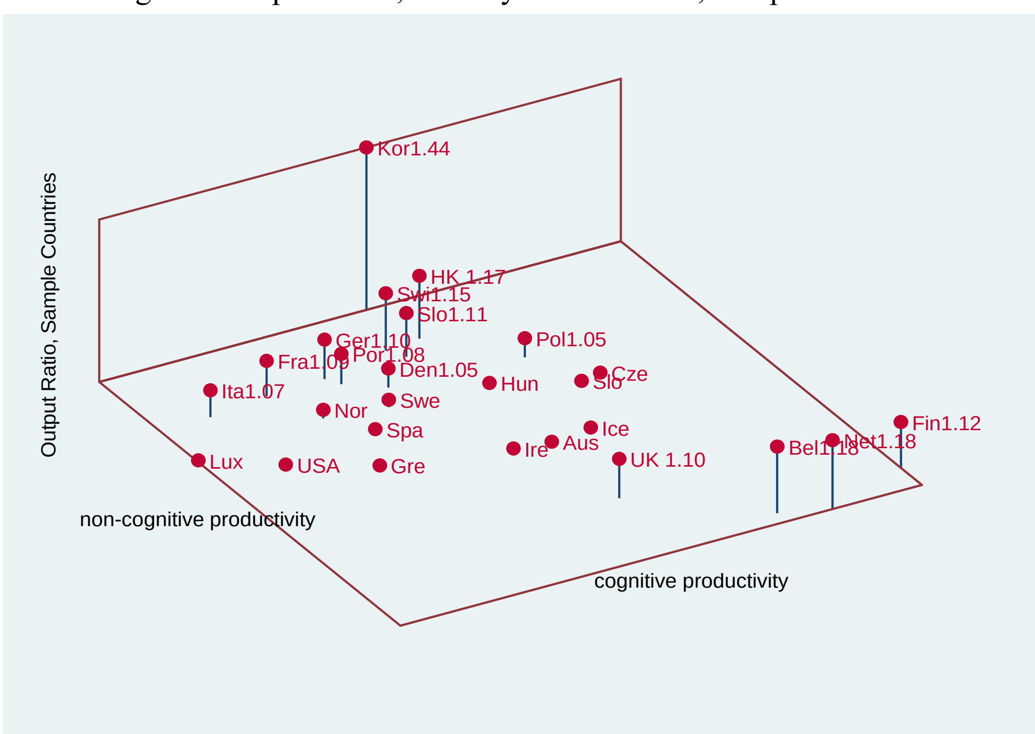

Figure 7 Overall Education Quality: Free-trade Counterfactual

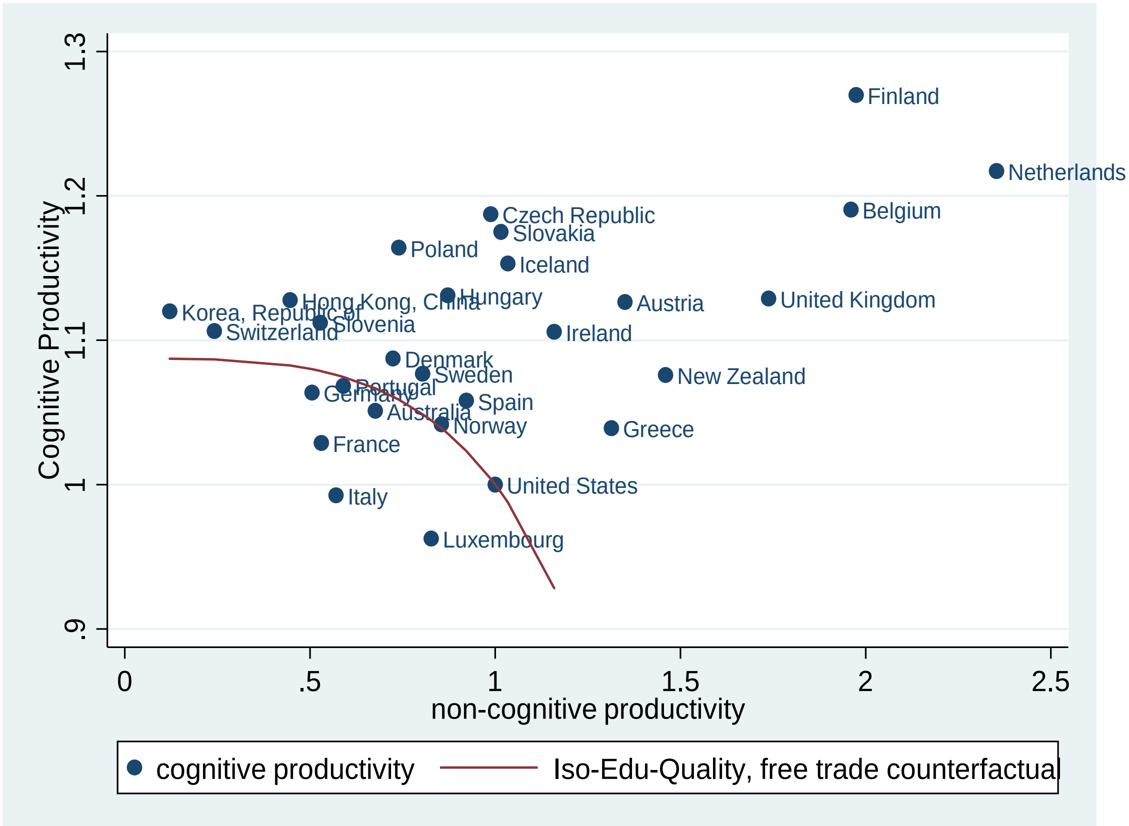

Figure 8 Output Gains, Autarky to Free Trade, Adding China & Japan

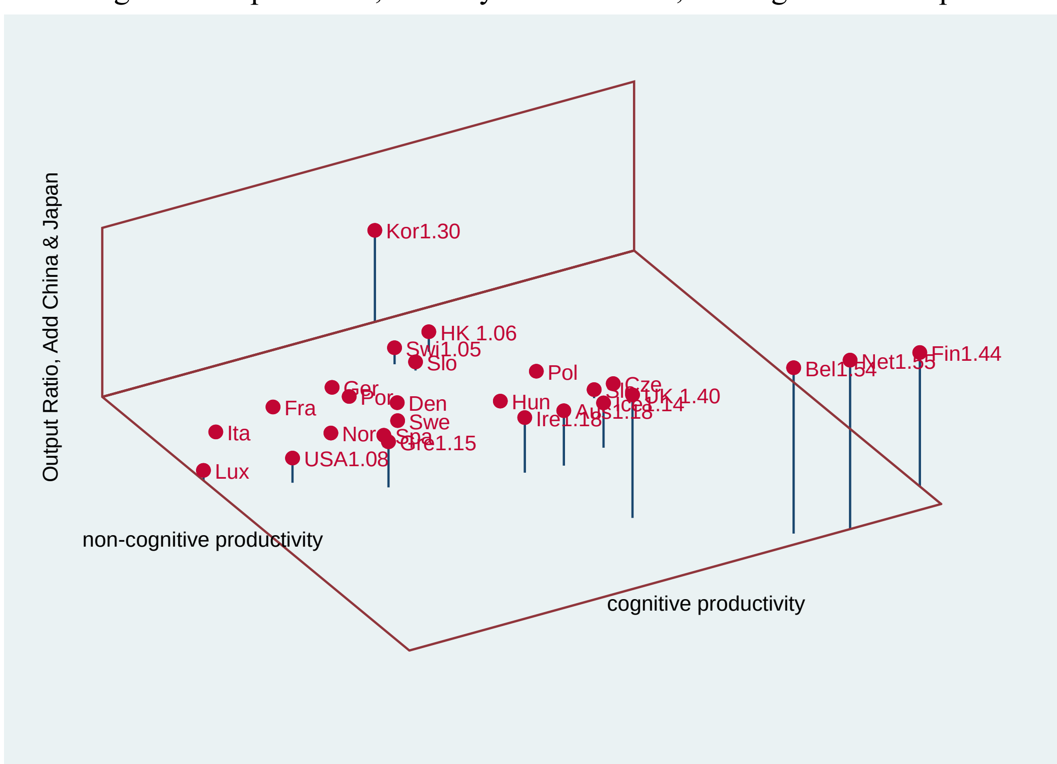

