ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Felipe, Jesus; Jin, Hongyuan; Mehta, Aashish

Working Paper

Education and the evolution of comparative advantage

ADB Economics Working Paper Series, No. 635

Provided in Cooperation with:
Asian Development Bank (ADB), Manila

Suggested Citation: Felipe, Jesus; Jin, Hongyuan; Mehta, Aashish (2021) : Education and the evolution of comparative advantage, ADB Economics Working Paper Series, No. 635, Asian Development Bank (ADB), Manila,
https://doi.org/10.22617/WPS210096-2

This Version is available at:
https://hdl.handle.net/10419/246712

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by/3.0/igo/

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

EDUCATION AND THE EVOLUTION OF COMPARATIVE ADVANTAGE

Jesus Felipe, Hongyuan Jin, and Aashish Mehta

NO. 635

March 2021

ADB ECONOMICS WORKING PAPER SERIES

ASIAN DEVELOPMENT BANK

ADB

ADB Economics Working Paper Series

Education and the Evolution of Comparative Advantage

Jesus Felipe, Hongyuan Jin, and Aashish Mehta

No. 635 | March 2021

Jesus Felipe (jfelipe@adb.org) is an advisor at the Asian Development Bank. Hongyuan Jin (hjin@umail.ucsb.edu) is a PhD student and Aashish Mehta (mehta@global.ucsb.edu) is an associate professor, at the University of California-Santa Barbara.

We thank for their comments and suggestions: Clement de Chaisemartin, Rana Hasan, John Morrow, Heather Royer, Yasuyuki Sawada, Asha Sundaram, Ashish Thapliyal, Yasuyuki Todo, and participants at the seminars at the Asian Development Bank and the National Graduate Institute for Policy Studies, Tokyo. Emmanuel Andal, Reynold Galope, and Caitlin Vejby provided excellent research assistance.

ASIAN DEVELOPMENT BANK

Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO)

© 2021 Asian Development Bank
6 ADB Avenue, Mandaluyong City, 1550 Metro Manila, Philippines
Tel +63 2 8632 4444; Fax +63 2 8636 2444
www.adb.org

Some rights reserved. Published in 2021.

ISSN 2313-5867 (print), 2313-5875 (electronic)
Publication Stock No. WPS210096-2
DOI: http://dx.doi.org/10.22617/WPS210096-2

The views expressed in this publication are those of the authors and do not necessarily reflect the views and policies of the Asian Development Bank (ADB) or its Board of Governors or the governments they represent.

ADB does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use. The mention of specific companies or products of manufacturers does not imply that they are endorsed or recommended by ADB in preference to others of a similar nature that are not mentioned.

By making any designation of or reference to a particular territory or geographic area, or by using the term "country" in this document, ADB does not intend to make any judgments as to the legal or other status of any territory or area.

This work is available under the Creative Commons Attribution 3.0 IGO license (CC BY 3.0 IGO) https://creativecommons.org/licenses/by/3.0/igo/. By using the content of this publication, you agree to be bound by the terms of this license. For attribution, translations, adaptations, and permissions, please read the provisions and terms of use at https://www.adb.org/terms-use#openaccess.

This CC license does not apply to non-ADB copyright materials in this publication. If the material is attributed to another source, please contact the copyright owner or publisher of that source for permission to reproduce it. ADB cannot be held liable for any claims that arise as a result of your use of the material.

Please contact pubsmarketing@adb.org if you have questions or comments with respect to content, or if you wish to obtain copyright permission for your intended use that does not fall within these terms, or for permission to use the ADB logo.

Corrigenda to ADB publications may be found at http://www.adb.org/publications/corrigenda.

Note:
In this publication, "$" refers to United States dollars.

The ADB Economics Working Paper Series presents data, information, and/or findings from ongoing research and studies to encourage exchange of ideas and to elicit comment and feedback about development issues in Asia and the Pacific. Since papers in this series are intended for quick and easy dissemination, the content may or may not be fully edited and may later be modified for final publication.

CONTENTS

TABLES AND FIGURES **iv**

ABSTRACT **v**

I. INTRODUCTION **1**

II. SPECIFICATION **3**

A. The Single-Stage Specification 3
B. A Two-Stage Approach 5

III. DATA **6**

IV. RESULTS **10**

A. Two-Stage Analysis 10
B. Single-Stage Analysis 16
C. Robustness Tests 20
D. Core and Peripheral Products 28

V. CONCLUSIONS **30**

APPENDIX **33**

REFERENCES **37**

TABLES AND FIGURES

TABLES

1	Summary Statistics	9
2	Explaining Shifts into Unfamiliar Products	12
3	Explaining Shifts into Complex Products	14
4	Explaining Shifts into Education-Intensive Products	15
5	Single-Stage Regression Results	17
6	Single-Stage Regression with Hanushek and Woessmann's (2009) Education Quality Measure	21
7	Alternative Estimates of Heckscher–Ohlin–Vanek Effects	22
8	Robustness to Alternative Measures of Product Sophistication	23
9	Robustness to Changes in Revealed Comparative Advantage Cutoffs	24
10	Robustness to Corrections for Institutions, Infrastructure, and Foreign Direct Investment	25
11	Correcting for Prior Industrial Dynamism	26
12	Correcting for Neighbors' Specialization Patterns	28
13	Regressions in Subsamples of Core and Peripheral Products	29
A1	List of Countries	34

FIGURES

1	Many Advanced Economies Have Few New Core Products to Develop	7
2	The Most Likely Effect of Education Is to Help Overcome Unfamiliarity	10
3	Countries with Better-Educated Workforces Experienced Less Path Dependence	11
4	Education and New Comparative Advantages in Complex Products	13
5	Education Expansions Were Not Associated with New Comparative Advantages in Education-Intensive Products	14
A1	Q-Q Plot for ln(RCA)	32
A2	Probability of Comparative Advantage in 2015 Conditional on Nonzero Revealed Comparative Advantages in 1995	33

ABSTRACT

We provide the first evidence that low- and middle-income countries with high education levels were more successful in developing comparative advantage in products unrelated to those they already export. In contrast, controlling for the relatedness of target products to these countries' exports, education appears unimportant for developing comparative advantage in products that are intrinsically complex or education intensive. These results are supported by analysis of the evolution of comparative advantage in 1,240 products from 49 low- and middle-income countries between 1995 and 2015. They are robust to corrections for measurement and specification errors, for institutional, infrastructure, and foreign direct investment-related factors, for regional specialization patterns, and for each economy's degree of industrial dynamism prior to 1995. These results suggest that the key role of education when seeking to shift from peripheral to core products is to help a country cope with unfamiliar challenges, and so overcome path dependence.

Keywords: comparative advantage, core, diversification, education, exports, path dependence, periphery, relatedness

JEL codes: I25, O11, O14

I. INTRODUCTION

Developing countries significantly improve their economic prospects by developing industries that produce complex, core products.[1] Unfortunately, these new industries require capabilities that must be acquired incrementally through practice, making industrial development difficult and path dependent.[2] Successful leaps into unrelated industries are rare, and the industrial policies responsible for many successes rely on institutional conditions that are difficult to replicate in other countries.[3] Seeking more neutral alternatives to industrial policy, this paper therefore asks whether education can facilitate the development of new export industries, and it characterizes the types of industrial transitions that education enables. While several papers suggest that education might help develop and sustain a salubrious export mix, these questions have not received focused attention in the literature.[4]

This is a noteworthy omission for at least three reasons. First, theory suggests that education could be important for developing new comparative advantages. The complexity literature and related work rooted in evolutionary theories (Nelson and Winter 1982, Stiglitz and Greenwald 2014, Hidalgo et al. 2018) regard the acquisition of tacit knowledge through learning by doing and the translation of knowledge across related domains to be key processes by which new capabilities are developed. Education has the potential to speed up learning by doing, and to permit the translation of knowledge across less related domains. In addition, education is thought to enhance actors' abilities to respond to emerging opportunities (Schultz 1975)—abilities that are crucial to the building of new industries. Second, while education, particularly high-quality education, has been shown to promote economic growth (Krueger and Lindahl 2001, Hanushek and Woessmann 2008), the mechanism underlying this relationship has not been clearly established. It is therefore useful to examine whether helping diversify toward core products might be such a mechanism. Third, some authors credit human capital with facilitating export-driven growth in East Asia, but do not provide comparative statistical evidence that it facilitates export transformation per se.[5]

---

1 Economies tend to grow faster when they export more complex products (Hidalgo et al. 2007, Hausmann et al. 2011) and a larger variety of products (Saviotti and Frenken 2008). A well-diversified and/or complex product mix has been linked to employment that grows faster and is more resilient to shocks (Frenken, Van Oort, and Verburg 2007), to shorter recessions (Hausmann, Rodríguez, and Wagner 2006), and to lower inequality (Felipe and Hidalgo 2014, Hartmann et al. 2017).

2 We apply standard definitions and concepts: "Capabilities" is shorthand for productive knowledge and practices embedded in individuals, firms, industries, supply chains, and institutions (Hausmann et al. 2011). Industries relying on overlapping capabilities (or inputs) are said to be related (Hidalgo et al. 2018). Path dependency means a tendency to develop new industries that are related to existing industries (Bahar et al. 2019). More complex products require more capabilities. Core products, like chemicals, sophisticated machinery, and advanced scientific instruments, are complex and are therefore related to many other products. Peripheral products, including many agricultural and mined commodities, are less complex and therefore less related to other products.

3 See, for example, Johnson (1982); Amsden (2001); Jomo (2003); Felker, Jomo, and Rasiah (2013); and Studwell (2013).

4 Previous studies show that: more rapid employment growth in skill-intensive industries in countries that had more highly educated workers and in those that expanded education faster (Ciccone and Papaioannou 2009); more educated countries are better able to maintain a diverse export mix in the face of terms-of-trade shocks (Agosin, Alvarez, and Bravo-Ortega 2012); and primary education attainment is a strong Bayesian predictor of national export diversification (Jetter and Ramírez Hassan 2015). None of this work examines the role of education in overcoming path dependence. Coniglio et al. (2018) come closest, showing that two crude proxies for education—scientific publication and educational expenditures—are associated with the development of unfamiliar industries in developing countries.

5 See, for example, World Bank (1993), Hobday (1995), and Stiglitz (1996). Others studying this history are more skeptical of education's role (Booth 1999, Asian Development Bank 2007, Chang 2012, Studwell 2013).

This paper examines the role of education in the evolution of comparative advantage using export data for 1,240 different goods for 49 low- and middle-income countries between 1995 and 2015.[6] In particular, we test three hypotheses, each about a role that education could play in altering a country's export mix.

Two of these three hypotheses are motivated by the theory of economic complexity. First, we ask whether countries whose workforces were more educated in 1995 were more likely to develop comparative advantage by 2015 in products that were unrelated to those they exported with comparative advantage in 1995 (i.e., whether they developed strengths in "unfamiliar" products). Second, motivated by the same theory, we ask whether countries with high education levels in 1995 were more likely to develop comparative advantage in products that are intrinsically more complex, controlling for those products' initial relatedness to their export mix. The third hypothesis derives from the Heckscher–Ohlin–Vanek intuition that education expansions should shift the export mix toward more education-intensive products.

Our key finding is that countries whose workforces were more educated in 1995 were indeed more likely to move into unfamiliar products. We also provide some evidence that, as might be expected, good quality basic education and high primary attainment facilitate movements into unfamiliar peripheral products, but not into unfamiliar core products. There is at best weak evidence to support the second hypothesized role of education, and none at all to support the third. We apply a two-step procedure which confirms that the lack of support for the latter two hypotheses is not driven by errors in the measurement of education or its change over time. We also demonstrate that our results are robust to errors in specification or operationalization; to biases owing to omitted variables related to institutional quality, infrastructure, foreign direct investment (FDI) receipt or regional specialization patterns; and to the fact that countries that underwent fast industrial development prior to 1995 tended to have both higher educational attainment in 1995, and more rapid industrial development between 1995 and 2015. While the instruments required to produce fully credible causal estimates of the effects of education on comparative advantage are not available, the robustness of our results to every alternative explanation suggests that they do provide a useful qualitative indication of the causal connections involved. The primary role of education in industrial development, at least among those we examine, is to help navigate the unfamiliar.

Although the most successful cases of industrialization in developing countries, with the exception of the People's Republic of China (PRC), predate 1995 (Felipe, Mehta, and Rhee 2019), we study the period 1995–2015 for three reasons. First, there is much more competition between nations for footholds in tradable industries now than in the past, so results from recent times are more relevant for policy. Second, the effectiveness of education should depend not only upon the quantity of schooling obtained, but also on its quality. We proxy for this using cognitive skills measures derived from international standardized tests that are only available beginning in the late 1990s. Third, trade policies vary less across countries after the structural adjustment era.

[6] Industrial development is a qualitatively different phenomenon in rich countries, many of which are deindustrializing. See Figure 1 in section III. The complete list of countries is found in Appendix Table A1.

The paper is structured as follows. We introduce our specifications and hypothesis tests in section II. We derive them in the appendix from a simple trade model that accounts for the connections between education, complexity, and path dependence.[7] We describe our data and variable definitions in section III and our results in section IV. Section V concludes.

