Trade and Divergence in Education Systems*

Pao-Li Chang and Fali Huang[1]

Singapore Management University, Singapore

Abstract

This paper presents a theory on the endogenous choice of education policy and the two-way causal relationship between trade and education systems. A country's education system determines its talent distribution and comparative advantage; the possibility of trade by raising the returns to the sector of comparative advantage in turn induces countries to further differentiate their education systems and reinforces the initial pattern of comparative advantage. Specifically, the Nash equilibrium choice of education systems by two countries interacting strategically are necessarily more divergent than their autarky choices, and yet less than what is socially optimal for the world.

JEL Classification: F16, I20, J24.

Key Words: Education System, Talent Distribution, Comparative Advantage, Trade Pattern.

Running Head: Trade and Divergence in Education

---

*Manuscript received November 2010; revised August 2012.

[1]We thank the editor, Charles Horioka, and three anonymous referees as well as Davin Chor, Alan Deardorff, Steven Durlauf, Andy Green, Hian-Teck Hoon, Katsuya Takii, and seminar participants at the 10th World Congress of the Econometric Society in Shanghai, the 16th World Congress of International Economic Association in Beijing, the 2012 Public Economic Theory Annual Conference in Taipei, the 2012 Asia Pacific Trade Seminars in Singapore, the 2012 DEGIT in Milan, International Trade Workshop at National Taiwan University, Academia Sinica, Osaka University, and SMU for constructive comments and suggestions. Huang gratefully acknowledges the financial support of SMU Research Grant C244/MSS6E019. Please address correspondence to: Pao-Li Chang, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Phone: +65-68280830. Fax: +65-68280833. E-mail: *plchang@smu.edu.sg*.

1. Introduction

In this era of a globalized knowledge economy, the education system, by shaping a country's human capital, may exert significant influences on its comparative advantages in international trade. And conversely, a country's trade pattern and intensity may affect how its education system is run. Such interactions between education and trade can be seen in many national reviews of education policies across the world. In the US, for example, the National Commission on Excellence in Education (1983) claimed in an influential report that America is at risk: "The risk is not only that the Japanese make automobiles more efficiently than Americans . . . , or that American machine tools . . . are being displaced by German products. It is also that these developments signify a redistribution of trained capability throughout the globe. . . . If only to keep and improve on the slim competitive edge we still retain in world markets, we must dedicate ourselves to the reform of our educational system ..."[2]

In spite of the clear importance and urgency of educational reform for its role in affecting countries' comparative advantages, we are not aware of any formal analysis in the economics literature that sheds light on this matter. This paper makes a first attempt at providing a theory on the two-way interaction between education policies and trade. In particular, we show that any initial difference in education policies across countries that contributes to their corresponding comparative advantage will be further enlarged when the same countries move from autarky to trade. The intuition is that international trade increases the returns to the sector of comparative advantage, and thus induces countries to further differentiate their education systems in order to maximize gains from specialization. As a result, a small difference in initial education systems across countries, possibly due to historical or cultural variation, will be further amplified by the increase in international trade.

[2]For related discussions on the US education, see Schaub and Baker (1991), Westbury (1992), Bracey (1996), Hanushek (2002), and Dillon (2007), among others. For recent educational reforms in other countries, see for example Takayama (2007) on Japan, Mok (2005) on East Asia, and Green (1999) for general discussions on the effects of globalization on education across countries.

In this paper, we focus on an important characterization of a country's education system, which is the degree of centralization or homogenization imposed on the curriculum, and its effect on a country's talent distribution. A more homogeneous structure of curricula across schools improves the likelihood that the same set of subjects are taught and delivered in a similar manner to the students; as a result, students are more likely to acquire a common set of skills. This type of education system is often associated with a centralized curriculum council that sets and enforces a uniform curriculum via textbooks, instructional guide to teachers, periodic curriculum evaluation, or national standardized tests, as is evident in, say, Japan and some East Asian countries. The resulting pressures to conform with the uniform standards in the education system tend to improve the talent homogeneity as well as the mean ability of the pupils. In contrast, if schools do not need to follow a standardized set of curricula or performance targets, this tends to introduce more variation in student performance, as students have more freedom to pursue their individual interests and realize their potentials under a flexible curriculum; without the necessary discipline, however, the less-talented students may fail to acquire the basic set of skills. This approach is often carried out in a decentralized education system as is exemplified by, say, the US system. Thus, relatively speaking, the Japanese style of education system promotes homogeneity in the distribution of skills, while the US education style leads to more diversity. Based on the Trends in International Mathematics and Science Study (TIMSS) dataset, we find some empirical support for the above relationship in the sample of OECD countries: a more centralized curriculum structure is indeed associated with a lower skill diversity (and a higher mean ability), and the relationship is statistically significant and consistent across different years of study.

We show in theory that by altering the resulting talent distribution, different educational approaches will lead to differences in comparative advantage and trade structure in countries with otherwise identical economic constraints. Specifically, the decentralized education system, as in the US, tends to promote talent diversity in its workforce, which enhances the productivity of industries that benefit from worker skill heterogeneity, e.g., software and movie; in contrast, with a centralized education system, the Japanese workforce tends to be more homogenous in their skills, which increases the productivity of industries characterized by skill complementarity, e.g., automobile and machinery. It then follows that countries with more decentralized education systems will have a comparative advantage in the software-type industries and countries with more centralized education systems a comparative advantage in the automobile-type industries.

Given the effects of the education system on production and a country's comparative advantage, we show that the endogenous choice of education system across countries will exhibit more divergence under trade than under autarky. This is because with trade, the equilibrium price will fall in between autarky prices and thus strengthen the incentives of a country to specialize more in the sector of its comparative advantage, not only via automatic resource reallocation across sectors with a given workforce, but also through active adjustment of education policies to reshape the composition of its workforce and the position of its production possibility frontier.

In particular, we identify the choice of education system under autarky by individual countries who may differ in their utility costs of implementing homogeneous curricula, but are otherwise identical in economic constraints and initial talent distributions. We then characterize the choice of education system under trade that is socially optimal for the world as a whole. This is compared with the noncooperative choice in the Nash equilibrium where each country maximizes its own welfare taking into account the terms-of-trade effect of its education policies. It is shown that the difference in education systems across countries under trade is larger than under autarky. However, the cross-country difference in education systems under the Nash equilibrium is less than what is socially optimal for the world. The intuition is that the incentive to specialize through more divergent education systems is weakened in each country by the accompanying terms-of-trade loss, which on the other hand cancels out for all countries in the world welfare calculation.

Our paper is connected with some existing strands of literatures on talent (human capital), trade and education. These literatures have contributed to our understanding of some aspects of the triangular relationship among talent, trade and education, but most of them have not systematically linked the three and studied their causal relationship in an integrated framework as we do in this paper. For example, in the literature pioneered by Grossman and Maggi (2000), the studies take talent distributions across countries as given, and analyze how talent composition affects nations' comparative advantage and trade pattern. Research in this area includes Grossman (2004), Bougheas and Riezman (2007), Ohnsorge and Trefler (2007), and Bombardini et al. (2012). In another strand of literature, researchers study how trade may affect human capital formulation, taking the education system as given. This includes, for example, Findlay and Kierzkowski (1983), and Bond et al. (2003). This line of work thus assumes away the possibility to change the human-capital production function via education policies. Third, there is an extensive literature on how the education system affects talent composition. This includes, among others, Bénabou (1996), Epple and Romano (1998), Fernández and Rogerson (1998), and Takii and Tanaka (2009). They examine, for example, how different education regimes (public versus private, ability tracking versus ability pooling) affect the dispersion of skills and aggregate output. These studies, however, often focus on *closed economies* and thus ignore potential ramifications of trade. Last but not least, there is a small strand of literature that explicitly links education with trade, but these studies focus on very different policy issues from ours; see, for example, Kim and Kim (2000), Falkinger and Grossmann (2005), and Bougheas et al. (2011).[3]

Our paper integrates these literatures and creates a unified conceptual framework. Particularly in our framework, we take into account how the education system shapes a country's talent com-

[3] Kim and Kim (2000) assume that school education enhances general human capital versus industry-specific skills, which together with international trade allows workers to move easily to the fastest-growing industry and hence facilitates economic growth. In an oligarchy society where landed elites have more political power, Falkinger and Grossmann (2005) show that public education investment, conducive to industrialization, is typically lower in an open economy than in autarky; this is similar in spirit to our result that an open economy may adopt a more extreme education approach than in autarky. Bougheas et al. (2011) also analyze the possible effect of trade on a country's education policy. They formulate the education policy, however, as a choice by a small open economy of whether to move up or down the skill chain, taking the terms of trade as given; this is in contrast with our focus on the education policy's role in affecting the diversity of human capital and the optimal allocation of talent across sectors. More importantly, our analysis takes into account the consequence of education policies on the equilibrium trade prices and patterns, which enables us to study the interactions of education policies across countries and their endogenous divergence.

position and thus its comparative advantage in trade. Most importantly, we also take into account the reverse causality of trade on the education system. Viewed from another perspective, our paper contributes to the literature in two ways: First, it demonstrates the possibility of education policy as a new source of comparative advantage in trade, and second, it emphasizes how trade can in turn affect a country's institutions such as education systems. To our knowledge, this endogenous determination of both education system and trade pattern as an equilibrium outcome has not been examined in the literature.

The education literature has only recently begun to assess the implications of globalization on education policies (Green, 1997; Burbules and Torres, 2000; Mok, 2005). Though "there is considerable convergence at the level of policy rhetoric and general policy objectives, there is less evidence of any systematic convergence at the level of structures and processes in different countries" (Green, 1999). This is consistent with our finding in this paper that differences in education systems across countries may be a persistent pattern reinforced by trade. To our knowledge, this result is new to the education-related literature and may provide a fresh perspective on how education policies are formed.

This paper is organized as follows. The model is set up in Section 2. The endogenous choice of education system is analyzed in Section 3. Section 4 discusses modeling choices, theoretical extensions, and possible empirical strategies to test the implications of our model. Section 5 concludes. Proofs for all lemmas and propositions are shown in Appendix A.1. The details on the empirical analysis of the relationship between curriculum centralization and talent distribution are provided in Appendix A.2. Appendix A.3 documents the data used. A technical appendix is available online[4] that provides proofs for alternative theoretical setups discussed in Section 4.

[4]http://www.mysmu.edu/faculty/plchang/papers/trade-edu-app.pdf

2. The Basic Model

2.1. The Education System

Suppose there is a unit measure of a continuum of pupils, whose innate abilities t_0 are not individually observable, but follow a distribution $G(\cdot)$ with support $[t_{l0}, t_{h0}] \subset (0, +\infty)$ and an initial mean ability $\bar{t}_0 \equiv \int_{t_{l0}}^{t_{h0}} t_0 \, \mathrm{d}G(t_0)$. All pupils have to go through an education system that is characterized by a parameter $\delta \in [0, 1]$, which indicates the degree of curriculum centralization imposed on each student. Specifically, a pupil with an innate ability t_0 will acquire a skill level t at graduation such that

$$t = \underbrace{[(1-\delta)t_0 + \delta\bar{t}_0]}_{\text{diversity effect}} \underbrace{\gamma^{\delta}}_{\text{mean effect}}, \tag{1}$$

where $\gamma > 1$. Each pupil's adulthood skill t is publicly observed. Let $\Phi(t;\delta)$ indicate the adulthood ability distribution. Note that $\Phi(t;\delta) \equiv G(\frac{\gamma^{-\delta}t - \delta\bar{t}_0}{1-\delta})$, which has a support $[t_l, t_h]$ with $t_h - t_l = \gamma^{\delta}(1-\delta)(t_{h0} - t_{l0})$. We will often write the adulthood distribution as $\Phi(t)$ to simplify presentation, bearing in mind that it depends on δ.

Thus, an education system with a higher δ will push all students' skills toward the middle and reduce the skill diversity, so there is a 'diversity effect' as reflected in (1). On top of that, a higher δ will shift the whole talent distribution rightward and increase the mean ability of the cohort (by a factor of γ^{δ}), exerting a 'mean effect' as well. The functional form adopted in (1) is broadly consistent with empirical observations. As shown in Appendix A.2, based on the Trends in International Mathematics and Science Study (TIMSS) dataset, we find empirical support for both a diversity and a mean effect in the sample of OECD countries. That is, a more centralized curriculum structure is indeed associated with a lower skill diversity and a higher mean ability; the relationship is statistically significant and consistent across different years of study (see Figure 3 in Appendix A.2).

This simple model of education attempts to capture the necessary tension between equipping all students with a common set of knowledge versus promoting talent diversity. The former goal is usually better achieved with a higher δ as exemplified by a more standardized set of curricula, where students have to go through the same subjects and numerous exams that test whether they have met required standards before they can go to the next level of study. These efforts tend to improve the mean ability and talent homogeneity of the pupils. The time and efforts committed to following the same curriculum and activities, however, often discourage the more talented students from exploring and acquiring new knowledge in their own ways, and hence may reduce the creativity component of human capital (Mayer et al., 1991). The opposite is true for pursuing the second goal, where a lower δ as characterized by flexible curricula and lenient standards leaves more freedom for individual exploration and hence may preserve more talent diversity; without the necessary discipline, however, the students may fail to acquire some basic set of skills, resulting in a lower average skill level. Education systems may vary across countries in their orientation toward these two goals (Cummings, 1999). Among industrial countries, Japan and the US are arguably two prominent examples at the opposite ends of the spectrum in terms of curriculum centralization (as indicated by Figure 3). In what follows, we will often use the US/Japan contrasting cases as rhetorical examples to illustrate the theory and predictions.

The innate or initial abilities t_0 are taken to be unobservable to educators or even to the student herself/himself. This is to underlie the difficulty of the education system to correctly evaluate each student's true talent and to tailor the teaching method as indicated by δ according to each individual's ability. For example, if the initial abilities were fully observable, the education resources would be best utilized to raise the skills of the less talented students with a more disciplined education method but those of the more talented students with a more flexible curriculum. Alternatively, if the innate abilities t_0 were known to the students, a menu of education methods could be offered such that students self select into different schemes.[5] The assumption of unobservability of initial abilities highlights the inevitable tradeoff of positive and negative effects of choosing a particular education style on the human capital of a country.

[5] Streaming by student ability, however, has other undesirable effects (for example, it may unduly discourage student self-esteem and erode social cohesion), and is unlikely to be the ideal policy for the majority of students in a country.