Figure 9 Overall Education Quality: Extended Sample

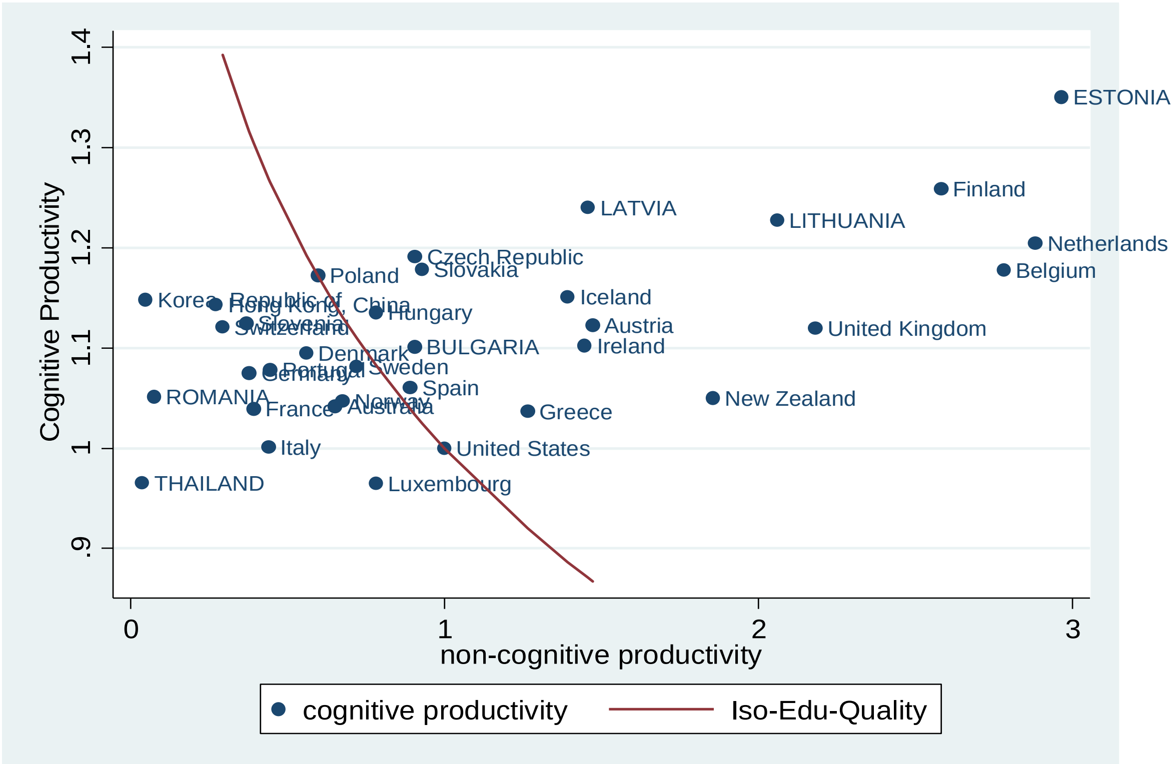

Table 1 Test Score and Wages of Non-cognitive and Cognitive Occupations

VARIABLES	(1) Replicate	(2) Non-Cog. SubSample	(3) Cog. SubSample	(4) Interaction	(5) 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)
$AFQT^2$	-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
R^2	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
$\lvert x_c^k \rvert$	28	0.0194	0.0156	0.0001	0.0626
$\lvert x_n^k \rvert$	28	0.0357	0.0348	0.0011	0.1396