II. SPECIFICATION

A. The Single-Stage Specification

We examine the relationship between education, familiarity, product characteristics, and the evolution of comparative advantage using a linear probability model, estimated on a pooled sample of countries and products, respectively indexed by c and p. Our generic, baseline specification, is:

$$CA_{c,p,t1} = \alpha_c + \alpha_p + f\left(RCA_{c,p,t0}\right) + \beta_F F_{c,p,t0} + \beta_{EF} E_{c,t0} F_{c,p,t0} + \gamma_{EC} E_{c,t0} C_{p,t0} + \delta_{EE} el_{p,t0} \Delta el_{c,t} + \delta_{KE} kl_{p,t0} \Delta kl_{c,t} + e_{c,p,t1} \quad (1)$$

Our initial time period, $t0$ is 1995, and the final time period, $t1$, is 2015. Here $RCA_{c,p,t} = \left(X_{c,p,t}/X_{c,t}\right)/\left(X_{p,t}/X_t\right)$ is Balassa's (1965) index of revealed comparative advantage (RCA) at time t, where X denotes exports. Let $CA_{c,p,t1} \equiv I\{RCA_{c,p,t1} \geq k\}$ indicate that country c has a comparative advantage in product p in 2015. Our baseline results use $k = 1$. Discretizing RCA sacrifices variation in the dependent variable, but is standard in the literature because it solves a range of econometric problems, and because varying the value of k allows us to check whether results are driven by information loss in particular RCA ranges (Bahar, Hausmann, and Hidalgo 2014; Bahar et al. 2019).[8]

We control for RCA in 1995 to capture long-run drivers of trade patterns, such as history and geography, as well as the availability of human and physical capital prior to $t0$. Controlling for lagged RCA also means that our coefficients capture the relationship between the independent variables and changes in comparative advantage. Country fixed effects allow for differences in the general level of development and diversification, while product fixed effects capture complexity, education intensity,

[7] This model combines a dynamic Ricardian trade model in the spirit of Redding (1999) with some simple assumptions about how the acquisition of tacit knowledge from production produces relatedness, and by extension, how education can facilitate this process. As discussed in section V, tacit knowledge is the most popular of several possible explanations of relatedness (Hausmann et al. 2011, Hidalgo et al. 2018).

[8] RCAs are nonnegative, often zero, and strongly right skewed, suggesting that a corner solution model is required if we are to treat them as continuous. Identification of these models relies on untestable distributional assumptions, and the product fixed effects required by theory also raise incidental parameter problems in a maximum likelihood context (Cameron and Trivedi 2005). Beyond these consistency problems, corner solution models yield nonlinear conditional expectations functions, which complicates hypothesis testing (Wooldridge 2002). Log-linearizing RCA, as required by the exponential Churdle model also results in findings being driven by differences close to RCA = 0, while using an inverse hyperbolic sin transformation would implicitly assume that starting to export a product poses similar challenges to increasing exports in an already exported product. In contrast, linear probability models are consistent and easy to interpret (Angrist and Pischke 2008). To ensure that our findings are not specific to the dynamics of comparative advantage around $RCA_{c,p,t1} = 1$, we follow Bahar, Hausmann, and Hidalgo (2014) in estimating the model after discretizing around $RCA_{c,p,t1} = 0.5$, 0.8, and 2 (see Table 9).

and other traits that make it more difficult to evolve comparative advantage in some products than in others.[9] Other than familiarity, $F_{c,p}$, the remaining variables are normalized to have a mean of 0 and a standard deviation of 1. The independent variables and interaction terms are motivated by our two theoretical approaches to the evolution of comparative advantage.

As noted, the economic complexity approach argues that it is difficult to develop comparative advantages in products with which a country is unfamiliar (Hidalgo et al. 2007, Hidalgo and Hausmann 2009, Hausmann et al. 2011). If this is true, then the evolution of comparative advantage will be path dependent, in the sense that whether a country's comparative advantage in a product grows over a time interval depends upon what else it exported at the start of that period. We will conclude that it is, on average, true for a country with initial education level $E_{c,t0}$ if $\beta_E + \beta_{EF}E_{c,t0} > 0$. A positive β_E indicates that a country endowed with average education experiences path dependence. The complexity literature also argues that it is more difficult to develop comparative advantages in intrinsically complex products, because doing so requires more know-how. We therefore utilize product fixed effects to permit the probability of comparative advantage to differ between products.

We turn next to our hypotheses regarding the role of education. First, high education levels may help countries to develop comparative advantage in unfamiliar products. This would be expected if education is useful for translating knowledge across domains, for identifying and acquiring required knowledge that was not already available from a country's initial product mix, or for responding to emerging business opportunities. This corresponds to a test of the alternative hypothesis that β_{EF} is negative. Second, holding familiarity constant, higher education levels may predispose countries to develop comparative advantage in intrinsically more complex products. This would be expected if larger amounts of knowledge are required to efficiently produce more complex products, and if education helps acquire this knowledge. If these two conditions hold, γ_{EC} should be positive. Appendix 1 derives the first row of specification (1) from an N-good Ricardian model, augmented to capture the role of education in promoting the acquisition of tacit and book knowledge over time.

The second row of specification (1) captures standard factor abundance considerations. We denote the initial education and capital intensity of products by $el_{p,t0}$ and $kl_{p,t0}$, and the growth in countries' per-worker endowments of education and capital by $\Delta el_{c,t}$ and $\Delta kl_{c,t}$. Our third hypothesis regarding education is that acquiring (rather than initially possessing) more education helps countries gain comparative advantage in education-intensive products. Producing education-intensive products involves more tasks in which educated workers have big productivity advantages over less educated workers. An increase in the number of educated workers is therefore predicted to increase production and export of education-intensive products to clear the market for educated workers (Leamer 1984). This corresponds to a test of the alternative hypothesis that $\delta_{EE} > 0$. One would also expect countries accumulating more physical capital per worker to shift more strongly toward capital-intensive products (i.e., $\delta_{KE} > 0$).

We measure education as a vector whose dimensions include quality and quantity, with quantity decomposable into contributions from primary, secondary, and college attainment. This permits us to test hypotheses regarding the roles of these dimensions of education in the evolution of

[9] More educated countries export a larger number of products with comparative advantage (Mehta and Felipe 2014).

comparative advantage. To examine whether and how the role of education varies with the type of product, we also reestimate this specification on subsamples of core and peripheral products.[10]

The estimated interaction coefficients reflect differences in the characteristics of the target products in which RCA is most often developed between more and less educated countries. They provide causal estimates of education's effect on the character of the export mix only if those differences are not explained by omitted variables that vary across country–product dyads.[11] While causal identification is challenged by the unavailability of suitable instruments for education and familiarity, we will show that our findings are robust to the inclusion of a wide variety of omitted country-product-level variables. To ensure they are robust to the omission of variables capturing institutional or infrastructure quality, openness to FDI, or industrial dynamism pre-1995, we estimate specifications that interact proxies for these national characteristics with familiarity and with product complexity. Finding, as we do, that our results are robust to this, and to several other potential sources of error, suggests that they do provide insight into the role that education has played.

B. A Two-Stage Approach

While the pooled, single-stage model produces efficient parameter estimates, the following two-stage approach is helpful for understanding why some interaction terms carry large, significant coefficients and others do not:

$$I\{RCA_{c,p,t1} \geq 1\} = \alpha_c + f(RCA_{c,p,t0}) + \beta_c F_{c,p,0} + \gamma_c C_{p,t0} + \delta_c\, el_{p,t0} + u_{c,p} \tag{2}$$

$$\hat{\beta}_c = b_0 + b_y y_{c,t0} + b_E E_{c,t0} + v_c \tag{3a}$$

$$\hat{\gamma}_c = c_0 + c_y y_{c,t0} + c_E E_{c,t0} + w_c \tag{3b}$$

$$\hat{\delta}_c = d_0 + d_y y_{c,t0} + d_E \Delta el_{c,t} + z_c \tag{3c}$$

The first-stage regression (2) is run separately for each country, with products as observations. A positive, significant coefficient on familiarity for a country, β_c, implies that its trade patterns evolved in a path-dependent fashion. The sign of γ_c summarizes whether country c's comparative advantage tended to shift toward more or less complex products—i.e., whether it climbed up or down the product ladder. The sign of δ_c indicates whether it shifted toward more or less education-intensive products. Our two-stage estimates scale $F_{c,p,0}$, $C_{p,t0}$, and $el_{p,t0}$, to have a mean of 0 and standard deviation of 1, so that the coefficients on each, are comparable to each other in magnitude.

The second-stage regressions (3a–3c) pool the estimated coefficients across countries and examine their relationship with initial per capita gross domestic product (GDP), $y_{c,t0}$, and education. As in specification (1), we hypothesize that high education quantity and quality in 1995 promote movements into unfamiliar and complex products between 1995 and 2015 (regressions 3a and 3b), and that increases in education levels between 1995 and 2015 promote movements into education-

[10] Our parameter estimates are summary statistics capturing the key historical differences between the export diversification experiences of better- and worse-educated countries. Given country and product fixed effects, as well as a rich array of corrections for theorized drivers of change in comparative advantage, they are useful for examining the plausibility of hypotheses about the role education has played in export diversification.

[11] Reverse causation is unlikely because the dependent variable is measured 15–20 years after the independent variables.

intensive products. Countries are weighted by the inverse of the standard errors of the relevant first-stage coefficients.

To see how this approach helps, suppose our estimates of (1) yield a negative result—that there is no tendency for more educated countries to develop comparative advantage in more complex products. In other words, suppose we fail to reject the null that $\gamma_{EC} = 0$. This could be because, controlling for the initial export mix, there is little variation between countries in their movements up or down the product ladder ($Var(\hat{\gamma}_c)$ is small), and therefore little in the historical record for differences in education across countries to explain. Or it could be because education does not correlate with these movements (i.e., $c_E = 0$ in regression 3b). Separating the stages will reveal which of these two explanations is relevant. This is important information if one is interested in economic history. Moreover, when the first explanation is relevant, the negative result cannot be attributed to measurement errors in our education variables.[12]

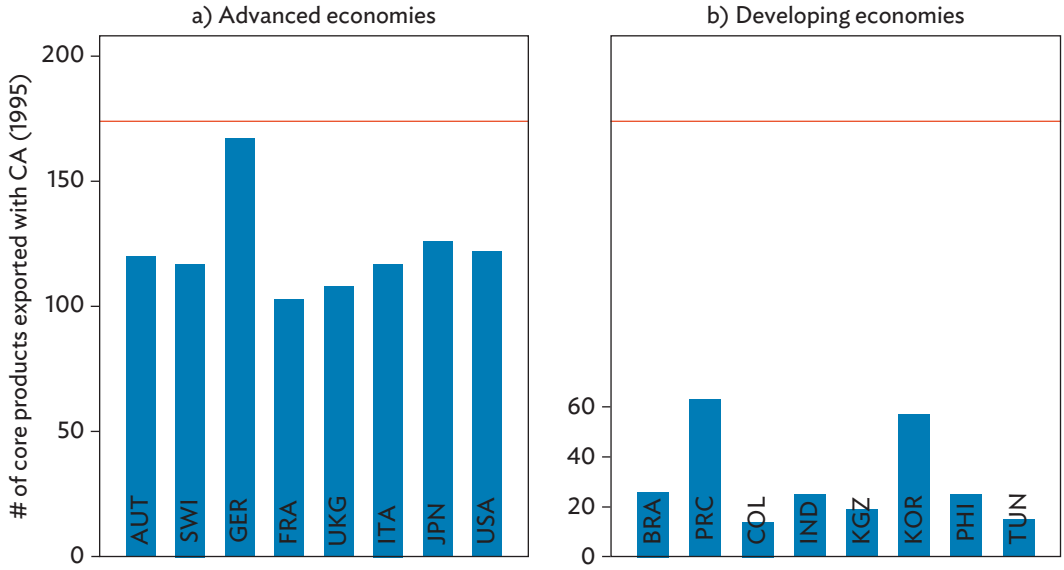
III. DATA

We use data from *The Atlas of Economic Complexity* on most countries' exports of 1,240 Harmonized System products in 1995 and 2015 (The Growth Lab at Harvard University 2019a). We exclude countries with per capita incomes in 1995 above $19,000. We do so because these countries had already developed comparative advantage in many core products by 1995 (Figure 1), making it difficult for them to establish a presence in many more core products during the period of our study. As advanced economies are among the most educated in the world, including them would lead us to underestimate the role of education among countries still attempting to move into the core of the global economy.

The main variable limiting our sample size is the quality of education. Hanushek and Woessmann (2009, henceforth, H&W) carefully calibrate and splice together the results of several international standardized mathematics and science tests administered to 15-year-olds to produce a cross-sectional dataset of countries' average student cognitive skills by the late 1990s. This calibration is performed relative to a group of Organisation for Economic Co-operation and Development countries that took multiple tests over time. Altinok, Diebolt, and Demeulemeester (2014) use slightly different criteria and procedures to assemble not only cross-sectional, but also time-series estimates of student cognitive skills for these and other countries between 1995 and 2012. In addition to a wider country coverage, these estimates are arguably more reliable than the H&W estimates for countries whose standardized test performances are most unlike those of Organisation for Economic Co-operation and Development countries that H&W use in their calibrations. Our sample includes 49 countries appearing in Altinok, Diebolt, and Demeulemeester (2014) cross-sectional dataset, 35 of which also appear in H&W, and our results are robust to switching to the H&W measures.

[12] We are motivated to deal with the possibility that measurement error drives negative results by previous work showing that inability to link growth to educational accumulation is at least partially explained by errors in measuring educational accumulation (Krueger and Lindahl 2001).