2.2. The Economy

Technology. The production technology follows Grossman and Maggi's (2000), but as discussed in Section 4, the paper's main prediction will hold in various alternative settings of production technology.

There are two industries, A (auto) and S (software). Production in each sector requires teamwork of two workers. Output by a pair of workers in sector $i \in \{A, S\}$ is $F^i(t_1, t_2)$, where t_j denotes the skill level of the worker performing task $j \in \{1, 2\}$. F^i is strictly increasing in t_j, symmetric, and has constant returns to talent. Sector A is characterized by a supermodular production technology such that $F_{12}^A > 0$; i.e., the workers' talents are 'complementary' in the production of good A. On the other hand, the production technology of industry S is submodular such that $F_{12}^S < 0$; i.e., the marginal product of talent is higher the less capable is one's partner. In this sense, workers' talents are 'substitutable' in sector S. Given constant returns to talent, this implies that there are decreasing returns to 'individual' talent in auto production ($F_{jj}^A < 0$) but increasing returns to individual talent in software production ($F_{jj}^S > 0$).

Given this contrast in the nature of production technology, competitive pressures will encourage self-matching in sector A and maximal cross-matching in sector S.

Lemma 1 (Grossman and Maggi (2000), Lemmas 1 and 3) *For a given output target Y_S, the output Y_A is maximized by: (i) allocating all workers with talents $t \leq \hat{t}$ and all workers with talents $t \geq m(\hat{t})$ to sector S, where $m(t)$ is defined implicitly by $\Phi[m(t)] = 1 - \Phi(t)$ and $\hat{t}$ solves $Y_S = \int_{t_l}^{\hat{t}} F^S[t, m(t)]\, d\Phi(t)$; (ii) allocating the remaining workers with $t \in [\hat{t}, m(\hat{t})]$ to sector A and matching worker abilities in each pair such that $t_1 = t_2$ in all teams; i.e., $Y_A = \frac{\lambda_A}{2} \int_{\hat{t}}^{m(\hat{t})} t\, d\Phi(t)$, where $\lambda_A = F^A(1, 1)$.*

As in Grossman and Maggi (2000), we assume symmetric talent distributions. This implies that $m(t) = 2\bar{t} - t$, where $\bar{t}$ is the adulthood mean ability level. Let $p \equiv p_S/p_A$ denote the relative price of software and $MRT \equiv -\frac{\partial Y_A/\partial \hat{t}}{\partial Y_S/\partial \hat{t}}$ the marginal rate of transformation of the production possibility frontier (PPF). The competitive profit condition $p = MRT$ pins down the equilibrium talent allocation $\hat{t}$. It is straightforward to verify that price movement has the following effects on talent allocation and industrial outputs.

Lemma 2 *As the relative price of software p increases,*
(i) marginal workers move to the software sector ($\frac{d\hat{t}}{dp} > 0$)*;*
(ii) the software output increases and the auto output decreases ($\frac{\partial Y_S}{\partial p} > 0 > \frac{\partial Y_A}{\partial p}$)*.*

Effects of Education Policy on PPF. The results above are conditional on a given talent distribution and hence a given PPF. A change in the education policy, however, will affect the composition of the talent pool and hence the PPF.

Define $Y_A^b(\delta, p)$ and $Y_S^b(\delta, p)$ as the auto and software outputs that exclude the mean effect of the education system. Similarly, label all corresponding variables excluding the mean effect with a decoration b (b for benchmark).

Lemma 3 *Excluding the mean effect, an increase in* δ *(the degree of curriculum centralization in the education system) will increase the output of autos but decrease the output of software at any given relative price level, that is,* $\frac{\partial Y_A^b(\delta,p)}{\partial \delta} > 0$ *and* $\frac{\partial Y_S^b(\delta,p)}{\partial \delta} < 0$ *at any given p.*

In essence, excluding the mean effect, a more centralized education policy (a higher δ) pushes the talent distribution toward the middle, increasing the proportion of population working in the auto sector. The higher density of population working in the auto sector contributes to the higher auto output. Not surprisingly, the lower density of workers engaged in the software sector has a negative effect on its output. In addition, for every pair of workers remaining in the software sector, the more talented worker's ability is lowered while the less talented worker's ability is lifted with a higher δ, but the negative effect of the former on the team's output dominates the positive effect of the latter (as tasks are symmetric and there are increasing returns to individual talent in software production). Thus, software output decreases on both accounts (a lower density of software workers and a lower average output by the remaining workers).[6]

[6] The results of Lemma 3, which is based on differential analysis, are consistent with Proposition 4 of Grossman and

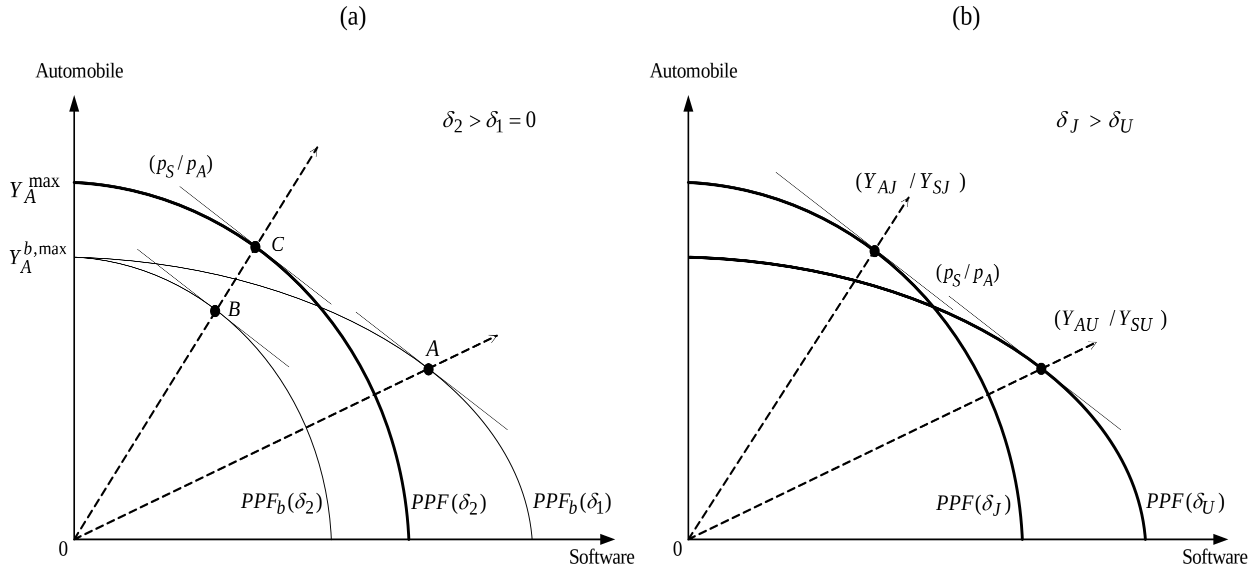


Figure 1: Production Possibility Frontiers and Education Systems

Figure 1(a) illustrates the effect of the education system on the PPF excluding the mean effect (indicated with a subscript b). Start with the case where $\delta = 0$. For any given relative price p, an increase in δ will raise the output of autos and lower that of software (illustrated by the shift from point A with $\delta_1 = 0$ to point B with $\delta_2 > \delta_1$). Note that when the whole population is employed in the auto sector, the economy attains its maximum auto output $Y_A^{b,max} \equiv \frac{\lambda_A}{2}\,\bar{t}_0$, which is independent of the education policy δ. On the other hand, the software sector's maximum output $Y_S^{b,max} \equiv \int_{t_{lb}}^{\bar{t}_b} F^S(t_b, 2\bar{t}_b - t_b)\mathrm{d}\Phi_b(t_b)$ is lower when δ is higher: although when all workers are engaged in the software sector, an increase in δ will not induce any further talent reallocation (the density effect disappears), it reduces the skill gap in each pair of workers, which lowers the average productivity in the software sector (the negative average effect remains).

This suggests that, based on the diversity effect alone, there is a pecking order among education policies, with a lower δ strictly dominating a higher δ (since the PPF corresponding to a more diverse talent distribution lies strictly outside the PPF corresponding to a less diverse one). Thus, in order

Maggi (2000) using discrete comparisons of two talent distributions. Note that we now index the talent distribution by a continuous education policy δ, and as a result, are able to provide exact analytical expressions for the changes in outputs (see Appendix A.1).

to arrive at a structure with inherent tradeoffs in setting education policies, a higher δ must bring about positive mean effects; in other words, $\gamma > 1$ is a necessary premise. Given the positive mean effect, the PPF shifts outward in a parallel fashion by the factor γ^δ, basically due to constant returns to overall talent of the production technologies. This is illustrated by the shift from point B to point C in Figure 1(a).

Lemma 4 *The mean effect shifts the PPF in a parallel fashion by the factor γ^δ; i.e., $Y_A(\delta, p) = \gamma^\delta Y_A^b(\delta, p)$ and $Y_S(\delta, p) = \gamma^\delta Y_S^b(\delta, p)$ for any given δ and p.*

However, as is clear from Figure 1(a), too large a mean effect (too large γ) will also eliminate the endogenous tradeoff of policy choice: when the positive mean effect of raising δ dominates the negative diversity effect on software output, it is Pareto superior to raise δ since a higher δ also raises the auto output. Thus, a cap $\widetilde{\gamma}$ on the mean effect parameter is necessary as well.

Definition 1 $\widetilde{\gamma} \equiv \min_{\{\delta,p\}} \gamma(\delta, p)$, *where* $\gamma(\delta, p) = \exp\left(-\frac{1}{Y_S^b(\delta,p)} \frac{\partial Y_S^b(\delta,p)}{\partial \delta}\right) > 1.$

Lemma 5 *For $1 < \gamma < \widetilde{\gamma}$, an increase in δ will increase the output of autos but decrease the output of software at any given relative price level, i.e., $\frac{\partial Y_A(\delta,p)}{\partial \delta} > 0$ and $\frac{\partial Y_S(\delta,p)}{\partial \delta} < 0$ at any given p.*

Figure 1(b) sums up our setup of education system and production technology. The PPFs corresponding to two different education systems with $\delta_J > \delta_U$ are illustrated. In essence, an economy with a more centralized education system has a PPF that is relatively skewed toward the auto output axis: with a higher δ, the maximum auto output $Y_A^{\max}$ is higher, while the maximum software output $Y_S^{\max}$ is lower. Given any relative price p, an economy with a higher δ will then produce relatively more autos and the other economy relatively more software. This framework has the desirable property that PPFs corresponding to different δ's cross each other, and thus, there are inherent tradeoffs in choosing one talent distribution (and education policy) against the others.

Preferences. Individuals have identical homothetic preferences represented by the Cobb-Douglas utility function $u(c_A, c_S) = c_A^\beta c_S^{1-\beta}$, where $0 < \beta < 1$, and c_A and c_S denote an individual's consumption of cars and software, respectively. The budget constraint is $c_A + pc_S = w(t)$, where $w(t)$ is an individual's wage income (the only source of income in the model) for a given talent level t.[7] Given the relative price and the income, the optimal consumption choices are thus $c_A(p, w(t)) = \beta w(t)$ and $c_S(p, w(t)) = \frac{(1-\beta)w(t)}{p}$, which lead to the indirect utility function

$$v(p, w(t)) = \beta^{\beta}(1-\beta)^{1-\beta}p^{-(1-\beta)}w(t). \tag{2}$$

2.3. Equilibrium Analysis

Autarky Equilibrium. In the autarky equilibrium, the domestic markets of both auto and software will clear so that the ratio of total supplies of cars and software is equal to the ratio of their total demands. That is,

$$\frac{Y_A(\delta, p^a)}{p^a Y_S(\delta, p^a)} = \frac{\int c_A(p^a, w(t))\mathrm{d}\Phi(t)}{p^a \int c_S(p^a, w(t))\mathrm{d}\Phi(t)} = \frac{\beta}{1-\beta}, \tag{3}$$

where p^a is the equilibrium price under autarky, and the second equality holds because consumption expenditure is proportional to wage incomes. This implies an autarky equilibrium price

$$p^a = \frac{1-\beta}{\beta}\frac{Y_A(\delta, p^a)}{Y_S(\delta, p^a)}, \tag{4}$$

which is unique because the LHS of (4) is strictly increasing in p^a, while the RHS strictly decreases in p^a by Lemma 2 for any given δ. Thus, (4) defines the autarky equilibrium price as a function of the education system δ.

Lemma 6 *The autarky equilibrium price $p^a(\delta)$ increases with the degree of curriculum centralization in the education system, that is, $\frac{dp^a}{d\delta} > 0$.*

Lemma 6 shows that the more centralized the education system is, the higher the autarky relative price for software in equilibrium. The intuition is obvious: since the relative supply of software is lower when δ is higher and as the preferences are homothetic, a closed economy with a higher δ will have a higher relative price for software.

[7] We refer readers to Grossman and Maggi (2000) for the specific wage structure consistent with the production technology. Note, however, the different numeraire good used in the current paper from Grossman and Maggi (2000).

Free Trade Equilibrium. Suppose that a world consists of two representative countries, Japan (J) and the US (U). They have the same economic structure as described above but different education systems ($\delta_J > \delta_U$), the endogenous choice of which will be analyzed in the next section. That is, Japan's education system is more centralized than the US's, and as a result, the adult talent distribution $\Phi(t)$ is more homogenous in Japan than in the US, though the talent distribution among children $G(t_0)$ is identical to begin with in the two countries.

Given the different education systems ($\delta_J > \delta_U$), Lemma 6 suggests that Japan will have a higher autarky relative price for software than the US ($p_J^a > p_U^a$). Thus, with the possibility of trade, Japan (the US) will have a comparative advantage in cars (software) and will export cars (software).

By similar arguments as in the autarky case, we know that the free trade equilibrium price p^f is uniquely determined by

$$p^f = \frac{1-\beta}{\beta} \frac{Y_A(\delta_J, p^f) + Y_A(\delta_U, p^f)}{Y_S(\delta_J, p^f) + Y_S(\delta_U, p^f)}, \tag{5}$$

where $Y_A(\delta_J, p^f)$ and $Y_A(\delta_U, p^f)$ are the auto outputs in Japan and the US, respectively, and $Y_S(\delta_J, p^f)$ and $Y_S(\delta_U, p^f)$ are the software outputs in the two countries. It follows that $p_J^a > p^f > p_U^a$ holds for any given education systems in the two countries with $\delta_J > \delta_U$.