Table 3 Sample Countries, Years and Rankings

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

Table 4 Value of θ

Dependent Variable = normalized test score, equation (33)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
lnp_c^k	0.717***	0.714***	0.696***	0.521***	0.512**	0.677*	0.565**
	(0.230)	(0.224)	(0.223)	(0.165)	(0.201)	(0.357)	(0.198)
ASNZ		0.213**	0.208***	0.189***	0.189**	0.175**	0.210**
		(0.0773)	(0.0574)	(0.0570)	(0.0695)	(0.0842)	(0.0682)
Constant	5.076***	5.075***	5.072***	5.032***	5.040***	5.032***	5.020**
	(0.0624)	(0.0607)	(0.0608)	(0.0448)	(0.0546)	(0.0784)	(0.0547)
Observations	26	28	28	28	28	28	34
R^2	0.288	0.347	0.393	0.384	0.292	0.196	0.302

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

Dependent Variable = normalized and adjusted output per worker, equation (34)

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln(1+\frac{p_n^k(1-x_n^k)}{p_c^k(1-x_c^k)})$	2.802**	2.784**	2.746**	2.616**	2.605**	3.418*	3.125**	3.872**
	(1.191)	(1.224)	(1.171)	(1.140)	(1.177)	(1.840)	(1.224)	(1.100)
ASNZ	-1.074**		-1.080***	-1.051**	-1.051**	-0.976**	-1.094**	-1.246***
	(0.425)		(0.311)	(0.407)	(0.420)	(0.434)	(0.423)	(0.379)
Constant	3.548***	3.553***	3.564***	3.590***	3.583***	3.554***	3.465***	3.004***
	(0.324)	(0.332)	(0.320)	(0.310)	(0.320)	(0.404)	(0.332)	(0.274)
Obs. No.	28	26	28	28	28	28	28	34
R2	0.263	0.177	0.342	0.264	0.251	0.209	0.282	0.363

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, (31)
θ	Dispersion of Innate Ability	2.0877~3.4965	Strength of selection effect, (32)
α	Sub Elasticity in Agg Production	1.4706~1.5549	Agg. production function, (33)
Θ^k	Output TFP	Table 7	Same as α, (33)
h_c^k	TFP, Cog. Human Cap.	Figures 3, 5, 7 & 9	Normalized test score and log cog. emp. share, (32)
h_n^k	TFP, Non-cog. Human Cap.	Figures 4, 5, 7 & 9	Revealed comp advantage by relative emp. share and trade, (20)

Table 7 Overall Education Quality and Output per Worker

		Closed Economy		Trade Cost		% Output, Closed-Econ	
	1	2	3	4	5	6	7
Countries	Output Per Worker	Contribution of Output TFP	Contribution of Overall Edu Quality	Contribution of Output TFP	Contribution of Overall Edu Quality	1% Rise in h_c^k	1% Rise in h_n^k
Austria	0.6434	0.5297	1.2147	0.5645	1.1399	0.84%	0.31%
Belgium	0.6892	0.4636	1.4867	0.5301	1.3001	0.79%	0.35%
Czech Republic	0.3293	0.2860	1.1513	0.2982	1.1045	0.87%	0.27%
Denmark	0.5979	0.6187	0.9664	0.7037	0.8496	0.89%	0.25%
Finland	0.5037	0.3259	1.5458	0.3326	1.5145	0.80%	0.34%
France	0.7329	0.8517	0.8606	0.8720	0.8405	0.91%	0.23%
Germany	0.6296	0.7126	0.8834	0.7088	0.8883	0.92%	0.23%
Greece	0.5190	0.4761	1.0901	0.5144	1.0089	0.84%	0.30%
H.K., China	0.6864	0.7724	0.8887	0.7840	0.8755	0.94%	0.21%
Hungary	0.3517	0.3292	1.0684	0.3375	1.0419	0.88%	0.27%
Iceland	0.5110	0.4168	1.2261	0.5139	0.9943	0.84%	0.30%
Ireland	0.6642	0.5583	1.1896	0.5136	1.2930	0.84%	0.31%
Italy	0.6761	0.7977	0.8476	0.7982	0.8470	0.90%	0.24%
S. Korea	0.4304	0.6027	0.7142	0.6267	0.6868	1.01%	0.13%
Luxembourg	1.4376	1.5674	0.9172	1.5674	0.9172	0.87%	0.28%
Netherlands	0.6712	0.4387	1.5300	0.4624	1.4515	0.79%	0.35%
Norway	0.7289	0.7589	0.9605	0.8704	0.8374	0.88%	0.26%
Poland	0.3045	0.2917	1.0438	0.2979	1.0219	0.90%	0.25%
Portugal	0.3845	0.4216	0.9121	0.4195	0.9165	0.91%	0.24%
Slovakia	0.2979	0.2600	1.1459	0.2732	1.0901	0.87%	0.27%
Slovenia	0.3929	0.4275	0.9191	0.4406	0.8918	0.92%	0.22%
Spain	0.6087	0.5913	1.0293	0.5979	1.0180	0.86%	0.28%
Sweden	0.5937	0.5917	1.0034	0.6389	0.9292	0.88%	0.26%