Figure 1: Many Advanced Economies Have Few New Core Products to Develop



AUT = Austria, BRA = Brazil, CA = comparative advantage, COL = Colombia, FRA = France, GER = Germany, IND = India, ITA = Italy, JPN = Japan, KGZ = Kyrgyz Republic, KOR = Republic of Korea, PHI = Philippines, PRC = People's Republic of China, SWI = Switzerland, TUN = Tunisia, UKG = United Kingdom, USA = United States.

Notes: Horizontal line is the maximum number of core products exported with comparative advantage by any country between 1995 and 2015 (a record held by Germany in 2010). Core products are those in the top one-third by connectedness and the top one-third by complexity. There are 238 core products.

Source: Authors' calculations.

We measure the quantity of education in 1995 and 2015 using data from Penn World Tables (PWT) (Feenstra, Inklaar, and Timmer 2015) on countries' average years of schooling. Data on primary, secondary, and college attainment rates in 1995 in the population aged 15 and above come from Barro and Lee (2010).

Define the proximity between products p and q by $\varphi_{p,q} \equiv \min\left[\Pr\left(CA_{c,p} = 1 \middle| CA_{c,q} = 1\right), \Pr\left(CA_{c,q} = 1 \middle| CA_{c,p} = 1\right)\right]$. Proximate products are presumed to rely on similar capabilities. We proxy for familiarity using "density," which measures how close a product is to the country's export basket, is calculated as $F_{c,p} \equiv \sum_{q \neq p}\left(CA_{c,q}\varphi_{p,q}\right) / \sum_{q \neq p}\left(\varphi_{p,q}\right)$, and must lie between 0 and 1.[13] Our measures of RCA and density are drawn from the *Atlas* data, and were calculated using all countries in the dataset.

Let M be a C x P matrix with each element equal to $CA_{c,p}$. We measure $C_{p,t0}$ by the product complexity index $\left(PCI_p\right)$, which utilizes information in M and a recursive method to rate as more complex those products that are exported with comparative advantage by fewer countries (uniqueness), especially when those countries' exports are themselves diverse (Hausmann et al. 2011, p. 24).

[13] This is in line with previous literature. Hidalgo et al. (2007) introduce density as a measure of the fraction of knowledge relevant to making product p that the country already possesses.

For any country-level factor endowment, Z_c, define $ProdZ_p \equiv \sum_c (RCA_{c,p} Z_c) / \sum_c (RCA_{c,p})$.[14] This infers, from the trade record, how intensive each product is in the use of this factor. We create three measures of the education content of products in this way. First, when Z_c is national average years of schooling, estimated from PWT, we obtain the measure $ProdYrs_p$. Second, when it is the share of a countries' population aged 15+ that completed college, taken from Barro and Lee's (2010) country data, we obtain $ProdColl_p$. Finally, when it is high schoolers' cognitive skill levels, using H&W's country data, we get $ProdCog_p$. We create each of these measures using data capturing conditions as close to 1995 as possible. We also use this procedure and PWT data to transform each country's ratio of capital to employment to produce our measure of products' relative capital intensities $kl_{p,t0}$. The same PWT data are used directly to measure log-changes in countries' endowments of education and capital per worker between 1995 and 2015 ($\Delta el_{c,t}$ and $\Delta kl_{c,t}$).

Next, we define the connectedness of each product as the sum of its proximities to all other products: $C_q \equiv \sum_{n \neq q} \varphi_{n,q}$. We classify products as "core" if they are in the top tercile of the distributions of both connectedness and PCI; and "peripheral" if they are in the bottom tercile of both distributions. To illustrate: most unprocessed agricultural and mined commodities, human hair, jute fibers, and electric power are revealed to be peripheral; jet engines, x-ray machines, watch movements, optical devices, and machine tools are core products; and paper, electric shavers, hats, copper wire, and wine are in-between. In our dataset of 1,240 products, 230 are core and 232 are peripheral. The remaining 778 are in-between.

The control variables used in this paper include multiple measures of the quality of countries' institutions and infrastructure, as well as the average FDI:Exports and FDI:GDP ratios between 1995 and 2015, drawn from the World Development Indicators. Finally, we use three proxy measures of countries' industrial dynamism prior to 1995: real per capita GDP growth from the World Development Indicators, labor productivity growth from PWT, and estimates of β_c and γ_c for the periods 1975–1995 and 1985–1995 from specification (2) using data on exports by 4-digit Standard International Trade Classification products (The Growth Lab at Harvard University 2019b).

Table 1 provides summary statistics. Usefully, the countries in our sample differ widely in educational attainment and quality, and RCAs in many industries (country–product dyads) demonstrate significant changes between 1995 and 2015.

[14] These measures are analogous to *PRODY*, introduced to measure products "income content" by Hausmann, Hwang, and Rodrik (2007).

Table 1: Summary Statistics

	Observations	Mean	SD	Minimum	Maximum
Country-level variables					
Quantity (Average years of schooling)	49	7.11	2.35	2.16	11.39
Edu Quality A (Altinok, Diebolt, and Demeulemeester 2014)	49	494.39	75.07	282.24	652.62
Edu Quality B (Hanushek and Woessmann 2009)	35	4.27	0.58	3.09	5.34
Primary (Primary attainment, aged 15+)	48	73.13	17.08	36.21	98.03
Secondary (Secondary attainment, aged 15+)	48	33.32	16.65	4.81	76.70
College (College attainment, aged 15+)	48	5.88	4.19	0.58	19.35
ΔYrsSch (Change in average years of schooling)	49	2.31	0.76	0.50	4.55
ΔQuality A (Change in Edu Quality A)	22	9.81	30.02	−36.60	63.47
ΔK/L (Change in capital intensity)	44	0.55	0.44	−0.51	2.02
Product-level variables					
PCI (Product Complexity Index)	1,242	0.00	1.00	−2.93	2.84
ProdYrs (Average years of schooling across product exporters)	1,242	9.02	1.18	5.04	12.59
ProdCog (Average education quality across product exporters)	1,242	4.75	0.22	3.89	5.20
ProdColl (Average college attainment across product exporters)	1,242	8.07	1.74	3.20	15.58
ProdKL (K/L averaged across product exporters)	1,242	−9.49	0.74	−13.11	−6.66
Country-product-level variables					
RCA (1995)	60,858	1.20	9.24	0.00	868.77
RCA (2015)	60,760	1.18	9.19	0.00	1,243.06
CA = I(RCA_2015>=1}	60,760	0.17	0.37	0.00	1.00
ΔRCA (Change in export RCA, 1995–2015)	60,760	−0.03	9.23	−866.65	476.23
Familiarity	60,858	0.16	0.09	0.00	0.70

CA = comparative advantage, RCA = revealed comparative advantage, SD = standard deviation.

Note: All country-level variables will be normalized to have a mean of 0 and standard deviation of 1 when used in regressions.

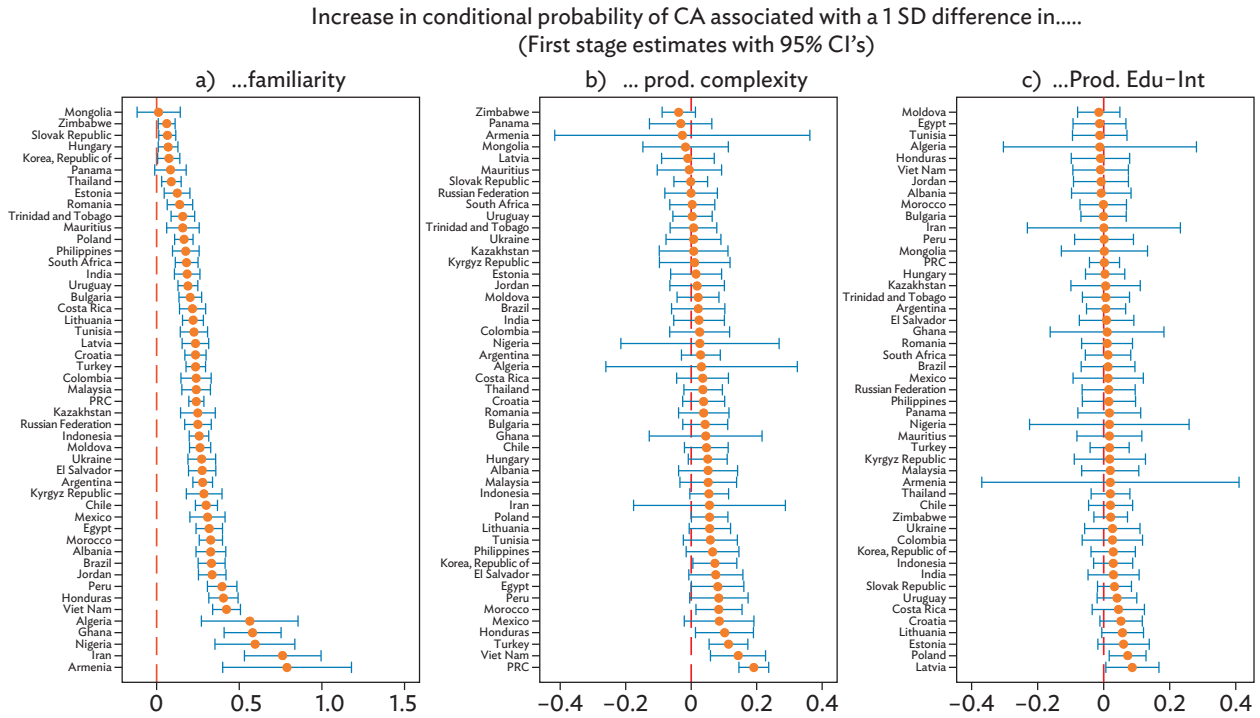
Source: Authors' calculations.

IV. RESULTS

A. Two-Stage Analysis

Figure 2 shows each country's first-stage coefficient estimates.[15] Panel (a) indicates that countries differed significantly in terms of their tendency to keep exporting familiar products. In Iran and Armenia, for example, a 1 standard deviation difference in familiarity between two products in 1995 is associated with a roughly 75 percentage point difference between them in the conditional probability of comparative advantage by 2015 (after conditioning on lagged RCA, product complexity, and education intensity). In Mongolia, the Republic of Korea, and Zimbabwe, on the other hand, familiarity played almost no role. It follows that there is significant variation in the degree of this type of path dependence across countries.

Figure 2: The Most Likely Effect of Education Is to Help Overcome Unfamiliarity



CA = comparative advantage, CI = confidence interval, PRC = People's Republic of China, SD = standard deviation.

Source: Authors' calculations.

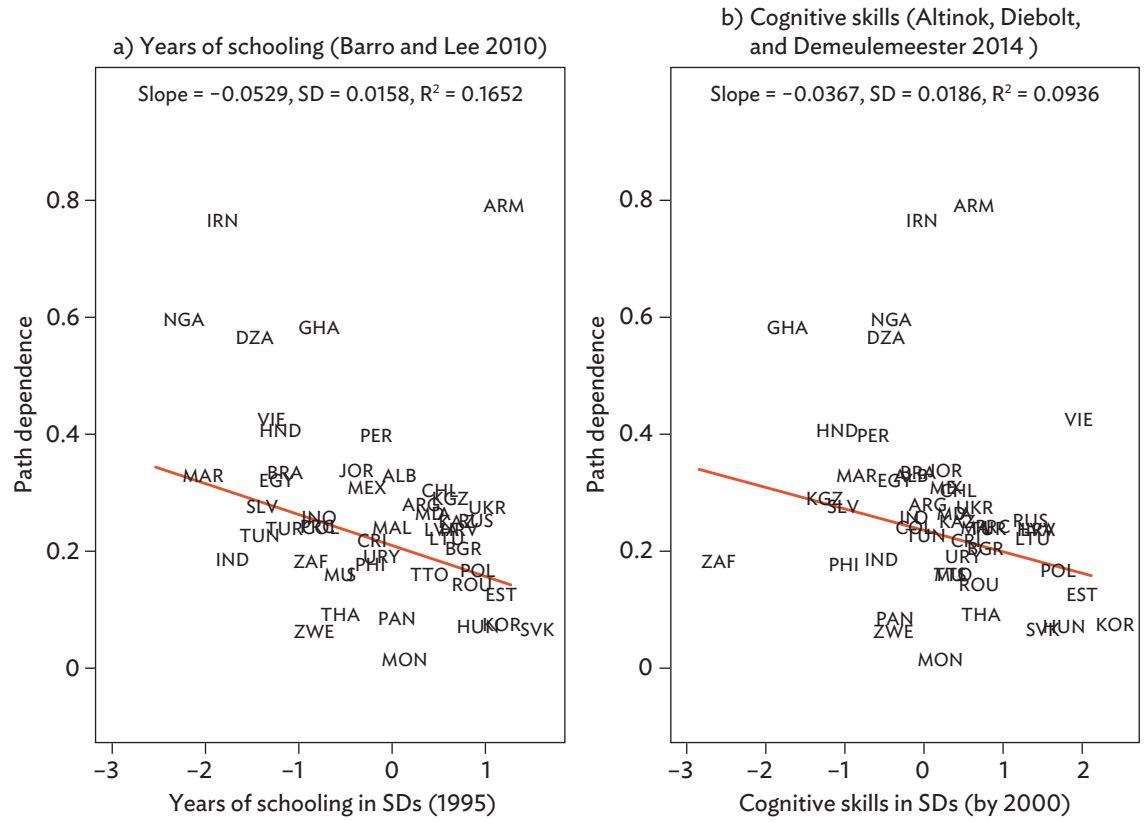
[15] The lagged RCA correction takes the form: $f(RCA) = f_0 * I\{RCA = 0\} + f_1 * [1 - I\{RCA = 0\}]g(RCA)$, where $g(RCA) = ln(RCA)$ when $RCA > 0$ and $g(RCA) = m$ when $RCA = 0$. This specification makes allowance for the possibility that exporting any of a product has different effects on the likelihood of future comparative advantage than does having a high RCA in it. Our coefficient estimates are invariant to the value chosen for the constant m by construction. Log-linearizing the nonzero values is recommended by the q-q plot of $ln(RCA)$ (Appendix Figure A1). We have also run our main regressions using a hyperbolic sine function in place of $g(\quad)$, but this is restrictive and does not alter our main results.