Lemma 7 *The free trade equilibrium price $p^f(\delta_J, \delta_U)$ increases with curriculum centralization in the education system of either country, that is,* $\frac{\partial p^f}{\partial \delta_J} > 0$ *and* $\frac{\partial p^f}{\partial \delta_U} > 0$.

The free trade equilibrium price p^f is higher when either δ_J or δ_U increases, as either change will increase the relative supply of cars in the world market, pushing up the relative price of software.

3. Endogenous Education System

In the previous section, we introduce education policies and demonstrate that these policies have a direct effect on the distribution of talent and thus a country's comparative advantage. We now analyze the endogenous choice of such education systems and show that the same pattern of trade in turn has important bearings on the optimal design of education systems. We begin the analysis by characterizing the socially optimal education system under autarky. Given different choices of education systems under autarky, we then analyze how countries will react to trade by changing their optimal choice of education system.

A higher degree of curriculum centralization in the education system usually corresponds to more rules and regulations imposed on the curriculum, the textbooks, the allocation of school hours, the frequency of tests, and the monitoring of student performance. Requirements such as these may translate into disutility for each student undergoing the system. The education literatures suggest that there are two broad types of utility cost on students that is associated with centralization in education. The first type is due to cultural diversity and the inability of centralized curricula to cater to local variations. For example, Weiler (1990) observes that "Except in very small or culturally very homogeneous societies, most countries show considerable variation across regions, communities, and language groups in terms of cultural and social frameworks of learning. The frames of reference for the study of history, botany, social studies, and other fields vary obviously and significantly between southern and northern Italy, the state of Alabama and the state of California, or Bavaria and Berlin." With differences like these, centralization in curriculum and method of teaching would then produce a mismatch between a student's specific learning environment (which tends to reflect local and regional cultures and traditions) on one hand, and a centrally defined learning agenda or curriculum on the other (Weiler, 1990). Such a mismatch is likely to make learning less enjoyable for students and impose a utility cost, since part of the things they have to learn in a standardized curriculum may not be familiar to them through their daily experiences and hence require more effort from them to understand and digest.

The second type of utility cost on students arises from having to take high-stakes exams that are often associated with centralization in education. As observed by Clune (1993) and shared by others, "high-stakes student examinations are a key component, perhaps the cornerstone, of the centralized version of systemic educational policy." Because the gateway to each level of education is guarded by examinations, the rewards for success and the penalties for failure in these examinations are substantial (Bray, 2007). "Exposing students to high-stakes assessments in which there must be winners and losers, in a compulsory education system in which the students have effectively no control or voice, is likely to produce detrimental effects" (Gregory and Clarke, 2003). For example, as cited in their paper, one in three primary Singaporean children finds life not worth living; nearly four out of five spend as many as three hours studying after school; and seven out of ten receive extra classes after school; in England, the introduction in the early 1990s of high-stakes exams brings pressures to both teachers and students, since the test results are very public and powerful labeling tools at many levels. In addition to huge pressures directly caused by exams and tests, the learning environment also becomes less enjoyable, because teachers tend to teach to the tests by narrowing the coverage of curriculum, shifting from holistic teaching to drilling through repetitive exercises focusing on exam questions (Smith et al., 1991; LeTendre, 1999).

The above discussions suggest that the utility cost is likely to increase with the degree of curriculum centralization in a given country, and the marginal increase may also differ across countries. For example, if a country is more homogeneous in terms of cultural composition, it may find it less costly to set and implement a common set of curriculum and standards, as pupils in the system may share similar ethnic, linguistic, behavioral or religious traits. To capture the above observations, we assume that imposing curriculum centralization at a level δ entails a utility cost of $K_j(\delta) \equiv k_j\delta$, which is borne by each individual pupil as they go through the education system in country j, where $j = \{J, U\}$, and $0 < k_J < k_U$. That is, the Japanese-type society J, given its relatively homogeneous cultural composition, is assumed to incur a lower marginal utility cost in conforming to a given degree of uniformity in education standards than the US-type society U with a relatively diverse cultural composition. As discussed in Section 4, the main results of the model shown below can be obtained under alternative settings without such utility cost or any initial differences across countries, albeit at the cost of more restrictions on the production technologies.

3.1. Education Systems under Autarky

Given the indirect utility function in (2), the net aggregate welfare of a country implementing an education system δ is

$$
\begin{aligned}
U(\delta) &= \int_{t_l}^{t_h} \left(v\left(p, w(t)\right) - k\delta\right) \mathrm{d}\Phi(t) \\
&= \beta^{\beta}(1-\beta)^{1-\beta} p^{-(1-\beta)} \int_{t_l}^{t_h} w(t)\mathrm{d}\Phi(t) - k\delta \\
&= \beta^{\beta}(1-\beta)^{1-\beta} p^{-(1-\beta)} \left(Y_A + pY_S\right) - k\delta, \qquad (6)
\end{aligned}
$$

where the second equality follows by plugging in the indirect utility function in (2) and also because the measure of population is one. The last equality holds since the total wage income should be equal to the total value of production in the economy due to perfect competition.

Let δ^a denote the optimal education choice under autarky that maximizes $U(\delta)$. The following first order condition (FOC) holds at δ^a:

$$
\begin{aligned}
\frac{\mathrm{d}U(\delta)}{\mathrm{d}\delta} &= \beta^{\beta}(1-\beta)^{1-\beta}(p^a)^{-(1-\beta)} \left[\frac{\partial Y_A}{\partial \delta} + p^a \frac{\partial Y_S}{\partial \delta} + \left(\frac{\partial Y_A}{\partial p^a} + p^a \frac{\partial Y_S}{\partial p^a}\right)\frac{\mathrm{d}p^a}{\mathrm{d}\delta} \right. \\
&\quad \left. + \left(-(1-\beta)(p^a)^{-1}Y_A + \beta Y_S\right)\frac{\mathrm{d}p^a}{\mathrm{d}\delta}\right] - k \\
&= \beta^{\beta}(1-\beta)^{1-\beta}(p^a)^{-(1-\beta)} \left(\frac{\partial Y_A}{\partial \delta} + p^a \frac{\partial Y_S}{\partial \delta}\right) - k = 0, \qquad (7)
\end{aligned}
$$

where the second equality follows, first because $\frac{\partial Y_A}{\partial p} + p\frac{\partial Y_S}{\partial p} = \left(\frac{\partial Y_A}{\partial \hat{t}} + p\frac{\partial Y_S}{\partial \hat{t}}\right)\frac{\mathrm{d}\hat{t}}{\mathrm{d}p} = 0$ by the competitive profit condition $p = MRT$, and second, by the autarky equilibrium condition (4). At the optimal choice, the second order condition (SOC) $\frac{\mathrm{d}^2U(\delta^a)}{\mathrm{d}\delta^2} < 0$ also holds. Thus, with $k_J < k_U$, it follows that the optimal degree of curriculum centralization will be higher in Japan than in the US under autarky, i.e., $\delta_J^a > \delta_U^a$.

3.2. Education Systems under Trade

Small Open Economy. We start with the optimal education system that would be chosen by a small open economy, that is, when a country takes the world price as given and does not take into account the effect of its education system on the world price. This analysis thus assumes away the terms-of-trade consideration and serves as a useful benchmark. As will become clear, many useful insights can be gained from comparing this scenario with others.

Proposition 1 *The optimal education system of a small open economy, $\delta^o(p)$, decreases with the given trade price, p.*

Proposition 1 implies that an increase in the relative price of software, p, will induce a small open economy to adjust downward the degree of curriculum centralization in its education system, shifting its economic structure in favor of the software sector. The opposite is true when p decreases.

As it turns out, the FOC for a small open economy is identical to (7), except that the price is taken as given (a parameter) and need not clear the domestic market. This suggests that the autarky optimal education system δ^a coincides with the optimal education system δ^o under trade if the given trade price happens to be the same as the autarky price. Recall that as countries move from autarky to trade, the free trade price is lower than the autarky price of the country with a more centralized education system but higher than that of the other country. In view of this, an important implication of Proposition 1 is that as countries move from autarky to trade, a small open economy initially having a more centralized education system will further raise its degree of curriculum centralization (as the relative price of software falls after trade compared to its autarky level) and a small open economy initially having a more decentralized education system will further lower its degree of curriculum centralization (as the relative price of software rises after trade).

World Optimal Choice. What would be the optimal education systems for the two countries if chosen by a world social planner that maximizes the joint welfare of the two countries? The objective function in this case is:

$$(8) \qquad U_w(\delta_J, \delta_U) = \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}\left(Y_{AJ} + pY_{SJ} + Y_{AU} + pY_{SU}\right) - k_J\delta_J - k_U\delta_U,$$

the derivation of which is similar to that of (6), with Y_{Aj} and Y_{Sj} indicating the auto and software outputs by country $j = \{J, U\}$. Given (8), the FOC for the optimal education system for Japan from the world social planner's viewpoint is:

$$\begin{aligned}\frac{\partial U_w(\delta_J,\delta_U)}{\partial \delta_J} &= \beta^{\beta}(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left[\frac{\partial Y_{AJ}}{\partial \delta_J}+p^f\frac{\partial Y_{SJ}}{\partial \delta_J}\right.\\ &\quad +\left(\frac{\partial Y_{AJ}}{\partial p^f}+p^f\frac{\partial Y_{SJ}}{\partial p^f}+Y_{SJ}+\frac{\partial Y_{AU}}{\partial p^f}+p^f\frac{\partial Y_{SU}}{\partial p^f}+Y_{SU}\right)\frac{\partial p^f}{\partial \delta_J}\\ &\quad \left.-(1-\beta)(p^f)^{-1}(Y_{AJ}+p^fY_{SJ}+Y_{AU}+p^fY_{SU})\frac{\partial p^f}{\partial \delta_J}\right]-k_J\\ &= \beta^{\beta}(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left(\frac{\partial Y_{AJ}}{\partial \delta_J}+p^f\frac{\partial Y_{SJ}}{\partial \delta_J}\right)-k_J=0, \qquad (9)\end{aligned}$$

where to obtain the second equality, we have used the competitive profit condition $p = MRT$ for both countries and the free trade equilibrium condition (5). Similarly, the FOC for the optimal education system for the US is:

$$\frac{\partial U_w(\delta_J,\delta_U)}{\partial \delta_U}=\beta^{\beta}(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left(\frac{\partial Y_{AU}}{\partial \delta_U}+p^f\frac{\partial Y_{SU}}{\partial \delta_U}\right)-k_U=0. \qquad (10)$$

Let $\delta_J^w(\delta_U)$ denote the solution of δ_J to (9) given δ_U, and similarly let $\delta_U^w(\delta_J)$ denote the solution of δ_U to (10) given δ_J. It is straightforward to verify that $\frac{\partial^2 U_w(\delta_J,\delta_U)}{\partial\delta_J\partial\delta_U}<0$; thus, it follows that $\frac{\mathrm{d}\delta_J^w(\delta_U)}{\mathrm{d}\delta_U}<0$ and $\frac{\mathrm{d}\delta_U^w(\delta_J)}{\mathrm{d}\delta_J}<0$. These two FOCs are illustrated in Figure 2. The optimal choice of education systems (δ_J^w,δ_U^w) by the world social planner corresponds to the intersection W of the two schedules $\delta_J^w(\delta_U)$ and $\delta_U^w(\delta_J)$. The assumption $k_J<k_U$ is reflected by the fact that the schedule $\delta_J^w(\delta_U)$ is further away from the origin than the schedule $\delta_U^w(\delta_J)$, and as a result, the optimal degree of curriculum centralization for Japan, δ_J^w, is higher than that for the US, δ_U^w.

Proposition 2 *The optimal choice of education systems* (δ_J^w,δ_U^w) *by a world social planner that maximizes the world welfare under trade coincides with the noncooperative equilibrium choice of education systems by individual countries if they behave as price takers in the world market. Relative to autarky, the difference in education systems is further enlarged after trade in the world socially optimal outcome:* $\delta_J^w>\delta_J^a>\delta_U^a>\delta_U^w$.

The intuition for the first result is that the terms-of-trade considerations in setting δ by countries who perceive their market powers are neutralized in a world social planner's problem, since one country's terms-of-trade gain is the other country's terms-of-trade loss. This can be seen in the

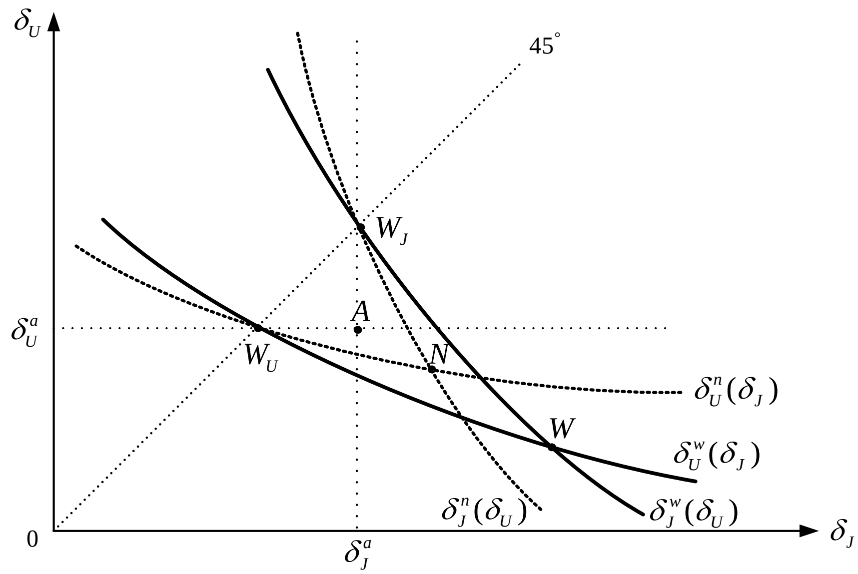


Figure 2: Endogenous Choices of Education System

derivations of (9), where the effects of δ_J on p^f and the effects of p^f on the joint income and welfare of the two countries are eliminated in the final expression. The only things that matter are the direct effect of δ_J on Japan's own production choice and that of δ_U on the US's. Thus, the FOCs and the optimal choice of education systems by a world social planner turn out to be the same as the best response functions and the noncooperative equilibrium choice of individual countries if they behave as price takers.