Table 7 Continued

		Closed Economy		Trade Cost		% Output, Closed-Econ	
	1	2	3	4	5	6	7
Countries	Output Per Worker	Contribution of Output TFP	Contribution of Overall Edu Quality	Contribution of Output TFP	Contribution of Overall Edu Quality	1% Rise in h_c^k	1% Rise in h_n^k
Switzerland	0.5855	0.6641	0.8816	0.9880	0.5926	0.93%	0.21%
United Kingdom	0.6349	0.4758	1.3345	0.5040	1.2597	0.81%	0.33%
United States	1.0000	1.0000	1.0000	1.0000	1.0000	0.85%	0.29%

Table 8 Change in Overall Education Quality and Gains from Trade

	Overall Edu. Quality			Gains From Trade		% Output, Free Trade, Sample Country		% Output, Free Trade, Add China & Japan	
Countries	Closed Economy (1)	Free Trade (2)	(3) = (2)/(1)	Sample Countries (4)	Adding China & Japan (5)	1% Rise in h_c^k (6)	1% Rise in h_n^k (7)	%1 Rise in h_c^k (8)	%1 Rise in h_n^k (9)
Austria	1.2147	1.2358	1.0174	1.0175	1.1810	0.67%	0.48%	0.32%	0.83%
Belgium	1.4867	1.7565	1.1814	1.1816	1.5453	0.29%	0.85%	0.09%	1.04%
Czech Republic	1.1513	1.1618	1.0091	1.0092	1.0232	0.97%	0.18%	0.68%	0.47%
Denmark	0.9664	1.0161	1.0514	1.0516	1.0003	1.07%	0.07%	0.92%	0.23%
Finland	1.5458	1.7387	1.1248	1.1250	1.4401	0.37%	0.78%	0.13%	1.01%
France	0.8606	0.9432	1.0960	1.0962	1.0158	1.10%	0.04%	1.02%	0.12%
Germany	0.8834	0.9783	1.1074	1.1075	1.0221	1.10%	0.03%	1.04%	0.11%
Greece	1.0901	1.1006	1.0096	1.0098	1.1502	0.72%	0.43%	0.36%	0.79%
Hong Kong	0.8887	1.0405	1.1709	1.1710	1.0662	1.13%	0.01%	1.10%	0.04%
Hungary	1.0684	1.0867	1.0172	1.0173	1.0114	1.00%	0.14%	0.75%	0.40%
Iceland	1.2261	1.2371	1.0090	1.0092	1.1474	0.73%	0.43%	0.36%	0.79%
Ireland	1.1896	1.2104	1.0175	1.0176	1.1816	0.67%	0.48%	0.31%	0.83%
Italy	0.8476	0.9094	1.0730	1.0732	1.0055	1.09%	0.05%	0.98%	0.17%
S. Korea	0.7142	1.0292	1.4411	1.4413	1.3012	1.14%	0.00%	1.14%	0.00%
Luxembourg	0.9172	0.9219	1.0051	1.0052	1.0332	0.95%	0.20%	0.64%	0.51%
Netherlands	1.5300	1.8167	1.1873	1.1875	1.5558	0.29%	0.84%	0.09%	1.03%
Norway	0.9605	0.9830	1.0235	1.0236	1.0061	1.02%	0.12%	0.79%	0.36%
Poland	1.0438	1.0982	1.0522	1.0523	1.0004	1.07%	0.07%	0.92%	0.23%
Portugal	0.9121	0.9872	1.0824	1.0825	1.0092	1.10%	0.04%	1.00%	0.15%
Slovakia	1.1459	1.1535	1.0067	1.0068	1.0286	0.96%	0.19%	0.66%	0.49%
Slovenia	0.9191	1.0280	1.1185	1.1186	1.0289	1.12%	0.02%	1.06%	0.09%
Spain	1.0293	1.0328	1.0033	1.0034	1.0397	0.93%	0.22%	0.62%	0.53%
Sweden	1.0034	1.0238	1.0204	1.0205	1.0084	1.01%	0.13%	0.77%	0.38%