In contrast, panel (b) provides much less evidence, controlling for product familiarity, that product complexity influenced the evolution of countries' comparative advantages, or that this influence varied across countries. Product complexity was significantly and positively associated with gains in comparative advantage in six countries (the Republic of Korea, Morocco, Honduras, Turkey, Viet Nam, and the PRC). But even in the PRC, a 1 standard deviation difference in product complexity was associated with less than 20 percentage point difference in the probability of acquiring a comparative advantage by 2015. No tendency to move up or down the complexity ladder is statistically discernible in the remaining 43 countries.

Panel (c) shows even less variation in the relationship between products' education intensity and their conditional probability of acquiring a comparative advantage by 2015. Indeed, there is only one country (Poland), where this relationship is significantly different from zero.

These results show that countries differ much more with respect to their success overcoming unfamiliarity than with their success overcoming complexity or education intensity. Our second-stage regressions examine whether education can account for the cross-country differences just documented.

Figure 3: Countries with Better-Educated Workforces Experienced Less Path Dependence



SD = standard deviation.

Note: See the list of countries in Appendix Table A1.

Source: Authors' calculations.

Figure 3 explores specification (3a) graphically. It shows that countries with more average years of schooling in 1995 and higher levels of cognitive skill by around 2000 were significantly less tethered to familiar products. Columns 1–3 of Table 2 show that these two relationships survive correcting for per capita GDP and switching measures of education quality—although in the case of the Altinok, Diebolt, and Demeulemeester (2014) measure, the relationship is not statistically significant. Columns 4 and 5 correct for both quantity, quality, and an interaction between the two. Whether quality or quantity have explanatory power is sensitive to the quality measure used. The quality–quantity interactions are statistically insignificant and their inclusion does not change these findings qualitatively.

Table 2: Explaining Shifts into Unfamiliar Products

	First-Stage Coefficient on Familiarity/Familiarity				
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Per capita GDP in 1995 (Constant 2010 $)	−0.026	−0.037	−0.024	−0.017	−0.012
	(0.023)	(0.025)	(0.026)	(0.025)	(0.034)
Quantity (Average years of schooling)	−0.042**			−0.036*	−0.007
	(0.017)			(0.021)	(0.034)
Quality A (Altinok, Diebolt, and Demeulemeester 2014)		−0.024		−0.017	
		(0.018)		(0.017)	
Quality B (Hanushek and Woessmann 2009)			−0.048***		−0.053**
			(0.015)		(0.020)
Quantity x Quality A				−0.016	
				(0.014)	
Quantity x Quality B					−0.022
					(0.013)
Constant	0.423**	0.536**	0.433*	0.370*	0.345
	(0.196)	(0.207)	(0.218)	(0.205)	(0.295)
Observations	49	49	35	49	35
R-squared	0.186	0.142	0.223	0.202	0.248

GDP = gross domestic product.

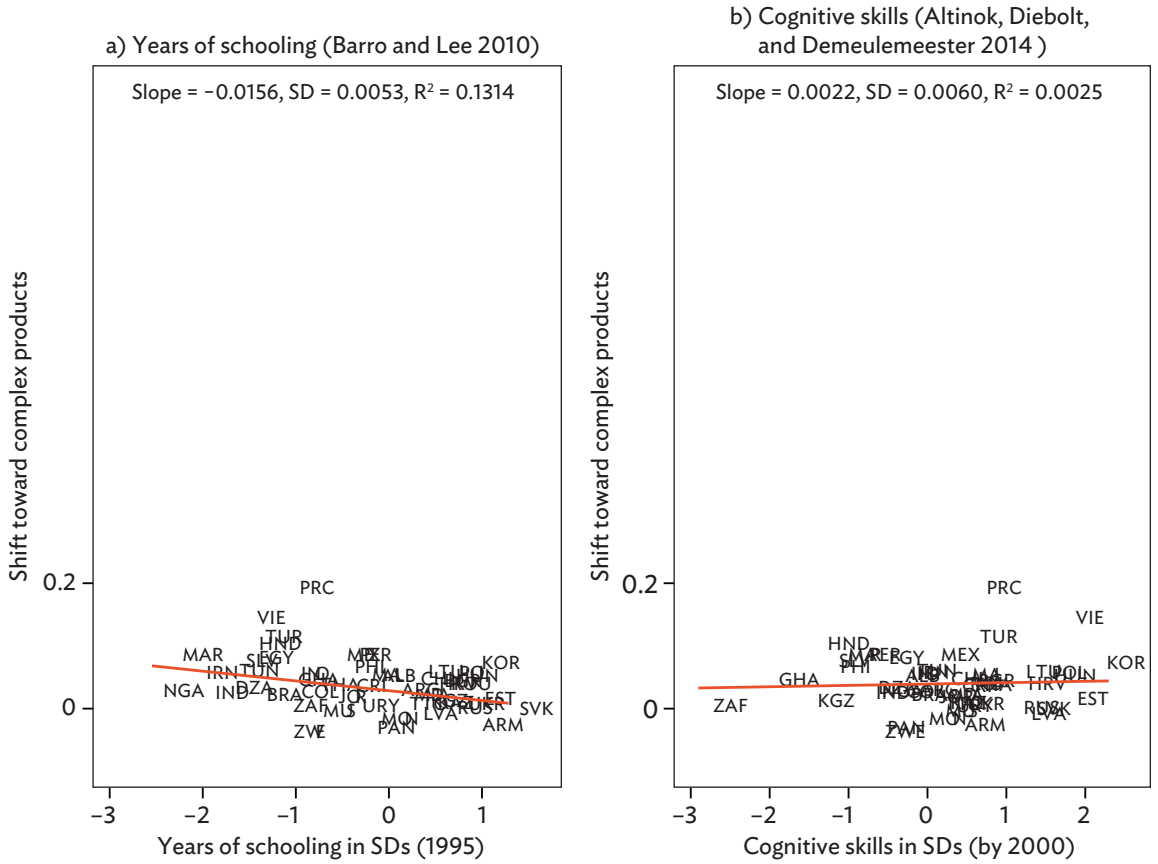
Notes: Weighted least squares estimates, per specification (3a). The dependent variable is each country's coefficient on familiarity from first-stage specification (2). Countries are weighted by the inverse of the standard error of that coefficient. Abbreviated variable names appear in Table 1. All standard errors are robust. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Figures 4 and 5 explore specifications (3b) and (3c), maintaining the same y-axis scale as Figure 3 for the sake of comparison. Neither provides strong support for the other two hypothesized roles of education. Figure 4a shows that higher initial years of schooling is associated with *less* movement toward complex products. Figure 4b shows a positive but extremely small relationship between the quality of education and movement into complex products. Figure 5 shows that

expansions in the quantity of education were associated with movements into *less* education-intensive products, and that there is no relationship between quality improvements and increased education intensity in the product mix. Tables 3 and 4, which again add per capita GDP corrections, confirm that most of these relationships have the wrong (negative) sign, or are small or statistically insignificant. The one partial exception (Table 3, column 4) is that when controlling for the quantity of education and the Altinok, Diebolt, and Demeulemeester (2014) quality measure simultaneously, the latter is associated with a slightly higher tendency to shift toward complex products. Table 4 shows that neither increases in education quantity (column 1) nor quality (in a much smaller sample of countries, column 2) are associated with movement into education-intensive products, and that the relationship with quantity expansions is not stronger if the education is of a higher quality (column 3).

Figure 4: Education and New Comparative Advantages in Complex Products

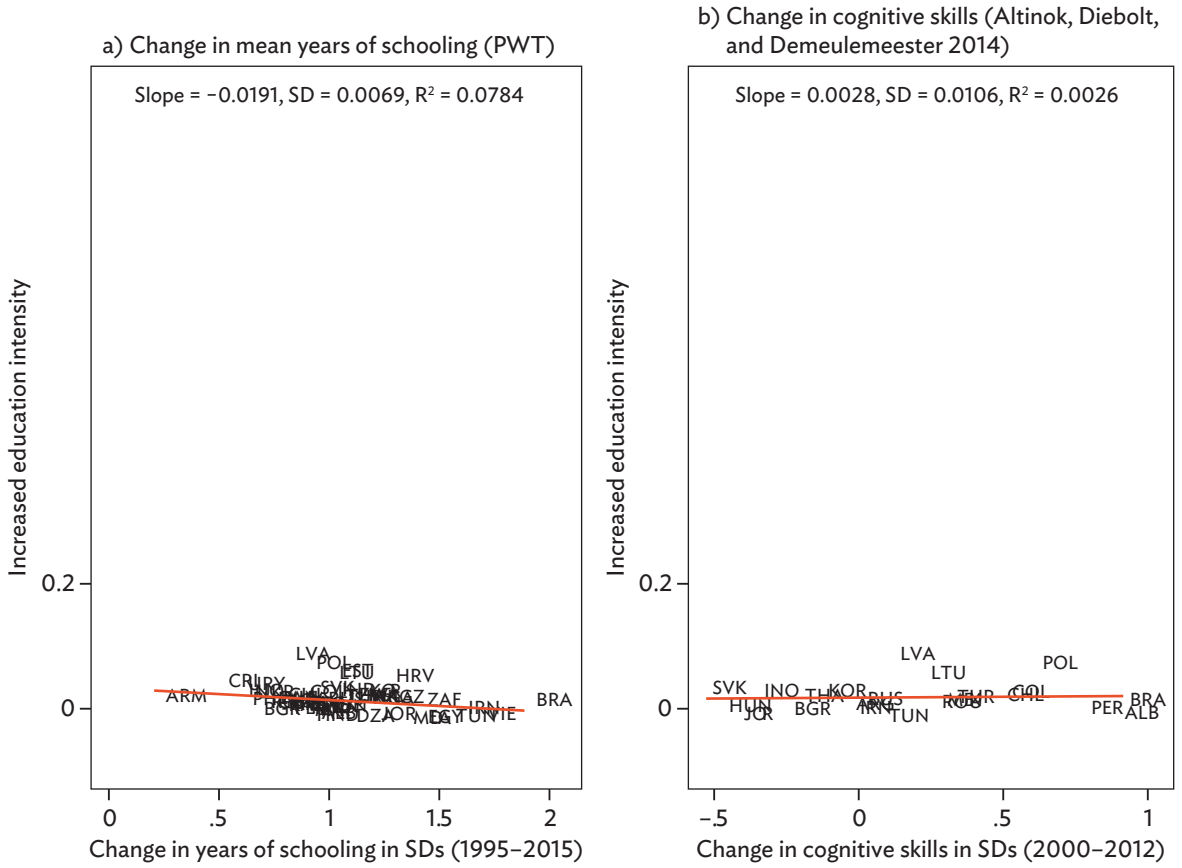


SD = standard deviation.

Note: See the list of countries in Appendix Table A1.

Source: Authors' calculations.

Figure 5: Education Expansions Were Not Associated with New Comparative Advantages in Education-Intensive Products



PWT = Penn World Table, SD = standard deviation.

Note: See the list of countries in Appendix Table A1.

Source: Authors' calculations.

Table 3: Explaining Shifts into Complex Products

	First-Stage Coefficient on Product Complexity				
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Per capita GDP in 1995 (Constant 2010 $)	0.001	−0.009	0.000	−0.005	0.007
	(0.008)	(0.012)	(0.014)	(0.009)	(0.011)
Quantity (Average years of schooling)	−0.016***			−0.025***	−0.016
	(0.005)			(0.007)	(0.010)
Quality A (Altinok, Diebolt, and Demeulemeester 2014)		0.006		0.017**	
		(0.008)		(0.008)	
Quality B (Hanushek and Woessmann 2009)			−0.003		0.003
			(0.010)		(0.011)

continued on next page

Table 3 *continued*

	First-Stage Coefficient on Product Complexity				
Dependent Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Quantity x Quality A				−0.002	
				(0.006)	
Quantity x Quality B					−0.005
					(0.010)
Constant	0.019	0.112	0.037	0.064	−0.021
	(0.070)	(0.100)	(0.114)	(0.073)	(0.093)
Observations	49	49	35	49	35
R-squared	0.132	0.028	0.006	0.242	0.109

GDP = gross domestic product.

Notes: Weighted least squares estimates, per specification (3b). The dependent variable is each country's coefficient on complexity from first-stage specification (2). Countries are weighted by the inverse of the standard error of that coefficient. Abbreviated variable names appear in Table 1. All standard errors are robust. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Table 4: Explaining Shifts into Education-Intensive Products

	First-Stage Coefficient on Education Intensity		
Dependent Variable	**(1)**	**(2)**	**(3)**
Per capita GDP in 1995 (Constant 2010 $)	0.009***	0.016**	0.003
	(0.003)	(0.007)	(0.004)
ΔQuantity	−0.018***		−0.019***
	(0.006)		(0.007)
ΔQuality A (Altinok, Diebolt, and Demeulemeester 2014, time-varying)		0.005	
		(0.009)	
Quality A (Altinok, Diebolt, and Demeulemeester 2014, time-varying)			0.020
			(0.013)
ΔQuantity x Quality A			−0.012
			(0.010)
Constant	−0.041*	−0.121*	0.009
	(0.024)	(0.061)	(0.031)
Observations	49	22	49
R-squared	0.200	0.128	0.303

GDP = gross domestic product.