To see the second result, note that δ_J at point W_J corresponds to Japan's optimal education system under autarky. The intuition is that W_J is the world socially optimal choice if both countries have Japan's disutility factor k_J; but then, this corresponds to the optimal choice of Japan under autarky, as p^f in this hypothetical scenario is identical to Japan's autarky price p_J^a. Similarly, δ_U at point W_U corresponds to the optimal choice of the US under autarky. Thus, the unilateral optimal choices under autarky in the two countries correspond to point A in Figure 2. It lies to the northwest of the world socially optimal choice under trade at point W. Thus, from the world's perspective, it is socially optimal to further enlarge the autarky difference in the education systems between Japan and the US to reinforce their initial pattern of comparative advantage and to maximize the gains from trade.

The above result has some interesting implications. With endogenous education policies (and PPFs), the output response to the possibility of trade becomes more elastic and the potential gains from trade are bigger than the classical trade theories (with given PPFs) would suggest. Not only does the world aggregate production increase as individual countries reallocate more productive resources to their sector of comparative advantage (corresponding to a movement along the given PPF), but it is further increased as individual countries restructure their education policies (and PPFs) to the advantage of the sector.

Nash Equilibrium Choice. If countries choose education systems unilaterally (as is likely the case in reality) and take into consideration the terms-of-trade effect of their education policies, the objective function is $U_J(\delta_J;\delta_U) = \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}(Y_{AJ}+pY_{SJ}) - k_J\delta_J$ for Japan, with p satisfying the free trade equilibrium condition (5). The resulting FOC for Japan's best response function $\delta_J^n(\delta_U)$ is

$$\begin{aligned} \frac{\partial U_J(\delta_J;\delta_U)}{\partial \delta_J} &= \beta^\beta(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left(\frac{\partial Y_{AJ}}{\partial \delta_J} + p^f\frac{\partial Y_{SJ}}{\partial \delta_J}\right. \\ &+ \left.\underbrace{\beta\frac{Y_{SJ}Y_{AU}-Y_{SU}Y_{AJ}}{Y_{AJ}+Y_{AU}}\frac{\partial p^f}{\partial \delta_J}}_{\text{TOT effect } <0 \text{ when } \delta_J>\delta_U}\right) - k_J = 0, \end{aligned} \tag{11}$$

which differs from the FOC for the world optimal choice $\delta_J^w(\delta_U)$ in (9) by a terms-of-trade (TOT) effect. The TOT effect is negative when $\delta_J > \delta_U$, because in this case $Y_{AJ} > Y_{AU}$ and $Y_{SU} > Y_{SJ}$ by Lemma 5.

Recall that in the optimization problem of either the individual autarky country or the world social planner, the effect of the endogenous price change on the autarky welfare or the joint world welfare is zero. A software price increase has a positive income effect scaled by the software output. At the same time, it entails a negative consumption effect. In autarky, the two effects offset each other, as production equals consumption. In the joint welfare calculation, the two effects also offset each other, as the two countries' joint production equals their joint consumption.

This is not the case when countries maximize individual welfare under trade. For an auto-exporting country, it produces less software than it consumes; thus, a software price increase will lead to a smaller income gain than consumption loss, which leads to an overall terms-of-trade loss. Hence, whenever $\delta_J > \delta_U$ holds, Japan as an exporter of cars would not want to raise δ_J as much as it would as a price taker (or in the world optimal outcome), since a higher δ_J, by increasing the auto output, depresses the world price of cars which it exports, and thus hurts its terms of trade.

Similarly, the FOC for the US's best response function $\delta_U^n(\delta_J)$ is:

$$\begin{aligned}\frac{\partial U_U(\delta_U;\delta_J)}{\partial \delta_U} &= \beta^\beta(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left(\frac{\partial Y_{AU}}{\partial \delta_U} + p^f\frac{\partial Y_{SU}}{\partial \delta_U}\right.\\ &+ \left.\underbrace{\beta\frac{Y_{SU}Y_{AJ}-Y_{SJ}Y_{AU}}{Y_{AJ}+Y_{AU}}\frac{\partial p^f}{\partial \delta_U}}_{\text{TOT effect }>0\text{ when }\delta_J>\delta_U}\right) - k_U = 0, \end{aligned} \tag{12}$$

where the TOT effect is equivalent in magnitude to that of its trading partner but of a positive sign. Thus, the US, as an exporter of software, would be reluctant to lower δ_U as much as in the world optimal outcome, because of a similar terms-of-trade disincentive.

The above results suggest that when $\delta_J > \delta_U$, the individual best response function of Japan $\delta_J^n(\delta_U)$ will lie to the left of the world optimal choice $\delta_J^w(\delta_U)$, as illustrated in Figure 2. The exact opposite is true for the US, with its best response function $\delta_U^n(\delta_J)$ located above the world optimal choice $\delta_U^w(\delta_J)$. Note that the more asymmetric the two countries are in their education systems (and as a result, in their production patterns), the stronger the terms-of-trade effect will be. This is reflected in Figure 2 by the larger distance between $\delta_J^n(\delta_U)$ and $\delta_J^w(\delta_U)$, and between $\delta_U^n(\delta_J)$ and $\delta_U^w(\delta_J)$, as one moves further away from the 45° line.

The Nash equilibrium (δ_J^n, δ_U^n) corresponds to the intersection (point N) of the two best response functions $\delta_J^n(\delta_U)$ and $\delta_U^n(\delta_J)$ in Figure 2. It necessarily lies to the southeast of the autarky choice (point A) but to the northwest of the world social planner's choice (point W). Thus, although the two countries will choose to differentiate more under trade at the Nash equilibrium than in autarky, the differentiation in their education systems is less than what is socially optimal.[8]

Proposition 3 *In a free trade Nash equilibrium where each country simultaneously chooses its education system taking as given the other country's choice, the education systems in Japan and the US diverge more than their autarky levels but still less than what is socially optimal:* $\delta_J^w > \delta_J^n > \delta_J^a > \delta_U^a > \delta_U^n > \delta_U^w$.

4. Discussions

Modeling Choices. The main results of the paper are robust to many alternative specifications of education and technology. Essentially, any specification that implies a single crossing of PPFs such that the education policy increases the output of one sector but decreases that of the other would deliver similar predictions.

In the current framework, we adopt the production technology of Grossman and Maggi (2000), and an education policy that affects not only the diversity but also the mean of talent distribution. In fact, the mean effect of education policy is not crucial under some alternative technology setups. For example, one can show that in a model with decreasing returns to overall talent in the auto

[8] In the Nash setting, we cannot exclude the possibility of multiple Nash equilibria. This can happen if the best response curves $\delta_J^n(\delta_U)$ and $\delta_U^n(\delta_J)$ are very wiggly. Given that the best response functions are downward sloping, if there are multiple equilibria, $(\delta_J^{n1}, \delta_U^{n1}), (\delta_J^{n2}, \delta_U^{n2}), \ldots$, they can be ordered such that if $\delta_J^{n1} < \delta_J^{n2} < \ldots$ holds, then $\delta_U^{n1} > \delta_U^{n2} > \ldots$ would also hold. We can also rank the set of multiple equilibria by their Pareto superiority: a Nash equilibrium with a greater divergence between the education systems is Pareto superior. To see this, note that $U_J(\delta_J^{n1}; \delta_U^{n1}) < U_J(\delta_J^{n1}; \delta_U^{n2})$ is true because the more different country U is from country J, the larger the gains from trade for J. Furthermore, note that $U_J(\delta_J^{n1}; \delta_U^{n2}) < U_J(\delta_J^{n2}; \delta_U^{n2})$ holds since δ_J^{n2} is the best response to δ_U^{n2}. It follows that $U_J(\delta_J^{n1}; \delta_U^{n1}) < U_J(\delta_J^{n2}; \delta_U^{n2})$. Similarly, it can be shown that $U_U(\delta_U^{n1}; \delta_J^{n1}) < U_U(\delta_U^{n2}; \delta_J^{n2})$.

In the cooperative setting, similarly, there could be multiple local maxima. The socially optimal choice is the one with the largest divergence in education systems, i.e., the most extreme combination of education systems.

Thus, the welfare loss of the Nash equilibrium relative to the socially optimal outcome could arise from both the terms-of-trade effect identified above, and also from the possibility of being trapped in an inferior equilibrium of the policy game.

production but increasing returns in the software production, change in diversity alone is sufficient to generate the same desirable property of single-crossing PPFs.

The education policy is chosen in the paper by the government maximizing the aggregate welfare of all individuals in a country. This seems to be a reasonable approximation to reality, where education policies are often determined by government bureaucracies rather than by a referendum or other voting process. Nonetheless, it is possible that they are indirectly affected by voter preferences, for example, through the choice of the ruling party or the education minister. In the online appendix, we develop a full-blown political economy model based on both the majority voting rule and the probabilistic voting approach, and show that our main result (i.e., educational choices become more divergent after trade) still holds under a variety of reasonable political scenarios.

In the model, the utility (psychic) cost of centralization is the only cost of centralization, which is borne by each individual pupil as they go through the education system. In other words, there are no pecuniary costs involved, and this is meant to simplify the exposition (otherwise such pecuniary costs would have to be deducted from the income, which complicates the indirect utility function). In the main text, we have used a linear function to capture such utility cost of curriculum centralization. A more general functional form can be adopted without changing our substantial results. For example, an alternative disutility functional form could be $K(\delta, k)$, where the utility cost increases with the level of curriculum centralization δ and a parameter k, with $\frac{\partial K(\delta,k)}{\partial\delta\partial k} > 0$ such that the marginal utility cost with respect to δ is higher when k is larger.

In the model, we assume heterogenous utility costs of centralization across countries in order to generate different "initial" education systems across countries. This is, however, not a necessary assumption for the paper's main results. We can arrive at similar results by assuming away differences in utility costs across countries or the existence of such costs altogether. However, in these alternative setups, increased divergence in education systems is either a 'possibility' (rather than an inevitable outcome as in the paper), or extra conditions on the production technologies have to be imposed. Full analysis is provided in the online appendix.

In the main text, we have focused on the scenario without perfect specialization. There is nevertheless a possibility of perfect specialization under trade in the current setup, because the slope of the PPF is bounded away from zero and infinity. The scenario with perfect specialization is analyzed in the online appendix. In essence, our main propositions are not affected by the possibility of perfect specialization, except that we need to modify the free-trade equilibrium condition and the FOCs of policy choice for the fact that a country may produce only one good, and that we need to allow weak inequalities in the main propositions in some cases.

Finally, we have taken the country size to be the same across countries (with a unit measure of population). As the production technologies in both sectors exhibit constant returns to *scale* (with respect to the measure of population, say, L) and the preferences are identical and homothetic in the two countries, the country sizes can differ without affecting the comparative advantage and trade pattern. Different country sizes will not affect the qualitative results regarding the endogenous choice of education systems, either.

Interpretations and Extensions. The education system in the model is characterized by a single parameter δ that measures the level of curriculum centralization imposed on students, where a more centralized education system decreases the skill gaps among students but by its very nature also reduces the diversity of talent. While we acknowledge that it is possible to increase the basic skill level in some dimensions (e.g., reading, math and science) for all students without reducing the desirable diversity in other dimensions, our key insight on the fundamental tradeoff among different sets of skills remains valid, since resources that can be devoted to all these dimensions of skills cannot be unlimited and thus some tough allocation decisions have to be made. Hence, it is almost inevitable that when countries vary in their emphasis on different combinations of skills, this will have a bearing on their comparative advantages in trade.

The education parameter δ in our model can be interpreted in alternative ways to represent different features of an education system. For example, it is possible to interpret δ as the degree of ability pooling in school, the prevalence of public schools, or the degree of income equality. The links between these three factors and skill diversity in a closed economy have been studied extensively in the literature of economics of education.[9] Given our main focus on the dynamics between two sectors and two trading countries, we have chosen not to explicitly model these factors, except for the part captured by δ.[10] By focusing on the degree of curriculum centralization, our paper also highlights a distinct aspect of the education system that has not received much attention in the literature but that could exert fundamental effects on a country's talent composition, sectoral specialization, and comparative advantage in trade. That said, it would be interesting in future research to consider a richer model that takes into account multiple features of an education system and to work out their interactions in shaping the talent distribution.

Finally, one may also interpret our framework in a more general perspective than education policy per se. In our model, δ could be affected by many elements that are related to culture; for example, the utility cost of centralization is arguably lower in a culture that stresses conformity. Thus, education policy choice is potentially one important manifestation of cultural differences across countries, and such differences may indeed be reinforced by increased trade.

Empirical Strategy. Our theory predicts a wider divergence in education policies in a world with two countries if they open to trade with each other. In the online appendix, we argue that it is possible to extend this general prediction to an empirical setting with many countries. Second, in reality, we do not observe the transition from autarky state to completely free trade as envisioned in the model. The mechanism at work is nonetheless similar in a setting with trade cost: when trade cost decreases and two countries trade more with each other, the relative price of exported goods increases, providing incentives for countries to further reinforce their existing education systems.

Given the above empirical considerations, one could potentially test the paper's theoretical

[9]For example, see Bénabou (1996), Epple and Romano (1998), Fernández and Rogerson (1998), and Takii and Tanaka (2009) among others.

[10]Allowing unequal educational resources to exert extra influence on skill diversity is not likely to change our main results. For example, starting from the same inequality of educational resources in both Japan and the US, a higher δ in Japan leads to a more homogenous skill pool, which will then lead to a lower income inequality that in turn contributes to even more homogenous skills in future generations in comparison to the US. Thus, the effect of δ on skill diversity is enlarged by allowing unequal resources.

prediction by looking at the degree of trade intensity over time between pairs of countries and its effect on their differences in education systems. For example, the education system distance between a pair of countries (i, j) at time t can be measured by $C_{ij,t} \equiv |central_{i,t} - central_{j,t}|$, with $central_{c,t}$ being the curriculum centralization measure for country c at time t. Define $\Delta C_{ij} \equiv C_{ij,t_1} - C_{ij,t_0}$: thus, the larger the measure, the more divergent the two countries' education policies have become from t_0 to t_1.[11] On the other hand, the trade intensity between a pair of countries (i, j) at time t can be measured by the average of their bilateral export intensity $X_{ij,t} \equiv (\frac{export_{ij,t}}{GDP_{i,t}} + \frac{export_{ji,t}}{GDP_{j,t}})/2$, with $export_{ij,t}$ indicating the amount of exports from country i to country j at time t. The time difference $\Delta X_{ij} \equiv X_{ij,t_1} - X_{ij,t_0}$ reflects the change in trade intensity between the pair of countries. One could then regress ΔC_{ij} on ΔX_{ij} for a cross section of country pairs and take into account the potential endogeneity of trade. A positive estimated slope coefficient would be consistent with our theory.