Table 8. Continued

	Overall Edu. Quality			Gains From Trade		% Output, Free Trade, Sample Country		% Output, Free Trade, Add China & Japan	
Countries	Closed Economy (1)	Free Trade (2)	(3) = (2)/(1)	Sample Countries (4)	Adding China & Japan (5)	1% Rise in h_c^k (6)	1% Rise in h_n^k (7)	%1 Rise in h_c^k (8)	%1 Rise in h_n^k (9)
Switzerland	0.8816	1.0188	1.1557	1.1558	1.0547	1.13%	0.01%	1.09%	0.05%
United Kingdom	1.3345	1.4772	1.1070	1.1071	1.4046	0.44%	0.67%	0.16%	0.92%
United States	1.0000	1.0000	1.0000	1.0001	1.0815	0.85%	0.30%	0.57%	0.52%

Table 9 Patterns of Trade

	Dep. Var. = net exp./(imp.+exp.)		
	(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

Table 10 Robustness Exercises

	Values		Correlation Coefficients			
	theta	alpha	Cog Productivity	Non-cog Productivity	Overall Edu Quality	Output TFP
Benchmark	3.4907	1.5549	1.0000	1.0000	1.0000	1.0000
Closed-economy	3.4907	1.4706	1.0000	0.9948	0.9675	0.9940
Alt. θ value	2.0887	1.5549	0.9184	0.9983	1.0000	1.0000
Alt. θ & α values	2.0887	2.0000	0.9184	0.9884	0.9778	0.9923
Ldshp. Principal Comp.	2.2866	1.5795	0.9798	0.9252	0.9425	0.9856
Mid. Income Countries	2.2975	1.3481	0.9926	0.9942	0.9985	0.9957

Figure A1 Non-Cognitive Employment Share Over Time for the Countries with Available Data