Notes: Weighted least squares estimates, per specification (3c). The dependent variable is each country's coefficient on education intensity from first-stage specification (2). Countries are weighted by the inverse of the standard error of that coefficient. Abbreviated variable names appear in Table 1. All standard errors are robust. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Together, these results suggest that: (i) the main effect of education on the evolution of comparative advantage is to help countries overcome unfamiliarity, (ii) the effects of quantity and quality in this regard are difficult to tease apart,[16] and (iii) it is unlikely that education helped counties move into more complex or education-intensive products. This is because while our education measures vary a lot across countries, countries did not vary much in their tendency to achieve comparative advantages in complex and education-intensive products (i.e., $Var(\hat{\gamma}_c)$ and $Var(\hat{\delta}_c)$ are relatively small).

B. Single-Stage Analysis

Table 5 builds up our baseline estimates of specification (1). Columns (1) and (2) include terms suggested by complexity theory. Column (1) measures the quantity of education in 1995 by average years of schooling, while column (2) measures it by the proportions of the population aged 15+ that completed primary, secondary, and college education. Column (3) includes only the explanatory variables suggested by a factor abundance approach. As we do not have attainment rates by level in 2015, or time series for most countries on changes in education quality, we focus on the effects of changing average years of schooling. The three-way interaction allows that increasing years of schooling would be more supportive of the development of comparative advantage in education-intensive products if that education is of a high quality. Columns (4) and (5) combine the two sets of coefficients. Other than familiarity and the terms derived from lagged *RCA*, every variable entering the table, whether on its own or interacted, is normalized to have a mean of 0 and standard deviation of 1.

As expected, the coefficients on the lagged dependent variable indicate positive relationships between having nonzero and larger RCAs in 1995 and the probability of comparative advantage in 2015.[17]

We use country-clustered standard errors throughout this paper to ensure conservative inferences. These standard errors are extremely conservative, given that we have not sampled a small number of countries from a large universe, but rather attempted to include every low- and middle-income country for which the relevant data are available (Abadie et al. 2017). We report in the text any instances in which using unclustered robust standard errors alters our qualitative inferences.

[16] The raw correlation between Quality A (Quality B) and years of schooling in 1995 is 0.559 (0.588).

[17] Postestimation calculation of $E(CA_{p,t}|CA_{p,t})$ as well as a kernel-weighted local polynomial regression of the same (Appendix Figure A2) confirm a nearly monotonic positive relationship.

Table 5: Single-Stage Regression Results

	Economic Complexity				Heckscher–Ohlin–Vanek		Both			
	(1)		(2)		(3)		(4)		(5)	
Corrections for lagged RCA										
I{RCA in 1995 = 0}	−0.203***	(0.013)	−0.199***	(0.013)	−0.263***	(0.016)	−0.202***	(0.014)	−0.195***	(0.014)
[1 - I{RCA in 1995 = 0}] *ln(RCA in 1995)	0.049***	(0.002)	0.049***	(0.002)	0.066***	(0.003)	0.050***	(0.002)	0.050***	(0.002)
Familiarity	1.885***	(0.191)	2.132***	(0.156)			1.917***	(0.182)	2.163***	(0.156)
Familiarity x Quantity	−0.672***	(0.172)					−0.655***	(0.165)		
Familiarity x Quality A	0.181	(0.157)	0.004	(0.181)			0.114	(0.132)	−0.078	(0.162)
Familiarity x Primary			−0.332	(0.232)					−0.400	(0.240)
Familiarity x Secondary			−0.084	(0.183)					0.072	(0.204)
Familiarity x College			−0.134	(0.246)					−0.169	(0.257)
PCI x Quantity	−0.024**	(0.010)					−0.029**	(0.012)		
PCI x Quality A	0.028***	(0.008)	0.023***	(0.008)			0.021**	(0.010)	0.015*	(0.008)
PCI x Primary			−0.010	(0.009)					−0.013	(0.010)
PCI x Secondary			−0.022*	(0.012)					−0.018	(0.014)
PCI x College			0.008	(0.012)					0.002	(0.013)
ΔK/L x ProdKL					−0.005	(0.018)	0.015	(0.010)	0.014	(0.011)
ΔQuantity x ProdYrs					−0.005	(0.010)	−0.014	(0.013)	−0.013	(0.012)
Quality A x ProdYrs					0.031	(0.020)	0.002	(0.023)	0.007	(0.021)
ΔQuantity x Quality A x ProdYrs					−0.014	(0.019)	0.005	(0.023)	0.002	(0.021)

continued on next page

Table 5 *continued*

	Economic Complexity		Heckscher–Ohlin–Vanek	Both	
	(1)	(2)	(3)	(4)	(5)
Country fixed effects	✓	✓	✓	✓	✓
Product fixed effects	✓	✓	✓	✓	✓
Observations	60,760	59,520	54,560	54,560	53,320
R-squared	0.266	0.266	0.242	0.264	0.264
Baseline without education variables					
Observations	60,760		54,560	54,560	
R-squared	0.261		0.239	0.259	

PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: Estimates based on linear probability specification (1). Country-clustered standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

The results in Table 5 are consistent with those in the previous section.[18] The positive significant coefficient on familiarity indicates that countries of average education quality and quantity are much more likely to develop comparative advantage in products that are more familiar to them. The negative sign on the familiarity–quantity interaction is consistent with education helping to develop comparative advantages in relatively unfamiliar products. Evidence of this effect is strongest for primary school, and indeed when using unclustered standard errors, the p-value on the interaction between familiarity and primary attainment is effectively zero.

This effect is of a modest size. Consider two products and two countries. Assume, for each product, that its lagged RCA and familiarity are the same in both countries, but that one product is 0.20 points (roughly 2 standard deviations, Table 1) more familiar than the other product in both countries. We also assume that both countries have average education quality, that country A has average education quantity, and that country B's years of schooling are 1 standard deviation higher than the mean. The estimates in column (4) then indicate that in country A, the probability of comparative advantage in 2015 is 38 percentage points higher in the more familiar product than in the unfamiliar product. However, in country B, this probability will only be 25 percentage points higher in the more familiar product. The 13 percentage point difference attributable to education is sizable compared to the mean probability of comparative advantage of 17 percentage points.[19] On the other hand, despite this large effect of education quantity, there are no countries in our sample with enough years of schooling to eliminate path dependence (i.e., there are no countries for which $\beta_E + \beta_{EF} E_{c,t0} \leq 0$).

As was the case with the second-stage results in Table 3, the single-stage results in Table 5 paint a mixed picture regarding our second hypothesis—that (controlling for familiarity) education helps develop comparative advantages in more complex products. Higher average years of schooling in 1995 are associated with developing comparative advantages in less complex products, but higher quality education is associated with developing comparative advantages in more complex products. The effect of school quality on comparative advantage in complex products is roughly one-third the size of the effect of school quantity on comparative advantage in unfamiliar products.[20]

Regressions (3)–(5) provide no significant evidence that growth in education shifts countries toward education-intensive products. The point estimates suggest that countries with larger increases in average years of schooling between 1995 and 2015 tended to develop comparative advantages in *less* education-intensive products. They also do not support the possibility that increasing years of schooling shifts comparative advantages toward education-intensive products more reliably in countries with higher quality education. In theory, this absence of measured effects of changes in education could be driven by attenuation biases arising from this differenced variable's low signal to noise ratio (Krueger and Lindahl 2001). However, the fact that countries vary so little in the extent to

---

[18] This was expected. Familiarity and product complexity are orthogonal to each other by definition (familiarity with a given product varies across countries, while the complexity of that product does not), practically guaranteeing that estimating familiarity and complexity interactions together will not alter the implied coefficients on either.

[19] In country A, the difference in probability of comparative advantage is 0.2 x 1.917 = 0.383. In country B, it is 0.2 x (1.917-0.655) = 0.252.

[20] Consider two products 2 standard deviations apart in complexity, and two countries 1 standard deviation apart in education quality but with the same years of schooling. Holding all other variables constant across products and countries, the difference between the probabilities of comparative advantage in the more and less complex products would be 4.2 percentage points (= 0.021 x 2 x 1) greater in the better-educated country. Compare this to the 13 percentage point effect of school quantity in unfamiliar products.

which they shifted into education-intensive products (Figure 2c) suggests that our estimate of δ_{EE} would have been small even in the absence of measurement error.

Table 5 also reports R-squared statistics for models that restrict the coefficients on all education terms to be zero in order to assess education's explanatory power. Allowing that education could be useful for overcoming unfamiliarity, product complexity or education intensity adds very little to model R-squared.[21] In combination with the large and statistically significant coefficient on the interaction between familiarity and education quantity, this indicates that education is useful for overcoming unfamiliarity, but that past specialization patterns, product and country characteristics, are still the main determinants of subsequent specialization patterns.

C. Robustness Tests

The previous two sections suggest that the primary beneficial effect of education when seeking to gain comparative advantage in new products is that it can help overcome a lack of familiarity with target products. The level of education most strongly associated with this effect is primary school. Conclusions about whether it is the quantity or the quality of education that matters depend upon the measure of education quality used. Next, we check whether these findings are robust.

Table 6 is analogous to Table 5, but replaces the Altinok, Diebolt, and Demeulemeester (2014) estimate of education quality with estimates on a smaller number of countries from Hanushek and Woessmann (2009). The results change very little. The most powerful role of education is still to overcome unfamiliarity.

[21] For example, the R-squared rises from 0.2614 to 0.2664 when the four education-related terms in column (1) are added. In the linear probability context, R-squared captures the difference in the predicted probabilities of "success"—in our case, a comparative advantage in the product—between observed cases of success and failure (Gronau 1998).

Table 6: Single-Stage Regression with Hanushek and Woessmann's (2009) Education Quality Measure

	Economic Complexity				Heckscher–Ohlin–Vanek		Both			
	(1)		(2)		(3)		(4)		(5)	
Corrections for lagged RCA										
I{RCA in 1995 = 0}	-0.220***	(0.016)	-0.211***	(0.016)	-0.288***	(0.019)	-0.219***	(0.018)	-0.208***	(0.018)
[1 - I{RCA in 1995 = 0}] **ln*(RCA in 1995)	0.053***	(0.003)	0.053***	(0.002)	0.071***	(0.004)	0.053***	(0.003)	0.053***	(0.003)
Familiarity	1.914***	(0.250)	2.193***	(0.191)			1.929***	(0.249)	2.247***	(0.195)
Familiarity x Quantity	-0.485*	(0.258)					-0.566**	(0.252)		
Familiarity x Primary			-0.130	(0.222)					-0.233	(0.248)
Familiarity x Secondary			-0.010	(0.203)					0.020	(0.200)
Familiarity x College			-0.303	(0.341)					-0.316	(0.335)
Familiarity x Quality B	-0.036	(0.234)	-0.251	(0.224)			-0.004	(0.230)	-0.220	(0.238)
PCI x Quantity	-0.019	(0.015)					-0.024	(0.016)		
PCI x Quality B	0.023*	(0.013)	0.022*	(0.012)			0.018	(0.015)	0.015	(0.014)
PCI x Primary			-0.008	(0.012)					-0.010	(0.014)
PCI x Secondary			-0.023	(0.015)					-0.023	(0.016)
PCI x College			0.009	(0.019)					0.000	(0.019)
ΔK/L x ProdKL					-0.003	(0.021)	0.018	(0.012)	0.019	(0.013)
ΔQuantity x ProdYrs					-0.005	(0.014)	-0.018	(0.022)	-0.014	(0.019)
Quality A x ProdYrs					0.011	(0.016)	-0.007	(0.019)	-0.009	(0.019)
ΔQuantity x Quality A x ProdYrs					0.003	(0.016)	0.012	(0.020)	0.015	(0.019)
Country fixed effects	✓		✓		✓		✓		✓	
Product fixed effects	✓		✓		✓		✓		✓	
Observations	43,400		42,160		40,920		40,920		39,680	
R-squared	0.276		0.276		0.251		0.274		0.274	

PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: The estimation is based on the linear probability model, per specification (1). Country-clustered standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Table 7 checks whether our inability to confirm a role for education accumulation in promoting the development of comparative advantage results from specification errors. Column 2 replicates Table 5, column 3 for comparison's sake. Column 3 strips it of quality effects, and column 5 checks to see whether quality accumulation over time is associated with moves into education-intensive products (note the sharply reduced sample size). Columns 1 and 4 provide restricted regressions that eliminate the education variables. Physical capital accumulation is associated with a movement into capital-intensive products in columns 4 and 5. However, the education effects are always statistically insignificant, usually of the wrong sign, and their inclusion adds little to R-squared. As Figure 2 demonstrates, their insignificance is unlikely to be driven by attenuation.