Our theoretical prediction is best understood as a long-run effect, because education systems and policies take time to effect changes; the endogenous talent formation and structural adjustment in industry and trade are also slow-moving processes. As implied by the discussions in Appendix A.2, measures of curriculum centralization comparable across countries over a long period of time are not available. One of our measures of curriculum centralization based on school-level questionnaires is available across time, but only for 1995 and 1999. In this short time frame, it is likely to reflect more of short-run variations in the school-level practices than long-run shifts in the national-level policy. In addition, it is a relatively noisy measure as documented in Appendix A.2. Thus, we leave a rigorous test of our hypothesis to future work, when data become available on comparable measures of curriculum centralization across countries at an extended period of time.

[11] To accommodate a swap in the relative ranking of centralization between a pair of countries, a more general measure, $sign(central_{i,t_1} - central_{j,t_1}) * [(central_{i,t_1} - central_{j,t_1}) - (central_{i,t_0} - central_{j,t_0})]$, can be used.

5. Concluding Remarks

While the importance of the role of education in enhancing a country's international competitiveness has been recognized by many countries, the effects of international trade on education policies are not well understood. In an attempt to shed light on this issue, this paper provides a theory on the simultaneous determination of a country's education system and its comparative advantage in trade. Countries have to face tradeoffs when choosing a particular style of education system (a more centralized curriculum structure tends to promote talent homogeneity and the mean talent level, while a more flexible one encourages talent diversity). Since talent distribution affects industry productivities unevenly, the choice of education system will inevitably have a bearing on a nation's comparative advantage. An interesting implication of the paper's analysis is that trade plays an important role in shaping cross-country differences in education systems, because trade enhances countries' incentives to become more specialized in their sectors of comparative advantage and to further differentiate their education policies.

The empirical analysis based on the TIMSS dataset finds a significant and negative effect of curriculum centralization on talent diversity (but a positive effect on the mean talent level) for participating OECD countries in various years of study, which supports our theoretical characterization of the education system. We propose some empirical strategies to test the theory's prediction of a positive relationship between divergence in education systems and increased trade intensity, although the implementation is constrained by the lack of comparable time-series data on curriculum centralization. We leave a rigorous test of this long-run hypothesis to future work when data become available.

The current paper focuses on how the trade pattern and education systems between two advanced industrial economies interact, and how, as a result, persistent differences in education systems across countries arise in the equilibrium. A fruitful topic for future research may be to study how the education system of a country, competing and trading in the world economy, evolves dynamically over time as it advances across development stages. Another interesting extension could be to allow for multiple features of an education system (e.g., the degree of curriculum centralization, the ratio of public versus private school, or the degree of ability streaming), and to study how these factors interact to determine a country's talent distribution and how they are set endogenously in open economies.

Appendix

A.1. Proofs

Proof of Lemma 2.

Proof. Define $\phi = \mathrm{d}\Phi/\mathrm{d}t$. By Lemma 1 and the symmetry of the talent distribution, it follows that

$$\frac{\partial Y_A}{\partial \hat{t}} = -\lambda_A \, \bar{t} \, \phi(\hat{t}) < 0, \tag{13}$$

$$\frac{\partial Y_S}{\partial \hat{t}} = F^S(\hat{t}, 2\bar{t} - \hat{t})\phi(\hat{t}) > 0, \tag{14}$$

which implies that the MRT of the PPF is

$$MRT = -\frac{\partial Y_A/\partial \hat{t}}{\partial Y_S/\partial \hat{t}} = \frac{\lambda_A \bar{t}}{F^S(\hat{t}, 2\bar{t} - \hat{t})}. \tag{15}$$

Given (15), note that

$$\frac{\partial MRT}{\partial \hat{t}} = -\frac{\lambda_A \bar{t}}{F^S(\hat{t}, 2\bar{t} - \hat{t})^2}(F_1^S - F_2^S) > 0, \tag{16}$$

where the inequality follows by the fact that $F_1^S - F_2^S < 0$, since tasks are symmetric and there are increasing returns to individual talent in sector S. Use the competitive equilibrium condition $p = MRT$; the result $\frac{\partial \hat{t}}{\partial p} > 0$ therefore follows. The results in Lemma 2(ii) follow by (13), (14), and Lemma 2(i). ■

Proof of Lemma 3.

Proof. Define $g = \mathrm{d}G/\mathrm{d}t_0$. We have

$$\begin{aligned} Y_A^b(\delta,p) &= \frac{\lambda_A}{2}\int_{\hat{t}_b}^{2\bar{t}_b-\hat{t}_b} t_b\phi_b(t_b)\mathrm{d}t_b \\ &= \frac{\lambda_A}{2}\int_{\hat{t}_0}^{2\bar{t}_0-\hat{t}_0}[(1-\delta)t_0+\delta\bar{t}_0]g(t_0)\mathrm{d}t_0 \\ &= \frac{\lambda_A}{2}\,\bar{t}_0\,[G(2\bar{t}_0-\hat{t}_0)-G(\hat{t}_0)] \end{aligned} \tag{17}$$

where the second equality follows by making the change of variable $t_b = (1-\delta)t_0 + \delta\bar{t}_0$ and the fact that $\bar{t}_b = \bar{t}_0$. The variable $\hat{t}_0 \equiv \frac{\hat{t}_b - \delta\bar{t}_0}{1-\delta}$ identifies the initial ability of the marginal workers who after education have the cutoff talent level $\hat{t}_b$. The maximum possible auto output $Y_A^{b,max} \equiv \frac{\lambda_A}{2}\,\bar{t}_0$ is attained when all workers are allocated to the sector ($\hat{t}_0 = t_{l0}$). Note that for $Y_A^b < Y_A^{b,max}$,

$$\begin{aligned} \frac{\partial Y_A^b(\delta,p)}{\partial\delta} &= \frac{\lambda_A}{2}\bar{t}_0[-g(2\bar{t}_0-\hat{t}_0)-g(\hat{t}_0)]\frac{\partial\hat{t}_0}{\partial\delta} \\ &= \lambda_A\,\bar{t}_0\,g(\hat{t}_0)\left(\frac{\bar{t}_0-\hat{t}_0}{1-\delta}\right) \\ &= \lambda_A\,\bar{t}_b\,\phi_b(\hat{t}_b)\left(\frac{\bar{t}_b-\hat{t}_b}{1-\delta}\right) > 0, \end{aligned} \tag{18}$$

where in the second equality, to derive $\partial\hat{t}_0/\partial\delta = -(\bar{t}_0-\hat{t}_0)/(1-\delta)$, we have used the fact that $\hat{t}_b$ is a constant regardless of δ for a given relative price p: the condition $p = MRT_b$ as shown in (21) pins down a unique cutoff value $\hat{t}_b$, since $\bar{t}_b = \bar{t}_0$ is a constant regardless of δ. The last equality is derived by making a change of variable again, noting that $g(\hat{t}_0) = (1-\delta)\phi_b(\hat{t}_b)$ and that $\bar{t}_0 - \hat{t}_0 = (\bar{t}_b - \hat{t}_b)/(1-\delta)$. Similarly, the software output can be expressed as

$$\begin{aligned} Y_S^b(\delta,p) &= \int_{t_{lb}}^{\hat{t}_b} F^S(t_b, 2\bar{t}_b - t_b)\phi_b(t_b)\mathrm{d}t_b \\ &= \int_{t_{l0}}^{\hat{t}_0} F^S[(1-\delta)t_0+\delta\bar{t}_0, 2\bar{t}_0-(1-\delta)t_0-\delta\bar{t}_0]g(t_0)\mathrm{d}t_0. \end{aligned} \tag{19}$$

Thus we obtain

$$\begin{aligned} \frac{\partial Y_S^b(\delta,p)}{\partial\delta} &= F^S[(1-\delta)\hat{t}_0+\delta\bar{t}_0, 2\bar{t}_0-(1-\delta)\hat{t}_0-\delta\bar{t}_0]g(\hat{t}_0)\frac{\partial\hat{t}_0}{\partial\delta} + \int_{t_{l0}}^{\hat{t}_0}(F_1^S - F_2^S)(\bar{t}_0 - t_0)g(t_0)\mathrm{d}t_0 \\ &= F^S[\hat{t}_b, 2\bar{t}_b-\hat{t}_b]\phi_b(\hat{t}_b)\left(\frac{\hat{t}_b-\bar{t}_b}{1-\delta}\right) + \int_{t_{lb}}^{\hat{t}_b}(F_1^S-F_2^S)\left(\frac{\bar{t}_b-t_b}{1-\delta}\right)\phi_b(t_b)\mathrm{d}t_b < 0, \end{aligned} \tag{20}$$

where the inequality follows by the fact that $\partial\hat{t}_0/\partial\delta < 0$ and that $(F_1^S - F_2^S) < 0$. ■

Proof of Lemma 4.

Proof. Consider the talent distribution $\Phi_b(t_b)$ excluding the mean effect and its corresponding PPF. We have

$$MRT_b = -\frac{\partial Y_A^b/\partial \hat{t}_b}{\partial Y_S^b/\partial \hat{t}_b} = \frac{\lambda_A \bar{t}_b}{F^S(\hat{t}_b, 2\bar{t}_b - \hat{t}_b)}. \tag{21}$$

Given the fact $\bar{t} = \gamma^\delta \bar{t}_b$ and constant returns to overall talent in production, the condition $MRT = MRT_b$ holds if and only if $\hat{t} = \gamma^\delta \hat{t}_b$. Given this and the fact that $\Phi(\gamma^\delta t_b) = \Phi_b(t_b)$, it follows that $Y_A = \frac{\lambda_A}{2}\int_{\hat{t}}^{2\bar{t}-\hat{t}} t\phi(t)\mathrm{d}t = \frac{\lambda_A}{2}\int_{\hat{t}_b}^{2\bar{t}_b-\hat{t}_b} \gamma^\delta t_b \phi_b(t_b)\mathrm{d}t_b = \gamma^\delta Y_A^b$ by the change of variable; and similarly, $Y_S = \gamma^\delta Y_S^b$. Thus, $MRT = MRT_b$ implies that $Y_A/Y_S = Y_A^b/Y_S^b$. ■

Proof of Lemma 5.

Proof. Given Lemma 4, the total effects of the education system on the auto and software outputs are

$$\frac{\partial Y_A(\delta,p)}{\partial \delta} = (\gamma^\delta \ln\gamma) Y_A^b(\delta,p) + \gamma^\delta \frac{\partial Y_A^b(\delta,p)}{\partial \delta}; \tag{22}$$

$$\frac{\partial Y_S(\delta,p)}{\partial \delta} = (\gamma^\delta \ln\gamma) Y_S^b(\delta,p) + \gamma^\delta \frac{\partial Y_S^b(\delta,p)}{\partial \delta}, \tag{23}$$

which consist of the mean and diversity effects (reflected by the first and second terms, respectively).

Note that (22) is positive for $\gamma > 1$ given Lemma 3. Next, note that (23) increases monotonically with γ. The function $\gamma(\delta, p)$ introduced in Definition 1 defines the critical value such that (23) is equal to zero. It is clear that $\gamma(\delta, p)$ is necessarily greater than one, as the diversity effect is negative for software. The cap $\widetilde{\gamma}$ as the minimum of $\gamma(\delta, p)$ over all possible values of δ and p ensures that (23) is negative for $\gamma < \widetilde{\gamma}$. Thus, for $1 < \gamma < \widetilde{\gamma}$, the result in Lemma 5 follows.

As a side note, for parameters δ and p such that the software output is zero, (23) is automatically negative regardless of γ. As $\gamma(\delta, p)$ is not defined in this case, the value of $\widetilde{\gamma}$ is not affected by this scenario of complete specialization. ■

Proof of Lemma 6.

Proof. Define $V(\delta, p) \equiv -(1-\beta)p^{-1}Y_A(\delta,p) + \beta Y_S(\delta,p)$. Condition (4) can be rewritten as $V(\delta, p^a) = 0$, based on which we get

$$\frac{\mathrm{d}p^a}{\mathrm{d}\delta} = -\frac{\partial V(\delta,p^a)/\partial\delta}{\partial V(\delta,p^a)/\partial p^a} > 0, \tag{24}$$

which follows because $\partial V(\delta, p^a)/\partial\delta = -(1-\beta)(p^a)^{-1}\frac{\partial Y_A(\delta,p^a)}{\partial\delta} + \beta\frac{\partial Y_S(\delta,p^a)}{\partial\delta} < 0$ by Lemma 5 and $\partial V(\delta, p^a)/\partial p^a = -(1-\beta)(p^a)^{-1}\frac{\partial Y_A(\delta,p^a)}{\partial p^a} + \beta\frac{\partial Y_S(\delta,p^a)}{\partial p^a} + \beta(p^a)^{-1}Y_S(\delta, p^a) > 0$ by Lemma 2. ■

Proof of Lemma 7.

Proof. Define $V(\delta_J, \delta_U, p) \equiv -(1-\beta)p^{-1}\left[Y_A(\delta_J, p) + Y_A(\delta_U, p)\right] + \beta\left[Y_S(\delta_J, p) + Y_S(\delta_U, p)\right]$. Condition (5) implies that $V(\delta_J, \delta_U, p^f) = 0$, based on which we get

$$\frac{\partial p^f}{\partial\delta_J} = -\frac{\partial V(\delta_J, \delta_U, p^f)/\partial\delta_J}{\partial V(\delta_J, \delta_U, p^f)/\partial p^f} > 0, \tag{25}$$

where the sign follows because $\partial V(\delta_J, \delta_U, p^f)/\partial\delta_J = -(1-\beta)(p^f)^{-1}\frac{\partial Y_A(\delta_J,p^f)}{\partial\delta_J} + \beta\frac{\partial Y_S(\delta_J,p^f)}{\partial\delta_J} < 0$ by Lemma 5, and $\partial V(\delta_J, \delta_U, p^f)/\partial p^f = -(1-\beta)(p^f)^{-1}\frac{\partial\left[Y_A(\delta_J,p^f)+Y_A(\delta_U,p^f)\right]}{\partial p^f} + \beta\frac{\partial\left[Y_S(\delta_J,p^f)+Y_S(\delta_U,p^f)\right]}{\partial p^f} + \beta(p^f)^{-1}\left[Y_S(\delta_J, p^f) + Y_S(\delta_U, p^f)\right] > 0$ by Lemma 2. The result $\frac{\partial p^f}{\partial\delta_U} > 0$ can be shown similarly. ■

Proof of Proposition 1.