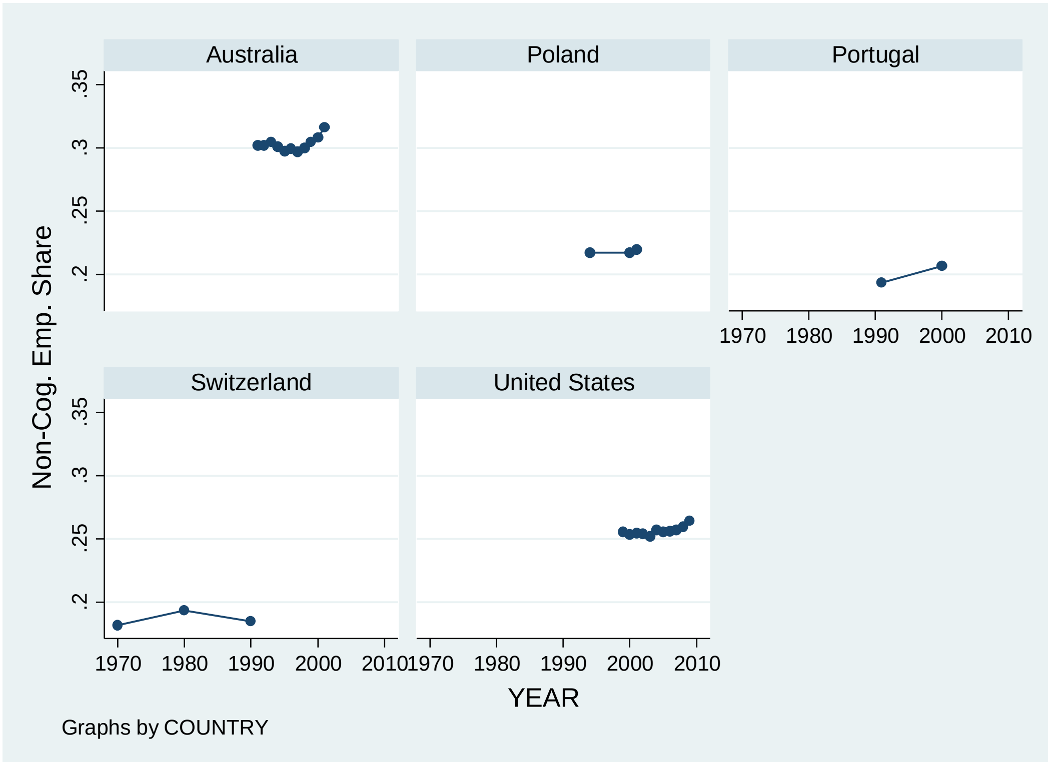

Table A1 Neal-Johnson Regressions for Alternative Measures of Non-Cognitive Skills

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 Neal-Johnson Regressions for Leadership Principal Component

VARIABLES	(1) Non-cog. Subsample	(2) Cog. Subsample	(3) Interaction
Black	-0.0740**	-0.0462**	-0.0656***
	(0.0355)	(0.0230)	(0.0191)
Hispanics	0.0276	0.0455*	0.0424**
	(0.0375)	(0.0251)	(0.0205)
Age	0.0465***	0.0286***	0.0320***
	(0.0126)	(0.00839)	(0.00687)
College			0.179***
			(0.0265)
AFQT	0.146***	0.183***	0.140***
	(0.0179)	(0.0114)	(0.0115)
$AFQT^2$	-0.0185	-0.00453	-0.0356***
	(0.0140)	(0.00971)	(0.00948)
AFQT x College			0.0582**
			(0.0245)
Non-cog. Occp.			0.131***
			(0.0163)
AFQT x Non-Cog.			-0.0541***
			(0.0160)
Constant	6.178***	6.276***	6.218***
	(0.199)	(0.133)	(0.109)
Obs. No.	973	2,237	3,210
R^2	0.130	0.167	0.217

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	HRG95	PWT_90	PWT_00
Ours	1						
HJ98	**0.5549**	1					
	0.0033						
KRC97	**0.4466**	0.8412	1				
	0.0424	0					
EK 96	**0.5858**	0.5348	0.7109	1			
	0.0171	0.0328	0.002				
HRG95	**0.6298**	0.5841	0.5394	0.068	1		
	0.0942	0.1284	0.1677	0.8729			
PWT_90	**0.6192**	0.8792	0.7401	0.6976	0.6126	1	
	0.0004	0	0.0001	0.0027	0.1064		
PWT_00	**0.6687**	0.6878	0.2565	0.3856	-0.5382	0.7089	1
	0.0001	0.0001	0.2617	0.1402	0.1688	0	
EK 02	**0.6222**	0.4159	0.4828	0.7655	0.3538	0.6114	0.4646
	0.0077	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).