Table 7: Alternative Estimates of Heckscher–Ohlin–Vanek Effects

	(1)	(2)	(3)	(4)	(5)
ΔK/L x ProdKL	0.001	−0.005	−0.000	0.046**	0.044*
	(0.019)	(0.018)	(0.019)	(0.021)	(0.021)
ΔQuantity x ProdYrs		−0.005	−0.010		−0.009
		(0.010)	(0.010)		(0.016)
Quality A x ProdYrs		0.031			
		(0.020)			
ΔQuantity x Quality A x ProdYrs		−0.014			
		(0.019)			
ΔQuality A x ProdYrs					0.004
					(0.016)
Country fixed effects	✓	✓	✓	✓	✓
Product fixed effects	✓	✓	✓	✓	✓
Corrections for lagged RCAs	✓	✓	✓	✓	✓
Observations	54,560	54,560	54,560	26,040	26,040
R-squared	0.239	0.242	0.239	0.259	0.259

RCA = revealed comparative advantage.

Notes: The estimation is based on linear probability model. Columns 1–3 include all countries in our sample; columns 4–5 are based on the sample with the time-varying Altinok quality data. Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Table 8 examines the possibility that educated countries develop comparative advantage in products that are sophisticated in ways not captured by the Product Complexity Index (PCI). To do so, columns (2)–(4) replace the PCI with sophistication measures derived from the education endowments of those countries that successfully export them. The interactions between these sophistication measures and education quantity continue to carry a negative coefficient, and making these substitutions does not qualitatively alter our findings on the interactions between education quantity and familiarity.

Table 8: Robustness to Alternative Measures of Product Sophistication

	Measure of Sophistication							
	PCI		ProdYrs		ProdCog		ProdColl	
	(1)		(2)		(3)		(4)	
Familiarity	1.917***	(0.182)	2.052***	(0.189)	2.005***	(0.181)	2.066***	(0.197)
Familiarity x Quantity	−0.655***	(0.165)	−0.404***	(0.121)	−0.366***	(0.124)	−0.312**	(0.127)
Familiarity x Quality A	0.114	(0.132)	−0.072	(0.112)	−0.070	(0.119)	−0.128	(0.120)
Soph x Quantity	−0.029***	(0.012)	−0.012**	(0.006)	−0.041*	(0.024)	−0.004*	(0.002)
Soph x Quality A	0.021**	(0.010)	0.007	(0.025)	0.043	(0.026)	−0.002	(0.002)
ΔK/L x ProdKL	0.015	(0.010)	0.014	(0.011)	0.013	(0.011)	0.013	(0.012)
ΔQuantity x ProdYrs	−0.014	(0.013)	−0.012	(0.014)	−0.007	(0.013)	−0.006	(0.013)
Quality A x ProdYrs	0.002	(0.023)			−0.005	(0.024)	0.000	(0.023)
ΔQuantity x Quality A x ProdYrs	0.005	(0.023)	0.009	(0.023)	0.011	(0.023)	0.014	(0.022)
Country fixed effects	✓		✓		✓		✓	
Product fixed effects	✓		✓		✓		✓	
Corrections for lagged RCAs	✓		✓		✓		✓	
Observations	54,560		54,560		54,560		54,560	
R-squared	0.264		0.263		0.263		0.263	

PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: The estimation is based on linear probability model, per specification (1). Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Table 9 shows that our results are not sensitive to most changes in the RCA cutoff used to decide whether a country exports a product with comparative advantage. The only discernable shift is that the interaction between familiarity and years of schooling declines as the bar for achieving comparative advantage is lowered.

Table 9: Robustness to Changes in Revealed Comparative Advantage Cutoffs

	Cutoff = 0.5		Cutoff = 0.8		Cutoff = 1		Cutoff = 2	
	(1)		(2)		(3)		(4)	
Familiarity	1.985***	(0.189)	2.017***	(0.184)	1.917***	(0.182)	1.376***	(0.166)
Familiarity x Quantity	−0.309*	(0.168)	−0.512***	(0.167)	−0.655***	(0.165)	−0.708***	(0.160)
Familiarity x Quality A	−0.068	(0.118)	0.019	(0.129)	0.114	(0.132)	0.099	(0.141)
PCI x Quantity	−0.026*	(0.014)	−0.028**	(0.013)	−0.029**	(0.012)	−0.023***	(0.008)
PCI x Quality A	0.033**	(0.012)	0.025**	(0.010)	0.021**	(0.010)	0.008	(0.006)
ΔK/L x ProdKL	0.018	(0.012)	0.017*	(0.010)	0.015	(0.010)	0.006	(0.008)
ΔQuantity x ProdYrs	−0.010	(0.016)	−0.014	(0.014)	−0.014	(0.013)	−0.016*	(0.009)
Quality A x ProdYrs	0.009	(0.029)	0.003	(0.024)	0.002	(0.023)	−0.004	(0.017)
ΔQuantity x Quality A x ProdYrs	−0.006	(0.029)	0.003	(0.024)	0.005	(0.023)	0.011	(0.017)
Country fixed effects	✓		✓		✓		✓	
Product fixed effects	✓		✓		✓		✓	
Corrections for lagged RCAs	✓		✓		✓		✓	
Observations	54,560		54,560		54,560		54,560	
R-squared	0.317		0.283		0.208		0.264	

PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: The estimation is based on linear probability model, per specification (1). Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

Table 10 checks whether our education-related coefficients might be biased by the omission of controls for national institutional and infrastructure quality or for openness to FDI.[22] This is done in a separate regression for each control variable by adding an interaction between the control variable and both familiarity and the PCI. It also includes, at the end, a regression controlling for the first principal component of all 10 institutional and infrastructure quality measures. None of these changes makes any difference to our conclusions regarding the role of education. In contrast to education, no infrastructure controls predict movement into unfamiliar or complex products. Only one institutional control—an index of regulatory quality—predicts developing comparative advantage in complex products, and none predict developing comparative advantage in unfamiliar products. And, countries receiving more FDI were more likely to develop comparative advantage in products with which they were already familiar.[23]

[22] FDI is widely considered to have been an important source of new productive knowledge for late industrializing countries (e.g., Felker, Jomo, and Rasiah 2013). We have not presented all of these regressions here in the interests of parsimony. Regressions controlling for access to information and communication technology, corruption, the number of procedures needed to open a small business, and kilometers of road per square kilometer yield the same qualitative results as those shown here. Mehta and Felipe (2014) list the sources of these institutional variables, while FDI data are drawn from the World Bank's World Development Indicators database.

[23] This is consistent with FDI flowing mainly toward industries relying on well-established capabilities. While intuitive, this idea cannot be tested without data on FDI receipts disaggregated by industry.

Table 10: Robustness to Corrections for Institutions, Infrastructure, and Foreign Direct Investment

	Institutional Control								
	Contracts	Infrastructure	Days	Internet	RQ	RL	FDI/Exports	FDI/GDP	Principal Component
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Familiarity	1.961***	2.018***	1.897***	3.322***	1.913***	1.986***	2.031***	1.949***	2.148***
Familiarity x Institutional control	−0.026	0.044	0.042	2.371	0.309	−0.069	0.271**	0.065	−0.206
Familiarity x Quantity	−0.663***	−0.675***	−0.646***	−0.749***	−0.738***	−0.659***	−0.640***	−0.683***	−0.920***
Familiarity x Quality A	0.083	0.089	0.121	0.096	0.059	0.112	0.080	0.094	0.129
PCI x Institutional control	0.000	0.002	−0.001	0.377	0.019***	0.011	−0.003	−0.006	−0.001
PCI x Quantity	−0.031*	−0.033**	−0.029**	−0.045***	−0.035***	−0.029**	−0.026**	−0.025*	−0.085***
PCI x Quality A	0.023	0.023*	0.021**	0.026**	0.018*	0.017	0.019*	0.020**	0.064***
ΔK/L x ProdKL	0.013	0.014	0.015	0.014	0.013	0.016	0.017*	0.015	0.010
ΔQuantity x ProdYrs	−0.019	−0.019*	−0.015	−0.026**	−0.014	−0.014	−0.015	−0.015	−0.067**
Quality A x ProdYrs	0.007	0.005	0.001	−0.003	0.001	−0.001	0.004	0.002	−0.004
ΔQuantity x Quality A x ProdYrs	0.001	0.003	0.006	0.009	0.007	0.008	0.004	0.006	0.013
Country fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Product fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Corrections for lagged RCAs	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	43,400	47,120	53,320	48,360	54,560	54,560	54,560	54,560	22,320
R-squared	0.271	0.273	0.264	0.274	0.266	0.265	0.265	0.265	0.303
F-test	4.834	7.241	6.275	6.008	6.019	6.891	8.691	7.981	17.658
Prob>F	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

FDI = foreign direct investment, GDP = gross domestic product, PCI = Product Complexity Index, RCA = revealed comparative advantage, RL = rule of law, RQ = regulatory quality.

Notes: FDI, exports, and GDP are nominal values, drawn from World Development Indicators. Sources of infrastructure and institutional variables appear in Mehta and Felipe (2014).

The estimation is based on linear probability model, per specification (1). Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.
Source: Authors' calculations.

Table 11: Correcting for Prior Industrial Dynamism

	Measure of Prior Industrial Dynamism Used					
	Real GDPPC Growth Rate		Labor Productivity Growth Rate		First-Stage Coefficients on Familiarity and PCI	
	(1)	(2)	(3)	(4)	(5)	(6)
Familiarity	1.753***	2.352***	1.983***	2.358***	1.710***	1.940***
Familiarity x Quantity	−0.830***		−0.691***		−0.702***	
Familiarity x Quality A	−0.107	−0.103	0.021	0.007	0.134*	−0.060
Familiarity x Primary		−0.605***		−0.438***		−0.331***
Familiarity x Secondary		0.144		−0.075		0.113
Familiarity x College		−0.163		−0.216**		−0.231**
Familiarity x Dynamism (1975–195)	0.857***	0.780***	0.583***	0.410**	−0.043	−0.650
Familiarity x Dynamism (1985–95)	−1.544***	−1.812***	−0.964**	−1.258***	−1.546***	−1.220***
PCI x Quantity	−0.032***		−0.020***		−0.019***	
PCI x Quality A	−0.009**	−0.019***	0.010**	0.004	0.019***	0.010***
PCI x Primary		−0.009**		−0.006*		0.004
PCI x Secondary		−0.033***		−0.017***		−0.009**
PCI x College		0.017***		−0.000		−0.009**
PCI x Dynamism (1975–95)	0.043***	0.057***	0.019***	0.021***	0.102***	0.112***
PCI x Dynamism (1985–95)	0.005	−0.019	0.040***	0.038***	0.669***	0.678***
ΔK/L x ProdKL	0.011***	0.008	0.011***	0.010*	0.010**	0.008
ΔQuantity x ProdYrs	−0.031***	−0.018***	−0.025***	−0.018***	−0.034***	−0.029***
Quality A x ProdYrs	0.003	0.006	0.010	0.005	0.020***	0.007
ΔQuantity x Quality A x ProdYrs	0.001	−0.000	−0.005	−0.001	−0.016***	−0.003
Country fixed effects	✓	✓	✓	✓	✓	✓
Product fixed effects	✓	✓	✓	✓	✓	✓
Corrections for lagged RCAs	✓	✓	✓	✓	✓	✓
Observations	33,480	32,240	38,440	37,200	42,160	40,920
R-squared	0.313	0.313	0.293	0.292	0.294	0.292

GDPPC = gross domestic product per capita, PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: GDP per capita data are from World Development Indicators, labor productivity from Penn World Tables, and first-stage coefficients are explained in the text. Estimates from a linear probability model of comparative advantage in 2015, per specification (1). Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

The correlations we report between education and post-1995 export dynamism (movements into unfamiliar and complex products) could reflect a tendency for intrinsically dynamic countries to invest in education. To check whether this is plausible, we first calculate four different measures of prior industrial dynamism. We calculate each of these over two prior time intervals: 1975–1995 and 1985–1995. The four measures are: (a) real per capita GDP growth; (b) real labor productivity growth; and the coefficients in country-by-country first-stage regressions (specification 2) using Standard International Trade Classification export data from the start and end of the time interval on (c) familiarity (i.e., prior $\hat{\beta}_c$) and (d) complexity (i.e., prior $\hat{\gamma}_c$).

The correlations between these years of schooling in 1995 and these four dynamism measures are higher in 1985–1995 than during 1975–1995. In 1985–1995 they are, in the same order: (a) 0.195, (b) 0.132, (c) 0.408, and (d) −0.237. Only the third of these correlations, between prior $\hat{\beta}_c$ and 1995 years of schooling, is statistically significant at even the 10% level.

Table 11 checks whether our finding that education helps to overcome familiarity reflects a spurious correlation owing to these relationships. Each regression interacts prior dynamism with familiarity and the PCI. Our preferred estimates are those appearing in columns 5 and 6, for two reasons: they use the same notion of dynamism in the earlier and later periods; and they include the dynamism measure that correlates most strongly with schooling (and so, a priori, is the likeliest source of omitted variables bias). With the limited exception of one of the four coefficients in the two less preferred specifications, the results do confirm that countries whose export mixes were more dynamic between prior to 1995 were more likely to move into unfamiliar and complex products between 1995 and 2015. However, perhaps because most measures of prior dynamism are not strongly correlated with education, these corrections do not alter our findings that countries with higher average years of schooling and higher primary attainment were more likely to develop comparative advantage in unfamiliar products. Moreover, the role of college and especially primary attainment in overcoming familiarity appears much stronger once these corrections are made.