Proof. In the case of a small open economy, the objective function is the same as (6) except that the price is taken as given (a parameter) and need not clear the domestic market. The resulting FOC is

$$\frac{\partial U(\delta; p)}{\partial\delta} = \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}\left(\frac{\partial Y_A}{\partial\delta} + p\frac{\partial Y_S}{\partial\delta}\right) - k = 0. \tag{26}$$

Note that

$$\begin{aligned}\frac{\partial U(\delta; p)}{\partial p} &= \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}\left[\left(\frac{\partial Y_A}{\partial p} + p\frac{\partial Y_S}{\partial p}\right) - (1-\beta)p^{-1}Y_A + \beta Y_S\right] \\ &= \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}[-(1-\beta)p^{-1}Y_A + \beta Y_S],\end{aligned}$$

where the second equality follows by the competitive profit condition $p = MRT$. Based on the above condition, we get

$$\frac{\partial^2 U(\delta; p)}{\partial\delta\partial p} = \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}\left[-(1-\beta)p^{-1}\frac{\partial Y_A}{\partial\delta} + \beta\frac{\partial Y_S}{\partial\delta}\right] < 0,$$

by Lemma 5. This implies

$$\frac{\mathrm{d}\delta^o}{\mathrm{d}p} = \frac{\partial^2 U(\delta^o; p)}{\partial\delta\partial p}/\left(-\frac{\partial^2 U(\delta^o; p)}{\partial\delta^2}\right) < 0,$$

where $\frac{\partial^2 U(\delta^o;p)}{\partial\delta^2} < 0$ holds by the SOC for δ^o. ■

Proof of Proposition 2.

Proof. If Japan behaves as a price taker, its FOC to maximize its aggregate welfare is (26) with $k = k_J$ and $p = p^f$ (which needs to hold in any trade equilibrium with two countries). The resulting condition is identical to the world social planner's FOC (9). Thus, $\delta_J^w(\delta_U)$ can also be regarded as Japan's best response function, when Japan behaves as a price taker. Similarly, $\delta_U^w(\delta_J)$ can be regarded as the US's best response function, when the US behaves as a price taker. The noncooperative equilibrium outcome, when each of these two countries maximizes their individual welfare but behaves as price takers, occurs at the intersection W of the two schedules $\delta_J^w(\delta_U)$ and $\delta_U^w(\delta_J)$, which is identical to the world social planner's choice. This proves the first part of the proposition.

To show the second part of the proposition, note that if k_U were to decrease to the level of k_J, $\delta_U^w(\delta_J)$ would shift out (not shown) in Figure 2 and intersect $\delta_J^w(\delta_U)$ at point W_J on the 45° line. This is the hypothetical world social planner's choice if both countries had identical disutility factors equal to k_J. But if both countries were identical, the equilibrium trade price would be equal to the autarky price, and in this case, Japan's autarky price. Given that the FOC for Japan under autarky (7) and the FOC for the world social planner's choice of Japan's education system (9) are identical if $p_J^a = p^f$, it follows that δ_J at point W_J corresponds to Japan's optimal choice of education system under autarky.

Analogously, if k_J were to increase to the level of k_U, $\delta_J^w(\delta_U)$ would shift in (not shown) and intersect $\delta_U^w(\delta_J)$ at point W_U on the 45° line in Figure 2. This is the hypothetical world social planner's choice if both countries had identical disutility factors equal to k_U. But then the equilibrium trade price would be equal to the US's autarky price. Hence, δ_U at point W_U corresponds to the US's optimal choice of education system under autarky, given that the FOC for the US under autarky (7) is the same as the FOC for the world social planner's choice of the US's education system (10) if $p_U^a = p^f$.

Thus, the combination of the socially optimal education systems under autarky in the two countries corresponds to point A in Figure 2. It lies to the northwest of the world social planner's choice under trade W. The result $\delta_J^w > \delta_J^a > \delta_U^a > \delta_U^w$ therefore follows. ■

Proof of Proposition 3.

Proof. To show the FOC for Japan's best response function $\delta_J^n(\delta_U)$, note that:

$$\begin{aligned}\frac{\partial U_J(\delta_J;\delta_U)}{\partial \delta_J} &= \beta^\beta(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left[\frac{\partial Y_{AJ}}{\partial \delta_J} + p^f\frac{\partial Y_{SJ}}{\partial \delta_J}\right.\\ &\quad\left. + \left(\frac{\partial Y_{AJ}}{\partial p^f} + p^f\frac{\partial Y_{SJ}}{\partial p^f} + Y_{SJ} - (1-\beta)(p^f)^{-1}(Y_{AJ}+p^fY_{SJ})\right)\frac{\partial p^f}{\partial \delta_J}\right] - k_J\\ &= \beta^\beta(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left(\frac{\partial Y_{AJ}}{\partial \delta_J} + p^f\frac{\partial Y_{SJ}}{\partial \delta_J} + T\frac{\partial p^f}{\partial \delta_J}\right) - k_J = 0,\end{aligned} \tag{27}$$

where $T \equiv -(1-\beta)(p^f)^{-1}Y_{AJ} + \beta Y_{SJ} = \beta\frac{Y_{SJ}Y_{AU}-Y_{SU}Y_{AJ}}{Y_{AJ}+Y_{AU}}$. The above results follow by the competitive profit condition $p = MRT$ and the trade equilibrium condition (5).

Note that the FOC for $\delta_J^n(\delta_U)$ in (27) differs from the FOC for $\delta_J^w(\delta_U)$ in (9) only by $T\frac{\partial p^f}{\partial \delta_J}$, the TOT effect. Given that $T \lesseqgtr 0$ when $\delta_J \gtreqless \delta_U$ by Lemma 5, Japan's best response function $\delta_J^n(\delta_U)$ lies to the left of the world optimal choice $\delta_J^w(\delta_U)$ in the area below the 45° line in Figure 2. The opposite is true for the area above the 45° line, where $\delta_J^n(\delta_U)$ lies to the right of $\delta_J^w(\delta_U)$. The two schedules $\delta_J^n(\delta_U)$ and $\delta_J^w(\delta_U)$ cross each other on the 45° line when $\delta_J = \delta_U$, as in this case, the two countries have the same production structures and the terms-of-trade effect is zero.

An analogous analysis applies to the US. In its case, the objective function is $U_U(\delta_U;\delta_J) = \beta^\beta(1-\beta)^{1-\beta}p^{-(1-\beta)}(Y_{AU}+pY_{SU}) - k_U\delta_U$, which implies the following FOC for $\delta_U^n(\delta_J)$, equivalent to what is given in the main text:

$$\frac{\partial U_U(\delta_U;\delta_J)}{\partial \delta_U} = \beta^\beta(1-\beta)^{1-\beta}(p^f)^{-(1-\beta)}\left(\frac{\partial Y_{AU}}{\partial \delta_U} + p^f\frac{\partial Y_{SU}}{\partial \delta_U} - T\frac{\partial p^f}{\partial \delta_U}\right) - k_U = 0. \tag{28}$$

Given this, the exact opposite occurs to the ranking between the Nash and the world optimal choice in the US's case: the best response function $\delta_U^n(\delta_J)$ lies above $\delta_U^w(\delta_J)$ in the area below the 45° line where $\delta_J > \delta_U$, and vice versa if $\delta_J < \delta_U$. The rest of the analysis is as shown in the main text. ■

A.2. Empirical Relevance

In this appendix, we explain how we measure curriculum centralization in education systems and analyze its effects on the mean and diversity of a nation's talent distribution. The measures are constructed based on the data from the Trends in International Mathematics and Science Study (TIMSS). Every four years since 1995, TIMSS conducts surveys of the education policies and practices, and also assessments of math and science knowledge of grade 4 and grade 8 students around the world. We will focus on grade 8 students, who presumably receive a larger influence of the underlying education system than grade 4 students (Hanushek and Kimko, 2000). Participating countries as well as survey questions vary across years. In this paper, we use the data from 1995, 1999, and 2003, which contain survey questions relevant to our curriculum centralization measures. Since talent distribution in reality may be affected by many factors in addition to curriculum, we restrict our study to the OECD countries, which are more homogeneous in terms of the initial talent pool, access to public education, investment in education, and the development stage, among other relevant factors. Our final sample consists of existing OECD members by 2003 (the last year of our study) who also participated in TIMSS as a single entity.[12]

Talent Mean and Diversity Measure. The data show a high correlation between the math and science scores; thus, we will focus on their average to simplify presentation. We measure the mean talent level of a nation by $(\mu_m + \mu_s)/2$, where μ_m and μ_s are the respective means of the math and science scores for grade 8 students. Following the literature, we measure a country's talent diversity by the standard deviations of the math/science scores divided by their means, $(\frac{\sigma_m}{\mu_m} + \frac{\sigma_s}{\mu_s})/2$, where σ_m and σ_s are the respective standard deviations of the math and science scores for the same batch of students.

Curriculum Centralization Measure. Given the survey questions in TIMSS, there are two plausible ways to measure the degree of curriculum centralization for an education system. The first measure is based on the response of a national representative to the *curriculum questionnaire*, where one set of questions ask whether any of the seven listed methods are "used to help implement the national mathematics (science) curriculum at grade 8". The measure is the sum of the response (yes=1; no=0) to each of the seven options if there exists such a national curriculum, and it is zero

[12]Table 1 in Appendix A.3 lists the participating OECD countries.

if no national curriculum exists.[13] This set of questions are, however, only available in the 2003 survey. Since the responses are often the same for both math and science curricula in our sample, we once again focus on their average. The scale is normalized to the unit interval. A larger measure indicates a higher degree of curriculum centralization at the national level.

Countries with similar stated centralization policies at the national level could potentially differ from each other in their actual implementation practices at the school level. The differences may also evolve over time even when the written policy does not change (Astiz et al., 2002). In view of this, our second measure of curriculum centralization incorporates the information of school-level practices, where we take the average of the national-level measure constructed above and the school-level implementation index discussed below.

In TIMSS 1995 and 1999, the *school questionnaire* addressed to school principals includes 15 questions regarding the importance of various factors in determining the curriculum. These 15 factors can be regarded to represent either centralizing or decentralizing forces in terms of our model. For instance, the first question asks "How much influence does the National Curriculum Council have in determining curriculum?" with responses 1=none, 2=a little, 3=some, and 4=a lot. In light of our theory, this factor represents a centralizing force as it tends to impose homogeneity on the curriculum structures across schools. Likewise, we consider National Subject Association and external examinations/standardized tests, asked in two other questions, to represent constituencies that are likely to exert centralizing effects. In contrast, the remaining 12 factors in the list refer to local or school forces which tend to introduce heterogeneity in curricula across schools.[14] The importance of each force (ranging from 1 to 4) is measured by the average response of all valid samples in a country. The school-level implementation index is then constructed as the ratio of the average importance of the three centralizing forces and that of the 12 decentralizing forces. It is again normalized to the unit interval to be consistent with the national-level measure. The school-level index is available for both math and science curricula (they are incidentally identical

[13] See Appendix A.3 for the seven listed methods asked in this set of questions.

[14] See Appendix A.3 for the list of 15 factors asked in these questions.

in the 1999 survey); we take the average of the two as in the case of the national-level measure.[15]

The national-level measure has good international comparability, as the curriculum questionnaire is answered by a national representative, and the relevant questions we use in constructing the measure are the same across countries. In contrast, some questions used in the school-level index could be deleted or modified by countries as deemed suitable for their national education systems. The adaptations could also differ across waves of surveys in the same country. As a result, the cross-country comparability of the school-level index by itself is noisy.[16] Thus, the average of the national-level measure and the school-level implementation index aims to capture the international comparability of the national-level measure and at the same time to incorporate the time-series variation in the degree of curriculum centralization observed at the school level. Since the national-level measure is only observed in 2003 while the school-level index is only observed in 1995 and 1999, the centralization measure for 1995 is constructed as the average of the 1995 school-level index and the 2003 national-level measure. The centralization measure for 1999 is constructed similarly.[17]

Curriculum Centralization and Talent Distribution. The effects of the education system on a nation's talent distribution are shown in Figure 3, where we regress the mean and diversity measures, respectively, on the curriculum centralization measure. The estimates and significance levels are indicated in the diagrams.[18]

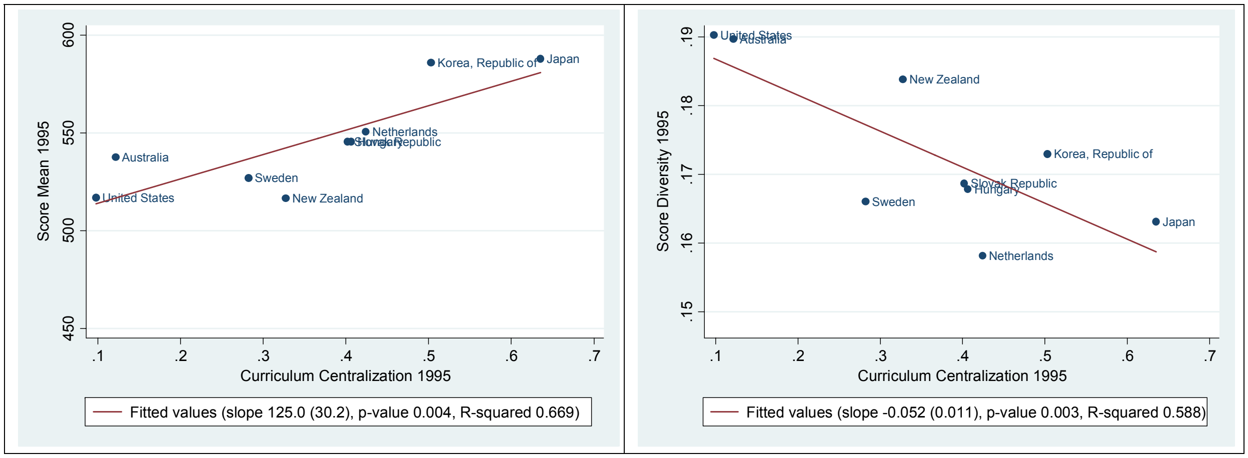
[15]This school-level index is highly correlated with the 1995 operational centralization measure of Astiz et al. (2002).