Table 12 checks our results for robustness to corrections for neighborhood effects—spillovers of knowledge and supply chains to nearby countries (Bahar, Hausmann, and Hidalgo 2014; Bahar et al. 2019). We capture these effects by correcting for the weighted average familiarity in 1995 of every other country with the product in question—where the weight is simply the inverse of that country's geographic distance from country c, and an analogous measure of neighbors' RCA in 1995. Column (1) reprises our main results, which do not change when correcting separately for neighbors' familiarity or RCA (columns 2 and 4), for these variables interacted with education (columns 3 and 5), or all of the above (column 6).[24]

[24] Education quality is associated with developing comparative advantage in products with which neighbors are familiar—consistent with quality education facilitating knowledge flows across borders, but that neighbors do not actively produce—consistent with it helping to overcome competition.

Table 12: Correcting for Neighbors' Specialization Patterns

	(1)	(2)	(3)	(4)	(5)	(6)
Familiarity	1.917***	1.916***	1.924***	1.918***	1.911***	1.921***
Familiarity x Quantity	−0.655***	−0.655***	−0.646***	−0.654***	−0.659***	−0.646***
Familiarity x Quality A	0.114	0.115	0.062	0.112	0.118	0.064
PCI x Quantity	−0.029**	−0.029**	−0.030**	−0.028**	−0.028**	−0.030**
PCI x Quality A	0.021**	0.021**	0.023**	0.021**	0.021**	0.023**
ΔK/L x ProdKL	0.015	0.015	0.013	0.015	0.015	0.013
ΔQuantity x ProdYrs	−0.014	−0.014	−0.016	−0.014	−0.014	−0.016
Quality A x ProdYrs	0.002	0.002	−0.004	0.002	0.002	−0.004
ΔQuantity x Quality A x ProdYrs	0.005	0.005	0.011	0.006	0.005	0.011
Neighbors' familiarity with the product		−0.001	−0.015			−0.010
Neighbors' familiarity x Quantity			−0.018			−0.025
Neighbors' familiarity x Quality A			0.068**			0.076**
Neighbor's RCA in the product				−0.002	−0.002	−0.003
Neighbor's RCA x Quantity					0.002	0.002
Neighbor's RCA x Quality A					−0.002	−0.003***
Country fixed effects	✓	✓	✓	✓	✓	✓
Product fixed effects	✓	✓	✓	✓	✓	✓
Corrections for lagged RCAs	✓	✓	✓	✓	✓	✓
Observations	54,560	54,560	54,560	54,560	54,560	54,560
R-squared	0.264	0.264	0.265	0.264	0.265	0.265

PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: Neighborhood familiarity is the average familiarity of all other countries in the world with product *p*, weighted by the inverse of each country's distance from country *c*. Neighborhood RCA is analogous. Estimates from a linear probability model of comparative advantage in 2015, per specification (1). Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

D. Core and Peripheral Products

Table 13 presents results on the subsets of core and peripheral products. Our main results, that countries with higher average years of schooling were substantially more likely to develop comparative advantage in unfamiliar products, holds for both subsets.

Table 13: Regressions in Subsamples of Core and Peripheral Products

	Core Products			Peripheral Products		
	(1)	(2)	(3)	(4)	(5)	(6)
Familiarity	2.906***	2.523***	2.935***	2.101***	1.837***	2.118***
Familiarity x Quantity		−0.700**			−0.901***	
Familiarity x Quality A		0.536	0.136		−0.300	−0.498*
Familiarity x Primary			−0.075			−0.677**
Familiarity x Secondary			−0.328			0.206
Familiarity x College			0.148			−0.182
PCI x Quantity		−0.025*			−0.024**	
PCI x Quality A		0.016	0.014		−0.012*	−0.015**
PCI x Primary			−0.024**			−0.027***
PCI x Secondary			−0.017			−0.002
PCI x College			0.019			0.007
ΔK/L x ProdKL	0.014	0.022	0.012	0.027***	0.038***	0.037***
Quality A x ProdYrs		−0.018	−0.007		−0.000	0.005
ΔQuantity x Quality A x ProdYrs		0.029	0.016		0.008	0.003
ΔQuantity x ProdYrs		0.011	0.007		−0.010	−0.011
Country fixed effect	✓	✓	✓	✓	✓	✓
Product fixed effect	✓	✓	✓	✓	✓	✓
Corrections for lagged RCAs	✓	✓	✓	✓	✓	✓
Observations	10,164	10,164	9,933	10,208	10,208	9,976
R-squared	0.2557	0.2583	0.2574	0.3189	0.3295	0.3300

PCI = Product Complexity Index, RCA = revealed comparative advantage.

Notes: The estimation is based on linear probability model, per specification (1). Significance using country-clustered standard errors: *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' calculations.

However, there are four differences between core and peripheral products. First, while countries with education of a higher quality exhibit a slight tendency to develop comparative advantages in unfamiliar peripheral products, this is reversed for core products.[25] Second, primary attainment was associated with learning to produce unfamiliar peripheral products, but this is not the case for core products. Third, capital accumulation is significantly associated with movements toward capital-intensive products among peripheral products, but not among core products. Fourth, the effects of familiarity are larger among core than among peripheral products. All of these results suggest that it is hard to develop comparative advantage in core products, and that general education is particularly important for moving up lower rungs of the product ladder.

[25] When using unclustered standard errors, the coefficients on the familiarity–quality interaction for core products in columns (2) and (3) have p-values of 0.008 and 0.092. The analogous p-values in the peripheral regressions (5) and (6) are 0.018 and 0.000.

V. CONCLUSIONS

We have analyzed the relationship between education and the evolution of comparative advantage among low- and middle-income countries. We found strong evidence consistent with education helping countries to develop comparative advantages in unfamiliar products—products that are unrelated to those in which they already have comparative advantages. In contrast, controlling for familiarity, more educated countries were not much more likely to develop comparative advantages in complex products, and those countries that experienced faster increases in education levels were not more likely to develop comparative advantages in education-intensive products. Taken together, these results are more obviously supportive of an approach to the development of comparative advantage that emphasizes relatedness between products than of a Heckscher–Ohlin–Vanek approach. The relatedness approach emphasizes that the process is evolutionary, with productive capabilities developing in path-dependent fashion. Education's core contribution to the process, it appears, is to reduce this path dependence by facilitating longer leaps into previously unfamiliar products.

While there are no plausible instruments for our key independent variables, several types of auxiliary evidence suggest that these results should be taken seriously. They are robust to many changes in specification, to changes in how product sophistication is operationalized, and to corrections for national, institutional, infrastructural, and FDI-related variables. We also find that primary education and education of a higher quality is most strongly associated with overcoming unfamiliarity when developing comparative advantage in peripheral products, but not in core products—which is consistent with the widely accepted idea that core products require the acquisition of more capabilities, but also with the idea that quality basic education is important for amassing basic capabilities. Perhaps most importantly, our results are *not* explained by the fact that the most educated countries in 1995 tended to be a little more dynamic in the prior decades. Lacking the micro data needed to test alternative mechanisms, we remain agnostic regarding *how* education might facilitate shifts toward unfamiliar products. Three possibly complementary mechanisms appear worth exploring in light of related research. First, developing new capabilities requires the acquisition of tacit knowledge and its transfer across less related activities (Hidalgo et al. 2018). Education may well facilitate both a more rapid acquisition of tacit knowledge through learning by doing, and greater efficiency in assembling teams with the right knowledge mix. Second, a more educated workforce might be better placed to acquire the additional pieces of book knowledge required to make inroads into new industries. Third, education may permit actors to identify and take advantage of new disequilibrium opportunities (Schultz 1975). There is evidence that new industrial strengths develop where downstream industries already exist (Bahar et al. 2019), which is consistent with changes in the organization of supply chains creating such disequilibrium opportunities.

These results do *not* mean that education is unimportant for developing countries seeking to move from peripheral products to complex, core products. They imply that education is important, but not because products in the core are complex. Education is important because core products are unfamiliar to these countries, given what they already produce. Put differently, education would be no less helpful in the unlikely scenario that a developing country wished to develop comparative advantage in an unfamiliar *peripheral* product.

Finally, we caution that education differences account for rather little of the cross-country variation in export diversification. Education variables have limited explanatory power overall, and while education is associated with less path dependence, it does not eliminate it. Even the most educated countries tend to develop comparative advantages in products that are related to those they already produce. This path dependence suggests that industrial development needs to be partly a deliberate process, with governments facilitating the development of a series of stepping-stone industries that the economy can traverse on its way to the core. Our findings suggest that investments in education allow the stepping stones to be spaced a little further apart.

APPENDIX: DERIVING THE ESTIMATION EQUATION FROM A SIMPLE EVOLUTIONARY MODEL OF EDUCATION, TACIT AND BOOK KNOWLEDGE, AND COMPARATIVE ADVANTAGE

Consider a standard 2-country, P-product Ricardian trade model, with labor, L, as the only factor of production. Value added in product p given by $Q_p = A_p L_p$ in Home and $Q_p^* = A_p^* L_p^*$ in the rest of the world. As usual, equilibrium in period t requires a ranking of products by Home's relative productivity, $\Omega_t(p) = A_{p,t}/A_{p,t}^*$, and that Home allocates its labor to meet global demand for the products at the top of this ranking. Thus, the likelihood that Home has a comparative advantage in p increases monotonically in $\Omega_t(p)$.

While relatedness can arise for many reasons (see footnotes 3 and 7), we adopt here the most widely discussed assumption in the literature—that it arises due to the importance of tacit knowledge for productivity. Let productivity in product p be log-separable in the tacit and book knowledge relevant to producing it: $A_{p,t} = B_{p,t}^{\theta_B} T_{p,t}^{\theta_T}$. We represent current tacit knowledge relevant to p as translog in prior familiarity with similar products, $F_{p,t0}$, and education, E, so that, ignoring quadratic terms: $ln(T_{p,t1}) = \tau_0 + \tau_1\, ln(F_{p,t0}) + \tau_2\, ln(E) + \tau_3\, ln(F_{p,t0})\, ln(E)$.Naturally, education and familiarity with p could increase the tacit knowledge relevant to p, so that $\tau_1, \tau_2 \geq 0$. If education permits tacit knowledge to be acquired in, or transferred from, less related production activities, then $\tau_3 < 0$. We can also represent book knowledge relevant to p as translog in the complexity of product p, C_p, and in education. Again, ignoring quadratic terms: $ln(B_{p,t}) = \beta_0 + \beta_1\, ln(C_p) + \beta_2\, ln(E) + \beta_3\, ln(C_p)\, ln(E)$. Education could make book knowledge easier to acquire $\beta_2 > 0$, and we hypothesize that this is especially true of the book knowledge required to produce the most complex products: $\beta_3 > 0$.[26]

Then, assuming that these knowledge production functions are the same in Home and the rest of the world, and using lower case to represent logarithms of variables described earlier, we have:

$$ln\Omega_{t1}(p) = \overbrace{[\theta_B\beta_2 + \theta_T\tau_2]}^{(+)}(e - e*) + \overbrace{\theta_B\beta_3}^{(+)}\, c_p(e - e*) \tag{A1}$$

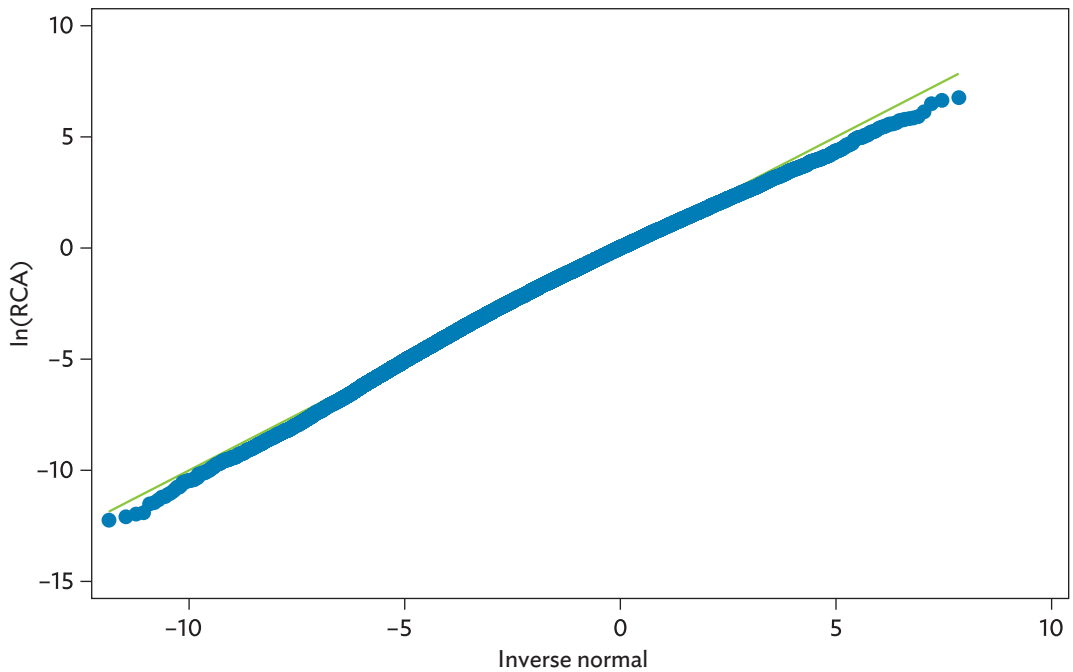
$$+ \underbrace{\theta_T\tau_1}_{(+)}\left(f_{p,t0} - f_{p,t0}^*\right) + \underbrace{\theta_T\tau_3}_{(-)}\left(ef_{p,t0} - e^* f_{p,t0}^*\right)$$

Notice that $f_{p,t0}$ and $f_{p,t0}^*$ are functions of the full vector of $\Omega_{t0}(p)$, so that the P instances of (A1) together determine how comparative advantage evolves between $t0$ and $t1$. As the probability that Home has a comparative advantage in p is monotonic in $\Omega_t(p)$, this equation becomes the right-hand side for a binary model of comparative advantage. In a multicountry setting, small-country assumptions make e^* and $f_{p,t0}^*$ constant across countries, and the first term in (A1) is absorbed into the country fixed effects. Assume that education is exponential in measured schooling, and normalize education so that $e^* \approx 0$. What remains becomes the first line on the right-hand side of specification (1), with the following expectations for the parameters in that specification:

[26] The restrictions on functional form just enumerated are either without loss, or are testable. Three of the four quadratic terms dropped from the tacit and book knowledge functions are absorbed into country and product fixed effects; and we have tested for robustness to correcting for familiarity-squared (available on request).