[16]For example, the US did not administer the question of National Curriculum Council in 1995, but replaced the option by Voluntary National Standards in 1999. In this case, we recode the 1999 score (of influence) for this factor to 1 (=none) as in 1995, judging that the substitute option does not constitute official curriculum standards with enforcement power. See TIMSS 1995 User Guide for the Primary and Middle School Years Supplement 3 (Section 7.8), and TIMSS 1999 User Guide Supplement 2 (Sections 4.6 to 4.8) for the complete list of national adaptations of the school questionnaire.

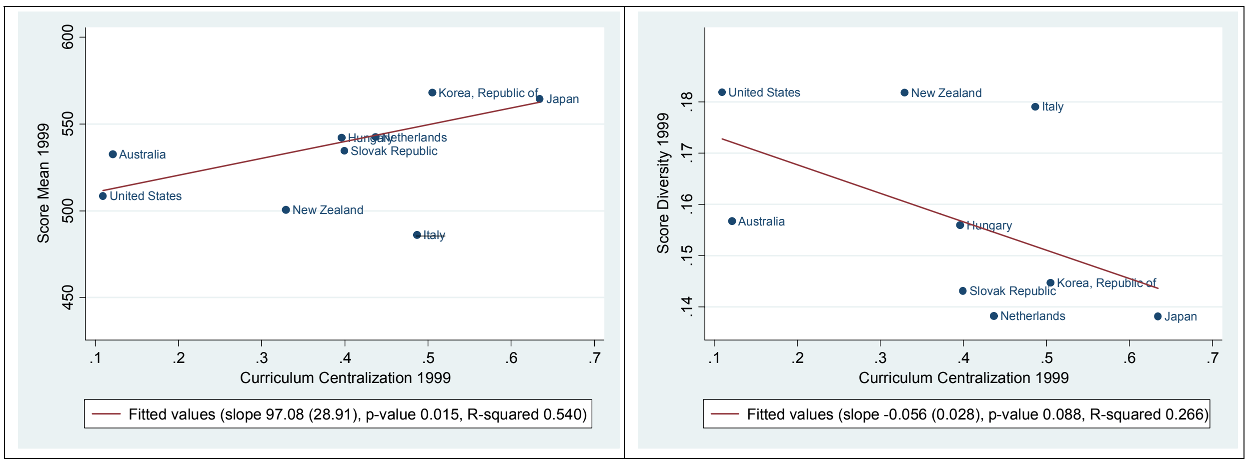
[17]Ideally, the national-level measure and the school-level index of the same year should be used in the average, but data for the 1995 and 1999 national-level measures are absent. It is reckoned that the national-level measure reflects the slow-moving component of education policies and is unlikely to change much in the short run.

[18]Robust standard errors are used in the estimation. The sample included in the regression is dictated by data availability.

(a) Year 1995



(b) Year 1999



(c) Year 2003

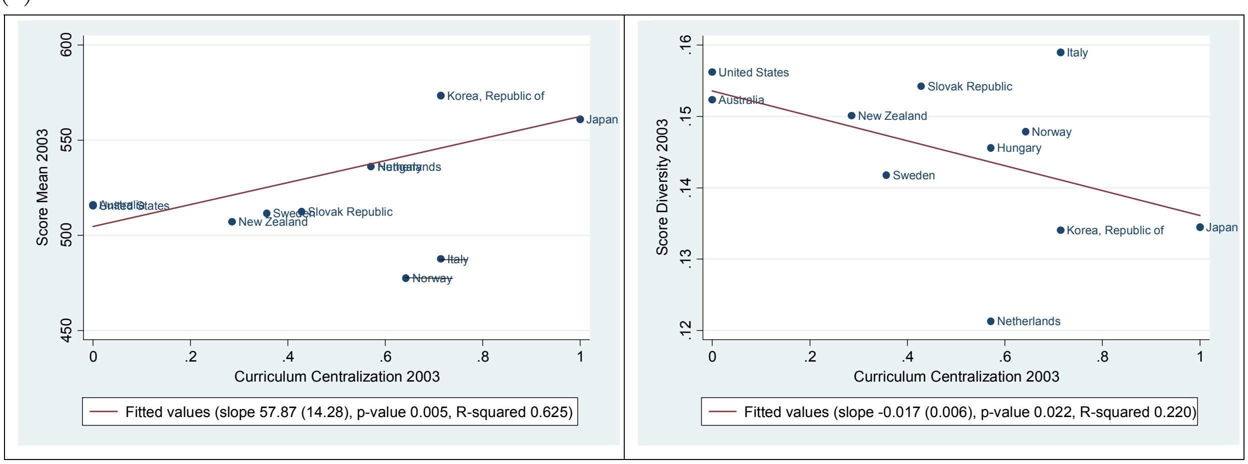


Figure 3: Score Mean and Diversity versus Curriculum Centralization

In all three years of surveys, there is strong evidence that the talent diversity is negatively correlated with the degree of curriculum centralization. The negative effect of curriculum centralization on talent diversity is highly significant and robust across years of study, even though the samples are slightly different across years and the measure of curriculum centralization in 2003 is based on the national-level measure alone. The findings are very similar if we regress the score diversity on curriculum centralization with a four-year lag (e.g., the score diversity of 1999 on the curriculum centralization of 1995) or with an eight-year lag.

There is also evidence that the average score tends to increase with the degree of curriculum centralization. The positive effect of curriculum centralization on the average score is highly significant among the sample of OECD countries that exclude Italy and Norway, two obvious outliers. We discuss in Appendix A.3 the potential idiosyncratic factors that could have a bearing on the average education outcome of these two countries but are not reflected by our centralization measure.[19]

The above results suggest that the more centralized a country's curriculum structure is, the more homogeneous its student performance tends to be and the higher the mean is. These are consistent with our theoretical characterization of the education system's effect on talent distribution. The data also show that the US and Japan indeed have the most extreme education systems in terms of curriculum centralization in the sample of OECD countries.

Stable Difference in the Education Systems of the US and Japan. Our theory suggests that the contrasting styles of education systems in the US and Japan could be a long-term equilibrium outcome that is compatible with and reinforced by their trade pattern. Their initial difference, due possibly to distinct cultural contexts, could be moderate but then gets reinforced over time and becomes difficult to reverse (short of dramatic shocks to the trade pattern).

A large degree of decentralization has long been a distinguishing feature of the US education system, and the evolution of this organization structure, dating back to the colonial era, has been "at least partially serendipitous" (Black and Sokoloff, 2006). The decentralized structure in financing

[19] It is also useful to note that our main propositions still hold in setups without the mean effect under alternative production technology specifications, as discussed in Section 4.

and administering schools by local or state authorities, through enhanced experimentation and flexibility and focused attention to local environments, has served the US quite well. Though in recent decades there are certain concerns about the relatively low performance of the US students in international tests compared with other developed countries, and some policies are adopted to address the "underperforming poor school districts" (Dillon, 2007, 2009), it does not appear that the distinctive features of the US education system, such as its decentralized manner, relatively low pressure on students, and emphasis on individual initiative and creativity, will change at all.

As a latecomer to the modern education, Japan in the Meiji era experimented with local funding and operation of education "only to discover that the people would not pay, so after only a few years of this experiment, the Meiji state took over the full burden of financing the public school system. ... The reliance on state support was associated with a uniform curriculum, a central system for exams and textbook production, and other centralizing tendencies" that deliver more homogeneous student outcomes than the decentralized US education system (Cummings, 1999). Similar experiences were repeated during the American occupation of Japan after the World War II; the decentralization effort initiated by the Americans was reversed years later to fit the traditional Japanese model more closely (Beauchamp, 1987). Most recently, in an effort to prevent "cram education" and boost individual potentials and the ability to think, Japan has experimented with various teaching methods under the so-called Yutori (Relaxed) Education. However, the new education guidelines have again met with strong resistance from education experts and parents (Takayama, 2007). These experiences illustrate the difficulty to reverse the initial pattern of a country's education style.

A.3. Data

The means and standard deviations of the test scores used in our study are retrieved from TIMSS 1995 Mathematics (Science) Achievement in the Middle School Years Table E.3, TIMSS 1999 International Mathematics (Science) Report Exhibit D.2, and TIMSS 2003 International Mathematics (Science) Report Exhibit D.2.

To construct the measure of curriculum centralization at the national level, we use Q.1A and

Q.3 in TIMSS 2003 Curriculum Questionnaire for Mathematics, and Q.1A and Q.4 in TIMSS 2003 Curriculum Questionnaire for Science. These questions ask whether there is a national curriculum that includes mathematics (science) at grade 8, and whether any of the following methods are used to help implement the national mathematics (science) curriculum at grade 8: (a) mandated or recommended textbook(s), (b) instructional or pedagogical guide, (c) ministry notes and directives, (d) curriculum evaluation during or after implementation, (e) specifically developed or recommended instructional activities, (f) national assessments based on student samples, and (g) a system of school inspection or audit. The survey data are retrieved from the files, BUGMATM3.xls and BUGSCIM3.xls, from TIMSS 2003.

In measuring the school-level implementation index, we use items SCQ2-13A to SCQ2-13O in TIMSS 1995 Population 2 School Questionnaire, and items SCQ2-9A to SCQ2-9O in TIMSS 1999 School Questionnaire. These questions ask how much influence each of the following 15 factors has in determining curriculum, with a rating from 1 to 4 (where 1=none, 2=a little, 3=some and 4=a lot): (A) National Curriculum Council, (B) National Subject Association, (C) Educational region or district, (D) School governing board, (E) Principal/head of school, (F) Teachers (collectively for the school), (G) Teachers (of same subject) as a group, (H) Each teacher individually, (I) Parents, (J) Students, (K) Church/religious groups, (L) Business community, (M) Textbook publishers, (N) External examinations/standardized tests, and (O) Teacher unions. We measure the importance of each factor (ranging from 1 to 4) by the average response of all valid samples in a country. If a factor was coded as N.A. for all schools in a country, we treat it as having no influence in the country and assign the factor a value of 1, since the most likely reason for this is that the question was not applicable to the country's context. For example, the US does not have a national curriculum council and Question (A) was not administered in 1995. See TIMSS 1995 User Guide for the Primary and Middle School Years (Chapter 7-30) for more discussions of N.A. entries. The above school questionnaire data are retrieved from BSALM92M1.TXT (for math) and BSALM42M1.TXT (for science) of TIMSS 1995, and bsalm3_m2.pdf (for math) and bsalm4_m2.pdf (for science) of TIMSS 1999.

Table 1: List of OECD countries participating in TIMSS

	TIMSS 1995			TIMSS 1999	TIMSS 2003		TIMSS 2007	
OECD countries	4th grade	8th grade	End of secondary school	8th grade	4th grade	8th grade	4th grade	8th grade
AUSTRALIA: 1971	o	o	o	o	o	o	o	o
AUSTRIA: 1961	o	o	o				o	
BELGIUM: 1961								
Belgium (Flemish)		o		o	o	o		
Belgium (French)		o						
CANADA: 1961	o	o	o	o				
DENMARK: 1961		o	o				o	
FINLAND: 1969				o				
FRANCE: 1961		o	o					
GERMANY: 1961		o	o				o	
GREECE: 1961	o	o	o					
ICELAND: 1961	o	o	o					
IRELAND: 1961	o	o						
ITALY: 1962	o	o	o	o	o	o	o	o
JAPAN: 1964	o	o		o	o	o	o	o
LUXEMBOURG: 1961								
NETHERLANDS: 1961	o	o	o	o	o	o	o	
NEW ZEALAND: 1973	o	o	o	o	o	o	o	
NORWAY: 1961	o	o	o		o	o	o	o
PORTUGAL: 1961	o	o						
SPAIN: 1961		o						
SWEDEN: 1961		o	o			o	o	o
SWITZERLAND: 1961		o	o					
TURKEY: 1961				o				o
UNITED KINGDOM: 1961								
England	o	o		o	o	o	o	o
Scotland	o	o			o	o	o	o
UNITED STATES: 1961	o	o	o	o	o	o	o	o
CZECH REPUBLIC: 1995	o	o	o	o			o	o
HUNGARY: 1996	o	o	o	o	o	o	o	
KOREA: 1996	o	o		o		o		o
MEXICO: 1994								
POLAND: 1996								
SLOVAK REPUBLIC: 2000		o		o		o	o	
CHILE: 2010				o		o		
SLOVENIA: 2010	o	o	o	o	o	o	o	o

Note: o = participating country. The first column lists OECD member countries and the year when they joined OECD. See http://nces.ed.gov/timss/countries.asp for a complete list of participating countries other than OECD members.

Norway. Norwegian pupils have performed relatively poorly on international comparative tests such as TIMSS. In fact, research has documented a downward trend in their achievements in mathematics and science over the 1990s and 2000s (Kjærnsli and Lie, 2002; Welle-Strand and Tjeldvoll, 2002; Tveit, 2009; Grønmo and Gustafsson, 2010).

Norway's education system is famous for its egalitarianism philosophy, where all pupils regardless of their endowments and socioeconomic backgrounds are entitled to 10 years of compulsory education. Equal right to education results in what is called the unified school, which has been at the heart of Norwegian education policies and a significant element of Norwegian welfare state policies since WWII (Braathe and Ongstad, 2001; Welle-Strand and Tjeldvoll, 2002). Characteristic of the system is a distaste for individual differentiation in academic achievements: there is no formal student assessment until lower secondary schools (grades 8–10), no permanent and structural streaming based on abilities, no repeating grades in the compulsory school years, and no choice for specialization in subjects before grade 12 (Braathe and Ongstad, 2001; Kjærnsli and Lie, 2002; Welle-Strand and Tjeldvoll, 2002; Tveit, 2009). In conjunction, there is a general tradition to let children be children, which is reflected in relatively short school hours in the primary years and few homework assignments after schools (Kjærnsli and Lie, 2002). This overall low academic pressure approach, with an aversion to competition and comparison in the primary school years, is likely one fundamental idiosyncratic factor for the outlying Norwegian performance observed at grade 8 in our study.