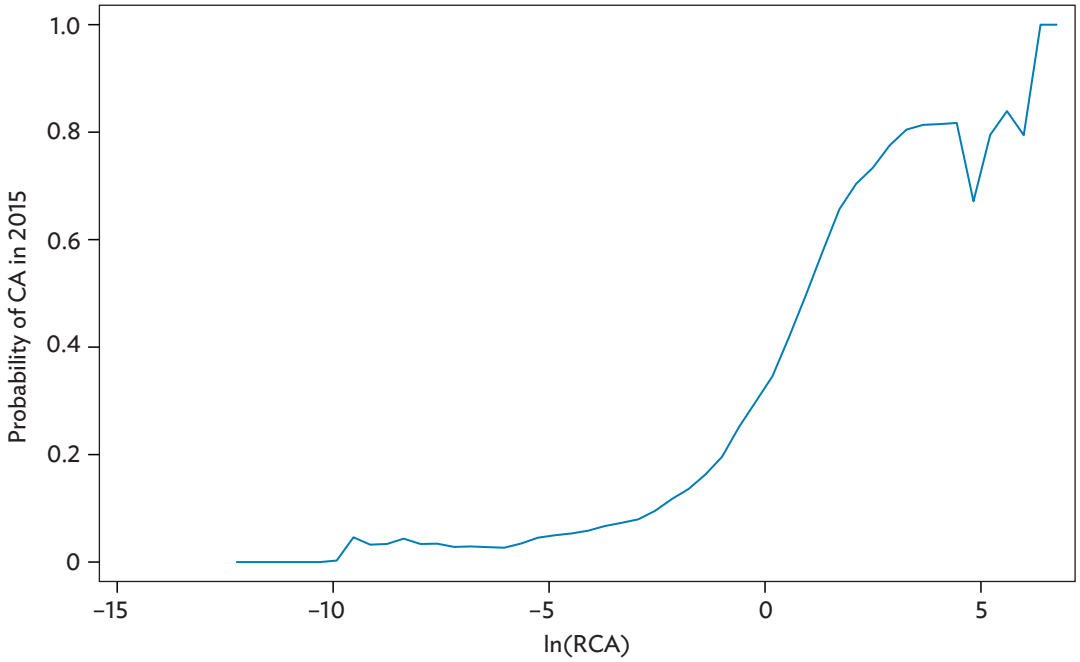
(i) Industrial development is path dependent, in the sense that $\beta_F = \theta_T \tau_1 > 0$.
(ii) Education can reduce path dependence by facilitating tacit knowledge acquisition in unfamiliar products: $\beta_{EF} = \theta_T \tau_3 < 0$.
(iii) Countries with more educated workforces are likely to export more complex products: $\gamma_{EC} = \theta_B \beta_3 > 0$.

Figure A1: Q–Q Plot for ln(RCA)



RCA = revealed comparative advantage.
Note: Distributional test for nonzero RCA values in 1995.
Source: Authors' calculations.

Figure A2: Probability of Comparative Advantage in 2015 Conditional on Nonzero Revealed Comparative Advantages in 1995



CA = comparative advantage, RCA = revealed comparative advantage.

Notes: Kernel-weighted local polynomial regression. For country-product pairs with zero exports in 1995, the probability of CA in 2015 is 0.046. Local mean smoothing is applied.

Source: Authors' calculations.

Table A1. List of Countries

Countries	ISO Codes	Countries	ISO Codes
Albania	ALB	Mauritius	MUS
Algeria	DZA	Mexico	MEX
Argentina	ARG	Moldova	MDA
Armenia	ARM	Mongolia	MON
Brazil	BRA	Morocco	MAR
Bulgaria	BGR	Nigeria	NGA
Chile	CHL	Panama	PAN
Colombia	COL	People's Republic of China	PRC
Costa Rica	CRI	Peru	PER
Croatia	HRV	Philippines	PHI
Egypt	EGY	Poland	POL
El Salvador	SLV	Republic of Korea	KOR
Estonia	EST	Romania	ROU
Ghana	GHA	Russian Federation	RUS
Honduras	HND	Slovak Republic	SVK
Hungary	HUN	South Africa	ZAF
India	IND	Thailand	THA
Indonesia	INO	Trinidad and Tobago	TTO
Iran	IRN	Tunisia	TUN
Jordan	JOR	Turkey	TUR
Kazakhstan	KAZ	Ukraine	UKR
Kyrgyz Republic	KGZ	Uruguay	URY
Latvia	LVA	Viet Nam	VIE
Lithuania	LTU	Zimbabwe	ZWE
Malaysia	MAL		

ISO = International Organization for Standardization.

Source: Authors' compilation.

REFERENCES

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors for Clustering?" NBER Working Paper No. W24003.

Agosin, Manuel R., Roberto Alvarez, and Claudio Bravo-Ortega. 2012. "Determinants of Export Diversification around the World: 1962–2000." *The World Economy* 35 (3): 295–315.

Altinok, Nadir, Claude Diebolt, and Jean-Luc Demeulemeester. 2014. "A New International Database on Education Quality: 1965–2010." *Applied Economics* 46 (11): 1212–47.

Amsden, Alice H. 2001. The Rise of "the Rest" : Challenges to the West from Late-Industrialization Economies. New York: Oxford University Press.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Asian Development Bank. 2007. *Asian Development Outlook 2007: Growth Amid Change*. Manila.

Bahar, Dany, Ricardo Hausmann, and César A. Hidalgo. 2014. "Neighbors and the Evolution of the Comparative Advantage of Nations: Evidence of International Knowledge Diffusion?" *Journal of International Economics* 92 (1): 111–23.

Bahar, Dany, Samuel Rosenow, Ernesto Stein, and Rodrigo Wagner. 2019. "Export Take-Offs and Acceleration: Unpacking Cross-Sector Linkages in the Evolution of Comparative Advantage." *World Development* 117 (C): 48–60.

Balassa, Bela. 1965. "Trade Liberalisation and 'Revealed' Comparative Advantage." *The Manchester School* 33 (2): 99–123.

Barro, Robert J., and Jong-Wha Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. W15902.

Booth, Anne. 1999. "Education and Economic Development in Southeast Asia: Myths and Realities." *ASEAN Economic Bulletin* 16 (3): 290–306.

Cameron, A. Colin, and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge University Press.

Chang, Ha-Joon. 2012. "23 Things They Don't Tell You about Capitalism?" New York: Bloomsbury Publishing.

Ciccone, Antonio, and Elias Papaioannou. 2009. "Human Capital, the Structure of Production, and Growth." *The Review of Economics and Statistics* 91 (1): 66–82.

Coniglio, Nicola D., Davide Vurchio, Nicola Cantore, and Michele Clara. 2018. "On the Evolution of Comparative Advantage: Path-Dependent versus Path-Defying Changes." Papers in Evolutionary

Economic Geography (PEEG) 1818, Utrecht University, Department of Human Geography and Spatial Planning, Group Economic Geography, revised Apr 2018.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105 (10): 3150–182.

Felipe, Jesus, and César A. Hidalgo. 2014. "Economic Diversification: Implications for Kazakhstan." Manuscript.

Felipe, Jesus, Aashish Mehta, and Changyong Rhee. 2019. "Manufacturing Matters…but It's the Jobs that Count." *Cambridge Journal of Economics* 43 (1): 139–68.

Felker, Greg, Jomo Kwame Sundaram, and Rajah Rasiah. 2013. *Industrial Technology Development in Malaysia: Industry and Firm Studies*. London: Routledge.

Frenken, Koen, Frank Van Oort, and Thijs Verburg. 2007. "Related Variety, Unrelated Variety and Regional Economic Growth." *Regional Studies* 41 (5): 685–97.

Gronau, Reuben. 1998. "A Useful Interpretation of R2; in Binary Choice Models (Or, Have We Dismissed the Good Old R2; Prematurely)." Princeton University IRS Working Paper No. 397.

Hanushek, Eric A., and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46 (3): 607–68.

———. 2009. "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation." NBER Working Paper No. W14633.

Hartmann, Dominik, Miguel R. Guevara, Cristian Jara-Figueroa, Manuel Aristarán, and César A. Hidalgo. 2017. "Linking Economic Complexity, Institutions, and Income Inequality." *World Development* 93: 75–93.

Hausmann, Ricardo, César A. Hidalgo, Sebastián Bustos, Michele Coscia, Sarah Chung, Juan Jimenez, Alexander Simoes, and Muhammed A. Yildirim. 2011. *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. Cambridge, MA: Center for International Development, Harvard University.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. "What You Export Matters." *Journal of Economic Growth* 12 (1): 1–25.

Hausmann, Ricardo, Francisco R. Rodríguez, and Rodrigo A. Wagner. 2006. "Growth Collapses." CID Working Paper No. 136.

Hidalgo, César A., Pierre-Alexandre Balland, Ron Boschma, Mercedes Delgado, Maryann Feldma, Koen Frenken, Edward Glaeser, Canfei He, Dieter Kogler, Andrea Morrison, Frank Neffke, David Rigby, Scott Stern, Siqi Zheng, and Shengjun Zhu. 2018. "The Principle of Relatedness." Papers in Evolutionary Economic Geography (PEEG) 1830, Utrecht University, Department of Human Geography and Spatial Planning, Group Economic Geography, revised Jul 2018.

Hidalgo, César A., and Ricardo Hausmann. 2009. "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences* 106 (26): 10570–75.

Hidalgo, César A., Bailey Klinger, Albert-Lazlo Barabási, and R. Hausmann. 2007. "The Product Space Conditions the Development of Nations." *Science* 317 (5837): 482–87.

Hobday, Michael. 1995. *Innovation in East Asia: The Challenge to Japan.* Aldershot: Edward Elgar.

Jetter, Michael, and Andrés Ramírez Hassan. 2015. "Want Export Diversification? Educate the Kids First." *Economic Inquiry* 53 (4): 1765–82.

Johnson, Chalmers. 1982. MITI and the Japanese Miracle: The Growth of Industrial Policy, 1925-1975. Stanford, CA: Stanford University Press.

Jomo, Kwame Sundaram. 2003. *Southeast Asian Paper Tigers?* London: Routledge.

Krueger, Alan B., and Mikael Lindahl. 2001. "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39 (4): 1101–136.

Leamer, Edward E. 1984. Sources of Comparative Advantage: Theory and Evidence. Cambridge, MA: MIT Press.

Mehta, Aashish, and Jesus Felipe. 2014. "Education and the Journey to the Core: Path Dependence or Leapfrogging?" https://dx.doi.org/10.2139/ssrn.2467895.

Nelson, Richard, and Sidney Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: The Belknap Press of Harvard University Press.

Redding, Stephen. 1999. "Dynamic Comparative Advantage and the Welfare Effects of Trade." *Oxford Economic Papers* 51 (1): 15–39.

Saviotti, Pier Paolo, and Koen Frenken. 2008. "Export Variety and the Economic Performance of Countries." *Journal of Evolutionary Economics* 18 (2): 201–18.

Schultz, Theodore W. 1975. "The Value of the Ability to Deal with Disequilibria." *Journal of Economic Literature* 13 (3): 827–46.

Stiglitz, Joseph E. 1996. "Some Lessons from the East Asian Miracle." *The World Bank Research Observer* 11 (2): 151–77.

Stiglitz, Joseph E., and Bruce C. Greenwald. 2014. Creating a Learning Society: A New Approach to Growth, Development, and Social Progress. Columbia University Press.

Studwell, Joe. 2013. *How Asia Works: Success and Failure in the World's Most Dynamic Region*. New York: Grove Press.

The Growth Lab at Harvard University. 2019a. International Trade Data (HS, 92). https://doi.org/10.7910/DVN/T4CHWJ, Harvard Dataverse, V3, UNF:6:NvC2E0DSAPV7VWI+FbltIg==[fileUNF].

The Growth Lab at Harvard University. 2019b. International Trade Data (SITC, Rev. 2). https://doi.org/10.7910/DVN/H8SFD2, Harvard Dataverse, V4, UNF:6:pIf+v0U5MmQ6GJjXXVeIug== [fileUNF].

Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

World Bank. 1993. *The East Asian Miracle*. Washington, DC.

Education and the Evolution of Comparative Advantage

The paper provides the first evidence that countries with high education levels were more successful in developing comparative advantage in products unrelated to those they already export. In contrast, education appears unimportant for developing comparative advantage in products that are intrinsically complex or education intensive.

About the Asian Development Bank

ADB is committed to achieving a prosperous, inclusive, resilient, and sustainable Asia and the Pacific, while sustaining its efforts to eradicate extreme poverty. Established in 1966, it is owned by 68 members —49 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.

ADB

ASIAN DEVELOPMENT BANK
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
www.adb.org