Some additional practices of the Norway education system may also have reinforced the above downside. Similarly reflecting the philosophy of the unified school, the Norwegian pedagogical rationale was that every pupil is entitled to a teaching method adapted to him or her. By some accounts (Braathe and Ongstad, 2001), the mathematics classroom is organized around students' activities: students spend large amounts of time working on worksheets or problem sets from the textbook, followed by rather brief periods of teacher-directed discussions and synthesis. This leads to a low content visibility in some lessons, and a reduced role for the teachers from a leader to a facilitator in the classroom. The teacher's role is also reflected in the low entry requirements for teacher colleges, as well as weak subject matter education in the colleges' own curriculum. For example, mathematics as a subject in teacher education for primary and secondary schools was not compulsory until 1992; more than 50 percent of teachers in primary schools who teach mathematics have no mathematics in their teacher education; a great portion of the teachers in primary schools who have no mathematics in their teacher education also have only the minimum of mathematics from upper secondary schools; nearly half of the teachers for grade 8 and below only have a short course in teaching methods in mathematics from the college (Braathe and Ongstad, 2001). In another source, it is documented that teacher college students are only able to solve 30 percent of tasks in mathematics curriculum for compulsory education (Welle-Strand and Tjeldvoll, 2002). Thus, there appears to be an internal conflict of goals in Norwegian basic school policies: by striving to offer adapted teaching to all children, it ends up with a depository school with a declining quality in terms of academic standards (Welle-Strand and Tjeldvoll, 2002).

Finally, some scholars have also identified the focus on practical knowledge instead of basic knowledge in the Norwegian primary education as one potential explanation for the less than spectacular performance of its pupils. In fact, the same unified school philosophy also has a distinct touch of anti-intellectualism: general knowledge is considered to be less important than local knowledge and practical skills (Welle-Strand and Tjeldvoll, 2002). More emphasis is placed on daily-life mathematics and science—such as measure, data and mountains—than on advanced and abstract concepts—such as algebra, number and atoms (Kjærnsli and Lie, 2002; Grønmo and Gustafsson, 2010). This implies that pupils may be less well trained in terms of systematic approach to knowledge synthesis and extension.

As an afterword, it is worthwhile to note that the Norwegian education system has in general shifted away from the tradition of the unified school toward a neo-liberalist approach after 2002, largely in response to the disappointing results of international assessment tests. In the new paradigm, quality is seen as a superior goal over equality, and differentiation and competition are accepted as the necessary means to achieve it. For example, a national test was introduced in 2004, and a new curricular framework, *The Knowledge Promotion*, was launched in 2006, with an emphasis on basic skills and subject competence. For example, national tests are now conducted in grade 5 and grade 8 to evaluate whether pupils have met the subject competence aims. It remains to be seen whether these educational reforms among others are effective in pushing up the Norwegian pupils' average performance in the future.

Italy. In Italy, primary schools last five years and lower secondary schools three years. Thus, grade 8 students (as surveyed by TIMSS in our study) are in their final year of lower secondary schools in Italy. Compulsory education was extended from five years to eight years in 1962, and to 10 years in 2007. The curriculum is the same for all schools until grade 8, after which students are streamed into three general courses of study by choice: general upper secondary schools, technical schools, and vocational schools (Checchi et al., 2013; De Simone, 2013; Ferrer-Esteban, 2011; Mocetti, 2012).

Italian students' performance in most international comparative tests is among the lowest in developed countries (Montanaro, 2008; Mocetti, 2012). Based on our reading, two common themes are frequently noted in studies of the Italian education system: first, students' achievement is strongly influenced by their family socioeconomic background, and second, there is a strong territorial divide between the North and the South in student achievement (with the South falling far behind the North). The learning divide among pupils, and across regions, respectively, appears to substantially widen during the lower secondary education (De Simone, 2013; Montanaro, 2008). Several studies highlight that the Italian education system has failed to reduce disparity in learning opportunities but allowed the social selection to develop implicitly at a stage as early as grade 6. This is manifested in within- and between-school social segregation (Ferrer-Esteban, 2011) along the gradient of socioeconomic backgrounds (or academic performance, as the two are highly correlated). In fact, the regional divide in academic performance in a way also reflects the influence of economic status, noting the disparity between the South and the North in their economic development levels.

This pattern of learning divide in Italy appears to have a significant bearing on both the level as well as the diversity of its student performance. Based on a study of the Italian education system by Ferrer-Esteban (2011), social dissimilarity between schools worsens pupils' academic results, all else being equal; furthermore, social segregation across classrooms (most likely derived from ability grouping practices) is also observed to adversely affect student performance. This is in line with the general findings of the education literature that class/school effect is larger for low achievers than for high achievers: grouping students by ability has a moderate positive effect on highly skilled students but has a strong negative impact on low-ability students. Thus, ability grouping enlarges the achievement diversity among children and also tends to reduce the average achievement level (Hanushek and Woessmann, 2010). This argument is also consistent with the observations on the differential academic performance along the territorial gradient of Italy, noting that the North is richer than the Center, and further richer than the South, of Italy. According to some studies (Siniscalco, 2005), the mean performance of the Italian Northern areas is slightly higher than the OECD average, that of the Centre of Italy about the OECD average, and that of the Southern areas being significantly much lower than the OECD average. Thus, the regional disparity in academic performance at the same time also lowers the national average.

Some other characteristics of the Italian schooling system may also have contributed to the lower academic performance of its grade-8 pupils. For example, the cycle of formal education is first split up at a relatively early stage (grade 5), coping with which could put the weaker students at a disadvantage; second, the full-time attendance is not practiced in all lower secondary schools, which again could put the students with less family resources at a disadvantage (Mocetti, 2012); third, it is documented by some studies that the Italian math lessons tend to focus more on rote memorization and citation of math definitions, than on the comprehension of the concepts (Santagata and Barbieri, 2005).

In spite of its seemingly highly-centralized education system (with homogeneous curriculum across students until grade 8 and with minimal school autonomy), the idiosyncratic factors discussed above may help account for Italian students' underperformance in international academic tests.

References

Astiz, M. F., A. W. Wiseman and D. P. Baker, "Slouching towards Decentralization: Consequences of Globalization for Curricular Control in National Education Systems," *Comparative Education Review* 46 (2002), 66–88.

Beauchamp, E. R., "The Development of Japanese Educational Policy, 1945–85," *History of Education Quarterly* 27 (1987), 299–324.

Bénabou, R., "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review* 86 (1996), 584–609.

Black, S. and K. Sokoloff, "Long Term Trends in Schooling: The Rise and Decline (?) of Public Education in the United States," in E. Hanushek and F. Welch, eds., *The Handbook of the Economics of Education* (Amsterdam: North Holland, 2006), 69–105.

Bombardini, M., G. Gallipoli and G. Pupato, "Skill Dispersion and Trade Flows," *American Economic Review* 102 (2012), 2327–48.

Bond, E. W., K. Trask and P. Wang, "Factor Accumulation and Trade: Dynamic Comparative Advantage with Endogenous Physical and Human Capital," *International Economic Review* 44 (2003), 1041–1060.

Bougheas, S., R. Kneller and R. Riezman, "Optimal Education Policies and Comparative Advantage," *Pacific Economic Review* 16 (2011), 538–552.

Bougheas, S. and R. Riezman, "Trade and the Distribution of Human Capital," *Journal of International Economics* 73 (2007), 421–433.

Braathe, H. J. and S. Ongstad, "Egalitarianism Meets Ideologies of Mathematical Education - Instances from Norwegian Curricula and Classrooms," *Zentralblatt für Didaktik der Mathematik* 33 (2001), 1–11.

Bracey, G. W., "International Comparisons and the Condition of American Education," *Educational Researcher* 25 (1996), 5–11.

Bray, M., *The Shadow Education System: Private Tutoring and its Implications for Planners*, 2nd edition (Paris: UNESCO, 2007).

Burbules, N. C. and C. A. Torres, eds., *Globalization and Education: Critical Perspectives* (London: Routledge, 2000).

Checchi, D., C. V. Fiorio and M. Leonardi, "Intergenerational Persistence of Educational Attainment in Italy," *Economics Letters* 118 (2013), 229–232.

Clune, W. H., "The Best Path to Systemic Educational Policy: Standard/ Centralized or Differentiated/Decentralized?," *Educational Evaluation and Policy Analysis* 15 (1993), 233–254.

Cummings, W. K., "The InstitutionS of Education: Compare, Compare, Compare!," *Comparative Education Review* 43 (1999), 413–437.

De Simone, G., "Render unto Primary the Things which are Primary's: Inherited and Fresh Learning Divides in Italian Lower Secondary Education," *Economics of Education Review* 35 (2013), 12–23.

Dillon, S., "Study Compares States' Math and Science Scores with Other Countries," New York Times (2007), November 14.

———, "Education Standards Likely to See Toughening," New York Times (2009), April 15.

Epple, D. and R. E. Romano, "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects," *American Economic Review* 88 (1998), 33–62.

Falkinger, J. and V. Grossmann, "Institutions and Development: The Interaction between Trade Regime and Political System," *Journal of Economic Growth* 10 (2005), 231–272.

Fernández, R. and R. Rogerson, "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review* 88 (1998), 813–833.

Ferrer-Esteban, G., "Beyond the Traditional Territorial Divide in the Italian Education System: Effects of System Management Factors on Performance in Lower Secondary School," FGA Working Paper No. 43, Giovanni Agnelli Foundation, 2011.

Findlay, R. and H. Kierzkowski, "International Trade and Human Capital: A Simple General Equilibrium Model," *Journal of Political Economy* 91 (1983), 957–978.

Green, A., *Education, Globalization and the Nation State* (London: Macmillan, 1997).

———, "Education and Globalization in Europe and East Asia: Convergent and Divergent Trends," *Journal of Education Policy* 14 (1999), 55–71.

Gregory, K. and M. Clarke, "High-Stakes Assessment in England and Singapore," *Theory Into Practice* 42 (2003), 66–74.

Grønmo, L. S. and J.-E. Gustafsson, "Student Achievement in Mathematics in Norway and Sweden as Evidenced by TIMSS," (The 4th IEA International Research Conference, 2010).

Grossman, G., "The Distribution of Talent and the Pattern and Consequences of International Trade," *Journal of Political Economy* 112 (2004), 209–39.

Grossman, G. M. and G. Maggi, "Diversity and Trade," *American Economic Review* 90 (2000), 1255–1275.

Hanushek, E. A., "The Long Run Importance of School Quality," NBER Working Paper Series 9071, National Bureau of Economic Research, 2002.

Hanushek, E. A. and D. D. Kimko, "Schooling, Labor-Force Quality, and the Growth of Nations," *American Economic Review* 90 (2000), 1184–1208.

Hanushek, E. A. and L. Woessmann, "The Economics of International Differences in Educational Achievement," NBER Working Paper Series 15949, National Bureau of Economic Research, 2010.

Kim, S.-J. and Y. J. Kim, "Growth Gains from Trade and Education," *Journal of International Economics* 50 (2000), 519–545.

Kjærnsli, M. and S. Lie, "TIMSS Science Results Seen from a Nordic Perspective," in D. F. Robitaille and A. E. Beaton, eds., *Secondary Analysis of the TIMSS Data* (Netherlands: Kluwer Academic Publisher, 2002), 193–208.

LeTendre, G. K., "The Problem of Japan: Qualitative Studies and International Educational Comparisons," *Educational Researcher* 28 (1999), 38–45.

Mayer, R. E., H. Tajika and C. Stanley, "Mathematical Problem Solving in Japan and the United States: A Controlled Comparison," *Journal of Educational Psychology* 83 (1991), 69–72.

Mocetti, S., "Educational Choices and the Selection Process: Before and After Compulsory Schooling," *Education Economics* 20 (2012), 189–209.

Mok, K. H., *Education Reform and Education Policy in East Asia* (New York: Routledge, 2005).

Montanaro, P., "Learning Divides across the Italian Regions: Some Evidence from National and International Surveys," Occasional Papers No. 14, Bank of Italy, 2008.

National Commission on Excellence in Education, "A Nation at Risk: The Imperative for Educational Reform," Washington, D.C.: U.S. Government Printing Office (1983).

Ohnsorge, F. and D. Trefler, "Sorting It Out: International Trade with Heterogeneous Workers," *Journal of Political Economy* 115 (2007), 868–892.

Santagata, R. and A. Barbieri, "Mathematics Teaching in Italy: A Cross-Cultural Video Analysis," *Mathematical Thinking and Learning* 7 (2005), 291–312.

Schaub, M. and D. P. Baker, "Solving the Math Problem: Exploring Mathematics Achievement in Japanese and American Middle Grades," *American Journal of Education* 99 (1991), 623–642.

Siniscalco, M. T., "Italy: Some Results from PISA 2003," Paper commissioned for the EFA Global Monitoring Report 2006: Literacy for Life, UNESCO, 2005.

Smith, M. L., C. Edelsky, K. Draper, C. Rottenberg and M. Cherland, "The Role of Testing in Elementary Schools," CSE Technical Report 321, UCLA Center for Research on Evaluation, Standards, and Student Testing, 1991.

Takayama, K., ""A Nation at Risk" Crosses the Pacific: Transnational Borrowing of the U.S. Crisis Discourse in the Debate on Education Reform in Japan," *Comparative Education Review* 51 (2007), 423–446.

Takii, K. and R. Tanaka, "Does the Diversity of Human Capital Increase GDP? A Comparison of Education Systems," *Journal of Public Economics* 93 (2009), 998–1007.

Tveit, S., "Educational Assessment in Norway—A Time of Change," in C. Wyatt-Smith and J. Cumming, eds., *Educational Assessment in the 21st Century* (Netherlands: Springer, 2009), 227–243.

Weiler, H. N., "Comparative Perspectives on Educational Decentralization: An Exercise in Contradiction?," *Educational Evaluation and Policy Analysis* 12 (1990), 433–448.

Welle-Strand, A. and A. Tjeldvoll, "The Norwegian Unified School – A Paradise Lost?," *Journal of Education Policy* 17 (2002), 673–686.

Westbury, I., "Comparing American and Japanese Achievement: Is the United States Really a Low Achiever?," *Educational Researcher* 21 (1992), 18–24.